

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tesisenxarxa.net) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tesisenred.net) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tesisenxarxa.net) service has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized neither its spreading and availability from a site foreign to the TDX service. Introducing its content in a window or frame foreign to the TDX service is not authorized (framing). This rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author



Tesis Doctoral

**AN OVERNIGHT PARCEL LOGISTICS COMPANY'S
CAPILLARY DISTRIBUTION NETWORK DESIGN BY
STOCHASTICAL MODELING**

Autor:

Oscar Rioja San Martín
Ingeniero en Organización Industrial

Director de la tesis:

Joaquim Lloveras Maciá

Trabajo para la obtención del título de Doctor Europeo

Programa de doctorado Proyectos de Innovación Tecnológica en la Ingeniería de Diseño
de Producto y Proceso

Escuela Técnica Superior de Ingeniería Industrial de Barcelona
Universidad Politécnica de Catalunya

Barcelona, Enero 2014



**AN OVERNIGHT PARCEL LOGISTICS COMPANY'S
CAPILLARY DISTRIBUTION NETWORK DESIGN BY
STOCHASTICAL MODELING**

Oscar Rioja San Martín

Memoria presentada para optar al título de Doctor Ingeniero Industrial Europeo
ETSEIB-UPC

Director de la tesis: Joaquim Lloveras Maciá

A Txiri

In this thesis, an overnight parcel logistics company's capillary distribution network will be modelled by stochastic techniques. In particular, the capillary distribution network an overnight parcel logistics company has in Sant Cugat del Vallés, belonging to the hub located in Barberá del Vallés, in Barcelona's metropolitan area.

The tool developed in this research will provide a stochastic modelling of this capillary distribution network. More specifically, the way the costs of each distribution areas are obtained for the design of the capillary network is by regression analysis. To predict the workload of each distribution area, the number of deliveries and collections, and the weight of each one, a model will be created by probability distribution functions.

The overnight parcel logistics company needs to design a capillary distribution network that is capable of collecting all the merchandise that their clients wish to deliver, and, at the same time, have the capacity to distribute all shipments to their destination. To reach this goal, the designers have to decide the type of vehicles (van, lorry or trailer) and the quantity of each type of vehicle the capillary distribution network needs to cover the area around each hub efficiently. This number of vehicles has to satisfy the delivery/collection quality requirements with the minimum cost to guarantee the maximum profit for the overnight parcel logistics company.

To design this network, the most powerful tool the managers have is reliable and accurate information about the nature of the goods needed to be collected and delivered, and consequently, about the randomness of the merchandise. This information can be obtained through a model. For an external observer, a model of a system is an object that the observer can use to answer any question about the system he is interested in. So, once a model is created, it can be used to represent the behaviour of the real system.

In the design of a capillary distribution network, the information that is needed to create the most profitable network is the number of deliveries and collections for each area, and the weight of the different deliveries and collections, i.e. the workload of each area. From the number of deliveries and collections, a manager will know the number of vehicles he needs for each area, and depending on the weight of the collections and the deliveries, the type of vehicles that are needed.

This workload is fundamental for another critical aspect of the overnight parcel logistics company; the delivery and collection cost. In parcel logistics companies these costs are as follows: the parcel logistics company pays a specific amount to the vehicle in charge of the collection, one part of which is due to pick up the shipment, the other which depends on the

weight of the shipment. The delivery cost is analogous; each delivery costs the parcel logistics company a specific amount which is divided into two parts; a fixed amount for serving the shipment, and another one based on the weight of the shipment.

For the design of a capillary distribution network it is important to consider that the delivery and collection costs for the overnight parcel logistics company are the income of the drivers of the network. If the income of a vehicle is not high enough, the continuity of the driver of this vehicle cannot be guaranteed. With this in mind, a continued change of drivers is not advisable for an overnight parcel logistics company due to the adaptation time the drivers need to learn each distribution area and clients' necessities. Therefore, the overnight parcel logistics company's managers need to find a balance between the minimum distribution costs and the continuity of the drivers.

To find this balance any capillary network developed has to be based on guarantee a minimum income thus to ensure the continuity of the drivers. This continuity maximises driver familiarity within their distribution area. With increased familiarity, driver performance improves due to ease in finding addresses and locations as well as efficiency in organising daily routes. Their capacity to make deliveries and collections increases, and therefore, so does their productivity. For this the workload and the driver's income in each distribution area has to be predicted.

This research will explore and analyse how to design the capillary distribution network, and will demonstrate that stochastic modelling is a valid technique to help an overnight parcel logistics company in the decision making related to the design a capillary distribution network. Most models related to logistics companies have been based on deterministic techniques so far. However, in this work the way the costs of each distribution areas are obtained for the design of the capillary network is by regression analysis. To predict the workload of each distribution area, the number of deliveries and collections, and the weight of each one, a model will be created by probability distribution functions.

The design of the capillary network will be based on assigning a distribution area to a vehicle as opposed to a determined route of collections and deliveries for each vehicle, as has been done until now in the vehicle routing problem. The aim of assigning each vehicle to specific distribution areas is to increase the familiarity of the driver with his distribution area, and also, guarantee a minimum income (the income being the cost for the overnight parcel logistics company), thus to ensure the continuity, of its driver.

The whole area the hub under study has to cover can be divided into smaller sub-areas. These sub-areas are postcodes. In the area the hub has to cover there are also several towns. It is possible that a town is compounded by more than one post code, or conversely, two small towns share the same post code.

The town under research, Sant Cugat del Vallés, has nearly 90.000 inhabitants, and is compounded of seven postcodes and through the stochastic tool developed in this research the workload and drivers' income of every postcode is analysed and modelled. With this information it will be possible to determine if each postcode can be assign to a driver in order to guarantee a minimum income for him, and therefore, his continuity. It can be also possible that the income of the driver in a postcode is not high enough to ensure his continuity. In these cases, more than one postcode can be assign to the same driver. By contrast, if the income and the workload in a postcode are excessive, more than one driver can be assigned to the same postcode.

In this way, it will be possible to determine if a postcode can be considered a distribution area for a driver, or conversely, more than one postcode will be needed to establish a distribution area.

The income of each postcode can be divided into two parts: the income related to the deliveries and the income related to the collections. As these incomes will be predicted by a regression analysis, the delivery income is considered the endogenous variable, and is estimated by the exogenous variables which are: the number of deliveries per day and the total delivered weight per day. The collection income is analogous, predicted by the exogenous variables: the number of collections per day and the total collected weight per day.

The way the regression models are fitted is by maximum likelihood. Also, by maximum likelihood different regression analysis will be compared and the one with best results will be used to predict the income of the vehicles. This comparison will be made by the likelihood ratio test. The different regression models analysed are ordinary least squares, generalised linear models and errors distributed with no normal probability distribution functions. In total, nine different delivery income regression analysis will be undertaken and compared in each postcode. Analogously, nine different collection income regression analysis will be undertaken and compared in each postcode.

As for the workload, it will be predicted by the adjustment of the corresponding variables by probability distribution functions. First the number per day of deliveries/collections is fitted.

Second, the weight of each delivery/collection is fitted. Finally, the model is applied by using the inverse-transform method to settle the numbers of deliveries/collections, and with the sum of the weight of every delivery/collection (obtained analogously) the total delivered/collected weight per day and area is determined.

Finally, the last step is the simulation of the models. With this simulation, the hypothesis made at the beginning of this research will be confirmed. For the workload model simulation, the model is applied by using the inverse-transform method to settle the numbers of deliveries/collections, and with the sum of the weight of every delivery/collection (obtained analogously) the total delivered/collected weight per day and area is determined.

As for the delivery and collection costs model simulation, the input data is the values of the number of delivery/collection and total delivered/collected weight random variables. In order to carry out a simulation using random inputs, their probability distribution functions have to be specified. And that is exactly what is done in the workload model. So this model will be the input data in the delivery and collection costs model.

En esta tesis, la red de distribución capilar de una empresa de paquetería industrial va a ser modelada mediante técnicas estocásticas. Más concretamente, la red de distribución capilar que una empresa de paquetería industrial posee en la población de Sant Cugat del Vallés, perteneciente al hub de la empresa situado en Barberá del Vallés, en el área metropolitana de Barcelona.

La herramienta desarrollada en esta investigación proporcionará un modelado estocástico de esta red de distribución capilar. Más en concreto, los costes de cada una de las áreas de distribución de esta red capilar van a ser modelados mediante regresión lineal. Para predecir las cargas de trabajo de cada una de las áreas de distribución, i.e. el número de repartos y de recogidas, así como el peso de cada uno de ellos, se ha creado un modelo basado en funciones de distribución de probabilidad.

Una red de distribución capilar diseñada para una empresa de paquetería industrial ha de ser capaz de recoger toda la mercancía que sus clientes desean distribuir, y al mismo tiempo, ha de tener la capacidad de repartir todos los envíos a sus correspondientes destinos. Para alcanzar este objetivo, los diseñadores han de decidir el tipo de vehículos que necesitan (furgoneta, camión carrozado o camión), y la cantidad de cada uno de ellos, de tal forma que la red de distribución capilar pueda cubrir de manera satisfactoria el área de influencia alrededor de su terminal de transporte, con el mínimo coste, para garantizar el máximo beneficio.

Para diseñar una red de distribución capilar, la herramienta más poderosa que los responsables de la misma pueden disponer es una información precisa y fiable de la naturaleza de los bienes que han de ser recogidos y/o repartidos, y por consiguiente, de la aleatoriedad de esta mercancía. En este trabajo, esta información se obtendrá a través de un modelo. Para un observador externo, un modelo de un sistema es un objeto que el observador puede emplear para responder cualquier pregunta sobre el sistema en el que está interesado. Por lo tanto, una vez un modelo este creado, éste puede ser empleado para representar el comportamiento del sistema real.

En el diseño de una red de distribución capilar, la información necesaria para crear la red de distribución más rentable posible es el número de recogidas y repartos para cada una de las áreas, así como el peso de cada uno de ellos. Dicho de otra de

manera, es necesario conocer la carga de trabajo en cada una de las áreas de distribución. A partir del número de repartos y recogidas, los responsables podrán conocer el número de vehículos que necesitan para cada área. Y en función del peso de éstos, el tipo de vehículo necesario que mejor se adapta a las necesidades.

Esta carga de trabajo es fundamental para otro aspecto crítico: los costes de reparto y recogida. En las empresas de paquetería industrial, los costes de recogida son como siguen. La empresa paga una cierta cantidad al vehículo de distribución por cada una de las recogidas que realiza. Esta cantidad se divide en dos partes, una fija por el hecho de haber realizado correctamente la recogida, y otra variable, que depende del peso de esta recogida.

El coste de reparto es análogo. Cada reparto supone un coste para la empresa de paquetería industrial, que también consta de dos partes. Una fija, que se paga al vehículo de distribución por realizar el reparto de manera adecuada, y otra variable que depende del peso de este reparto.

A la hora de diseñar una red de distribución capilar, es importante considerar que los gastos de recogida y reparto de la empresa de paquetería industrial son los ingresos de los vehículos de distribución. Si los ingresos de un vehículo en particular por realizar los repartos y las recogidas no son lo suficientemente elevados, no es posible garantizar la continuidad de este conductor. Teniendo esto en cuenta, un cambio continuado de vehículos de distribución no es aconsejable para una empresa de paquetería industrial, debido al tiempo de adaptación que un nuevo vehículo de distribución necesita para conocer tanto su nueva área de distribución como las necesidades de los clientes que se dan en esa área. Por lo tanto, los responsables de la empresa de paquetería industrial han de encontrar un equilibrio entre el mínimo coste de reparto y recogida posible y la continuidad de los conductores de los vehículos de distribución.

Para encontrar este equilibrio, toda red de distribución capilar a diseñar ha de garantizar unos ingresos mínimos que aseguren la continuidad de los vehículos de distribución. Esta continuidad maximiza la familiaridad de los conductores con sus áreas de distribución. A mayor familiaridad de un conductor con su área de distribución, mayor será el rendimiento del vehículo de distribución, ya que el conductor del mismo

incrementa su capacidad de encontrar direcciones y localizaciones, así como su habilidad de organizar sus rutas diarias. Un mejor y mayor conocimiento de un área de distribución permite realizar más repartos y más recogidas, incrementando la productividad del vehículo de distribución de esa área.

Esta investigación, además de explorar y analizar como diseñar una red de distribución capilar, demostrará que un modelado estocástico es una técnica válida para ayudar a las empresas de paquetería industrial en la toma de decisiones sobre sus redes de distribución capilar. La mayoría de los modelos desarrollados hasta el momento para empresas logísticas han sido modelos deterministas. Sin embargo, en este trabajo los costes de cada una de las áreas de distribución en los que se basa el diseño de la misma, se van modelar mediante regresión lineal. Para predecir las cargas de trabajo de cada una de las áreas de distribución, el número de repartos y de recogidas, así como el peso de cada uno de ellos, se ha creado un modelo basado en funciones de distribución de probabilidad.

El diseño propuesto en este trabajo para la red capilar de distribución, consiste en asignar un área de distribución a un vehículo de distribución, en contra posición a asignar una ruta determinada de repartos y recogidas para cada vehículo como se ha hecho hasta ahora en problema de encaminamiento de vehículos. El objetivo de asignar a cada vehículo a una determinada área de distribución es incrementar la familiaridad del conductor con esta área de distribución y garantizar un ingreso mínimo (siendo el ingreso el coste para la empresa de paquetería industrial), y de este modo, asegurar la continuidad del conductor.

Toda el área que el terminal de transporte de la empresa de paquetería industrial ha de cubrir puede dividirse en zonas más pequeñas, denominadas sub-áreas. Estas sub-áreas son los códigos postales. Esta área a cubrir por el hub, puede estar también formada, según la particularidad de cada caso, por más de una población. También es posible, en función de cada caso, que una población esté compuesta por más de un código postal. Y puede darse el caso, que de forma contraria, que un código postal incluya más de una población.

La población que se va a estudiar en este trabajo, Sant Cugat del Vallés, tiene casi 90.000 habitantes y está compuesta por siete códigos postales. Mediante la herramienta desarrollada en esta investigación, la carga de trabajo y los ingresos de los conductores de cada uno de estos códigos postales será analizado y estudiado. Con esta información, va a ser posible determinar si cada una de estas áreas de distribución puede ser asignada a un único conductor y asegurarle unos ingresos mínimos, que a su vez, permitan su continuidad. Puede ocurrir que los ingresos en un código postal no sean lo suficientemente altos como para asegurar la continuidad del conductor. En estos casos, más de un código postal se asignará al mismo conductor. Por el contrario, puede darse también el caso de que la carga de trabajo y los ingresos que esta produce sean excesivos para un único conductor, por lo que más de un conductor será asignado al mismo código postal.

De esta manera se podrá determinar si un código postal puede considerarse un área de distribución por sí mismo, o por el contrario, si más de un código postal son necesaria para conformar un área de distribución.

Los ingresos de un vehículo de distribución en cada código postal se pueden dividir en dos partes. Los ingresos debidos a las recogidas y los ingresos debido a los repartos. Como estos ingresos se predecirán mediante regresión lineal, el ingreso debido a los repartos se considera la variable endógena, y será estimado mediante las variables exógenas, número de repartos por día y peso total repartido por día. Los ingresos de recogida son análogos, siendo el ingreso de recogido la variable endógena, y el número de recogidas por día y el peso total recogido por día las variables exógenas.

El ajuste de los modelos de regresión en este trabajo se hará mediante la técnica de máxima verosimilitud. Esta técnica presenta una segunda ventaja, ya que además del ajuste de los modelos de regresión, el empleo de esta técnica permite comparar diferentes análisis de regresión, y aquel con mejores resultados será el que se emplee para predecir el los ingresos de cada área de distribución mediante el coeficiente de verosimilitud. Los diferentes modelos de regresión analizados son: método de mínimos cuadrados, modelos lineales generalizados y modelos cuyos errores no se distribuyen con funciones de distribución de probabilidad no normales. En total, nueve modelos de regresión de ingresos de reparto diferentes se van a llevar a cabo y serán comparados en

cada uno de los códigos postales. Análogamente, se van a implementar nueve modelos de regresión de ingresos de recogidas diferentes en cada código postal, comparándolos entre ellos y eligiendo el que mejor resultados de.

Respecto a la carga de trabajo, ésta va a ser predicha a través del ajuste de las correspondientes variables mediante funciones de distribución de probabilidad. En primer lugar se ajustará el número de recogidas/repartos por día. Posteriormente, el peso de cada recogida/reparto se ajustado. Finalmente, se aplicará el modelo empleando el método de transformada inversa para fijar el número de repartos/recogidas por día, y con la suma del peso de cada estos (obtenido de manera análoga en ambos casos), se determina el peso por día total recogido/repartido en cada una de las áreas.

Como último paso, solo quedaría la simulación de los modelos. Esta simulación será la herramienta empleada para confirmar las hipótesis hechas al comienzo de esta investigación. Para simular el modelo de carga de trabajo, como se ha explicado, se empleará el método de la transformada inversa. Respecto al modelo de costes de reparto y recogida, los datos de entrada para simulación serán los valores obtenidos de la simulación del modelo de carga de trabajo, por lo que se puede decir que el modelo de carga de trabajo será los datos de entradas del modelo de costes de reparto y recogida y su simulación.

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisor Joaquim Lloveras for his untiring guidance, patience and sincere and helpful advice. Without his support and orientation, this work would never have arrived to a safe port, never. I am very grateful also to Universitat Politècnica de Catalunya, because of giving me the opportunity to start a carrier in this fantastic world that is the science and the research, and to develop my skills and consider myself as a researcher in these hard times for science in Spain.

I would also like to express my gratitude to Benjamin Heydecker, my tutor during my stay at Centre for Transport Studies at University College of London. His time and dedication, and all the knowledge that he shared in our weekly meetings are a key factor in achieving this work. His sincere interest and dedication undoubtedly increased the value and ensured the success of this project. My gratitude to University College of London for opening its doors and for making me room at Centre for Transport Studies department.

Of course, I cannot forget all of the university colleagues that, by the end of this journey, have become true friends. To all of them I would to say thank you, and tell them that those moments, both in Barcelona and London, are unforgettable. I would like to give special thanks to Carlos, the “Doctor”, whose enthusiasm, optimism and happiness encouraged all of us to try new things and continue with our respective researches. It is impossible to forget the long hours in the trains to Girona and Castellon, where the endless conversations, Mexican songs and jokes made everything easier.

It is essential not to forget my family, who have always been there supporting me throughout these five years, and especially my mother, who has asked every week without fail how my thesis is going. And also to Vincent, for receiving me in his house and making me feel part of his family.

And finally, I would like to thank Miriam for being Miriam. For her determination in making our trip to London happen, her companionship at the other side of the table writing her novel while I was writing my thesis, and for her patience with the English corrections in the endless drafts of papers and texts. I would not be writing these acknowledgements if she had not been with me all this time.

ÍNDICE

	<u>Página</u>
I. INTRODUCTION	1
1.1 Problem statement	1
1.2 Research hypothesis	3
1.3 Justification for research	3
1.4 Objectives	4
1.5 Scope	5
1.6 Methodology	6
II. CASO DE ESTUDIO	10
2.1. Introducción	10
2.2. Redes de distribución	10
2.3. Clasificación de las redes de distribución	10
2.4. Rol del sector de la paquetería industrial	11
2.5. Estructura de una empresa de paquetería industrial	12
2.6. Descripción del caso de estudio	14
2.6.1. Características de la red capilar de la empresa de paquetería industrial del caso de estudio	15
2.6.2. Costes de reparto y recogida de la empresa de paquetería del caso de estudio	19
2.6.3. Municipio de Sant Cugat del Vallés	20
III. ESTADO DEL ARTE	22
3.1. Introducción	22
3.2. Definición de modelo y sus funciones.	23
3.3. Tipos de modelos	24
3.3.1. Modelos mentales	24

3.3.2. Modelos físicos	24
3.3.3. Modelos matemáticos	25
3.4. Definición de simulación	26
3.5. Ventajas y desventajas de la creación de modelos y su simulación	27
3.6. Tipos de modelos matemáticos	29
3.6.1. Modelos matemáticos estáticos/dinámicos	29
3.6.2. Modelos matemáticos deterministas/estocásticos	29
3.6.3. Modelos matemáticos continuos/discretos	30
3.7. Tipos de modelos matemáticos deterministas	31
3.7.1. Programación lineal y no lineal	31
3.7.2. Programación entera	32
3.7.3. Programación multiobjetivo	32
3.7.4. Modelos de redes	32
3.7.5. Programación heurística	33
3.8. Tipos de modelos matemáticos estocásticos	34
3.8.1. Cadenas de Markov	34
3.8.2. Análisis de decisión	34
3.8.3. Teoría de la utilidad	36
3.8.4. Teoría de juegos	37
3.8.5. Proceso de Poisson	38
3.8.6. Teoría de colas	39
3.8.7. Análisis de Regresión	40
3.9. Técnicas de regresión	42
3.9.1. Método de mínimos cuadrados	43
3.9.2. Regresión robusta	46
3.9.3. Regresión ortogonal	47
3.9.4. Modelos lineales generalizados	49
3.9.5. Método de máxima verisimilitud	50
3.9.5.1. <i>Test del cociente de verosimilitud</i>	51
3.10. Aspectos de la simulación	52
3.10.1. Monte Carlo	52
3.10.1.1. <i>Números aleatorios</i>	53
3.11. Aplicaciones de modelos de simulación matemáticos en el sector del transporte	55
 IV. METODOLOGÍA	 58
4.1. Introducción	58
4.2. Identificación de variables	59
4.3. Toma de muestras	60
4.4. Creación del modelo	64
4.4.1. Modelo de los costes de distribución capilar	64
4.4.1.1. <i>Análisis de regresión para el modelo de los costes de reparto y recogida</i>	66
4.4.1.1.1. <i>Elaboración del modelo</i>	66
- Distribuciones de probabilidad de los errores	67
- Modelos lineales generalizados	70
4.4.1.1.2. <i>Selección de variables</i>	70
4.4.1.1.3. <i>Adecuación del modelo</i>	73
4.4.1.2. <i>Validación del modelo</i>	75
4.4.2. Modelo de la carga de trabajo	77
4.4.2.1. <i>Validación del modelo</i>	77

4.4.2.1.1. Test Chi cuadrado	78
4.4.2.1.2. Test de Kolmogorov-Smirnov	78
4.4.2.1.3. Distribuciones empíricas	79
4.5. Simulación	80
4.5.1. Método de la transformada inversa	80
4.5.2. Simulación del modelo de carga de trabajo	81
4.5.3. Simulación de los costes de reparto y recogida	82
V. ANÁLISIS DE RESULTADOS	83
5.1. Resultados de las entrevistas	83
5.2. Estudio de cada código postal	85
5.2.1. Código postal 08198	85
5.2.1.1. Vehículo de 1500 kg. para el código postal 08198	86
5.2.1.2. Vehículo de 3500 kg. de peso máximo autorizado para el código postal 08198	116
5.2.1.3. Vehículo de más de 3500 kg. de peso máximo autorizado para el código postal 08198	145
5.2.2. Código postal 08197	160
5.2.3. Código postal 08196	189
5.2.4. Código postal 08195	217
5.2.5. Código postal 08174	246
5.2.6. Código postal 08173	275
5.2.7. Código postal 08172	305
5.3. RESUMEN DE LOS RESULTADOS	333
VI. CONCLUSIONS	335
VII. BIBLIOGRAFÍA	339
ANEXO I	A-1

ÍNDICE DE TABLAS

Página

CASO DE ESTUDIO

Tabla 1. Cuadro de ingresos de reparto y recogida en función del peso	20
---	----

ESTADO DEL ARTE

Tabla 2: funciones cadena más comunes	50
---------------------------------------	----

METODOLOGÍA

Tabla 3: Costes por reparto, repartos por día y peso total repartido por día	65
Tabla 4: valores para el test Kolmogorov-Smirnov	79

ANÁLISIS DE RESULTADOS

Tabla 5: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08198	87
Tabla 6: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08198	88
Tabla 7: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08198	95
Tabla 8: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08198	96
Tabla 9: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08198	97
Tabla 10: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08198	98
Tabla 11: VMA modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08198	100
Tabla 12: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08198	100
Tabla 13: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08198	100
Tabla 14: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08198	100
Tabla 15: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08198	100

Tabla 16: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08198	101
Tabla 17: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08198	101
Tabla 18: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08198	101
Tabla 19: VMA modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08198	101
Tabla 20: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08198	104
Tabla 21: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08198	105
Tabla 22: VMA modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08198	105
Tabla 23: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08198	105
Tabla 24: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08198	106
Tabla 25: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08198	106
Tabla 26: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08198	106
Tabla 27: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08198	106
Tabla 28: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08198	106
Tabla 29: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08198	106
Tabla 30: VMA modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08198	107
Tabla 31: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08198	109
Tabla 32: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08198	109
Tabla 33: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08198	110
Tabla 34: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08198	111
Tabla 35: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08198	112
Tabla 36: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08198	112
Tabla 37: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal	114
Tabla 38: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08198	114
Tabla 39: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08198	115
Tabla 40: Resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08198	115
Tabla 41: Toma de muestras de los repartos del vehículo de 3500 kg. PMA para el código postal 08198	117
Tabla 42: Toma de muestras de las recogidas del vehículo de 3500 kg. PMA para el código postal 08198	118
Tabla 43: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 3500 kg. PMA código postal 08198	125
Tabla 44: Muestra de calibración y de predicción de repartos del vehículo de 3500 kg. PMA código postal 08198	126
Tabla 45: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 3500 kg. PMA código postal 08198	127
Tabla 46: Muestra de calibración y de predicción de recogidas del vehículo de 3500 kg. PMA código postal 08198	128

Tabla 47: VMA modelo de costes de reparto con distribución de probabilidad normal vehículo 3500 kg. PMA código postal 08198	129
Tabla 48: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 3500 kg. PMA código postal 08198	129
Tabla 49: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 3500 kg. PMA código postal 08198	129
Tabla 50: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 3500 kg. PMA código postal 08198	129
Tabla 51: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 3500 kg. PMA código postal 08198	130
Tabla 52: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 3500 kg. PMA código postal 08198	130
Tabla 53: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 3500 kg. PMA código postal 08198	130
Tabla 54: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 3500 kg. PMA código postal 08198	130
Tabla 55: VMA modelo de costes de reparto con distribución de probabilidad logística vehículo 3500 kg. PMA código postal 08198	130
Tabla 56: Resultado prueba Welch para el modelo de costes de reparto del vehículo 3500 kg. PMA para el código postal 08198	133
Tabla 57: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 3500 kg. para el código postal 08198	133
Tabla 58: VMA modelo de costes de recogida con distribución de probabilidad normal vehículo 3500 kg. PMA código postal 08198	134
Tabla 59: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 3500 kg. PMA código postal 08198	134
Tabla 60: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 3500 kg. PMA código postal 08198	134
Tabla 61: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 3500 kg. PMA código postal 08198	134
Tabla 62: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 3500 kg. PMA código postal 08198	135
Tabla 63: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 3500 kg. PMA código postal 08198	135
Tabla 64: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 3500 kg. PMA código postal 08198	135
Tabla 65: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 3500 kg. PMA código postal 08198	135
Tabla 66: VMA modelo de costes de recogida con distribución de probabilidad logística vehículo 3500 kg. PMA código postal 08198	135
Tabla 67: Resultado prueba Welch para el modelo de costes de recogida del vehículo 3500 kg. PMA para el código postal 08198	138
Tabla 68: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08198	138
Tabla 69: Prueba Chi Cuadrado número de repartos para el vehículo 3500 kg. PMA código postal 08198	139
Tabla 70: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 3500 kg. PMA código postal 08198	140
Tabla 71: Prueba Chi Cuadrado número de recogidas para el vehículo 3500 kg. PMA código postal 08198	141
Tabla 72: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 3500 kg. PMA código postal 08198	141
Tabla 73: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 3500 kg. PMA en el código postal 08198	143
Tabla 74: Resumen de la simulación de los costes de reparto del vehículo de 3500 kg. de PMA en el código postal 08198	143
Tabla 75: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 3500 kg. PMA en el código postal 08198	144
Tabla 76: Resumen de la simulación de los costes de recogida del vehículo de 3500 kg. de PMA en el código postal 08198	144
Tabla 77: Toma de muestras de las recogidas del vehículo de más de 3500 kg. PMA para el código postal 08198	146

Tabla 78: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo de más de 3500 kg. PMA código postal 08198	150
Tabla 79: Muestra de calibración y de predicción de recogidas del vehículo de más de 3500 kg. PMA código postal 08198	151
Tabla 80: VMA modelo de costes de recogida con distribución de probabilidad normal vehículo 3500 kg. PMA código postal 08198	152
Tabla 81: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo de más de 3500 kg. PMA código postal 08198	152
Tabla 82: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo de más de 3500 kg. PMA código postal 08198	152
Tabla 83: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo de más de 3500 kg. PMA código postal 08198	152
Tabla 84: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo de más de 3500 kg. PMA código postal 08198	153
Tabla 85: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo de más de 3500 kg. PMA código postal 08198	153
Tabla 86: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo de más de 3500 kg. PMA código postal 08198	153
Tabla 87: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo de más de 3500 kg. PMA código postal 08198	153
Tabla 88: VMA modelo de costes de recogida con distribución de probabilidad logística vehículo de más de 3500 kg. PMA código postal 08198	153
Tabla 89: Resultado prueba Welch para el modelo de costes de recogida del vehículo de más de 3500 kg. PMA para el código postal 08198	155
Tabla 90: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de más de 3500 kg. para el código postal 08198	156
Tabla 91: Prueba Chi Cuadrado número de recogidas para el vehículo de más de 3500 kg. PMA código postal 08198	157
Tabla 92: Test Kolmogorov-Smirnov número de recogida para el vehículo de más de 3500 kg. PMA código postal 08198	157
Tabla 93: Distribución empírica número de recogidas para el vehículo de más de 3500 kg. PMA código postal 08198	158
Tabla 94: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo de más de 3500 kg. PMA código postal 08198	158
Tabla 95: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo de más de 3500 kg. PMA en el código postal 08198	159
Tabla 96: Resumen de la simulación de los costes de recogida del vehículo de más de 3500 kg. de PMA en el código postal 08198	159
Tabla 97: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08197	161
Tabla 98: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08197	162
Tabla 99: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08197	169
Tabla 100: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08197	170
Tabla 101: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08197	171
Tabla 102: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08197	172
Tabla 103: VMA modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08197	173
Tabla 104: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08197	173
Tabla 105: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08197	173
Tabla 106: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08197	173
Tabla 107: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08197	174

Tabla 108: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08197	174
Tabla 109: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08197	174
Tabla 110: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08197	174
Tabla 111: VMA modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08197	174
Tabla 112: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08197	177
Tabla 113: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08197	177
Tabla 114: VMA modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08197	178
Tabla 115: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08197	178
Tabla 116: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08197	178
Tabla 117: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08197	178
Tabla 118: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08197	179
Tabla 119: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08197	179
Tabla 120: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08197	179
Tabla 121: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08197	179
Tabla 122: VMA modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08197	179
Tabla 123: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08197	182
Tabla 124: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08197	183
Tabla 125: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08197	184
Tabla 126: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08197	184
Tabla 127: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08197	185
Tabla 128: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08197	186
Tabla 129: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08197	187
Tabla 130: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08197	187
Tabla 131: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08197	188
Tabla 132: Resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08197	188
Tabla 133: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08196	190
Tabla 134: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08196	191
Tabla 135: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08196	198
Tabla 136: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08196	199
Tabla 137: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08196	200
Tabla 138: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08196	201

Tabla 139: VMA modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08196	202
Tabla 140: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08196	202
Tabla 141: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08196	202
Tabla 142: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08196	202
Tabla 143: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08196	202
Tabla 144: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08196	203
Tabla 145: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08196	203
Tabla 146: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08196	203
Tabla 147: VMA modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08196	203
Tabla 148: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08196	205
Tabla 149: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08196	206
Tabla 150: VMA modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08196	206
Tabla 151: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08196	206
Tabla 152: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08196	206
Tabla 153: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08196	207
Tabla 154: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08196	207
Tabla 155: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08196	207
Tabla 156: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08196	207
Tabla 157: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08196	207
Tabla 158: VMA modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08196	207
Tabla 159: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08196	210
Tabla 160: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08196	210
Tabla 161: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08196	211
Tabla 162: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08196	210
Tabla 163: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08196	213
Tabla 164: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08196	213
Tabla 165: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08196	214
Tabla 166: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08196	215
Tabla 167: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08196	216
Tabla 168: Resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08196	216
Tabla 169: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08195	218

Tabla 170: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08195	219
Tabla 171: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08195	226
Tabla 172: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08195	227
Tabla 173: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08195	228
Tabla 174: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08195	229
Tabla 175: VMA modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08195	230
Tabla 176: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08195	230
Tabla 177: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08195	230
Tabla 178: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08195	230
Tabla 179: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08195	231
Tabla 180: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08195	231
Tabla 181: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08195	231
Tabla 182: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08195	231
Tabla 183: VMA modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08195	231
Tabla 184: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08195	234
Tabla 185: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08195	234
Tabla 186: VMA modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08195	235
Tabla 187: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08195	235
Tabla 188: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08195	235
Tabla 189: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08195	235
Tabla 190: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08195	235
Tabla 191: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08195	236
Tabla 192: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08195	236
Tabla 193: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08195	236
Tabla 194: VMA modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08195	236
Tabla 195: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08195	238
Tabla 196: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08195	239
Tabla 197: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08195	240
Tabla 198: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08195	241
Tabla 199: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08195	241
Tabla 200: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08195	242

Tabla 201: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08195	243
Tabla 202: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08195	244
Tabla 203: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08195	244
Tabla 204: Resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08195	245
Tabla 205: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08174	247
Tabla 206: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08174	248
Tabla 207: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08174	255
Tabla 208: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08174	256
Tabla 209: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08174	257
Tabla 210: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08174	258
Tabla 211: VMA modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08174	259
Tabla 212: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08174	259
Tabla 213: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08174	259
Tabla 214: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08174	259
Tabla 215: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08174	260
Tabla 216: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08174	260
Tabla 217: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08174	260
Tabla 218: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08174	260
Tabla 219: VMA modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08174	260
Tabla 220: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08174	263
Tabla 221: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08174	263
Tabla 222: VMA modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08174	264
Tabla 223: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08174	264
Tabla 224: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08174	264
Tabla 225: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08174	264
Tabla 226: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08174	265
Tabla 227: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08174	265
Tabla 228: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08174	265
Tabla 229: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08174	265
Tabla 230: VMA modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08174	265
Tabla 231: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08174	268

Tabla 232: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08198	268
Tabla 233: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08174	269
Tabla 234: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08174	270
Tabla 235: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08174	271
Tabla 236: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08174	271
Tabla 237: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08174	273
Tabla 238: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08174	273
Tabla 239: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08174	274
Tabla 240: resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08174	274
Tabla 241: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08173	276
Tabla 242: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08173	277
Tabla 243: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08173	284
Tabla 244: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08173	285
Tabla 245: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08173	286
Tabla 246: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08173	287
Tabla 247: VMA modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08173	288
Tabla 248: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08173	288
Tabla 249: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08173	288
Tabla 250: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08173	288
Tabla 251: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08173	289
Tabla 252: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08173	289
Tabla 253: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08173	289
Tabla 254: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08173	289
Tabla 255: VMA modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08173	289
Tabla 256: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08173	292
Tabla 257: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08173	293
Tabla 258: VMA modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08173	294
Tabla 259: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08173	294
Tabla 260: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08173	294
Tabla 261: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08173	294
Tabla 262: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08173	294
Tabla 263: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08173	294

Tabla 264: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08173	295
Tabla 265: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08173	295
Tabla 266: VMA modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08173	295
Tabla 267: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08173	297
Tabla 268: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08173	298
Tabla 269: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08173	299
Tabla 270: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08173	300
Tabla 271: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08173	300
Tabla 272: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08173	301
Tabla 273: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08173	302
Tabla 274: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08173	302
Tabla 275: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08173	303
Tabla 276: Resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08173	303
Tabla 277: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08172	306
Tabla 278: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08172	307
Tabla 279: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08172	314
Tabla 280: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08172	315
Tabla 281: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08172	316
Tabla 282: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08172	317
Tabla 283: VMA modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08172	318
Tabla 284: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08172	318
Tabla 285: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08172	318
Tabla 286: VMA modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08172	318
Tabla 287: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08172	319
Tabla 290: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08172	319
Tabla 291: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08172	319
Tabla 292: VMA modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08172	319
Tabla 293: VMA modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08172	319
Tabla 294: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08172	321
Tabla 295: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08172	322
Tabla 296: VMA modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08172	323
Tabla 297: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08172	323

Tabla 298: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08172	323
Tabla 299: VMA modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08172	323
Tabla 300: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08172	323
Tabla 301: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08172	323
Tabla 302: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08172	324
Tabla 303: VMA modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08172	324
Tabla 304: VMA modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08172	324
Tabla 305: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08172	326
Tabla 306: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08198	327
Tabla 307: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08172	328
Tabla 308: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08172	328
Tabla 309: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08172	329
Tabla 310: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08172	329
Tabla 311: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08172	331
Tabla 312: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08172	331
Tabla 313: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08172	332
Tabla 314: Resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08172	332
Tabla 315: Resumen de los modelos elaborados para cada uno de los códigos postales	333

ÍNDICE DE FIGURAS

Página

INTRODUCTION

Figura 1: Summary of the methodology followed during the development of the stochastic model.

7

CASO DE ESTUDIO

Figura 2: Ejemplo simplificado de una empresa de transporte

Figura 3: Costes acumulados de un envío en una empresa de paquetería industrial

Figura 4: Distribución de códigos postales de la ciudad de Barcelona

Figura 5: Distribución de códigos postales en el municipio de Sant Cugat del Vallés

14

14

16

21

ESTADO DEL ARTE

Figura. 6: Gráfica de la función de utilidad

Figura 7: Gráfica influencia con efecto de los residuales limitado

36

47

ANÁLISIS DE RESULTADOS

Figura 8: Grafica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08198

89

Figura 9: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08198

90

Figura 10: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08198

90

Figura 11: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08198

91

Figura 12: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08198

91

Figura 13: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08198

91

Figura 14: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08198

92

Figura 15: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08198

92

Figura 16: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08198

93

Figura 17: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08198

93

Figura 18: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08198	94
Figura 19: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08198	94
Figura 20: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08198	102
Figura 21: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08198	108
Figura 22: Grafica de autocorrelación de coste de reparto por día vehículo 3500 kg. PMA código 08198	119
Figura 23: Grafica de autocorrelación de número de repartos por día vehículo 3500 kg. PMA código 08198	120
Figura 24: Grafica de autocorrelación del peso total repartido por día vehículo 3500 kg. PMA código 08198	120
Figura 25: Grafica de dispersión de coste de reparto por día vehículo 3500 kg. PMA código 08198	121
Figura 26: Grafica de dispersión de número de repartos por día vehículo 3500 kg. PMA código 08198	121
Figura 27: Grafica de dispersión del peso total repartido por día vehículo 3500 kg. PMA código 08198	122
Figura 28: Grafica de autocorrelación de coste de recogida por día vehículo 3500 kg. PMA código 08198	122
Figura 29: Grafica de autocorrelación del peso total recogido por día vehículo 3500 kg. PMA código 08198	123
Figura 30: Grafica de autocorrelación del peso total recogido por día vehículo 3500 kg. PMA código 08198	123
Figura 31: Grafica de dispersión de coste de recogida por día vehículo 3500 kg. PMA código 08198	124
Figura 32: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08198	124
Figura 33: Grafica de dispersión de peso total recogido por día vehículo 3500 kg. PMA código 08198	124
Figura 34: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 3500 kg. PMA código postal 08198	132
Figura 35: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 3500 kg. PMA código postal 08198	137
Figura 36: Grafica de autocorrelación de coste de recogida por día vehículo de más de 3500 kg. PMA código 08198	147
Figura 37: Grafica de autocorrelación del peso total recogido por día vehículo de más de 3500 kg. PMA código 08198	147
Figura 38: Grafica de autocorrelación del peso total recogido por día vehículo de más de 3500 kg. PMA código 08198	148
Figura 39: Grafica de dispersión de coste de recogida por día vehículo de más de 3500 kg. PMA código 08198	148
Figura 40: Grafica de dispersión de número de recogidas por día vehículo de más de 3500 kg. PMA código 08198	149
Figura 41: Grafica de dispersión de peso total recogido por día vehículo de más de 3500 kg. PMA código 08198	149
Figura 42: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo de más de 3500 kg. PMA código postal 08198	154
Figura 43: Grafica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08197	163
Figura 44: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08197	164
Figura 45: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08197	164
Figura 46: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08197	165
Figura 47: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08197	165
Figura 48: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08197	165
Figura 49: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08197	166
Figura 50: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08197	166

Figura 51: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08197	167
Figura 52: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08197	167
Figura 53: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08197	168
Figura 54: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08197	168
Figura 55: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08197	176
Figura 56: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08197	181
Figura 57: Grafica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08196	192
Figura 58: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08196	193
Figura 59: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08196	193
Figura 60: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08196	194
Figura 61: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08196	194
Figura 62: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08196	194
Figura 63: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08196	195
Figura 64: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08196	195
Figura 65: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08196	196
Figura 66: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08196	196
Figura 67: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08196	197
Figura 68: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08196	197
Figura 69: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08196	204
Figura 70: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08196	209
Figura 71: Grafica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08195	220
Figura 72: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08195	221
Figura 73: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08195	221
Figura 74: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08195	222
Figura 75: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08195	222
Figura 76: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08195	222
Figura 77: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08195	223
Figura 78: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08195	223
Figura 79: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08195	224
Figura 80: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08195	224
Figura 81: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08195	225
Figura 82: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08195	225

Figura 83: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08195	233
Figura 84: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08195	237
Figura 85: Grafica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08174	249
Figura 86: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08174	250
Figura 87: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08174	250
Figura 88: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08174	251
Figura 89: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08174	251
Figura 91: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08174	252
Figura 92: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08174	252
Figura 93: Grafica de autocorrelación de número de recogidas por día vehículo 1500 kg. PMA código 08174	253
Figura 94: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08174	253
Figura 95: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08174	254
Figura 96: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08174	254
Figura 97: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08174	254
Figura 98: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08174	262
Figura 99: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08174	267
Figura 100: Grafica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08173	278
Figura 101: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08173	278
Figura 102: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08173	279
Figura 103: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08173	279
Figura 104: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08173	280
Figura 105: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08173	280
Figura 106: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08173	281
Figura 107: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08173	281
Figura 108: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08173	282
Figura 109: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08173	282
Figura 110: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08173	283
Figura 111: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08173	283
Figura 112: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08173	291
Figura 113: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08173	296
Figura 114: Grafica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08172	308
Figura 115: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08172	309

Figura 116: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08172	309
Figura 117: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08172	310
Figura 118: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08172	310
Figura 119: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08172	310
Figura 120: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08172	311
Figura 121: Grafica de autocorrelación de número de recogidas por día vehículo 1500 kg. PMA código 08172	311
Figura 122: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08172	312
Figura 123: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08172	312
Figura 124: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08172	313
Figura 125: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08172	313
Figura 126: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08172	321
Figura 127: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08172	325

INTRODUCTION

1. Problem statement

One of the main factors that affect a parcel logistics company is the randomness of the goods it has to distribute. This randomness has a big effect on all the sections of the parcel logistics company, but the greatest is on the capillary distribution network.

The nature of the merchandise given by different clients of a logistics company can be completely different. Even the nature of the merchandise of the same client on different days can differ considerably. For instance, a client can have unique single shipment of 1000 kg. for a particular destination on one day, it is possible to have 100 shipments of 1 kg. each, for 75 different destinations on the next day. In an overnight parcel logistics company, this dynamic environment has more considerable effect, because all the merchandise the clients consign has to be collected on one day and delivered on the following day.

The environment where a parcel logistics company has to work is dynamic in the short term, as it has been showed, but is dynamic in the long term as well. The number of clients that a logistics company has is changing continuously and the battle for gaining new clients between the different logistics companies is tough. In the same way the logistics company gains new clients due to its commercial activity, it also loses some because of the commercial activity of the competition. From one year to the next the locations of clients can vary, being more concentrated or more dispersed.

The parcel logistics company's managers have to design a capillary distribution network to collect all the merchandise their clients give them to distribute, and, at the same, this capillary distribution network is the one that has to deliver all the shipments originated from the collected merchandise. To reach this goal, the managers have to decide the type of vehicles (van, lorry or trailer) and the number of each type of vehicle the capillary distribution network needs to cover the area around each hub efficiently.

Of course, this number of vehicle has to satisfy the delivery/collection quality requirements with the minimum cost to guarantee the maximum profit for the logistics company.

To design this network, the strongest and most powerful tool the managers have is reliable and accurate information about the nature of the goods needed to be collected and delivered, and consequently, about the randomness of this merchandise. The way this information can be obtained is through a model. For an external observer, a model of a system is an object that the observer can use to answer any question about the system he is interested in (Minsky, 1965).

So, once a model is created, it can be used to represent the behaviour of the real system. In the design of a capillary distribution network, the information needed to create the most profitable network is the number of deliveries and collections for each area, and the weight of the different deliveries and collections, i.e. the workload of each vehicle. From the number of deliveries and collections, a manager will know the number of vehicles he needs for each area, and depending on the weight of the collections and the deliveries, the type of vehicles that are needed.

This workload is fundamental for another critical aspect of the parcel logistics company; the delivery and collection cost. These costs depend on the workload. In parcel logistics companies these costs are as follows: the parcel logistics company has to pay a specific amount of money to the vehicle in charge of the collection, one part of which is due to pick up the shipment, the other which depends on the weight of the shipment. The delivery cost is analogous; the parcel logistics company has to pay a specific amount of money to the vehicle in charge of the delivery, which is divided into two parts; a fixed amount for serving the shipment, and another one based on the weight of the shipment.

For the design of a capillary distribution network it is important to consider that these delivery and collection costs are the income for the drivers of the capillary distribution networks. If the income of a driver is not high enough, the continuity of this driver cannot be guaranteed. A continued change of drivers is not advisable for a parcel logistics company due to the adaptation time the drivers need to learn each distribution area and clients' necessities. Therefore, the parcel's logistics company's managers have to find a balance between the minimum distribution costs and the continuity of the drivers.

To find this balance any capillary network developed has to be based on guarantee a minimum income thus to ensure the continuity of the drivers. This continuity maximises driver familiarity within their distribution area. With increased familiarity, driver performance improves due to ease in finding addresses and locations as well as efficiency in organising daily routes. Their capacity to make deliveries and collections increases, and therefore, so does their productivity. For this the workload and the driver's income in each distribution area has to be predicted (Zhong et al, 2004).

Due to the peculiarity of the delivery cost (like the collection cost, because they are analogous), the relationship between this cost and the number of the deliveries and their weight cannot be predicted exactly. The model created to predict the behaviour of these costs has to be able to deal with their specific characteristics of uncertainty. Because of the uncertainty and random components, a stochastic model is required (Guasch et al, 2002).

In the literature related to logistics companies, regardless of the nature of the goods they distribute or the delivery time they offer to their clients, most of the models are created by deterministic

methods. Mostly, all the aspects of logistics companies have been modelled by deterministic models, e.g. numbers of vehicles, route selection, number of hubs and their locations and truck loading times and flows (cross docking).

Stochastic methods have not been used extensively in researches related to logistics companies. Only demand models and customer models have been done using stochastic modelling (Gendreau et al, 1996). In most cases, these stochastic models have been used as a complement of the deterministic models, adding certain stochastic characteristics. An example of this is the Vehicle Routing Problem (VRP), which was renamed the Stochastic Vehicle Routing Problem (SVRP) when certain stochastic parameters were included.

In this work, the workload of each vehicle is predicted by the adjusted of the variables by probability distribution functions. However, the way the drivers' incomes model of each distribution areas is obtained is by regression analysis. The regression analysis is the most used technique nowadays and because of its multiple uses, regression applications are numerous and there are in almost any field (Montgomery et al, 2004). In the reviewed literature, cases of capillary distribution networks modelled by regression analysis have not been found.

Once the model has been created, a simulation will be run with it to observe its behaviour and infer the conclusions about the system. A simulation is the process of design of a series of experiments with a computerized model, with the aim of understanding the system and evaluating different strategies that the system can work with (Shannon, 1988).

2. Research hypothesis

The research hypothesis for this PhD is if a stochastic model is a valid tool to design a parcel logistics company's capillary distribution network.

In particular, if the workload of each vehicle can be design by the adjustment of the different variables by a probability distribution function, and if the income of the drivers can be adjusted by regression analysis.

3. Justification for research

The tool developed in this research is created by using stochastic methods. This tool allows the managers of an overnight capillary distribution network to design a new distribution network, but also, to study and optimize an already existing network.

Most of the distribution networks have a stochastic behaviour by nature. However, due to the difficulty of solving these kind of problems, the models used to predict the performance of these networks are formulated in such a way that they can eventually be reduced to a deterministic problem (Laporte, 1992), or have been used as a complement of the deterministic models, adding certain stochastic characteristics (Gendreau et al, 1996).

In this research, the system will not be simplified, and the stochastic model created will be able to predict the costs and the workload related to each distribution area with better accuracy.

Moreover, the stochastic techniques presented in this research, the adjustment of variables by probability distribution functions and the regression analysis, are simpler than the deterministic ones, and can be found in commercial software. This feature makes its introduction in companies more affordable and faster.

This tool can be used not only for the design of a new capillary distribution network, but also to describe and to analyse a previously created capillary distribution network. Analysing a capillary network with this tool gives the opportunity of detecting if the costs of the different distribution areas are equilibrated and well balanced, and if there is an area with not enough income.

Another important characteristic of this tool is the help it provides to advance solutions to any change that happens in the network. As was said previously in the problem statement, parcel logistics companies are gaining and losing clients constantly. When an alteration occurs in the network, either by gaining new clients in an area or by losing some in a particular area, managers can know how the workload will be affected and modify the operational planning if needed.

Because delivery and collection costs are modelled, managers can investigate the effect of modifying the charges for services in the capillary distribution network. The managers can predict the savings for the parcel logistics company in case they decide to reduce them; or they can advance how much the costs can increase if they have to pay more for them.

4. Objectives

In the topic of this research, which is overnight parcel logistics companies, the research made is less common, and most of the research about modelling capillary distribution networks has been made by deterministic methods.

Therefore, the main objectives of this study are, firstly, to research about overnight parcel logistics companies due to the lack of information and secondly, to develop a stochastic tool to demonstrate that stochastic techniques are appropriate techniques to model a parcel logistics company's capillary distribution network.

This second objective is the principal aims of this research. This research is a useful tool to help managers of parcel logistics companies in the decision making of how to design a new capillary distribution network. By using this tool, managers can gain accurate information about the workload in each distribution area, which helps in the decision making about the resources needed and how to organize them among the different distribution areas. Also, as costs are modelled by this stochastic tool, after running a simulation, the delivery and collection costs of this new capillary distribution network is at hand for the managers, and they can see if there is going to be any continuity problem with any vehicle driver.

5. Scope

The scope of this research is to create and develop a stochastic tool that can model an overnight parcel logistics company's capillary distribution network and help in the decision making related to the design of a capillary distribution network.

To achieve this aim, a case of study is analysed in this research. The case of study for this research is the capillary distribution network in a town called Sant Cugat del Valles, belonging to a hub of a parcel logistics company located in Barberá del Vallés, in Barcelona's metropolitan area.

The whole area this hub has to cover can be divided into smaller sub-areas. These sub-areas are postcodes. In the area the hub has to cover there are also several towns. It is possible that a town is compounded by more than one post code, or conversely, two small towns share the same post code.

In this research a town, called Sant Cugat del Vallés, compound of seven postcodes will be studied. This town has 83.337 inhabitants and is located to eleven kilometres from the hub.

Through the stochastic tool developed in this research the workload and drivers' income of every postcode is analysed and modelled. With this information it will be possible to determine if each postcode can be assign to a driver in order to guarantee a minimum income for him, and therefore, , his continuity. It can be also possible that the income of the driver in a postcode is not high enough to ensure his continuity. In these cases, more than one postcode can be assign to the same driver. By contrast, if the

income and the workload in a postcode are excessive, more than one driver can be assigned to the same postcode.

In this way, it will be possible to determine if a postcode can be considered a distribution area for a driver, or conversely, more than one postcode will be needed to establish a distribution area.

In this research the main variables that affect a capillary network, the relationship between them and the grade of each one affects the system will be identified. However, the scope of this research will be never to optimize this capillary distribution network.

The model created, as the conclusions drawn in this research, can be used as a starting point of a possible research about optimization of a capillary distribution network. It is not intended that the obtained model for the case of study to be extrapolated to other geographical areas, or other similar companies. What is intended is to check the good performance of the developed tool, and use it in different geographical areas, or companies.

6. Methodology

In this section a brief summary is illustrated about the methodology followed during the development of the stochastic model (Figure 1). The methodology followed in this research, allows the managers of an overnight capillary distribution network to create a stochastic tool through which it is possible to design a new distribution network, but also, to study and optimize an already existing network. All the methodology will be explained more deeply in chapter four of this thesis.

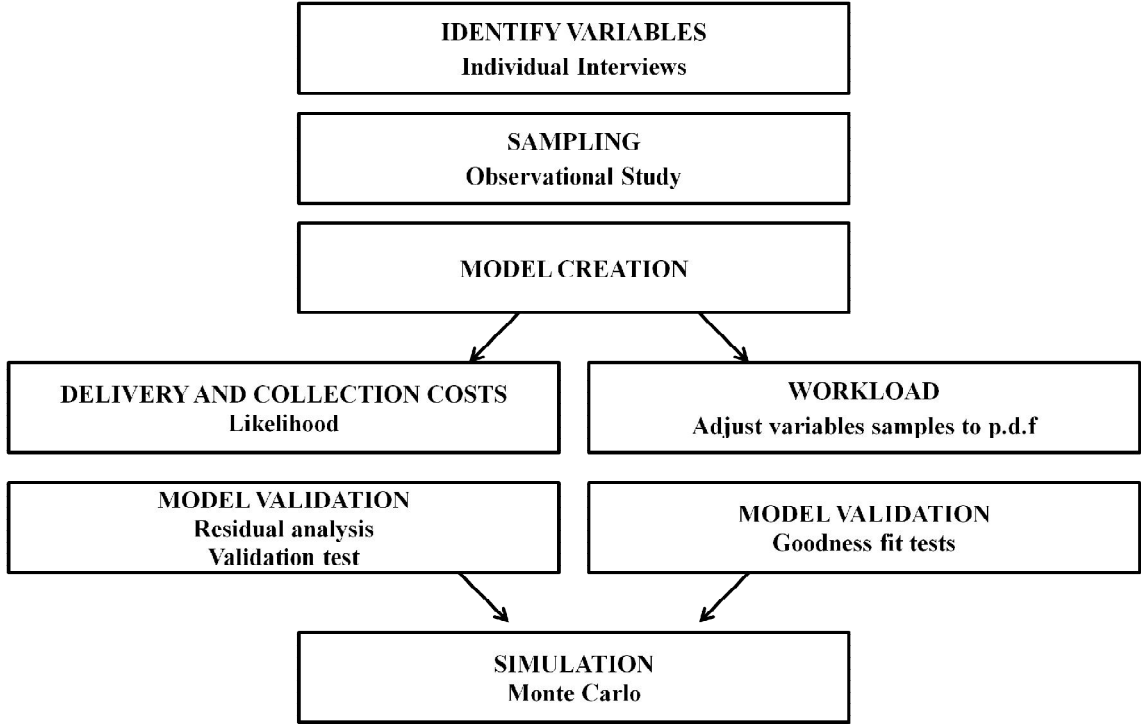


Figure 1: Summary of the methodology followed during the development of the stochastic model.

This stochastic model will determine the workload and the distribution costs of each postcode, to know if each one can be considered a distribution area.

The capillary distribution network that can be developed through this stochastic model will be based on assigning a distribution area to a vehicle as opposed to a determined route of collections and deliveries for each vehicle, as has been done until now in the vehicle routing problem.

The first step in the creation of the model is the identification of the variables that affect the capillary distribution network. For this purpose, individual interviews are made with different people who work in the parcel logistics company’s hub. The people selected to be interviewed are people who work closely with the drivers of the vehicles of the capillary distribution network.

The people interviewed are two managers, the person in charge of assigning the deliveries to the drivers every day and the person in charge of the assignation of the collections of the drivers every day. Some drivers are also interviewed to know their point of view and contrast their opinions with those that determine the delivery and collection areas.

After analysing the interviews, and contrasting the different opinions, the main variables that are needed to model a capillary distribution network are identified. The next step is sampling of every variable. The sampling is obtained from the computer system of the parcel logistics company, therefore,

the sampling is going to be an observational study, where the system is not affected by the sampling (Shannon, 1988).

Once the sampling of all variables is done, the information needed to construct the model is available. The model is compounded of two main parts: the first one which models the delivery and collection costs of the capillary distribution network; and the second one which models the workload.

6.1 The delivery and collection costs model

This model is created by linear regression. Linear regression is an appropriate technique to determine the nature in the relation of two or more quantities (Navidi, 2006), or two or more variables. This model gives the delivery cost of each distribution area per day, knowing the amount of deliveries and the total delivered weight in that area in a day. As for collection cost, the model works in an analogous way, predicting that cost knowing the number of collections made in a day and the total collected weight. The values of the parameters which show the relationship between the different variables to predict the output of the model are obtained by maximum likelihood.

Maximum likelihood enables the modeller to develop varied regression models under different assumptions, and compare them. In this research the assumptions made are that the residuals of the regression model can be distributed by different probability distribution functions. Also, the exogenous variables are transformed under the generalized linear model criteria and the regression models obtained are compared with the previous ones.

The model is analyzed to show how good it is, and ensure its adequacy. With this aim there are some performance measures that can be used to compare two different models, as adjusted R^2 . But to ensure the goodness of the model and that a different sample does not lead to a completely different model, there are some diagnostic methods based on the study of the residuals (Montgomery, et al, 2004). Finally, to determine the model is an accurate representation of the real system, the system and the model are compared with historical system input data and with new real system data (Law and Kelton, 1991).

6.2 Workload model

This model is created by probability distribution functions. First the number per day of deliveries/collections is fitted. Second, the weight of each delivery/collection is fitted. Finally, the model is applied by using the inverse-transform method (Law and Kelton, 1991) to settle the numbers of deliveries/collections, and with the sum of the weight of every delivery/collection (obtained analogously) the total delivered/collected weight per day and area is determined.

To ensure the adequacy of this model, non parametric goodness of fit tests are undertaken. These tests are the Chi Squared test for discrete distributions and Kolmogorov-Smirnov test for continuous distributions.

Finally, the last step of this methodology is the simulation of the models. The workload model simulation has been explained above. As for the delivery and collection costs model simulation, the input data is the values of the number of delivery/collection and total delivered/collected weight random variables. In order to carry out a simulation using random inputs, their probability distribution functions have to be specified (Law and Kelton, 1991). And that is exactly what is done in the workload model. So this model will be the input data in the delivery and collection costs model.

CASO DE ESTUDIO

1. Introducción

En este segundo capítulo de la tesis se va a presentar el caso de estudio sobre el que se desarrollará la investigación en esta tesis, la red de distribución capilar de la población de Sant Cugat del Valles, correspondiente a una empresa de paquetería industrial situada en el área metropolitana de Barcelona. Sin embargo, previamente se analizarán ciertos aspectos relacionados con las redes de transporte y distribución de mercancía por carretera.

2. Redes de distribución

Las redes de transporte son un sistema que soluciona la necesidad inherente de proveer movimiento tanto a personas como a bienes de consumo. En un entorno industrial en el que el transporte de personas no se realiza, es necesario conectar los bienes producidos desde su punto de origen en el que se han elaborado hasta su destino final, el cliente. En este caso las redes de transporte pasan a llamarse redes de distribución y en función del modo de transporte seleccionado, las redes de distribución pueden ser aéreas, terrestres o marítimas.

3. Clasificación de las redes de distribución

En general las redes de distribución se pueden clasificar en envíos directos, desde un punto geográfico hasta otro, o en envíos indirectos, en los que la mercancía se clasifica, se almacena, se maneja y se consolida en las distintas terminales de tránsito, de carga o *hubs* (Manivannan, 1998).

El segundo tipo de redes de distribución definido por Manivannan, ha aumentado considerablemente debido a la externalización del proceso de distribución de muchas de empresas productoras. Los operadores logísticos, las empresas externas encargadas de realizar este tipo de envíos indirectos, prestan sus servicios a más de un empresa productora para consolidar la mercancía de varias de ellas en una misma ruta de transporte, reduciendo de esta manera los costes unitarios de distribución.

Los operadores logísticos se pueden clasificar siguiendo otros criterios como la tipología de servicio, número de clientes y plazo temporal de entrega (Estrada, 2007).

De acuerdo a esta clasificación el primer tipo de empresa son las empresas de paquetería industrial (Second Part Logistics, 2PL). Estas empresas ofrecen plazos de entrega muy pequeños. Si el plazo de entrega es de 24 horas o menor, estas empresas de paquetería son consideradas empresas de paquetería urgente.

El volumen de cada uno de los envíos no es muy elevado. Pueden transportar todo tipo de bienes de consumo con un empaquetado adecuado. Existen empresas especializadas en el transporte de mercancía de volumen muy reducido, incluida documentación y correspondencia conocidas como empresas de transporte courier.

Debido al reducido volumen de la mercancía transportada por estas empresas, la consolidación de la mercancía de muchos clientes es un factor fundamental para asegurar la viabilidad de las compañías.

El segundo tipo de operador logístico en esta clasificación son aquellas empresas que ofrecen los servicios de almacenaje, empaquetado y otro tipo de valor añadido conjuntamente con el servicio de transporte y distribución (Third Part Logistics, 3PL). Los volúmenes de transporte de estas empresas son mayores que en el caso de las empresas de paquetería y puede que empleen más de un modo de transporte. La consolidación de la mercancía se lleva a cabo en los modos de transporte de mayor capacidad y que realizan distancias más largas (generalmente ferrocarril y marítimo).

Por último, se han desarrollado otros operadores logísticos conocidos como Fourth Part Logistics, 4PL, en los que el operador logístico y la empresa de producción desarrollan una estrategia común, de modo que ambas comparten riesgos y beneficios en toda la cadena de suministro de los productos.

4. Rol del sector de paquetería industrial

El sector del transporte de paquetería industrial es un sector de muy rápido crecimiento en la economía mundial, y ha aportado una contribución significativa a la economía global: emplea a 1,3 millones de personas de manera directa (un tercio más que la industria del refinado de petróleo, por poner un ejemplo) y de manera indirecta a 2,75 millones; y realizó una aportación directa al producto interior bruto mundial de 80.000 millones de dólares (Oxford Economics, 2009).

Un amplio abanico de empresas emplean los servicios de las empresas de paquetería industrial, pero las empresas que principalmente contratan estos servicios son empresas cuyos productos tienen un muy alto valor añadido, como por ejemplo productos de alta tecnología, componente tecnológicos, aparatos informáticos y de telecomunicaciones, compañías farmacéuticas y empresas financieras (Oxford Economics, 2009). Una de las mayores aportaciones que hacen estas empresas de este sector, además de permitir que empresas de tamaño no muy grande puedan competir de manera eficaz en un mercado cada vez mayor y más globalizado, es aumentar la competitividad de las empresas productoras al permitirles añadir valor a sus productos, como reducción de inventarios, reducción del plazo de entrega al cliente final, mayor costumización de los productos y aporte de tecnología (Dennis, 2011).

Cuatro compañías, DHL, FedEx, TNT y UPS, son las empresas líderes a nivel mundial (Oxford Economics, 2009). En el caso del estado español, a diferencia de la mayoría de países, el sector de la paquetería industrial está dominado por empresas nacionales (Martínez y Martínez, 2007), aunque esto no es óbice para que empresas internacionales no tengan presencia en este sector tan competitivo.

5. Estructura de una empresa de paquetería industrial

La estructura de las empresas de paquetería industrial más importantes, o de mayor tamaño, está compuesta por terminales de carga o *hubs* y centros locales de distribución para la distribución capilar (Dennis, 2011). A modo de ejemplo la estructura de DHL en Alemania está compuesta por 33 centros de carga o hubs, los cuales están interconectados. Desde estos *hubs*, los paquetes son distribuidos a los centros de distribución local (Zustellbasen), de los cuales hay 256 (Sodiper, 2011).

Este tipo de estructura es debido a que las empresas de paquetería industrial han de consolidar la mercancía que los diferentes clientes le entregan para su distribución para poder garantizar así una productividad mínima. La consolidación consiste en agrupar la mercancía con destino común en vehículos de mayor capacidad para de esta manera reducir el coste unitario por envío.

Debido a esta necesidad de consolidación las empresas de paquetería industrial se han visto obligadas a jerarquizar sus instalaciones en una serie de almacenes, delegaciones y terminales de consolidación de tamaños y características diferentes entre sí.

Esta jerarquización permite la clasificación de la red de transporte de una empresa de paquetería industrial en dos grupos (Crainic, 2003). Un ejemplo simplificado de una empresa de paquetería industrial se muestra en la Figura 2:

- La red troncal o entre terminales y centros locales de distribución, que se compone de aquellas rutas y vehículos que comunican únicamente las distintas terminales entre sí, sin servir directamente a los clientes. Estos vehículos son vehículos de gran capacidad que permiten la consolidación de la mercancía y posibilitan economías de escala. Las rutas realizadas por estos vehículos son rutas largas con un número de paradas reducido. Las empresas más importantes de distribución de paquetería industrial, habitualmente subcontratan este tipo de red troncal (Dennis, 2011).
- La red capilar, que está compuesta por las rutas y vehículos que realizan la distribución en los cascos urbanos de cada uno de los municipios que se encuentran dentro del área de alcance de cada centro local de distribución. Ocasionalmente, los terminales de carga también pueden poseer sus propias redes capilares. Los vehículos que se encargan de efectuar el reparto de la mercancía desde los centros a los clientes finales (en el caso de las recogidas, las rutas son de sentido inverso) son de capacidad menor a los de la red troncal. Las rutas de estos vehículos

se circunscriben en el área de influencia del centro correspondiente al que pertenecen y el número de paradas que realizan es muy alto.

En función del número de habitantes y de la extensión de cada municipio, el número de vehículos asignados a cada población puede ser uno o más. Puede darse el caso contrario también, en el que más de un municipio de población muy reducida y que no estén muy distantes unos de otros sean asignados al mismo vehículo de la red capilar.

Al contrario que en las redes troncales, las redes capilares suelen ser propias de las empresas de distribución de paquetería (Dennis, 2011).

Por lo tanto, cuando a una empresa de paquetería se le asigna el envío de una mercancía desde un cliente origen (cliente *o*) a otro cliente destino (cliente *d*) el proceso es el siguiente:

En la primera fase del envío, un vehículo de la red capilar recogerá el envío en el domicilio del cliente *o* y lo transportará al centro de distribución local *O* que el cliente tiene asociado. En el centro la mercancía será etiquetada, separada y clasificada en función del destino en el que tenga que ser entregada.

En la segunda fase, todos los envíos que en esa delegación haya para un mismo destino serán consolidados en un vehículo de la red troncal. El envío de este ejemplo será cargado en el vehículo de la red troncal con destino al centro de distribución local *D*. Este vehículo puede que realice un viaje directo al centro *D*, o puede que haga alguna parada en un terminal de carga o *hub* intermedio en la que se consolidará la mercancía de más centros locales de distribución con diferentes orígenes, pero con destino común *D*.

En la tercera fase, un vehículo de la red capilar del centro local de distribución *D* transportará el envío desde el centro hasta el domicilio del cliente *d*.

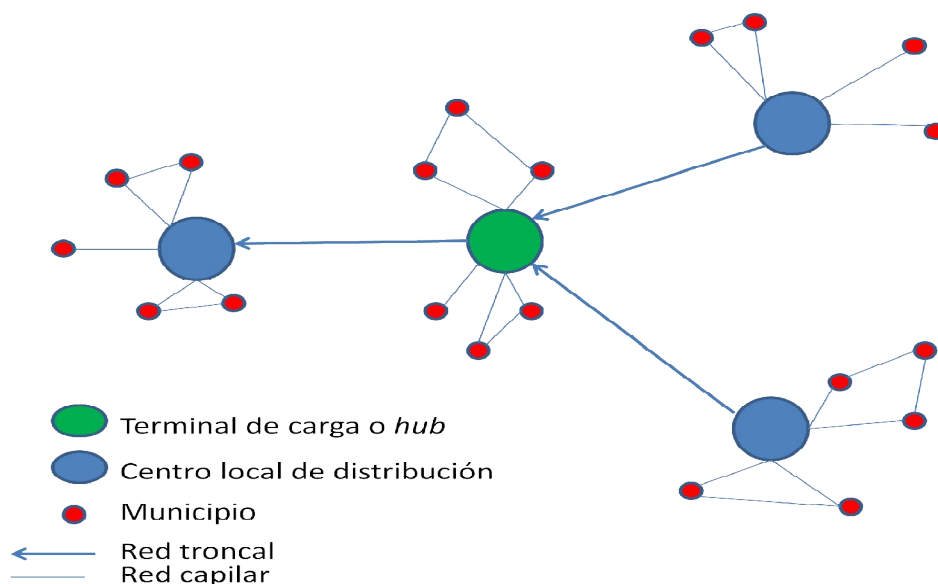


Figura 2: ejemplo simplificado de una empresa de transporte

Los costes acumulados derivados del transporte de este envío para la empresa de paquetería industrial se pueden observar en la siguiente Figura 3 (Estrada, 2007). Los costes derivados de la red troncal son menores que los de la red capilar, ya que debido a la consolidación de envíos los costes unitarios se reducen.

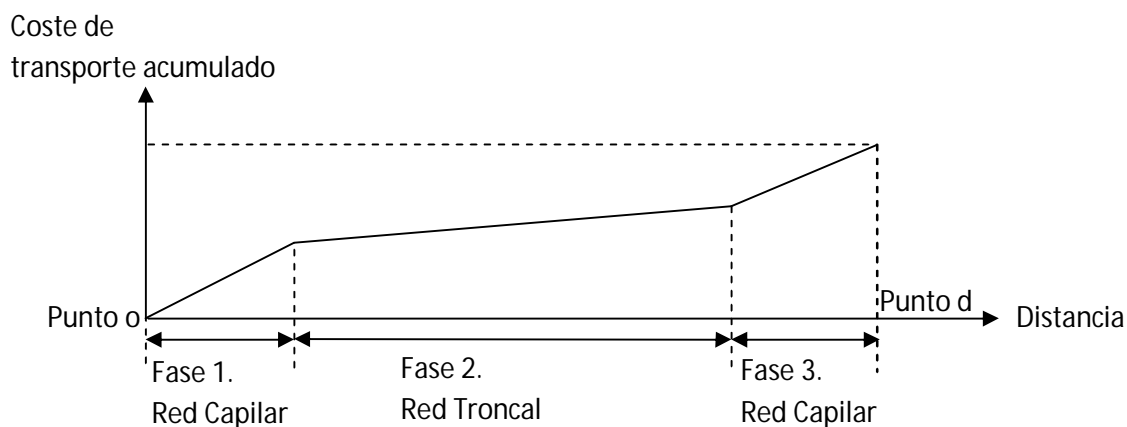


Figura 3. Costes acumulados de un envío en una empresa de paquetería industrial

6. Descripción del caso de estudio

La red de distribución capilar que se modelará en este trabajo de investigación es la que una empresa de paquetería industrial situada en Barberá del Vallés, en la zona metropolitana de la ciudad de Barcelona (España), tiene asignada en el municipio de Sant Cugat del Valles (Provincia de Barcelona).

El municipio de Sant Cugat se dividirá en siete zonas, y se estudiará cada una de ellas de manera individual, teniendo en cuenta solamente las características propias de cada zona. Una vez modelada cada

zona, se tomarán las decisiones oportunas para cada una de las zonas, o para el conjunto de algunas de ellas.

Con este trabajo de modelado, no se pretende que el modelo obtenido para este caso de estudio se extrapole a otras zonas geográficas, u otras compañías similares. Lo que se pretende es comprobar el buen funcionamiento de la metodología empleada, y utilizar ésta para modelar la red de distribución de otras áreas geográficas diferentes a la de Sant Cugat en esta misma empresa, o incluso, extenderse a otras empresas. Esta técnica de dividir el área a cubrir por una red de distribución capilar en sub-áreas más pequeñas, y estudiar cada una de estas sub-áreas de manera particular, es una práctica habitual. En el estudio de cómo organizar el la distribución logística en la ciudad de Potenza (Basilicata, Italia), la superficie de la ciudad se dividió en 33 sub-áreas, y se modeló cada una de ellas (Frosini, 2005). Un estudio similar, de cómo organizar la distribución capilar de mercancía en la ciudad de Zaragoza, dividió ésta en siete zonas diferentes, y estudió cada una de ellas de manera independiente, y analizó diferentes soluciones para cada una de las zonas (Val et al, 2007). Y un caso extremo de esta práctica es el caso de DHL en la ciudad de Bonn, donde el estudio, y el diseño de la red de distribución capilar se ha dado hasta por calles (Sodiper, 2011).

6.1. Características de la red capilar de la empresa de paquetería industrial del caso de estudio

Los plazos de entrega que esta empresa de paquetería tiene contratados con sus clientes son de 24 horas, por lo que se considera una empresa de transporte urgente. Esta empresa cuenta en el centro local de distribución de Barberá del Vallés con 42 vehículos de reparto/recogida en su red capilar, 8 vehículos de gran tonelaje los cuales unen directamente este centro local de distribución con otros de la misma compañía, y otros 6 vehículos de gran tonelaje los cuales transportan mercancía a un terminal de carga que realiza labores de consolidación de mercancía de varios centros de distribución local.

La tipología de la mercancía que esta empresa de paquetería industrial distribuye, principalmente son envíos de peso reducido, en el intervalo de 1 a 40 kg. Esto no es óbice para que no distribuya envíos de mayor peso, pudiendo realizar envíos de hasta 3500 kg. La paquetería está compuesta por cajas de cartón de 6 caras, con caras paralelas dos a dos; mientras, los envíos de mayor peso están paletizados y compuestos por varias cajas formando un único bulto cada palet, o por una caja por palet.

La forma en que las empresas de paquetería industrial reparten su carga de trabajo entre los diferentes vehículos de la red capilar es diferente de una a otra, en función de cómo crea cada una que es la mejor forma de rentabilizar su negocio. En el caso particular de esta empresa de paquetería, los

vehículos de reparto tienen asignados uno o varios códigos postales. Un ejemplo de cómo está distribuida la zona metropolitana de Barcelona en códigos postales se puede observar en la Figura 4.

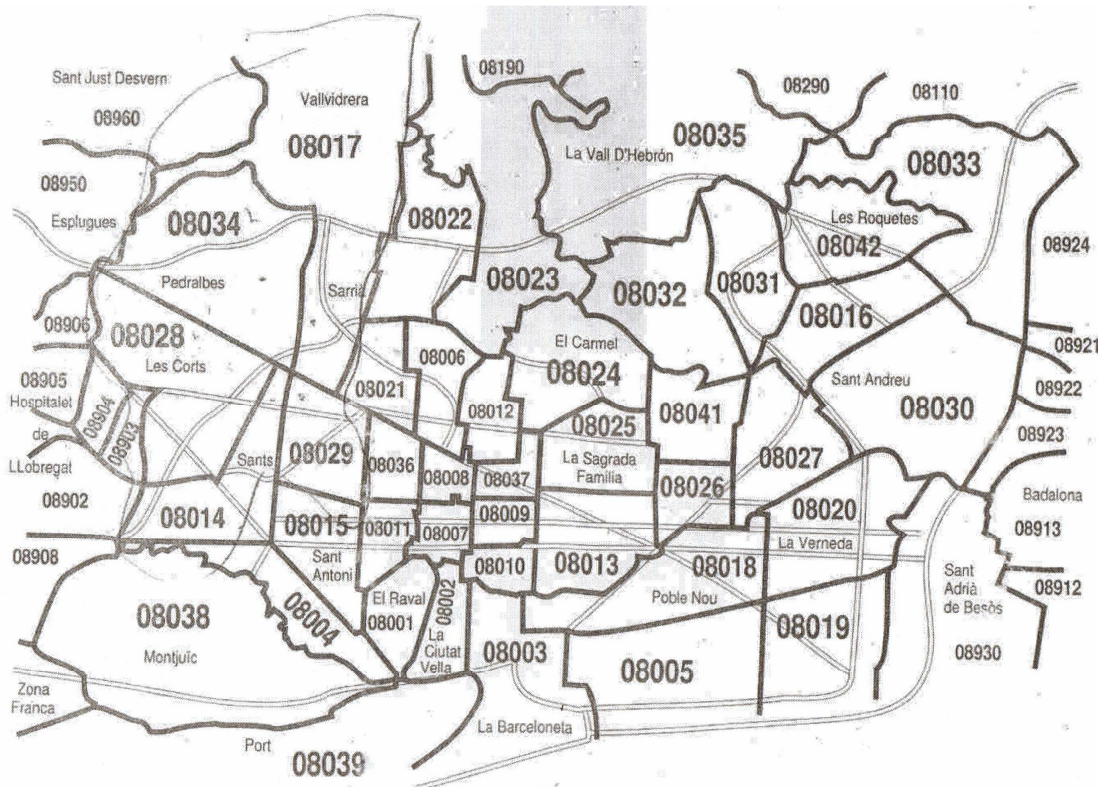


Figura 4: Distribución de códigos postales de la ciudad de Barcelona

En función de la carga de trabajo que exista en cada uno de los códigos postales, cada una de las zonas es repartida entre los diferentes vehículos. Si en un código postal la carga de trabajo es superior a la que un único vehículo es capaz de absorber para garantizar unos estándares mínimos de calidad de entrega y recogida, este código postal se dividirá en las partes que sean necesarias para una distribución eficiente, asignando un vehículo a cada parte. Los factores que determinan que en una zona haya más de un vehículo son el número de repartos a distribuir, el número de recogidas a realizar y el peso de estos repartos y recogidas.

El número de repartos es un factor importante a analizar, ya que al ser una empresa con plazo de entrega de 24 horas tiene el compromiso de entregar todos los repartos recepcionados en la jornada anterior. Por tanto, su red capilar debe de estar diseñada de tal manera que todos de los envíos que recibe puedan ser cargados en los vehículos de reparto y se intente la entrega de los mismos en la jornada siguiente a su recepción.

El número de recogidas es otro factor a considerar. Es necesario el disponer de vehículos con capacidad suficiente para recoger todas las recogidas por dos motivos:

- i) generar la confianza en los clientes asegurando que siempre que lo necesiten tendrán sus envíos en los lugares de destino cuando sus clientes lo precisen. Para ello se ha de disponer del número de vehículos con capacidad suficiente para que en ningún momento la mercancía de ningún cliente se quede sin recoger.
- ii) las empresas de transporte ingresan una cierta cantidad de dinero en función de la mercancía que transportan. Muchas de empresas productoras o manufactureras tienen más de un proveedor de transporte. Si en un momento dado una recogida no se realiza, es muy probable que la mercancía que tenían previsto entregar a esta empresa de paquetería se la asignen a otro proveedor con el que trabajan, sin esperar al día siguiente para que se recoja. Este tipo de casos son ingresos que se pierden y que va directamente a una empresa de competencia directa.

Por último, cada uno de los vehículos que circulan por la red viaria tienen limitado por ley el peso máximo que pueden transportar, por motivos de seguridad. En función del tipo de vehículo y las características del mismo el peso máximo a transportar varía. Los tipos de vehículos que esta empresa transporte de paquetería dispone son tres:

- Furgoneta normal o furgoneta carrozada. Estos dos tipos de vehículos pueden transportar hasta 1500 kg. de peso máximo autorizado.
- Camión. Este tipo de vehículos pueden transportar hasta 3500 kg. de peso máximo autorizado.
- Trailer o camión articulado. Este tipo de vehículo puede transportar hasta 24000 kg.

Existen otros tipos de vehículos con peso máximos autorizados distintos a los que se han expuesto, pero en la actualidad esta empresa no dispone de ellos. Si tras el estudio de las cargas de trabajo se observa que otros vehículos son necesarios, cabe la posibilidad de cambiar alguno de los existentes por otros más adecuados a cada situación.

Este peso máximo autorizado limita el número de repartos y recogidas a realizar, ya que el peso de los repartos y recogidas que los vehículos lleven cargados no podrá superar estos valores máximos. Para contrarrestar esta posible falta de capacidad, esta empresa de paquetería industrial, a nivel particular, obliga a todos sus vehículos a regresar a la terminal a mitad de la jornada laboral. De este modo, los vehículos podrán ser descargados de las primeras recogidas hechas durante la mañana, y de los repartos que no han podido ser realizados. De esta manera, disponen de toda la capacidad de sus vehículos durante la tarde, que es cuando se realizan el grueso de las recogidas.

Por otro lado, anteriormente se ha comentado el caso en el que la carga de trabajo de un código postal sea suficientemente grande como para que dos vehículos compartan un mismo código postal. Sin embargo, puede darse el caso contrario en el que la carga de trabajo en un código postal no sea tan elevada, incluso que sea demasiado pequeña para un único vehículo. Como se explicará más adelante con más detalle, los vehículos que trabajan en esta empresa de paquetería trabajan en régimen de autónomo y a producción. Es decir, no son trabajadores asalariados o por cuenta ajena, los cuales perciben unos ingresos por realizar sus labores de reparto y recogida, sino que sus ingresos dependen del número de recogidas y repartos que hagan (más adelante se explicará con más detalle como los ingresos también depende del peso de cada uno de estos repartos y recogidas). Por lo tanto, el que en un código postal no haya una carga de trabajo elevada y que uno de los vehículos esté ocioso es un inconveniente para la empresa de transporte. Si el conductor del vehículo no ingresa lo suficiente como para subsanar todos los gastos que le origina el vehículo (gasóleo, mantenimiento, amortización del vehículo) y asegurarse unos ingresos mínimos para su propia manutención, el conductor dejará de trabajar con la empresa transporte de paquetería.

El cambiar de chofer en una zona de reparto presenta varias desventajas. La primera es que a un chofer nuevo le es necesario un periodo de adaptación para conocer el funcionamiento de una empresa nueva, y lo más importante conocer las características y particularidades de una nueva zona. Debe conocer las nuevas calles y carreteras, las ubicaciones de los clientes habituales, sus horarios de recogida y reparto, y lo que cada uno de los clientes espera del repartidor. Todo este proceso de adaptación hace que su eficacia disminuya y que la calidad del servicio ofrecido por parte de la empresa de paquetería merme, pudiendo darse el caso de que todos los repartos no se hagan en el plazo previsto o que alguna recogida no se realice en tiempo y forma, con la consiguiente posibilidad de pérdida de algún cliente.

La segunda es la pérdida de imagen que la empresa de paquetería industrial muestra ante sus clientes. Si los clientes ven que se cambia mucho de chofer, lo interpretan como que la empresa a la que pagan por unos servicios no es tan seria y competente como consideraban y puede afectar a las relaciones comerciales.

Por lo tanto, cuando la carga de trabajo de un código postal no es lo suficientemente alta como para garantizar unos ingresos mínimos para el conductor de un vehículo, lo que se hace es asignar más de un código postal al mismo vehículo. Evidentemente, estos códigos postales han de ser contiguos, o muy cercanos entre sí, para evitar desplazamientos innecesarios y excesivos de los vehículos que resten productividad a los mismos.

En definitiva, la carga de trabajo de cada uno de los códigos postales ha de ser estudiada para asegurar unos estándares de calidad mínimos del servicio de recogida y reparto, y por otro lado, garantizar unos ingresos mínimos para cada uno de los conductores que permitan su continuidad.

6.2. Costes de reparto y recogida de la empresa de paquetería del caso de estudio

Cada empresa de transporte de paquetería tiene a los conductores de los vehículos de la red capilar en régimen que cree más conveniente para sus intereses. En este caso particular, el régimen en el que trabajan los conductores es en régimen de autónomo y a producción.

Este régimen presenta una serie de ventajas para la empresa de paquetería. Al ser un régimen de autónomo la empresa no debe realizar inversiones en vehículos de la red capilar, ya que estos son propiedad de los conductores, y los gastos de mantenimiento y de amortización de los mismos corren a cargo de los conductores. Por otro lado, los conductores al no ser trabajadores por cuenta ajena o asalariados, la vinculación contractual de los mismos puede ser rescindida sin los gastos derivados de este tipo de relaciones laborales.

Como contrapartida, a la hora de asignar unos ingresos mínimos a los conductores se ha de tener en cuenta los gastos que tienen, y los ingresos deben de ser tales que puedan sufragar los gastos y les queden algunos beneficios.

El régimen de trabajo a producción implica que los conductores no ingresan una cantidad fija por acudir al puesto de trabajo y realizar las funciones que se les asigna, sino que sus ingresos dependen del número de repartos y recogidas que hagan. A mayor número de repartos y recogidas que hagan, mayores serán sus ingresos. Se explicará el caso de los ingresos de un conductor por cada reparto, siendo análogo el caso de las recogidas.

No todos los repartos suponen el mismo ingreso para los conductores. En este caso particular, se les asigna una cantidad fija por cada reparto que realizan, y otra variable en función del peso del reparto. Las cantidades, tanto de la parte fija como de la variable, dependen a su vez del peso del envío, estando tabuladas en estratos. Un ejemplo de estas tablas pueden verse en la tabla 1.

Tabla 1. Cuadro de ingresos de reparto y recogida en función del peso

TABLAS DE COSTES DE REPARTO Y RECOGIDA					
RECOGIDA			REPARTO		
PESO	€/ RECOGIDA	€/ kg	PESO	€/ REPARTO	€/ kg
1	1,943314	0,017939	1	1,833237	0,018074
2	1,943314	0,017939	2	1,852534	0,018074
3	1,943314	0,017939	3	1,871831	0,018074
4	1,943314	0,017939	4	1,891128	0,018074
5	1,943314	0,017939	5	1,910426	0,018074
10	2,391771	0,017802	10	2,183169	0,017939
20	2,921765	0,016529	20	2,563665	0,017464
30	3,485733	0,016036	30	3,019011	0,017326
40	3,485733	0,016036	40	3,087626	0,017326
50	3,485733	0,016036	50	3,156240	0,017326
60	3,757526	0,015928	60	3,219865	0,016647
70	3,757526	0,015752	70	3,309799	0,016647
80	3,757526	0,015752	80	3,440444	0,016647
90	3,757526	0,015752	90	3,580654	0,016647
100	3,757526	0,015752	100	3,628426	0,016647
200	4,063292	0,014650	200	3,913806	0,015091
500	4,382649	0,013860	500	4,233162	0,014362
1.000	4,735977	0,013355	1.000	4,566106	0,013701
2.000	4,953410	0,013132	2.000	5,150936	0,013319
3.000	16,056537	0,007666	3.000	6,996578	0,012401

Como se puede observar en la tabla 1, por el reparto de un envío de 3 kg. de peso, al chofer se le asignará un ingreso de 1.943314 euros por realizar el reparto y 0.017939 euros por cada uno de los tres kg. de los que consta el envío. Para un reparto de 25 kg. se le asignará un ingreso de 2.921765 euros por el reparto y 0.016529 euros por cada uno de los 25 kg. de los que consta el envío.

Como se puede observar la variación de la cantidad asignada de un estrato a otro no es lineal, y las cantidades para los mismos estratos de los repartos no son los mismos que los de las recogidas (a pesar de ser la forma de calcularlos análoga). Por otro lado, se pueden también entender como los ingresos que un conductor de un vehículo de una zona de reparto obtiene, son los costes de reparto y recogida de la empresa de paquetería industrial tiene en esa zona.

6.3. Municipio de Sant Cugat del Vallés

La red de distribución capilar de la empresa de paquetería industrial del caso de estudio correspondiente al municipio de Sant Cugat del Vallés. Sant Cugat del Vallés es una población de la zona

metropolitana de Barcelona y de acuerdo al Instituto Nacional de Estadística, en el censo 2012 tenía 84.946 habitantes y una extensión de 48 km².

Este municipio está compuesto de 7 códigos postales: 08172, 08173, 08174, 08195, 08196, 08197 y 08198, siendo la distribución de los mismos la que se muestra en la Figura 5.

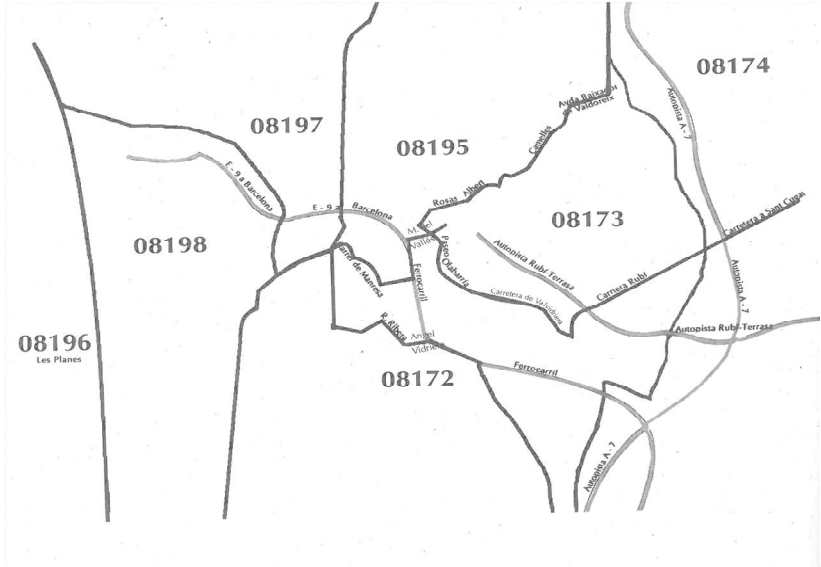


Figura 5: Distribución de códigos postales en el municipio de Sant Cugat del Vallés

Esta población se encuentra a 11,3 km. del centro de distribución local de la empresa de paquetería. Está comunicado con ella por la autopista A-7, por lo que los vehículos destinados a esta zona de reparto pueden ir de la zona de reparto a al centro de distribución en más de una ocasión, en caso de que tengan problemas de capacidad, sin que ello les repercute de manera considerable en la calidad de entrega y recogida.

Por tanto, se va a desarrollar una herramienta estocástica para cada uno de estos códigos postales la cual predecirá la carga de trabajo, tanto de reparto como de recogida, así como los costes asociados a esta carga de trabajo.

ESTADO DEL ARTE

1. Introducción

De acuerdo con la hipótesis de investigación y los objetivos enunciados en la introducción de esta tesis, se va a presentar en este tercer capítulo un estudio del estado del arte sobre los tipos de modelos existentes, así como los modelos matemáticos más empleados en la actualidad en el sector de transporte de mercancía por carretera y sus distintas técnicas de creación y elaboración.

El objetivo de este capítulo, además de describir los tipos de modelos y presentar un resumen de la literatura existente hasta este momento sobre estos temas, es mostrar los fundamentos por los cuales un modelo estocástico puede ser una técnica válida con la que crear una herramienta que facilite la toma de decisiones relacionada con una red de distribución capilar en una empresa de transporte de paquetería.

Tanto los directivos como los ingenieros de una empresa de transporte de paquetería se ven obligados a tomar una gran cantidad de decisiones a diario en sus respectivos puestos de trabajo. Estas decisiones buscan un único objetivo, ser la más adecuada. Por lo tanto, toda herramienta que ayude a tomar cualquier tipo de decisión será de gran utilidad para el trabajo habitual de las personas que operan en una empresa de este tipo.

Estas decisiones están orientadas a la búsqueda de ideales o niveles de la excelencia teniendo como objetivos básicos:

- i) buscar los mejores niveles óptimos de economicidad
- ii) incrementar la productividad y la calidad total
- iii) incrementar la rentabilidad de los sistemas.

Debido a la complejidad de los sistemas que el hombre ha desarrollado, la toma de decisiones sobre estos sistemas, es muy complicada. Por tanto, una información adecuada es la mejor arma de la que pueden disponer los encargados de dirigir las empresas, para que la toma de decisiones sea más sencilla, y por consiguiente, reducir la posibilidad de que éstas sean erróneas.

Una de las herramientas más empleada para la obtención de información es la creación de modelos que se comporten de manera análoga a los sistemas reales que se quieren analizar, para que a partir de la observación y estudio de éstos, obtener la información deseada. Una vez obtenida la información, se estudiará y se sacarán las conclusiones pertinentes, para inferir éstas al sistema real.

Por tanto, se puede decir que a través de estos modelos simulamos la realidad para conocerla mejor, y aplicar este conocimiento en la gestión de la toma de decisiones sobre el sistema.

2. Definición de modelo y sus funciones.

Antes de definir un modelo se va a definir en primer lugar que es un sistema. Según Robert Shannon *“Sistema es el grupo o conjunto de objetos unidos por alguna forma de interacción o interdependencia regular, para realizar una función específica”* (Shannon, 1988).

Se trata de una definición sencilla pero que pone de manifiesto los caracteres relevantes de lo que constituye el denominado enfoque sistémico: contemplación del todo y no de las partes aisladamente, acento en las relaciones entre las partes y consideración teleológica al tener en cuenta los propósitos u objetivos del sistema, especialmente válida para los sistemas creados por el hombre.

Una vez definido sistema, se van a analizar las definiciones que distintos autores han dado al concepto modelo para intentar conocer cuáles son las funciones que un buen modelo ha de poseer.

De acuerdo con Minsky, *“un objeto A es un modelo de un objeto B para un observador C, si C puede usar A para contestar preguntas que le interesen a C sobre B”* (Minsky, 1965).

Robert E. Shannon define un modelo de la siguiente manera: *“un modelo es una representación de un objeto, sistema, o idea, de forma diferente a la de identidad misma. Usualmente, su propósito es ayudarnos a explicar, entender o mejorar un sistema”* (Shannon, 1988).

A partir de estas definiciones se pueden deducir cuales son las dos principales funciones que un modelo debe satisfacer. La primera de ellas es la función predictiva. Mediante la simulación de los modelos se llevan a cabo experimentos con el objetivo de conocer el comportamiento del modelo bajo diferentes condiciones de trabajo, y de esta manera poder responder de manera satisfactoria a las cuestiones que el observador quiera plantear sobre el sistema, a través de inferir las conclusiones obtenidas del modelo al sistema. En definitiva, a través de los experimentos llevados a cabo con el modelo bajo ciertas condiciones de trabajo, el comportamiento del sistema debe de ser predicho bajo las mismas suposiciones con las que se llevaron a cabo los experimentos.

La segunda función es la función descriptiva. La función descriptiva de los modelos no se aprecia de manera tan clara como la predictiva, pero se deriva de ella. Para que esta función predictiva sea fiable es necesario que tanto el modelo con el que se realiza la simulación como las condiciones en las que se realizan dichos experimentos sean los correctos (Rubinstein y Melamed, 1998).

Para que estos dos factores sean correctos, es necesario conocer muy bien y en profundidad el sistema que queremos modelar y su entorno. La mejor manera de conocer tanto el sistema como su entorno es hacer una descripción profunda de ambos. Es necesario un amplio conocimiento de todos los elementos que componen el sistema, las interrelaciones entre los diferentes elementos, las variables (tanto internas como externas) que afectan a cada uno de los elementos, así como todos los aspectos relacionados con estas variables para poder entender y/o explicar el sistema, y así poder modelarlo. Sí que es cierto que en ocasiones no se desea modelar el sistema en su totalidad, sino solamente una sección del mismo. Esto no es óbice para que el conocimiento sobre esta sección no sea tan exhaustivo como el que se necesitaría para modelar el sistema en su totalidad.

3. Tipos de modelos

Existen en la literatura más de una clasificación sobre de los diferentes tipos de modelos. Pero autores como Shannon, Vincent o Guasch diferencian los tipos de modelos en tres grandes grupos: modelos mentales, modelos físicos y modelos matemáticos.

3.1. Modelos mentales

El ser humano en su día a día emplea constantemente modelos mentales. Las decisiones y acciones de cualquier ser humano no están basadas en la verdad universal, sino en imágenes mentales del estado del mundo, cómo están relacionadas dichas imágenes, y en la manera en que sus acciones influyen en el sistema.

Los modelos mentales tienen ventajas muy poderosas. El modelo mental puede tomar una gran cantidad de información, no sólo datos numéricos. Además es muy flexible, adaptándose a situaciones nuevas y modificándose conforme llega nueva información.

Sin embargo, los modelos mentales no son fácilmente entendibles para el resto de personas. Las suposiciones de un modelo mental son difíciles de expresar. Se puede dar el caso de diferencia de interpretaciones, ambigüedades y contradicciones. Y lo más peligroso, son subjetivos. Puede que no representen la realidad, sino lo que alguien supone que es la realidad.

3.2. Modelos físicos

Los modelos físicos tratan de reproducir fenómenos mediante escenarios físicos simplificados. El objetivo principal de los mismos es descubrir las leyes que gobiernan el comportamiento de los sistemas.

En un modelo físico de un sistema, los atributos del sistema son representados por ciertas mediciones, como pueden ser un voltaje o la posición de un marcador. Las actividades del sistema son el reflejo de las leyes físicas que rigen el modelo. Por ejemplo, el rango en el cual un marcador indica la corriente continua que un motor consume depende del voltaje aplicado al motor. Si el voltaje aplicado representa la velocidad del vehículo, entonces el número de vueltas que la aguja da puede considerarse la distancia que el vehículo ha recorrido. Cuanto mayor sea el voltaje (y la velocidad), mayor será el número de vueltas (y mayor será la distancia recorrida) en un tiempo dado.

Los modelos físicos se pueden clasificar, a su vez, en modelos físicos estáticos ó dinámicos. Los modelos físicos estáticos solo son capaces de mostrar los valores de los atributos del sistema cuando el sistema está en equilibrio. Este tipo de modelo es también conocido como modelo *icónico*, y se definen como aquellos que representan las propiedades relevantes del fenómeno real mediante la misma propiedad, con una simple transformación a escala. Los modelos dinámicos, a diferencia de los estáticos, sufren cambios a lo largo del tiempo como resultado de las actividades del sistema. Los modelos físicos dinámicos están basados en alguna analogía entre el sistema en estudio y algún otro sistema de distinta naturaleza. Esta analogía depende de alguna similitud subyacente en las fuerzas que gobiernan el comportamiento del sistema.

3.3. Modelos matemáticos

Se definen, de manera general, como una formulación o una ecuación que expresa las características esenciales de un sistema físico o de un proceso en términos matemáticos. Para el estudio de sistemas de ingeniería o de gestión, la gran mayoría de los modelos creados son modelos matemáticos (Law y Kelton, 1991).

Los modelos matemáticos emplean notaciones simbólicas y ecuaciones matemáticas para representar al sistema. La variable dependiente o variable endógena refleja el comportamiento o estado de un sistema y son variables que dependen del modelo; las variables independientes o exógenas representan los atributos del sistema y no están determinadas por el modelo. Estas variables independientes pueden ser controlables o no controlables. Además de las variables del modelo se pueden distinguir los parámetros. Los parámetros son cantidades que afectan a las variables endógenas, pero que a diferencia de las variables exógenas, son constantes. Las actividades del sistema se representan mediante funciones matemáticas, las cuales indican cómo se interrelacionan las variables (Kleijnen, 1974).

Una vez desarrollado el modelo matemático para poder solucionar cualquier cuestión que se desee plantear sobre el sistema, es necesario conocer que valores tomarán las variables endógenas ante ciertos valores de las variables exógenas, para poder inferir estos resultados al sistema. En función de la

metodología aplicada para resolver el modelo se puede hacer la siguiente clasificación: técnicas de solución analítica y técnicas de solución numérica.

Aplicar técnicas analíticas hace referencia al empleo del razonamiento deductivo. Los modelos analíticos utilizan ecuaciones en las que la variable cuyo valor pretende obtenerse es una función directa de otras variables en forma de una ecuación o sistema de ecuaciones clásico. De esta forma, dando valores a las variables independientes podrá calcularse, analíticamente, el valor de la variable dependiente objeto de la modelización.

En la práctica, solo ciertos tipos de ecuaciones pueden ser resueltos mediante técnicas analíticas. Por consiguiente el problema de emplear técnicas analíticas reside primero en encontrar un modelo que sea posible resolver, y segundo que el modelo se ajuste bien al sistema en estudio. Por ejemplo, las ecuaciones diferenciales lineales son posibles de resolver analíticamente. Conociendo esto, un modelo que restringe la descripción de un sistema a estas ecuaciones diferenciales puede ser resuelto analíticamente.

Los métodos numéricos son aquellos en los que se emplean técnicas o procedimientos computacionales para resolver las ecuaciones. El factor que diferencia los métodos analíticos y los métodos numéricos es que los métodos analíticos producen soluciones generales de manera directa, mientras que los métodos numéricos son técnicas iterativas, es decir, la solución obtenida en cada paso es una mejor solución que la ofrecida en el paso anterior, cuyos resultados fueron el punto de partida del siguiente paso.

En la mayoría de los casos en los que se desarrollaba un modelo matemático la primera opción con la que solucionar dicho modelo era siempre mediante técnicas analíticas. Cuando esto no era posible se recurría a procedimientos numéricos para resolver las ecuaciones del modelo. Pero a pesar de que los procedimientos numéricos se ha considerado como el último recurso, recientes avances en las metodologías de simulación, en los análisis de sensibilidad, en la optimización estocástica y la disponibilidad de nuevos y más eficaces software, han conseguido que esta técnica sea una de las herramientas más aceptadas y utilizadas en el análisis de sistemas e investigación de operaciones (Rubinstein y Melamed, 1998).

4. Definición de simulación

En la literatura no existe uniformidad a la hora de definir simulación, por lo que a continuación se exponen las definiciones que varios autores dan de ella para conocer cuales son sus principales características.

Robert E. Shannon define simulación como “*el proceso de diseñar y desarrollar un modelo computerizado de un sistema o proceso y conducir experimentos con este modelo con el propósito de entender el comportamiento del sistema o evaluar varias estrategias con las cuales se puede operar el sistema*” (Shannon, 1988).

Otros autores como H. Masiel y G. Gnugnoli definen simulación “*como una técnica numérica para realizar experimentos en una computadora digital. Estos experimentos involucran ciertos tipos de modelos matemáticos y lógicos que describen el comportamiento del sistema de negocios, económicos, sociales, biológicos, físicos o químicos a través de largos periodos de tiempo*” (Coos, 1993).

La definición J. Banks para simulación es “*la creación de una historia artificial de un sistema y la observación de ésta, para poder inferir las características operativas del sistema real que ha sido representado*” (Banks, 1998).

Por último, la definición que D. Ríos da sobre simulación es la siguiente “*la simulación consiste en construir modelos informáticos que describen la parte esencial del comportamiento de un sistema de interés, así como en diseñar y realizar experimentos con el modelo y extraer conclusiones de sus resultados para apoyar la toma de decisiones*” (Ríos, 1997).

5. Ventajas y desventajas de la creación de modelos y su simulación

En este apartado se van a enumerar ciertos aspectos favorables y otros no tanto, que la creación de un modelo y su posterior simulación pueden proporcionar.

La simulación permite chequear todos los efectos que un determinado cambio, tanto interno como externo, puede ocasionar en el sistema, sin llegar a comprometer los recursos del propio sistema (Guasch et al, 2003). Esto es un aspecto muy favorable, ya que una vez tomada la decisión de modificar el sistema, cometer una corrección del cambio realizado puede resultar muy costoso. La simulación permite testear alteraciones del modelo sin comprometer la adquisición de recursos.

La técnica de simulación puede ser empleada para experimentar con nuevas situaciones sobre las que se tiene poca o ninguna información. Se pueden explorar nuevos procedimientos operativos o métodos sin tener que experimentar con el sistema real; simplemente se aplican al modelo desarrollado y se observan cómo funcionan. Esta posibilidad de estudiar las diferentes alternativas de funcionamiento de los procesos es una de mayores ventajas que la simulación aporta (Ayyub y McCuen, 2003).

Mediante la simulación se puede “comprimir” o “expandir” el tiempo. Esto es así, debido a que en la simulación se puede acelerar o desacelerar un fenómeno de modo que se puede analizar a fondo. Se puede examinar el proceso de días en unos minutos, o poder emplear horas en estudiar lo que durante un minuto ha ocurrido en el proceso (Law y Kelton, 1991).

También se puede conocer el por qué de ciertos fenómenos, ya que, se pueden reconstruir escenarios con las condiciones particulares que se deseen y analizar el fenómeno que ocurre en ese caso (Banks, 1998).

Otra de las ventajas de la simulación proviene de su carácter descriptivo. Como ya se ha comentado anteriormente, es necesario conocer en profundidad el sistema para poder realizar un modelo de él. En ocasiones las decisiones que se toman están basadas en los pensamientos o creencias de alguna persona en función de cómo cree él que el sistema opera. El análisis y estudio previo necesario para poder crear el modelo, proporcionan el conocimiento y entendimiento necesario de cómo realmente el sistema opera, y corrobora o desmiente las creencias que se tenían.

El poder crear un modelo y hacer con él una simulación permite una observación detallada del sistema. Mediante esta observación entender mejor las interacciones entre las variables que conforman el sistema. En la mayoría de los casos estos sistemas son muy complejos, y el diagnosticar los problemas y lograr ver la importancia que estas variables realmente tienen, incrementa el conocimiento que se posee del sistema. Así, se puede entender también mejor la importancia e influencia de cualquier modificación sobre la ejecución del sistema global. En ocasiones el adquirir todo este conocimiento sobre el sistema puede ser tan valioso como los resultados obtenidos tras la simulación (Rubinstein y Melamed, 1998).

Por estas razones, a la hora de elegir entre las diferentes opciones que se presentan a la hora tomar una decisión, la simulación crea una opinión objetiva. Se seleccionan de entre todas las pruebas aquellas que mejores resultados aportan, ya que es mucho más sencillo aceptar los datos obtenidos a partir de una simulación fiable (en la cual se han modelado, testado, validado y representado visualmente los datos), que la opinión que una persona tiene sobre como el sistema debe funcionar una vez adoptados los cambios en él.

Desafortunadamente, la simulación también tiene sus desventajas. Las simulaciones siempre serán aproximaciones del comportamiento del sistema, ya que el modelo siempre será más simple que el sistema que representa (Lewis y Smith, 1979).

La creación de modelos requiere de un entrenamiento especial. Es una arte el cual se aprende con el tiempo y a través de la experiencia. Y en determinadas ocasiones puede consumir muchos recursos, tanto humanos como de tiempo, el realizar un buen modelo y una buena simulación (Banks, 1998).

Los resultados obtenidos de la simulación pueden ser en ocasiones difíciles de interpretar. La mayoría de estos resultados serán variables aleatorias (porque las entradas son también variables aleatorias), y la simulación resultante también tendrá un componente de incertidumbre.

6. Tipos de modelos matemáticos

Una vez conocidas las ventajas, y también los inconvenientes, que la simulación computacional posee, y que el uso de los modelos matemáticos de simulación se ha extendido considerablemente en los últimos años, en este apartado se van a exponer los diferentes tipos de modelos matemáticos de simulación existentes.

6.1. Modelos matemáticos estáticos/dinámicos

Al igual que el caso de los modelos físicos, los modelos de simulación matemáticos se pueden clasificar en modelos estáticos y en modelos dinámicos. Los modelos matemáticos estáticos expresan la relación entre las variables del sistema cuando el sistema está en equilibrio (no evoluciona respecto al tiempo). Si el punto de equilibrio es alterado, modificando el valor de alguna de las variables exógenas o de los parámetros, el modelo permite deducir los nuevos valores de todas las variables y parámetros, pero no indica el modo en el que esos nuevos valores son alcanzados.

En un modelo dinámico, por el contrario, alguno/s de los elementos que intervienen en la modelización no permanecen invariables, sino que se consideran como funciones del tiempo, describiendo trayectorias temporales. Por lo tanto, un modelo matemático dinámico permite deducir el cambio del valor de las variables exógenas y de los parámetros a lo largo del tiempo.

El análisis de un modelo dinámico tiene por objeto el estudio de la trayectoria temporal específica de alguno/s de sus elementos.

6.2. Modelos matemáticos deterministas/estocásticos

Otra clasificación que es necesaria hacer dentro de los modelos de simulación matemáticos, es la diferenciación entre modelos de simulación deterministas y modelos de simulación estocásticos. La característica fundamental que permite clasificar los modelos de simulación en deterministas o

estocásticos es el uso o no de variables aleatorias para formalizar la evolución del sistema que se representa (Ríos et al, 2004).

En los modelos deterministas las relaciones que rigen las variables, y las propias variables no pueden tomar valores aleatorios; todos los datos pertinentes se conocen con certeza. De esta forma un modelo se considera determinista si su nuevo estado puede ser completamente definido a partir del estado previo y de sus entradas (Guasch et al, 2003).

Un ejemplo típico de un modelo estocástico es un sistema de ecuaciones diferenciales complejo e irresoluble analíticamente que representa una reacción química.

La palabra estocástico proviene del griego *stokhos* que significa centro del blanco, aludiendo a la posibilidad de error cuando se lanzan dardos a un tablero. El hecho de que un modelo sea estocástico implica dos cosas: una sucesión ordenada y una función de probabilidad que genera la sucesión. En estos modelos se presupone que algunas variables importantes, llamadas variables aleatorias, no tendrán valores conocidos antes de que se tomen las decisiones correspondientes, y que ese desconocimiento debe ser incorporado al modelo. Por lo tanto, los modelos estocásticos (a diferencia de los deterministas), no producen simples resultados únicos, pues debido a la función de probabilidad, las aplicaciones sucesivas del mismo modelo pueden producir resultados completamente diferentes.

En los modelos estocásticos el proceso tiene una salida indefinida y aleatoria para una entrada determinada, y las relaciones funcionales entre la entrada y la salida se representan por ecuaciones matemáticas que se obtienen del análisis estadístico matemático (Shannon, 1988).

6.3. Modelos matemáticos continuos/discretos

Modelos continuos son aquellos en los que las variables pueden tomar valores en el conjunto de los números reales o en un subconjunto de él con la potencia del continuo. Es decir, que dentro de dos valores arbitrarios $[a,b]$ hay infinitos valores.

Modelos enteros o discretos, son modelos, donde sus variables sólo pueden tomar valores enteros. Por ejemplo, si una variable representa el número de personas que se necesita para cierta tarea, sólo los valores enteros tienen sentido, número personas.

A pesar de lo que pueda parecer al principio, los modelos enteros, son por mucho, más difíciles de resolver que los problemas o modelos continuos. Requieren muchísimo más recursos de máquina para resolver que los problemas continuos.

7. Tipos de modelos matemáticos deterministas

Como se podrá ver más adelante en este capítulo en el apartado aplicaciones de modelos de simulación matemáticos en el sector del transporte, la gran mayoría de modelos empleados en la simulación de los diferentes campos en el sector del transporte por carretera son modelos deterministas (Laporte, 1992), y en algunas ocasiones se le han añadido ciertos complementos estocásticos (Gendreau et al, 1996). Es por ello, que antes de analizar estas aplicaciones se va a describir los diferentes tipos de modelos deterministas y estocásticos existentes.

Según Moder y Elmaghraby los modelos deterministas se pueden clasificar de la siguiente manera:

7.1. Programación lineal y no lineal

La programación lineal es una clase especial de modelos de programación matemática que se desarrolló a partir de la Segunda Guerra Mundial para resolver cierto tipo de problemas de asignación de recursos entre distintas actividades. Las aplicaciones posteriores a una amplia variedad de problemas han sido numerosas y esto ha llevado a que los modelos de optimización lineal constituyan una herramienta básica a la hora de tomar decisiones.

En esta técnica se considera un conjunto de variables denominadas variables de decisión que representan entradas al modelo, controlables por el modelizador o decisor, y una función objetivo que será una función lineal en las variables de decisión, que representa algún criterio o meta importante. Además, existen ciertas limitaciones prácticas que constituyen las restricciones del modelo y que se representan como ecuaciones o inecuaciones lineales en las variables de decisión. La resolución de un programa o modelo lineal conlleva la obtención de valores para las variables de decisión que optimicen la función objetivo sujetas a las limitaciones debidas a las restricciones impuestas (Taha, 1978). La función objetivo especifica las estimaciones de cualquier combinación de las variables, es decir, de las decisiones tomadas por el modelizador.

Los algoritmos creados en la programación lineal se pueden resolver bien gráficamente o bien mediante alguna técnica analítica.

La facilidad con que se puede alcanzar una solución óptima en programación lineal ha convertido esta técnica en un medio atractivo para afrontar muchos problemas. Sin embargo, el mundo está lleno de sistemas que no se comportan linealmente, como pueden ser problemas de planificación en economía, redes eléctricas, ingeniería o control óptimo.

Aunque en algunas situaciones particulares la aproximación lineal de estos problemas está justificada, sobre todo motivada por la eficacia de los métodos de soluciones lineales, en otros será necesario manejar problemas de programación no lineal.

Debido a la gran cantidad de problemas matemáticos que pueden ser catalogados como no lineales, no se dispone de un único algoritmo que los resuelva todos. En su lugar, se han desarrollado algoritmos para algunas clases de programas no lineales (Taha, 1978).

7.2. Programación entera

Un problema de programación lineal en el cual algunas, o todas las variables tienen que ser números enteros no negativos, se llama un problema de programación entera (Hillier y Lieberman, 1967).

En el caso en el que todas las variables tengan que ser números enteros, se denomina problema de programación entera pura. En el caso el cual solamente algunas de las variables tienen que ser números enteros, se le llama problema de programación entera mixta (Hu, 1969).

Aunque en la realidad se dan muchas situaciones que se puede formular mediante programación entera, desafortunadamente la resolución de ésta es más complicada que la de la programación lineal.

7.3. Programación multiobjetivo

La programación multiobjetivo es otra variante de la programación lineal. En la resolución de muchos casos de la vida real es necesaria la consideración simultánea de varias funciones objetivo, que en muchas ocasiones podrían ser conflictivas, no siendo deseable o posible reducirlas a una única función objetivo y, por tanto, teniendo que tratarlas de una manera conjunta. Este enfoque más flexible y con una metodología más práctica para la modelización, resolución y análisis de los problemas reales, se denomina programación multiobjetivo, o de forma más general, decisión multicriterio (House, 1978).

Existen muchas aplicaciones en las que la modelización de sistemas reales conduce a la construcción de modelos con varios objetivos, que pueden ser totalmente o parcialmente conflictivos de manera que la mejora en algunos de ellos conlleva el empeoramiento de otros (Taha, 1978).

7.4. Modelos de redes

En el mundo real hay una gran multitud de situaciones que se pueden modelar y resolver como redes (nodos conectados por ramas). Un aspecto notable de los métodos de solución de estos problemas,

que se puede catalogar como modelos combinatorios (consistentes en determinar entre un número finito de alternativas una que optimice el valor de la función objetivo), es su complejidad computacional. La cantidad elevada de variables de decisión y restricciones requeridas hacen necesario un esfuerzo computacional elevado. Sin embargo, existen técnicas de solución que explotan la estructura especial de cada problema y que son generalmente más eficientes (Ríos et al, 2004). Entre los diferentes problemas posibles de las redes se encuentra el problema del transporte.

El denominado problema del transporte es uno de los primeros problemas que se formuló como un modelo de redes y en su día tuvo una gran motivación debido al escaso desarrollo de los medios de transporte hasta ese momento, aunque posteriormente también se han desarrollado otras aplicaciones distintas.

El citado problema consiste en determinar las cantidades que hay que transportar de un producto homogéneo desde un conjunto de orígenes (por ejemplo, fabricas) a un conjunto de destinos (por ejemplo, puntos de venta) con coste mínimo, al mismo tiempo que se satisfagan tanto unas disponibilidades o límites de las ofertas en los orígenes como unos requerimientos de demanda de los destinos.

En el problema general hay m orígenes o fábricas y n destinos o puntos de venta, cada uno de ellos representado mediante un nodo. Cada origen lleva asociado una capacidad o disponibilidad en unidades de producto, mientras que cada destino lleva asociado una demanda. Entre cada origen y destino hay más de una ruta que lleva asociada un coste por unidad transportada diferente de una ruta a otra. El modelo asume que el coste de envío a través de una determinada ruta entre un origen y un destino es directamente proporcional al número de unidades enviadas por esa ruta (Szandtner, 1978).

7.5. Programación heurística

La programación heurística es un procedimiento mediante el cual es posible obtener una solución factible muy buena, pero que no sea la óptima para el problema especificado. La programación heurística no garantiza una solución óptima pero sí una que sea casi óptima o concluir que una solución óptima no existe.

El procedimiento heurístico es un procedimiento iterativo, en el cual cada iteración implica la realización de una nueva búsqueda de una solución que puede ser mejor que la solución calculada previamente. Cuando el procedimiento finaliza después de un tiempo prudencialmente razonable, la solución dada por el modelo será la mejor de todas las obtenidas en cualquiera de las iteraciones.

Estos métodos generalmente están basados en ideas relativamente sencillas y de sentido común de cómo buscar una buena solución. Estas ideas han de ser ajustadas cuidadosamente al problema que se pretende solucionar. Por esta razón, cada procedimiento heurístico es diseñado para ajustarse a un problema en concreto más que para que pueda ser aplicado a un conjunto de problemas agrupados por algún elemento común (Hillier y Lieberman, 2010).

8. Tipos de modelos matemáticos estocásticos

En el siguiente apartado se analizan los principales modelos estocásticos que se encuentran en la literatura.

8.1. Cadenas de Markov

El proceso de Markov indica que en un sistema en el que un momento dado solo puede tomar uno, y nada más que uno, de los valores finitos o infinitos de la secuencia X_0, X_1, \dots, X_n , el estado actual X_n es suficiente para predecir el estado siguiente X_{n+1} , y este nuevo estado es independiente del resto de estados tomados por el sistema, excepto del inmediatamente anterior (Smith, 1978).

Esto es, las predicciones del futuro del proceso, una vez conocido el estado actual, no pueden mejorar con conocimiento adicional del pasado. La probabilidad de que el sistema pase del estado X_i al estado X_j es p_{ij} , y el conjunto de probabilidades de cambio de estado se puede agrupar en una matriz denominada matriz transición.

Los procesos de Markov son una herramienta eficiente, poderosa y flexible para la descripción y el análisis de las propiedades de los sistemas dinámicos, ya que permiten obtener fácilmente medidas de rendimiento y de fiabilidad de los sistemas.

Un proceso estocástico con espacio de estados discretos se denomina normalmente cadena. Por lo tanto, un proceso estocástico en tiempo discreto con espacio de estados discreto es una cadena en tiempo discreto y un proceso estocástico en tiempo continuo con espacio de estados discreto es una cadena en tiempo continuo.

8.2. Análisis de decisión

El análisis de decisión es un marco con una serie de procedimientos sistemáticos para racionalmente analizar decisiones sobre problemas en los que están involucradas salidas con incertidumbre (Keeney, 1978).

En general, la persona encargada de tomar una decisión (decisor) debe elegir una alternativa entre un conjunto de posibles opciones. En este conjunto deben encontrarse todas las alternativas viables que permitan proceder con el problema en cuestión.

La elección de la alternativa deseada se realiza en un contexto de incertidumbre, ya que la salida se verá afectada por una serie de factores aleatorios que la persona que toma la decisión no puede controlar. Estos factores determinan que situación se dará al mismo tiempo que se elige y se ejecuta la alternativa seleccionada. Cada una de estas posibles situaciones se denomina posible estado de la naturaleza.

Para cada combinación de alternativa seleccionada y estado de la naturaleza se produce un pago. El decisor conoce el pago resultante de cada una de las posibles combinaciones. El pago se define como una medida cuantitativa de la consecuencia de la combinación de una alternativa y un estado de la naturaleza. Si esta medida no se puede asegurar con certeza a pesar de conocer el estado de la naturaleza, el pago se convierte en un valor esperado, en sentido estadístico, de la consecuencia. Las medidas de todas las consecuencias se tabulan en una matriz denominada tabla de consecuencias.

Por último, puede darse el caso de que el decisor tenga cierta información acerca de la probabilidad relativa de los posibles estados de la naturaleza. Esta información puede ser traducida a una distribución de probabilidad, convirtiendo de esta manera al estado de la naturaleza en una variable aleatoria. Esta distribución de probabilidad se conoce como distribución a priori, y generalmente está basada en la experiencia de la persona y es totalmente subjetiva.

Una opción para reducir la subjetividad y mejorar las estimaciones previas sobre el valor la distribución de probabilidad a priori, es la realización de ensayos o experimentación. Esta experimentación implica una serie de costes que pueden hacer desechar esta posibilidad. Las distribuciones de probabilidad obtenidas mediante la experimentación se denominan distribuciones a posteriori, y al proceso de toma de decisión, toma de decisión con experimentación.

Una vez construida la tabla, el decisor puede emplear el criterio de valor esperado para seleccionar la mejor alternativa. El valor esperado es la suma ponderada de los pagos correspondientes a la alternativa de decisión. El factor de ponderación de cada pago es la probabilidad de ocurrencia del estado de la naturaleza asociado a ese pago.

8.3. Teoría de la utilidad

En el análisis de decisión se ha definido pago como una medida cuantitativa de la consecuencia de la combinación de una alternativa y un estado de la naturaleza. Se ha considerado pago como la única medida a la hora de asignar un determinado valor a una posible consecuencia.

Sin embargo, el término “valor” puede ser empleado en dos sentidos. El primero, explicado anteriormente, es el sentido cuantitativo ó numérico, en el cual se asigna una representación numérica a una consecuencia. El segundo sentido es el cualitativo o comparativo, el cual se puede observar en expresiones como “valoro esto más que aquello” ó “prefiero esto a aquello”. A pesar de estas diferencias los dos sentidos están relacionados e interaccionan, siendo la representación numérica de valores una expresión cuantitativa apropiada para expresar la relación cualitativa preferencial del decisor (Fishburn, 1978).

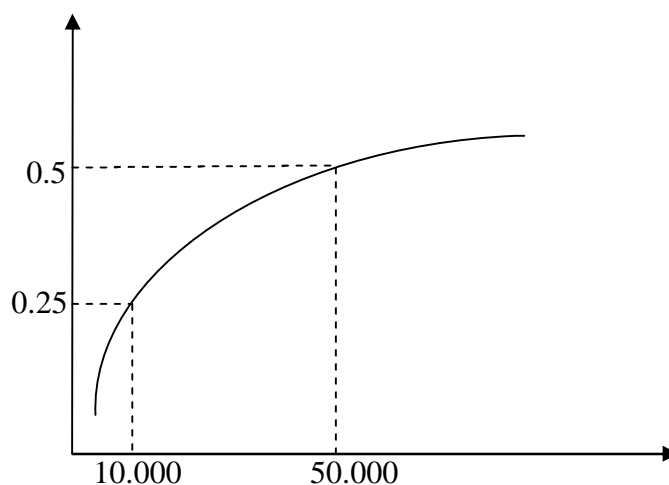


Figura 6: Gráfica de la función de utilidad

Hay ocasiones en las cuales antes las mismas posibles alternativas, diferentes individuos no seleccionan la misma. Ante la posibilidad de ganar 30 unidades monetarias con un 100% de seguridad, o la posibilidad de ganar 100 unidades o perder las 100 tras acertar el lado sobre en el que una moneda lanzada al aire caerá, es un ejemplo claro en el que diferentes individuos seleccionarían alternativas opuestas.

Para transformar el sentido cuantitativo del pago en una escala que refleje de manera adecuada las preferencias del decisor, se emplea la función de utilidad (Fig. 3).

La función de utilidad representada en la figura 3, muestra como este decisor valora el obtener 50.000 unidades monetarias el doble que obtener 10.000. La curva de la gráfica muestra un pendiente ascendente a medida que la cantidad de unidades monetarias aumenta, por lo que a este tipo de gráficas se

las denomina ley de utilidad marginal creciente. Estos individuos prefieran obtener valores monetarios altos.

Sin embargo, pueden existir individuos que tienen sus necesidades cubiertas con valores monetarios bajos, y no valoran el conseguir mayores ingresos. En estos casos la pendiente la gráfica de la función de utilidad es descendente a medida que el valor de las unidades monetarias aumenta, y la gráfica se conoce como ley de utilidad marginal decreciente.

Un caso intermedio entre los dos mencionado anteriormente, es aquel en el que el individuo valora la utilidad de forma proporcional al número de unidades monetarias.

Aunque se han mostrado tres tipos de individuos claramente diferenciados, lo más habitual es que la actitud ante el riesgo varíe en función de la circunstancias, adoptando características de los tres individuos expuestos en momentos diferentes. También la actitud de un individuo ante el riesgo es diferente cuando la decisión involucra más persona, en el caso de un consejo de administración, ó si la consecuencia afecta a bienes propios o corporativos.

8.4. Teoría de juegos

Existen una gran variedad de sistemas en los cuales se da un conflicto o competición. Ejemplos en los cuales se dan adversarios en un conflicto son campañas políticas, de marketing y publicidad, batallas militares, etc. Una característica básica que se da en muchas de las situaciones en estos sistemas es que la salida final depende de las estrategias que los adversarios adoptan. La teoría de juegos es una teoría matemática que de manera abstracta intenta conocer el proceso de toma de decisiones del adversario en situaciones que incluyen conflicto y cooperación (William, 1981).

La teoría de juegos da por sentado que un conflicto surgirá de manera natural cuando varios participantes tienen preferencias diferentes, y pone las bases de cómo estudiar cuantitativamente estos problemas. Muchos de los aspectos abstractos que se dan en estos problemas son comunes en diferentes situaciones competitivas, y mediante esta teoría, pueden ser estudiados de manera científica.

La teoría de juegos indica cómo se debería de proceder para alcanzar la mejor solución conociendo las opciones que el oponente tiene a su disposición. En la teoría de juegos se dan dos o más participantes que se ven envueltos en una serie de encuentros estratégicos. La principal característica que diferencia esta técnica del análisis de decisión, es precisamente el hecho de que haya más de un participante. En estos modelos no existirá un único participante frente elementos o situaciones aleatorias, sino que se enfrentará a otro participante con intereses contrarios a los suyos. Como en cada uno de estos

encuentros cada participante considera que es valioso para él, la teoría de juegos constantemente emplea la teoría de utilidad.

El alcance de la teoría de juegos es muy amplio y se deben hacer ciertas acotaciones para poder acometer un problema. Los diferentes factores que se pueden dar en una situación de conflicto ofrecen un abanico muy amplio de posibles juegos. En 1944 von Neumann y Morgenstern describieron tres modelos abstractos ó formas de estudiar los juegos: la forma extensiva, la forma normal y la forma de función característica. La gran mayoría de los modelos empleados actuales en la teoría de juegos están basados en estos tres tipos básicos, a los que se le ha hecho una generalización o una extensión (William, 1981).

Un juego de n personas de forma extensiva es un gráfico que no tiene bucles y con un nodo principal del cual se descuelgan el resto de nodos mediante arcos. Cada nodo representa una posición en el juego y cada arco un movimiento. Asociado a cada movimiento existe una información del pago por cada uno de los movimientos, tanto del propio jugador como de los oponentes. Cada movimiento puede ser asignado a un jugador o puede ser un movimiento aleatorio. Si el movimiento es realizado por un jugador, las opciones y los pagos están a disposición del jugador. Si el movimiento es aleatorio una distribución de probabilidad es especificada (Owen, 1978). Esta forma de juego es muy descriptiva, pero no resulta de gran utilidad debido a su complicado manejo computacional.

Un juego de n personas de forma normal está compuesto por una serie de números enteros y una función del pago F de dimensión n . El vector F describe el pago de los n jugadores en el momento que cada uno realiza su elección. Se puede incluso añadir una serie de movimientos aleatorios, determinando una distribución de probabilidad para cada uno de esos movimientos.

Un juego de n personas en la forma de función característica se determina por una función característica de valores reales v en un conjunto de 2^n de todos los subconjuntos de un conjunto finito N . El vector v asigna el número real $v(s)$ a cada subconjunto S de N . N representa el conjunto de n jugadores y $v(s)$ representa el valor que la coalición puede obtener si los jugadores colaborasen.

8.5. Proceso de Poisson

En ocasiones es necesario conocer una secuencia de puntos aleatorios en el tiempo, como la ocurrencia del i -ésimo evento en el tiempo t_i , así como, la distribución de tiempos de los diferentes eventos. A este tipo de procesos estocásticos se les denomina proceso de llegadas, e indica como los diferentes elementos llegan a un determinado servicio.

Uno de los modelos más empleados para este tipo de situaciones es el proceso de Poisson, el cual es un proceso de llegadas en el que los tiempos entre las diferentes llegadas se distribuyen de manera exponencial (Law y Kelton, 1991) y también es conocido como procesos de rara ocurrencia (Montgomery et al, 2004).

Se dice que un proceso de llegadas es un proceso de Poisson si se cumple:

- Las llegadas se producen de una en una
- El número de llegadas en un intervalo de tiempo es independiente al número de llegadas hasta ese momento.

$$\{N(t + s) - N(t)\} \text{ es independiente de } \{N(u), 0 \leq u \leq t\}$$

- La distribución de las llegadas es independiente del tiempo.

$$\{N(t + s) - N(t)\} \text{ es independiente de } t \text{ para todo } t, \text{ siendo } s \geq 0$$

Si estas tres condiciones se cumplen, las llegadas a un determinado servicio siguen una distribución de Poisson de parámetro λ , y el tiempo entre los diferentes eventos ocurridos sigue una distribución exponencial de parámetro $1/\lambda$ (Çinlar, 1975).

8.6. Teoría de colas

En numerosas ocasiones hay que esperar en una cola o línea de espera. La espera se produce porque hay más demanda de servicio que el disponible. La espera para obtener un servicio puede suponer una merma importante de la productividad de cualquier actividad comercial. Ejemplos de esta pérdida son el tiempo de espera de una máquina en ser reparada, el tiempo de espera de un vehículo en ser descargado, el tiempo de espera de los aviones en un aeropuerto para poder despegar, y muchos otros.

Sin embargo, ampliar esta capacidad de servicio no siempre es la solución adecuada. Por un lado, esta ampliación se puede acometer con una inversión económica, pero podríamos llegar a una situación en la que raramente se formasen colas y los servicios estarían ociosos gran parte del tiempo, infrautilizando los medios disponibles. Por otro lado, si casi todos los clientes deben ingresar en la cola, probablemente perderíamos algunos que irían a satisfacer su servicio a otros sistemas. Por tanto, se trata de compensar un nivel adecuado de servicio con unos gastos no excesivos.

Para llegar a una solución hay que resolver ciertas preguntas como ¿cuánto tiempo debe esperar un cliente?, ¿cuántos clientes se acumularán en la cola? o ¿cuántos clientes llegan por unidad de tiempo? A partir de las respuestas a estas preguntas, se podrá considerar varios sistemas alternativos y tratar de evaluar su funcionamiento. La teoría de colas proporciona diversos modelos que dan respuesta a cuestiones de como emplear el diseño de colas para minimizar el coste total de servicio y espera.

En un modelo de colas los clientes, que provienen de una población o fuente llegan al sistema para recibir algún tipo de servicio. El término cliente se usa en sentido genérico, pudiendo ser tanto una persona que espera en la cola del cine, como un avión esperando para despegar o un programa esperando para ser ejecutado. El dispositivo de servicio del sistema ofrece un conjunto (limitado) de servidores o recursos, a veces llamados canales, para satisfacer las peticiones de los clientes. Si cuando el cliente llega al sistema, todos los servidores están ocupados, deberá esperar en la cola antes de empezar a recibir servicio (Winston, 1994).

8.7. Análisis de Regresión

El análisis de regresión es una metodología estadística que emplea una relación entre dos o más variables cuantitativas, de tal manera que una variable respuesta o salida esté relacionada con una o varias de estas variables cuantitativas (Kutner et al, 2004). En el análisis de regresión la relación entre distintas variables se puede investigar y modelar mediante una relación algebraica, y se dice que una variable depende de otras.

Cuando la variable endógena o dependiente está relacionada con una única variable exógena o independiente, el análisis de regresión se denomina regresión simple. Sin embargo, cuando son varias las variables que muestran el comportamiento de la variable dependiente se dice que el análisis de regresión es múltiple (Webster, 2000). La expresión algebraica mediante la que se puede estimar el comportamiento de la variable dependiente se conoce como ecuación de regresión, y es

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad \text{Ecuación (1)}$$

donde y es la variable endógena, x_i son las variables exógenas, β_i son los parámetros de regresión y ε el error. Mediante esta ecuación, se puede determinar la variación de la variable endógena cada vez que se varía en una unidad una de las variables exógenas, manteniendo el resto de ellas constantes.

Esta expresión algebraica representa la recta que mejor indica el valor de la variable endógena para cualquier combinación de los valores de las variables exógenas. Desafortunadamente, no lo hace de manera exacta, y de ahí que haya que incorporar en la expresión el término ε , el cual describe la perturbación aleatoria o error del modelo (Myers, 1990). Por lo tanto, se puede describir el error como la variable aleatoria que cuantifica la diferencia entre la ecuación de regresión y el valor de los datos.

Conocido esto, la ecuación de la variable endógena se puede dividir en dos bloques. El primero, $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k$, el cual es constante y se considera el valor medio de la respuesta para una combinación de valores fijos de las variables exógenas. Y el segundo, el término ε , el cual es una variable aleatoria. Esto implica que hay una distribución de los valores y_i para cada

combinación de valores x_i , y las propiedades estadísticas de estas distribuciones serán las mismas que las del error.

Dos aspectos importantes en el análisis de regresión son el determinar que variables exógenas son las que realmente determinan el comportamiento de la variable endógena de manera sustancial, y estimar los valores desconocidos de los parámetros de regresión. Estas dos acciones se conocen como ajuste del modelo a los datos, obteniendo de esta manera el modelo de regresión lineal. El análisis de regresión es un proceso iterativo, en el cual los datos conducen al modelo a través del ajuste de los datos. Una vez el ajuste se ha realizado hay que comprobar la calidad del mismo, y si este análisis no es satisfactorio habrá que modificar el modelo, ajustándolo de nuevo.

El modelo de regresión, al igual que el resto de modelos, no es más que una aproximación del sistema, por lo tanto, cada vez que pretendemos estimar el valor de la variable endógena se comete un error. Esto es debido a que los parámetros de regresión calculados al ajustar el modelo a los datos no son más que estimaciones de los mismos, ya que los verdaderos valores de β_i son desconocidos. Y el modelo de regresión queda como sigue:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_k x_k + e_i \quad \text{Ecuación (2)}$$

Donde \hat{y}_i es el valor estimado que la variable endógena toma para una combinación de las variables exógenas. Es importante diferenciar entre los valores verdaderos β_i y los valores estimados $\hat{\beta}_i$. Los valores verdaderos β_i son constantes que no se conocen y los valores estimados $\hat{\beta}_i$ son cantidades que se estiman a partir de los datos (Navidi, 2006). Como provienen de los datos, cada vez que seleccionemos una muestra diferente, las cantidades calculadas serán diferentes, y por tanto los $\hat{\beta}_i$ serán variables aleatorias.

También es importante diferenciar el error de la ecuación de regresión del residual del modelo de regresión e_i . El residual es la diferencia entre el valor i que la variable dependiente toma en la muestra, y el valor i ajustado en el modelo de regresión. Mientras que el error es un valor desconocido de la falta de exactitud de la ecuación de regresión (Navidi, 2006). Si no existiese error, los valores verdaderos β_i y los valores estimados $\hat{\beta}_i$ serían idénticos.

Otra diferenciación que hay que realizar en los modelos de regresión es la de modelos de regresión lineales y los modelos de regresión no lineales. El modelo de regresión lineal es aquel que es lineal respecto a los parámetros, no respecto a las variables exógenas (Seber y Lee, 2003). Aunque esta definición de linealidad parezca muy restrictiva, no lo es tanto, ya que las variables endógenas pueden tomar cualquier tipo de función, bien sea sinusoidal, exponencial, polinomial o de cualquier otro tipo.

Los usos para los cuales se puede emplear un análisis de regresión son cuatro: predicción, identificación de variables, especificación del modelo y estimación de los parámetros (Myers, 1990).

- En la predicción, el análisis de regresión busca conocer la respuesta de la variable endógena ante ciertos valores de las variables exógenas con un nivel de aceptación adecuado.
- En la identificación de variables, la formulación del modelo se emplea como un instrumento para detectar el grado de importancia que cada una de las variables exógenas tiene en la explicación de la variación de la variable endógena. De esta manera, las variables que realmente influyen en la respuesta pueden ser identificadas para futuros análisis, mientras que las no influyentes no son consideradas en futuros estudios.
- Mediante la especificación del modelo el sistema en estudio puede ser descrito, y por tanto el conocimiento sobre él aumenta.
- Por último, en ciertas ocasiones mediante el análisis de regresión se puede resolver problemas de estimación de parámetros.

Debido a todos estos posibles usos, las aplicaciones de los análisis de regresión son numerosas y las hay en casi cualquier campo. De hecho, puede ser que el análisis de regresión sea la técnica estadística más empleada (Montgomery et al, 2004).

A pesar de todas estas ventajas, hay que tener en cuenta que hay dos aspectos a considerar. El primero es que exista correlación no implica causalidad. Es decir, que se puede determinar unos parámetros de regresión entre una variable endógena y otras exógenas no implica que realmente exista causalidad entre ellas. Es por ello que es muy importante el analizar detenidamente el sistema antes de realizar el análisis de regresión, para que las variables involucradas en el análisis realmente sean las que determinan la variable endógena.

El segundo es que no se deben extrapolar datos fuera del rango de valores de las variables exógenas. Esto es así, ya que no conocemos cómo evoluciona el sistema fuera de este rango, pudiendo llegar a ser complementa diferente, y por tanto, las inferencias realizadas completamente erróneas.

9. Técnicas de regresión

Como uno de los principales objetivos de este trabajo de investigación es demostrar que mediante el análisis de regresión se puede desarrollar un modelo de una red de distribución capilar, a continuación se va a describir las principales técnicas con las que es posible ajustar el modelo a los datos.

9.1. Método de mínimos cuadrados

Cuando se ajusta un modelo a una serie de datos, una de las acciones a realizar es determinar el valor de los diferentes parámetros de regresión. Existen diferentes técnicas mediante las cuales se pueden calcular estos parámetros, cada una con sus características particulares.

La primera que se va a analizar es la estimación de los parámetros mediante mínimos cuadrados. Este procedimiento matemático estima el valor de los parámetros de regresión de tal manera que la recta de regresión se extiende por el centroide de los puntos de datos, aproximándose a todos ellos mejor que ninguna otra recta. Para garantizar esta aseveración, la recta obtenida mediante mínimos cuadrados, también conocida como recta de mínimos cuadrados, es tal que la suma de todos los residuales es cero, y minimiza el sumatorio del cuadrado de todos los residuales (Webster, 2000), es decir, $\sum e_i = 0$ y $\min \sum e_i^2$.

La calidad de los valores estimados mediante mínimos cuadrados de los parámetros de regresión $\hat{\beta}_i$ se puede contrastar mediante el teorema de Gauss-Markov:

“Los valores estimados de $\hat{\beta}_i$ mediante mínimos cuadrados son insesgados y tienen varianza mínima en comparación con todos los demás estimadores insesgados que sean combinaciones lineales de los valores y_i , si el valor medio de los residuales es cero, la varianza es constante y los residuales no están correlacionados” (Shao, 2003). Es decir, si se cumplen las premisas, los valores estimados de $\hat{\beta}_i$ se consideran los que mejor ajustan ya que su varianza es la mínima.

El ajuste de mínimos aporta una serie de propiedades útiles:

- La suma de los residuales de un modelo con intercepto, modelo con una ordenada en el origen, siempre es igual a cero.
- La suma de los valores observados es igual a la suma de valores ajustados.
- La suma de los residuales, ponderados por el valor correspondiente de la variable exógena, siempre es igual a cero.
- La suma de los residuales, ponderados por el valor ajustado correspondiente, siempre es igual a cero.

Para la obtención de los valores estimados de parámetros de regresión vamos a suponer que se disponen $n > p$ observaciones, siendo p el número de variables exógenas, y todas las observaciones se van agrupar en forma matricial en lugar de la forma de la ecuación (2).

$$Y = X\beta + \varepsilon \quad \text{Ecuación (3)}$$

Donde Y es un vector de $n \times 1$ de las observaciones, X es una matriz de $n \times p$ de los niveles de las variables endógenas, β es un vector de $p \times 1$ de los parámetros de regresión y ε es un vector de $n \times 1$ de errores aleatorios.

Como una de las premisas del método de mínimos cuadrados es minimizar la suma de los errores al cuadrado, según Graybill los parámetros de regresión se pueden obtener de la siguiente manera:

$$\sum_{i=1}^n \varepsilon_i^2 = \varepsilon' \varepsilon = (Y - X\beta)'(Y - X\beta) = S(\beta) \quad \text{Ecuación (4)}$$

Donde $S(\beta)$ se puede expresar como sigue:

$$S(\beta) = Y'Y - \beta'X'Y - Y'X\beta + \beta'X'X\beta \quad \text{Ecuación (5)}$$

Como $\beta'X'Y$ y $Y'X\beta$ son matrices 1×1 , es decir, son escalares, y la traspuesta de una es idéntica a la otra matriz, la expresión queda:

$$S(\beta) = Y'Y - 2\beta'X'Y + \beta'X'X\beta \quad \text{Ecuación (6)}$$

Para que esta expresión sea mínima, y de esta manera se cumpla la premisa expuesta de minimizar la suma de los errores al cuadrado, se ha de derivar $S(\beta)$ tomando como variable los parámetros de regresión β e igualar dicha derivada a cero:

$$\frac{\partial S}{\partial \beta} = -2X'Y + 2X'X\beta = 0 \quad \text{Ecuación (7)}$$

Si sustituimos β por los valores estimados de los parámetros de regresión, y resolvemos la ecuación tenemos:

$$X'X\hat{\beta} = X'Y \quad \text{Ecuación (8)}$$

Para resolver las ecuaciones se multiplican ambos lados de la ecuación anterior por la inversa de $X'X$, y el estimador de β por mínimos cuadrados es:

$$\hat{\beta} = (X'X)^{-1}X'Y \quad \text{Ecuación (9)}$$

Para que este sistema de ecuaciones tenga solución debe de existir la matriz inversa $(X'X)^{-1}$, y esta matriz siempre existe si los valores que toman las variables exógenas en cada observación son linealmente independientes del resto de observaciones, es decir, si ninguna fila de la matriz X es una combinación lineal del resto de filas (Montgomery et al, 2004).

Una vez que se conocen los valores de los estimados de los parámetros de regresión, los valores estimados de la variable endógena se pueden calcular mediante la ecuación (10):

$$\hat{Y} = X\hat{\beta} = X(X'X)^{-1}X'Y = HY \quad \text{Ecuación (10)}$$

Donde H se conoce como la matriz sombrero. También es sencillo a partir de esta expresión, el calcular los residuales que los valores estimados de la variable endógena originan respecto a los datos originales.

$$e = Y - \hat{Y} = Y - HY = (I - H)Y \quad \text{Ecuación (11)}$$

El determinar el valor de los residuales es importante ya que el estimar los parámetros de regresión mediante el método de los mínimos cuadrados se basa en la asunción de ciertos supuestos, como ya se indicó al expresar el teorema de Gauss-Markov. Estos supuestos son las siguientes (Kleinbaum et al, 1988):

- Todos los errores e_i tienen media cero.
- Todos los errores e_i están distribuidos normalmente.
- Todos los errores e_i son aleatorios e independientes. La magnitud de cualquier error e_i no influye en el valor del siguiente error e_{i+1} .
- Todos los errores e_i cumplen el principio de homoscedasticidad. Es decir, la varianza de todos los errores es constante para todas las observaciones.

Si los errores no tienen una varianza constante la utilidad del método de mínimos cuadrados para estimar la varianza se reduce hasta el punto de que el estimado de la varianza del obtenido a partir de los residuales no se considera como un estimador insesgado de la misma.

La varianza del error, y por tanto de la variable endógena, no se puede conocer porque el error es desconocido. Sin embargo, si se cumplen los supuestos anteriormente mencionados, se puede estimar esta varianza a partir de los residuales obtenidos mediante el método de mínimos cuadrados (Graybill, 1961). El estimador de la varianza se puede calcular de la siguiente forma:

$$\hat{\sigma}^2 = \frac{e'e}{n-p} = \frac{Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta}}{n-p} \quad \text{Ecuación (12)}$$

Como $X'X\hat{\beta} = X'Y$, la ecuación queda

$$\hat{\sigma}^2 = \frac{e'e}{n-p} = \frac{Y'Y - \hat{\beta}'X'Y}{n-p} \quad \text{Ecuación (13)}$$

Siendo $\hat{\sigma}^2$ un estimador insesgado de la varianza del error. Este estimado de la varianza mide el grado de dispersión de los valores estimados de la variable endógena \hat{Y} respecto a los valores Y . Esta estimación también se empleará a la hora de calcular intervalos de confianza de los parámetros de regresión.

9.2. Regresión robusta

Como se ha analizado previamente, los valores estimados de los parámetros de regresión mediante mínimos cuadrados son los estimadores insesgados más eficientes cuando los errores se distribuyen normalmente. Sin embargo, se pueden dar casos en los que los errores sean de tal magnitud, que afecten considerablemente al modelo de regresión. Estos errores con valores nominales grandes se denominan valores atípicos.

El método de mínimos cuadrados hace que sea mínima la suma de los cuadros de los residuales. Por tanto, si existe un solo residual de valor nominal grande, hará que su cuadrado sea grande, y la curva del modelo de mínimos cuadrados tienda a acercarse al valor atípico para reducir este residual y minimizar así la suma total de residuales, desviando significativamente la curva de mínimos cuadrados debido a un solo residual.

El procedimiento de regresión robusta, sin embargo, es aquel que amortigua el efecto de las observaciones que serían muy influyentes si se usaran mínimos cuadrados. Mediante este procedimiento el análisis de regresión creado no dependerá de un subconjunto pequeño de datos (Peña, 2002).

Como se mostró anteriormente, mediante el método de mínimos cuadrados la suma de los residuales, ponderados por el valor correspondiente de la variable exógena, siempre es igual a cero. Es decir,

$$\sum_{i=1}^n e_i x_i = 0 \quad \text{Ecuación (14)}$$

Esta función se puede generalizar de la siguiente manera

$$\sum_{i=1}^n \varphi \left[\frac{e_i}{\sigma} \right] x_i = 0 \quad \text{Ecuación (15)}$$

Donde φ se denomina función de influencia. En el procedimiento de regresión robusta se elige la función de influencia φ de tal manera que los puntos con residuales grandes no ejerzan una influencia indebida. La solución de los coeficientes de regresión en la función de influencia se conoce como estimadores M.

Huber en 1973 desarrollo una función la cual limitaba el efecto de los residuales de la siguiente manera:

$$\begin{aligned} \varphi \left[\frac{e_i}{\sigma} \right] &= \left[\frac{e_i}{\sigma} \right] & \text{si} & & -r \leq \left[\frac{e_i}{\sigma} \right] \leq r \\ \varphi \left[\frac{e_i}{\sigma} \right] &= r & \text{si} & & r < \left[\frac{e_i}{\sigma} \right] \\ \varphi \left[\frac{e_i}{\sigma} \right] &= -r & \text{si} & & \left[\frac{e_i}{\sigma} \right] < -r \end{aligned}$$

Siendo la gráfica la Figura 7:

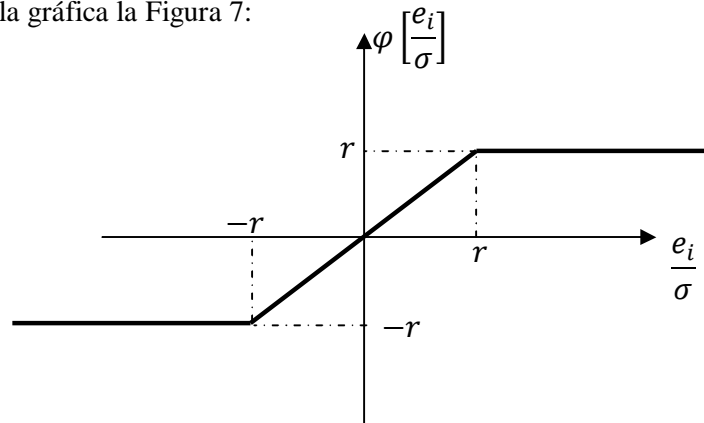


Figura 7: Gráfica influencia con efecto de los residuales limitado

Para obtener los estimadores M que cumplan con estas características, las ecuaciones a resolver son no lineales y deben emplearse procedimientos iterativos para su solución.

Por otro lado, el valor de σ debe estimarse. Existen más de una posibilidad para estimar este valor pero sea cual sea la elección, ha de ser una estimación robusta de la escala. Myers indica que una elección razonable puede ser

$$\hat{\sigma} = 1.5 \text{ mediana } |e_i| \quad \text{Ecuación (16)}$$

9.3. Regresión ortogonal

La técnica de regresión ortogonal, o regresión múltiple por etapas, es un procedimiento de análisis de regresión para calcular los parámetros de regresión cuando las variables exógenas están correlacionadas. Si las variables exógenas son dependientes, al calcular los valores de los parámetros de regresión mediante el método de mínimos cuadrados, la pendiente en la regresión entre la variable endógena y una variable exógena x_i , mide el efecto debido a x_i y una parte de los efectos debidos a x_j (Peña, 2002).

El procedimiento para realizar la regresión por etapas con variables exógenas dependientes, suponiendo que tienen media cero, es el siguiente:

- Realizar el análisis de regresión con la primera variable

$$\hat{Y} = \hat{b}X_1 + e_{y,1} \quad \text{Ecuación (17)}$$

Donde \hat{b} representa la suma de los efectos debidos a la variable x_1 más los efectos debidos al resto de variables.

- Eliminar de la siguiente variable la parte común con la primera variable, mediante un análisis de regresión simple.

$$X_2 = \hat{c}X_1 + e_{2,1} \quad \text{Ecuación (18)}$$

- Relacionar $e_{y,1}$, parte de y no explicada por x_1 , con $e_{2,1}$, parte de x_2 independiente de x_1 , mediante regresión simple.

$$e_{y,1} = \hat{d}e_{2,1} + e_{y,12} \quad \text{Ecuación (19)}$$

- Por último, sustituimos en la primera ecuación (17), las ecuaciones (18) y (19), y tras operar se obtiene la ecuación de regresión múltiple para dos variables (Ecuación (20)).

$$\hat{Y} = \hat{b}X_1 + \hat{d}(X_2 - \hat{c}X_1) + e_{y,12} \quad \text{Ecuación (20)}$$

Para el caso de tener más de dos variables exógenas, el proceso es análogo. Para demostrar que la ecuación (20) es válida como para considerarla la ecuación de regresión múltiple, el residuo $e_{y,12}$ debe ser ortogonal a ambas variables exógenas.

Si multiplicamos la ecuación (19) por el vector X'_1 en ambos lados de la igualdad, se tiene:

$$X'_1 e_{y,1} = \hat{d}X'_1 e_{2,1} + X'_1 e_{y,12} \quad \text{Ecuación (21)}$$

Debido a la propiedad enunciada anteriormente, por la cual la suma de los residuales, ponderados por el valor correspondiente de la variable exógena, siempre es igual a cero, los términos de la expresión $X'_1 e_{y,1}$ y $\hat{d}X'_1 e_{2,1}$ son cero. Y por tanto, para que el término $X'_1 e_{y,12}$ sea igual a cero, implica que X'_1 es ortogonal a $e_{y,12}$.

Por otro lado, se ha cumplir que

$$e'_{2,1} e_{y,12} = 0 \quad \text{Ecuación (22)}$$

Para ello, se puede sustituir $e'_{2,1}$ por la ecuación (19)

$$(X_2 - \hat{c}X_1)' e_{y,12} = 0 \quad \text{Ecuación (23)}$$

Como el término $X'_1 e_{y,12}$ es igual a cero, el término $X'_2 e_{y,12}$ también será cero, y por tanto, $e_{y,12}$ es ortogonal respecto a X_2 .

9.4. Modelos lineales generalizados

Los modelos lineales generalizados (GLM de sus siglas en inglés Generalized Linear Models) proporcionan una aproximación sistemática de la especificación de modelos de regresión que pueden alojar relaciones no lineales en sus respuestas (McCullagh y Nelder, 1983).

El GLM es un procedimiento que unifica los análisis de regresión lineal y no lineal. Se realizan una serie de transformaciones en las variables exógenas, y tras esta transformación se estiman los parámetros de regresión como si de un modelo lineal se tratase. Además, permite incorporar distribuciones de respuesta no normales. Con que la distribución de probabilidad de la variable endógena sea miembro de la familia exponencial, siendo la distribución normal una de ellas, es suficiente para poder aplicar el GLM. Las distribuciones que son miembros de la familia exponencial tienen la forma (Myers, 1990):

$$f(y_i, \theta_i, \varphi) = \exp \left\{ \frac{[y_i \theta_i - b(\theta_i)]}{a(\varphi)} + h(y_i, \varphi) \right\} \quad \text{Ecuación (24)}$$

Donde φ es un parámetro de escala, y θ_i se denomina parámetro natural.

Los modelos lineales generalizados están compuestos de tres elementos:

- Una distribución de probabilidad para la variable endógena, que dependa de la media.

$$y_i = \mu_i + \varepsilon \quad \text{Ecuación (25)}$$

- Un predictor lineal en las variables exógenas, que corresponde a la forma

$$\eta_i = g[E(y_i)] = g(\mu_i) = X'_i \beta \quad \text{Ecuación (26)}$$

donde η_i se denomina cadena canónica. La expresión anterior es equivalente a

$$E(y_i) = g^{-1}(\eta_i) = g^{-1}(X'_i \beta) \quad \text{Ecuación (27)}$$

- Una función cadena la cual relaciona la media de la respuesta con el predictor lineal. Es decir, es la transformación que se realiza a las variables exógenas para convertirla en μ_i . Esta función cadena ha de ser continua y monótona, y por tanto, diferenciable.

Las funciones cadenas más comunes se ilustran en la tabla 2:

Tabla 2: funciones cadena más comunes

Cadena canónica		
Cadena identidad	$\eta_i = \mu_i$	$\mu_i = \eta_i$
Cadena logística	$\eta_i = \ln \mu_i$	$\mu_i = \exp \eta_i$
Cadena logarítmica	$\eta_i = \ln \frac{\mu_i}{(1 - \mu_i)}$	$\mu_i = \frac{e^{\eta_i}}{(1 + e^{\eta_i})}$
Cadena recíproca	$\eta_i = \frac{1}{\mu_i}$	$\mu_i = \frac{1}{\eta_i}$
Cadena log-log	$\eta_i = \ln\{\ln[1 - E(y_i)]\}$	$\mu_i = \exp \eta_i$
Cadena raíz cuadrada	$\eta_i = \sqrt{\mu_i}$	$\mu_i = \eta_i^2$

La base teórica para estimar los parámetros de regresión es el método de máxima verosimilitud, aunque también hay casos en los que los parámetros se estiman mediante mínimos cuadrados ponderados. Y como en cualquier otro procedimiento de ajuste de modelo, se puede hacer un análisis de residuales.

Una característica importante de los modelos lineales generalizados es que la varianza de la variable endógena no es constante, por lo que para probar hipótesis y establecer intervalos de confianza para los parámetros del análisis de regresión será necesario emplear procedimientos alternativos a los de mínimos cuadrados, como por ejemplo la inferencia de Wald (Gonzales, 2001).

9.5. Método de máxima verosimilitud

El método de máxima verosimilitud es un procedimiento iterativo, en el cual se disponen los valores muestrales y_i de una variable aleatoria Y , cuya probabilidad de distribución es conocida excepto por los valores de parámetro θ . A partir de los datos observados y_i de la muestra, se pueden realizar afirmaciones o aseveraciones acerca de la distribución de probabilidad Y (Hocking, 2003). Para una muestra de observaciones independientes la ecuación mediante la cual se define máxima verosimilitud es:

$$L(\theta) = f(y; \theta) = \prod_{j=1}^n f(y_j; \theta) \quad \text{Ecuación (28)}$$

Esta expresión se conoce como función de máxima verosimilitud. La asunción de que f es conocida excepto por la incertidumbre sobre θ , reduce el problema a realizar aseveraciones sobre qué valores de θ son plausibles, dado que los valores y_i son dados (Davison, 2003). La función de máxima verosimilitud cuantifica la posibilidad de que θ genere los valores muestrales observados.

Como las observaciones son independientes, la probabilidad total de todas las observaciones se puede obtener como la multiplicación de la probabilidad individual de cada observación. Lógicamente, los valores de L serán mayores para valores de θ próximos a los que generaron los datos. Por lo tanto, se busca conocer que valores de θ hacen que el valor de L sea mayor. El valor de θ para el cual L es máximo

se conoce como estimado de máxima verosimilitud, MLE (Maximum likelihood Estimator) (Seber y Lee, 2003).

En el análisis de regresión, hasta el momento se ha considerado que los errores se distribuyen de manera normal. Sin embargo, puede darse el caso en el que los errores (y por tanto, la variable endógena, como ya se ha explicado anteriormente) no sigan esta distribución. Mediante el método de máxima verosimilitud se podrá determinar los valores de los parámetros de regresión cuando los errores siguen cualquier otra función de distribución de probabilidad diferente a la normal. En estos casos la función de máxima verosimilitud será la siguiente:

$$L(f(e; \hat{\beta}, \hat{\sigma}^2)) \quad \text{Ecuación (29)}$$

Pudiendo probar con diferentes funciones de distribución de probabilidad f para ver cuál es la que ofrece un mayor valor de L . Si calculamos los estimados de parámetros de regresión con una función f de distribución de probabilidad normal, los resultados obtenidos serán exactamente los mismos que mediante el método de mínimos cuadrados (Myers, 1990).

Para calcular los estimados de máxima verosimilitud $\hat{\beta}$ y $\hat{\sigma}^2$, es necesario derivar la función de máxima verosimilitud y ver donde esta primera derivada se anula. Sin embargo, por definición, la función de máxima verosimilitud es un producto de probabilidades, lo cual puede ser bastante difícil de derivar. Matemáticamente puede ser más interesante transformar esta función en una escala logarítmica. Como la función logaritmo es una función monótona, es decir, creciente, el máximo del logaritmo de máxima verosimilitud será equivalente al máximo de verosimilitud. Una vez calculado los puntos en los que la primera derivada se anula, será necesario realizar la segunda derivada de la función de verosimilitud, y comprobar que realmente se trata de un máximo y no de un mínimo en los puntos estimados de máxima verosimilitud.

El método de máxima verosimilitud da lugar a estimadores consistentes, eficientes y suficientes. Por otro lado, los estimadores obtenidos mediante este método presentan la propiedad invariante, es decir, que no cambian al aplicarles un conjunto de transformaciones, aspecto que generalmente no ocurre en el método de mínimos cuadrados (Graybill, 1961).

9.5.1. Test del cociente de verosimilitud

Otra característica importante del método de máxima verosimilitud es que nos permite la comparación de modelos. Mediante este test se comparan los valores de máxima verosimilitud para dos análisis de regresión. Cada uno de estos análisis no incluirá el mismo número de variables exógenas. Por lo tanto, este test nos permitirá determinar si la inclusión o no de ciertas variables aumenta la exactitud del análisis de regresión a la hora de ajustar el modelo.

El test del cociente de verosimilitud permite cuantificar la falta de ajuste del modelo cuando son eliminadas ciertas variables del modelo. Es decir, muestra si el modelo es una representación válida de los datos cuando se compara con la representación de otro modelo completo, el cual incluye todas las variables exógenas (Myers, 1990). O visto de otra manera, el empleo de este test permite comprobar si la adición de variables al modelo proporciona un mejor ajuste a los datos.

El estadístico empleado para este test es:

$$\lambda(\beta) = -2 \ln \left(\frac{L(\text{modelo reducido})}{L(\text{modelo completo})} \right) \quad \text{Ecuación (30)}$$

O lo que es lo mismo

$$\lambda(\beta) = -2 \ln(L(\text{modelo reducido}) - L(\text{modelo completo})) \quad \text{Ecuación (31)}$$

Según el teorema de Wilks, tanto el numerador como el denominador siguen asintóticamente una distribución X^2 con $n-p-1$ grados de libertad, siendo p el número variables de los que consta el modelo (Halls, 1990). Por lo tanto, el término $\lambda(\beta)$, denominado desviación, seguirá una distribución de probabilidad X^2 con grados de libertad igual a la diferencia de variables entre el modelo completo y el modelo reducido.

10. Aspectos de la simulación

10.1. Monte Carlo

Bajo el nombre de Método Monte Carlo o Simulación Monte Carlo se agrupan una serie de procedimientos que analizan distribuciones de variables aleatorias usando simulación de números aleatorios (Ríos et al, 2004). Generalmente en estadística los modelos aleatorios se usan para simular fenómenos que poseen algún componente aleatorio. Pero en el método Monte Carlo, por otro lado, el objeto de la investigación es el objeto en sí mismo, un suceso aleatorio o pseudo-aleatorio se usa para estudiar el modelo. Por lo tanto, bajo el nombre de método de Monte Carlo se agrupan aquellas simulaciones que emplean el uso de números aleatorios.

Según Kleijnen las principales áreas de aplicación del método Monte Carlo son las siguientes:

La primera área de aplicación es la solución de problemas determinísticos. Un problema determinístico se puede solucionar mediante Monte Carlo si el problema tiene la misma expresión formal que un proceso estocástico. En estos casos un parámetro determinista del problema se expresa como una distribución aleatoria y se simula dicha distribución. Halton hizo en 1971 una colección de 251 de

ejemplos en los que mediante Monte Carlo se encontró la solución a diferentes problemas matemáticos y físicos.

La segunda área es la formada por la toma de muestras del modelo. Cuando la variable endógena es una variable aleatoria, que a su vez es función de otras variables exógenas aleatorias, se puede conocer su valor estimando a partir los valores de las variables exógenas. Esta prueba se puede repetir tantas veces como sea necesario para poder estimar los valores de la distribución de probabilidad de la variable endógena. Mediante esta técnica se han comprobado por ejemplo la robustez de algunos estadísticos ó el rendimiento de varias técnicas de regresión lineal en modelos econométricos (Kleijnen, 1974).

La tercera área de aplicación es la simulación. A diferencia de Hammersley y Handscomb, los cuales solo consideran Monte Carlo una técnica de resolución matemática Kleijnen considera el método Monte Carlo una técnica válida con la que realizar experimentos con un modelo para poder conocer más acerca del sistema.

Para poder aplicar una simulación mediante el método de Monte Carlo, el primer paso es conocer las funciones de distribución de probabilidad acumuladas las variables aleatorias, y los parámetros de las mismas.

Una vez conocidas estas distribuciones de probabilidad acumuladas, se generan una serie de números aleatorios cuyos valores se encuentren entre 0 y 1. Con estos números aleatorios hay que determinar el valor que la variable aleatoria toma para ese número aleatorio mediante el método de la transformada inversa, el cual obtiene este valor a través de la distribución de probabilidad acumulada. Se realizarán tantas iteraciones de esta operación como muestras necesitemos del modelo. Esta técnica se denomina método de las repeticiones (Guasch et al, 2003).

10.1.1. Números aleatorios

Como el método de Monte Carlo se basa en el empleo de números aleatorios, en este apartado se analizan en más detalle. En la actualidad existen multitud de técnicas para obtención de números aleatorios, algunos basados en fenómenos aleatorios, y otros en procedimientos deterministas recurrentes. Los números aleatorios generados por estos procedimientos deterministas no son estrictamente aleatorios, ya que son procedimientos recursivos, y son conocidos como números pseudo-aleatorios, aunque se les denomina igualmente números aleatorios. Sin embargo, estos números pseudo-aleatorios se pueden emplear como números aleatorios si cumplen las siguientes características (Rubinstein, 1981):

- Uniformidad de la distribución: Los números aleatorios generados deberán distribuirse de forma uniforme a lo largo de todo el rango de valores posibles que pueda tomar la serie, de forma que no aparezcan tramos no recubiertos ni agrupaciones no deseadas.
- Independencia estadística: No se puede establecer ninguna relación relevante entre una serie de números con los restantes números de la serie.
- Reproductividad: Los números aleatorios han de ser estables y no degenerar a lo largo de su utilización en la simulación.

La técnica de obtención de números aleatorios más conocida actualmente son los generadores congruenciales, los cuales pueden ser lineales o multiplicativos.

Los generadores congruenciales lineales, introducidos originalmente por Lehmer, son una secuencia de números enteros X_i definida por la congruencia lineal

$$X_{i+1} = (aX_i + c) \pmod{m} \quad \text{Ecuación (32)}$$

donde m se denomina modulo, a es el factor multiplicativo, c es el factor aditivo y X_0 es la semilla o valor inicial, y son todos enteros no negativos. Además, deben satisfacer las siguientes condiciones $0 < m$, $a < m$, $c < m$ y $X_0 < m$. De esta forma se cumple que X_i va estar entre 0 y $m-1$. El término \pmod{m} es igual a mk_i , donde k_i es

$$k_i = \frac{(aX_i + c)}{m} \quad \text{Ecuación (33)}$$

Y los números aleatorios en el intervalo $[0,1]$ se obtienen

$$U_i = \frac{X_i}{m} \quad \text{Ecuación (34)}$$

Si en un determinado momento se repite un número, entonces se repite toda la serie. A esta longitud de serie se la denomina ciclo del generador. Este ciclo está directamente relacionado con el valor del modulo.

Los generadores congruenciales multiplicativos son aquellos en los que el parámetro aditivo c es igual a cero, y el factor multiplicador mayor que uno. Su expresión queda

$$X_{i+1} = (aX_i) \pmod{m} \quad \text{Ecuación (35)}$$

A la hora obtener estos números aleatorios hay que tener en cuenta ciertas consideraciones:

- Elegir el mayor número posible primo de m
- Elegir convenientemente a para obtener un ciclo $m-1$, ya que todos los valores de a no son convenientes

11. Aplicaciones de modelos de simulación matemáticos en el sector del transporte

A la hora de modelar redes de distribución capilares en el sector del transporte terrestre por carretera de mercancías, los modelos de simulación matemáticos más empleados son los modelos deterministas.

Baldarcci et al, emplean un algoritmo exacto en programación entera para optimizar las rutas de servicio de un grupo de vehículos localizados en un depósito central (Baldarcci et al, 2011). Al igual que Dong et al, en su trabajo de diseño de una red de recogida y traslado de pasajeros a un aeropuerto (Dong et al, 2009).

Subramanian et al, en su trabajo de modelado de una red de distribución de logística inversa, desarrollan un modelo mediante programación heurística (Subramanian et al, 2010). También Berbeglia, et al, desarrollan un algoritmo heurístico para optimizar una red de distribución capilar one-to-one, en las que se puede incluir las redes de distribución capilar de una empresa de paquetería industrial (Berbeglia, et al, 2010). Al igual que Batarra et al, para modelar los costes de su red one-to-one (Batarra et al, 2010).

Del mismo modo, Hoff et al, utilizan también un algoritmo heurístico para modelar su red de distribución capilar de una empresa de distribución de cervezas (Hoff et al, 2009). Karacik en su trabajo de investigación, empleó métodos heurísticos para determinar el número óptimo de camiones y conductores de camión que debería tener cada terminal de transporte para reducir al máximo los costes y optimizar así el rendimiento de dichas terminales (Karacik, 2007).

Tang y Dieguez, en su trabajo, emplean un algoritmo heurístico para solucionar el problema de reparto y recogida simultaneo en una red capilar (Tang y Dieguez, 2006).

Leturia Infante, por su parte, analizó en su trabajo de investigación, analizó como optimizar una red de distribución capilar de una empresa de distribución minorista y la modeló mediante programación lineal mixta (Leturia, 2007)). Bowers, en su tesis doctoral, minimizó el número de vehículos necesarios por una empresa de transporte empleando también los algoritmos de la programación lineal (Bowers, 2005). Palmgren, empleó el principio de generación de columnas de programación lineal para optimizar una red de distribución, en este caso, de una industria maderera en su trabajo (Palmgren, 2005). Más ejemplos en los que se ha empleado programación lineal para modelar redes de distribución por carretera en trabajos de investigación son Gnanasekaran en 2004; Kuyzu en 2007; Ribiero en 2004 y Ashish en 2004, entre otros.

Todas las redes de distribución modeladas en los ejemplos anteriores pueden considerarse muy similares a las redes de distribución capilar que una empresa de paquetería industrial posee, y que por tanto, se desean modelar en este trabajo, aunque ninguna es estrictamente una red de distribución capilar de una empresa de paquetería industrial. Pero, tal y como se apuntó anteriormente, la metodología que se desea emplear para este trabajo estará basada en técnicas estocásticas.

A pesar de que la técnica más empleada en el modelado de redes de transporte es el modelo matemático determinista, y de que se ha empleado esta técnica en casi todos los aspectos relacionados con el transporte por carretera, como pueden ser elaboración y selección de rutas, determinación del número de vehículos, y optimización de los tiempos y flujos de carga, también se han empleado modelos matemáticos estocásticos para modelar ciertos aspectos de las redes de distribución.

Entre los diferentes modelos estocásticos empleados para simular redes de distribución por carretera se encuentran algunos llevados a cabo mediante la técnica de teoría de juegos. Örsan, en su trabajo de investigación empleó esta técnica de modelado para determinar los costes en el transporte por carretera de camiones (Örsan, 2008). Ozgun, en su trabajo modela mediante la teoría de juegos los problemas logísticos de una cadena de suministros (Ozgun, 2007).

En otras ocasiones, los modelos estocásticos se han empleado como complementos a los modelos deterministas, añadiendo ciertas características estocásticas. Un ejemplo es el problema de asignación de rutas a vehículos (VRP de sus siglas en inglés, vehicle routing problem) el cual paso a denominarse a denominarse SVRP, problema de asignación de rutas a vehículos estocástico, al añadir ciertos parámetros estocásticos. Los modelos VRP se basan en que un cierto número de localizaciones han de ser visitadas, y las características de las mismas, como por ejemplo, el peso de los envíos y recogidas es conocido. Sin embargo, en la mayoría de los casos esto no es así. Parámetros como la demanda de los clientes, el tiempo de trayecto entre clientes, e incluso las localizaciones a visitar son de carácter estocástico. Debido a estos aspectos, surgieron los modelos SVRP.

En este tipo de modelos, que son una combinación entre modelos deterministas y estocásticos, la parte estocástica está basada en procesos de decisión Markovianos (Gendreau et al, 1996). Bertsimas, añadió al problema de los VRP una demanda estocástica (Bertsimas, 1991). En 2004, el propio Bertsimas mejoró su trabajo anterior al incluir, además de una demanda estocástica, la aparición de clientes de manera estocástica (Bertsimas et al, 2004). Otros ejemplos con demanda y clientes estocásticos son Benton y Rossetti, en 1991, y Laporte y Louveaux en 1993. Al introducir estos aspectos estocásticos, el objetivo principal de los SVRP es la reducción de costes en las redes de distribución.

El problema de asignación de rutas a vehículos, siguió evolucionando. Así, después del SVRP se desarrolló el VRP con aplicaciones meta-heurísticas (Cordeau et al, 2002), en los que se empleaban técnicas como búsquedas Tabu, técnicas deterministas o algoritmos genéticos (Eksioglu et al, 2009). Posteriormente se desarrolló el problema de asignación de rutas dinámico (Eksioglu et al, 2009) el cual consiste en que algunos de los datos de entrada en el modelo son revelados o actualizados durante los periodos de tiempo en los cuales las operaciones de transporte se llevan a cabo (Berbeglia, et al, 2010). Estos modelos de asignación de rutas a vehículos dinámicos también se resuelven con algoritmos heurísticos, salvo en algunos casos en los que se introducen nuevamente parámetros estocásticos como elementos de entrada al algoritmo como en el trabajo de Novoa y Storer (Novoa y Storer, 2009), o el de Lei et al, (Lei et al, 2011).

Como ya se ha comentado, la técnica que se pretende emplear en este trabajo de investigación es el análisis de regresión. Uno de los principales objetivos de esta tesis es demostrar que esta técnica es válida para modelar redes de distribución capilar, ya que hasta ahora no se ha empleado esta técnica para modelar dichas redes de distribución capilar. A modo de ejemplo se presentan unos trabajos de investigación en los que se emplea regresión lineal para modelar ciertos aspectos relacionados con el sector del transporte. Kunkel y Schwind mediante un análisis de regresión son capaces de determinar la tarifa que hay que asignar a un nuevo cliente de una empresa de paquetería industrial (Kunkel y Schwind, 2011)

Kennedy empleó regresión lineal para determinar el volumen de vehículos y de peatones que circulaban simultáneamente por intersecciones, y para determinar el número de posibles atropellos en esas intersecciones en su trabajo de investigación (Kennedy, 2008). Fu en su tesis doctoral modeló mediante regresión lineal como se vería afectado el tráfico de vehículos en situaciones de emergencia como la de un huracán (Fu, 2004).

A pesar de no haberse empleado hasta ahora regresión lineal para modelar redes de distribución capilar, es una técnica que ofrece total garantías. Para corroborar esta aseveración se ha estudiado el trabajo de investigación de Xuelei en que analiza la solidez, la robustez y lo adecuado de los modelos de regresión (Xuelei, 2006).

METODOLOGÍA

1. Introducción

Tras presentar el caso de estudio y la revisión de la literatura en el estado del arte, en este cuarto capítulo se va a analizar la metodología seguida en este trabajo de investigación para alcanzar los objetivos marcados en él.

En el diagrama de flujo ilustrado en la figura 1 se puede observar las principales etapas que se llevarán a cabo para completar la investigación y resolver de esta manera las hipótesis de investigación lanzadas en la introducción. Cada una de estas etapas se analizará en profundidad en este apartado.

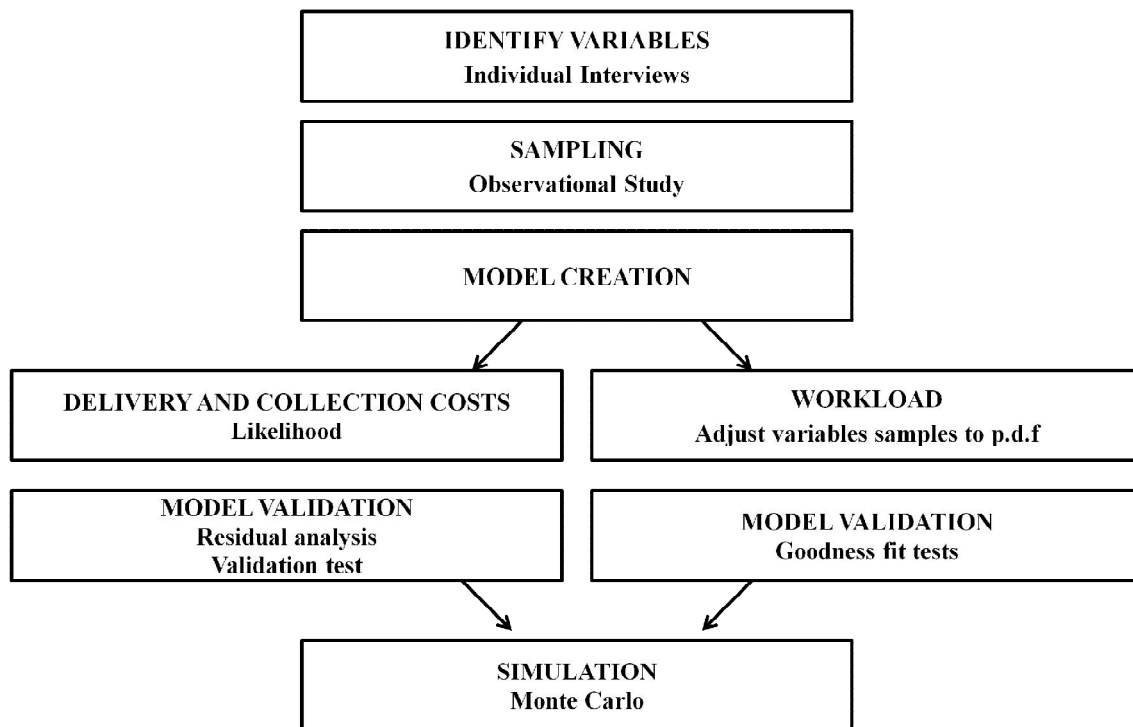


Figure 1: Summary of the methodology followed during the development of the stochastic model.

Para llevar a cabo esta investigación se va a desarrollar un modelo matemático con el que poder realizar simulaciones con el objetivo de conocer mejor el funcionamiento el sistema a estudio. El modelo elegido es un modelo matemático, debido a la gran versatilidad que este tipo de modelo ofrece. Esta versatilidad ha hecho que sean aplicados a múltiples sistemas de muy diferente índole, llegando a ser una de las tres metodologías más empleados por ingenieros industriales, científicos y en investigación operativa (Banks, 1998).

La hipótesis de investigación en este trabajo es comprobar si se puede emplear un análisis de regresión para modelar una red de distribución capilar. Se ha decidido emplear esta técnica porque hasta ahora la gran mayoría de modelos elaborados para modelar estas redes han sido deterministas. A pesar de que la elección de esta técnica ha sido arbitraria, esto no es óbice para que a priori esta técnica no sea válida. La mayoría de las redes de distribución tienen un comportamiento estocástico por naturaleza, sin embargo, es muy complicado el resolver este tipo de problemas estocásticos. Debido a esta razón, estos modelos son formulados de tal manera que finalmente puedan ser reducidos a un problema determinista (Laporte, 1992). Si las redes de distribución en general son de carácter estocástico por naturaleza, hay que buscar métodos estocásticos capaces de modelar dichas redes. La técnica de análisis de regresión es la técnica estadística más empleada en la actualidad y debido a sus múltiples usos, las aplicaciones de regresión son numerosas y las hay en casi cualquier campo (Montgomery et al, 2004). Es por estas razones por lo que se ha considerado oportuno el intentar demostrar que el análisis de regresión es una técnica válida para modelar redes de distribución capilar.

2. Identificación de variables

Como el objetivo principal de este trabajo de investigación es demostrar que la técnica de regresión lineal es una técnica válida con la que modelar una red de distribución capilar. Recordemos por un momento la definición de regresión lineal: “El análisis de regresión es una metodología estadística que emplea una relación entre dos o más variables cuantitativas, de tal manera que una variable respuesta o salida esté relacionada con una o varias de estas variables cuantitativas” (Kutner et al, 2004).

De esta definición la primera conclusión que se obtiene es que para poder realizar un análisis de regresión en primer lugar es necesario conocer que variable es la que se desea modelar, y en segundo lugar, que otras variables están relacionadas con la variable que se desea modelar, para posteriormente, se debe determinar cómo es esta relación.

Por lo tanto, el primer paso de esta metodología será averiguar cuál es la variable salida que estamos interesados en modelar, para poder estimar sus valores, y cuáles son las variables que están relacionadas con la anterior, mediante las cuales va a ser posible estimar la variable respuesta. Para poder determinar estas variables se ha empleado una aproximación al método Delphi, el cual se basa en consultar a expertos familiarizados con el sistema en análisis para conocer en profundidad el mismo (Marasco, 2008). Para ello se han realizado una serie de entrevistas a personas que trabajan en la empresa de paquetería industrial en estudio. Las personas seleccionadas para estas entrevistas son empleados de la empresa de paquetería industrial que su quehacer diario está relacionado y ligado con la red de distribución capilar que se quiere modelar. Las personas seleccionadas para estas entrevistas son el director de la terminal, el subdirector de la terminal, el jefe de tráfico (persona encargada de asignar

diariamente los repartos a los vehículos) y el responsable de recogidas (persona encargada de asignar diariamente las recogidas a los vehículos). También se ha realizado entrevistas a dos conductores para conocer sus puntos de vista, y contrastarlos con la opinión de aquellos que crean y diseñan sus zonas de reparto.

Las entrevistas se han realizado de manera individual y asegurando la confidencialidad de las mismas, para que el entrevistado pueda expresarse con la seguridad de que sus comentarios no llegarán a sus superiores y no coartar así sus opiniones. El diseño de la entrevista es de suma importancia, ya que en función de cómo se elaboren éstos los datos obtenidos pueden estar sesgados. Las preguntas deben ser cortas, precisas y deben estar expresadas de tal manera que sean fáciles de entender. Debe cuidarse el orden de las preguntas para evitar fomentar determinadas respuestas (Santos et al, 2003).

Estas entrevistas constan de dos partes. En la primera se les presentan cinco campos (Conductor, vehículo, delegación y personal, zona de reparto/recogida y política comercial), y mediante una pregunta abierta se le pide a cada entrevistado que conteste a la siguiente pregunta: *“¿Qué es necesario en cada uno de estos campos para que cada zona de reparto/recogida funcione correctamente?”* A partir de estas entrevistas, se obtienen las variables que cada entrevistado considera más importantes, y en la segunda parte, se les muestran todas las variables indicadas por todos los entrevistados en cada campo y se les pide que las puntúen en función de la importancia que ellos creen que tienen, independientemente que sean del mismo campo.

En el siguiente capítulo de esta tesis, análisis de resultados, se expondrá detalladamente cada una de las entrevistas realizadas y como se identifican las variables a considerar para poder modelar una red de distribución capilar. Pero para poder entender ciertos pasos o decisiones que se exponen en esta metodología es necesario conocer cuáles son estas variables. Estas variables son:

- i) los ingresos de los conductores por cada zona de reparto/recogida, que es lo mismo que los costes de la red de distribución para la empresa de paquetería industrial,
- ii) la carga de trabajo, la distancia a la que se encuentra la zona de reparto/recogida,
- iii) la extensión de la zona de reparto/recogida.

3. Toma de muestras

Una vez identificadas todas las variables que afectan al funcionamiento de cada una de las zonas de reparto y recogida, el siguiente paso es la toma de muestras de cada una de estas variables. La toma de muestras es uno de las etapas más importantes y laboriosa en la creación y la elaboración de un modelo. Cualquier modelo será tan bueno como los datos en los que esté basado (Vincent, 1998).

Para la obtención de muestras existen dos tipos de estudios que se pueden llevar a cabo, el estudio observacional y un experimento diseñado.

En los estudios observacionales, una vez conocidas las variables que se desean medir, se obtienen las muestras a través de la observación del sistema sin perturbar ni interaccionar con el proceso. Si este estudio se planea con anterioridad y se realiza adecuadamente puede proporcionar datos exactos, completos y fiables. Por el contrario, este tipo de estudios proporcionan información muy limitada acerca de las relaciones entre los datos, y puede conducir a problemas de colinealidad (Ryan, 2007).

Por el contrario, en los experimentos diseñados se recolectan los datos de las variables que se desean estudiar mediante una estrategia bien definida, la cual consiste en una serie de experimentos, en los cuales se puede modificar los valores de cada factor o variable. El principal inconveniente que presenta este tipo de estudios es que en ocasiones puede resultar muy complejo y costoso el llevar a cabo una serie de experimentos diseñados (Montgomery et al, 2004).

A la hora de elegir entre un estudio u otro se debe alcanzar un compromiso entre la complejidad de obtener los datos y el potencial explicativo de los mismos (Kleinbaum et al, 1988). En este caso, se ha decidido llevar a cabo un estudio observacional ya que resulta de gran rapidez y sencillez el obtener todos los datos del sistema informático de la empresa de paquetería industrial. Otra ventaja que ofrece en este caso el estudio observacional es que conociendo de antemano las variables que se quieren medir, es posible el desarrollar una aplicación informática mediante la cual los datos pueden ser presentados en la forma más interesante para realizar el análisis de regresión.

Antes de explicar cómo se lleva a cabo la toma de muestras es interesante recordar que una zona de reparto/recogida puede estar compuesta por un único código postal, por varios códigos postales, o un código postal puede estar compartido por más de un chofer. Por lo tanto, para poder analizar adecuadamente cada una de estas cuatro variables identificadas en el apartado anterior es necesario realizar el estudio por código postal de estas variables, y posteriormente, determinar si cada uno de los códigos postales analizados puede conformar por sí solo una zona de reparto/recogida, si es necesario el conjunto de más de un código postal para crear una zona de reparto, o por el contrario si es necesario el dividir un código postal entre varios conductores. De esta manera, al realizar este estudio por códigos postales se puede conocer los ingresos que un chofer tendría en el caso hipotético de que se le asignase ese código postal. De la misma manera, se podrá conocer la carga de trabajo en ese código postal, y la distancia y la extensión de ese código postal.

De las cuatro variables que se han identificado en el apartado anterior, tanto la distancia a la zona de reparto como la extensión de la misma, no se han considerado variables aleatorias, por lo que no es

necesario tomar muestras de las mismas. Solamente será necesario realizar una tabla descriptiva con estas características. De las otras dos, sin embargo, sí que será necesario.

En el caso de estudio ya se analizó cómo se obtienen los costes de distribución en la red capilar de esta empresa de transporte de paquetería industrial. Por un lado, están los costes debidos a los repartos, y por otro, los costes asociados a las recogidas. Estos costes son independientes los unos de los otros. Tanto los costes de reparto como de recogida son análogos. Hay una parte fija por realizar un reparto, y otra variable en función del peso del reparto. Las cantidades de la parte fija como de la variable, no son las mismas para todos los repartos, están tabuladas en función del peso de los mismos. Cuanto mayor es el peso, mayor es la parte fija; mientras que a mayor peso, menor es la parte variable (ver tabla 1). Por lo tanto, se puede observar como los costes de reparto dependen del número de repartos y del peso de cada uno de esos repartos. Y en el caso de las recogidas, del número de recogidas y del peso de cada una de ellas. Como los costes de recogida y reparto son independientes, se realizará por cada código postal un análisis de regresión de los costes de reparto, y otro de los costes de recogida.

Debido a que los conductores trabajan en un régimen a producción, estos costes de distribución dependen de los días trabajados. Como los días laborales varían de un mes a otro, no es posible el crear un modelo que estime los costes mensuales de manera adecuada, por lo que se ha decidido realizar un modelo que estime los costes diarios. Estos costes dependen del número de repartos realizados por día y del peso total repartido por día. Este peso total repartido por día es la suma unitaria de los pesos de cada uno de los repartos. Siendo el caso de las recogidas análogo.

Por lo tanto, las variables costes diarios de reparto, peso total recogido por día, número de repartos por día, costes diarios de recogida, peso total recogido por día y número de recogidas por día se consideran como las variables más representativas para la toma de muestras y determinar así los costes de distribución.

Tras comprobar que no existen variaciones estacionales en ninguna de estas variables mediante pruebas de comparación de medias, el tamaño de la muestra seleccionado será de cuatro meses, lo que reportará 84 observaciones.

Una vez que se disponen de las 84 observaciones, el siguiente paso es evaluar la independencia de éstas e investigar si existe algún tipo de correlación entre ellas. Para realizar estas comprobaciones se han empleado dos técnicas heurísticas (Guasch et al, 2003):

- Gráficas de autocorrelación: Estas gráficas evalúan la independencia de las observaciones de una muestra calculando los valores de autocorrelación. Si se considera un retardo j , se calcula la correlación de todos los pares de observaciones O_i y O_{i+j} . Si la muestra tiene n pares, la autocorrelación de retardo j se calcula empleando $n-j$ pares de observaciones. Por lo tanto, i va desde 1 hasta $n-j$.

Los valores obtenidos para los retardos entre 1 y 10 son los que mayor información aportan, mientras que a partir del retardo 20 no aportan ninguna información. Los valores de autocorrelación se denominan coeficientes de autocorrelación, y se calculan mediante la ecuación 36.

$$R_j = \frac{C_j}{S^2(n)} \quad \text{Ecuación (36)}$$

Donde $S^2(n)$ es la varianza muestral, y C_j (Ecuación 37) es la función de autocovariancia

$$C_j = \frac{\sum_{i=1}^{n-j} [O_i - \bar{O}(n)][O_{i+j} - \bar{O}(n)]}{n-j} \quad \text{Ecuación (37)}$$

Para considerar la muestra independiente, los valores de R_j deben estar entre -1 y 1.

- Diagrama de dispersión: Los diagramas de dispersión es una gráfica de los pares (O_i, O_{i+1}) , para todo $i=1,2,\dots, n-1$. Si las observaciones son independientes, los puntos estarán distribuidos aleatoriamente en la gráfica. Si por el contrario están correlacionados, los puntos se distribuirán siguiendo una línea, bien con pendiente positiva, o bien con pendiente negativa.

Para la toma de muestras de la carga de trabajo, hay que indicar que ésta se compone del número de repartos y recogidas, y del peso total repartido y recogido. Como las muestras de estas variables ya se tomaron para el caso de la variable costes de distribución, esas mismas muestras serán las que se empleen para el estudio de la variable carga de trabajo.

Como ya se ha expuesto, se dispone de 84 observaciones obtenidas del estudio observacional, y se ha comprobado la independencia de todas las observaciones de las muestras. Tras esta comprobación, estas 84 observaciones se han de separar en dos grupos. El primer grupo de observaciones se denominará muestra de calibración, y se empleará para realizar el análisis de regresión y crear el modelo. Mientras tanto, el segundo se denominará muestra de predicción, y servirá para comprobar la capacidad de predicción de los datos y la validez del modelo (Snee, 1977). La separación de estas dos muestras se realizará mediante asignación aleatoria. Se calcularán una serie de 84 números aleatorios o *stream*, que siga una distribución uniforme entre 0 y 1. A cada uno de los números aleatorios le corresponderá una

observación. Si el número aleatorio es menor que 0.5, su observación correspondiente será incluida en la muestra de predicción. Si el número aleatorio es mayor o igual que 0.5, la observación será incluida en la muestra de calibración. Para este trabajo, la generación de los números aleatorios para la separación de las muestras se lleva a cabo mediante el software Minitab.

4. Creación del modelo

Una vez que se disponen las muestras debidamente tomadas, y se conoce que son adecuadas, el siguiente paso es la elaboración del modelo.

Para este trabajo de investigación, se ha decidido realizar un estudio por código postal de cada una de las variables, y crear así un modelo por cada código postal. Sin embargo; como las variables a modelar son dos, costes de distribución y carga de trabajo por código postal, es necesario crear dos modelos diferentes por código postal, uno para estimar los costes de distribución por día y el otro para estimar la carga de trabajo por día.

4.1. Modelo de los costes de distribución capilar

Tal y como se ha explicado en el apartado 3, los costes debidos a los repartos y los costes debidos a las recogidas son independientes. Al ser éstos independientes también por cada código postal es posible realizar un análisis de regresión con el que crear un modelo para poder estimar los costes de reparto por día, y otro modelo con el que poder estimar los costes de recogida por día. Las salidas de estos modelos, dependerán del número de reparto y del peso total de estos repartos, y del número de recogidas y del peso total de estas recogidas, respectivamente. A modo de ejemplo se presenta en la tabla 3 las muestras obtenidas de los costes de reparto para un determinado código postal.

En la tabla 3 se observa, como para el día 1, en este código postal había asignados 34 repartos, con peso total a repartir ese día de 1.212 kg., y con un coste para la empresa de paquetería industrial a 106 euros. El peso total repartido por día es la suma del peso unitario de cada uno de los 34 repartos. Mientras que los 106 euros de coste para ese día 1, es la suma del coste unitario de cada uno de los 34 repartos.

Tabla 3: Costes por reparto, repartos por día y peso total repartido por día

Día	REPARTOS			Día	REPARTOS		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	34	1.212	106	43	19	975	68
2	55	2.076	177	44	27	1.122	90
3	50	1.671	155	45	24	1.143	82
4	42	1.259	133	46	27	576	75
5	34	1.764	124	47	30	1.449	107
6	37	1.440	120	48	35	1.103	108
7	39	1.070	117	49	20	1.004	73
8	40	1.121	115	50	31	1.271	100
9	42	1.643	142	51	30	810	88
10	23	556	66	52	40	1.799	145
11	34	902	101	53	37	1.208	120
12	33	975	99	54	32	1.246	111
13	30	1.080	93	55	37	1.306	120
14	47	1.365	143	56	26	718	80
15	32	1.175	104	57	41	1.558	137
16	36	1.348	116	58	28	820	88
17	45	1.600	146	59	29	1.051	93
18	29	1.288	97	60	25	968	86
19	41	1.183	123	61	32	1.379	110
20	34	875	101	62	20	899	71
21	48	2.057	166	63	24	1.067	85
22	32	1.120	107	64	26	1.423	96
23	38	1.265	122	65	25	1.077	83
24	30	1.227	102	66	28	885	88
25	36	1.145	114	67	29	1.101	96
26	28	1.105	94	68	27	1.224	94
27	11	262	31	69	31	1.402	109
28	51	1.241	147	70	22	1.021	80
29	55	1.758	168	71	40	1.391	129
30	49	2.070	170	72	28	1.217	96
31	47	1.627	156	73	31	868	96
32	40	1.602	134	74	24	1.333	87
33	40	1.543	137	75	27	784	86
34	41	1.055	121	76	34	763	100
35	39	2.303	141	77	23	569	71
36	27	1.138	90	78	7	86	18
37	46	1.203	143	79	42	1.651	139
38	39	1.860	130	80	31	2.133	114
39	35	947	102	81	25	1.527	101
40	38	1.222	118	82	25	800	82
41	44	1.653	141	83	31	1.255	103
42	25	1.088	86	84	37	1.490	125

Como los datos de cada observación de la muestra obtenidos mediante el estudio observacional es por día, el análisis de regresión estimará los costes por día. De esta manera, mediante los modelos de

regresión creados se podrán estimar los costes para la empresa de paquetería (y por tanto, los ingresos para los conductores), derivados del reparto en un determinado código postal por día. El caso de las recogidas es análogo, con una tabla con la misma forma y mismas características. Se creará otro modelo de regresión que permitirá estimar los costes de recogida en un determinado código postal por día.

Por último queda explicar cómo en función de los pesos de los repartos y las recogidas, se elaborará una tabla similar a la tabla 3 para cada tipo de vehículo, en función de su peso máximo autorizado de los mismos. En el caso particular de esta empresa de paquetería los repartos se dividen entre los diferentes tipos de vehículos. Todos los repartos de menos de 300 kg. los realizan los vehículos de 1500 kg. de peso máximo autorizado, y los repartos de más de 300 kg. lo realizan los vehículos de 3500 kg. Por lo general, no suele haber repartos de más de 3500 kg. de peso máximo autorizado. Si se presentase alguno de manera excepcional, se buscará una solución al reparto de este tipo de envíos con un vehículo de 3500 kg., haciéndolo en varias tandas.

El caso de las recogidas es distinto, ya que todas las recogidas de hasta 700 kg. las realizan los vehículos de 1500 kg. de peso máximo autorizado. Las recogidas entre 700 y 3500 kg las realizan los vehículos de 3500 kg. de peso máximo autorizado. Y para todas aquellas recogidas de más de 3500 kg. es necesario vehículos de peso máximo autorizado de 24000 kg.

Debido a esta clasificación, en cada código postal es necesario conocer los pesos de los diferentes repartos y recogidas, y realizar una tabla como la tabla 3 para cada tipo de vehículo. Por lo tanto, para cada código postal puede que sea necesario el realizar 5 análisis de regresión: por cada vehículo de 1500 kg. de peso máximo autorizado uno de reparto y otro de recogida; si se dan repartos de más de 300 kg. y recogidas de más de 700 kg., hay que realizar otros dos análisis de regresión para los vehículos de 3500 kg. de peso máximo autorizado; y si hay recogidas de más de 3500 kg., otro análisis de regresión, en este caso sólo de recogida, para los vehículos de 24000 kg.

4.1.1 Análisis de regresión para el modelo de los costes de reparto y recogida

Tras explicar los diferentes análisis de regresión que se deben realizar para poder estimar los costes de reparto y recogida de cada código postal, y de esta manera modelar la red de distribución capilar, en la sección 4.1.1.1. se va exponer como se realizarán estos análisis.

4.1.1.1. Elaboración del modelo

Con la muestra de calibración obtenida para la elaboración del modelo, el siguiente paso es el ajuste del modelo. En este trabajo de investigación la elaboración del modelo, o ajuste del modelo, se va a

realizar mediante el método de máxima verosimilitud. El empleo de este procedimiento presenta la principal ventaja de que realizando unas modificaciones no muy complejas, permite ajustar los modelos de regresión con características diferentes. Por ejemplo, el método de mínimos cuadrados asume que los errores se distribuyen según una distribución normal. Mediante el método de máxima verosimilitud, es posible ajustar el modelo bajo esa asunción, pero también permite ajustarlos suponiendo que los errores se distribuyen con cualquier otra distribución, y comparar si estos ajustes con distribuciones de probabilidad no normales aumentan la exactitud del análisis de regresión ó no, mediante la comparación de los valores de máxima verosimilitud. Incluso, mediante máxima verosimilitud se puede ajustar el modelo bajo los principios de los modelos lineales generalizados y comparar los modelos con estos ajustes con los realizados previamente, comprobando de esta manera si un análisis de regresión no lineal es mejor que uno lineal.

En definitiva, el método de máxima verosimilitud ha sido elegido por la flexibilidad que ofrece a la hora de ajustar diferentes modelos de regresión y la posibilidad de poder comparar unos con otros observando los valores de máxima verosimilitud.

De acuerdo con lo expuesto en el apartado 9.5 del capítulo 2, la función de máxima verosimilitud que se empleará para el ajuste los modelos se ilustra en la siguiente ecuación:

$$L(\hat{\beta}, \hat{\sigma}^2) = f(y; \hat{\beta}, \hat{\sigma}^2) = \prod_{i=1}^n f(y_i; \hat{\beta}, \hat{\sigma}^2) \quad \text{Ecuación (38)}$$

donde los estimados de máxima verosimilitud son los estimados de los parámetros $\hat{\beta}$ que hacen que la primera derivada de $f(y; \hat{\beta}, \hat{\sigma}^2)$ se anule para las diferentes funciones de distribución de probabilidad que los errores puedan tomar. Como las derivadas de un producto son más complicadas de resolver que las de una suma, esta función se transformará en una escala logarítmica, transformando de esta manera el productorio en un sumatorio.

En este trabajo en lugar de obtener los estimados a través de la primera derivada, se empleará un método iterativo mediante una hoja de cálculo Excel para obtener dichos estimados, ya que los cálculos se agilizan de esta manera (Law y Kelton, 1991).

Distribuciones de probabilidad de los errores

En esta sección se expondrán las diferentes funciones de distribución de probabilidad de los errores empleadas en este trabajo de investigación. Éstas han sido tres, la distribución normal, distribución doble exponencial y la distribución logística. Se han seleccionado las funciones de

distribución de probabilidad doble exponencial y logística, además de la normal, porque sus dominios se extiende desde $-\infty$ a $+\infty$.

Función de distribución de probabilidad normal. Método de mínimos cuadrados.

En primer lugar, el modelo de regresión se va ajustar suponiendo que los errores se distribuyen de acuerdo a una distribución normal. En este caso, si los errores se distribuyen de manera normal, los estimados de los parámetros de regresión calculados mediante máxima verosimilitud son los mismos que si se hubiera aplicado el método de los mínimos cuadrados, y además estos parámetros de regresión son los de menor varianza (Shao, 2003), de acuerdo con el teorema de Gauss-Markov.

La ecuación 39 representa la función de densidad de una distribución de probabilidad normal para la variable endógena es

$$f(y_i) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1(y_i-\mu_i)^2}{2\sigma^2}} \quad \text{Ecuación (39)}$$

La función de máxima verosimilitud queda

$$L(y_i; \hat{\beta}, \hat{\sigma}^2) = \prod_1^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1(y_i-\mu_i)^2}{2\hat{\sigma}^2}} \quad \text{Ecuación (40)}$$

Y transformándola a la escala logarítmica

$$\ln L(y_i; \hat{\beta}, \hat{\sigma}^2) = \ln \prod_1^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1(y_i-\mu_i)^2}{2\hat{\sigma}^2}} = \ln \sum_i^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1(y_i-\mu_i)^2}{2\hat{\sigma}^2}} \quad \text{Ecuación (41)}$$

El valor medio de la variable endógena es igual a $X\hat{\beta}$ (sección 8.7 del capítulo 2), por lo que si se sustituye en la ecuación 39 se obtiene la función 40, en la cual las variables independientes son $\hat{\sigma}^2$ y $\hat{\beta}$.

$$\ln L(y_i; \hat{\beta}, \hat{\sigma}^2) = \ln \sum_i^n \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1(y_i-X_i\hat{\beta})^2}{2\hat{\sigma}^2}} = \ln \frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2} \sum_1^n \frac{(y_i-X_i\hat{\beta})^2}{\hat{\sigma}^2} \quad \text{Ecuación (42)}$$

Por lo tanto, solo hay que descubrir que valores de estas variables independientes son las que consiguen un valor de máxima verosimilitud mayor. Para calcularlos se empleará un proceso iterativo.

Función de distribución de probabilidad doble exponencial

Si los errores no se distribuyen de acuerdo a una distribución normal, es probable que sea debido a que la distribución tenga unas colas más pesadas que la de una distribución normal. Una de las

distribuciones con colas más pesadas que la normal es la distribución doble exponencial (Montgomery et al, 2004).

La ecuación 43 representa la función de densidad de una distribución de probabilidad doble exponencial para la variable endógena

$$f(y_i) = \frac{1}{2\sigma} e^{-\frac{|y_i - \mu_i|}{\sigma}} \quad \text{Ecuación (43)}$$

La función de máxima verosimilitud queda

$$L(y_i; \hat{\beta}, \hat{\sigma}) = \prod_1^n \frac{1}{2\hat{\sigma}} e^{-\frac{|y_i - \mu_i|}{\hat{\sigma}}} \quad \text{Ecuación (44)}$$

Y transformándola a la escala logarítmica

$$\ln L(y_i; \hat{\beta}, \hat{\sigma}) = \ln \prod_1^n \frac{1}{2\hat{\sigma}} e^{-\frac{|y_i - \mu_i|}{\hat{\sigma}}} = \ln \sum_i^n \frac{1}{2\hat{\sigma}} e^{-\frac{|y_i - \mu_i|}{\hat{\sigma}}} \quad \text{Ecuación (45)}$$

Al igual que en el caso anterior, sustituimos μ_i por $X_i\hat{\beta}$.

$$\ln L(y_i; \hat{\beta}, \hat{\sigma}) = \ln \sum_i^n \frac{1}{2\hat{\sigma}} e^{-\frac{|y_i - X_i\hat{\beta}|}{\hat{\sigma}}} = -\ln 2\hat{\sigma} - \sum_i^n \frac{|y_i - X_i\hat{\beta}|}{\hat{\sigma}} \quad \text{Ecuación (46)}$$

En esta ecuación las variables independientes son $\hat{\beta}$ y $\hat{\sigma}$, y mediante el mismo proceso iterativo se determinan los valores que maximizan el valor de máxima verosimilitud.

Función de distribución de probabilidad logística

Otra de las distribuciones de probabilidad cuyas colas son más pesadas que la de la normal es la distribución logística. La ecuación 47 representa la función de densidad de una distribución de probabilidad doble exponencial para la variable endógena

$$f(y_i) = \frac{e^{-\frac{y_i - \alpha}{\vartheta}}}{\vartheta \left(1 + e^{-\frac{y_i - \alpha}{\vartheta}}\right)^2} \quad \text{Ecuación (47)}$$

Y transformándola a la escala logarítmica

$$\ln L(y_i; \hat{\vartheta}, \hat{\alpha}) = \ln \prod_1^n \frac{e^{-\frac{y_i - \hat{\alpha}}{\hat{\vartheta}}}}{\hat{\vartheta} \left(1 + e^{-\frac{y_i - \hat{\alpha}}{\hat{\vartheta}}}\right)^2} = \ln \sum_i^n \frac{e^{-\frac{y_i - \hat{\alpha}}{\hat{\vartheta}}}}{\hat{\vartheta} \left(1 + e^{-\frac{y_i - \hat{\alpha}}{\hat{\vartheta}}}\right)^2} \quad \text{Ecuación (48)}$$

En el caso de la distribución logística, el parámetro $\hat{\alpha}$ es el valor medio, por lo que se sustituye éste por $X_i\hat{\beta}$ en la ecuación 48, y de la ecuación resultante 49 calculamos los valores de los parámetros de regresión $\hat{\beta}$ y $\hat{\vartheta}$ que maximizan L .

$$\ln L(y_i; \hat{\vartheta}, \hat{\alpha}) = \ln \sum_i^n \frac{e^{-\frac{y_i - X_i\hat{\beta}}{\hat{\vartheta}}}}{\hat{\vartheta} \left(1 + e^{-\frac{y_i - X_i\hat{\beta}}{\hat{\vartheta}}}\right)^2} = -\frac{y_i - X_i\hat{\beta}}{\hat{\vartheta}} - \ln \hat{\vartheta} - 2 \ln \left(1 + e^{-\frac{y_i - X_i\hat{\beta}}{\hat{\vartheta}}}\right)^2 \quad \text{Ecuación (49)}$$

Modelos lineales generalizados

Tras el estudio de los valores de máxima verosimilitud para las diferentes distribuciones de probabilidad descritas en la sección 4.1.1.1., posteriormente se aplicará bajo los principios de los modelos lineales generalizados ciertas transformaciones a las variables exógenas para verificar si los valores de máxima verosimilitud obtenidos tras estas transformaciones son mayores. En este trabajo de investigación se van a emplear las funciones de cadena logarítmica, recíproca y raíz cuadrada. Recordar que si tras alguna de las transformaciones el valor de máxima verosimilitud es mayor, los procedimientos para estimar intervalos de confianza de los parámetros de regresión, como el de la respuesta del modelo, han de ser diferentes, ya que la propiedad de homocedasticidad no se cumple. En estos casos se puede aplicar la inferencia de Wald (Gonzales, 2001).

Los modelos lineales generalizados solo se aplican a distribuciones de la familia exponencial, por lo que en este trabajo solo se aplicarán a la distribución normal y doble exponencial, y no a la logística.

4.1.1.2. Selección de variables

Una vez que se conoce el procedimiento con el que se va a elaborar el modelo, hay que considerar también el procedimiento mediante el cual se van a determinar que las variables exógenas que realmente influyen en el comportamiento de la variable endógena de manera sustancial. Es decir, el modelo ha de cumplir el principio de parsimonia, para lo cual, éste debe de contener el menor número de variables exógenas necesario para ajustar los datos. Es decir, el mejor modelo es el modelo más simple que explica los hechos observados (Navidi, 2006). Por tanto, aquellas variables que no tienen influencia en la predicción de la variable endógena deben ser excluidas del modelo.

El método de máxima verosimilitud ofrece la posibilidad de determinar en qué medida la adición de una nueva variable al modelo aporta un mejor ajuste a los datos, mediante el test del cociente de verosimilitud. En cada uno de los ajustes explicados en el apartado 4.1.1.1., se va introducir las variables exógenas de una en una, y se estimará los valores de los parámetros de regresión que maximizan la

función de verosimilitud. El valor máximo de verosimilitud obtenido con $k+1$ variables se ha de comparar con el obtenido para k variables.

Para ello, se empleará es estadístico desviación, ecuación 50

$$\lambda(\beta) = \lambda\left(\frac{\text{modelo con } k + 1 \text{ variables}}{\text{modelo con } k \text{ variables}}\right) =$$
$$= \lambda(\beta_{\text{modelo con } k + 1 \text{ variables}}) - \lambda(\beta_{\text{modelo con } k \text{ variables}}) \quad \text{Ecuación (50)}$$

Como la adición de variables se realiza de una en una, la desviación $\lambda(\beta)$ se comportará como una distribución X^2 con un grado de libertad. Para determinar si una variable exógena más aporta un mejor ajuste, se realizará un contraste de hipótesis en la que la hipótesis nula (H_o) es que la nueva variable no aporta un mejor ajuste y la hipótesis alternativa (H_I) que la nueva variable si aporta un mejor ajuste.

Y los criterios de decisión son

Si $\lambda(\beta) \geq X_{\alpha,1}^2$ se rechaza la hipótesis nula

Si $\lambda(\beta) < X_{\alpha,1}^2$ no se rechaza la hipótesis nula

Resumiendo, el método de máxima de verosimilitud se ha empleado tanto para ajustar el modelo como para determinar que variables introducir en el mismo para que cumpla el principio de parsimonia.

4.1.1.3. Adecuación del modelo

Tras el ajuste del modelo con el mínimo número de variables exógenas, habrá que comprobar la adecuación del modelo y crear unos intervalos de confianza para los estimados de los parámetros de regresión estimados. Los parámetros de regresión y los estimados de la variable endógena son variables aleatorias, y las estimaciones puntuales de sus valores no serán iguales a los valores verdaderos; por tanto será necesario el poder ofrecer un intervalo de confianza de estos valores para un nivel de significancia dado, con el objetivo de poder determinar estos valores.

En este trabajo, al emplear el método de máxima verosimilitud, se han comparado diferentes modelos de regresión, buscando aquel que ajuste mejor los datos. En función de que el modelo que mejor ajuste sea un modelo lineal generalizado, de mínimos cuadrados o con distribuciones de probabilidad diferentes a la normal los pasos para comprobar la adecuación del modelo y calcular los intervalos de confianza son diferentes.

Método de mínimos cuadrados

De acuerdo con el teorema de Gauss-Markov, mediante el método de mínimos cuadrados se obtiene los estimados de los parámetros de regresión de menor varianza, siempre y cuando se cumplan los siguientes supuestos:

- Todos los errores e_i tienen media cero.
- Todos los errores e_i están distribuidos normalmente.
- Todos los errores e_i son aleatorios e independientes. La magnitud de cualquier error e_i no influye en el valor del siguiente error e_{i+1} .
- Todos los errores e_i cumplen el principio de homoscedasticidad. Es decir, la varianza de todos los errores es constante para todas las observaciones.

Por tanto, para comprobar la adecuación del modelo es necesario verificar que estos supuestos se cumplan. Un análisis gráfico de los residuales es un medio eficaz para diagnosticar violaciones de las asunciones anteriores (Navidi, 2006). Estos análisis gráficos son tres. El primero es una gráfica de probabilidad normal, la cual está diseñado de tal manera que si los errores se distribuyen de manera normal, la distribución normal acumulada que siguen se grafica como una línea recta.

El segundo es una gráfica de residuales en función de los valores ajustados \hat{y}_i . Con esta gráfica se puede verificar la homoscedasticidad y la linealidad. Si los residuales se encuentran dentro de una banda horizontal, y no tiene forma de embudo o doble embudo, los supuestos de homoscedasticidad y linealidad se cumplen.

Y por último, hay que llevar a cabo una gráfica de residuales contra su orden de observación. Esta gráfica permite determinar si los errores en un periodo de tiempo están relacionados con los de otros periodos, es decir, muestran si son independientes. Para que los errores sean independientes, en esta gráfica los residuales se distribuyen a lo largo de la gráfica sin ningún tipo de patrón determinado.

Una vez que estos supuestos son verificados, se está en disposición de calcular los intervalos de confianza de los parámetros de regresión, así como otros parámetros que indican la calidad del modelo.

El primero de ellos es la estimación de la varianza del error σ^2 . Debido a que la magnitud de la varianza se refleja en el grado de dispersión de los puntos alrededor de la recta de mínimos cuadrados, se tiene que midiendo esta dispersión se puede calcular la varianza. El primer paso para calcularla es obtener la suma de cuadrados del error, o de los residuales, el cual indica el error de predicción del modelo debido a la falta de ajuste del mismo respecto a los datos (Webster, 2000). Y se calcula mediante la ecuación 51

$$SSE = \sum_{i=1}^n (y_i - \hat{y}_i)^2 = \sum_{i=1}^n e_i^2 = \mathbf{e}'\mathbf{e} \quad \text{Ecuación (51)}$$

Y partir de la ecuación 51, se calcula un estimador insesgado de la varianza, ecuación 52 (Seber, 2000).

$$\hat{\sigma}^2 = \frac{SSE}{n-p} \quad \text{Ecuación (52)}$$

Donde p es el número de parámetros en el modelo de regresión.

Otro parámetro que se emplea para contrastar la calidad del modelo es R^2 ajustada, el cual indica la proporción de la variación explicada por las variables exógenas, pero con la particularidad que penalizada la adición de términos que no son útiles. Para calcular este parámetro es necesario conocer primero la suma total de cuadrados (SST) el cual indica la diferencia entre cada uno de los valores individuales y el valor medio de éstos (ecuación 53) (Webster, 2000).

$$SST = \sum_i^n (y_i - \bar{y})^2 \quad \text{Ecuación (53)}$$

Siendo R^2 ajustada se expresa mediante la ecuación 54,

$$R^2_{adj} = 1 - \frac{SSE/(n-p)}{SST/(n-1)} \quad \text{Ecuación (54)}$$

Una vez calculados estos parámetros, se calculan los intervalos de confianza de los parámetros de regresión. Los estimados de los parámetros de regresión son variables aleatorias y, por tanto, tendrán un grado de dispersión. Conocer este grado de dispersión es necesario para poder estimar unos intervalos de confianza de los mismos. Como el estimado de $\hat{\beta}$ por mínimos cuadrados es una combinación lineal de las observaciones, también está distribuido normalmente, con vector medio β y matriz de covarianzas $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$. Esto implica que un parámetro de regresión $\hat{\beta}_j$ es normal, con media β_j y varianza $\sigma^2 C_{jj}$, donde C_{jj} es el j -ésimo elemento de la diagonal de la matriz $(\mathbf{X}'\mathbf{X})^{-1}$ (Montgomery et al, 2004), siendo la función 55 su expresión

$$se(\hat{\beta}_j) = \sqrt{\hat{\sigma}^2 C_{jj}} \quad \text{Ecuación (55)}$$

Por tanto, el intervalo de confianza de $100(1-\alpha)$ por ciento para cualquier parámetro de regresión queda como ilustra la ecuación 56,

$$\hat{\beta}_j - t_{\frac{\alpha}{2}, n-p} \sqrt{\hat{\sigma}^2 C_{jj}} \leq \beta_j \leq \hat{\beta}_j + t_{\frac{\alpha}{2}, n-p} \sqrt{\hat{\sigma}^2 C_{jj}} \quad \text{Ecuación (56)}$$

Por último, obtenidos los intervalos de confianza de los parámetros de regresión, solo queda calcular los intervalos de predicción de la estimación de la variable endógena. Aquí se pueden dar dos casos, uno en el que se quiera obtener el valor medio de la variable endógena para unos determinados

valores de las variables exógenas, y el segundo en el que se quiera obtener el estimado puntual de la variable endógena para unos determinados valores de las variables exógenas.

Si se considera el vector \mathbf{X}_0 como los valores de las variables exógenas para los cuales se quiere determinar los estimados de la variable endógena \hat{y}_0 , para el caso de la respuesta media, el intervalo de predicción de $100(1-\alpha)$ por ciento de nivel de confianza se calcula mediante la función 57

$$\hat{y}_0 - t_{\frac{\alpha}{2}, n-p} \sqrt{\hat{\sigma}^2 \mathbf{X}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_0} \leq y_0 \leq \hat{y}_0 + t_{\frac{\alpha}{2}, n-p} \sqrt{\hat{\sigma}^2 \mathbf{X}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_0} \quad \text{Ecuación (57)}$$

Lográndose el intervalo para el estimado puntual mediante la ecuación 58

$$\hat{y}_0 - t_{\frac{\alpha}{2}, n-p} \sqrt{\hat{\sigma}^2 (1 + \mathbf{X}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_0)} \leq y_0 \leq \hat{y}_0 + t_{\frac{\alpha}{2}, n-p} \sqrt{\hat{\sigma}^2 (1 + \mathbf{X}'_0 (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}_0)} \quad \text{Ecuación (58)}$$

Distribuciones de probabilidad no normales

En el método de mínimos cuadrados se ha supuesto que los errores se distribuyen normalmente y con varianza constante, y partir de esas suposiciones se han estimado la varianza, el parámetro R^2 ajustada, y los intervalos de confianza de los parámetros de regresión de los estimados de la variable endógena.

Sin embargo, si los errores no se distribuyen de manera normal, pero son independientes y están idénticamente distribuidos, las inferencias bajo la teoría de normalidad son asintóticamente correctas por el teorema central del límite (Owen, 1991). Es decir, los procedimientos empleados en el apartado anterior para el método de los mínimos cuadrados son válidos siempre y cuando los errores se distribuyan independiente e idénticamente. Sin embargo, el principio de homoscedasticidad si que ha de ser respetado.

Por tanto, una vez verificados estos dos supuestos, que los errores sean independientes e idénticamente distribuidos, y que la varianza sea constante, los pasos a seguir serán idénticos a los del apartado anterior.

Modelos lineales generalizados

En el caso de los modelos lineales generalizados el principio de homoscedasticidad no se cumple, por lo que los procedimientos empleados en los dos apartados anteriores no son válidos. Para poder realizar inferencias e intervalos de confianza de los parámetros de regresión se ha de emplear la inferencia de Wald (Gonzales, 2001).

Sea la matriz \mathbf{G} (ecuación 59), denominada hessiana o de Hess, de las segundas derivadas parciales de la función logaritmo de verosimilitud, de tamaño $p \times p$,

$$\mathbf{G}(\hat{\boldsymbol{\beta}}) = \frac{\partial^2 L(\hat{\boldsymbol{\beta}})}{\partial \hat{\beta}_i \partial \hat{\beta}_j} \quad \text{Ecuación (59)}$$

Para muestras grandes, la matriz de covarianzas de los parámetros de regresión queda de acuerdo a la ecuación 60.

$$\text{Var}(\hat{\boldsymbol{\beta}}) = -\mathbf{G}(\hat{\boldsymbol{\beta}})^{-1} \quad \text{Ecuación (60)}$$

Siendo las raíces cuadradas de los elementos de la diagonal de esta matriz los errores estándar de los parámetros de regresión para muestras grandes, $se(\hat{\beta}_j)$.

Quedando los intervalos de confianza de $100(1-\alpha)$ por ciento para los parámetros de regresión según se puede observar en la ecuación 61

$$\hat{\beta}_j - Z_{\frac{\alpha}{2}} se(\hat{\beta}_j) \leq \beta_j \leq \hat{\beta}_j + Z_{\frac{\alpha}{2}} se(\hat{\beta}_j) \quad \text{Ecuación (61)}$$

Por último, obtenidos los intervalos de confianza de los parámetros de regresión, solo queda calcular los intervalos de confianza de la estimación de la variable endógena. Si se considera el vector \mathbf{X}_0 como los valores de las variables exógenas para los cuales se quiere determinar los intervalos de confianza para la respuesta media de la variable endógena \hat{y}_0 , se tiene la ecuación 62

$$\hat{y}_0 = \hat{\mu}_0 = g^{-1}(\mathbf{X}'_0 \hat{\boldsymbol{\beta}}) \quad \text{Ecuación (62)}$$

Considerando la matriz $\boldsymbol{\Sigma}$ (ecuación 63) como la matriz de varianza-covarianza asintótica para $\hat{\boldsymbol{\beta}}$, que es

$$\boldsymbol{\Sigma} = a(\varphi)(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \quad \text{Ecuación (63)}$$

La varianza asintótica del predictor lineal estimado \mathbf{X}_0 es

$$\text{Var}(\hat{\eta}_0) = \text{Var}(\mathbf{X}'_0 \hat{\boldsymbol{\beta}}) = \mathbf{X}'_0 \boldsymbol{\Sigma} \mathbf{X}_0 \quad \text{Ecuación (64)}$$

Así, un estimado de esta varianza es $\mathbf{X}'_0 \hat{\boldsymbol{\Sigma}} \mathbf{X}_0$, donde $\hat{\boldsymbol{\Sigma}}$ es la matriz estimada de varianza-covarianza de $\hat{\boldsymbol{\beta}}$ (Montgomery et al, 2004). El intervalo de confianza para un nivel de significancia de $100(1-\alpha)$ por ciento según la ecuación 65

$$g^{-1}\left(\mathbf{X}'_0 \hat{\boldsymbol{\beta}} - Z_{\frac{\alpha}{2}} \mathbf{X}'_0 \hat{\boldsymbol{\Sigma}} \mathbf{X}_0\right) \leq \mu(\mathbf{X}_0) \leq g^{-1}\left(\mathbf{X}'_0 \hat{\boldsymbol{\beta}} + Z_{\frac{\alpha}{2}} \mathbf{X}'_0 \hat{\boldsymbol{\Sigma}} \mathbf{X}_0\right) \quad \text{Ecuación (65)}$$

4.1.2. Validación del modelo

Tras la elaboración y la comprobación de la adecuación del modelo, la siguiente etapa en el desarrollo del modelo es la de su validación. Para llevar a cabo esta etapa se emplearán los datos de la

muestra de predicción en lugar de la muestra de calibración. El objetivo de la validación del modelo consiste en verificar que el modelo cumple los requisitos de acuerdo a los criterios objetivos y subjetivos de un buen modelo de aproximación (Johansson, 1993).

Un modelo se considera válido si responde correctamente a las preguntas que se desean formular sobre el sistema real, es decir, si es útil (Guasch et al, 2003). Todos los modelos no son más que aproximaciones de sistemas reales muy complejos, por lo que las preguntas que pretendan resolver han de estar en concordancia con los objetivos iniciales en los que se planteó la creación del modelo.

En la validación de los modelos de regresión de costes de reparto y recogida, se estudia la respuesta de variable endógena y se presentan los resultados en términos de su intervalo de confianza. Para ello, los datos de las variables exógenas de la muestra de predicción se introducen en el modelo y se comparan las salidas proporcionadas por éste con los datos de la variable endógena de la muestra de predicción que corresponden a cada conjunto de variables exógenas. Es decir, es una prueba pareada, con el mismo número de estimaciones que observaciones con las que compararlas, y se da una correspondencia uno a uno entre los valores estimados y los valores empleados para la validación. Como las observaciones de la muestra de predicción son independientes, se pueden aplicar métodos estadísticos clásicos. En este caso particular se va a emplear la prueba *t* pareada o prueba de Welch (Law y Kelton, 1991).

En esta prueba se construye un intervalo de confianza directamente con la diferencia entre cada par de valores para un nivel de significancia dado. El intervalo de confianza se va a llevar a cabo con *n* observaciones. Para cada pareja de datos se calcula su diferencia Z_j , y se define el valor medio de esta diferencia de acuerdo a la ecuación 66

$$\bar{Z}(n) = \frac{\sum_{j=1}^n Z_j}{n} \quad \text{Ecuación (66)}$$

Si las observaciones de las muestras son independientes e idénticamente distribuidas, $\hat{\sigma}^2$ se define según la ecuación 67

$$\hat{\sigma}^2[\bar{Z}(n)] = \frac{\sum_{j=1}^n (Z_j - \bar{Z}(n))^2}{n(n-1)} \quad \text{Ecuación (67)}$$

Quedando los intervalos de confianza de $100(1-\alpha)$ por ciento

$$\bar{Z}(n) - t_{1-\frac{\alpha}{2}, n-1} \sqrt{\hat{\sigma}^2[\bar{Z}(n)]} \leq Z(n) \leq \bar{Z}(n) + t_{1-\frac{\alpha}{2}, n-1} \sqrt{\hat{\sigma}^2[\bar{Z}(n)]} \quad \text{Ecuación (68)}$$

Si el intervalo incluye el valor cero, se considera que no hay diferencia significativa entre los valores estimados y los valores originales, por lo que se considera válido el modelo.

4.2. Modelo de la carga de trabajo

La carga de trabajo se define como el número de repartos y recogidas que deben realizarse en un día en un determinado código postal, así como el peso total que todos estos repartos y recogidas suponen.

Para modelar la carga de trabajo de cada código postal, cada una de las variables aleatorias de las que consta se va a ajustar a una distribución de probabilidad, identificando los parámetros característicos de la distribución a la que se ajuste. Para el caso de los repartos se va a ajustar las variables aleatorias el número de repartos por día, el peso unitario de cada reparto y el peso total repartido por día. En el caso de las recogidas es análogo, siendo las variables aleatorias el número de recogidas por día, el peso unitario de cada recogida y el peso total recogido por día. Para ello se va a emplear el software Minitab (Minitab Inc), el cual proporciona los valores de los parámetros característicos de cada distribución de probabilidad.

Las variables número de repartos y número de recogidas son variables aleatorias discretas. Debido a su naturaleza, en primer lugar se intentará ajustar éstas a alguna distribución de probabilidad discreta. Al tratarse de un conteo de un elemento, una distribución discreta razonable con la que tratar de ajustar una muestra de datos es la distribución de Poisson (Montgomery et al, 2004). Si el ajuste a una distribución de Poisson resulta insatisfactorio, se comprobará si mediante alguna distribución de probabilidad continua es posible llevar a cabo algún ajuste satisfactorio.

Para todas las variables aleatorias que se ajustan mediante distribuciones de probabilidad continuas, siempre que sea posible, se elegirán aquellas cuyo rango de aplicación sea mayor que cero. Esto es debido a que por la naturaleza de las variables aleatorias en estudio, todas toman valores positivos. El número de repartos y recogidas, así como los pesos totales repartidos y recogidos, nunca pueden ser negativos. Si las distribuciones de probabilidad seleccionadas son aquellas con dominio positivo, se evita el tener que emplear distribuciones de probabilidad truncadas, de aquellas distribuciones de probabilidad con rango de aplicación de menos infinito a más infinito.

4.2.1. Validación del modelo

Para comprobar la adecuación de los ajustes en este trabajo se van a realizar otros test no paramétricos. Para el caso de las distribuciones discretas se realizará el test de Chi cuadrado y para las distribuciones continuas el test de Kolmogorov-Smirnov. Estos test se emplean para validar si una cierta función de distribución puede utilizarse como patrón para generar valores con las mismas propiedades estadísticas que la muestra con la que se está trabajando.

4.2.1.1. Test Chi cuadrado

Este test los datos de la muestra con los datos hipotéticos de la distribución de probabilidad esperada. La ecuación 68 representa el estadístico empleado para llevar a cabo esta comprobación

$$X^2 = \sum_{i=1}^r \frac{(O_i - E_i)^2}{E_i} \quad \text{Ecuación (69)}$$

Donde O_i es la frecuencia observada de valores en el intervalo i , E_i es la frecuencia de las observaciones esperada si la función de distribución fuese correcta, y r es el número de intervalos empleados. Los pasos a seguir en la evaluación de la bondad del ajuste son (Milton y Arnold, 2004):

- Separar los datos por intervalos. Hay que tener en cuenta que en cada intervalo el número valores observados ha de ser mayor de cinco.
- Calcular las frecuencias esperadas observadas y las frecuencias esperadas para cada uno de los intervalos. Las frecuencias esperadas se calculan de acuerdo con la distribución de probabilidad que se quiere comprobar la bondad del ajuste.
- Calcular el valor del estadístico X^2 .
- Seleccionar un nivel de significancia α para la prueba, y evaluar la hipótesis de que las muestras siguen la distribución de probabilidad seleccionada. Para ello se ha de comparar el valor del estadístico X^2 con el valor de la distribución $X_{\alpha, \gamma}^2$, donde γ es igual a $r-p-1$ y son los grados de libertad de la distribución Chi cuadrado
- Si $X^2 \leq X_{\alpha, \gamma}^2$, se acepta la hipótesis de que los datos siguen la función de probabilidad con la que se ha comparado.

4.2.1.2. Test de Kolmogorov-Smirnov

Este test se basa en la discrepancia máxima entre los valores hipotéticos de la distribución de probabilidad acumulada esperada con la distribución de probabilidad acumulada obtenida de la muestra, para determinar si la distribución de probabilidad esperada se ajusta satisfactoriamente a los datos de la muestra. Este test también es conocido como test EDF ó test de distancia (D'Agostino y Stephens, 1986). Los pasos a seguir en este test de bondad de ajuste son:

- Ordenar todos los valores de la muestra de menor a mayor.
- Construir la función de distribución acumulada empírica $S(x)$ a partir de los valores de la muestra.
- Calcular la función de distribución de probabilidad acumulada hipotética $F(x)$ para cada una de los valores de la muestra.

- Para cada uno de los valores de la muestra calcular la diferencia entre

$$F(x_i) - S(x_i)$$

- Verificar la máxima discrepancia de las diferencias anteriores

$$D = \max|F(x_i) - S(x_i)|$$

- Escoger un valor de significancia α , y aceptar la hipótesis de que los datos se ajustan a la distribución de probabilidad hipotética si

$$D_n \leq c_{1-\alpha}$$

Los valores de $c_{1-\alpha}$ se pueden obtener de la siguiente tabla 4 (Law y Kelton, 1991).

Tabla 4: valores para el test Kolmogorov-Smirnov

Caso	D_n	$c_{1-\alpha}$				
		0.850	0.900	0.950	0.975	0.990
Parámetros conocidos	$\left(\sqrt{n} + 0.12 + \frac{0.11}{\sqrt{n}}\right)D$	0.138	1.224	1.358	1.480	1.628
Parámetros estimados	$\left(\sqrt{n} - 0.01 + \frac{0.85}{\sqrt{n}}\right)D$	0.775	0.819	0.895	0.955	1.035

4.2.1.3. Distribuciones empíricas

Si tras realizar las pruebas no paramétricas anteriores los datos de una muestra no se pueden ajustar a ninguna función de distribución de probabilidad, estos datos se ajustarán de acuerdo a alguna distribución de probabilidad empírica.

Si la distribución empírica es continua, se emplearán los valores de las observaciones para crearlas. Para ello se ordenaran estas observaciones de menor a mayor, siendo la ecuación 70 la distribución F la distribución empírica

$$F(x) = \frac{i-1}{n-1} + \frac{x-X_i}{(n-1)(X_{i+1}-X_i)} \quad \text{si } X_i \leq x < X_{i+1} \quad \text{Ecuación (70)}$$

Siendo $F(x)=0$ para $x < X_1$ y $F(x)=1$ para $X_n \leq x$

Si la distribución de probabilidad empírica es discreta, para cada posible valor de x la distribución de probabilidad queda definida como una función de masa $p(x)$ de acuerdo a la proporción de valores X_i iguales a x .

La principal desventaja que presenta el uso de funciones de distribución empíricas es que los valores aleatorios generados a partir de ella no pueden estar fuera de los límites de la muestra, por lo tanto solo se podrán estimar valores dentro de los límites de la muestra.

5. Simulación

Según Kleijnen el método de Monte Carlo es un procedimiento de simulación adecuado mediante el cual se puede conocer más acerca del sistema. Como la obtención de observaciones de una determinada distribución de probabilidad requiere el uso de números aleatorios, como más adelante se verá, este tipo de simulaciones estocásticas son conocidas como simulaciones Monte Carlo (Rubinstein, 1981).

En este trabajo de investigación, mediante número aleatorios se generarán los valores de variables aleatorias del modelo de carga de trabajo a través del método de la transformada inversa. Estos valores simulados de la carga de trabajo son los valores que se emplearán como datos de entrada para realizar la simulación del modelo de costes de reparto y recogida, y de esta manera obtener un estimado de la variable endógena. Si los estimados que se obtienen de diferentes repeticiones del proceso son independientes, se pueden aplicar la estadística clásica. Este proceso se repetirá un número suficiente de veces como para que sea posible aplicar el teorema central del límite. Todo este proceso se conoce como el método de las repeticiones (Guasch et al, 2003). En este caso se va a repetir el proceso 250 veces, que son los días laborables en un año completo, para modelar de esta manera los costes de reparto y recogida para un año completo.

5.1. Método de la transformada inversa

Mediante este método se va a poder generar una muestra aleatoria que sigue una determinada distribución de probabilidad, aprovechando las posibilidades de inversión de las funciones de distribución acumuladas.

Las funciones de distribución acumulada del modelo de carga de trabajo son del tipo $y=F(x)$, que realizan una aplicación del rango de valores de la variable x en el rango $[0,1]$.

$$F: R \rightarrow [0,1]$$

La inversa de estas funciones distribución acumuladas permite realizar una aplicación de un conjunto de valores definidos en el rango $[0,1]$ a valores del rango en el que se encuentra definida la variable aleatoria x .

$$F^{-1}: [0,1] \rightarrow R$$

De este modo, si se dispone de un conjunto de datos uniformemente distribuidos en el rango $[0,1]$, es posible mapearlos mediante la inversa de la función de distribución acumulada en un conjunto de datos en el rango para el cual se encuentra definida la variable aleatoria x , tales que presentan las mismas propiedades estadísticas que la variable aleatoria x .

Por tanto, generando una variable aleatoria u_i que siga una distribución uniforme en entre los valores 0 y 1, y conociendo las distribuciones de probabilidad acumuladas de las variables aleatorias a modelar, es posible realizar una simulación de las mismas. En este caso en particular, las series de números aleatorios con los cuales comienza la simulación son aportados por Minitab.

5.2. Simulación del modelo de carga de trabajo

A partir de los números aleatorios proporcionados por Minitab, y empleando el método de la transformada inversa, se va a simular la carga de trabajo de cada uno de los códigos postales, proporcionando para cada uno de los 250 días de los consta la simulación el número de repartos, el peso total repartido, el número de recogidas y peso total recogido. Para calcular tanto el peso total repartido como el peso total recogido por día, si es posible el determinar una distribución de probabilidad que se ajuste satisfactoriamente a estas dos variables aleatorias, se empleará ésta para simular el peso unitario de cada reparto y cada recogida, obteniendo el peso total como la suma unitaria de cada reparto y de cada recogida respectivamente. El número de repartos y recogidas para los cuales habrá que calcular el peso se habrá obtenido previamente.

Este número de repartos y recogidas, son variables aleatorias discretas. Si el ajuste de las muestras a una distribución de probabilidad sólo es posible mediante una distribución continua, a la hora de aplicar el método de la transformada inversa, el valor resultante no será un número entero. Como los resultados del número de repartos y recogidas se han de expresar como número enteros, los valores decimales obtenidos se redondearán a valor entero más próximo.

Si no es posible el ajustar ninguna distribución de probabilidad a la variable peso de cada reparto y peso de cada recogida, el peso total repartido y peso total recogido se calculará a través del método de la transformada inversa igualmente, pero empleando en este caso la distribución de probabilidad de la variables peso total de repartido y peso total recogido. Es por esta razón, por la que se realizó el ajuste de la variable peso total repartido y peso total recogido, para que en los casos en los que no fuese posible ajustar las variables aleatorias peso de cada reparto y peso de cada recogida, poder simular el peso total repartido y recogido.

Los resultados que se presentan para la simulación de la carga de trabajo son el valor medio, la desviación estándar, el valor máximo y el valor mínimo para cada una de las variables aleatorias.

5.3. Simulación de los costes de reparto y recogida

Una vez que se dispone de los valores simulados de la carga de trabajo de cada código postal para cada uno de los 250 días de los que consta la simulación, se calculan los costes de reparto y recogida para cada uno de estos días. Para ello, se introducen en los modelos realizados previamente para los costes de reparto y coste de recogida los valores de la carga de trabajo. En los casos que haga falta, también se introducen estos valores elevados a la segunda potencia. De esta manera están disponibles los costes de reparto y recogida para cada uno de los 250 días.

Para cada una de las salidas de estos modelos, de acuerdo con lo descrito en el apartado 4.1.1.3., se determina un intervalo de predicción. De esta manera se puede determinar entre que dos valores estará la predicción de cada salida, es decir, en que intervalo estará la salida para la carga de trabajo simulada. Una vez que se disponen los 250 intervalos de predicción, se realizará un estudio para los límites inferior y superior de estos intervalos. Se presentará el valor medio, la varianza y la desviación estándar de estos límites, así como un intervalo de confianza de los valores que estos límites pueden tomar para un nivel de confianza del 95%.

ANALISIS DE RESULTADOS

En este quinto capítulo, se muestra los resultados obtenidos en este trabajo de investigación de acuerdo con los pasos expuestos en la metodología. En primer lugar se expondrán las diferentes entrevistas llevadas a cabo para identificar las variables que hay que modelar mediante el análisis de regresión. Una vez conocidas éstas, se presentará para cada uno de los códigos postales los diferentes modelos efectuados mediante análisis de regresión con los que simular los costes de la red de distribución, explicando en cada caso las decisiones tomadas, así como los modelos de carga de trabajo mediante probabilidades de distribución. Por último, se expondrá para cada uno de los códigos postales los resultados de la simulación de los diferentes modelos.

1. Resultados de las entrevistas

Las entrevistas llevadas a cabo en este trabajo de investigación, debido al gran volumen que ocupan se han colocado en el anexo I de esta tesis. En este apartado se muestra un resumen de las conclusiones derivadas de ellas.

En las entrevistas llevadas a cabo al director, al subdirector de la terminal y al encargado de recogidas, los tres coinciden en que la principal variable para que una zona de reparto y recogida funcione adecuadamente es la continuidad del conductor del vehículo de la zona, ligada muy directamente a los ingresos que el conductor obtiene en su trabajo. Tras la entrevista al preguntarles el por qué, argumentaban que cuando un chofer lleva mucho tiempo en una zona desarrolla una afinidad con los clientes que es vital para la continuidad de los clientes, ya que son los propios conductores los que miman a los clientes, siendo estos clientes en cierta manera de los conductores, ya que corresponden a su zona de reparto. Los clientes confían en los conductores, y los conductores conocen las necesidades cada cliente y lo que cada cliente desea y espera de ellos. Al conseguir la continuidad de los clientes, los ingresos de los conductores no se ven reducidos, e incluso aumentan, y la motivación de conductores aumenta, al trabajar a producción. Por tanto, es como un círculo vicioso. Si los ingresos de los conductores son adecuados su continuidad se garantiza, y al garantizar esta continuidad, sus ingresos no se ven mermados, o incluso aumentan, lo que a su vez aumenta su motivación y su deseo de realizar más repartos y recogidas, y de manera más adecuada.

Tras indicar que la principal variable a considerar es los ingresos de los choferes, también coinciden que las siguientes variables a considerar son la distancia de la zona de reparto, la extensión de los mismos y el número de repartos y recogidas asignados a cada conductor por día. Respecto a la distancia a la zona de reparto, argumentaron que todo el tiempo que el conductor tarda en llegar a la zona

de reparto y volver es tiempo efectivo que pierde para realizar repartos y recogidas, por lo que cuanto mayor sea, menos tiempo dispondrá para hacer entregas y recogidas.

Si la extensión de la zona es muy grande, es posible que no lleguen a tiempo de poder realizar todos los repartos y recogidas, ya que pierden tiempo efectivo de reparto y recogida en el desplazamiento de una parte a otra de la zona de reparto. Por otro lado, si la zona es muy extensa imposibilita a los conductores a volver a intentar repartos en caso de que la persona que recibe un envío esté ausente en el momento de un primer intento.

Por último, si a un conductor se le asigna un número de repartos y/o recogidas a realizar excesivo y desproporcionado es comprensible que no pueda llegar a realizar todos.

Por otro lado, el encargado de repartos coincide con los otros tres entrevistados pero en orden inverso. Las principales variables a considerar son la extensión de la zona, la distancia a la zona de reparto/recogida y el número de repartos y/o recogidas, alegando los mismos argumentos que sus compañeros. Tras estas tres variables, considera la continuidad y los ingresos de los conductores como las siguientes variables en importancia.

Como resultado de las entrevistas, se considera que las variables que hay que analizar y considerar para que una zona de reparto de una red de distribución capilar funcione adecuadamente son los ingresos de los conductores de reparto (que son los costes de la red de distribución para la empresa de paquetería industrial), la carga de trabajo de cada zona de reparto/recogida, la distancia a la misma y la extensión de ella.

Debido a la naturaleza de las variables identificadas como significativas para el correcto funcionamiento de una zona de reparto y recogida de una red de distribución capilar, se ha decidido no modificar la forma de trabajar de esta empresa de paquetería industrial, y asignar a los conductores una zona en lugar de una serie de clientes. El organizar las zonas de reparto/recogida de esta manera facilita la familiaridad de los conductores con los territorios en los que tienen que repartir y recoger mercancía. Al aumentar la familiaridad, la capacidad de trabajo aumenta debido a que se facilita la localización de direcciones y localizaciones en edificios, y la eficiencia al organizar sus rutas de reparto y recogida diarias. Es decir, se aumenta la eficiencia ante la aparición estocástica de clientes. Estas conclusiones no son descabelladas, ya que Zhong, Hall y otros llegaron a las mismas conclusiones en su trabajo en 2004.

Al asignar una zona de reparto en lugar de una serie de clientes, hay que analizar la carga de trabajo de esa zona y los ingresos que el conductor obtendrá debido a esta carga de trabajo, independientemente de los clientes que tenga que visitar cada día, asegurando una cantidad mínima de ingresos. Por tanto, las zonas de reparto se modelarán en función de los ingresos que los conductores

tengan y no en función de los clientes que tenga que visitar, de tal forma que se garantice la continuidad de los conductores.

2. Estudio de cada código postal

Tal y como se explicó en el capítulo de la metodología, las variables identificadas en apartado anterior se van a analizar por código postal. De esta manera, es posible conocer si un código postal puede conformar por si solo una zona de reparto y recogida, generando por si mismo unos ingresos tales que puedan asegurar la continuidad del conductor. Por otro lado, también hemos de conocer si la carga de trabajo que genera estos ingresos no es tan elevada como para que la calidad del servicio se vea mermada. Si se ha de conformar una zona de reparto/recogida con más de un código postal, se ha de considerar la extensión de la zona de reparto con el objetivo de que el chofer sea capaz de abarcarla en su totalidad sin que la calidad del servicio se vea disminuida.

A continuación se va a exponer por cada uno de los códigos postales de la población en estudio, los diferentes resultados obtenidos del estudio de cada una de las variables identificadas en el apartado 1 de este capítulo.

Como todos los códigos postales que se van a modelar pertenecen al mismo municipio la distancia a la zona de reparto será prácticamente la misma para los 7 códigos postales en estudio. Tal y como se vio en el caso de estudio, esta distancia es de 7,3 km. Al ser esta distancia tan reducida, permite a los conductores ir a la terminal en más de una ocasión por día, por lo que se reduce considerablemente cualquier problema de capacidad del vehículo.

2.1. Código postal 08198

El tipo de repartos que se dan en este código postal son mayores de 300 kg., y las recogidas las hay superiores de 700 kg. y superiores de 3500 kg. Esto indica que serán necesarios tres tipos de vehículos diferentes para poder satisfacer la demanda del mismo, uno vehículo de 1500 kg. de peso máximo autorizado, otro de 3500 kg. de peso máximo autorizado, y por último, uno de más de 3500 kg. Esto quiere decir que habrá que realizar un estudio completo para cada uno de los tres tipos de vehículos por separado.

2.1.1. Vehículo de 1500 kg. para el código postal 08198

2.1.1.1. Toma de muestras

De acuerdo a lo expuesto en la metodología, el primer paso para realizar el modelo estocástico para cada uno de los códigos postales, es la toma de muestras de las variables que se han identificado como las que han de ser modeladas. La primera que se quiere modelar es el coste de reparto y recogida por día en este código postal, y este tipo de vehículo. Los costes de reparto dependen del número de repartos y del peso total repartido en un día, mientras que los costes de recogida dependen del número de recogidas y el peso total recogido por día. Tras el estudio observacional realizado, las 84 observaciones relativas a los repartos se puede observar en la tabla 5:

Tabla 5: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08198

Día	REPARTOS			Día	REPARTOS		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	34	1.212	106	43	37	1.490	125
2	55	2.076	177	44	19	975	68
3	50	1.671	155	45	27	1.122	90
4	42	1.259	133	46	24	1.143	82
5	34	1.764	124	47	27	576	75
6	37	1.440	120	48	30	1.449	107
7	39	1.070	117	49	35	1.103	108
8	40	1.121	115	50	20	1.004	73
9	42	1.643	142	51	31	1.271	100
10	23	556	66	52	30	810	88
11	34	902	101	53	40	1.799	145
12	33	975	99	54	37	1.208	120
13	30	1.080	93	55	32	1.246	111
14	47	1.365	143	56	37	1.306	120
15	32	1.175	104	57	26	718	80
16	36	1.348	116	58	41	1.558	137
17	45	1.600	146	59	28	820	88
18	29	1.288	97	60	29	1.051	93
19	41	1.183	123	61	25	968	86
20	34	875	101	62	32	1.379	110
21	48	2.057	166	63	20	899	71
22	32	1.120	107	64	24	1.067	85
23	38	1.265	122	65	26	1.423	96
24	30	1.227	102	66	25	1.077	83
25	36	1.145	114	67	28	885	88
26	28	1.105	94	68	29	1.101	96
27	11	262	31	69	27	1.224	94
28	51	1.241	147	70	31	1.402	109
29	55	1.758	168	71	22	1.021	80
30	49	2.070	170	72	40	1.391	129
31	47	1.627	156	73	28	1.217	96
32	40	1.602	134	74	31	868	96
33	40	1.543	137	75	24	1.333	87
34	41	1.055	121	76	27	784	86
35	39	2.303	141	77	34	763	100
36	27	1.138	90	78	23	569	71
37	46	1.203	143	79	7	86	18
38	39	1.860	130	80	42	1.651	139
39	35	947	102	81	31	2.133	114
40	38	1.222	118	82	25	1.527	101
41	44	1.653	141	83	25	800	82
42	25	1.088	86	84	31	1.255	103

Los costes de recogida son análogos a los de reparto, y las 84 observaciones del estudio observacional se pueden ver en la tabla 6.

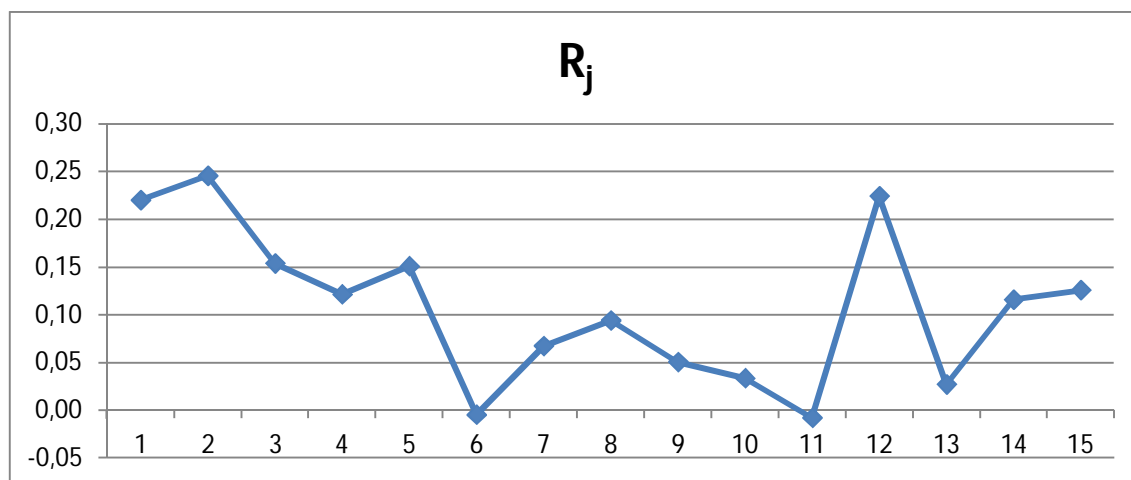
Tabla 6: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08198

Día	RECOGIDAS			Día	RECOGIDAS		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	18	668	62	43	17	1.695	81
2	17	2.725	99	44	10	1.314	55
3	14	2.552	89	45	13	1.267	64
4	16	1.921	79	46	13	1.456	65
5	19	1.991	91	47	12	1.739	68
6	15	1.749	77	48	16	1.249	70
7	17	2.175	87	49	17	1.592	82
8	11	1.635	63	50	12	1.423	61
9	20	2.413	100	51	19	1.914	88
10	18	2.115	90	52	15	1.447	69
11	17	1.532	79	53	15	1.658	77
12	17	2.080	87	54	22	2.064	104
13	18	2.045	89	55	16	2.295	92
14	17	2.078	90	56	18	3.370	114
15	19	1.238	72	57	25	1.627	101
16	25	1.631	99	58	21	2.269	98
17	19	2.239	98	59	22	1.364	84
18	17	1.032	69	60	16	1.327	71
19	19	1.996	95	61	22	3.270	121
20	17	2.425	92	62	17	3.087	104
21	17	2.637	96	63	19	2.720	105
22	15	2.305	84	64	18	2.256	99
23	23	2.210	106	65	15	2.077	81
24	17	1.677	77	66	17	3.543	114
25	17	1.004	70	67	17	1.971	81
26	7	436	29	68	14	1.015	62
27	15	3.222	103	69	16	2.382	90
28	18	2.442	96	70	19	2.065	93
29	14	1.535	65	71	10	1.921	66
30	19	2.282	100	72	16	1.137	64
31	19	1.979	94	73	21	2.096	97
32	22	1.606	94	74	17	2.129	90
33	15	2.368	89	75	19	1.741	92
34	23	1.569	100	76	21	2.157	100
35	16	2.191	88	77	9	1.068	49
36	18	2.137	94	78	17	2.757	98
37	20	2.665	106	79	6	69	18
38	16	1.019	71	80	18	1.657	85
39	19	2.588	103	81	14	1.963	80
40	14	1.596	69	82	20	2.133	98
41	16	2.334	92	83	17	1.819	83
42	19	1.712	85	84	11	880	53

2.1.1.2. Independencia de las muestras

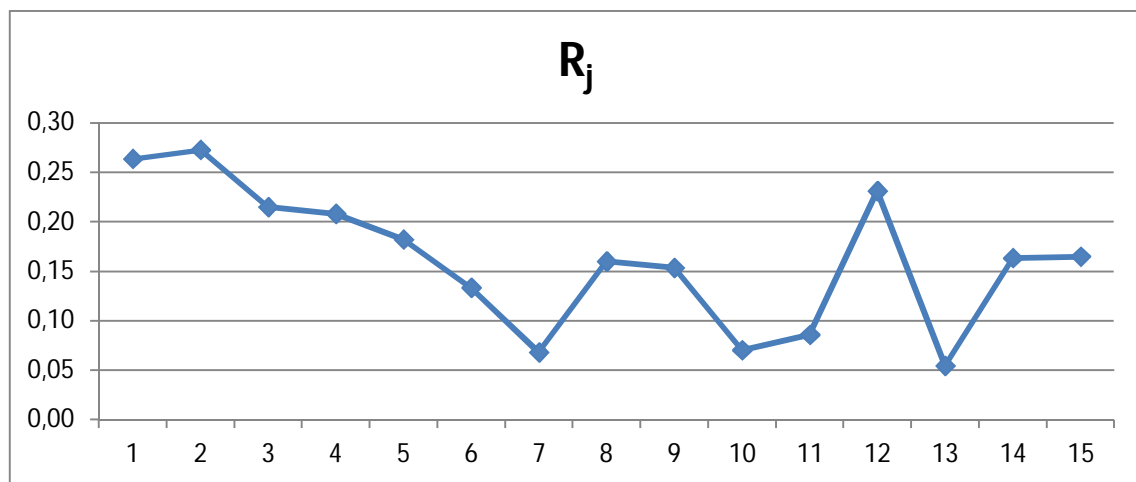
Tras la obtención de las muestras, hay que comprobar si las observaciones de las mismas son independientes. Que las observaciones de las muestras sean independientes es fundamental, ya que para emplear el método de máxima verosimilitud en la elaboración de los modelos de costes de reparto y recogida, las muestras utilizadas han de ser independientes. Por otro lado, los test no paramétricos de bondad de ajuste de las distribuciones de probabilidad también presuponen que las muestras empleadas son independientes.

En primer lugar se analiza la independencia de la muestras de los repartos. Como esta muestra se compone de tres variables, coste de reparto por día, número de repartos por día y peso total repartido por día, es necesario comprobar la independencia de cada una de ellas. En primer lugar se muestran las gráficas de autocorrelación. Aplicando las expresiones enunciadas en la metodología, se tabulan los valores R_j en función de los retardos j , para comprobar que sus valores se encuentran entre 1 y -1. La Figura 8 corresponde a la gráfica de autocorrelación de la muestra de los costes de reparto por día. La Figura 9 corresponde a la gráfica de autocorrelación del número de repartos por día, y la Figura 10 corresponde a la gráfica de autocorrelación del peso total repartido por día



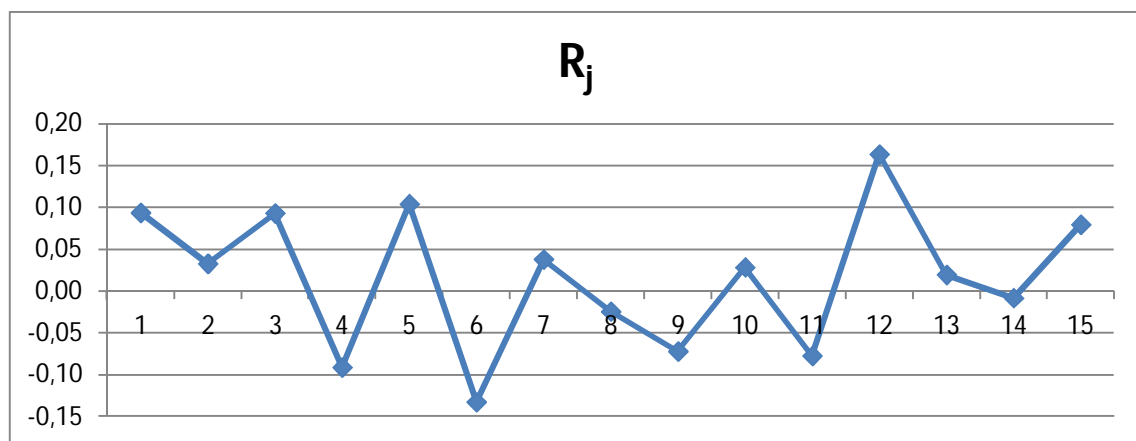
R_j	0,22	0,25	0,15	0,12	0,15	0,00	0,07	0,09	0,05	0,03	-0,01	0,22	0,03	0,12	0,13
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	183	204	128	101	126	-4	56	78	42	28	-7	187	23	97	105

Figura 8: Grafica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08198



R_j	0,26	0,27	0,22	0,21	0,18	0,13	0,07	0,16	0,15	0,07	0,09	0,23	0,05	0,16	0,16
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	21,8	22,6	17,8	17,2	15,1	11,0	5,6	13,3	12,7	5,8	7,1	19,1	4,5	13,5	13,7

Figura 9: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08198



R_j	0,09	0,03	0,09	-0,09	0,10	-0,13	0,04	-0,02	-0,07	0,03	-0,08	0,16	0,02	-0,01	0,08
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	14867	5192	14759	-14471	16538	-21064	6018	-3889	-11454	4497	-12286	25944	3103	-1330	12612

Figura 10: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08198

A continuación se muestran los gráficos de dispersión, que es la segunda prueba para comprobar la independencia de las observaciones. En ellos, para considerar las observaciones independientes los puntos se han de distribuir aleatoriamente en el plano. Al igual que para las gráficas de autocorrelación, se realizará una por cada una de las variables.

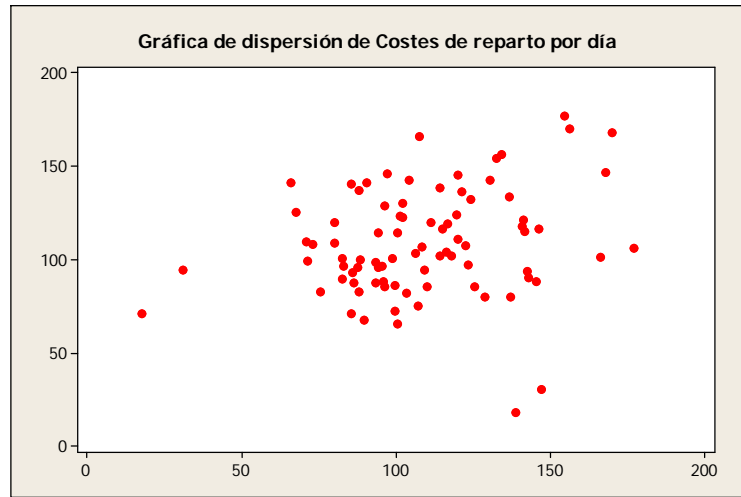


Figura 11: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08198

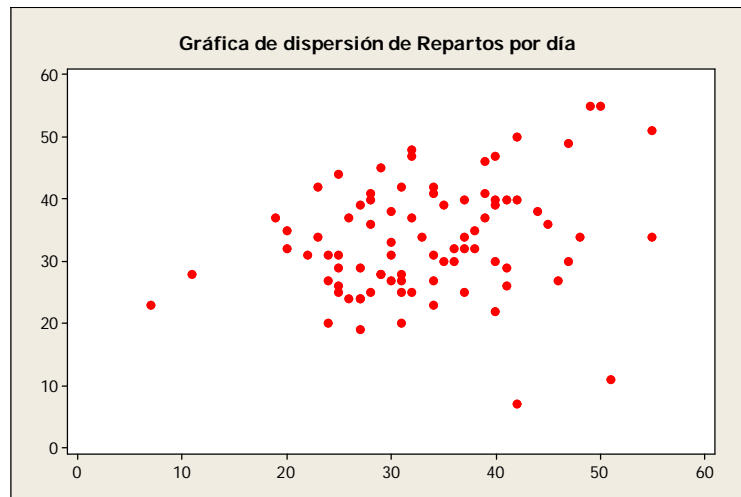


Figura 12: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08198

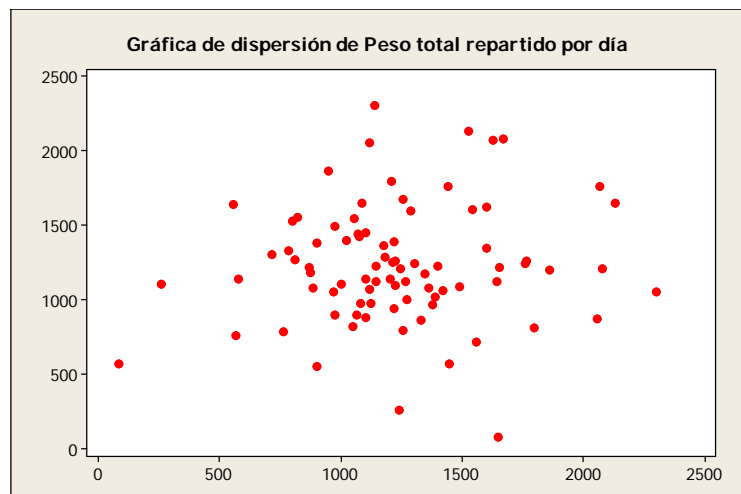
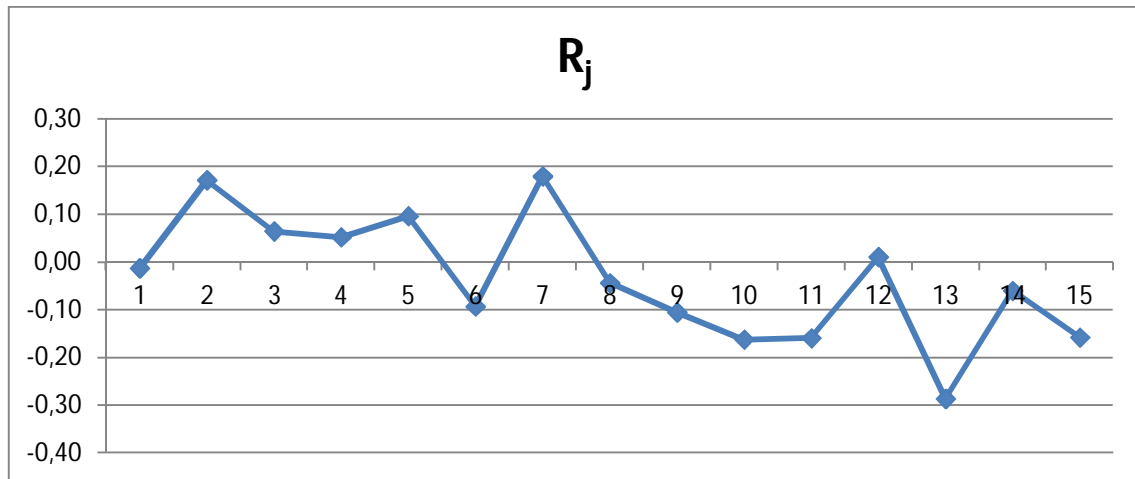


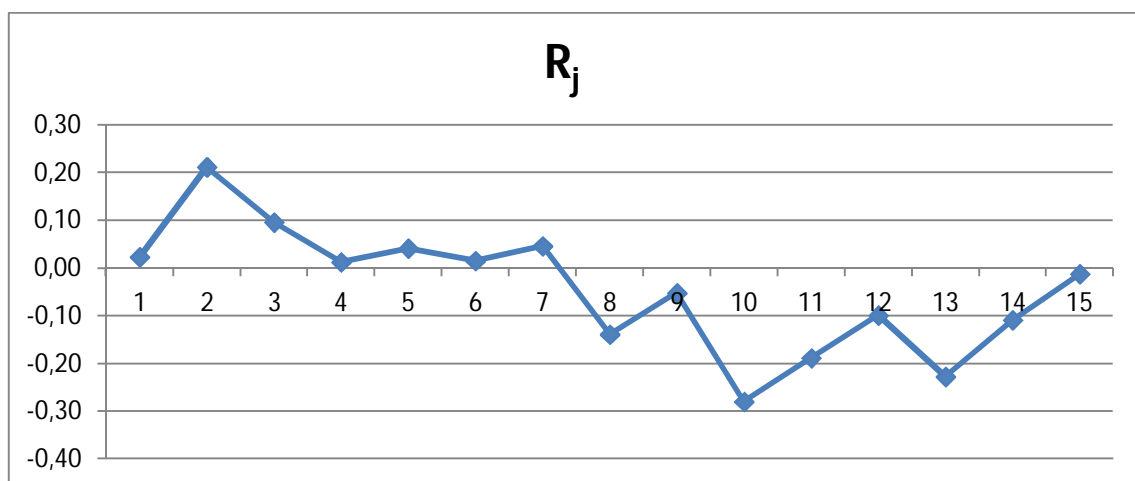
Figura 13: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08198

En segundo lugar, se analiza la independencia de la muestra de las recogidas. Al igual que en el caso de los repartos también existen tres variables y habrá que comprobar la independencia de las tres, costes de recogida por día, número de recogidas por día y peso total recogido por día. Primero se muestran las gráficas de autocorrelación:



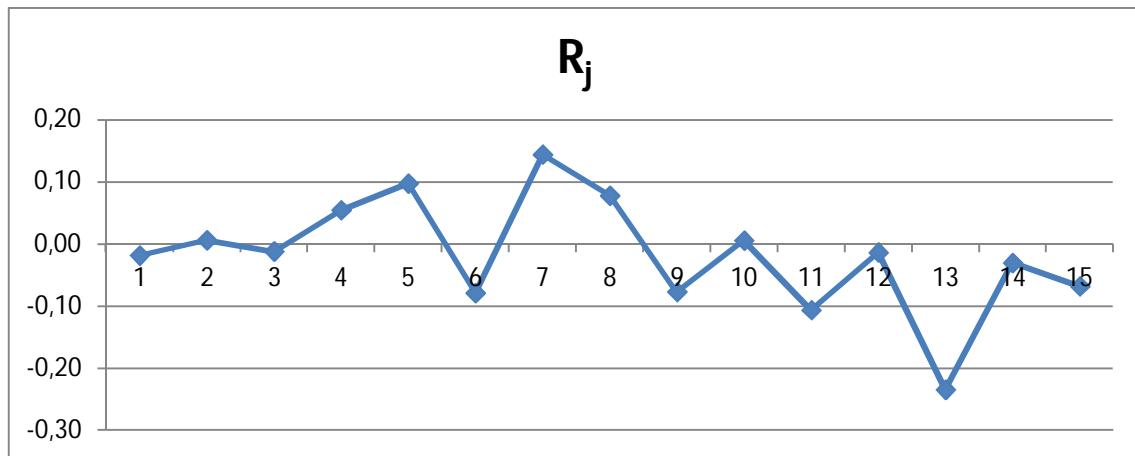
R_j	-0,01	0,17	0,06	0,05	0,10	-0,09	0,18	-0,04	-0,11	-0,16	-0,16	0,01	-0,29	-0,06	-0,16
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-4	55	21	17	31	-30	58	-14	-34	-52	-51	3	-92	-19	-51

Figura 14: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08198



R_j	0,02	0,21	0,10	0,01	0,04	0,02	0,05	-0,14	-0,05	-0,28	-0,19	-0,10	-0,23	-0,11	-0,01
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	0,3	2,7	1,2	0,2	0,5	0,2	0,6	-1,8	-0,7	-3,5	-2,4	-1,3	-2,9	-1,4	-0,2

Figura 15: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08198



R_j	-0,02	0,01	-0,01	0,05	0,10	-0,08	0,14	0,08	-0,08	0,01	-0,11	-0,01	-0,24	-0,03	-0,07
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-7504	2427	-4949	22045	39373	-32020	58259	31414	-31298	2231	-43159	-5596	-95088	-12336	-27668

Figura 16: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08198

Tras las gráficas de autocorrelación se exponen los diagramas de dispersión:

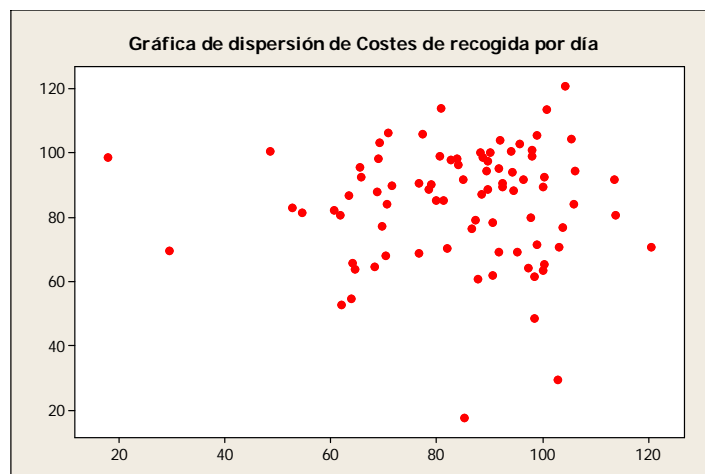


Figura 17: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08198

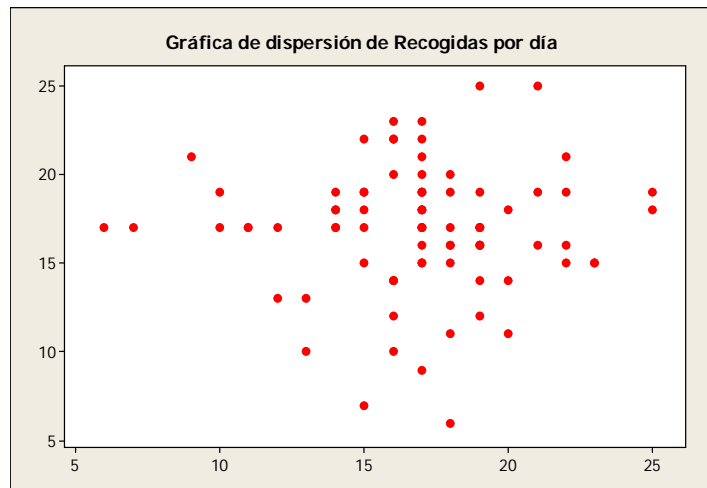


Figura 18: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08198

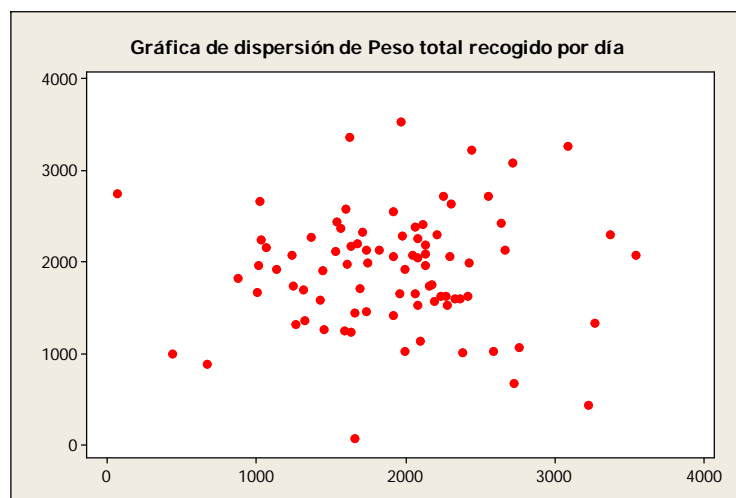


Figura 19: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08198

2.1.1.3. Separación de las muestras de calibración y predicción

Una vez que se concluye que las 84 observaciones de reparto y de recogida son independientes, éstas han de ser separadas en la muestra de calibración y la muestra de predicción. En este apartado de empleará la muestra de calibración, la cual se empleará para ajustar el modelo.

Esta separación se realizará mediante asignación aleatoria. Se asigna un número aleatorio proveniente de una distribución uniforme a cada una de las observaciones, y si éste es mayor o igual que 0,5, esta observación se incluirá en la muestra de calibración. Por el contrario, si es menor de 0,5, se incluye en la muestra de predicción.

La asignación de números aleatorios a cada una de las observaciones de reparto se muestra en la tabla 7.

Tabla 7: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08198

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,382	22	0,356	43	0,064	64	0,405
2	0,101	23	0,910	44	0,358	65	0,552
3	0,596	24	0,466	45	0,487	66	0,712
4	0,899	25	0,426	46	0,511	67	0,555
5	0,885	26	0,304	47	0,373	68	0,181
6	0,958	27	0,976	48	0,986	69	0,970
7	0,014	28	0,807	49	0,041	70	0,687
8	0,407	29	0,991	50	0,231	71	0,529
9	0,863	30	0,256	51	0,005	72	0,797
10	0,139	31	0,952	52	0,926	73	0,806
11	0,245	32	0,053	53	0,100	74	0,262
12	0,045	33	0,705	54	0,257	75	0,178
13	0,032	34	0,817	55	0,776	76	0,867
14	0,164	35	0,973	56	0,680	77	0,115
15	0,220	36	0,466	57	0,809	78	0,060
16	0,017	37	0,300	58	0,724	79	0,762
17	0,285	38	0,750	59	0,085	80	0,738
18	0,343	39	0,351	60	0,132	81	0,986
19	0,554	40	0,776	61	0,756	82	0,926
20	0,357	41	0,074	62	0,627	83	0,904
21	0,372	42	0,198	63	0,174	84	0,545

A partir de estas asignaciones, las muestras de calibración y de predicción para el modelo de costes de reparto quedan como sigue. En la tabla 8 se puede observar como la muestra de calibración consta de 40 observaciones, y la de predicción 44, debido a la aleatoriedad de la asignación de las observaciones en las muestras.

En la tabla 8 se puede observar, como para el día 1 de la muestra de calibración, para este código postal había asignados 50 repartos, con peso total para repartir ese día de 1.671 kg., y que el coste para la empresa de paquetería industrial por repartir esos 50 repartos ascendería a 155 euros. El peso total repartido por día es la suma del peso unitario de cada uno de los 50 repartos. Mientras que los 155 euros de coste para ese día 1, es la suma del coste unitario de cada uno de los 34 repartos.

Tabla 8: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08198

Día	REPARTOS. MUESTRA DE CALIBRACIÓN			Día	REPARTOS. MUESTRA DE PREDICCIÓN		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	50	1.671	155	1	34	1.212	106
2	42	1.259	133	2	55	2.076	177
3	34	1.764	124	3	39	1.070	117
4	37	1.440	120	4	40	1.121	115
5	42	1.643	142	5	23	556	66
6	41	1.183	123	6	34	902	101
7	38	1.265	122	7	33	975	99
8	11	262	31	8	30	1.080	93
9	51	1.241	147	9	47	1.365	143
10	55	1.758	168	10	32	1.175	104
11	47	1.627	156	11	36	1.348	116
12	40	1.543	137	12	45	1.600	146
13	41	1.055	121	13	29	1.288	97
14	39	2.303	141	14	34	875	101
15	39	1.860	130	15	48	2.057	166
16	38	1.222	118	16	32	1.120	107
17	24	1.143	82	17	30	1.227	102
18	30	1.449	107	18	36	1.145	114
19	30	810	88	19	28	1.105	94
20	32	1.246	111	20	49	2.070	170
21	37	1.306	120	21	40	1.602	134
22	26	718	80	22	27	1.138	90
23	41	1.558	137	23	46	1.203	143
24	25	968	86	24	35	947	102
25	32	1.379	110	25	44	1.653	141
26	26	1.423	96	26	25	1.088	86
27	25	1.077	83	27	37	1.490	125
28	28	885	88	28	19	975	68
29	27	1.224	94	29	27	1.122	90
30	31	1.402	109	30	27	576	75
31	22	1.021	80	31	35	1.103	108
32	40	1.391	129	32	20	1.004	73
33	28	1.217	96	33	31	1.271	100
34	27	784	86	34	40	1.799	145
35	7	86	18	35	37	1.208	120
36	42	1.651	139	36	28	820	88
37	31	2.133	114	37	29	1.051	93
38	25	1.527	101	38	20	899	71
39	25	800	82	39	24	1.067	85
40	31	1.255	103	40	29	1.101	96
				41	31	868	96
				42	24	1.333	87
				43	34	763	100
				44	23	569	71

Por otro lado, la asignación de números aleatorios a cada una de las observaciones de recogida se muestra es la tabla 9.

Tabla 9: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08198

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,378	22	0,044	43	0,138	64	0,145
2	0,209	23	0,866	44	0,694	65	0,081
3	0,180	24	0,854	45	0,842	66	0,560
4	0,814	25	0,551	46	0,245	67	0,658
5	0,963	26	0,908	47	0,122	68	0,066
6	0,794	27	0,716	48	0,864	69	0,298
7	0,669	28	0,480	49	0,125	70	0,044
8	0,835	29	0,638	50	0,567	71	0,560
9	0,773	30	0,265	51	0,990	72	0,028
10	0,736	31	0,236	52	0,696	73	0,256
11	0,768	32	0,810	53	0,498	74	0,128
12	0,642	33	0,011	54	0,295	75	0,973
13	0,414	34	0,986	55	0,899	76	0,167
14	0,523	35	0,782	56	0,780	77	0,566
15	0,604	36	0,350	57	0,522	78	0,409
16	0,578	37	0,906	58	0,259	79	0,323
17	0,169	38	0,585	59	0,320	80	0,068
18	0,864	39	0,642	60	0,332	81	0,415
19	0,865	40	0,270	61	0,166	82	0,007
20	0,632	41	0,236	62	0,197	83	0,694
21	0,425	42	0,555	63	0,596	84	0,647

Y las muestras de calibración y predicción para las recogidas de este tipo de vehículo en este código postal se separan de manera análoga a los repartos, pudiendo ver en la tabla 10 cada una de las muestras. En este caso, debido a la aleatoriedad de la asignación, se tienen 45 observaciones en la muestra de calibración y 39 en la muestra de predicción.

Tabla 10: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08198

Día	RECOGIDAS. MUESTRA DE CALIBRACIÓN			Día	RECOGIDAS. MUESTRA DE PREDICCIÓN		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	16	1.921	79	1	18	668	62
2	19	1.991	91	2	17	2.725	99
3	15	1.749	77	3	14	2.552	89
4	17	2.175	87	4	18	2.045	89
5	11	1.635	63	5	19	2.239	98
6	20	2.413	100	6	17	2.637	96
7	18	2.115	90	7	15	2.305	84
8	17	1.532	79	8	18	2.442	96
9	17	2.080	87	9	19	2.282	100
10	17	2.078	90	10	19	1.979	94
11	19	1.238	72	11	15	2.368	89
12	25	1.631	99	12	18	2.137	94
13	17	1.032	69	13	14	1.596	69
14	19	1.996	95	14	16	2.334	92
15	17	2.425	92	15	17	1.695	81
16	23	2.210	106	16	13	1.456	65
17	17	1.677	77	17	12	1.739	68
18	17	1.004	70	18	17	1.592	82
19	7	436	29	19	15	1.658	77
20	15	3.222	103	20	22	2.064	104
21	14	1.535	65	21	21	2.269	98
22	22	1.606	94	22	22	1.364	84
23	23	1.569	100	23	16	1.327	71
24	16	2.191	88	24	22	3.270	121
25	20	2.665	106	25	17	3.087	104
26	16	1.019	71	26	18	2.256	99
27	19	2.588	103	27	15	2.077	81
28	19	1.712	85	28	14	1.015	62
29	10	1.314	55	29	16	2.382	90
30	13	1.267	64	30	19	2.065	93
31	16	1.249	70	31	16	1.137	64
32	12	1.423	61	32	21	2.096	97
33	19	1.914	88	33	17	2.129	90
34	15	1.447	69	34	21	2.157	100
35	16	2.295	92	35	17	2.757	98
36	18	3.370	114	36	6	69	18
37	25	1.627	101	37	18	1.657	85
38	19	2.720	105	38	14	1.963	80
39	17	3.543	114	39	20	2.133	98
40	17	1.971	81				
41	10	1.921	66				
42	19	1.741	92				
43	9	1.068	49				
44	17	1.819	83				
45	11	880	53				

2.1.1.4. Modelo de costes de reparto

Como ya se ha expuesto en la metodología, los costes de reparto y recogida son independientes, por lo que es posible el realizar dos modelos diferentes, uno para los costes de reparto y otro para los costes de recogida. En este apartado se analiza la elaboración del modelo de costes de reparto para el vehículo de 1500 kg. de peso máximo autorizado para este código postal.

Se va a realizar un análisis de regresión, en el cual se van a determinar los parámetros de regresión empleando el método de máxima verosimilitud. Como este método aplicado mediante un proceso iterativo ofrece la posibilidad de ajustar los modelos de regresión bajo diferentes premisas, como diferentes distribuciones de probabilidad de los errores y/o comportamientos no lineales de las variables exógenas, se van enumerar cuales se han aplicado antes de mostrar los resultados de cada una.

La primera distribución de probabilidad de los errores que se va a considerar es la distribución normal. Si con esta premisa el valor de máxima verosimilitud es mayor que con el resto, se pueden aplicar los principios del método de mínimos cuadrados. Tras calcular el valor máximo de verosimilitud con distribución de probabilidad normal, se van a aplicar ciertas funciones cadena para comprobar si modelos no lineales ofrecen un mejor ajuste a los datos. Estas funciones cadena serán la logarítmica, la recíproca y la raíz cuadrada. Para aplicar la cadena logarítmica, los valores de máxima verosimilitud se obtienen con el proceso iterativo de la misma manera que en el caso de la distribución normal, salvo que en este caso, el valor medio de la respuesta es la potencia del número e . Para el caso de la cadena recíproca, el valor medio en el proceso iterativo será el inverso del empleado en el caso de la distribución normal. Por último, en el caso de la cadena raíz cuadrada, el valor medio se eleva al cuadrado en el proceso iterativo para calcular los parámetros de regresión.

Después de considerar que la distribución de probabilidad de los errores es normal, a continuación se va a analizar los valores de máxima verosimilitud suponiendo que los errores se distribuyen con otras funciones de distribución de probabilidad, la distribución de probabilidad doble exponencial y la distribución de probabilidad logística. En el caso de la distribución doble exponencial, los cálculos son análogos al caso de la distribución normal, solo que hay que en lugar de la expresión de la distribución de probabilidad normal de los errores, hay que emplear la expresión de la distribución doble exponencial. Los cálculos para los modelos no lineales serán análogos al caso anterior.

Por último, para el caso de la distribución logística sólo se puede aplicar en el supuesto de que los errores siguen esta distribución, sin suposiciones de que las variables exógenas pueden tomar un carácter no lineal, ya que el método de los modelos lineales generalizados solo aplica las suposiciones de no linealidad de las variables exógenas a distribuciones de probabilidad de la familia exponencial.

Los valores de máxima verosimilitud para cada una de los supuestos anteriores se muestran a continuación, en las siguientes tablas.

Tabla 11: VMV modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
110,188					-193,3		0 %	30,77
9,884	3,001				-137,8	55,519	93,6 %	7,78
3,071	2,386	0,021			-104,4	33,423	98,8 %	3,42
-2,755	2,874	0,020	-0,007		-101,7	2,621	98,9 %	3,25
-2,734	2,206	0,039	0,001	-0,00001	-97,7	4,079	99,1 %	2,97

Tabla 12: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
4,702					-193,3		0 %	30,77
3,841	0,025				-153,5	39,846	86 %	11,51
3,709	0,020	0,0002			-138,8	14,663	93,1 %	8,08
3,077	0,060	0,0002	-0,001		-119,4	19,354	97,3 %	5,05
3,077	0,060	0,0002	-0,001	0	-119,4	0,000	97,2 %	5,12

Tabla 13: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,009					-193,3		0 %	30,77
0,0158	-0,0002				-162,9	30,454	77,6 %	14,56
0,016	-0,0002	0			-162,9	0,000	77 %	14,75
0,016	-0,0002		0		-162,7	0,115	77,1 %	14,71

Tabla 14: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-10,497					-193,3		0 %	30,77
-5,757	-0,139				-146,7	46,643	90 %	9,71
-5,196	-0,112	-0,001			-125,3	21,372	96,5 %	5,77
-3,244	-0,243	-0,001	0,002		-108,5	16,779	98,4 %	3,84
-3,244	-0,243	-0,001	0,002	0	-108,5	0,000	98,4 %	3,90

Tabla 15: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
111,718					-194,0		0 %	30,80
13,927	2,871				-140,3	53,749	93,4 %	7,89
2,655	2,246	0,025			-105,6	34,741	98,6 %	3,70
0,652	2,487	0,024	-0,004		-103,5	2,039	98,6 %	3,61
0,652	2,487	0,024	-0,004	0,00000	-103,5	0,000	98,6 %	3,66

Tabla 16: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
4,733					-194,0		0 %	30,96
3,886	0,024				-151,0	43,057	85,5 %	11,73
3,740	0,020	0,00021			-129,4	21,593	92,6 %	8,39
3,711	0,022	0,00021	-0,00003		-127,9	1,499	93,1 %	8,08
3,711	0,022	0,00021		0	-131,1	-3,157	92,3 %	8,54

Tabla 17: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-19,015					-255,8		-1808,90%	134,42
-19,015	0				-255,1	0,755	-1340,68%	116,77
-19,015	0	0			-255,1	0,000	-1340,68%	116,77

Tabla 18: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-10,677					-194,0		0 %	31,01
-6,043	-0,132				-146,1	47,930	89,7 %	31,41
-5,473	-0,106	-0,001			-119,6	26,500	96,2 %	31,83
-5,473	-0,106	-0,001	0		-119,6	0,000	96,1 %	32,27
-5,473	-0,106	-0,001		0	-119,6	0,000	96,1 %	32,27

Tabla 19: VMV modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
111,809					-192,5		0 %	30,81
10,089	2,999				-138,6	53,876	93,6 %	7,78
3,009	2,331	0,023			-105,0	33,608	98,7 %	3,46
-4,796	2,970	0,021	-0,009		-101,6	3,466	98,9 %	3,26
-3,062	2,321	0,037	-0,001	-0,00001	-97,8	4,248	99 %	3,02

En las tablas se puede observar como el valor de máxima verosimilitud es mayor para la distribución normal y para la distribución logística. El valor es casi idéntico para ambas, pero en el caso de la distribución normal el valor R² ajustado es un poco mayor que para la distribución logística. Por otro lado, la varianza es menor para la distribución normal, algo que es lógico ya que los estimados de los parámetros de regresión cuando los errores se distribuyen de manera normal son los de menor varianza, de acuerdo al teorema de Gauss-Markov. Por tanto, el modelo que se empleará para estimar los costes de reparto de este vehículo en este código postal será el elaborado considerando que los errores se distribuyen normalmente.

2.1.1.4.1. Selección de variables para el modelo de costes de reparto

Si se vuelve a observar la tabla 11, en la columna referente al cociente de verosimilitud se puede observar que la diferencia en el valor máximo de verosimilitud cada vez que se añade una nueva variable

al modelo es mayor que dos, por lo que la adición de cada una de estas variables se considera significativa, siendo el ajuste del modelo mejor con la adición de las variables que sin ella.

Las variables que se emplearan en este modelo son el número de repartos por día, el peso total repartido por día, el número de repartos por día elevado a la segunda potencia, y el peso total repartido por día elevado a la segunda potencia. Los parámetros de regresión correspondientes a cada variable son 2,206 para número de repartos, 0,039 para el peso total repartido, 0,001 para el número de repartos por día elevado a la segunda potencia, y -0,00001 para el peso total repartido por día elevado a la segunda potencia. El valor umbral para este modelo es -2,734.

2.1.1.4.2. Adecuación del modelo de costes de reparto

Como el mayor valor máximo de verosimilitud es el que aporta una distribución de probabilidad normal, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores.

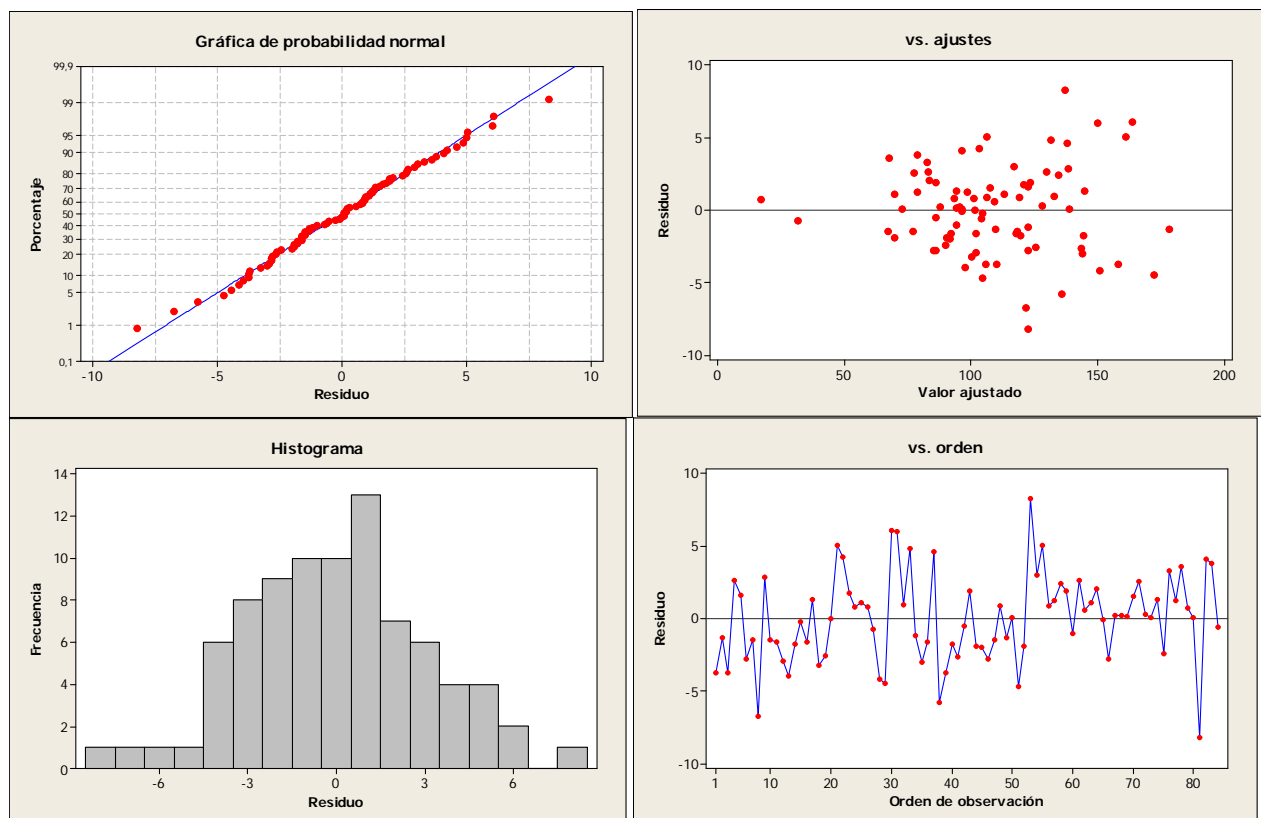


Figura 20: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08198

En la primera gráfica de la Figura 20 se puede observar como los errores se distribuyen de manera normal ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los

residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y , lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en esta cuarta gráfica se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.1.1.4.3. Validación del modelo de costes de reparto

Para comprobar que el modelo predice de manera adecuada los valores de los costes de reparto ante nuevos valores de las variables exógenas, se realiza la prueba t pareada o prueba de Welch. Estos nuevos valores de las variables exógenas son los valores de la muestra de predicción que se elaboró en el apartado 2.1.1.3., y la validación del modelo se lleva a cabo comparando uno a uno los valores de la variable endógena estimados por el modelo tras introducir los nuevos valores de las variables exógenas, con los valores correspondientes de los costes de reparto de la muestra de predicción. Los datos para esta prueba se obtienen de la tabla 20:

Tabla 20: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08198

β_0	β_1	β_2	β_3	β_4	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
-3,062	2,321	0,037	-0,001	-0,00001			
	34	1212	1156	1468944	110,77	106,27	4,50
	55	2076	3025	4309776	172,90	177,06	-4,16
	39	1070	1521	1144900	118,51	116,69	1,81
	40	1121	1600	1256641	121,98	115,09	6,89
	23	556	529	309136	68,41	65,76	2,65
	34	902	1156	813604	103,05	100,51	2,53
	33	975	1089	950625	102,73	98,81	3,92
	30	1080	900	1166400	98,65	93,49	5,16
	47	1365	2209	1863225	143,01	142,57	0,45
	32	1175	1024	1380625	105,43	104,20	1,22
	36	1348	1296	1817104	118,27	116,29	1,99
	45	1600	2025	2560000	143,32	146,13	-2,80
	29	1288	841	1658944	101,29	96,99	4,29
	34	875	1156	765625	102,32	101,39	0,93
	48	2057	2304	4231249	157,32	166,13	-8,81
	32	1120	1024	1254400	104,11	107,32	-3,21
	30	1227	900	1505529	102,15	101,92	0,24
	36	1145	1296	1311025	113,66	114,11	-0,46
	28	1105	784	1221025	94,77	94,28	0,49
	49	2070	2401	4284900	159,69	169,83	-10,14
	40	1602	1600	2566404	132,30	133,94	-1,65
	27	1138	729	1295044	93,32	90,38	2,94
	46	1203	2116	1447209	137,19	142,74	-5,54
	35	947	1225	896809	106,47	101,94	4,53
	44	1653	1936	2732409	142,09	140,92	1,17
	25	1088	625	1183744	87,59	85,63	1,97
	37	1490	1369	2220100	123,46	125,42	-1,97
	19	975	361	950625	71,16	67,73	3,42
	27	1122	729	1258884	92,93	89,59	3,35
	27	576	729	331776	78,05	75,45	2,59
	35	1103	1225	1216609	110,41	108,32	2,09
	20	1004	400	1008016	74,17	72,76	1,41
	31	1271	961	1615441	105,40	99,73	5,67
	40	1799	1600	3236401	135,76	145,42	-9,65
	37	1208	1369	1459264	117,37	119,95	-2,58
	28	820	784	672400	87,37	87,95	-0,58
	29	1051	841	1104601	95,69	93,42	2,27
	20	899	400	808201	71,43	71,04	0,39
	24	1067	576	1138489	84,82	85,40	-0,58
	29	1101	841	1212201	96,92	95,86	1,06
	31	868	961	753424	95,42	95,60	-0,18
	24	1333	576	1776889	91,01	87,39	3,62
	34	763	1156	582169	99,23	99,60	-0,37
	23	569	529	323761	68,80	71,21	-2,40

Y la tabla 21 muestra el resumen de esta validación.

Tabla 21: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08198

	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	107,74	107,32	0,420	Límite superior I.C	$t_{0,025;40}$
Varianza	624,74	744,28	14,55	1,547	1,96
Varianza/n			0,331		
Máximo	172,90	177,06		Límite inferior I.C	
Mínimo	68,41	65,76		-0,708	

En ella se puede observar como el modelo estima los costes de reparto de manera adecuada, ya que el intervalo de confianza creado incluye el cero. Para elaborarlo se ha empleado la distribución t de Student, con un nivel de significancia del 95% y 40 grados de libertad, cuyo valor es 1.96.

2.1.1.5. Modelo de costes de recogida

Una vez finalizado el modelo de los costes de reparto, hay que realizar el modelo de costes de recogida. La creación este modelo es análogo al de los costes de reparto, por lo que los pasos a seguir son idénticos.

Al igual que en el modelo anterior, en primer lugar se ajustan mediante máxima verosimilitud el modelo de regresión para diferentes distribuciones de probabilidad y distintas funciones cadenas, que examina comportamientos no lineales de las variables exógenas. Los valores de máxima verosimilitud son los siguientes.

Tabla 22: VMV modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
82,696					-194,5		0 %	18,44
19,434	3,771				-171,4	23,130	63,4 %	11,16
3,808	2,751	0,018			-104,4	66,924	98,1 %	2,55
3,921	2,734	0,018	0,001		-104,4	0,001	98 %	2,58
5,237	2,786	0,016		0,000001	-104,1	0,304	98 %	2,60

Tabla 23: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
4,415					-194,5		0 %	18,44
3,698	0,042				-174,5	20,017	58 %	11,95
3,394	0,036	0,0002			-118,5	56,011	96,4 %	3,48
3,081	0,076	0,0002	-0,001		-112,2	6,277	97,2 %	3,07
3,081	0,076	0,0002	-0,001	0	-112,2	0,000	97,2 %	3,11

Tabla 24: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-19294742					-263,6		-2057%	85,64
32,325	-2,020				-261,4	2,248	-1897%	82,41
32,375	-2,022	-0,000002			-261,3	0,046	-1941%	83,30
32,184	-2,035		0,001		-261,3	-0,029	-1944%	83,35

Tabla 25: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-9,094					-194,5		0 %	18,44
-5,688	-0,201				-172,9	21,580	60,8 %	11,55
-4,540	-0,160	-0,001			-109,1	63,780	97,6 %	2,83
-3,862	-0,250	-0,001	0,003		-106,5	2,612	97,8 %	2,70
-3,862	-0,250	-0,001	0,003	0	-106,5	0,000	97,8 %	2,74

Tabla 26: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
86,720					-197,5		0 %	18,88
15,009	3,958				-170,9	26,578	62,7 %	11,26
4,294	2,690	0,018			-103,0	67,882	98,1 %	2,56
4,294	2,690	0,018	-0,000001		-103,0	0,0002	98 %	2,59
4,294	2,690	0,018		0	-103,0	0,00003	98 %	2,63

Tabla 27: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
4,463					-197,5		0 %	18,88
3,676	0,043				-174,7	22,728	57,9 %	11,97
3,428	0,035	0,0002			-116,6	58,135	96,3 %	3,53
3,428	0,035	0,0002	0		-116,6	0,000	96,3 %	3,57
3,428	0,035	0,0002		0	-116,6	0,000	96,3 %	3,61

Tabla 28: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,012					-197,5		0 %	18,88
0,021	-0,0005				-177,9	19,518	51 %	12,81
0,021	-0,0005	0			-177,9	0,000	50 %	12,96
0,021	-0,0005		0		-177,9	0,000	50 %	12,96

Tabla 29: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-9,312					-197,5		0 %	18,88
-5,286	-0,218				-173,0	24,436	59,1 %	11,79
-4,759	-0,152	-0,001			-108,2	64,844	97,4 %	2,95
-4,759	-0,152	-0,001	0		-108,2	0,000	97,4 %	2,99
-4,759	-0,152	-0,001		0	-108,2	0,000	97,3 %	3,02

Tabla 30: VMV modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
83,605					-195,0		0 %	18,46
18,248	3,790				-170,6	24,428	63,2 %	11,19
4,060	2,726	0,018			-104,0	66,574	98,1 %	2,55
3,039	2,873	0,018	-0,004		-104,0	0,038	98 %	2,59
3,036	2,786	0,018		-0,0000001	-104,4	-0,402	98 %	2,62

En las tablas mostradas se puede observar como el mayor valor de máxima verosimilitud se obtiene para el modelo creado suponiendo que los errores se distribuyen de acuerdo a una distribución de probabilidad doble exponencial. Las diferencias, entre este supuesto y los supuestos de que los errores se distribuyen normalmente o con una distribución logística son muy pequeñas, pero se empleará el modelo ajustado con errores distribuidos según una distribución doble exponencial.

2.1.1.5.1 Selección de variables para el modelo de costes de recogida

Tal y como se puede observar en la tabla 26, al introducir en el modelo la variable número de recogidas elevado a la segunda potencia, se puede comprobar cómo la diferencia entre el valor de máxima verosimilitud entre este modelo con tres variables y el modelo el que incluye sólo las dos primeras es de 0,0002. Al ser menor de 2 esta diferencia, esta tercera variable no aporta un mejor ajuste a los datos en combinación con las dos anteriores, por lo que no se incluirá en el modelo.

Como esta variable no se incluye en el modelo, al calcular el valor de máxima verosimilitud incluyendo la cuarta variable, peso total recogido elevado a la segunda potencia, junto con las dos primeras, no se emplean los valores de la tercera variable exógena. En la tabla se puede comprobar cómo el valor de máxima verosimilitud con las variables número de recogidas, peso total recogido por día y peso total recogido elevado a la segunda potencia, se diferencia solamente en 0,00003 unidades respecto al de únicamente dos variables. Como esta diferencia es también menor de 2 unidades, esta cuarta variable tampoco se incluirá en el modelo.

Por tanto, para este modelo de costes de recogida solo incluye dos variables, número de recogidas y peso total recogido por día, con parámetros de regresión iguales a 2,69 y 0,018 respectivamente.

2.1.1.5.2 Adecuación del modelo de costes de recogida

Cuando los errores se distribuyen con una distribución de probabilidad no normal, si éstos son independientes y se distribuyen idénticamente, las inferencias bajo la teoría de normalidad son asintóticamente correctas. Por tanto, los pasos para comprobar la adecuación del modelo serán los del

método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores.

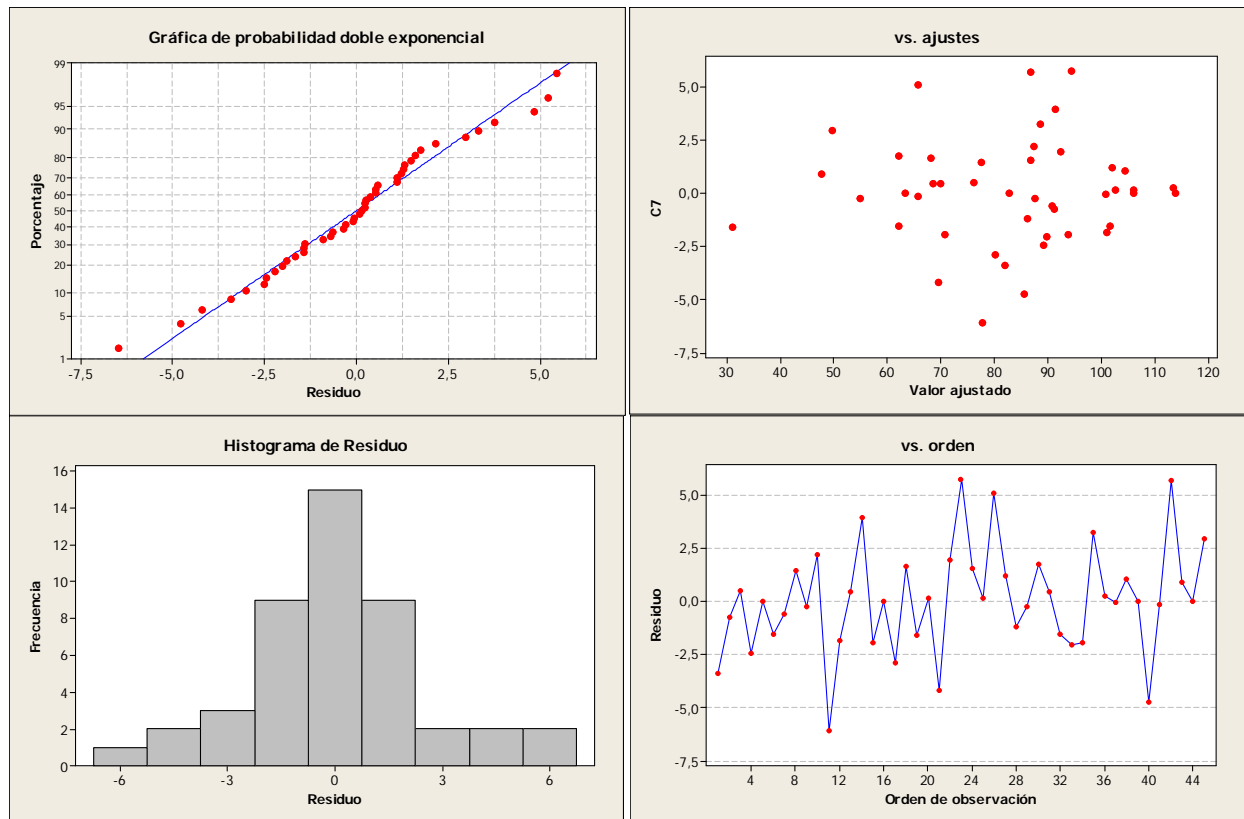


Figura 21: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08198

En la primera gráfica se puede observar como los errores se distribuyen según una distribución doble exponencial ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.1.1.5.3 Validación del modelo de costes de recogida

La validación del modelo de costes de recogida se realizará empleando la muestra de predicción. Se comparan los costes de recogida de esta muestra, con los resultados que el modelo ofrece tras introducirle los valores de variables exógenas de la misma muestra. Los valores con los que se lleva a cabo es validación se pueden observar en la tabla 31.

Tabla 31: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
4,294	2,690	0,018			
	18	668	64,75	61,71	3,04
	17	2725	99,09	98,51	0,58
	14	2552	87,91	88,77	-0,86
	18	2045	89,54	88,56	0,98
	19	2239	95,72	98,09	-2,36
	17	2637	97,51	96,39	1,12
	15	2305	86,15	83,95	2,20
	18	2442	96,69	95,58	1,11
	19	2282	96,50	100,35	-3,86
	19	1979	91,04	93,95	-2,91
	15	2368	87,29	89,46	-2,18
	18	2137	91,20	94,46	-3,26
	14	1596	70,69	69,20	1,50
	16	2334	89,36	91,63	-2,27
	17	1695	80,55	81,28	-0,73
	13	1456	65,48	64,60	0,89
	12	1739	67,89	68,19	-0,30
	17	1592	78,69	82,00	-3,30
	15	1658	74,50	76,72	-2,22
	22	2064	100,64	103,78	-3,14
	21	2269	101,64	98,02	3,62
	22	1364	88,04	83,92	4,12
	16	1327	71,23	70,68	0,55
	22	3270	122,36	120,65	1,71
	17	3087	105,61	104,36	1,25
	18	2256	93,34	98,94	-5,60
	15	2077	82,05	80,58	1,46
	14	1015	60,23	61,86	-1,63
	16	2382	90,23	90,47	-0,25
	19	2065	92,59	92,55	0,05
	16	1137	67,81	64,06	3,75
	21	2096	98,53	97,36	1,17
	17	2129	88,36	89,54	-1,18
	21	2157	99,63	100,31	-0,68
	17	2757	99,67	98,48	1,19
	6	69	21,68	17,73	3,95
	18	1657	82,55	85,33	-2,78
	14	1963	77,30	79,95	-2,65
	20	2133	96,50	97,69	-1,18

La tabla en la que se resume esta validación es la tabla 32.

Tabla 32: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08198

	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	85,91	86,15	-0,234	Límite superior I.C	$t_{0,025;37}$ 1,96
Varianza	282,59	304,60	5,70		
Varianza/n			0,146		
Máximo	122,36	120,65		Límite inferior I.C	-0,983
Mínimo	21,68	17,73			

El modelo de costes de recogida de este código postal se considera válido ya que el intervalo obtenido en la prueba de Welch contiene el cero.

2.1.1.6 Modelo de carga de trabajo

En este modelo se va a identificar que distribución de probabilidad siguen las variables aleatorias con las que se puede modelar la carga la de trabajo, así como los valores de los parámetros que las caracterizan. Estas variables son el número de repartos por día y el peso de cada uno de los repartos, en el caso de los repartos, y el número de recogidas por día y el peso de cada recogida, en el de las recogidas.

Las muestras que se emplean para las variables número de repartos y número de recogidas son las mismas muestras que las muestras de calibración que se emplearon para ajustar los modelos de los costes de reparto y recogida. Las muestras para las variables peso de cada reparto y peso de cada recogida, es una muestra con los pesos unitarios de cada uno de los repartos y de cada una de las recogidas.

2.1.1.6.1. Número de repartos por día

Minitab busca los parámetros para cada uno de los distintos tipos de distribuciones de probabilidad, siendo estos parámetros tales que la distribución de probabilidad que caracterizan, es la que más se aproxima a los datos de la muestra. Una vez que se tienen estos datos hay que comprobar que la distribución de probabilidad que conforma se ajuste suficientemente bien a los datos. Para realizar esta comprobación se realizan los denominados test no paramétricos de bondad de ajuste. Debido a la naturaleza de esta variable aleatoria, la cual es un conteo de algún elemento, la primera función de distribución de probabilidad que se va a tratar de comprobar el ajuste es la distribución de Poisson. Como esta distribución es discreta se emplea el test no paramétrico de bondad de ajuste de Chi cuadrado. El resultado de este test se puede ver en la tabla 33.

Tabla 33: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08198

Observaciones	O _i	f(O _i)	F(e _i)	E _i	Chi-cuadrado	
7-25	5	0,125	0,0806	3	0,9768	alfa=0,05 k=5 Est. par. =1 n=5-1-1=3
26-29	9	0,225	0,1978	8	0,1502	
30-34	8	0,200	0,3311	13	2,0755	
35-42	14	0,350	0,3528	14	0,0009	
43-55	4	0,100	0,0626	3	0,8944	
	40				4,10	7,85

En la tabla se puede observar como el valor límite para aceptar la hipótesis de que la una distribución de Poisson con lambda 33,43 ajusta adecuadamente la variable número de repartos por día es 7,85. Este valor corresponde a una distribución Chi cuadrado con 3 grados de libertad y nivel de significancia del 5%. Los grados de libertad es igual al número de intervalos de los que consta el test, al que hay que restar el número de parámetros estimados (para el caso de la distribución de Poisson es únicamente el parámetro lambda) y otra unidad. Como en este caso el valor del test es 4,10, menor al

valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.1.1.6.2. Peso de cada reparto

La segunda variable a analizar para el modelo de carga de trabajo. Para esta variable continua, las distribuciones de probabilidad que se comprueban si ajustan adecuadamente a los datos de la muestra son distribuciones continuas. Y el test de bondad de ajuste a realizar es el test Kolmogorov-Smirnov. Los datos dados por Minitab de los parámetros de las distribuciones de probabilidad y los resultados de los test de Kolmogorov-Smirnov se pueden ver en la tabla 34.

Tabla 34: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08198

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,247	alfa	0,092	alfa	0,075	Media	0,045	Lambda	0,253
36,51	7,323	0,6645	2,718	0,7469	2,212	2,681	1,319	25,93	7,486
Desv. Est.	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	Desv. Est.	$c_{\alpha} = 1,358$	Theta	$c_{\alpha} = 1,358$
52,26	$\alpha = 0,05$	54,94	$\alpha = 0,05$	29,99	$\alpha = 0,05$	1,454	$\alpha = 0,05$	23,13	$\alpha = 0,05$

En la tabla 33 se puede ver como únicamente la distribución lognormal supera satisfactoriamente el test Kolmogorov-Smirnov, ya que el valor obtenido para esta distribución es 1,319 cuando el valor límite para superar el test es de 1,358.

2.1.1.6.3. Número de recogidas por día

Esta variable al igual que el número de repartos por día, es un conteo de algún elemento, y por tanto, es una distribución discreta. Por ser una distribución discreta, la primera distribución que se va a comprobar si ajusta adecuadamente a los datos es la distribución de Poisson. El resultado de la prueba de bondad de ajuste Chi cuadrado se puede observar en la tabla 35.

Tabla 35: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08198

Observaciones	O _i	f(O _i)	F(e _i)	E _i	Chi-cuadrado	
7-13	8	0,178	0,2157	10	0,3005	alfa=0,05
14-17	20	0,444	0,3695	17	0,6849	k=4
18-19	10	0,222	0,1688	8	0,7612	Est. par. =1
20-25	7	0,156	0,2240	10	0,9409	n=4-1-1=2
	45				2,69	5,991

El valor límite para aceptar la hipótesis de que la una distribución de Poisson con media 16,78 ajusta adecuadamente la variable número de recogidas por día. Este valor corresponde a una distribución Chi cuadrado con 2 grados de libertad y nivel de significancia del 5%. Como en este caso el valor del test es 2.69, menor al valor límite 5,991, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de recogidas por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.1.1.6.4. Peso de cada recogida

Los resultados de los test Kolmogorov-Smirnov para las distribuciones con las que se pretenden ajustar a los datos esta variable aleatoria continua, y los valores dados por Minitab para cada uno de los parámetros de estas distribuciones son los de la tabla 36.

Tabla 36: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08198

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,237	alfa	0,089	alfa	0,060	Media	0,044	Lambda	0,252
114,2	4,718	0,5801	1,764	0,6882	1,198	3,667	0,868	81,58	5,019
Desv. Est.	c _α = 1,358	beta	c _α = 1,358	beta	c _α = 1,358	Desv. Est.	c _α = 1,358	Theta	c _α = 1,358
159,7	α = 0,05	196,8	α = 0,05	87,91	α = 0,05	1,673	α = 0,05	74,92	α = 0,05

Las dos únicas distribuciones que han superado satisfactoriamente el test Kolmogorov-Smirnov para esta variable han sido la distribución Weibull y la distribución lognormal, siendo la que menor discrepancia presenta respecto a los valores de la muestra es la distribución Lognormal. Ambas distribuciones han obtenido un valor menor de 1,358 del valor límite de aceptación del test.

2.1.1.7 Simulación de los costes de reparto y recogida, y de la carga de trabajo

En este apartado se llevará a cabo la simulación de los modelos creados para el estudio del vehículo de 1500 kg. de peso máximo autorizado en el código postal 08198. En primer lugar se simulará la carga de trabajo. Para ello hay que determinar el número de repartos y de recogidas para cada uno de los días simulados, y el peso total repartido y el peso total recogido por día. Empleando los ajustes de las distribuciones de probabilidad de Poisson obtenidos en los apartados 2.1.1.6.1 y 2.1.1.6.3. se simularán el número de repartos y el número de recogidas para cada uno días que se desean simular.

Posteriormente, se simulará el peso correspondiente de cada uno de los repartos y las recogidas mediante las distribuciones de probabilidad calculadas en los apartados 2.1.1.6.2. y 2.1.1.6.4. Sumando el peso unitario de todos los repartos simulados en cada uno de los días, se calculará el peso total repartido por día. Con el peso unitario de las recogidas simulado para cada uno de los días, se calculará el peso total recogido por día.

Finalmente, se simularán los modelos de costes de reparto y costes de recogida. Como valores de entrada a estos modelos se emplearán los datos simulados para el modelo de carga de trabajo. El número de repartos y el peso total repartido por día serán los datos de entrada para el modelo de costes de reparto. Análogamente, el número de recogidas y el peso total recogido por día serán los valores de entrada para el modelo de costes de recogida.

2.1.1.7.1 Simulación de los costes de reparto

La carga de trabajo relativa a los repartos para este vehículo del código postal 08198 es la siguiente: el número de repartos puede ser descrito mediante una distribución de probabilidad Poisson con parámetro lambda igual a 33,43. Mientras que el peso de cada reparto sigue una distribución de probabilidad lognormal con media 2,681 y desviación estándar 1,454.

Un resumen de la carga de trabajo simulada se puede ver en la tabla 37, en el cual se puede observar el promedio de repartos por día simulados, su desviación estándar, el número máximo de repartos a realizar en un día, así como el número mínimo. Y los mismos datos para la variable peso total repartido.

Tabla 37: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08198

	Repartos por día	Peso total repartido por día
Promedio	33	1078
Desv. Estan.	6	331
Máximo	48	2017
Mínimo	19	417

La simulación de la carga de trabajo en 250 días bajo las condiciones previamente enunciadas, son las entradas para la simulación en el modelo de costes de reparto. Para cada uno de los días se obtiene un estimado puntual del coste de reparto para la carga de trabajo simulada correspondiente. A partir de cada estimado puntual se desarrolla un intervalo de predicción, teniendo un total de 250 intervalos de predicción. Estos intervalos de predicción se calculan tal y como se explicó en el apartado 4.1.1.3 del capítulo de metodología. Los resultados que aquí se exponen son el valor medio de límite inferior de las 250 simulaciones, la varianza y desviación estándar de este límite inferior, así como un intervalo de confianza de este límite inferior. Para el límite superior del intervalo de predicción los resultados que se presentan son análogos. Estos resultados se pueden ver en la tabla 38:

Tabla 38: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08198

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	98,55	111,82
Varianza	327	330
Desv. Estan.	18,07	18,17
I.C	96,66	109,93
	100,43	113,72

En la tabla se puede observar como el límite inferior de los estimados puntuales de los costes de repartos se distribuyen con un valor medio de 98,55 euros por día y una desviación estándar de 18,07. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 111,82 y una desviación estándar de 18,17 euros por día.

2.1.1.7.2 Simulación de los costes de recogida

La carga de trabajo relativa a las recogidas para este vehículo del código postal 08198 es la siguiente: el número de repartos puede ser descrito mediante una distribución de probabilidad Poisson con parámetro lambda igual a 16,78. Mientras que el peso de cada reparto sigue una distribución de probabilidad lognormal con media 3,667y desviación estándar 1,673.

Un resumen de la carga de trabajo simulada se puede ver en la tabla 39, en el cual se puede observar el promedio de recogidas por día simulado, su desviación estándar, el número máximo de

recogidas a realizar en un día, así como el número mínimo. Y los mismos datos para la variable peso total recogido por día.

Tabla 39: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08198

	Recogidas por día	Peso total recogido por día
Promedio	16	1439
Desv. Estan.	4	652
Máximo	29	3816
Mínimo	7	219

Para la simulación de los costes de recogida, los resultados que se exponen son idénticos al caso de los costes de reparto. Se exponen los valores medios, varianza y desviación estándar, así como un intervalo de confianza de los límites superior e inferior de los límites de predicción de los 250 días simulados a partir de la carga de trabajo previamente simulada bajo las condiciones expuestas en el primer párrafo de este subapartado. Los resultados son los de la tabla 40.

Tabla 40: Resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08198

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	69,28	79,75
Varianza	426	422
Desv. Estan.	20,64	20,55
I.C	67,12	77,60
	71,43	81,89

En la tabla se puede observar como el límite inferior de los estimados puntuales de los costes de repartos se distribuyen con un valor medio de 69,28 euros por día y una desviación estándar de 20,64. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 79,75 y una desviación estándar de 20,55 euros por día.

2.1.2 *Vehículo de 3500 kg. de peso máximo autorizado para el código postal 08198*

Como en este código postal se dan repartos de más de 300 kg. de peso y recogidas de más de 700 kg. existen tanto repartos como recogidas para un vehículo de 3500 kg. de peso máximo autorizado. Hay que analizar por tanto el análisis de los costes de reparto y recogida, así como las carga de trabajo para un vehículo de este peso máximo autorizado.

2.1.2.1 Toma de muestras

Al igual que para el caso del vehículo de 1500 kg. el primer paso para la realización de los modelos, una vez identificadas la variables, es la toma de muestras de cada una de las variables. Las variables serán las mismas que para el caso anterior, y para el resto de modelos que se realicen, y son los costes de reparto y los costes de recogida. Las variables exógenas que determinan estas variables endógenas son el número de repartos y el peso total repartido, para el modelo de los costes de reparto, y el número de recogidas y el peso total recogido para el caso de la variable costes de recogida. Las 84 observaciones recopiladas para el modelo de costes de reparto como para el modelo costes de recogida se pueden observar en las tablas 41 y 42 respectivamente.

Tabla 41: Toma de muestras de los repartos del vehículo de 3500 kg. PMA para el código postal 08198

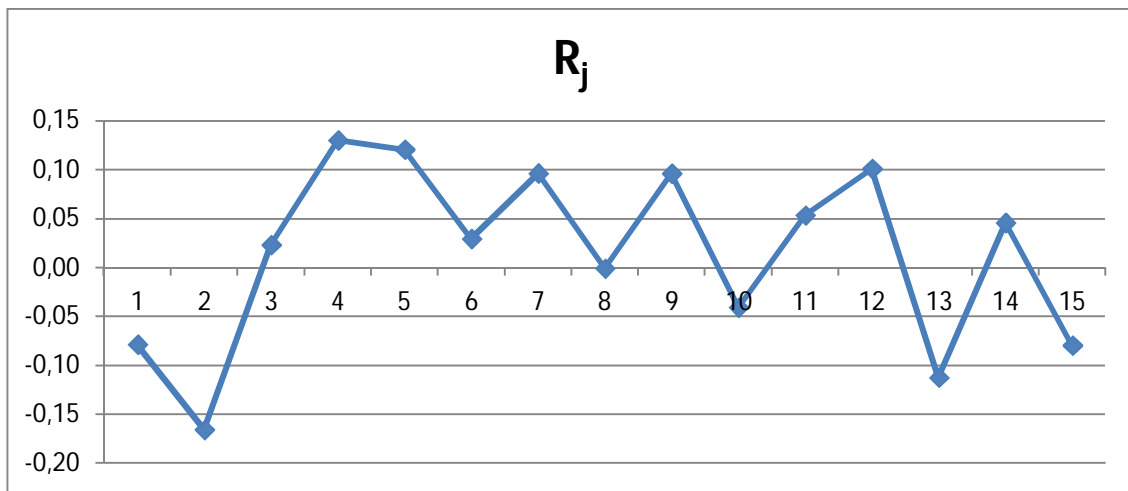
Día	REPARTOS			Día	REPARTOS		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	0	0	0	43	4	1.644	41
2	3	1.500	34	44	1	440	11
3	1	392	10	45	1	320	9
4	1	379	10	46	3	2.410	47
5	3	1.442	33	47	3	1.991	41
6	1	610	13	48	0	0	0
7	1	1.024	19	49	2	5.344	80
8	0	0	0	50	1	379	10
9	1	586	13	51	0	0	0
10	0	0	0	52	1	591	13
11	3	2.603	49	53	1	326	9
12	1	330	9	54	3	1.593	35
13	0	0	0	55	3	1.510	34
14	3	1.820	39	56	2	749	19
15	1	314	9	57	1	473	11
16	0	0	0	58	1	2.184	34
17	0	0	0	59	3	1.515	34
18	0	0	0	60	3	2.287	45
19	4	3.262	63	61	4	1.518	39
20	0	0	0	62	2	1.419	29
21	2	739	19	63	3	3.107	56
22	1	643	13	64	2	961	22
23	3	2.366	46	65	0	0	0
24	1	800	16	66	3	4.911	81
25	1	608	13	67	8	4.543	98
26	2	1.557	30	68	0	0	0
27	0	0	0	69	2	1.093	24
28	0	0	0	70	2	888	21
29	1	1.691	28	71	2	1.494	30
30	2	4.009	63	72	2	1.906	35
31	2	718	19	73	1	724	14
32	0	0	0	74	2	1.104	24
33	0	0	0	75	0	0	0
34	1	3.161	46	76	5	2.916	63
35	4	2.434	51	77	1	1.865	30
36	0	0	0	78	1	396	10
37	2	913	22	79	0	0	0
38	1	400	10	80	2	1.144	25
39	4	2.342	50	81	1	2.008	32
40	1	1.195	21	82	0	0	0
41	4	1.918	44	83	5	4.067	78
42	2	596	17	84	3	1.422	33

Tabla 42: Toma de muestras de las recogidas del vehículo de 3500 kg. PMA para el código postal 08198

Día	RECOGIDAS			Día	RECOGIDAS		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	3	5.000	77	43	1	1.391	23
2	3	3.591	62	44	1	1.423	24
3	1	3.049	39	45	2	3.485	54
4	1	3.865	43	46	4	8.365	116
5	2	4.990	69	47	3	3.978	66
6	1	3.979	44	48	2	4.166	62
7	1	3.590	42	49	1	3.103	40
8	1	3.316	41	50	4	5.080	82
9	2	3.942	61	51	1	896	17
10	3	7.622	103	52	3	5.525	81
11	1	2.584	36	53	0	0	0
12	1	4.560	47	54	3	8.180	106
13	2	4.549	62	55	1	1.101	19
14	2	3.362	53	56	1	3.297	41
15	3	4.441	73	57	0	0	0
16	1	3.174	40	58	2	2.621	44
17	1	3.133	40	59	3	6.895	94
18	3	7.737	101	60	2	4.323	57
19	3	5.690	85	61	3	5.980	89
20	3	3.989	67	62	3	5.294	81
21	2	4.758	69	63	2	5.199	72
22	3	6.210	91	64	2	5.530	73
23	2	4.543	62	65	3	4.955	77
24	4	7.768	113	66	3	3.919	66
25	3	4.805	76	67	3	5.475	84
26	1	3.049	39	68	2	4.287	63
27	2	6.136	78	69	2	3.681	54
28	5	9.690	143	70	3	4.680	72
29	4	8.857	125	71	2	4.733	63
30	1	1.958	31	72	4	5.216	88
31	2	3.366	52	73	1	3.222	40
32	4	9.275	133	74	1	1.767	28
33	0	0	0	75	4	7.798	114
34	1	2.737	37	76	4	8.648	118
35	1	1.453	24	77	3	6.903	96
36	2	3.000	48	78	1	1.660	27
37	2	4.129	58	79	0	0	0
38	2	4.374	59	80	2	1.993	36
39	3	6.200	93	81	3	6.325	94
40	2	2.974	48	82	3	4.536	72
41	3	6.074	87	83	2	3.383	51
42	0	0	0	84	4	6.205	94

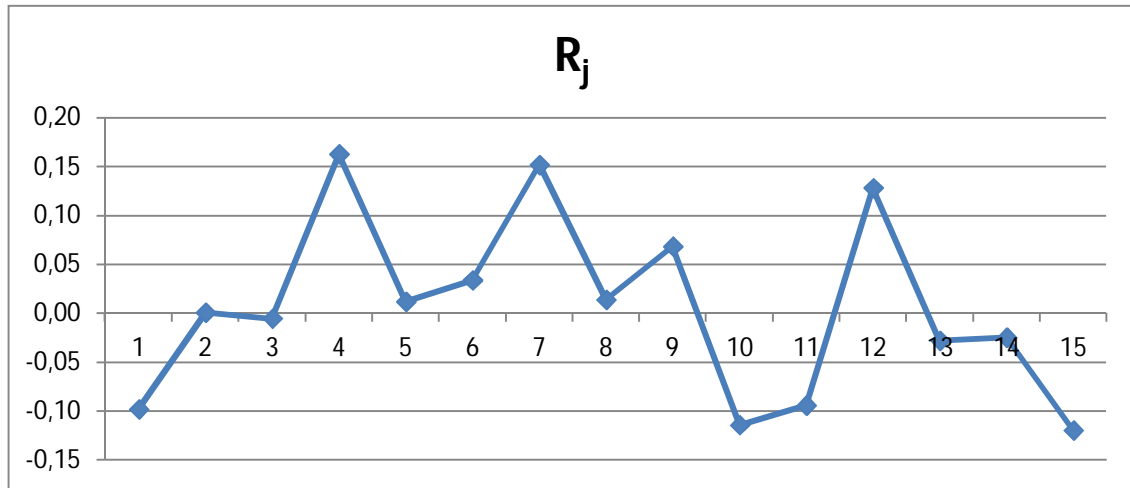
2.1.2.2 Independencia de la muestras

Tras la obtención de las muestras de las variables mediante el estudio observacional diseñado, el siguiente paso es comprobar la independencia de cada una de las observaciones de las muestras para poder aplicar con garantías el método de máxima verosimilitud y los test no paramétricos de bondad de ajuste. En primer lugar se comprobará la independencia de las observaciones de la muestra correspondiente a los repartos. Para cada una de las variables, una endógena y dos exógenas, se realizan las gráficas de autocorrelación a cada una. La Figura 22 corresponde a la gráfica de autocorrelación de la muestra de los costes de reparto por día. La Figura 23 corresponde a la gráfica de autocorrelación del número de repartos por día, y la Figura 24 corresponde a la gráfica de autocorrelación del peso total repartido por día



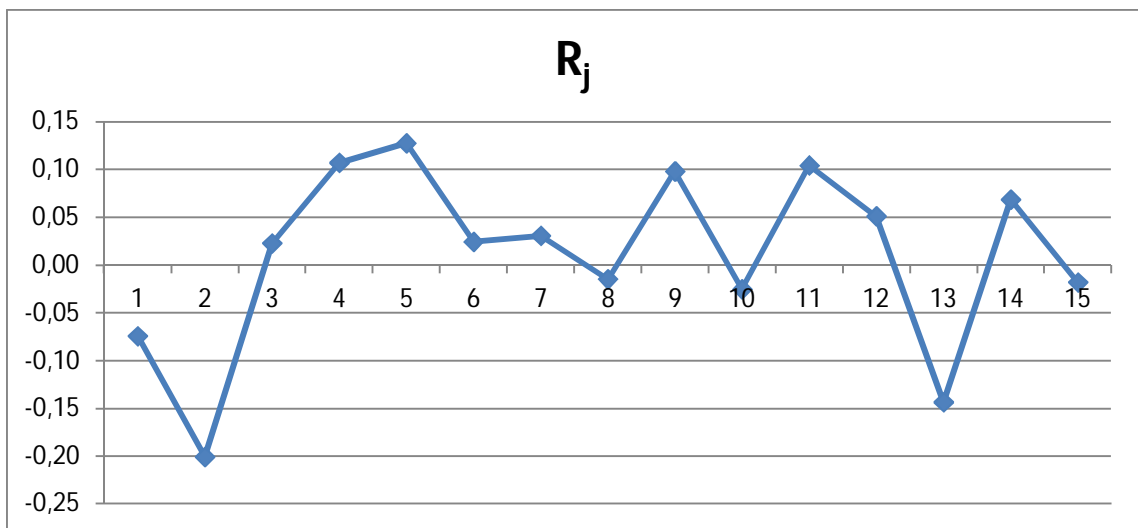
R_j	-0,08	-0,17	0,02	0,13	0,12	0,03	0,10	0,00	0,10	-0,04	0,05	0,10	-0,11	0,05	-0,08
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-39	-83	12	65	60	15	48	0	48	-20	27	51	-56	23	-40

Figura 22: Grafica de autocorrelación de coste de reparto por día vehículo 3500 kg. PMA código 08198



R_j	-0,10	0,00	-0,01	0,16	0,01	0,03	0,15	0,01	0,07	-0,11	-0,09	0,13	-0,03	-0,02	-0,12
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-0,2	0,0	0,0	0,4	0,0	0,1	0,3	0,0	0,2	-0,3	-0,2	0,3	-0,1	-0,1	-0,3

Figura 23: Grafica de autocorrelación de número de repartos por día vehículo 3500 kg. PMA código 08198



R_j	-0,07	-0,20	0,02	0,11	0,13	0,02	0,03	-0,01	0,10	-0,03	0,10	0,05	-0,14	0,07	-0,02
J	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-113589	-306799	34849	163689	195058	37164	46893	-22461	149805	-38607	159349	78051	-219226	104990	-27575

Figura 24: Grafica de autocorrelación del peso total repartido por día vehículo 3500 kg. PMA código 08198

En las tres gráficas se puede observar como en ninguna de ellas los valores de R_j se encuentran en el rango de -1 y 1, y que en ninguna de ellas se puede apreciar un patrón determinado. Por tanto, esta prueba indica la independencia de las observaciones de las muestras.

A continuación, una vez expuestas las gráficas de autocorrelación, se muestra las gráficas de dispersión para las mismas variables, como segunda comprobación de la independencia de las observaciones. Al igual que para las gráficas de autocorrelación, se realizará una por cada una de las variables.

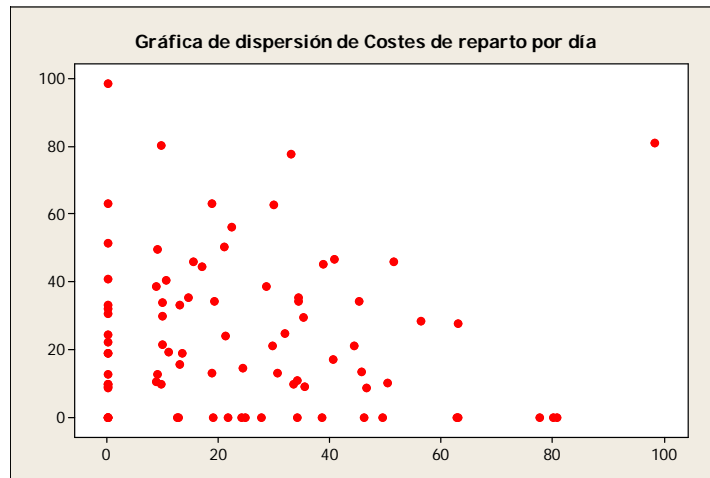


Figura 25: Grafica de dispersión de coste de reparto por día vehículo 3500 kg. PMA código 08198

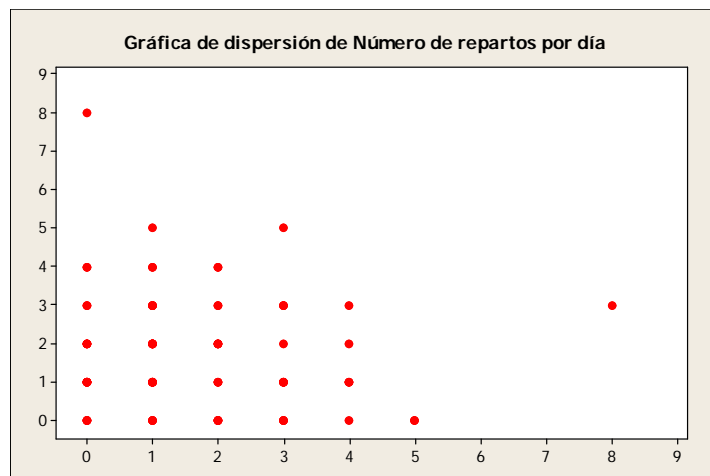


Figura 26: Grafica de dispersión de número de repartos por día vehículo 3500 kg. PMA código 08198

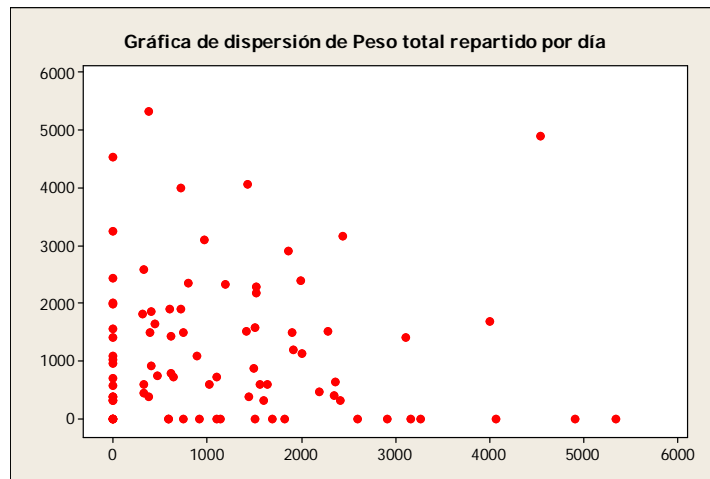
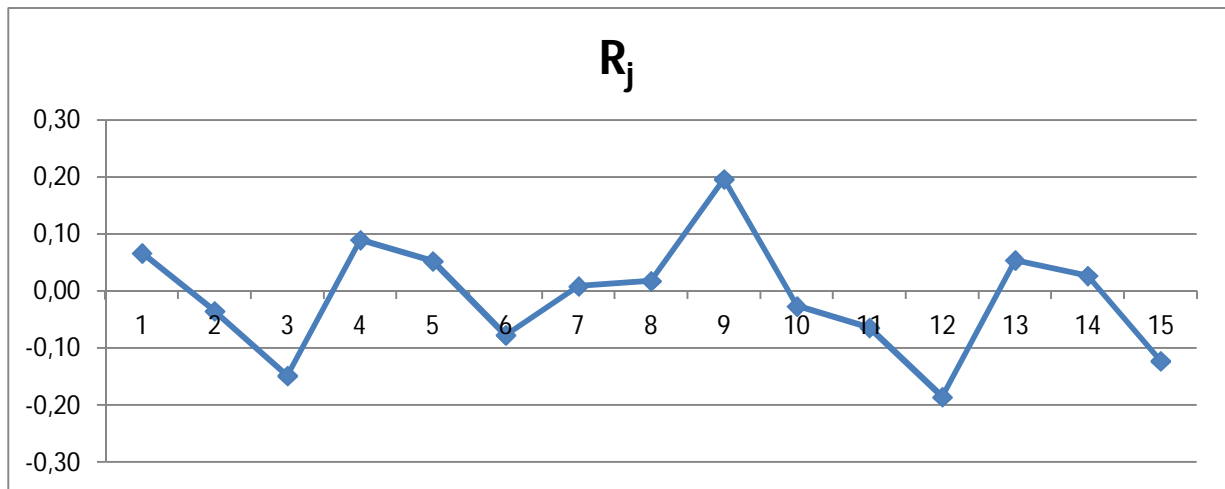


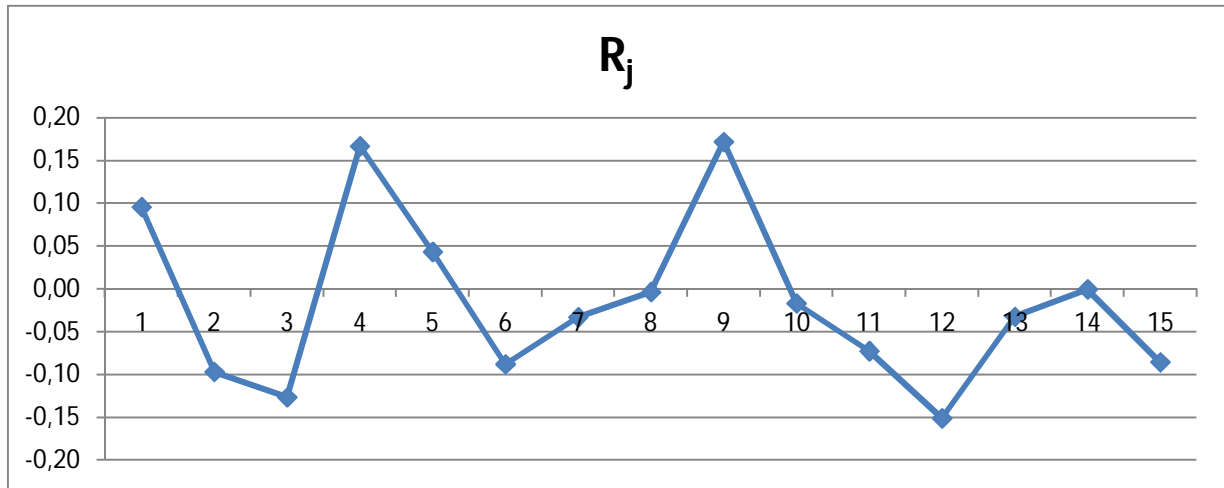
Figura 27: Grafica de dispersión del peso total repartido por día vehículo 3500 kg. PMA código 08198

En estas gráficas se puede comprobar cómo los puntos están distribuidos aleatoriamente por los planos, lo cual indica que no existe correlación entre las observaciones y éstas son independientes. Una vez comprobado la independencia de la muestra correspondiente a los repartos, se comprobará la independencia de la muestra de las recogidas. De la misma manera que para los repartos, en primer lugar se realizarán las gráficas de autocorrelación.



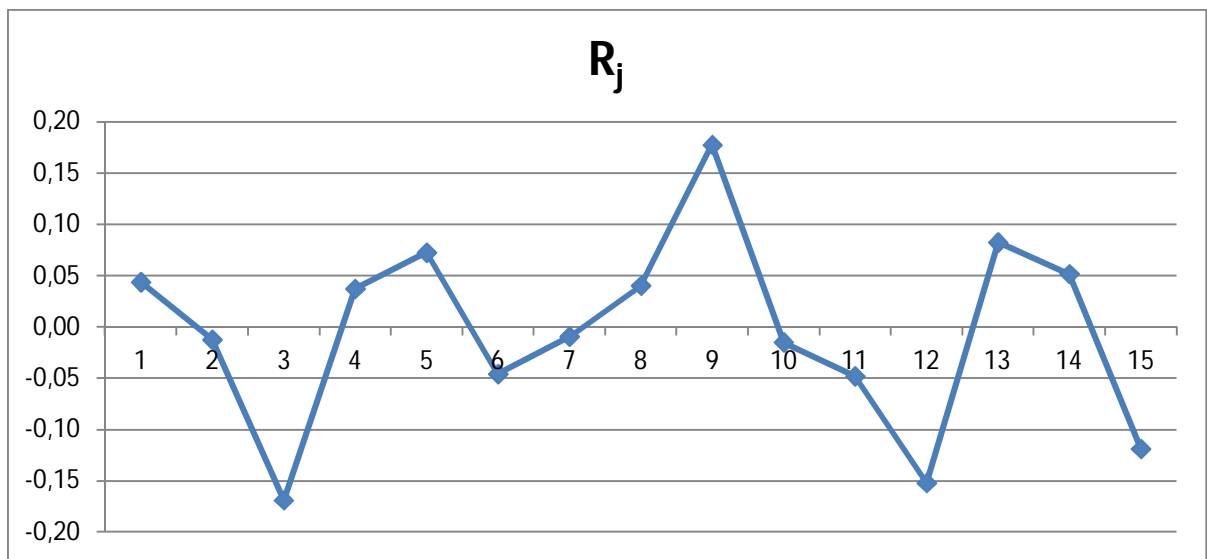
R_j	0,07	-0,03	-0,15	0,09	0,05	-0,08	0,01	0,02	0,20	-0,03	-0,06	-0,19	0,05	0,03	-0,12
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	67	-35	-149	90	53	-77	9	18	197	-26	-64	-186	55	27	-123

Figura 28: Grafica de autocorrelación de coste de recogida por día vehículo 3500 kg. PMA código 08198



R_j	0,10	-0,10	-0,13	0,17	0,04	-0,09	-0,03	0,00	0,17	-0,02	-0,07	-0,15	-0,03	0,00	-0,09
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	0,1	-0,1	-0,2	0,2	0,1	-0,1	0,0	0,0	0,2	0,0	-0,1	-0,2	0,0	0,0	-0,1

Figura 29: Grafica de autocorrelación del peso total recogido por día vehículo 3500 kg. PMA código 08198



R_j	0,04	-0,01	-0,17	0,04	0,07	-0,05	-0,01	0,04	0,18	-0,02	-0,05	-0,15	0,08	0,05	-0,12
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	216815	-62417	-840219	184628	360317	-228301	-47302	199938	881998	-75273	-238796	-756472	410021	256762	-591239

Figura 30: Grafica de autocorrelación del peso total recogido por día vehículo 3500 kg. PMA código 08198

En este caso también las gráficas muestran como no existe ningún patrón determinado y los valores R_j no superan en ningún caso los valores -1 y 1, por lo que esta prueba determina que las observaciones de muestra correspondiente a las recogidas son independientes. Las gráficas de dispersión para estas observaciones se muestran a continuación, en las gráficas 31,32 y 33.

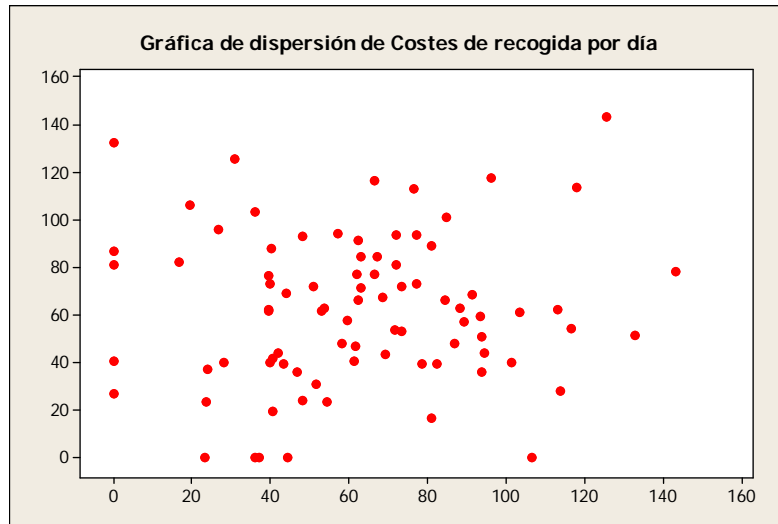


Figura 31: Grafica de dispersión de coste de recogida por día vehículo 3500 kg. PMA código 08198

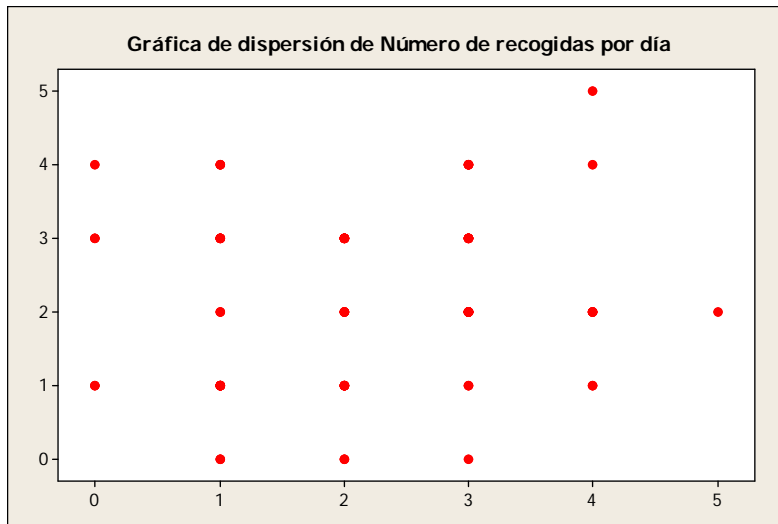


Figura 32: Grafica de dispersión de número de recogidas por día vehículo 3500 kg. PMA código 08198

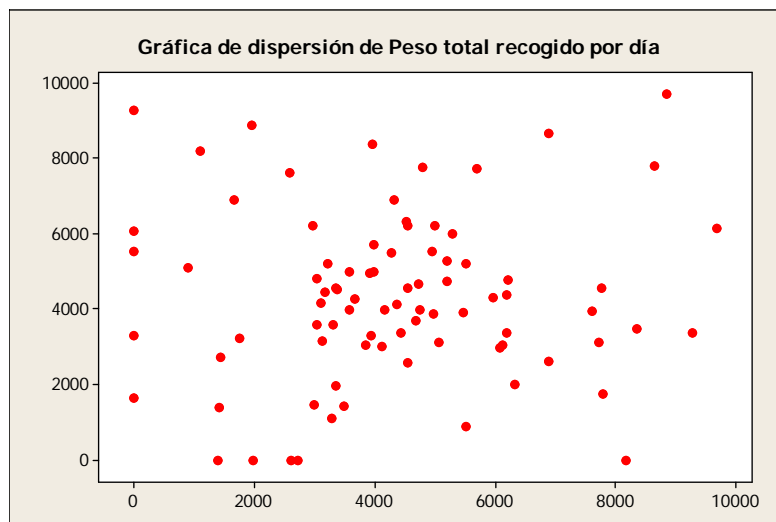


Figura 33: Grafica de dispersión de peso total recogido por día vehículo 3500 kg. PMA código 08198

Como los puntos de todos los gráficos se distribuyen de manera aleatoria a lo largo de cada plano correspondiente, se concluye que las observaciones de la muestra son independientes.

2.1.2.3 Separación de las muestras de calibración y predicción

Para la elaboración de los modelos de costes de reparto y costes de recogida es necesario el separar las 84 observaciones en la muestra de calibración y la muestra de predicción. La primera se emplea para la creación del modelo y la segunda para la validación del mismo. Esta separación se realiza mediante asignación aleatoria. Se asigna un número aleatorio proveniente de una distribución uniforme a cada una de las observaciones, y si éste es mayor o igual que 0,5, esta observación se incluirá en la muestra de calibración. Por el contrario, si es menor de 0,5, se incluye en la muestra de predicción.

La asignación de números aleatorios a cada una de las observaciones de reparto se muestra en la tabla 43.

Tabla 43: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 3500 kg. PMA código postal 08198

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,729	22	0,171	43	0,551	64	0,841
2	0,280	23	0,129	44	0,263	65	0,644
3	0,495	24	0,274	45	0,847	66	0,751
4	0,342	25	0,573	46	0,973	67	1,000
5	0,060	26	0,666	47	0,630	68	0,400
6	0,608	27	0,213	48	0,027	69	0,395
7	0,331	28	0,509	49	0,395	70	0,176
8	0,894	29	0,460	50	0,471	71	0,621
9	0,490	30	0,261	51	0,112	72	0,199
10	0,173	31	0,702	52	0,647	73	0,469
11	0,075	32	0,669	53	0,369	74	0,470
12	0,436	33	0,133	54	0,058	75	0,559
13	0,125	34	0,409	55	0,259	76	0,629
14	0,890	35	0,128	56	0,054	77	0,487
15	0,097	36	0,876	57	0,475	78	0,239
16	0,887	37	0,929	58	0,655	79	0,629
17	0,159	38	0,820	59	0,493	80	0,890
18	0,781	39	0,023	60	0,403	81	0,494
19	0,669	40	0,078	61	0,819	82	0,327
20	0,333	41	0,745	62	0,661	83	0,092
21	0,210	42	0,645	63	0,307	84	0,864

Quedando la muestra de calibración y la de predicción de acuerdo tal y como se puede ver en la tabla 44.

Tabla 44: Muestra de calibración y de predicción de repartos del vehículo de 3500 kg. PMA código postal 08198

Día	REPARTOS. MUESTRA DE CALIBRACIÓN			Día	REPARTOS. MUESTRA DE PREDICCIÓN		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	0	0	0	1	3	1.500	34
2	1	610	13	2	1	392	10
3	0	0	0	3	1	379	10
4	3	1.820	39	4	3	1.442	33
5	0	0	0	5	1	1.024	19
6	0	0	0	6	1	586	13
7	4	3.262	63	7	0	0	0
8	1	608	13	8	3	2.603	49
9	2	1.557	30	9	1	330	9
10	0	0	0	10	0	0	0
11	2	718	19	11	1	314	9
12	0	0	0	12	0	0	0
13	0	0	0	13	0	0	0
14	2	913	22	14	2	739	19
15	1	400	10	15	1	643	13
16	4	1.918	44	16	3	2.366	46
17	2	596	17	17	1	800	16
18	4	1.644	41	18	0	0	0
19	1	320	9	19	1	1.691	28
20	3	2.410	47	20	2	4.009	63
21	3	1.991	41	21	0	0	0
22	1	591	13	22	1	3.161	46
23	1	2.184	34	23	4	2.434	51
24	4	1.518	39	24	4	2.342	50
25	2	1.419	29	25	1	1.195	21
26	2	961	22	26	1	440	11
27	0	0	0	27	0	0	0
28	3	4.911	81	28	2	5.344	80
29	8	4.543	98	29	1	379	10
30	2	1.494	30	30	0	0	0
31	0	0	0	31	1	326	9
32	5	2.916	63	32	3	1.593	35
33	0	0	0	33	3	1.510	34
34	2	1.144	25	34	2	749	19
35	3	1.422	33	35	1	473	11
				36	3	1.515	34
				37	3	2.287	45
				38	3	3.107	56
				39	0	0	0
				40	0	0	0
				41	2	1.093	24
				42	2	888	21
				43	2	1.906	35
				44	1	724	14
				45	2	1.104	24
				46	1	1.865	30
				47	1	396	10
				48	1	2.008	32
				49	0	0	0

Se puede observar como para este caso, debido a la asignación aleatoria, la muestra de calibración a quedado conformada por 35 observaciones, mientras que la predicción a esta compuesta por 49.

Para el caso de la separación de las muestras de calibración y predicción relativas a las recogidas del vehículo de peso máximo autorizado de 3500 kg. en este código postal, la asignación de números aleatorios para cada una de las observaciones es la que aparece en la tabla 45.

Tabla 45: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 3500 kg. PMA código postal 08198

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,170	22	0,884	43	0,221	64	0,766
2	0,923	23	0,131	44	0,543	65	0,183
3	0,289	24	0,600	45	0,072	66	0,330
4	0,497	25	0,278	46	0,455	67	0,806
5	0,157	26	0,355	47	0,728	68	0,661
6	0,454	27	0,030	48	0,978	69	0,297
7	0,314	28	0,536	49	0,758	70	0,382
8	0,198	29	0,392	50	0,055	71	0,090
9	0,531	30	0,151	51	0,135	72	0,645
10	0,718	31	0,214	52	0,680	73	0,134
11	0,654	32	0,962	53	0,658	74	0,521
12	0,957	33	0,249	54	0,945	75	0,734
13	0,534	34	0,844	55	0,519	76	0,005
14	0,588	35	0,317	56	0,502	77	0,411
15	0,028	36	0,255	57	0,796	78	0,574
16	0,463	37	0,094	58	0,427	79	0,924
17	0,049	38	0,799	59	0,915	80	0,241
18	0,866	39	0,697	60	0,523	81	0,609
19	0,377	40	0,584	61	0,754	82	0,986
20	0,118	41	0,720	62	0,141	83	0,827
21	0,037	42	0,646	63	0,559	84	0,905

Y la correspondiente tabla de calibración y predicción es la número 46.

Tabla 46: Muestra de calibración y de predicción de recogidas del vehículo de 3500 kg. PMA código postal 08198

Día	RECOGIDAS. MUESTRA DE CALIBRACIÓN			Día	RECOGIDAS. MUESTRA DE PREDICCIÓN		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	3	3.591	62	1	3	5.000	77
2	2	3.942	61	2	1	3.049	39
3	3	7.622	103	3	1	3.865	43
4	1	2.584	36	4	2	4.990	69
5	1	4.560	47	5	1	3.979	44
6	2	4.549	62	6	1	3.590	42
7	2	3.362	53	7	1	3.316	41
8	3	7.737	101	8	3	4.441	73
9	3	6.210	91	9	1	3.174	40
10	4	7.768	113	10	1	3.133	40
11	5	9.690	143	11	3	5.690	85
12	4	9.275	133	12	3	3.989	67
13	1	2.737	37	13	2	4.758	69
14	2	4.374	59	14	2	4.543	62
15	3	6.200	93	15	3	4.805	76
16	2	2.974	48	16	1	3.049	39
17	3	6.074	87	17	2	6.136	78
18	0	0	0	18	4	8.857	125
19	1	1.423	24	19	1	1.958	31
20	3	3.978	66	20	2	3.366	52
21	2	4.166	62	21	0	0	0
22	1	3.103	40	22	1	1.453	24
23	3	5.525	81	23	2	3.000	48
24	0	0	0	24	2	4.129	58
25	3	8.180	106	25	1	1.391	23
26	1	1.101	19	26	2	3.485	54
27	1	3.297	41	27	4	8.365	116
28	0	0	0	28	4	5.080	82
29	3	6.895	94	29	1	896	17
30	2	4.323	57	30	2	2.621	44
31	3	5.980	89	31	3	5.294	81
32	2	5.199	72	32	3	4.955	77
33	2	5.530	73	33	3	3.919	66
34	3	5.475	84	34	2	3.681	54
35	2	4.287	63	35	3	4.680	72
36	4	5.216	88	36	2	4.733	63
37	1	1.767	28	37	1	3.222	40
38	4	7.798	114	38	4	8.648	118
39	1	1.660	27	39	3	6.903	96
40	0	0	0	40	2	1.993	36
41	3	6.325	94				
42	3	4.536	72				
43	2	3.383	51				
44	4	6.205	94				

Teniendo 44 observaciones para la muestra de calibración y 40 para la de predicción, debido a la asignación aleatoria.

2.1.2.4 Modelo de costes de reparto

Al igual que para el caso del vehículo de 1500 kg. de peso máximo autorizado, los parámetros de regresión del modelo de costes de reparto se calculan mediante el método de máxima verosimilitud. Se va a estudiar qué modelos ofrecen mayores valores de máxima verosimilitud, bajo los supuestos de diferentes tipos de distribuciones de probabilidad para los errores, y suposiciones de no linealidad de las variables exógenas. Estos supuestos son los mismos que para el caso del vehículo anterior.

Tabla 47: VMV modelo de costes de reparto con distribución de probabilidad normal vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
34,885					-112,6		0 %	22,28
3,146	12,022				-95,2	17,306	73,9 %	11,39
-0,045	4,535	0,014			19,1	114,329	100 %	0,12
-0,089	4,567	0,014	-0,004		19,4	0,269	100 %	0,12
-0,089	4,567	0,014		0	15,1	-4,244	100 %	0,14

Tabla 48: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
3,552					-112,6		0 %	22,28
2,918	0,219				-98,5	14,060	66,1 %	12,97
2,794	0,070	0,0003			-79,7	18,779	92,1 %	6,26
2,270	0,362	0,0003	-0,028		-65,4	14,282	97,4 %	3,62
2,270	0,362	0,0003	-0,028	0	-65,4	0,000	97,2 %	3,71

Tabla 49: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
40,000					-128,4		-1%	0,072
0,297	-0,069				-124,0	4,375	-270%	42,85
0,302	-0,068	-0,000004			-123,7	0,313	-287%	43,82
-0,008	0,072		-0,016		-122,1	1,636	-307%	44,91

Tabla 50: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-5,906					-112,6		0	22,28
-3,562	-0,839				-96,5	16,040	71,1 %	11,98
-3,221	-0,306	-0,001			-65,4	31,063	97,5 %	3,54
-2,450	-0,779	-0,001	0,050		-51,6	13,841	99,1 %	2,08
-2,450	-0,779	-0,001	0,050	0	-51,6	0,000	99,1 %	2,13

Tabla 51: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
30,490					-111,8		0 %	22,72
0,429	12,252				-89,4	22,394	72,9 %	11,60
-0,005	4,498	0,014			20,6	109,999	100 %	0,13
-0,005	4,498	0,014	0,000001		20,6	0,000	100 %	0,13
-0,005	4,498	0,014		0	20,6	0,000	100 %	0,13

Tabla 52: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
3,417					-111,8		0 %	22,72
2,842	0,218				-96,0	15,773	64,2 %	13,33
2,800	0,095	0,0003			-82,2	13,758	90,3 %	6,94
2,445	0,292	0,0003	-0,021		-67,7	14,549	96,9 %	3,91
2,445	0,292	0,0003	-0,021	0	-67,7	0,000	96,8 %	4,01

Tabla 53: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
40,0000					-131,1		0%	42,00
0,042	-0,004				-100,0	31,108	56%	14,80
0,042	-0,004	0,000000001			-100,0	-0,003	54%	15,14
0,042	-0,004		0		-100,0	0,003	54%	15,14

Tabla 54: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-5,522					-111,8		0 %	22,72
-3,219	-0,843				-92,7	19,039	67,9 %	12,62
-3,221	-0,389	-0,001			-65,8	26,918	96,3 %	4,27
-2,273	-0,865	-0,001	0,056		-50,1	15,688	99 %	2,25
-2,273	-0,865	-0,001	0,056	0	-50,1	0,000	98,9 %	2,30

Tabla 55: VMV modelo de costes de reparto con distribución de probabilidad logística vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,388					-128,9		0 %	41,66
1,210	12,096				-91,6	37,320	73,2 %	11,53
-0,012	4,515	0,014			20,2	111,808	100 %	0,12
-0,023	4,524	0,014	-0,001		20,3	-0,020	100 %	0,12
-0,023	4,524	0,014		0	19,9	0,325	100 %	0,13

Se puede observar como los valores de máxima verosimilitud para los supuestos de que los errores se distribuyen según una distribución de probabilidad normal, doble exponencial y logística son casi idénticos. Son también casi idénticos los valores R² ajustada y la varianza. Pero a pesar de esta pequeña diferencia, el valor de máxima verosimilitud mayor es el de la distribución logística. Es por ello que se empleará como modelo de costes de reparto el análisis de regresión bajo el supuesto de que los errores se distribuyen bajo una distribución logística.

2.1.2.4.1 Selección de variables para el modelo de costes de reparto

Si se observa la tabla del apartado anterior correspondiente al supuesto de distribución logística de los errores, se puede observar como al añadir la variable número de repartos por día, el cociente de verosimilitud es superior a dos unidades, por lo que adición de esta variable aporta un mejor ajuste al modelo con su inclusión. Con la variable peso total repartido por día ocurre exactamente lo mismo, ya que al introducirla en el modelo el cociente de verosimilitud aumenta en 111 unidades. Sin embargo, al introducir la variable número de repartos por día elevada a la segunda potencia, el cociente de verosimilitud es de $-0,020$, por lo que el modelo al incluir esta variable es peor que el obtenido solamente con las variables número de repartos por día y peso total repartido por día. Por tanto, como esta variable no aporta un mejor ajuste a los datos, al calcular el valor de máxima verosimilitud con las dos primeras variables más la variable peso total repartido elevada a la segunda potencia, no se debe incluir la variable número de repartos por día elevada al cuadrado. Al obtener el valor de máxima verosimilitud se observa como el cociente de verosimilitud es de tan solo $0,325$, por lo que la adición de esta variable al modelo no aporta un mejor ajuste.

En definitiva, este modelo de costes de reparto se compone de dos variables, número de repartos por día, con un parámetro de regresión igual $4,515$, y peso total repartido por día, con su correspondiente parámetro de regresión igual a $0,014$. El valor umbral para este modelo es igual a $-0,023$.

2.1.2.4.2 Adecuación del modelo de costes de reparto

Cuando los errores se distribuyen con una distribución de probabilidad no normal, si éstos son independientes y se distribuyen idénticamente, las inferencias bajo la teoría de normalidad son asintóticamente correctas. Por tanto, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. Es análisis se puede observar en la figura 34.

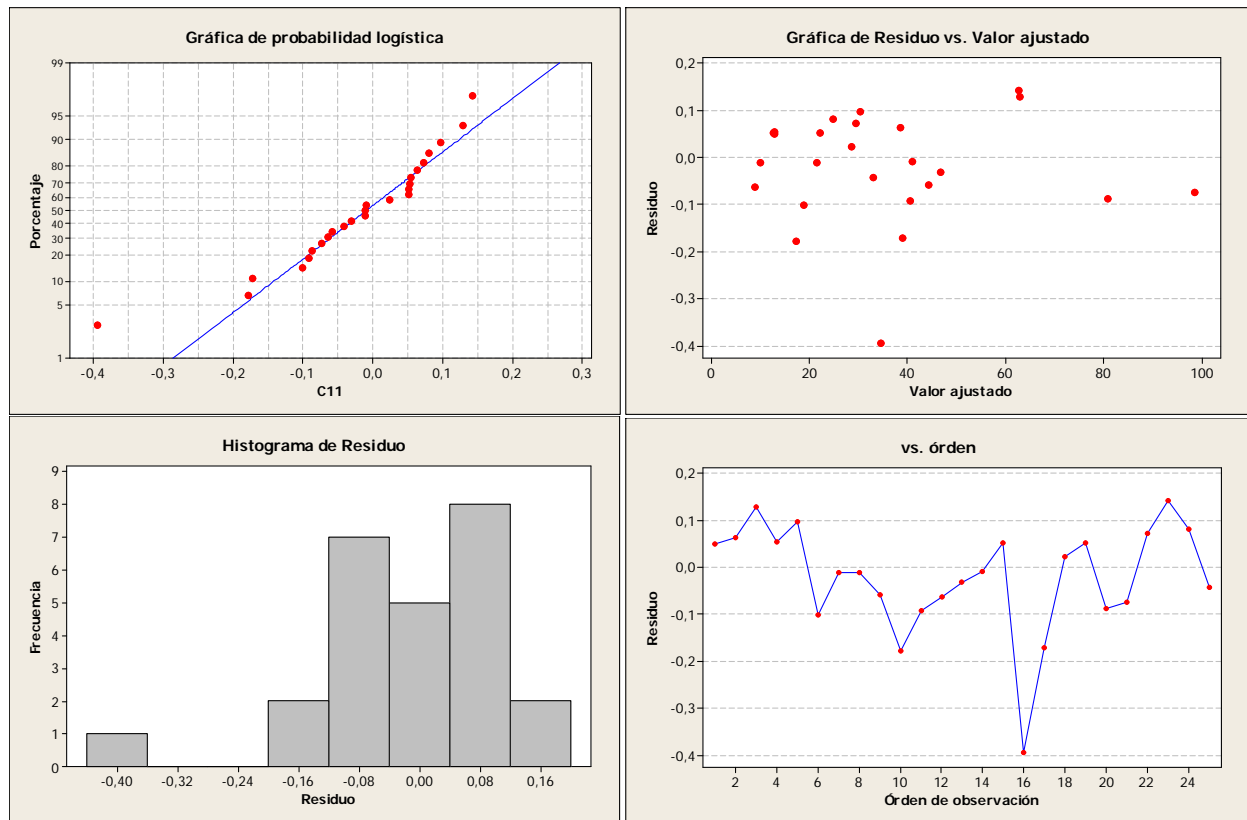


Figura 34: Análisis gráficos de los residuales para modelo de costes de reparto vehículo 3500 kg. PMA código postal 08198

En la primera gráfica se puede observar como los errores se distribuyen de acuerdo a una distribución logística ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.1.2.4.3 Validación del modelo de costes de reparto

La validación del modelo de costes de reparto para este tipo de vehículo se lleva a cabo mediante la prueba *t* pareado ó prueba de Welch. En ella se comparan uno a uno los valores de variable endógena dadas por el modelo, tras introducir los valores de variable exógena de la muestra de predicción, con los valores correspondientes de los costes de reparto de cada día de la muestra de predicción. El proceso es análogo al del tipo de vehículo anterior. Los datos de la prueba de Welch son los siguientes, en la cual se han eliminado aquellas observaciones de la muestra de predicción en las que los costes son cero, quedando 38 observaciones. La tabla en la que se resume la validación es la tabla 56.

Tabla 56: Resultado prueba Welch para el modelo de costes de reparto del vehículo 3500 kg. PMA para el código postal 08198

	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	29,59	28,28	0,105	Límite superior I.C	$t_{0,025;36}$ 1,96
Varianza	387,88	314,09	0,23		
Varianza/n			0,006	Límite inferior I.C	
Máximo	82,28	80,16		-0,047	
Mínimo	8,79	8,74			

Y la tabla de donde se obtiene los valores Z para la prueba es la tabla 57.

Tabla 57: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 3500 kg. para el código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
-0,045	4,535	0,014			
	3	1500	34,12	34,05	0,07
	1	392	9,86	9,86	0,00
	1	379	9,69	9,68	0,01
	3	1442	33,33	33,33	-0,01
	1	1024	18,53	18,79	-0,26
	1	586	12,52	12,60	-0,07
	3	2603	49,24	49,46	-0,21
	1	330	9,01	8,97	0,04
	1	314	8,79	8,74	0,05
	2	739	19,16	19,08	0,08
	1	643	13,30	13,38	-0,08
	3	2366	45,99	45,77	0,22
	1	800	15,46	15,53	-0,07
	1	1691	27,67	27,67	0,00
	2	4009	63,98	62,96	1,02
	1	3161	47,82	46,09	1,73
	4	2434	51,46	51,40	0,06
	4	2342	50,20	50,24	-0,04
	1	1195	20,87	21,07	-0,20
	1	440	10,52	10,55	-0,03
	2	5344	82,28	80,16	2,12
	1	379	9,69	9,68	0,01
	1	326	8,96	8,92	0,04
	3	1593	35,40	35,46	-0,06
	3	1510	34,26	34,22	0,04
	2	749	19,29	19,22	0,07
	1	473	10,97	11,03	-0,05
	3	1515	34,33	34,31	0,02
	3	2287	44,91	45,20	-0,29
	3	3107	56,15	56,24	-0,08
	2	1093	24,01	24,02	-0,02
	2	888	21,20	21,22	-0,02
	2	1906	35,15	35,25	-0,09
	1	724	14,41	14,49	-0,07
	2	1104	24,16	24,26	-0,10
	1	1865	30,06	29,99	0,07
	1	396	9,92	9,92	0,00
	1	2008	32,02	31,90	0,12

El modelo de costes de reparto del tipo de vehículo 3500 kg. de peso máximo autorizado para este código postal se considera válido ya que el intervalo obtenido en la prueba de Welch contiene el cero.

2.1.2.5 Modelo de costes de recogida

Tras la consecución del modelo de costes de reparto, a continuación se crea el modelo de costes de recogida. Los pasos a seguir son análogos a los del modelo de costes de reparto. Los valores de máxima verosimilitud para cada uno de los supuestos se pueden observar en las siguientes tablas:

Tabla 58: VMV modelo de costes de recogida con distribución de probabilidad normal vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máxima veros.	Coci. de veros.	R ² adj.	σ
71,743					-193,2		0%	30,02
7,574	26,191				-153,8	39,451	85%	11,50
-0,387	9,700	0,010			-91,2	62,581	99%	2,50
-1,116	10,399	0,010	-0,135		-91,2	0,001	99%	2,53
1,285	9,733	0,009		0,0000001	-91,3	-0,060	99%	2,57

Tabla 59: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máxima veros.	Coci. de veros.	R ² adj.	σ
4,273					-192,3		0%	30,01
3,424	0,325				-158,5	33,820	81%	13,05
3,424	0,325	0,000004			-158,0	0,465	81%	13,07
2,955	0,670		-0,056		-154,1	3,900	84%	11,86
2,955	0,670		-0,056	0	-154,1	0,000	83%	12,02

Tabla 60: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máxima veros.	Coci. de veros.	R ² adj.	σ
41,000					-230,8		0%	78,59
0,023	-0,003				-166,3	64,477	72%	15,88
0,023	-0,003	0			-166,3	0,000	71%	16,10
0,0350	-0,0111		0,0011		-156,8	9,565	82%	12,85

Tabla 61: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máxima veros.	Coci. de veros.	R ² adj.	σ
-8,470					-192,3		0%	30,01
-4,648	-1,503				-155,1	37,242	84%	11,98
-4,169	-0,567	-0,001			-111,9	43,128	98%	4,13
-3,246	-1,331	-0,001	0,134		-94,7	17,253	99%	2,72
-3,246	-1,331	-0,001	0,134	0	-94,7	0,000	99%	2,76

Tabla 62: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máxima veros.	Coci. de veros.	R ² adj.	σ
67,591					-195,9		0%	30,30
9,732	25,982				-153,8	42,093	85%	11,60
0,504	9,064	0,010			-91,7	62,051	99%	2,53
0,504	9,065	0,010	0,000002		-91,7	0,001	99%	2,56
0,504	9,065	0,010		0	-91,7	0,000	99%	2,56

Tabla 63: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máxima veros.	Coci. de veros.	R ² adj.	σ
4,216					-195,9		0%	30,28
3,436	0,324				-160,9	34,968	81%	13,08
3,349	0,132	0,00011			-133,4	27,483	94%	7,19
2,946	0,406	0,00012	-0,047		-109,4	24,055	98%	4,07
2,946	0,406	0,00012	-0,047	0	-109,4	0,000	98%	4,13

Tabla 64: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máxima veros.	Coci. de veros.	R ² adj.	σ
35,000					-238,6		0%	78,58
0,023	-0,004				-170,3	68,311	71%	16,20
0,023	-0,004	0			-170,3	0,000	70%	16,42
0,035	-0,012		0,001		-158,2	12,104	81%	13,05

Tabla 65: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máxima veros.	Coci. de veros.	R ² adj.	σ
-8,142					-196,4		0%	30,51
-4,833	-1,456				-156,0	40,376	83%	12,05
-4,107	-0,626	-0,001			-113,1	42,931	98%	4,18
-3,540	-1,078	-0,001	0,095		-97,1	16,013	99%	2,91
-3,540	-1,078	-0,001	0,095	0	-97,1	0,000	99%	2,95

Tabla 66: VMV modelo de costes de recogida con distribución de probabilidad logística vehículo 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máxima veros.	Coci. de veros.	R ² adj.	σ
0,098					-233,0		0%	78,52
0,184	24,042				-218,6	14,398	66%	17,48
0,184	22,661	0,002			-155,4	63,170	86%	11,06
0,092	9,239	0,0098	0,014		-91,2	64,230	99%	2,53
0,092	9,239	0,0098	0,014	0	-91,2	0,000	99%	2,57

El modelo que mayor valor de máxima verosimilitud posee es el que supone que los errores se distribuyen según una distribución de probabilidad logística, con una diferencia muy pequeña respecto al modelo que supone que los errores se distribuyen de manera normal. Por tanto, se empleará el modelo

creado bajo el supuesto de que los errores se distribuyen de acuerdo a una distribución de probabilidad logística.

2.1.2.5.1 Selección de variables para el modelo de costes de recogida

El modelo seleccionado para llevar a cabo la simulación sobre los costes de recogida para el vehículo de 3500 kg. de peso máximo autorizado en el código postal 08198, consta de tres variables: número de recogidas por día, peso total repartido por día y número de recogidas elevada a la segunda potencia. Como se puede observar en la tabla correspondiente en el apartado anterior, al introducir una a una estas variables en el modelo, el cociente de verosimilitud es superior a dos unidades, por lo que la adición de cada una de ellas aporta un mejor ajuste. No ocurre lo mismo al introducir la variable peso total repartido elevada a la segunda potencia, ya que el cociente de verosimilitud es cero, y por tanto, esta variable no aporta un mejor ajuste datos.

Los parámetros de regresión son 9,239 para la variable número de recogidas, 0,0098 para la variable peso total repartido, y 0,014 para el número de recogidas elevado al cuadrado, siendo el valor umbral para este modelo de regresión igual a 0,092.

2.1.2.5.2 Adecuación del modelo de costes de recogida

Cuando los errores se distribuyen con una distribución de probabilidad no normal, si éstos son independientes y se distribuyen idénticamente, las inferencias bajo la teoría de normalidad son asintóticamente correctas. Por tanto, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. Este análisis se puede ver en la figura 35.

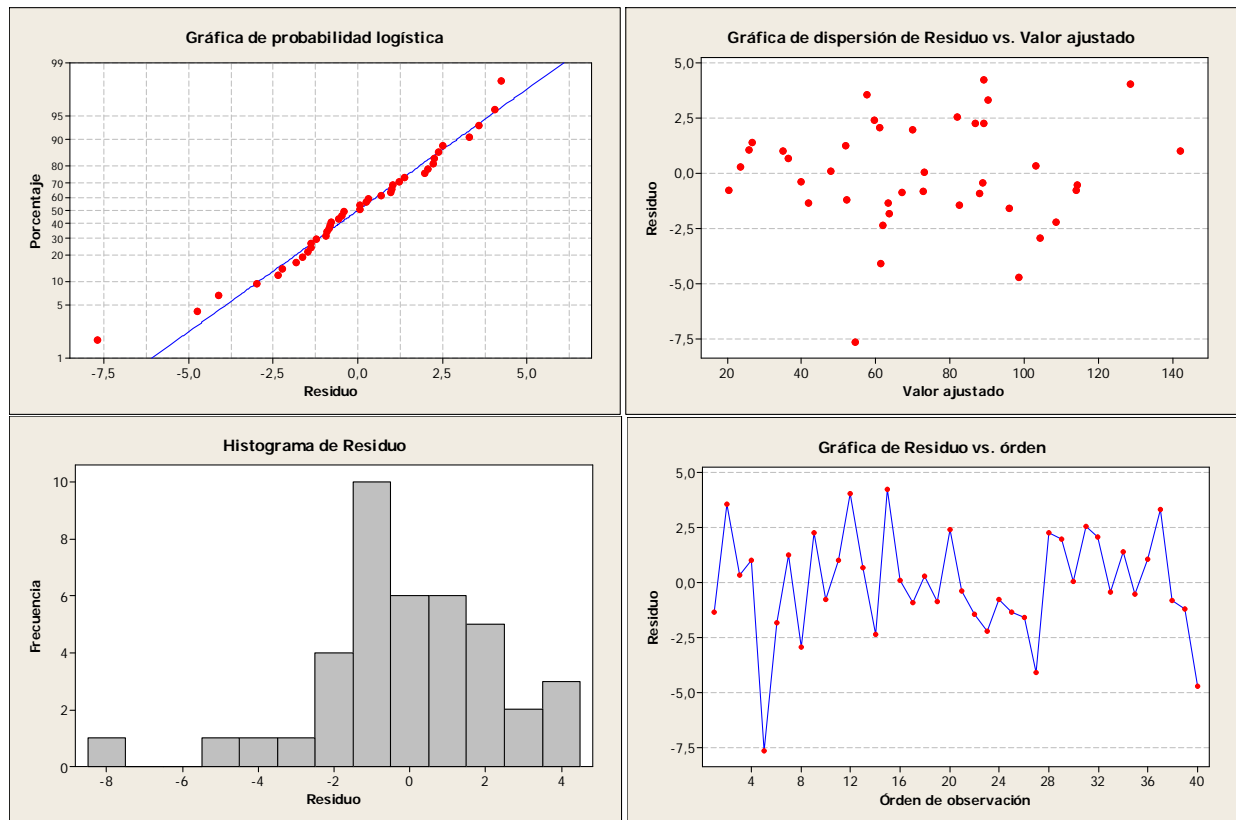


Figura 35: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 3500 kg. PMA código postal 08198

En la primera gráfica se puede observar como los errores se distribuyen según una distribución logística ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.1.2.5.3 Validación del modelo de costes de recogida

Para la validación del modelo se va a emplear la muestra de predicción obtenida anteriormente en el apartado 2.1.2.3. Con esta muestra se va a realizar una prueba *t* pareada, comparando la respuesta dada por el modelo con los costes de recogida de la muestra de predicción. El cuadro que resume la prueba *t* pareada es la tabla 67.

Tabla 67: Resultado prueba Welch para el modelo de costes de recogida del vehículo 3500 kg. PMA para el código postal 08198

	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	61,08	60,59	0,486	Límite superior I.C	$t_{0,025;36}$
Varianza	667,66	663,31	3,55	1,078	1,96
Varianza/n			0,091		
Máximo	124,07	125,45		Límite inferior I.C	
Mínimo	18,13	16,70		-0,105	

Como la prueba pareada incluye el cero, se concluye que el modelo predice de manera adecuada para nuevos valores. Los resultados de la diferencia entre los valores estimados por el modelo y los de la muestra de predicción, que son los valores Z que se emplean en la elaboración de la prueba *t* pareada, se muestran en la tabla 68.

Tabla 68: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 3500 kg. para el código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
0,092	9,239	0,0098	0,014			
	3	5000	9	76,94	76,98	-0,04
	1	3049	1	39,23	39,29	-0,07
	1	3865	1	47,22	43,27	3,95
	2	4990	4	67,53	69,08	-1,56
	1	3979	1	48,34	43,83	4,51
	1	3590	1	44,53	41,93	2,59
	1	3316	1	41,84	40,60	1,25
	3	4441	9	71,46	73,30	-1,84
	1	3174	1	40,45	39,90	0,55
	1	3133	1	40,05	39,70	0,34
	3	5690	9	83,70	84,73	-1,04
	3	3989	9	67,03	67,24	-0,22
	2	4758	4	65,25	68,59	-3,33
	2	4543	4	63,15	62,37	0,78
	3	4805	9	75,02	76,44	-1,41
	1	3049	1	39,23	39,29	-0,07
	2	6136	4	78,76	78,36	0,40
	4	8857	16	124,07	125,45	-1,38
	1	1958	1	28,53	30,67	-2,13
	2	3366	4	51,61	51,52	0,10
	1	1453	1	23,58	24,03	-0,45
	2	3000	4	48,03	48,24	-0,21
	2	4129	4	59,09	57,96	1,13
	1	1391	1	22,98	23,22	-0,24
	2	3485	4	52,78	54,43	-1,65
	4	8365	16	119,25	116,45	2,80
	4	5080	16	87,06	82,25	4,80
	1	896	1	18,13	16,70	1,42
	2	2621	4	44,31	44,27	0,04
	3	5294	9	79,82	80,98	-1,17
	3	4955	9	76,49	77,27	-0,77
	3	3919	9	66,34	66,30	0,05
	2	3681	4	54,70	53,76	0,94
	3	4680	9	73,80	71,57	2,23
	2	4733	4	65,01	63,08	1,93
	1	3222	1	40,92	40,14	0,78
	4	8648	16	122,02	117,77	4,26
	3	6903	9	95,58	96,02	-0,44
	2	1993	4	38,16	36,03	2,13

2.1.2.6 Modelo de carga de trabajo

En este modelo se van a estudiar la distribución de probabilidad que siguen las variables aleatorias que conforman la carga la de trabajo, así como sus parámetros. Al igual que en el caso del vehículo de 1500 kg. de peso máximo autorizado estas variables son el número de repartos por día y el peso de cada uno de los repartos, en el caso de los repartos, y el número de recogidas por día y el peso de cada recogida, en el de las recogidas.

Las muestras que se emplean para las variables número de repartos y número de recogidas son las mismas muestras que las muestras de calibración que se emplearon para ajustar los modelos de los costes de reparto y recogida. Las muestras para las variables peso de cada reparto y peso de cada recogida, es una muestra con los pesos unitarios de cada uno de los repartos y de cada una de las recogidas.

2.1.2.6.1. Número de repartos por día

A través de Minitab se obtienen los parámetros para cada una de las distribuciones de probabilidad que hacen que éstas sean las que mejor se ajusten a los valores de la muestra. Esta variable aleatoria es un conteo de algún elemento, y por tanto, una variable discreta. Al ser una variable discreta, la primera distribución de probabilidad que se comprueba es la distribución de Poisson. El test no paramétrico con el que se comprueba si el ajuste para las distribuciones discretas es adecuado es el test Chi cuadrado. Para los datos de la muestra de calibración de esta variable, Minitab indica que la distribución de Poisson a comprobar es aquella con parámetro lambda igual a 1,89. Los resultados del test paramétrico para esta distribución de Poisson son los que se pueden ver en la tabla 69.

Tabla 69: Prueba Chi Cuadrado número de repartos para el vehículo 3500 kg. PMA código postal 08198

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
0	10	0,286	0,1517	5	4,1418	alfa=0,05 k=5 Par. Est. =1 n=5-1-1=3
1	6	0,171	0,2861	10	1,6087	
2	8	0,229	0,2698	9	0,2200	
3	5	0,143	0,1696	6	0,1472	
4-8	6	0,171	0,1229	4	0,6718	
	35				6,79	

En la tabla se puede observar como el valor límite para aceptar la hipótesis de que la una distribución de Poisson con lambda 1,89 ajusta adecuadamente la variable número de repartos por día es 7,85. Este valor corresponde a una distribución Chi cuadrado con 3 grados de libertad y nivel de significancia del 5%. Los grados de libertad es igual al número de intervalos de los que consta el test, al que hay que restar el número de parámetros estimados (para el caso de la distribución de Poisson es únicamente el parámetro lambda) y otra unidad. Como en este caso el valor del test es 6,79, menor al

valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.1.2.6.2. Peso de cada reparto

La segunda variable a analizar para el modelo de carga de trabajo es el peso de cada reparto. Esta variable es una variable aleatoria continua, y al igual que para el resto de variables continuas, las distribuciones de probabilidad a analizar serán también continuas. Los valores de los parámetros de las distintas distribuciones dados por Minitab y los resultados del test Kolmogorov-Smirnov para cada una de las distribuciones analizadas se pueden ver en la tabla 70.

Tabla 70: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 3500 kg. PMA código postal 08198

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,256	alfa	0,181	alfa	0,202	Media	0,135	Lambda	0,190
639,2	2,083	3,517	1,462	1,703	1,634	6,311	1,092	553,1	1,539
Desv. Est.	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	Desv. Est.	$c_{\alpha} = 1,358$	Theta	$c_{\alpha} = 1,358$
424,3	$\alpha = 0,05$	181,8	$\alpha = 0,05$	724,6	$\alpha = 0,05$	0,5065	$\alpha = 0,05$	194,8	$\alpha = 0,05$

En la tabla mostrada se puede ver como la única de las distribuciones que ajusta satisfactoriamente los datos de la muestra de la variable es la distribución de probabilidad lognormal. Por tanto, en la simulación de la carga de trabajo se empleará esta distribución para la calcular el peso de cada reparto, y a partir de éstos, el peso total repartido por día.

2.1.2.6.3. Número de recogidas por día

Al igual que para la variable número de repartos por día, esta variable es un conteo de algún elemento. Por ello, la primera distribución de probabilidad que se va a comprobar si ajusta adecuadamente a los valores de la muestra de esta variable es la distribución de probabilidad discreta de Poisson. Al ser una distribución discreta, el test no paramétrico a realizar es el test Chi cuadrado. El resultado de este test se puede ver en la tabla 71.

Tabla 71: Prueba Chi Cuadrado número de recogidas para el vehículo 3500 kg. PMA código postal 08198

Observaciones	O _i	f(O _i)	f(E _i)	E _i	Chi-cuadrado	
0	4	0,091	0,1078	5	0,1167	alfa=0,05 k=5 Par. Est. =1 n=5-1-1=3
1	9	0,205	0,2401	11	0,2323	
2	11	0,250	0,2674	12	0,0500	
3	14	0,318	0,1986	9	3,1714	
4-5	6	0,136	0,1598	7	0,1513	
	44				3,72	

El valor lambda dado por Minitab para una distribución de Poisson con los datos de la muestra de esta variable aleatoria es 2,23. En la tabla se puede observar como el valor límite para aceptar la hipótesis de que la una distribución de Poisson con lambda 2,23 ajusta adecuadamente la variable número de repartos por día es 7,85. Este valor corresponde a una distribución Chi cuadrado con 3 grados de libertad y nivel de significancia del 5%. Los grados de libertad es igual al número de intervalos de los que consta el test, al que hay que restar el número de parámetros estimados (para el caso de la distribución de Poisson es únicamente el parámetro lambda) y otra unidad. Como en este caso el valor del test es 3,72, menor al valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.1.2.6.4. Peso de cada recogida

Esta variable aleatoria es continua, y se determinan los valores de cada distribución, así como la bondad de ajuste de las mismas de la misma manera que las variables continuas anteriores. Los resultados son los de la tabla 72.

Tabla 72: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 3500 kg. PMA código postal 08198

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,109	alfa	0,082	alfa	0,089	Media	0,097	Lambda	0,095
2027	1,092	5,016	0,817	2,497	0,890	7,511	0,970	1988	0,952
Desv. Est.	c _α = 1,358	beta	c _α = 1,358	beta	c _α = 1,358	Desv. Est.	c _α = 1,358	Theta	c _α = 1,358
887,5	α = 0,05	404	α = 0,05	2293	α = 0,05	0,4686	α = 0,05	531,7	α = 0,05

En esta ocasión es posible ajustar la variable aleatoria peso de cada recogida mediante alguna distribución de probabilidad. Más concretamente las cinco distribuciones de probabilidad que se han empleado han superado satisfactoriamente el test de Kolmogorov-Smirnov, y podrían emplearse para

modelar esta variable. La que menor discrepancia ha mostrado ha sido la distribución gamma. Como su rango de aplicación es solamente para valores mayores que cero, también posibilita que no sea necesario el emplear una distribución truncada.

2.1.2.7 Simulación de los costes de reparto y recogida, y de la carga de trabajo

En este apartado se llevará a cabo la simulación de los modelos creados para el estudio del vehículo de 3500 kg. de peso máximo autorizado en el código postal 08198. En primer lugar se simulará la carga de trabajo. Para ello hay que determinar el número de repartos y de recogidas para cada uno de los días simulados, y el peso total repartido y el peso total recogido por día. Empleando los ajustes de las distribuciones de probabilidad de Poisson obtenidos en los apartados 2.1.2.6.1 y 2.1.2.6.3. se simularán el número de repartos y el número de recogidas para cada uno días que se desean simular.

Posteriormente, se simulará el peso correspondiente de cada uno de los repartos y las recogidas mediante las distribuciones de probabilidad calculadas en los apartados 2.1.2.6.2. y 2.1.6.6.4. Sumando el peso unitario de todos los repartos simulados en cada uno de los días, se calculará el peso total repartido por día. Con el peso unitario de las recogidas simulado para cada uno de los días, se calculará el peso total recogido por día.

Finalmente, se simularán los modelos de costes de reparto y costes de recogida. Como valores de entrada a estos modelos se emplearán los datos simulados para el modelo de carga de trabajo. El número de repartos y el peso total repartido por día serán los datos de entrada para el modelo de costes de reparto. Análogamente, el número de recogidas y el peso total recogido por día serán los valores de entrada para el modelo de costes de recogida.

2.1.2.7.1 Simulación de los costes de reparto

La carga de trabajo relativa a los repartos para el vehículo de 3500 kg de peso máximo autorizado del código postal 08198 es la siguiente: el número de repartos puede ser descrito mediante una distribución de probabilidad de Poisson con parámetro lambda igual a 1,89. Mientras que el peso de cada reparto sigue una distribución de probabilidad lognormal con valor medio igual a 6,311 y con desviación estándar igual a 0,5065.

Un resumen de la carga de trabajo simulada se puede ver en la tabla 73, en el cual se puede observar el promedio de repartos por día simulados, su desviación estándar, el número máximo de repartos a realizar en un día, así como el número mínimo. Y los mismos datos para la variable peso total repartido.

Tabla 73: Resumen de la simulación de la carga de trabajo relativo a los repartos del vehículo 3500 kg. PMA en el código postal 08198

	Repartos por día	Peso total repartido por día
Promedio	2	1141
Desv. Estan.	1	983
Máximo	7	6757
Mínimo	0	0

La simulación de la carga de trabajo en 250 días bajo las condiciones previamente enunciadas, son las entradas para la simulación en el modelo de costes de reparto. Para cada uno de los días se obtiene un estimado puntual del coste de reparto para la carga de trabajo simulada correspondiente. A partir de cada estimado puntual se desarrolla un intervalo de predicción, teniendo un total de 250 intervalos de predicción. Estos intervalos de predicción se calculan tal y como se explicó en el apartado 4.1.1.3. de la metodología. Los resultados que aquí se exponen son el valor medio de límite inferior de las 250 simulaciones, la varianza y desviación estándar de este límite inferior, así como un intervalo de confianza de este límite inferior. Para el límite superior del intervalo de predicción los resultados que se presentan son análogos. Estos resultados se pueden ver en la tabla 74.

Tabla 74: Resumen de la simulación de los costes de reparto del vehículo de 3500 kg. de PMA en el código postal 08198

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	23,72	24,20
Varianza	357	358
Desv. Estan.	18,91	18,91
I.C	21,74	22,23
	25,69	26,17

En la tabla 75 se puede ver como el límite inferior oscila entre los valores 21,74 y 25,69 euros por día, mientras que el límite superior va desde 22,23 hasta 26,17 euros por día.

2.1.2.7.2 Simulación de los costes de recogida

La carga de trabajo relativa a las recogidas para este código postal queda como sigue: el número de recogidas puede ser expresado mediante una distribución de probabilidad Poisson con parámetro lambda 2,23. El peso de cada recogida, por otro lado, puede ser descrito mediante una distribución de probabilidad gamma con valor alfa igual a 5,016, y un valor beta igual a 404. Un cuadro resumen de la carga de trabajo simulada es la tabla 75, en el cual se puede observar el promedio de recogidas por día simulados, su desviación estándar, el número máximo de recogidas a realizar en un día, así como el número mínimo. Y los datos análogos para la variable peso total recogido.

Tabla 75: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 3500 kg. PMA en el código postal 08198

	Recogidas por día	Peso total recogido por día
Promedio	2	4105
Desv. Estan.	1	2731
Máximo	6	13400
Mínimo	0	0

Para la simulación de los costes de recogida, los resultados que se exponen son idénticos al caso de los costes de reparto. Se exponen los valores medios, varianza y desviación estándar, así como un intervalo de confianza de los límites superior e inferior de los límites de predicción de los 250 días simulados a partir de la carga de trabajo previamente simulada bajo las condiciones expuestas en el párrafo anterior. Los resultados son los de la tabla 76.

Tabla 76: Resumen de la simulación de los costes de recogida del vehículo de 3500 kg. de PMA en el código postal 08198

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	56,11	65,72
Varianza	1422	1557
Desv. Estan.	37,71	39,46
I.C	52,18	61,60
	60,05	69,84

En la tabla se puede observar como el límite inferior de los estimados puntuales de los costes de recogidas se distribuyen con un valor medio de 56,11 euros por día y una desviación estándar de 37,71. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 65,72 y una desviación estándar de 39,46 euros por día.

2.1.3. Vehículo de más de 3500 kg. de peso máximo autorizado para el código postal 08198

Como para este código postal se dan recogidas de más de 3500 kg. de peso, es necesario disponer de vehículos que sean capaces de realizar este tipo de recogidas. Sin embargo, no se dan repartos de más de 3500 kg. de peso, por lo que solamente se confeccionará el modelo de costes de recogida, así como el modelo de carga de trabajo relativo a las recogidas.

2.1.3.1. Toma de muestras

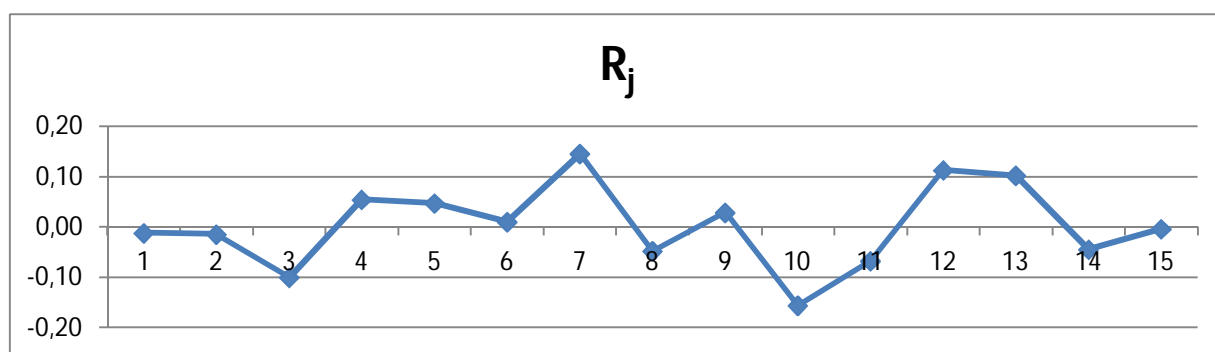
El primer paso, al igual que en los dos tipos de vehículos anteriormente analizados, para comenzar la elaboración del modelo de los costes de recogida es la toma de muestras de las variables identificadas como las necesarias a modelar. Como para este tipo de vehículo sólo se dan recogidas, se han de tomar muestras de los costes de recogida por día, del número de recogidas por día y del peso total recogido por día. Las 84 observaciones obtenidas en el estudio observacional llevado a cabo son las de la tabla 77.

Tabla 77: Toma de muestras de las recogidas del vehículo de más de 3500 kg. PMA para el código postal 08198

Día	RECOGIDAS			Día	RECOGIDAS		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	1	5.956	53	43	1	3.352	41
2	1	5.735	52	44	0	0	0
3	1	4.259	45	45	1	4.679	47
4	2	8.002	88	46	0	0	0
5	0	0	0	47	1	6.811	58
6	2	8.690	91	48	1	4.024	44
7	2	7.518	86	49	1	4.369	46
8	1	5.576	52	50	2	11.351	104
9	1	3.680	42	51	3	11.151	128
10	0	0	0	52	1	5.091	49
11	2	12.683	110	53	2	17.911	136
12	2	10.018	98	54	1	6.269	55
13	1	11.661	81	55	2	11.143	103
14	1	7.641	62	56	1	5.665	52
15	1	7.699	62	57	2	7.674	86
16	1	6.158	54	58	1	7.441	61
17	1	7.449	61	59	1	4.194	45
18	1	5.649	52	60	2	15.053	122
19	1	2.913	38	61	1	10.081	74
20	1	5.506	51	62	1	4.122	45
21	0	0	0	63	0	0	0
22	0	0	0	64	0	0	0
23	1	3.535	42	65	2	17.966	136
24	1	7.000	59	66	1	9.367	70
25	1	4.281	45	67	3	12.393	134
26	2	10.687	101	68	1	7.302	60
27	0	0	0	69	1	9.785	72
28	0	0	0	70	0	0	0
29	2	10.705	101	71	1	9.050	69
30	1	5.069	49	72	1	7.695	62
31	0	0	0	73	1	13.427	90
32	0	0	0	74	2	11.465	105
33	2	13.434	114	75	0	0	0
34	1	6.222	55	76	0	0	0
35	1	4.910	48	77	1	9.566	71
36	1	4.089	44	78	1	13.967	93
37	1	5.194	50	79	0	0	0
38	1	8.724	67	80	2	12.861	112
39	0	0	0	81	1	5.054	49
40	1	7.200	60	82	1	8.616	66
41	1	6.737	57	83	1	5.323	50
42	2	9.638	96	84	2	9.876	97

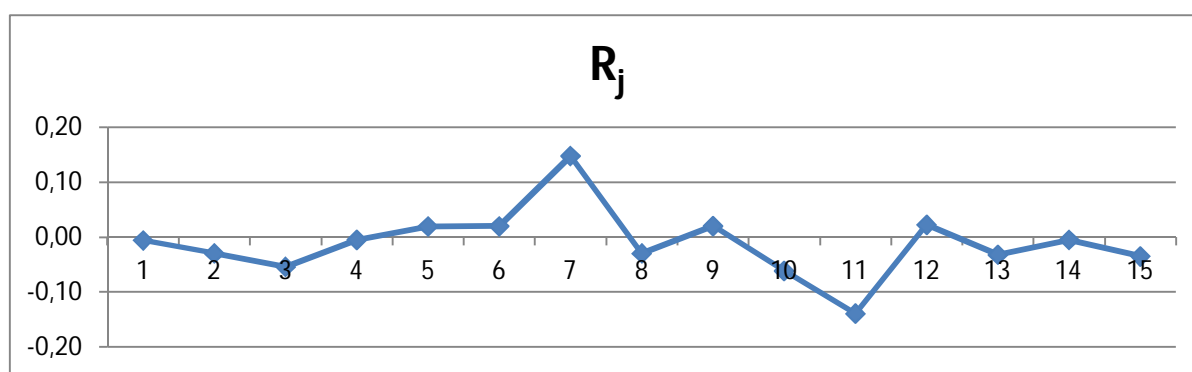
2.1.3.2. Independencia de las muestras

Como en los otros tipos de vehículos analizados en este código postal, tras obtener las muestras pertinentes de cada variable, se procede a comprobar la independencia de las observaciones de cada una de las muestras, para poder aplicar con garantías el método de máxima verosimilitud y los test no paramétricos de bondad de ajuste. Al igual que en los casos anteriores, la independencia se examina mediante las gráficas de autocorrelación y las gráficas de dispersión. Al tener que analizar tres variables diferentes, se lleva a cabo una gráfica de autocorrelación y una gráfica de dispersión para cada una de ellas. La Figura 36 corresponde a la gráfica de autocorrelación de la muestra de los costes de recogida por día. La Figura 37 corresponde a la gráfica de autocorrelación del número de recogidas por día, y la Figura 38 corresponde a la gráfica de autocorrelación del peso total recogido por día.



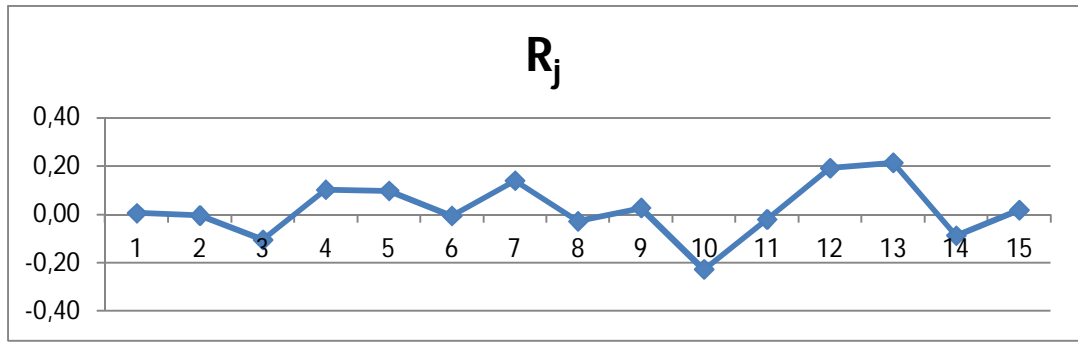
R_j	-0,01	-0,01	-0,10	0,05	0,05	0,01	0,15	-0,05	0,03	-0,16	-0,07	0,11	0,10	-0,04	0,00
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-16	-20	-142	78	68	15	207	-67	41	-221	-95	161	146	-63	-5

Figura 36: Grafica de autocorrelación de coste de recogida por día vehículo de más de 3500 kg. PMA código 08198



R_j	-0,01	-0,03	-0,05	0,00	0,02	0,02	0,15	-0,03	0,02	-0,06	-0,14	0,02	-0,03	0,00	-0,03
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	0,0	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	-0,1	0,0	0,0	0,0	0,0

Figura 37: Grafica de autocorrelación del peso total recogido por día vehículo de más de 3500 kg. PMA código 08198



R_j	0,00	0,00	-0,11	0,10	0,10	-0,01	0,14	-0,03	0,03	-0,23	-0,02	0,19	0,21	-0,09	0,02
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	100737	-85371	-2127610	2061663	1978984	-115538	2821313	-574709	544432	-4592140	-405958	3872798	4316271	-1764991	356742

Figura 38: Gráfica de autocorrelación del peso total recogido por día vehículo de más de 3500 kg. PMA código 08198

Para ninguna de las variables los valores de R_j supera los valores de -1 y 1, y no se da un patrón determinado en ninguna de las gráficas, por lo que de acuerdo a este test, se consideran independientes las observaciones de cada una de las muestras.

La segunda prueba para la comprobación de la independencia de las observaciones de las muestras es la gráfica de dispersión. A continuación se muestran la gráfica de dispersión para cada una de las variables.

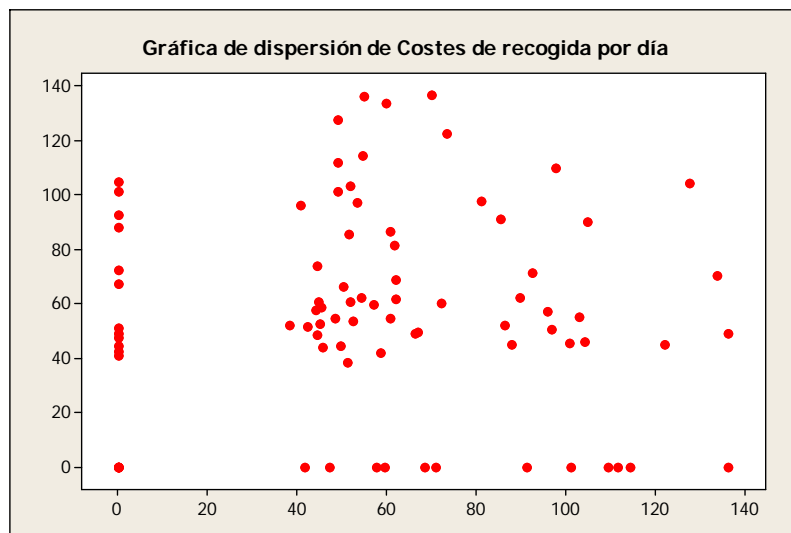


Figura 39: Gráfica de dispersión de coste de recogida por día vehículo de más de 3500 kg. PMA código 08198

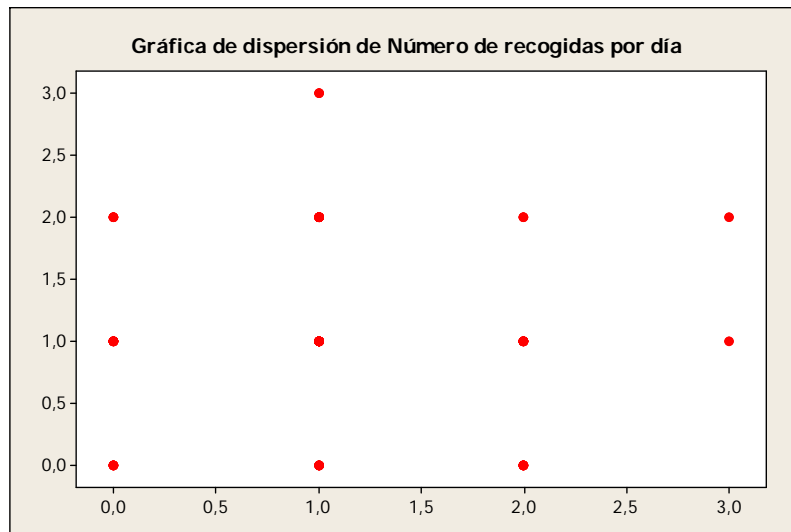


Figura 40: Grafica de dispersión de número de recogidas por día vehículo de más de 3500 kg. PMA código 08198

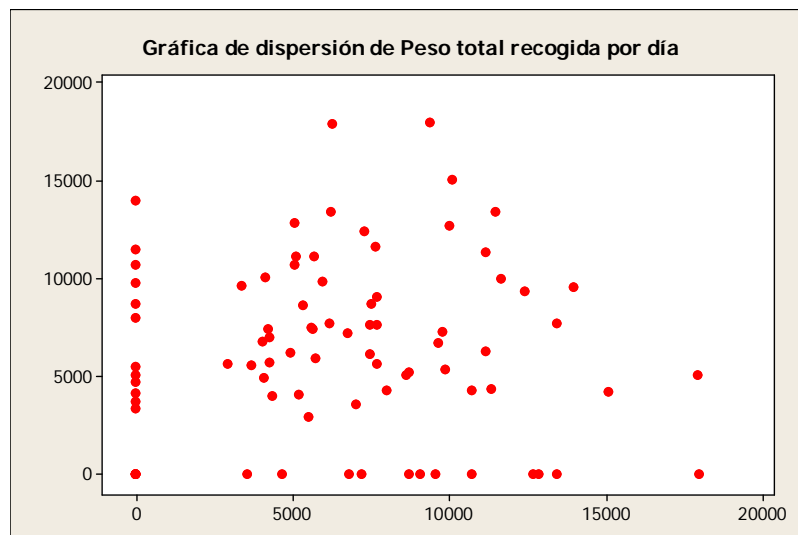


Figura 41: Grafica de dispersión de peso total recogido por día vehículo de más de 3500 kg. PMA código 08198

Como los puntos se distribuyen de manera aleatoria en las tres gráficas, se determina que las observaciones de cada una de las muestras son independientes.

2.1.3.3. Separación de las muestras de calibración y predicción

Tras comprobar que las 84 observaciones de las muestras son independientes, es necesario separar éstas en las muestras de calibración y predicción. La primera se emplea para la elaboración del modelo de costes de recogida, y la segunda para la validación del mismo. Al igual que para los otros tipos de vehículos de este código postal, esta separación se hace mediante asignación aleatoria. La asignación de números aleatorios para cada una de las observaciones es la que se puede ver en la tabla 78.

Tabla 78: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo de más de 3500 kg. PMA código postal 08198

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,702	22	0,719	43	0,210	64	0,917
2	0,950	23	0,882	44	0,350	65	0,955
3	0,853	24	0,904	45	0,741	66	0,020
4	0,963	25	0,430	46	0,885	67	0,889
5	0,225	26	0,536	47	0,750	68	0,904
6	0,343	27	0,090	48	0,871	69	0,467
7	0,907	28	0,334	49	0,419	70	0,341
8	0,099	29	0,423	50	0,494	71	0,030
9	0,662	30	0,651	51	0,347	72	0,930
10	0,960	31	0,471	52	0,449	73	0,459
11	0,782	32	0,104	53	0,311	74	0,825
12	0,836	33	0,591	54	0,409	75	0,590
13	0,175	34	0,347	55	0,968	76	0,921
14	0,288	35	0,410	56	0,981	77	0,292
15	0,627	36	0,435	57	0,256	78	0,080
16	0,803	37	0,683	58	0,098	79	0,643
17	0,123	38	0,724	59	0,995	80	0,089
18	0,888	39	0,528	60	0,286	81	0,793
19	0,994	40	0,730	61	0,748	82	0,061
20	0,914	41	0,916	62	0,747	83	0,126
21	0,106	42	0,976	63	0,722	84	0,292

A partir de estas asignaciones, las muestras de calibración y de predicción para el modelo de costes de reparto quedan como se ve en la tabla 79. Se puede observar como la muestra de calibración consta de 46 observaciones, y la de predicción 38, debido a la aleatoriedad de la asignación de las observaciones en las muestras.

Tabla 79: Muestra de calibración y de predicción de recogidas del vehículo de más de 3500 kg. PMA código postal 08198

Día	RECOGIDAS. MUESTRA DE CALIBRACIÓN			Día	RECOGIDAS. MUESTRA DE PREDICCIÓN		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	1	5.956	53	1	0	0	0
2	1	5.735	52	2	2	8.690	91
3	1	4.259	45	3	1	5.576	52
4	2	8.002	88	4	1	11.661	81
5	2	7.518	86	5	1	7.641	62
6	1	3.680	42	6	1	7.449	61
7	0	0	0	7	0	0	0
8	2	12.683	110	8	1	4.281	45
9	2	10.018	98	9	0	0	0
10	1	7.699	62	10	0	0	0
11	1	6.158	54	11	2	10.705	101
12	1	5.649	52	12	0	0	0
13	1	2.913	38	13	0	0	0
14	1	5.506	51	14	1	6.222	55
15	0	0	0	15	1	4.910	48
16	1	3.535	42	16	1	4.089	44
17	1	7.000	59	17	1	3.352	41
18	2	10.687	101	18	0	0	0
19	1	5.069	49	19	1	4.369	46
20	2	13.434	114	20	2	11.351	104
21	1	5.194	50	21	3	11.151	128
22	1	8.724	67	22	1	5.091	49
23	0	0	0	23	2	17.911	136
24	1	7.200	60	24	1	6.269	55
25	1	6.737	57	25	2	7.674	86
26	2	9.638	96	26	1	7.441	61
27	1	4.679	47	27	2	15.053	122
28	0	0	0	28	1	9.367	70
29	1	6.811	58	29	1	9.785	72
30	1	4.024	44	30	0	0	0
31	2	11.143	103	31	1	9.050	69
32	1	5.665	52	32	1	13.427	90
33	1	4.194	45	33	1	9.566	71
34	1	10.081	74	34	1	13.967	93
35	1	4.122	45	35	2	12.861	112
36	0	0	0	36	1	8.616	66
37	0	0	0	37	1	5.323	50
38	2	17.966	136	38	2	9.876	97
39	3	12.393	134				
40	1	7.302	60				
41	1	7.695	62				
42	2	11.465	105				
43	0	0	0				
44	0	0	0				
45	0	0	0				
46	1	5.054	49				

2.1.3.4. Modelo de costes de recogida

Como para los modelos de los costes de reparto y recogida de los otros dos tipos de vehículos de este código postal, el modelo de costes de recogida para este tipo de vehículo se realiza mediante un análisis de regresión, en el cual los parámetros de regresión de cada una de las variables exógenas se determinan mediante el método de máxima verosimilitud. Al igual que anteriormente, se ajustarán diferentes modelos bajo diferentes premisas, como son diferentes distribuciones de probabilidad para los errores o transformaciones no lineales de las variables exógenas. Los valores de máxima verosimilitud para cada uno de los supuestos se pueden observar en las siguientes tablas:

Tabla 80: VMV modelo de costes de recogida con distribución de probabilidad normal vehículo de más de 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
68,645					-174,3		0%	27,27
5,734	47,493				-139,0	35,324	85%	10,64
0,066	24,403	0,005			9,6	148,567	100%	0,19
0,296	24,060	0,005	0,098		9,9	0,357	100%	0,20
0,066	24,403	0,005		0	9,6	-0,361	100%	0,20

Tabla 81: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo de más de 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
4,229					-174,3		0%	27,27
3,489	0,530				-144,8	29,469	79%	12,47
3,489	0,530	0			-144,8	0,000	78%	12,65
2,866	1,309		-0,211		-137,5	7,323	85%	10,54

Tabla 82: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo de más de 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,015					-174,3		0%	27,27
0,022	-0,005				-152,9	21,408	68%	15,51
0,022	-0,005	0			-152,9	0,000	67%	15,73
0,035	-0,020		0,004		-137,5	15,385	86%	10,38

Tabla 83: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo de más de 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-8,284					-174,3		0%	27,27
-4,799	-2,555				-141,2	33,075	83%	11,31
-4,354	-1,348	-0,0003			-78,9	62,350	99%	2,13
-3,921	-1,992	-0,0003	0,174		-74,1	4,787	100%	1,90
-3,921	-1,992	0,000	0,174	0	-74,1	0,000	100%	1,93

Tabla 84: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo de más de 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
57,638					-173,7		0%	29,46
3,958	48,093				-138,4	35,223	85%	10,70
0,007	24,419	0,005			54,8	193,240	100%	0,20
0,007	24,417	0,005	0,003		58,1	3,295	100%	0,21
0,007	24,417	0,005	0,003	0	58,1	0,000	100%	0,21

Tabla 85: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo de más de 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
4,054					-173,7		0%	29,46
3,340	0,619				-143,2	30,433	76%	13,38
3,340	0,619	0			-143,2	0,000	75%	13,57
2,882	1,267		-0,199		-136,3	6,936	85%	10,61

Tabla 86: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo de más de 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
20					-219,1		0%	74,68
0,023	-0,006				-152,5	66,567	66%	15,82
0,023	-0,006	0			-152,5	0,000	65%	16,05
0,023	-0,006		0		-152,5	0,000	65%	16,05

Tabla 87: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo de más de 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-7,592					-173,7		0%	29,46
-4,548	-2,668				-140,2	33,411	82%	11,45
-4,236	-1,190	-0,0003			-74,4	65,887	99%	3,01
-4,236	-1,190	-0,0003	0		-74,4	0,000	99%	3,06
-4,236	-1,190	-0,0003		0	-74,4	0,000	99%	3,06

Tabla 88: VMV modelo de costes de recogida con distribución de probabilidad logística vehículo de más de 3500 kg. PMA código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,248					-213,5		0%	74,50
0,250	0,053				-213,4	0,131	0%	75,47
0,250		0,010			-192,1	21,262	83%	11,33
0,250		0,010		-0,0000001	-131,6	-29,885	89%	9,13

Tal y como se puede comprobar en las tablas, la premisa bajo la cual el valor de máxima verosimilitud obtenido es mayor es la de que los errores se distribuyen de acuerdo a una distribución doble exponencial. Además, la diferencia con el resto de premisas es considerable, por lo que para la simulación se los parámetros de regresión calculados bajo esta premisa.

2.1.3.4.1. Selección de variables

En la tabla correspondiente a la distribución de probabilidad doble exponencial, se puede comprobar cómo cada vez que se introduce las variables número de recogidas, peso total recogido y número de recogidas elevado a la segunda potencia, la diferencia en el valor de máxima verosimilitud es superior a dos. Debido a que esta diferencia es mayor que dos, la inclusión de estas variables aporta un mejor ajuste del modelo a los datos. Sin embargo, al introducir la variable peso total recogido elevada a la segunda potencia, el valor de máxima verosimilitud no sufre modificaciones, por lo que no se incluye esta variable en el modelo. En definitiva, el modelo de costes de recogida para este tipo de vehículo incluye las variables anteriormente mencionadas con los siguientes parámetros de regresión: 24,417, 0,005 y 0,003, con un valor umbral igual a 0,007.

2.1.3.4.2. Adecuación del modelo

Como los errores se distribuyen según una distribución de probabilidad doble exponencial, si éstos son independientes y están idénticamente distribuidos, las inferencias bajo la teoría de la normalidad son asintóticamente correctas. Por tanto, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. Este análisis se puede ver en la figura 42.

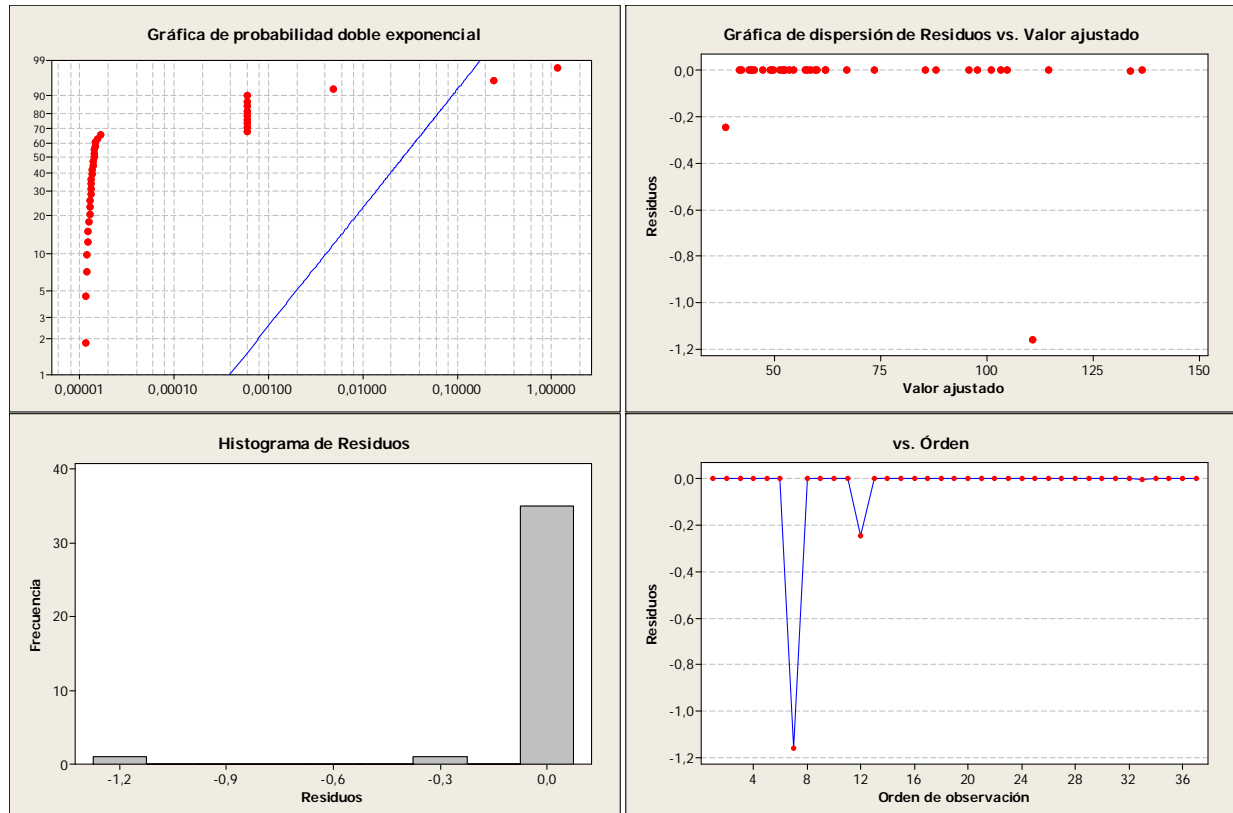


Figura 42: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo de más de 3500 kg. PMA código postal 08198

En la primera gráfica se puede observar como los errores no caen sobre una línea recta, tal y como ocurría en el resto de casos vistos hasta este momento. Esto es debido a que de los 37 residuos obtenidos en el ajuste del modelo, 35 su valor es prácticamente cero. En la segunda, en la que aparecen los residuos frente a sus ajustados, no se aprecia que los residuos se distribuyen de manera aleatoria en todo el eje y, por el mismo motivo. De los 37 residuos calculados en el ajuste del modelo, 35 tienen prácticamente valor cero. En la tercera, se puede observar un histograma de la distribución de los residuales, y como solo hay dos residuos cuyo valor es distinto de cero. Y por último, en la cuarta gráfica, como en el resto, al tener 35 residuos con valor cero, la gráfica es un poco diferente a las vistas hasta ahora para comprobar la independencia de los residuos.

2.1.3.4.3. Validación del modelo de costes de recogida

Para comprobar que el modelo predice de manera adecuada los valores de los costes de reparto ante nuevos valores de las variables exógenas, se realiza la prueba *t* pareada o prueba de Welch, comparando los valores dados por el modelo al introducirle los valores de variables exógenas de la muestra de predicción, y comparándolos con los valores de la variable endógena de la muestra de predicción. Siendo el cuadro en el que se resume esta validación la tabla 89.

Tabla 89: Resultado prueba Welch para el modelo de costes de recogida del vehículo de más de 3500 kg. PMA para el código postal 08198

	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	59,43	59,43	0,001	Límite superior I.C	$t_{0,025;35}$
Varianza	1534,28	1534,44	0,00	0,002	1,96
Varianza/n			0,000		
Máximo	136,19	136,19		Límite inferior I.C	
Mínimo	0,01	0,00		0,001	

En él se puede observar como en la prueba de Welch, el intervalo de confianza no incluye el cero, pero ambos valores del intervalo de confianza están tan próximos a cero, que prácticamente tienen valor cero.

Se puede comprobar cómo en 30 de las 38 ocasiones en las que se compara el valor estimado por el modelo con el valor de la muestra de predicción, la diferencia es cero, siendo en las otras ocho, 0,01. Por tanto, a pesar de que el intervalo no incluye el cero, se considera que el modelo obtenido es válido, debido al grado de exactitud con la que es capaz de predecir los costes de recogida.

Los datos a partir de los que se ha calculado esta prueba *t* pareada provienen de la tabla 90.

Tabla 90: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de más de 3500 kg. para el código postal 08198

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
0,007	24,417	0,005	0,003			
	0	0	0	0,01	0,00	0,01
	2	8690	4	91,23	91,23	0,00
	1	5576	1	51,62	51,62	0,00
	1	11661	1	81,29	81,29	0,00
	1	7641	1	61,69	61,69	0,00
	1	7449	1	60,75	60,75	0,00
	0	0	0	0,01	0,00	0,01
	1	4281	1	45,30	45,30	0,00
	0	0	0	0,01	0,00	0,01
	0	0	0	0,01	0,00	0,01
	2	10705	4	101,05	101,05	0,00
	0	0	0	0,01	0,00	0,01
	0	0	0	0,01	0,00	0,01
	1	6222	1	54,77	54,77	0,00
	1	4910	1	48,37	48,37	0,00
	1	4089	1	44,37	44,37	0,00
	1	3352	1	40,77	40,77	0,00
	0	0	0	0,01	0,00	0,01
	1	4369	1	45,73	45,73	0,00
	2	11351	4	104,20	104,20	0,00
	3	11151	9	127,66	127,65	0,00
	1	5091	1	49,25	49,25	0,00
	2	17911	4	136,19	136,19	0,00
	1	6269	1	55,00	55,00	0,00
	2	7674	4	86,27	86,27	0,00
	1	7441	1	60,71	60,71	0,00
	2	15053	4	122,25	122,25	0,00
	1	9367	1	70,10	70,10	0,00
	1	9785	1	72,14	72,14	0,00
	0	0	0	0,01	0,00	0,01
	1	9050	1	68,56	68,56	0,00
	1	13427	1	89,90	89,90	0,00
	1	9566	1	71,07	71,07	0,00
	1	13967	1	92,53	92,53	0,00
	2	12861	4	111,57	111,57	0,00
	1	8616	1	66,44	66,44	0,00
	1	5323	1	50,38	50,38	0,00
	2	9876	4	97,01	97,01	0,00

2.1.3.5. Modelo de carga de trabajo

En este modelo se va a identificar que distribución de probabilidad siguen las variables aleatorias con las que se puede modelar la carga de trabajo, así como los parámetros que las caracterizan. Estas variables son el número de recogidas por día y el peso de cada recogida.

Con la ayuda de Minitab, se obtendrá los parámetros de cada distribución siendo estos parámetros tales que la distribución de probabilidad que caracterizan, es la que más se aproxima a los datos de la muestra para cada una de las variables.

2.1.3.5.1. Número de recogidas por día

Esta variable aleatoria representa el conteo de algún elemento, y por tanto, se considera una variable discreta. Por esta razón, la primera distribución de probabilidad de la cual se comprueba el ajuste es la distribución discreta de Poisson. Al tratarse de una distribución discreta, el test de bondad de ajuste a realizar es el test Chi cuadrado. Los resultados del test son los siguientes:

Tabla 91: Prueba Chi Cuadrado número de recogidas para el vehículo de más de 3500 kg. PMA código postal 08198

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	alfa=0,05 k=3 Est. par. =1 n=3-1-1=1
0	9	0,196	0,3447	16	2,9631	
1	26	0,565	0,3671	17	4,9164	
2-3	11	0,239	0,2882	13	0,3846	
	46				8,26	3,841

El valor límite para aceptar la hipótesis de que la una distribución de Poisson con media 1,07 ajusta adecuadamente la variable número de repartos por día es 3,841. Este valor corresponde a una distribución Chi cuadrado con 1 grado de libertad y nivel de significancia del 5%. Como el resultado del test es 8,26, superior al valor límite, se rechaza la hipótesis de que esta distribución ajusta adecuadamente los datos de la muestra, y hay que buscar otras distribuciones que ajusten los datos.

Las siguientes distribuciones que se comprobarán son, distribuciones continuas, por lo que el test de bondad de ajuste que se emplea para este tipo de distribuciones es el test de Kolmogorov-Smirnov. Los resultados de estos test para las distribuciones normal y logística son los la tabla 92. No se emplea la distribución gamma, Weibull y lognormal como en el resto de casos, ya que su rango de aplicación es mayor que cero, y no admite el valor cero.

Tabla 92: Test Kolmogorov-Smirnov número de recogidas para el vehículo de más de 3500 kg. PMA código postal 08198

Normal	Test K - S	Logística	Test K - S
Media	0,297	Lambda	0,291
1,065	2,050	1,047	2,005
Desv. Est.	c _α = 1,358	Theta	c _α = 1,358
0,7119	α = 0,05	0,3945	α = 0,05

En la tabla 92 se puede observar como ninguna de las distribuciones supera satisfactoriamente el test de Kolmogorov-Smirnov. Como no se ha podido determinar que alguna distribución de probabilidad ajuste adecuadamente a los datos de la muestra de la variable número de recogidas por día para este tipo de vehículo, se empleará una distribución de probabilidad empírica para describir el comportamiento de esta variable. Esta distribución empírica se puede ver en la tabla 93.

Tabla 93: Distribución empírica número de recogidas para el vehículo de más de 3500 kg. PMA código postal 08198

Número de recogidas	Observaciones	f(Observ.)	F(Observ.)
0	9	0,196	0,196
1	26	0,565	0,761
2	10	0,217	0,978
3	1	0,022	1,000

46

En la tabla 93 se puede observar la probabilidad de masa de cada uno de los valores que la variable aleatoria puede tomar, así como la probabilidad de masa acumula, $F(Observ.)$. De esta manera, al realizar la simulación se obtendrá un número aleatorio que siga una distribución uniforme, y si el número aleatorio se encuentra entre los valores 0 y 0,196 se asigna a ese caso 0 recogidas. Si el valor del número aleatorio está comprendido entre 0,196 y 0,761, el número de recogidas que asigna es 1. Y así sucesivamente, con todos los números aleatorios creados, por cada una de las simulaciones que se deseen elaborar.

2.1.3.5.2. Peso de cada recogida

Esta variable aleatoria es continua, y se determinan los valores de cada distribución, así como los resultados de los test la bondad de ajuste de las mismas de la misma manera que las variables continuas anteriores. Los resultados son los siguientes.

Tabla 94: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo de más de 3500 kg. PMA código postal 08198

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,159	alfa	0,121	alfa	0,149	Media	0,120	Lambda	0,121
5665	1,133	9,989	0,862	2,969	1,060	8,591	0,853	5410	0,858
Desv. Est.	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	Desv. Est.	$c_{\alpha} = 1,358$	Theta	$c_{\alpha} = 1,358$
1966	$\alpha = 0,05$	567,2	$\alpha = 0,05$	6338	$\alpha = 0,05$	0,3128	$\alpha = 0,05$	1038	$\alpha = 0,05$

Como se puede observar en la tabla 94 sí que es posible ajustar la variable aleatoria peso de cada recogida mediante alguna distribución de probabilidad. Para ser exactos, las cinco distribuciones de probabilidad que se han empleado han superado satisfactoriamente el test de Kolmogorov-Smirnov, y podrían emplearse para modelar esta variable. La que menor discrepancia ha mostrado ha sido la distribución lognormal. Como su rango de aplicación es solamente para valores mayores que cero, también posibilita que no sea necesario el emplear una distribución truncada.

2.1.3.6. Simulación de los costes de recogida, y del carga de trabajo

La carga de trabajo relativa a las recogidas de este tipo de vehículo se puede describir como sigue: el número de recogidas se describe de acuerdo a la distribución de probabilidad empírica expresada mediante la tabla 93 del apartado anterior, y el peso de cada una de las recogida mediante una distribución de probabilidad lognormal con media igual a 8,591 y desviación estándar 0,3128.

Un resumen de la carga de trabajo simulada se expresa en la tabla 95, en el cual se puede observar el promedio de repartos por día simulados, su desviación estándar, el número máximo de repartos a realizar en un día, así como el número mínimo. Y los mismos datos para la variable peso total repartido.

Tabla 95: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo de más de 3500 kg. PMA en el código postal 08198

	Recogidas por día	Peso total recogido por día
Promedio	1	5874
Desv. Estan.	1	4258
Máximo	3	18040
Mínimo	0	0

Para la simulación de los costes de recogida, los resultados que se exponen son idénticos a los casos de los vehículos anteriores. Se exponen los valores medios, varianza y desviación estándar, así como un intervalo de confianza de los límites superior e inferior de los límites de predicción de los 250 días simulados a partir de la carga de trabajo previamente simulada bajo las condiciones expuestas en el párrafo anterior. Los resultados son los de la tabla 96.

Tabla 96: Resumen de la simulación de los costes de recogida del vehículo de más de 3500 kg. de PMA en el código postal 08198

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	53,92	54,76
Varianza	1326	1327
Desv. Estan.	36,42	36,42
I.C	50,11	50,96
	57,72	58,57

En la tabla 96 se puede observar como el límite inferior de los estimados puntuales de los costes de recogida se distribuyen con un valor medio de 53,92 euros por día y una desviación estándar de 36,42. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 54,76 y una desviación estándar de 36,42 euros por día.

2.2. Código postal 08197

En este segundo código postal analizado, no se dan repartos superiores a 300 kg. de peso, ni recogidas superiores a 700 kg., por lo que en él solo existe volumen de trabajo para vehículos de 1500 kg. de peso máximo autorizado.

Los pasos a seguir en la elaboración del modelo de costes de la red de distribución capilar de este código postal son análogos a los empleados en el código postal anterior. Tal y como se llevó a cabo en los tres tipos de vehículos del código postal anterior, el primer paso en la elaboración del modelo es la toma de muestras.

2.2.1. Toma de muestras

Se toman muestras de aquellas variables que se identificaron como las determinantes para el buen funcionamiento de la red de distribución capilar. Estas son el coste de reparto por día y el coste de recogida por día. Éstas a su vez dependen del número de repartos por día y del peso repartido por día, y del número de recogidas por día y del peso total repartido por día respectivamente. Esta toma de muestras se ha llevado a cabo mediante un estudio observacional, y se han obtenido 84 observaciones para cada una de ellas. Las muestras obtenidas para las variables relacionadas con los repartos se pueden observar en la tabla 97, mientras que en la tabla 98 se pueden ver las relacionadas con las recogidas.

Tabla 97: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08197

Día	REPARTOS			Día	REPARTOS		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	5	28	11	43	5	208	18
2	4	222	16	44	9	279	27
3	4	63	11	45	2	149	9
4	7	193	22	46	8	347	25
5	4	174	14	47	3	55	8
6	2	102	7	48	7	271	24
7	3	199	11	49	8	110	21
8	4	195	15	50	4	39	10
9	5	144	15	51	5	172	16
10	5	141	16	52	7	272	21
11	6	83	16	53	16	200	39
12	6	37	11	54	5	275	18
13	5	244	17	55	7	299	23
14	9	213	24	56	6	204	18
15	3	10	6	57	4	190	13
16	7	132	19	58	10	124	25
17	1	40	4	59	5	146	15
18	2	17	5	60	5	186	17
19	9	112	23	61	4	121	11
20	8	461	30	62	2	110	7
21	6	108	17	63	4	191	14
22	5	405	22	64	8	270	23
23	5	286	18	65	4	204	16
24	5	110	14	66	2	127	9
25	11	414	35	67	3	35	8
26	4	125	12	68	11	284	34
27	1	64	4	69	3	14	6
28	7	303	24	70	8	272	23
29	10	544	35	71	2	4	4
30	10	434	32	72	4	166	14
31	5	114	13	73	3	153	10
32	4	136	12	74	8	242	22
33	7	296	25	75	6	105	17
34	9	286	27	76	5	119	15
35	1	37	4	77	4	26	9
36	14	325	40	78	9	448	30
37	4	90	11	79	9	157	25
38	4	130	13	80	4	38	9
39	9	349	31	81	6	119	15
40	4	137	13	82	6	160	18
41	9	268	27	83	7	199	21
42	5	165	16	84	3	129	11

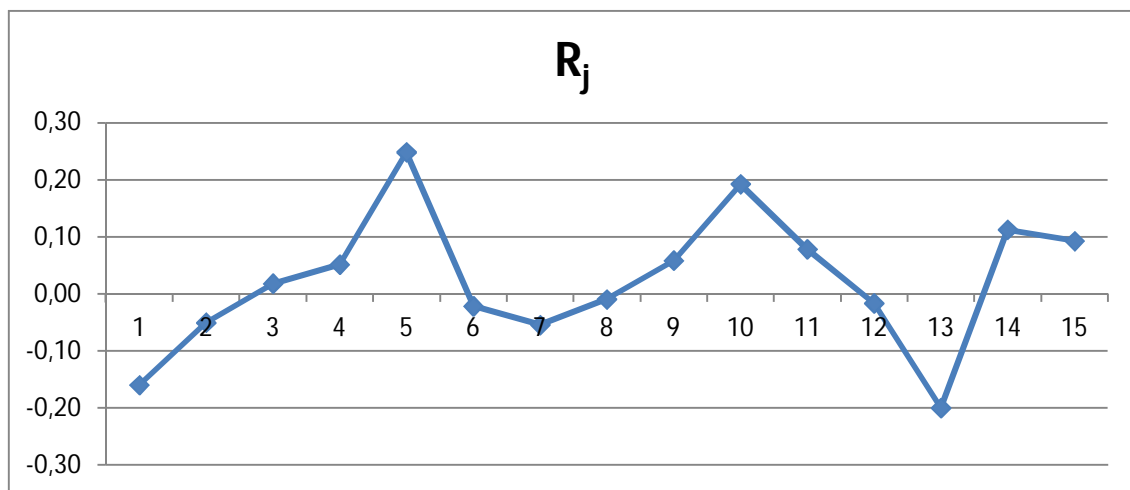
Tabla 98: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08197

Día	RECOGIDAS			Día	RECOGIDAS		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	4	41	11	43	2	140	10
2	4	79	11	44	0	0	0
3	2	9	4	45	2	25	6
4	2	15	5	46	4	59	11
5	4	254	18	47	4	60	13
6	2	110	9	48	2	22	5
7	2	29	6	49	3	57	10
8	3	149	13	50	2	35	6
9	2	145	10	51	2	29	6
10	3	188	14	52	1	3	2
11	2	137	8	53	2	45	7
12	3	41	9	54	3	25	8
13	4	94	14	55	1	9	3
14	4	416	18	56	1	33	4
15	1	89	5	57	2	81	9
16	3	38	8	58	3	48	9
17	3	72	10	59	3	57	10
18	1	60	5	60	2	292	12
19	3	34	8	61	5	348	21
20	3	109	10	62	2	133	8
21	3	30	8	63	2	49	8
22	1	18	3	64	2	230	11
23	1	340	9	65	3	55	9
24	4	182	16	66	4	49	11
25	2	101	8	67	2	66	7
26	2	138	8	68	1	32	4
27	3	32	8	69	5	67	13
28	2	69	8	70	2	83	8
29	8	481	33	71	4	70	12
30	3	110	11	72	1	30	4
31	2	51	6	73	3	38	8
32	2	255	10	74	2	572	16
33	4	23	10	75	2	314	13
34	5	210	19	76	4	378	20
35	1	40	4	77	4	297	18
36	2	305	12	78	5	64	13
37	4	411	21	79	1	2	2
38	3	21	7	80	1	57	5
39	3	206	14	81	1	20	3
40	4	485	21	82	4	78	13
41	3	60	9	83	3	142	13
42	3	26	7	84	2	33	7

2.2.2. Independencia de la muestras

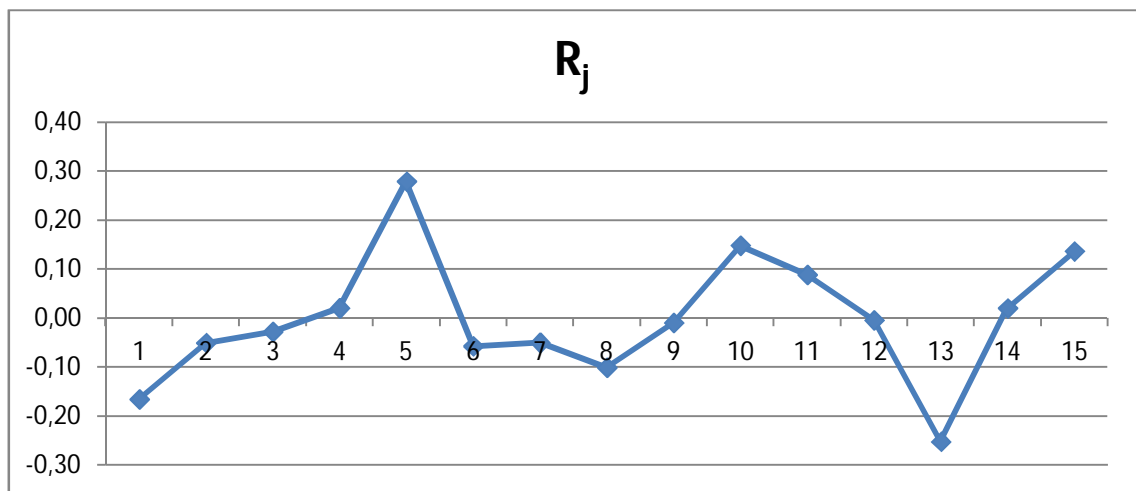
Tras la obtención de las muestras, hay que comprobar si las observaciones de las mismas son independientes. Que las observaciones sean independientes es fundamental, ya que para emplear el método de máxima verosimilitud en la elaboración de los modelos de costes de reparto y recogida, las muestras utilizadas han de ser independientes. Por otro lado, los test no paramétricos de bondad de ajuste de las distribuciones de probabilidad también presuponen que las muestras empleadas son independientes.

En primer lugar se analiza la independencia de la muestras de los repartos. Como esta muestra se compone de tres variables, coste de reparto por día, número de repartos por día y peso total repartido por día, es necesario comprobar la independencia de cada una de ellas. En primer lugar se muestran las gráficas de autocorrelación. La Figura 43 corresponde a la gráfica de autocorrelación de la muestra de los costes de reparto por día. La Figura 44 corresponde a la gráfica de autocorrelación del número de repartos por día, y la Figura 45 corresponde a la gráfica de autocorrelación del peso total repartido por día.



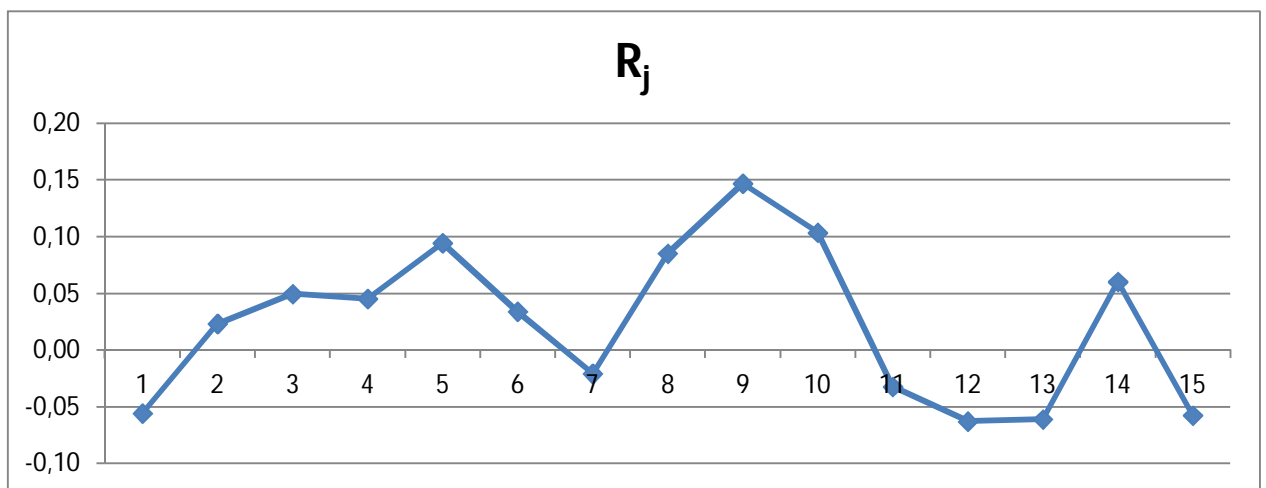
R_j	-0,16	-0,05	0,02	0,05	0,25	-0,02	-0,05	-0,01	0,06	0,19	0,08	-0,02	-0,20	0,11	0,09
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_k	-11	-4	1	4	18	-2	-4	-1	4	14	6	-1	-14	8	7

Figura 43: Gráfica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08197



R_j	-0,17	-0,05	-0,03	0,02	0,28	-0,06	-0,05	-0,10	-0,01	0,15	0,09	0,00	-0,25	0,02	0,14
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-1,4	-0,4	-0,2	0,2	2,3	-0,5	-0,4	-0,8	-0,1	1,2	0,7	0,0	-2,1	0,2	1,1

Figura 44: Gráfica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08197



R_j	-0,06	0,02	0,05	0,05	0,09	0,03	-0,02	0,09	0,15	0,10	-0,03	-0,06	-0,06	0,06	-0,06
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-740	311	659	602	1256	452	-273	1134	1951	1377	-428	-834	-809	802	-763

Figura 45: Gráfica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08197

En las tres gráficas se puede observar como en ninguna de ellas los valores de R_j se encuentran en el rango de -1 y 1, y que en ninguna de ellas se puede apreciar un patrón determinado. Por tanto, esta prueba indica la independencia de las observaciones de las muestras.

Tras las gráficas de autocorrelación, se muestran a continuación las gráficas de dispersión para cada una de las variables.

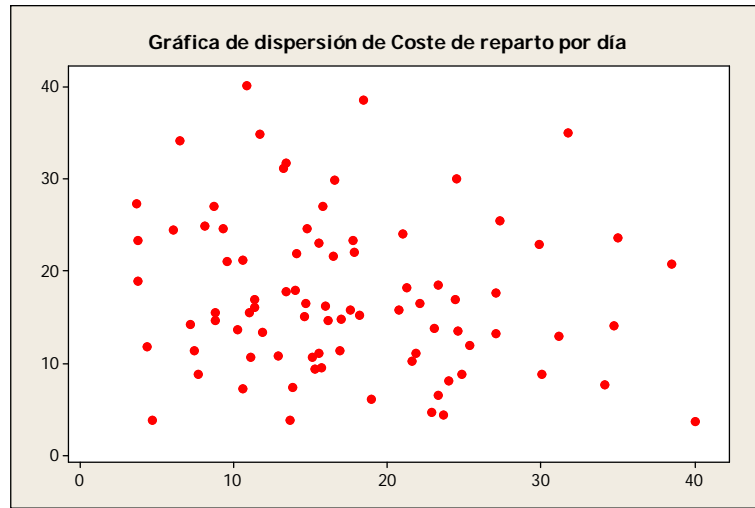


Figura 46: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08197

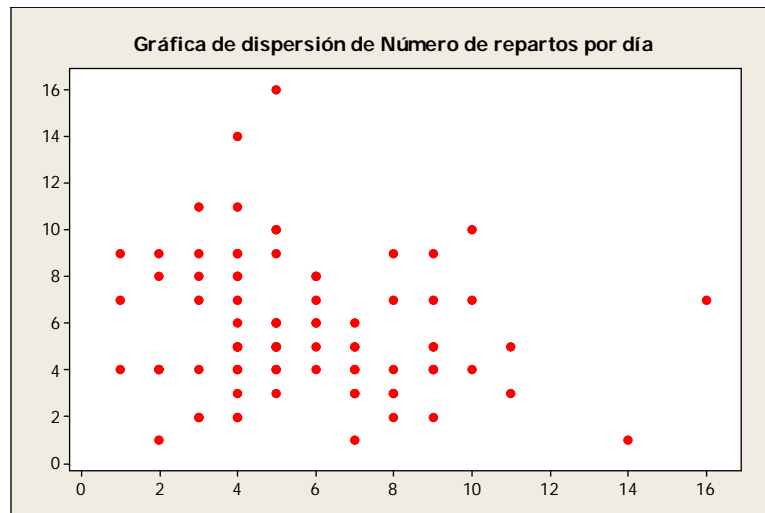


Figura 47: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08197

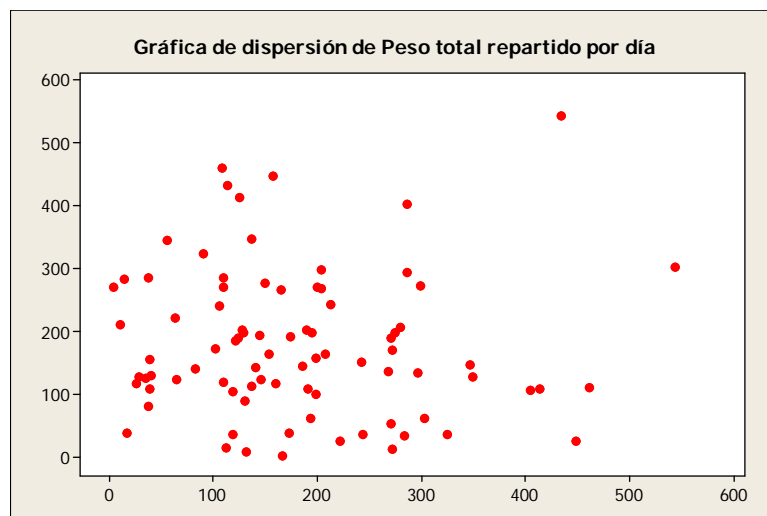
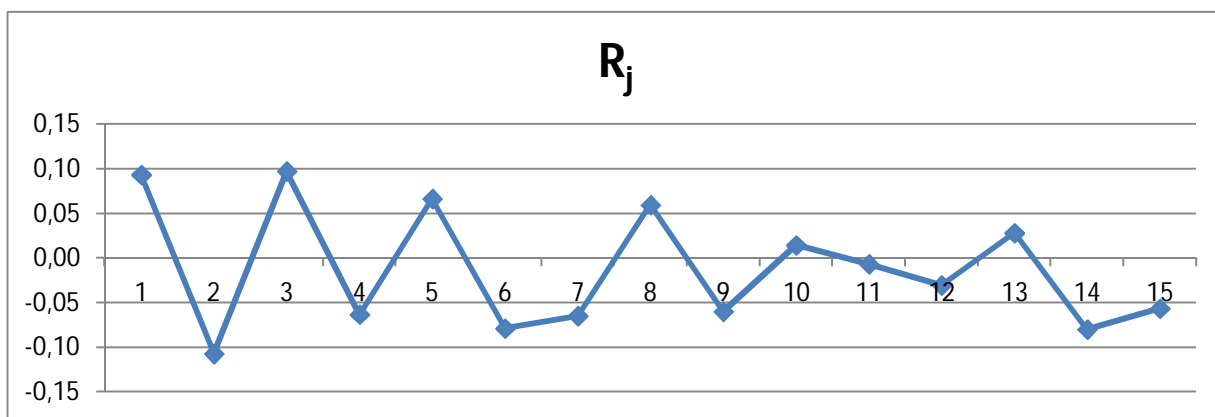


Figura 48: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08197

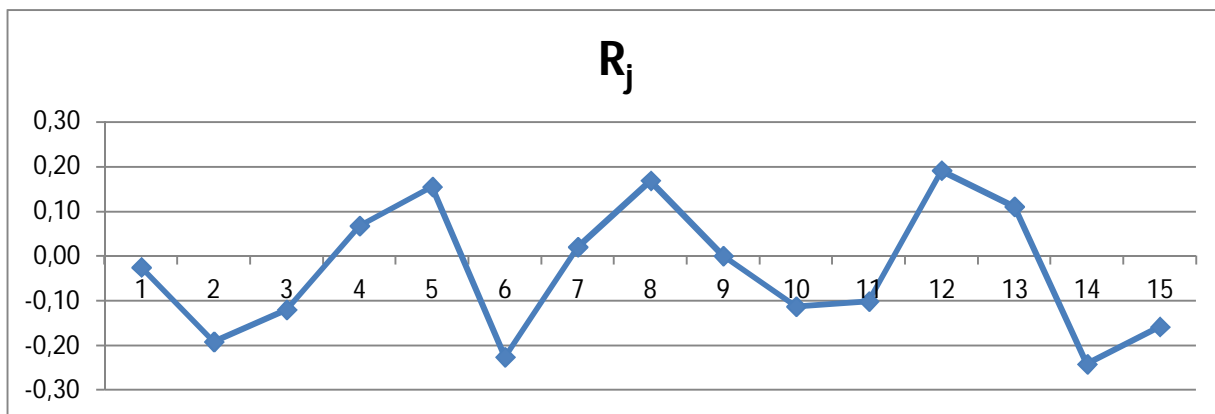
Como en las tres gráficas los puntos se distribuyen de manera aleatoria a lo largo de todo el plano, esta prueba también concluye que las observaciones son independientes.

Tras comprobar que las muestras relativas a los repartos son independientes, hay que analizar si las muestras correspondientes a las recogidas también lo son. Al igual que para los repartos se comenzará mostrando las gráficas de autocorrelación.



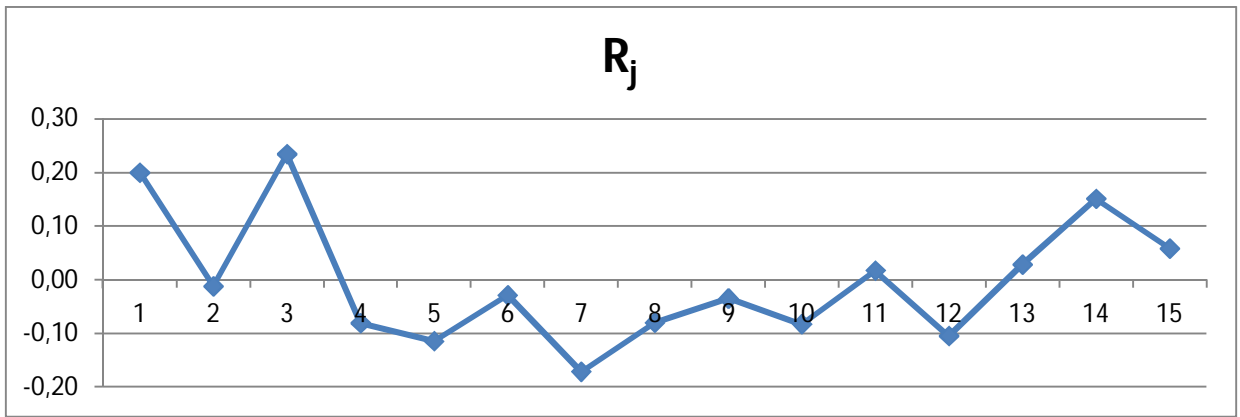
R_j	0,09	-0,11	0,10	-0,06	0,07	-0,08	-0,06	0,06	-0,06	0,01	-0,01	-0,03	0,03	-0,08	-0,06
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	3	-3	3	-2	2	-2	-2	2	-2	0	0	-1	1	-2	-2

Figura 49: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08197



R_j	-0,03	-0,19	-0,12	0,07	0,15	-0,23	0,02	0,17	0,00	-0,11	-0,10	0,19	0,11	-0,24	-0,16
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	0,0	-0,3	-0,2	0,1	0,3	-0,4	0,0	0,3	0,0	-0,2	-0,2	0,3	0,2	-0,4	-0,3

Figura 50: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08197



R_j	0,20	-0,01	0,23	-0,08	-0,11	-0,03	-0,17	-0,08	-0,03	-0,08	0,02	-0,10	0,03	0,15	0,06
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	3271	-196	3838	-1325	-1874	-470	-2807	-1310	-570	-1356	283	-1713	467	2476	957

Figura 51: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08197

En las tres gráficas se puede ver como los valores de R_j no superan los valores 1 y -1, y que no se da un patrón determinado en ninguna de ellas por lo que se concluye que las observaciones de cada muestra son independientes. Tras estudiar las gráficas de autocorrelación, a continuación se verifica la independencia de las observaciones mediante las gráficas de dispersión.

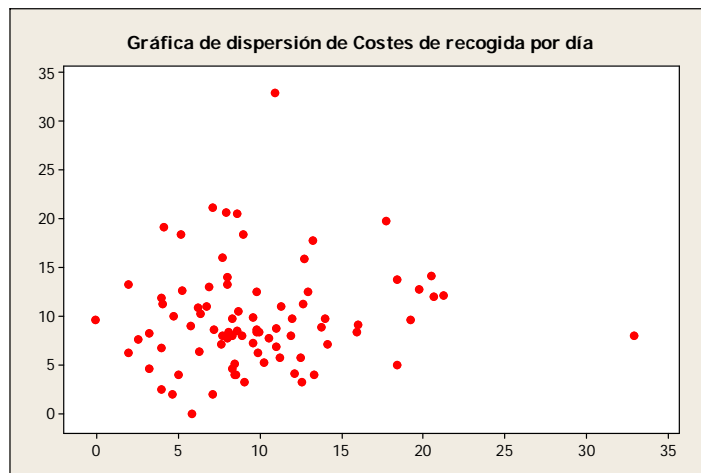


Figura 52: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08197

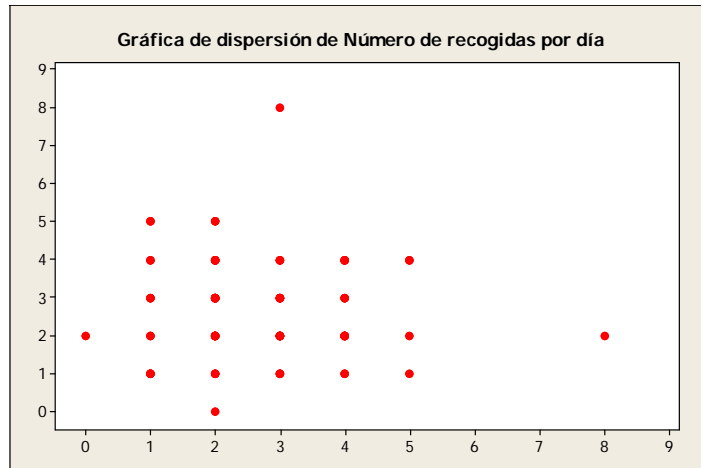


Figura 53: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08197

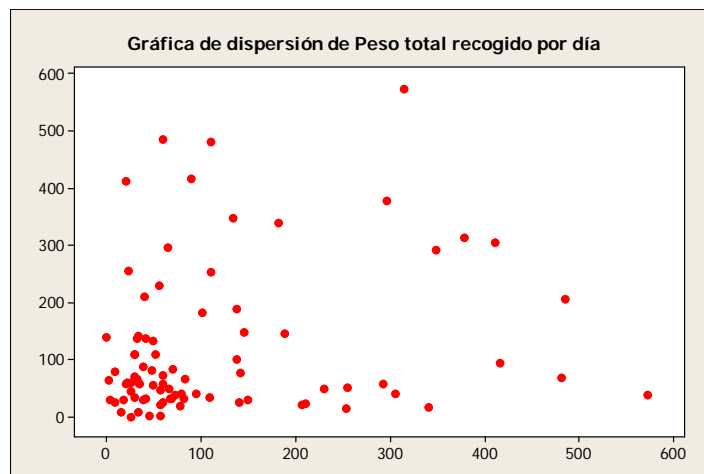


Figura 54: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08197

En las tres gráficas se puede observar como todos los puntos se distribuyen de manera aleatoria en los planos correspondientes, pudiéndose concluir que las observaciones de las muestras son independientes.

2.2.3. Separación de las muestras de calibración y predicción

Tras verificar que las 84 observaciones obtenidas del estudio observacional de cada una de las variables son independientes, hay que separar estas 84 observaciones en dos muestras diferentes, la de calibración y la de predicción. La de calibración se emplea para la creación y elaboración de los modelos necesarios para el estudio de los costes de la red de distribución, mientras que la de predicción se emplea para comprobar que los modelo elaborados mediante el análisis de regresión estiman los costes de reparto y recogida de manera adecuada. Esta separación se lleva a cabo mediante una asignación de un número aleatorio a cada una de las observaciones, y si éste es mayor o igual que el valor 0,5, se asigna a la

muestra de calibración. Si por el contrario, este número aleatorio es menor que 0,5, se asigna a la muestra de predicción. El listado de números aleatorios asignados a cada una de las observaciones relativas al reparto se puede ver en la tabla 99.

Tabla 99: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08197

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,774	22	0,452	43	0,064	64	0,405
2	0,184	23	0,178	44	0,358	65	0,552
3	0,150	24	0,840	45	0,487	66	0,712
4	0,754	25	0,426	46	0,511	67	0,555
5	0,610	26	0,304	47	0,373	68	0,181
6	0,571	27	0,976	48	0,986	69	0,970
7	0,055	28	0,807	49	0,041	70	0,687
8	0,398	29	0,991	50	0,231	71	0,529
9	0,507	30	0,256	51	0,005	72	0,797
10	0,232	31	0,952	52	0,926	73	0,806
11	0,076	32	0,053	53	0,100	74	0,262
12	0,033	33	0,705	54	0,257	75	0,178
13	0,023	34	0,817	55	0,776	76	0,867
14	0,211	35	0,973	56	0,680	77	0,115
15	0,786	36	0,466	57	0,809	78	0,060
16	0,447	37	0,300	58	0,724	79	0,762
17	0,858	38	0,750	59	0,085	80	0,738
18	0,117	39	0,351	60	0,132	81	0,986
19	0,451	40	0,776	61	0,756	82	0,926
20	0,783	41	0,074	62	0,627	83	0,904
21	0,434	42	0,198	63	0,174	84	0,545

Quedando las muestras de calibración y predicción de acuerdo a esta asignación como se muestra en la tabla 100.

Tabla 100: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08197

Día	REPARTOS. MUESTRA DE CALIBRACIÓN			Día	REPARTOS. MUESTRA DE PREDICCIÓN		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	5	28	11	1	4	222	16
2	7	193	22	2	4	63	11
3	4	174	14	3	3	199	11
4	2	102	7	4	4	195	15
5	5	144	15	5	5	141	16
6	3	10	6	6	6	83	16
7	1	40	4	7	6	37	11
8	8	461	30	8	5	244	17
9	5	110	14	9	9	213	24
10	1	64	4	10	7	132	19
11	7	303	24	11	2	17	5
12	10	544	35	12	9	112	23
13	5	114	13	13	6	108	17
14	7	296	25	14	5	405	22
15	9	286	27	15	5	286	18
16	1	37	4	16	11	414	35
17	4	130	13	17	4	125	12
18	4	137	13	18	10	434	32
19	8	347	25	19	4	136	12
20	7	271	24	20	14	325	40
21	7	272	21	21	4	90	11
22	7	299	23	22	9	349	31
23	6	204	18	23	9	268	27
24	4	190	13	24	5	165	16
25	10	124	25	25	5	208	18
26	4	121	11	26	9	279	27
27	2	110	7	27	2	149	9
28	4	204	16	28	3	55	8
29	2	127	9	29	8	110	21
30	3	35	8	30	4	39	10
31	3	14	6	31	5	172	16
32	8	272	23	32	16	200	39
33	2	4	4	33	5	275	18
34	4	166	14	34	5	146	15
35	3	153	10	35	5	186	17
36	5	119	15	36	4	191	14
37	9	157	25	37	8	270	23
38	4	38	9	38	11	284	34
39	6	119	15	39	8	242	22
40	6	160	18	40	6	105	17
41	7	199	21	41	4	26	9
42	3	129	11	42	9	448	30

En este caso, debido a la asignación aleatoria, se tienen 42 observaciones para la muestra de calibración y 42 observaciones para la muestra de predicción relativas a los repartos. Una vez elaboradas estas muestras para los repartos, se realiza un proceso análogo para las recogidas. La asignación de números aleatorios a cada una de las 84 observaciones relativas a las recogidas es la de la tabla 101.

Tabla 101: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08197

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,057	22	0,218	43	0,748	64	0,666
2	0,515	23	0,221	44	0,413	65	0,641
3	0,875	24	0,168	45	0,821	66	0,434
4	0,905	25	0,804	46	0,931	67	0,185
5	0,420	26	0,023	47	0,989	68	0,500
6	0,238	27	0,015	48	0,046	69	0,059
7	0,513	28	0,862	49	0,883	70	0,564
8	0,490	29	0,206	50	0,488	71	0,352
9	0,986	30	0,435	51	0,021	72	0,656
10	0,044	31	0,285	52	0,708	73	0,363
11	0,551	32	0,939	53	0,147	74	0,430
12	0,566	33	0,347	54	0,883	75	0,217
13	0,542	34	0,424	55	0,133	76	0,466
14	0,665	35	0,621	56	0,285	77	0,032
15	0,336	36	0,800	57	0,186	78	0,897
16	0,332	37	0,447	58	0,033	79	0,080
17	0,055	38	0,286	59	0,252	80	0,291
18	0,089	39	0,923	60	0,373	81	0,552
19	0,624	40	0,177	61	0,691	82	0,768
20	0,425	41	0,292	62	0,404	83	0,765
21	0,671	42	0,403	63	0,412	84	0,894

A partir de este listado de números aleatorios se confeccionan las muestras de calibración y predicción relativas a las recogidas. Estas muestras se pueden observar en la tabla 102.

Tabla 102: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08197

Día	RECOGIDAS. MUESTRA DE CALIBRACIÓN			Día	RECOGIDAS. MUESTRA DE PREDICCIÓN		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	4	79	11	1	4	41	11
2	2	9	4	2	4	254	18
3	2	15	5	3	2	110	9
4	2	29	6	4	3	149	13
5	2	145	10	5	3	188	14
6	2	137	8	6	1	89	5
7	3	41	9	7	3	38	8
8	4	94	14	8	3	72	10
9	4	416	18	9	1	60	5
10	3	34	8	10	3	109	10
11	3	30	8	11	1	18	3
12	2	101	8	12	1	340	9
13	2	69	8	13	4	182	16
14	2	255	10	14	2	138	8
15	1	40	4	15	3	32	8
16	2	305	12	16	8	481	33
17	3	206	14	17	3	110	11
18	2	140	10	18	2	51	6
19	2	25	6	19	4	23	10
20	4	59	11	20	5	210	19
21	4	60	13	21	4	411	21
22	3	57	10	22	3	21	7
23	1	3	2	23	4	485	21
24	3	25	8	24	3	60	9
25	5	348	21	25	3	26	7
26	2	230	11	26	0	0	0
27	3	55	9	27	2	22	5
28	2	83	8	28	2	35	6
29	1	30	4	29	2	29	6
30	5	64	13	30	2	45	7
31	1	20	3	31	1	9	3
32	4	78	13	32	1	33	4
33	3	142	13	33	2	81	9
34	2	33	7	34	3	48	9
				35	3	57	10
				36	2	292	12
				37	2	133	8
				38	2	49	8
				39	4	49	11
				40	2	66	7
				41	1	32	4
				42	5	67	13
				43	4	70	12
				44	3	38	8
				45	2	572	16
				46	2	314	13
				47	4	378	20
				48	4	297	18
				49	1	2	2
				50	1	57	5

2.2.4. Modelo de costes de reparto

Como para el resto de modelos de costes de reparto elaborados de este trabajo de investigación, los parámetros de regresión se calculan empleando el método de máxima verosimilitud. Se estudiará el valor de máxima verosimilitud obtenido bajo diferentes premisas, como diferentes distribuciones de probabilidad, ó transformaciones de las variables exógenas. Los parámetros de regresión del modelo que mayor valor de verosimilitud proporcionen, serán los que se empleen en la simulación del modelo para estimar los costes de reparto. Los supuestos que se van a analizar son los mismos que los modelos de reparto anteriores, y el estudio de máxima verosimilitud completo para cada uno de estos supuestos se exponen en las siguientes tablas.

Tabla 103: VMV modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
15,554					-145,9		0%	7,89
0,215	3,039				-95,2	50,633	91%	2,39
0,439	2,087	0,027			-53,3	41,981	99%	0,89
0,091	2,250	0,028	-0,016		-53,0	0,282	99%	0,90
-0,044	2,053	0,035		-0,00002	-50,2	2,775	99%	0,84

Tabla 104: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,744					-145,9		0%	7,89
1,797	0,172				-104,1	41,815	86%	2,95
1,820	0,136	0,001			-91,9	12,161	92%	2,24
1,173	0,365	0,001	-0,018		-70,6	21,342	97%	1,36
1,192	0,305	0,002	-0,013	0	-68,7	1,848	97%	1,32

Tabla 105: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,064					-145,9		0%	7,89
0,111	-0,008				-115,7	30,218	76%	3,89
0,110	-0,007	0,000			-111,4	4,219	80%	3,57
0,190	-0,032	0,000	0,002		-82,3	29,106	95%	1,81
0,190	-0,032	0,000	0,002	0	-82,3	0,000	95%	1,83

Tabla 106: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-3,944					-145,9		0%	7,89
-1,951	-0,375				-97,8	48,072	90%	2,54
-2,000	-0,277	-0,003			-74,9	22,926	96%	1,49
-1,395	-0,517	-0,003	0,020		-62,4	12,480	98%	1,12
-1,482	-0,393	-0,005	0,011	0,000005	-55,6	6,796	98%	0,97

Tabla 107: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
14,099					-148,8		-3%	8,03
0,861	2,920				-95,6	53,237	91%	2,41
0,537	2,054	0,028			-54,6	41,040	99%	0,90
0,537	2,054	0,028	-0,0026		-54,5	0,035	99%	0,90
0,537	2,054	0,028		0	-54,5	0,000	99%	0,90

Tabla 108: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,646					-148,8		-3%	8,03
1,816	0,174				-103,0	45,802	85%	3,02
1,828	0,129	0,001			-92,7	10,364	91%	2,31
1,209	0,361	0,001	-0,017		-73,5	19,156	97%	1,40
1,209	0,361	0,001	-0,017	0	-73,5	0,017	97%	1,42

Tabla 109: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,071					-148,8		-4%	8,04
0,109	-0,008				-116,2	32,634	75%	3,91
0,108	-0,007	-0,00003			-112,7	3,537	77%	3,81
0,108	-0,007	-0,00003	0		-112,7	0,001	76%	3,86
0,108	-0,007	-0,00003		0	-112,7	-0,001	76%	3,86

Tabla 110: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-3,754					-148,8		-3%	8,03
-1,911	-0,384				-97,2	51,660	90%	2,56
-2,034	-0,252	-0,003			-77,4	19,777	96%	1,57
-2,034	-0,252	-0,003	0		-77,4	0,000	96%	1,59
-2,035	-0,252	-0,003		0	-76,7	0,704	96%	1,51

Tabla 111: VMV modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,656					-175,7		-273%	15,26
0,393	3,011				-95,3	80,303	91%	2,40
0,409	2,078	0,028			-53,6	41,783	99%	0,89
0,189	2,190	0,028	-0,011		-53,4	0,145	99%	0,90
-0,086	2,050	0,036		-0,00002	-49,6	3,787	99%	0,84

El supuesto que mayor valor de máxima verosimilitud ha proporcionado ha sido el que considera que los errores se distribuyen de acuerdo a una distribución logística. Por lo tanto, se empleará como modelo de costes de reparto el análisis de regresión bajo el supuesto de que los errores se distribuyen bajo una distribución logística.

2.2.4.1. Selección de variables del modelo de costes de reparto

En la tabla referente al estudio de máxima verosimilitud con los errores bajo la suposición de que se distribuyen de acuerdo a función logística, se puede observar como al introducir las variables exógenas número de repartos y peso total repartido por día, el cociente de verosimilitud es mayor que dos unidades cuando se introducen cada una de estas variables de una en una en el modelo. Sin embargo, al introducir la variable exógena número de repartos elevada a la segunda potencia, el cociente de verosimilitud es de 0,145 unidades, por lo que la adición de esta variable no aporta un mejor ajuste del modelo. Por último, al calcular el valor de máxima verosimilitud con las dos primeras variables exógenas más la variable peso total repartido elevada al cuadrado, el cociente de verosimilitud entre el modelo con dos variables y el modelo con estas tres variables es casi de 4 unidades, por lo que esta variable sí que aporta un mejor ajuste del modelo, y se empleará en el modelo de costes de reparto que se empleará posteriormente en las simulación. Al calcular el valor de máxima verosimilitud con estas tres variables, no se incluyeron los valores de la variable número de repartos elevado al cuadrado.

En definitiva, el modelo de costes de reparto para el código postal 08197 constará de las variables número de repartos por día, peso total repartido por día y peso total repartido por día elevado al cuadrado, con los siguientes parámetros de regresión respectivamente, 2,05, 0,036 y -0,00002, con un valor umbral igual a -0,086.

2.2.4.2. Adecuación del modelo de costes de reparto

Cuando los errores se distribuyen con una distribución de probabilidad no normal, si éstos son independientes y se distribuyen idénticamente, las inferencias bajo la teoría de normalidad son asintóticamente correctas. Por tanto, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. El análisis gráfico de los errores para el modelo de costes de reparto del vehículo de 1500 kg. de peso máximo autorizado en el código postal 08197 se puede ver en la figura 55.

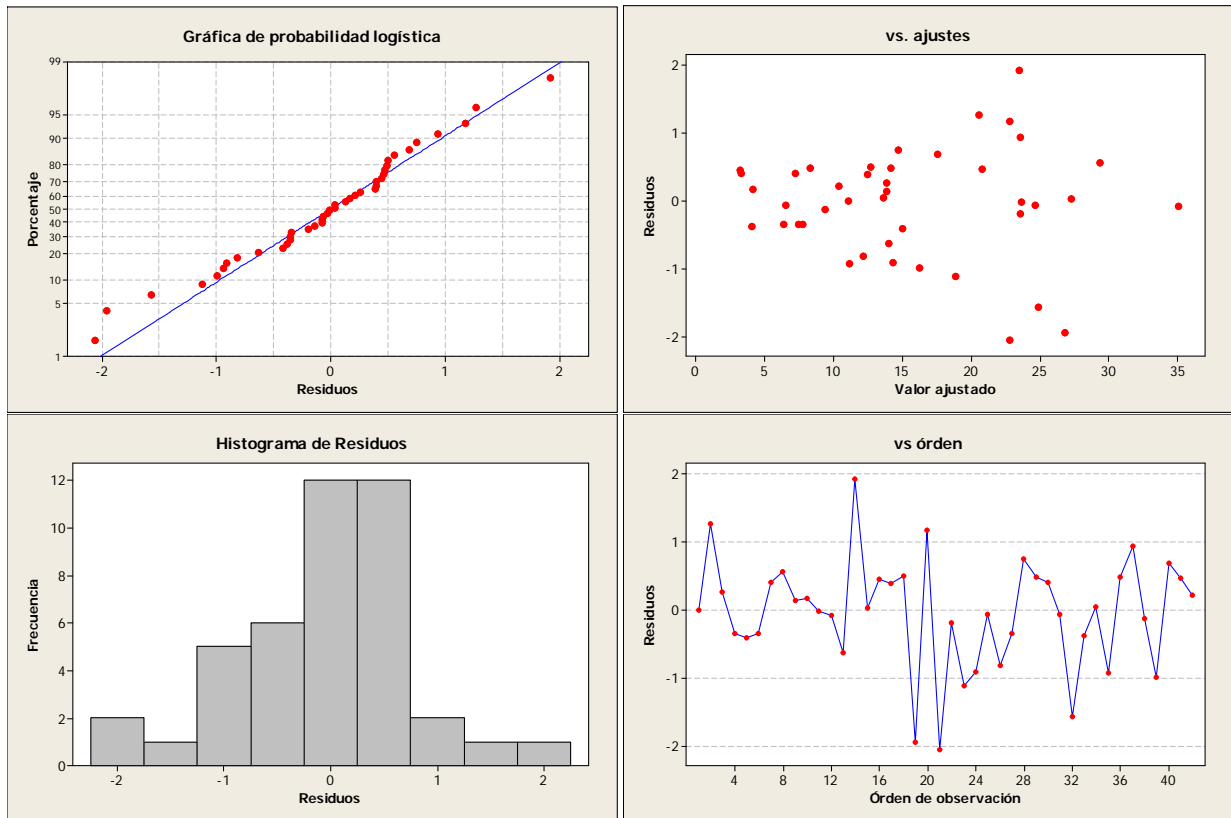


Figura 55: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08197

En la primera gráfica se puede observar como los errores se distribuyen de acuerdo a distribución logística ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.2.4.3. Validación del modelo de costes de reparto

Para comprobar que el modelo predice de manera adecuada los valores de los costes de reparto ante nuevos valores de las variables exógenas, se realiza la prueba t pareada o prueba de Welch. Los nuevos valores de las variables exógenas son los valores de la muestra de predicción, y la validación del modelo se lleva a cabo comparando uno a uno los valores de la variable endógena estimados por el modelo tras introducir los nuevos valores de las variables exógenas, con los valores correspondientes de los costes de reparto de la muestra de predicción. Los valores Z que se emplean para realizar en esta prueba de Welch se obtienen de la tabla 112.

Tabla 112: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_4$	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
-0,086	2,050	0,036	-0,00002			
	4	222	49284	15,31	15,54	-0,24
	4	63	3969	10,33	11,08	-0,75
	3	199	39601	12,59	10,61	1,98
	4	195	38025	14,52	15,11	-0,59
	5	141	19881	14,93	16,17	-1,24
	6	83	6889	15,10	16,03	-0,93
	6	37	1369	13,53	11,37	2,16
	5	244	59536	17,98	16,96	1,02
	9	213	45369	25,30	24,45	0,85
	7	132	17424	18,74	18,94	-0,20
	2	17	289	4,62	4,71	-0,09
	9	112	12544	22,20	22,88	-0,67
	6	108	11664	15,92	16,56	-0,63
	5	405	164025	22,03	22,11	-0,08
	5	286	81796	19,12	17,90	1,22
	11	414	171396	34,53	34,81	-0,27
	4	125	15625	12,37	11,74	0,63
	10	434	188356	32,92	31,76	1,16
	4	136	18496	12,72	11,93	0,79
	14	325	105625	38,58	40,09	-1,51
	4	90	8100	11,23	10,86	0,37
	9	349	121801	28,92	31,19	-2,26
	9	268	71824	26,84	27,05	-0,22
	5	165	27225	15,67	15,85	-0,18
	5	208	43264	16,95	17,61	-0,66
	9	279	77841	27,13	27,07	0,07
	2	149	22201	9,03	8,76	0,26
	3	55	3025	8,00	8,17	-0,17
	8	110	12100	20,09	21,03	-0,95
	4	39	1521	9,50	9,58	-0,08
	5	172	29584	15,88	15,72	0,16
	16	200	40000	39,27	38,51	0,76
	5	275	75625	18,83	18,50	0,33
	5	146	21316	15,08	14,81	0,27
	5	186	34596	16,30	17,00	-0,69
	4	191	36481	14,40	13,83	0,57
	8	270	72900	24,84	23,11	1,73
	11	284	80656	31,37	34,15	-2,78
	8	242	58564	24,07	21,61	2,46
	6	105	11025	15,83	16,54	-0,72
	4	26	676	9,04	8,85	0,19
	9	448	200704	31,16	30,08	1,09

Siendo la tabla 113 que resume los resultados de esta prueba de Welch.

Tabla 113: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08197

	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	19,11	19,06	0,051	Límite superior I.C	$t_{0,025;39}$ 1,96
Varianza	71,64	74,67	1,16		
Varianza/n			0,028	Límite inferior I.C	-0,275
Máximo	39,27	40,09			
Mínimo	4,62	4,71			

El intervalo de confianza obtenido para esta prueba contiene el valor cero, por lo que se concluye que el modelo estima nuevos costes de reparto de manera adecuada. Para elaborarlo se ha empleado la distribución *t* de Student, con un nivel de significancia del 95% y 39 grados de libertad, cuyo valor es 1.96.

2.2.5. *Modelo de costes de recogida*

Una vez finalizado el modelo de los costes de reparto, hay que realizar el modelo de costes de recogida. La creación este modelo es análogo al de los costes de reparto, por lo que los pasos a seguir son idénticos.

Tabla 114: VMV modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
9,358					-95,8		0%	4,11
1,247	3,064				-77,0	18,793	66%	2,40
0,563	2,463	0,022			-43,59	33,428	95%	0,91
-0,490	3,313	0,022	-0,147		-42,8	0,834	95%	0,91
-0,156	2,475	0,038		-0,00004	-37,5	5,281	96%	0,78

Tabla 115: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,236					-95,8		0%	4,11
1,422	0,290				-78,2	17,570	63%	2,49
1,439	0,222	0,002			-60,6	17,646	87%	1,51
0,853	0,627	0,002	-0,062		-54,8	5,793	90%	1,29
0,481	0,677	0,006	-0,067	-0,00001	-38,43	16,371	96%	0,79

Tabla 116: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
23,500					-127,1		0%	10,31
0,180	-0,025				-80,1	46,923	59%	2,63
0,171	-0,018	-0,0001			-70,6	9,574	77%	1,99
0,247	-0,066	-0,0001	0,007		-65,0	5,560	82%	1,74
0,247	-0,066	-0,0001	0,007	0	-65,0	0,000	81%	1,77

Tabla 117: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-3,059					-95,8		0%	4,11
-1,753	-0,478				-77,5	18,282	65%	2,44
-1,693	-0,380	-0,003			-52,5	25,062	92%	1,19
-1,058	-0,841	-0,003	0,074		-47,7	4,804	94%	1,05
-0,884	-0,793	-0,007	0,063	0,00001	-37,8	9,859	97%	0,79

Tabla 118: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
8,820					-96,3		-2%	4,15
1,304	2,822				-78,9	17,426	-5%	4,21
0,495	2,381	0,023			-45,5	33,367	-8%	4,28
0,490	2,383	0,023	0,001		-45,5	0,002	-12%	4,35
0,470	2,331	0,029		0	-42,1	3,441	-12%	4,35

Tabla 119: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,175					-96,3		-2%	4,15
1,338	0,300				-80,6	15,711	62%	2,54
1,495	0,198	0,0016			-61,5	19,064	86%	1,57
1,495	0,198	0,0016	0		-61,5	0,000	85%	1,59
1,495	0,198	0,0016		0	-61,5	-0,003	85%	1,59

Tabla 120: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,114					-96,3		-2%	4,16
0,203	-0,031				-81,3	14,970	55%	2,77
0,180	-0,021	-0,0001			-72,4	8,950	74%	2,09
0,178	-0,023	0,000	0,001		-70,1	2,317	77%	1,99
0,178	-0,023	0,000	0,001	0	-70,1	0,000	76%	2,03

Tabla 121: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-2,959					-96,3		-2%	4,16
-1,561	-0,500				-79,9	16,369	61%	2,56
-1,785	-0,335	-0,003			-54,8	25,172	91%	1,25
-1,785	-0,335	-0,003	0		-54,8	0,000	90%	1,27
-1,785	-0,335	-0,003		0	-54,8	0,000	90%	1,27

Tabla 122: VMV modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
9,181					-95,6		0%	4,12
1,358	2,937				-77,9	17,748	65%	2,42
0,596	2,438	0,022			-44,5	33,342	95%	0,91
-0,265	3,137	0,023	-0,122		-43,9	0,596	95%	0,91
-0,169	2,461	0,039		0	-38,1	5,823	96%	0,78

En las tablas se puede observar como el valor de máxima verosimilitud es mayor para la distribución normal. El valor obtenido para las transformaciones de las variables exógenas con función cadena logarítmica y raíz, cuando los errores se distribuyen de manera normal, es muy parecido al valor con la función cadena unidad, pero un poco más alto. Por tanto, para la simulación se empleará el modelo con la suposición de que los errores se distribuyen de manera normal, y sin transformaciones lineales de las variables exógenas.

2.2.5.1. Selección de las variables para el modelo de costes de recogida

Si se observa el cuadro en el que los errores se distribuyen según una distribución de probabilidad normal, se puede observar como al introducir las variables número de recogidas y peso total recogido, el valor de máxima verosimilitud al introducir cada una de estas variables una a una, difiere en más de dos unidades, por lo que el modelo que incluye estas variables ajusta mejor a los datos que aquel que no las tiene. Sin embargo, al introducir la variable número de recogidas por día elevada al cuadrado el valor de máxima verosimilitud sólo se incrementa en 0,834 unidades, por lo que la adición de esta variable no aporta un mejor ajuste. Por el contrario, al calcular el valor de máxima verosimilitud con la variable peso total recogido a la segunda potencia junto con las dos primeras variables, se puede ver como la diferencia respecto al mismo valor solo con dos variables es de seis unidades, por lo que esta variable será la tercera en el modelo de costes de recogida.

Resumiendo, en este modelo de costes de recogida se incluirán las variables número de recogidas por día, peso total recogido por día, y peso total recogido por día con los parámetros de regresión 3,313, 0,022 y -0,00004 respectivamente.

2.2.5.2. Adecuación del modelo de costes de recogida

Como el mayor valor máximo de verosimilitud es el que aporta una distribución de probabilidad normal, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. Este análisis gráfico se puede ver en la figura 56.

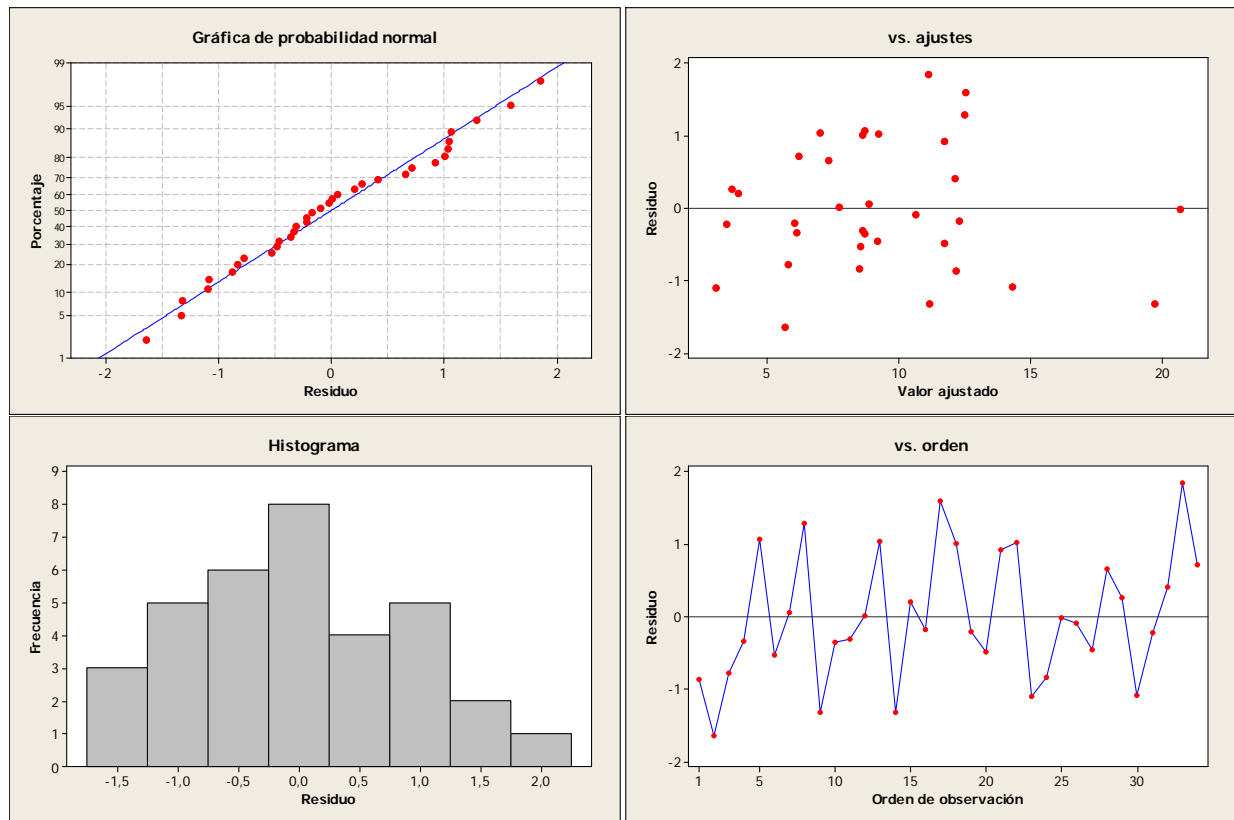


Figura 56: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08197

En la primera gráfica se puede observar como los errores se distribuyen de manera normal ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.2.5.3. Validación del modelo de costes de recogida

Para comprobar que el modelo predice de manera adecuada los valores de los costes de reparto ante nuevos valores de las variables exógenas, se realiza la prueba t pareada o prueba de Welch, con los datos de la muestra de predicción previamente obtenida. Los valores Z con los que se elabora este cuadro provienen de la tabla 123.

Tabla 123: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08197

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_4$	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
-0,156	2,475	0,038	-0,00004			
	4	41	1681	11,22	10,97	0,25
	4	254	64516	16,58	18,38	-1,80
	2	110	12100	8,43	8,98	-0,56
	3	149	22201	11,94	12,54	-0,59
	3	188	35344	12,86	13,97	-1,12
	1	89	7921	5,34	5,16	0,18
	3	38	1444	8,64	8,45	0,19
	3	72	5184	9,76	9,98	-0,21
	1	60	3600	4,43	4,71	-0,29
	3	109	11881	10,87	9,83	1,04
	1	18	324	2,98	3,22	-0,24
	1	340	115600	10,23	9,10	1,14
	4	182	33124	15,20	16,03	-0,83
	2	138	19044	9,19	8,03	1,16
	3	32	1024	8,43	8,35	0,08
	8	481	231361	27,97	32,89	-4,93
	3	110	12100	10,90	10,94	-0,04
	2	51	2601	6,61	6,25	0,35
	4	23	529	10,59	9,59	0,99
	5	210	44100	18,26	19,22	-0,95
	4	411	168921	18,07	21,21	-3,14
	3	21	441	8,04	7,16	0,88
	4	485	235225	18,05	20,53	-2,48
	3	60	3600	9,38	8,61	0,76
	3	26	676	8,22	7,25	0,97
	0	0	0	-0,16	0,00	-0,16
	2	22	484	5,60	5,23	0,37
	2	35	1225	6,06	6,42	-0,36
	2	29	841	5,85	6,32	-0,47
	2	45	2025	6,40	7,14	-0,73
	1	9	81	2,66	2,55	0,10
	1	33	1089	3,52	4,01	-0,50
	2	81	6561	7,57	8,54	-0,97
	3	48	2304	8,98	8,61	0,37
	3	57	3249	9,28	9,83	-0,55
	2	292	85264	12,18	11,98	0,21
	2	133	17689	9,06	7,97	1,09
	2	49	2401	6,54	7,76	-1,22
	4	49	2401	11,49	11,03	0,46
	2	66	4356	7,10	6,75	0,35
	1	32	1024	3,48	4,00	-0,52
	5	67	4489	14,55	13,33	1,23
	4	70	4900	12,17	11,89	0,28
	3	38	1444	8,64	8,45	0,19
	2	572	327184	12,49	15,93	-3,45
	2	314	98596	12,45	12,73	-0,28
	4	378	142884	17,93	19,76	-1,83
	4	297	88209	17,20	17,75	-0,55
	1	2	4	2,39	1,98	0,42
	1	57	3249	4,33	4,67	-0,34

La tabla 124 resume la prueba de Welch.

Tabla 124: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08197

	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	9,62	10,12	-0,321	Límite superior I.C	$t_{0,025;47}$
Varianza	25,55	35,44	1,49	0,017	1,96
Varianza/n			0,030		
Máximo	27,97	32,89		Límite inferior I.C	
Mínimo	-0,16	0,00		-0,658	

En él se puede observar como el modelo estima los costes de recogida de manera adecuada, ya que el intervalo de confianza creado incluye el cero. Para elaborarlo se ha empleado la distribución t de Student, con un nivel de significancia del 95% y 47 grados de libertad, cuyo valor es 1.96.

2.2.6. Modelo de carga de trabajo

En este modelo se va a identificar que distribución de probabilidad siguen las variables aleatorias con las que se puede modelar la carga la de trabajo, así como los valores de los parámetros que las caracterizan. Estas variables son el número de repartos por día y el peso de cada uno de los repartos, en el caso de los repartos, y el número de recogidas por día y el peso de cada recogida, en el de las recogidas.

Las muestras que se emplean para las variables número de repartos, peso total repartido, número de recogidas y peso total recogido son las mismas muestras que las muestras de calibración que se emplearon para ajustar los modelos de los costes de reparto y recogida. Las muestras para las variables peso de cada reparto y peso de cada recogida, es una muestra con los pesos unitarios de cada uno de los repartos y de cada una de las recogidas.

2.2.6.1. Número de repartos por día

Minitab busca los parámetros para cada una de los distintos tipos de distribuciones de probabilidad, siendo estos parámetros tales que la distribución de probabilidad que caracterizan, es la que más se aproxima a los datos de la muestra. Una vez que se tienen estos datos hay que comprobar que la distribución de probabilidad que conforma se ajuste suficientemente bien a los datos. Para realizar esta comprobación se realizan los denominados test no paramétricos de bondad de ajuste. Debido a la naturaleza de esta variable aleatoria, la cual es un conteo de algún elemento, la primera función de distribución de probabilidad que se va a tratar de comprobar el ajuste es la distribución de Poisson. Como esta distribución es discreta se emplea el test no paramétrico de bondad de ajuste de Chi cuadrado. El resultado de este test se puede ver en la tabla 125.

Tabla 125: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08197

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
1-2	7	0,167	0,1207	5	0,7353	alfa=0,05
3-4	13	0,310	0,3115	13	0,0005	k=5
5-6	8	0,190	0,3230	14	2,2840	Par. Est. =1
7-10	14	0,333	0,2302	10	1,9391	n=4-1-1=2
	42				4,96	5,991

El valor lambda dado por Minitab para una distribución de Poisson con los datos de la muestra de esta variable aleatoria es 5,05. En la tabla 124 se puede observar como el valor límite para aceptar la hipótesis de que la una distribución de Poisson con lambda 5,05 ajusta adecuadamente la variable número de repartos por día es 5,991. Este valor corresponde a una distribución Chi cuadrado con 2 grados de libertad y nivel de significancia del 5%. Como en este caso el valor del test es 4,96, menor al valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.2.6.2. Peso de cada reparto

Para esta variable continua se presentan los valores aportados por Minitab para cada uno de los parámetros de cada distribución, así como los resultados de los test de Kolmogorov-Smirnov en la tabla 126.

Tabla 126: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08197

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,293	alfa	0,164	alfa	0,124	Media	0,076	Lambda	0,283
33,05	4,304	0,5631	2,389	0,6676	1,810	2,39	1,112	22,3	4,138
Desv. Est.	c _α = 1,358	beta	c _α = 1,358	beta	c _α = 1,358	Desv. Est.	c _α = 1,358	Theta	c _α = 1,358
50,08	α = 0,05	58,7	α = 0,05	24,07	α = 0,05	1,556	α = 0,05	23,53	α = 0,05

Como se puede ver la única distribución que supera satisfactoriamente el test de Kolmogorov-Smirnov es la distribución lognormal. Por lo tanto, para la simulación se empleará esta distribución para estimar los pesos de cada uno de los repartos.

2.2.6.3. Número de recogidas por día

Esta variable al igual que el número de repartos por día, es un conteo de algún elemento, y por tanto, es una distribución discreta. Por ser una distribución discreta, la primera distribución que se va a comprobar si ajusta adecuadamente es la distribución de Poisson. El resultado de la prueba de bondad de ajuste Chi cuadrado es el que se puede ver en la tabla 127.

Tabla 127: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08197

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
1-2	18	0,529	0,5067	17	0,0347	alfa=0,05
3	8	0,235	0,2190	7	0,0410	k=5
4-5	8	0,235	0,1293	4	2,9530	Par. Est. =1
	34				3,03	n=3-1-1=1
						3,841

El valor límite para aceptar la hipótesis de que la una distribución de Poisson con media 2,65 ajusta adecuadamente la variable número de repartos por día es 3,841. Este valor corresponde a una distribución Chi cuadrado con 1 grados de libertad y nivel de significancia del 5%. Este test solo tiene un grado de libertad ya que tanto el número de observaciones para los valores 1 como 5 es menor que 5 unidades, y es necesario tener un mínimo de 5 observaciones por intervalo. Por esta razón, las observaciones relativas al valor 1 se han agrupado con las del valor 2, y las del valor 5 con las del valor 4; es de aquí por lo que solo se tienen 3 intervalos a pesar de tener 5 valores de observación diferentes.

Como el resultado del test es 3,03, superior al valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente los datos de la muestra. Como se dispone de una distribución discreta que ajusta adecuadamente los datos, no es necesario el comprobar que distribuciones continuas ajustan convenientemente.

2.2.6.4. Peso de cada recogida

Por último, para finalizar con la elaboración del modelo de carga de trabajo para el código postal 08197 se va a analizar la variable peso de cada recogida. Esta variable es una variable continua, y las distribuciones a analizar son también continuas. Los parámetros de las distintas distribuciones analizadas, así como los resultados de los test de Kolmogorov-Smirnov de cada una se pueden ver en la tabla 128.

Tabla 128: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08197

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,285	alfa	0,158	alfa	0,120	Media	0,059	Lambda	0,261
38,41	2,740	0,7017	1,514	0,7603	1,150	2,787	0,560	24,62	2,498
Desv. Est.	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	Desv. Est.	$c_{\alpha} = 1,358$	Theta	$c_{\alpha} = 1,358$
65,06	$\alpha = 0,05$	54,74	$\alpha = 0,05$	31,37	$\alpha = 0,05$	1,298	$\alpha = 0,05$	24,02	$\alpha = 0,05$

Como se puede observar las únicas distribuciones que han superado el test de Kolmogorov-Smirnov han sido las distribuciones Weibull y lognormal, siendo la distribución lognormal la que menos discrepancia presenta. Por lo tanto, para la simulación se empleará esta distribución para estimar los pesos de cada uno de las recogidas.

2.2.7. Simulación de los costes de reparto y recogida, y de la carga de trabajo

En este apartado se llevará a cabo la simulación de los modelos creados para el estudio del vehículo de 1500 kg. de peso máximo autorizado en el código postal 08198. En primer lugar se simulará la carga de trabajo. Para ello hay que determinar el número de repartos y de recogidas para cada uno de los días simulados, y el peso total repartido y el peso total recogido por día. Empleando los ajustes de las distribuciones de probabilidad de Poisson obtenidos en los apartados 2.2.6.1 y 2.2.6.3. se simularán el número de repartos y el número de recogidas para cada uno días que se desean simular.

Posteriormente, se simulará el peso correspondiente de cada uno de los repartos y las recogidas mediante las distribuciones de probabilidad calculadas en los apartados 2.2.6.2. y 2.2.6.4. Sumando el peso unitario de todos los repartos simulados en cada uno de los días, se calculará el peso total repartido por día. Con el peso unitario de las recogidas simulado para cada uno de los días, se calculará el peso total recogido por día.

Finalmente, se simularán los modelos de costes de reparto y costes de recogida. Como valores de entrada a estos modelos se emplearán los datos simulados para el modelo de carga de trabajo. El número de repartos y el peso total repartido por día serán los datos de entrada para el modelo de costes de reparto. Análogamente, el número de recogidas y el peso total recogido por día serán los valores de entrada para el modelo de costes de recogida.

2.2.7.1. Simulación de los costes de reparto

La carga de trabajo relativa a los repartos para este código postal es la siguiente: el número de repartos puede ser descrito mediante una distribución de probabilidad Poisson con parámetro lambda igual a 5,05. Mientras que el peso de cada reparto sigue una distribución de probabilidad lognormal con media 2,39 y desviación estándar 1,556.

Un resumen de la carga de trabajo simulada es el de la tabla 129, en el cual se puede observar el promedio de repartos por día simulados, su desviación estándar, el número máximo de repartos a realizar en un día, así como el número mínimo. Y los mismos datos para la variable peso total repartido, siendo este peso total repartido la suma de los pesos simulados para cada una de los repartos.

Tabla 129: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08197

	Repartos por día	Peso total repartido por día
Promedio	5	163
Desv. Estan.	2	205
Máximo	11	1560
Mínimo	0	0

Esta simulación de la carga de trabajo se emplea como entrada para el modelo de costes de reparto. Como se han simulado 250 días de carga de trabajo, se obtendrán los estimados puntuales de costes de reparto para 250 días. Los resultados que aquí se exponen son el valor medio de límite inferior de las 250 simulaciones, la varianza y desviación estándar de este límite inferior, así como un intervalo de confianza de este límite inferior. Para el límite superior del intervalo de predicción los resultados que se presentan son análogos. Estos resultados son los de la tabla 130.

Tabla 130: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08197

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	11,99	15,90
Varianza	48	57
Desv. Estan.	6,94	7,53
I.C	11,27	15,11
	12,72	16,68

En la tabla 130 se puede observar como el límite inferior de los estimados puntuales de los costes de repartos se distribuyen con un valor medio de 11,99 euros por día y una desviación estándar de 6,94. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 15,9 y una desviación estándar de 7,53 euros por día.

2.2.7.2. Simulación de los costes de recogida

La carga de trabajo relativa a las recogidas para este código postal queda como sigue: el número de recogidas puede ser expresado mediante una distribución de probabilidad Poisson con un parámetro lambda igual a 2,65. Por otro lado, el peso de cada recogida puede ser descrito mediante una distribución lognormal con un valor medio igual a 2,787 y desviación estándar 1,298. La tabla 131 resume de la carga de trabajo simulada es el siguiente, en el cual se puede observar el promedio de recogidas por día simulados, su desviación estándar, el número máximo de recogidas a realizar en un día, así como el

número mínimo. Y los datos análogos para la variable peso total recogido, siendo este peso total recogido la suma de los pesos simulados para cada una de las recogidas.

Tabla 131: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08197

	Recogidas por día	Peso total recogido por día
Promedio	3	88
Desv. Estan.	2	117
Maximo	7	921
Mínimo	0	0

Y la simulación de los costes de recogida correspondiente a esta carga de trabajo para 250 días se resume en la tabla 132.

Tabla 132: Resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08197

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	7,13	10,57
Varianza	31	33
Desv. Estan.	5,58	5,76
I.C	6,54	9,97
	7,71	11,17

Donde se puede comprobar como el límite inferior de los estimados puntuales de los costes de recogida se distribuyen con un valor medio de 7,13 euros por día y una desviación estándar de 5,58. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 10,57 y una desviación estándar de 5,76 euros por día.

2.3. Código postal 08196

El tipo de repartos que se dan en este código postal son menores de 300 kg., y las recogidas son inferiores de 700 kg. por lo que en él solo existe volumen de trabajo para vehículos de 1500 kg. de peso máximo autorizado.

Los pasos a seguir en la elaboración del modelo de costes de la red de distribución capilar de este código postal son análogos a los empleados en los códigos postales anteriores.

2.3.1. Toma de muestras

Al igual que para el resto de modelo elaborados hasta el momento, el primer paso es la toma de muestras de las variables identificadas como significativas en el primer apartado de este capítulo. Éstas son los costes de reparto y los costes de recogida por día, los cuales dependen a su vez, del número de repartos por día y del peso total repartido, y del número de recogidas por día y del peso total recogido por día, respectivamente. Para ello, se llevado a cabo un estudio observacional del cual se han obtenido 84 observaciones las cuales se pueden observar en las tablas 133 y 134.

Tabla 133: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08196

Día	REPARTOS			Día	REPARTOS		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	21	976	77	43	21	609	68
2	19	571	59	44	6	471	25
3	18	454	52	45	25	1.173	88
4	25	1.166	89	46	32	1.984	122
5	34	1.614	120	47	26	733	80
6	17	559	54	48	33	1.410	114
7	20	824	67	49	23	1.128	80
8	32	968	99	50	26	1.471	100
9	20	663	65	51	24	953	81
10	24	1.329	91	52	22	766	71
11	26	679	75	53	16	686	53
12	21	319	55	54	22	817	75
13	27	724	82	55	22	612	69
14	23	1.126	82	56	33	1.555	118
15	20	260	52	57	32	1.436	107
16	25	1.012	86	58	20	481	58
17	28	1.715	109	59	20	1.044	73
18	35	1.602	126	60	25	1.775	108
19	25	1.248	89	61	25	703	75
20	23	521	67	62	26	1.028	87
21	22	798	74	63	27	897	86
22	29	1.332	101	64	23	1.143	84
23	24	1.004	83	65	28	1.211	98
24	35	1.238	112	66	27	1.148	88
25	25	1.116	83	67	23	807	75
26	17	700	57	68	29	1.099	94
27	17	545	54	69	22	676	67
28	34	1.404	112	70	27	1.015	89
29	24	718	75	71	24	1.296	90
30	30	1.377	103	72	22	1.175	84
31	29	1.270	102	73	23	494	66
32	36	1.480	119	74	26	1.200	92
33	19	1.210	71	75	27	1.113	95
34	30	1.042	97	76	20	751	67
35	23	1.039	81	77	28	744	81
36	12	464	39	78	18	503	54
37	37	992	112	79	9	923	41
38	22	961	76	80	23	1.063	83
39	31	792	96	81	16	1.002	64
40	24	575	70	82	26	1.023	82
41	28	794	88	83	27	1.087	91
42	21	517	61	84	18	1.070	69

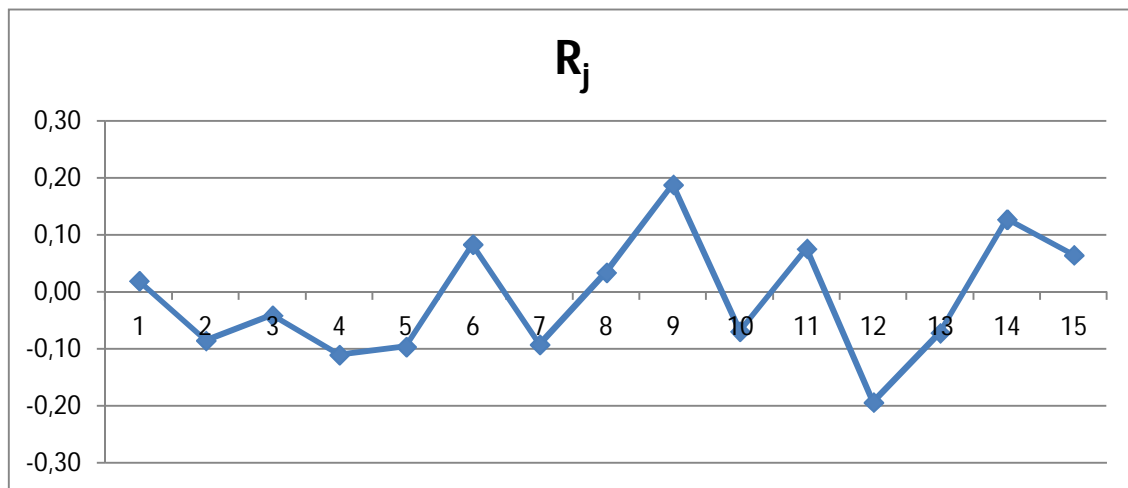
Tabla 134: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08196

Día	RECOGIDAS			Día	RECOGIDAS		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	16	2.453	89	43	15	2.343	87
2	4	901	28	44	22	1.501	87
3	18	1.700	78	45	25	3.310	129
4	16	2.009	80	46	25	2.284	106
5	24	3.126	132	47	16	1.836	79
6	17	3.085	107	48	18	2.632	95
7	13	1.865	72	49	16	1.990	82
8	20	1.950	87	50	21	1.469	86
9	14	2.069	76	51	14	2.271	81
10	20	2.023	93	52	17	1.488	79
11	17	2.072	85	53	17	2.951	104
12	16	1.974	80	54	17	1.860	82
13	13	956	55	55	11	2.393	75
14	22	2.826	116	56	23	3.308	128
15	17	2.229	94	57	21	1.873	94
16	14	1.237	63	58	12	2.095	69
17	17	2.826	104	59	19	1.474	86
18	15	2.264	84	60	16	1.768	78
19	16	2.449	83	61	20	1.754	89
20	9	1.102	43	62	21	1.791	97
21	14	1.197	64	63	19	2.839	105
22	13	2.580	81	64	14	1.965	77
23	16	1.457	72	65	18	1.800	83
24	18	2.482	94	66	19	3.069	108
25	16	1.032	57	67	11	1.940	63
26	13	883	52	68	3	55	9
27	17	1.376	70	69	20	2.494	99
28	20	2.146	94	70	21	2.210	98
29	15	1.478	72	71	13	1.292	58
30	14	1.210	61	72	12	1.630	66
31	16	1.503	70	73	10	1.151	49
32	17	3.102	102	74	18	2.333	89
33	12	1.262	61	75	16	1.383	69
34	13	1.642	67	76	18	2.251	94
35	14	1.732	68	77	8	996	45
36	12	1.025	53	78	15	2.234	84
37	7	725	34	79	11	1.112	49
38	14	2.047	74	80	20	2.458	104
39	17	985	64	81	15	1.347	61
40	18	798	68	82	17	2.136	86
41	11	1.969	65	83	11	2.280	68
42	21	2.051	95	84	20	1.690	87

2.3.2. Independencia de la muestras

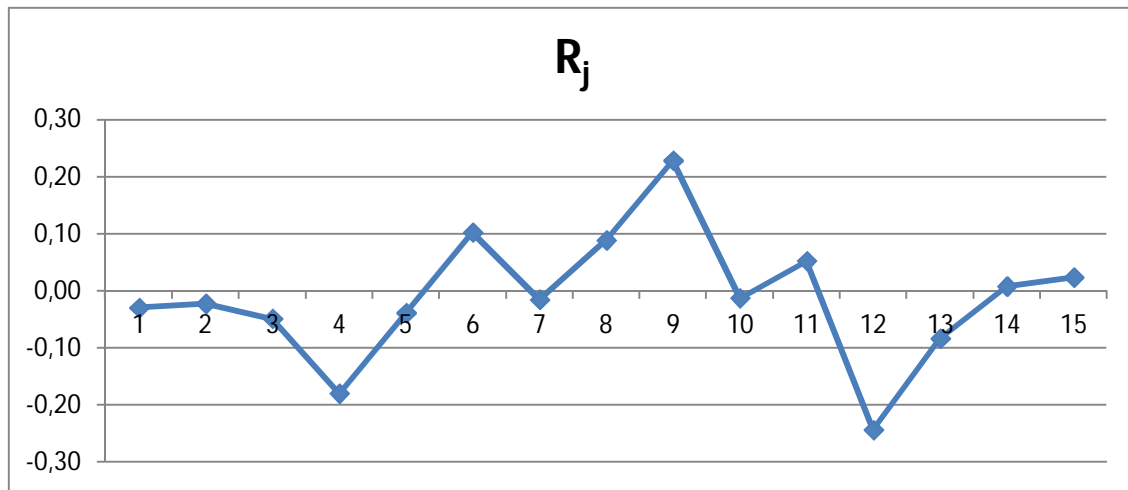
Tras la obtención de las muestras, hay que comprobar si las observaciones de las mismas son independientes. La independencia de estas observaciones es necesaria para poder aplicar el método de máxima verosimilitud, así como los test no paramétricos de bondad de ajuste, ya que ambos presuponen la independencia de las mismas.

Para comprobar que estas observaciones son independientes en primer lugar se elaboran las gráficas de autocorrelación, y posteriormente las gráficas de dispersión. A continuación se muestran las gráficas de autocorrelación relativas a las muestras de los repartos. La Figura 57 corresponde a la gráfica de autocorrelación de la muestra de los costes de reparto por día. La Figura 58 corresponde a la gráfica de autocorrelación del número de repartos por día, y la Figura 59 corresponde a la gráfica de autocorrelación del peso total repartido por día.



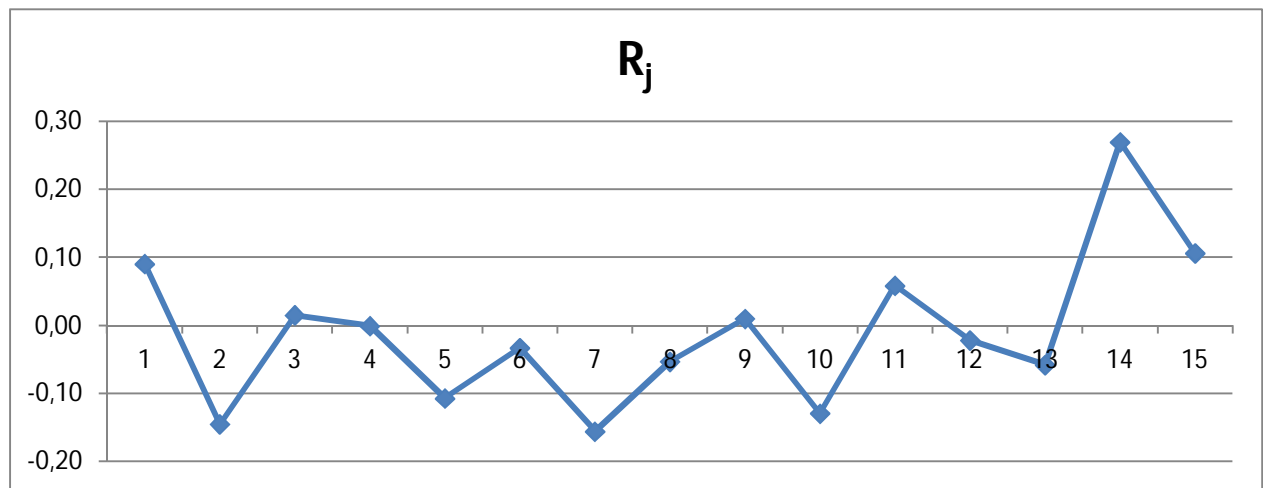
R_j	0,02	-0,08	-0,04	-0,11	-0,10	0,08	-0,09	0,03	0,19	-0,07	0,08	-0,19	-0,07	0,13	0,07
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	8	-35	-17	-46	-40	35	-38	14	79	-29	32	-81	-30	53	27

Figura 57: Grafica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08196



R_j	-0,03	-0,02	-0,05	-0,18	-0,04	0,10	-0,02	0,09	0,23	-0,01	0,05	-0,24	-0,08	0,01	0,02
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-1,0	-0,7	-1,6	-5,9	-1,3	3,4	-0,5	2,9	7,5	-0,4	1,7	-8,0	-2,8	0,3	0,8

Figura 58: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08196



R_j	0,09	-0,14	0,02	0,00	-0,11	-0,03	-0,16	-0,05	0,01	-0,13	0,06	-0,02	-0,06	0,27	0,11
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	11446	-18275	1926	-87	-13510	-4135	-19665	-6679	1273	-16312	7407	-2748	-7306	34045	13446

Figura 59: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08196

Como en ninguna de las gráficas los valores de R_j superan los valores de -1 y 1, ni se da ningún tipo de patrón determinado, se puede concluir que de acuerdo a esta prueba las observaciones de cada muestra son independientes. Una vez analizadas las gráficas de autocorrelación, se estudiarán las graficas de dispersión.

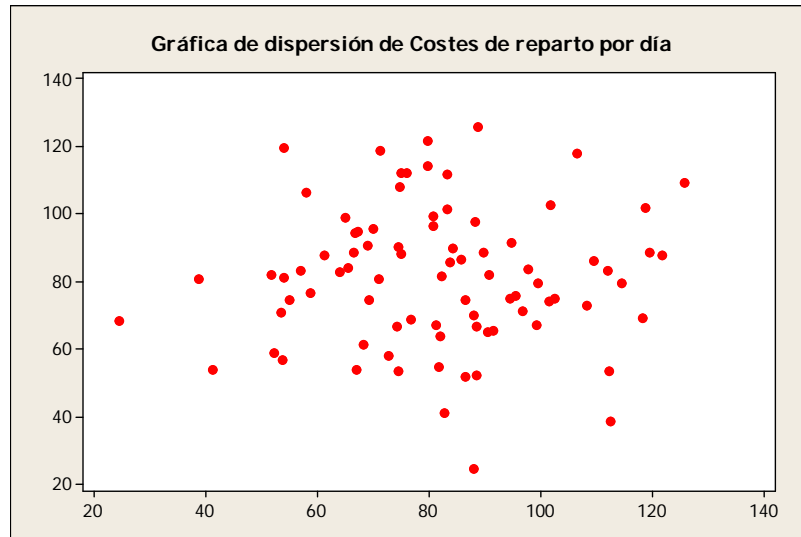


Figura 60: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08196

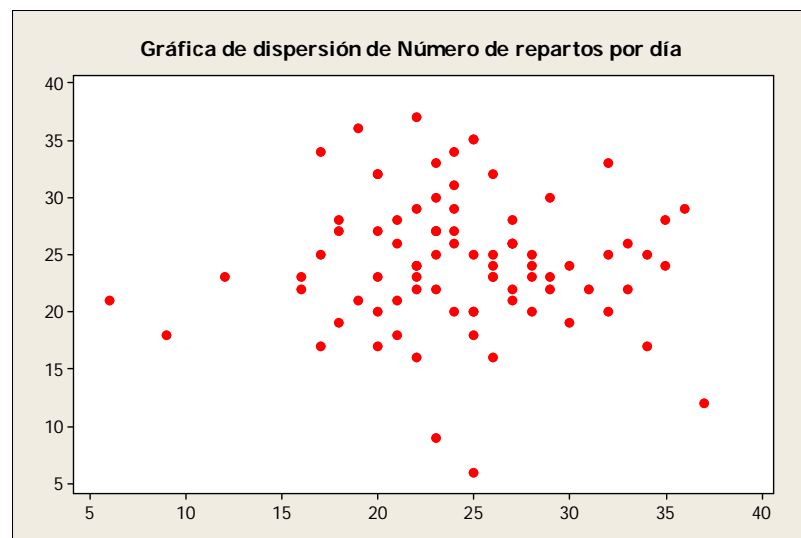


Figura 61: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08196

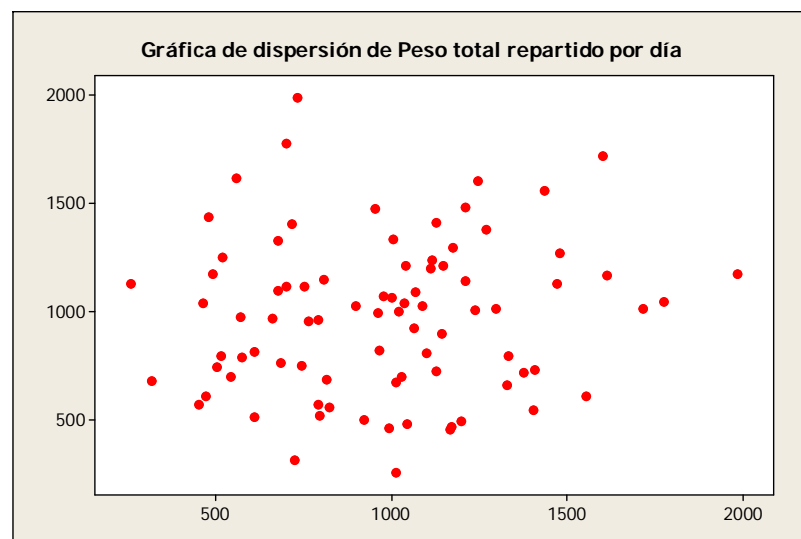
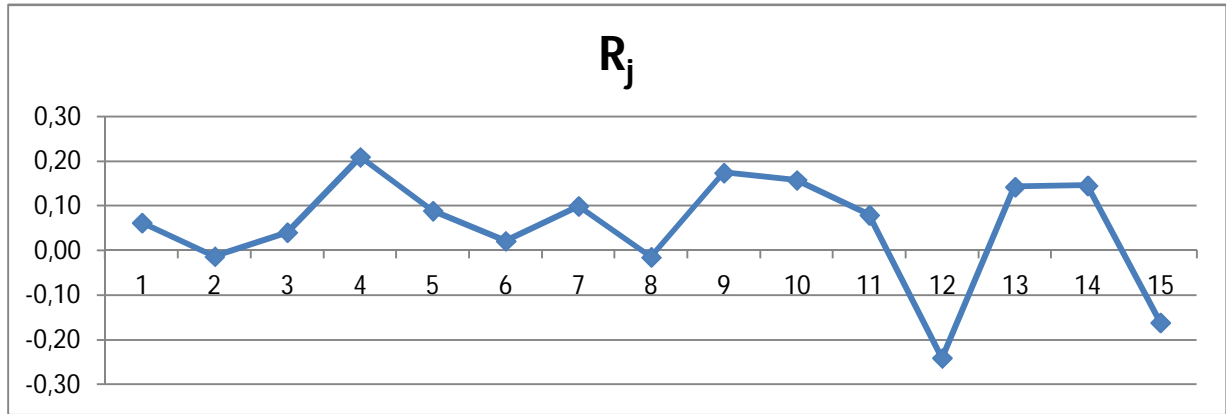


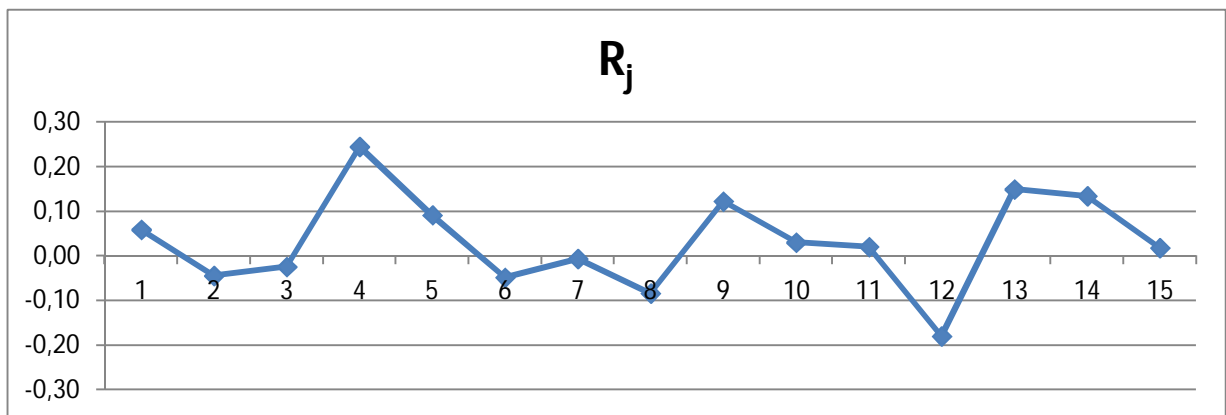
Figura 62: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08196

Como los puntos en las tres gráficas se distribuyen de manera aleatoria, se puede concluir que las observaciones de las tres muestras son independientes. Una vez analizadas las observaciones de las muestras relativas a los repartos, se hará el proceso para las muestras relativas a las recogidas.



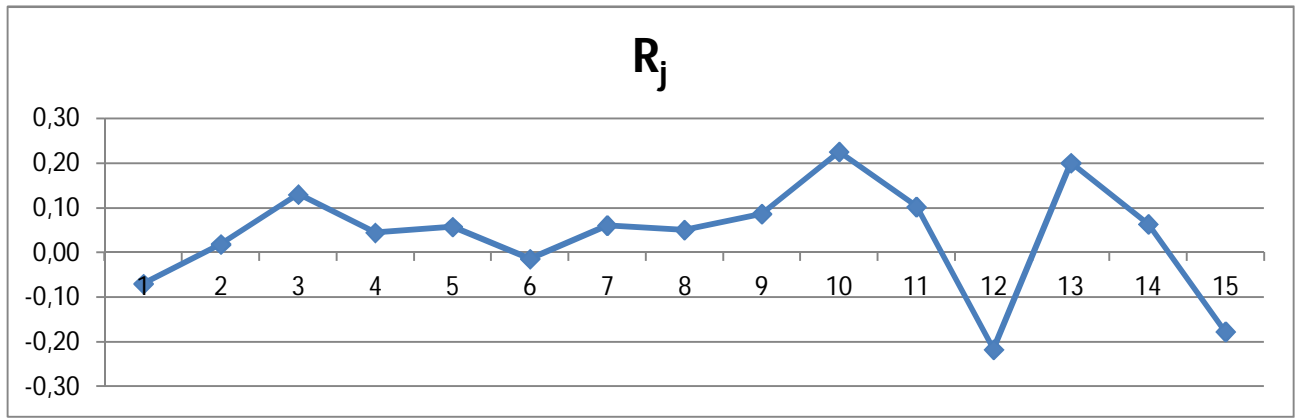
R_j	0,06	-0,01	0,04	0,21	0,09	0,02	0,10	-0,01	0,17	0,16	0,08	-0,24	0,14	0,15	-0,16
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	29	-6	19	97	41	10	46	-7	81	73	37	-111	66	67	-74

Figura 63: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08196



R_j	0,06	-0,04	-0,02	0,24	0,09	-0,05	-0,01	-0,08	0,12	0,03	0,02	-0,18	0,15	0,13	0,02
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	1,0	-0,8	-0,4	4,3	1,6	-0,9	-0,1	-1,5	2,2	0,5	0,4	-3,2	2,6	2,4	0,3

Figura 64: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08196



R_j	-0,07	0,02	0,13	0,05	0,06	-0,01	0,06	0,05	0,09	0,23	0,10	-0,22	0,20	0,06	-0,18
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-30419	8334	57188	19769	25509	-6019	26845	22425	38151	99159	45171	-94971	88223	28117	-77540

Figura 65: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08196

Como en las tres gráficas no se supera los valores de -1 y 1 para el parámetro R_j , y no se da ningún patrón en ellas, se puede concluir que de acuerdo con esta prueba, que las observaciones de las muestras son independientes. A continuación se muestran las gráficas de dispersión de las mismas muestras.

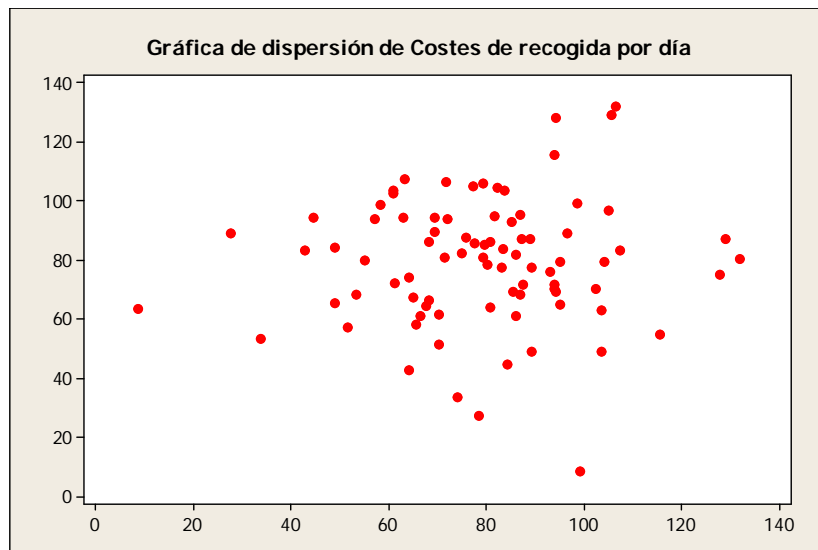


Figura 66: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08196

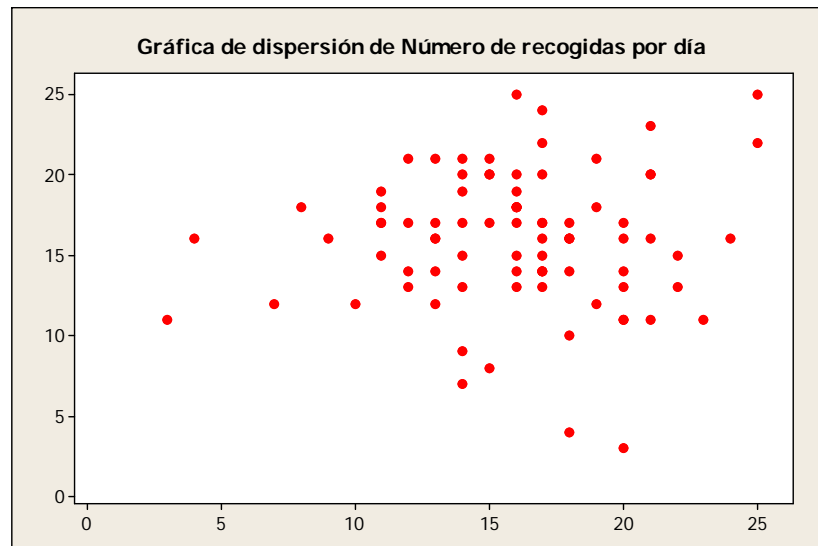


Figura 67: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08196

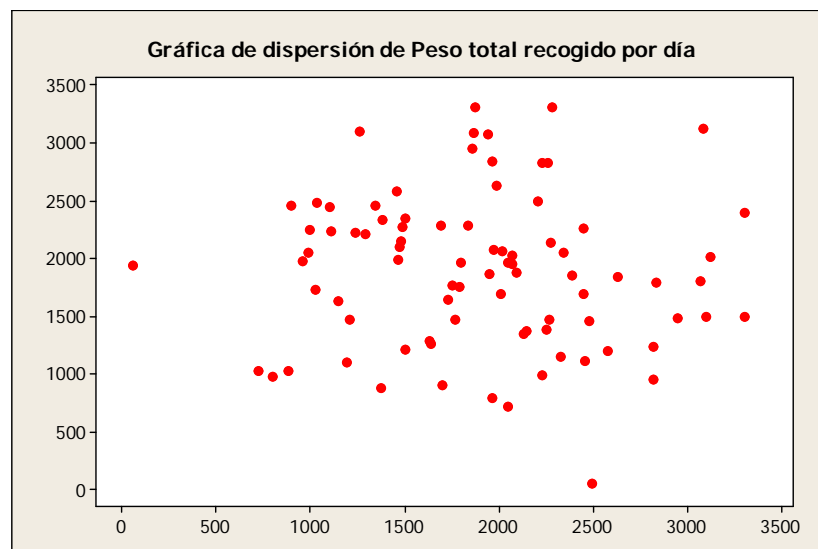


Figura 68: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08196

Como los puntos en las tres gráficas se distribuyen de manera aleatoria en los planos, sin seguir ningún tipo de línea, se puede concluir que de acuerdo a esta prueba, que las observaciones relativas a las muestras de las variables de las recogidas son independientes.

2.3.3. Separación de las muestras de calibración y predicción

Una vez que se concluye que las 84 observaciones de reparto y de recogida son independientes, éstas han de ser separadas en la muestra de calibración y la muestra de predicción. Esta separación se realizará mediante asignación aleatoria. Se asigna un número aleatorio proveniente de una distribución uniforme a cada una de las observaciones, y si éste es mayor o igual que 0,5, esta observación se incluirá en la muestra de calibración. Por el contrario, si es menor de 0,5, se incluye en la muestra de predicción.

La asignación de números aleatorios a cada una de las observaciones de reparto se muestra en la tabla 135.

Tabla 135: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08196

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,005	22	0,120	43	0,760	64	0,147
2	0,194	23	0,013	44	0,119	65	0,748
3	0,458	24	0,057	45	0,389	66	0,906
4	0,054	25	0,381	46	0,702	67	0,413
5	0,671	26	0,069	47	0,665	68	0,980
6	0,016	27	0,511	48	0,628	69	0,638
7	0,066	28	0,068	49	0,344	70	0,323
8	0,319	29	0,973	50	0,772	71	0,418
9	0,778	30	0,809	51	0,164	72	0,917
10	0,370	31	0,032	52	0,106	73	0,605
11	0,196	32	0,193	53	0,921	74	0,318
12	0,148	33	0,223	54	0,779	75	0,937
13	0,876	34	0,088	55	0,350	76	0,638
14	0,677	35	0,888	56	0,058	77	0,937
15	0,827	36	0,676	57	0,547	78	0,643
16	0,130	37	0,084	58	0,118	79	0,723
17	0,633	38	0,760	59	0,554	80	0,182
18	0,894	39	0,597	60	0,736	81	0,927
19	0,487	40	0,891	61	0,555	82	0,850
20	0,180	41	0,576	62	0,774	83	0,443
21	0,060	42	0,104	63	0,698	84	0,708

A partir de estas asignaciones, las muestras de calibración y de predicción para el modelo de costes de reparto quedan como se puede ver en la tabla 136.

Tabla 136: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08196

Día	REPARTOS. MUESTRA DE CALIBRACIÓN			Día	REPARTOS. MUESTRA DE PREDICCIÓN		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	34	1.614	120	1	21	976	77
2	20	663	65	2	19	571	59
3	27	724	82	3	18	454	52
4	23	1.126	82	4	25	1.166	89
5	20	260	52	5	17	559	54
6	28	1.715	109	6	20	824	67
7	35	1.602	126	7	32	968	99
8	17	545	54	8	24	1.329	91
9	24	718	75	9	26	679	75
10	30	1.377	103	10	21	319	55
11	23	1.039	81	11	25	1.012	86
12	12	464	39	12	25	1.248	89
13	22	961	76	13	23	521	67
14	31	792	96	14	22	798	74
15	24	575	70	15	29	1.332	101
16	28	794	88	16	24	1.004	83
17	21	609	68	17	35	1.238	112
18	32	1.984	122	18	25	1.116	83
19	26	733	80	19	17	700	57
20	33	1.410	114	20	34	1.404	112
21	26	1.471	100	21	29	1.270	102
22	16	686	53	22	36	1.480	119
23	22	817	75	23	19	1.210	71
24	32	1.436	107	24	30	1.042	97
25	20	1.044	73	25	37	992	112
26	25	1.775	108	26	21	517	61
27	25	703	75	27	6	471	25
28	26	1.028	87	28	25	1.173	88
29	27	897	86	29	23	1.128	80
30	28	1.211	98	30	24	953	81
31	27	1.148	88	31	22	766	71
32	29	1.099	94	32	22	612	69
33	22	676	67	33	33	1.555	118
34	22	1.175	84	34	20	481	58
35	23	494	66	35	23	1.143	84
36	27	1.113	95	36	23	807	75
37	20	751	67	37	27	1.015	89
38	28	744	81	38	24	1.296	90
39	18	503	54	39	26	1.200	92
40	9	923	41	40	23	1.063	83
41	16	1.002	64	41	27	1.087	91
42	26	1.023	82				
43	18	1.070	69				

Se puede observar como la muestra de calibración consta de 43 observaciones, y la de predicción 41, debido a la aleatoriedad de la asignación de las observaciones en las muestras. Por otro lado, la asignación de números aleatorios a cada una de las observaciones de recogida es la que se puede ver en la tabla 137.

Tabla 137: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08196

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,315	22	0,885	43	0,804	64	0,558
2	0,916	23	0,815	44	0,247	65	0,978
3	0,355	24	0,939	45	0,966	66	0,294
4	0,641	25	0,568	46	0,098	67	0,463
5	0,108	26	0,702	47	0,508	68	0,027
6	0,983	27	0,392	48	0,607	69	0,253
7	0,676	28	0,744	49	0,647	70	0,959
8	0,413	29	0,339	50	0,847	71	0,091
9	0,508	30	0,985	51	0,041	72	0,249
10	0,189	31	0,664	52	0,313	73	0,920
11	0,184	32	0,312	53	0,107	74	0,433
12	0,480	33	0,959	54	0,384	75	0,855
13	0,998	34	0,780	55	0,670	76	0,704
14	0,845	35	0,472	56	0,454	77	0,343
15	0,077	36	0,650	57	0,773	78	0,944
16	0,143	37	0,381	58	0,762	79	0,186
17	0,847	38	0,112	59	0,386	80	0,122
18	0,853	39	0,304	60	0,193	81	0,647
19	0,701	40	0,205	61	0,708	82	0,718
20	0,325	41	0,293	62	0,492	83	0,479
21	0,366	42	0,396	63	0,642	84	0,523

Y las muestras de calibración y predicción para las recogidas de este tipo de vehículo en este código postal se separan de manera análoga a los repartos, pudiendo ver en la tabla 138 cada una de las muestras. En este caso, debido a la aleatoriedad de la asignación, se tienen 42 observaciones en la muestra de calibración y 42 en la muestra de predicción.

Tabla 138: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08196

Día	RECOGIDAS. MUESTRA DE CALIBRACIÓN			Día	RECOGIDAS. MUESTRA DE PREDICCIÓN		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	22	1.501	87	1	15	2.343	87
2	25	2.284	106	2	25	3.310	129
3	18	2.632	95	3	16	1.836	79
4	16	1.990	82	4	21	1.469	86
5	14	2.271	81	5	17	1.488	79
6	11	2.393	75	6	17	2.951	104
7	23	3.308	128	7	17	1.860	82
8	19	1.474	86	8	21	1.873	94
9	16	1.768	78	9	12	2.095	69
10	20	1.754	89	10	21	1.791	97
11	14	1.965	77	11	19	2.839	105
12	18	1.800	83	12	20	2.494	99
13	19	3.069	108	13	13	1.292	58
14	11	1.940	63	14	18	2.333	89
15	3	55	9	15	8	996	45
16	21	2.210	98	16	11	1.112	49
17	12	1.630	66	17	20	2.458	104
18	10	1.151	49	18	15	1.347	61
19	16	1.383	69	19	17	2.136	86
20	18	2.251	94	20	11	2.280	68
21	15	2.234	84	21	20	1.690	87
22	16	2.453	89	22	4	901	28
23	18	1.700	78	23	16	2.009	80
24	24	3.126	132	24	14	2.069	76
25	17	3.085	107	25	20	2.023	93
26	13	1.865	72	26	17	2.072	85
27	20	1.950	87	27	16	1.974	80
28	13	956	55	28	22	2.826	116
29	17	2.229	94	29	17	2.826	104
30	14	1.237	63	30	15	2.264	84
31	16	2.449	83	31	9	1.102	43
32	14	1.197	64	32	18	2.482	94
33	13	2.580	81	33	16	1.032	57
34	16	1.457	72	34	13	883	52
35	20	2.146	94	35	17	1.376	70
36	16	1.503	70	36	15	1.478	72
37	12	1.262	61	37	14	1.210	61
38	13	1.642	67	38	17	3.102	102
39	12	1.025	53	39	14	1.732	68
40	17	985	64	40	7	725	34
41	18	798	68	41	14	2.047	74
42	21	2.051	95	42	11	1.969	65

2.3.4. Modelo de costes de reparto

Al igual que para los dos códigos postales anteriormente analizados, los parámetros de regresión del modelo de costes de reparto se calculan mediante el método de máxima verosimilitud. Se va a estudiar qué modelos ofrecen mayores valores de máxima verosimilitud, bajo los supuestos de diferentes tipos de distribuciones de probabilidad para los errores, y suposiciones de no linealidad de las variables exógenas. Estos supuestos son los mismos que para los casos anteriores.

Tabla 139: VMV modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
81,692					-191,5		0%	21,03
-0,156	3,378				-152,3	39,236	83%	8,55
-0,501	2,348	0,026			-95,5	56,720	99%	2,31
-6,463	2,874	0,026	-0,012		-94,5	1,060	99%	2,29
-1,456	2,348	0,028		0	-95,4	-0,944	99%	2,34

Tabla 140: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
4,403					-191,5		0%	21,03
3,338	0,043				-150,7	40,815	85%	8,24
3,409	0,029	0,0003			-120,0	30,627	96%	4,09
2,869	0,073	0,0003	-0,001		-106,6	13,407	98%	3,03
2,869	0,073	0,0003	-0,001	0	-106,6	0,000	98%	3,07

Tabla 141: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,012					-191,5		0%	21,03
0,024	-0,0005				-153,6	37,858	82%	8,83
0,025	-0,0005	0			-153,6	0,000	82%	8,94
0,032	-0,001		0,00001		-152,9	0,717	83%	8,79

Tabla 142: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-9,039					-191,5		0%	21,03
-4,269	-0,194				-150,7	40,816	85%	8,24
-4,456	-0,133	-0,001			-107,3	43,402	98%	3,04
-3,161	-0,242	-0,001	0,002		-99,3	7,963	99%	2,56
-3,161	-0,242	-0,001	0,002	0	-99,3	0,000	98%	2,59

Tabla 143: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
81,374					-193,2		0%	21,03
-2,179	3,412				-154,2	38,975	83%	8,64
-1,952	2,384	0,026			-96,0	58,259	99%	2,34
-1,953	2,383	0,026	0		-96,0	0,007	99%	2,37
-1,953	2,383	0,026		0	-96,0	-0,015	99%	2,37

Tabla 144: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
4,399					-193,2		0%	21,03
3,311	0,043				-150,3	42,922	84%	8,32
3,455	0,028	0,0003			-120,3	29,980	96%	4,20
3,455	0,028	0,0003	0		-120,3	0,003	96%	4,26
3,455	0,028	0,0003		0	-120,3	-0,003	96%	4,26

Tabla 145: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,012					-193,2		0%	21,03
0,024	0,000				-153,2	40,034	82%	9,02
0,024	0,000	0			-153,0	0,206	83%	8,70
0,024	0,000		0		-153,2	-0,206	81%	9,14

Tabla 146: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-9,021					-193,2		0%	21,03
-4,022	-0,202				-151,7	41,517	84%	8,31
-4,567	-0,129	-0,001			-108,7	42,955	98%	3,12
-4,567	-0,129	-0,001	0,0000003		-108,7	0,004	98%	3,16
-4,567	-0,129	-0,001		0	-108,7	-0,005	98%	3,16

Tabla 147: VMV modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
22,183					-248,2		-820%	63,79
33,005	0,556				-230,2	17,931	-271%	40,50
-0,926	2,363	0,026			-96,0	134,192	99%	2,32
-7,478	2,952	0,026	-0,014		-94,7	1,292	99%	2,29
-2,563	2,367	0,029		-0,000002	-95,6	-0,868	99%	2,35

Se puede observar como los valores de máxima verosimilitud para los supuestos de que los errores se distribuyen según una distribución de probabilidad normal. Por lo tanto, en la simulación del modelo de costes de reparto por día se emplearán los parámetros de regresión calculados bajo el supuesto de que los errores se distribuyen normalmente.

2.3.4.1. Selección de variables para el modelo de costes de repartos

Si se observa la tabla del apartado anterior correspondiente al supuesto de la distribución normal, se puede observar como al añadir la variable número de repartos por día, el cociente de verosimilitud es superior a dos unidades, por lo que adición de esta variable aporta un mejor ajuste al modelo con su inclusión. Con la variable peso total repartido por día ocurre exactamente lo mismo, ya que al introducirla en el modelo el cociente de verosimilitud aumenta en 56 unidades. Sin embargo, al introducir las variables número de repartos por día elevada a la segunda potencia y peso total repartido por día también elevada a la segunda potencia, los cocientes de verosimilitud para cada uno de los casos no superan el

valor límite de dos unidades, por lo que la inclusión de estas variables en el modelo no aporta un mejor ajuste a los datos. En definitiva, el modelo de costes de reparto para el vehículo de 1500 kg. de peso máximo autorizado del código postal 08196, consta de dos variables, número de repartos y peso total repartido por día, con parámetros de regresión igual a 2,348 y 0,026 respectivamente. Siendo el valor umbral igual a -0,501.

2.3.4.2. Adecuación del modelo de costes de reparto

Como el mayor valor máximo de verosimilitud es el que aporta una distribución de probabilidad normal, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. Este análisis se puede ver en la figura 69.

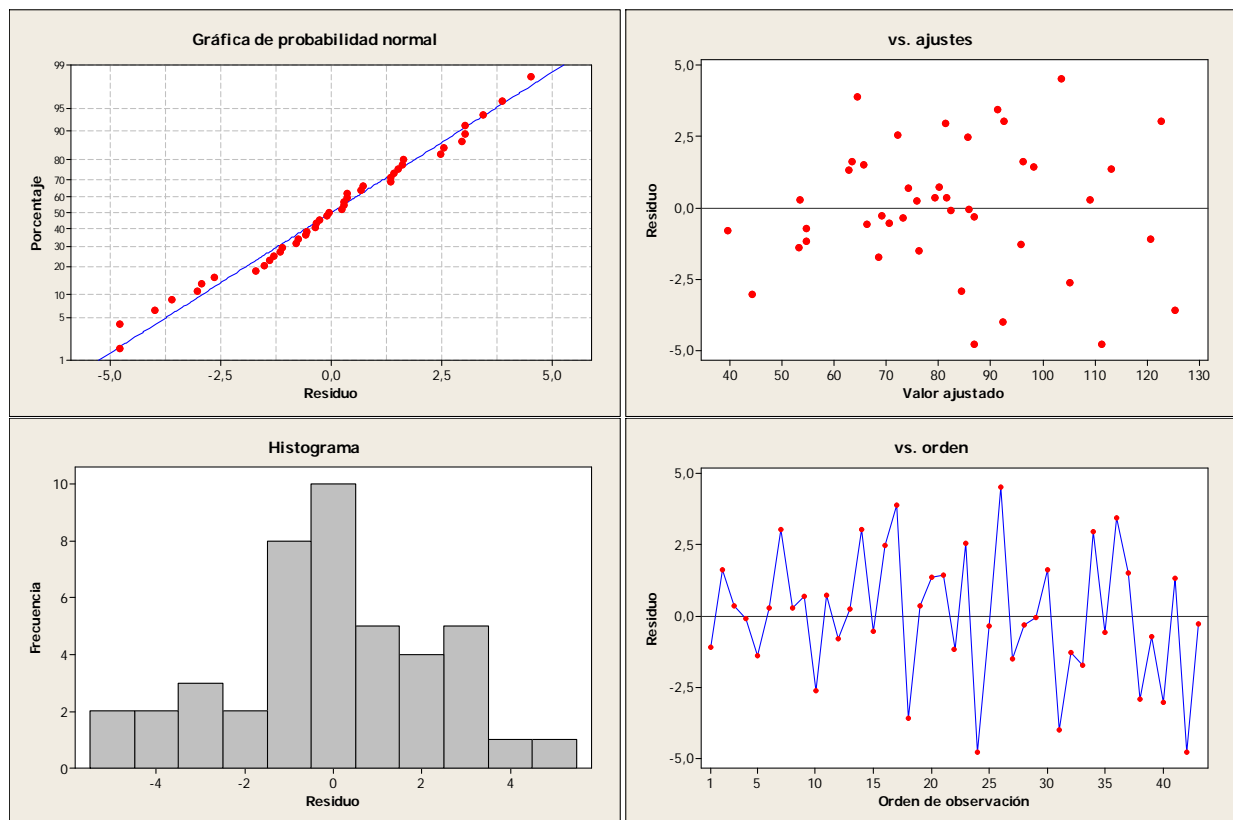


Figura 69: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08196

En la primera gráfica se puede observar como los errores se distribuyen de acuerdo a una distribución logística ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la

distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.3.4.3. Validación del modelo de costes de reparto

La validación del modelo de costes de reparto para este tipo de vehículo se lleva a cabo mediante la prueba *t* pareado ó prueba de Welch. Los valores que se emplean en esta prueba se obtienen de la tabla 148.

Tabla 148: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
-0,501	2,348	0,026			
	21	976	73,79	76,64	-2,85
	19	571	58,73	58,82	-0,08
	18	454	53,39	52,29	1,10
	25	1166	88,05	88,55	-0,50
	17	559	53,73	54,04	-0,31
	20	824	67,56	67,05	0,50
	32	968	99,42	99,22	0,21
	24	1329	89,87	90,51	-0,64
	26	679	77,94	74,54	3,40
	21	319	56,98	54,98	2,00
	25	1012	84,11	86,41	-2,30
	25	1248	90,15	88,69	1,46
	23	521	66,85	66,59	0,25
	22	798	71,59	74,32	-2,73
	29	1332	101,69	101,50	0,20
	24	1004	81,56	83,21	-1,66
	35	1238	113,38	111,99	1,39
	25	1116	86,77	83,26	3,51
	17	700	57,34	57,01	0,32
	34	1404	115,28	112,32	2,96
	29	1270	100,11	101,80	-1,70
	36	1480	121,92	118,81	3,10
	19	1210	75,08	71,31	3,78
	30	1042	96,62	96,74	-0,12
	37	992	111,78	112,40	-0,62
	21	517	62,05	61,32	0,73
	6	471	25,64	24,58	1,07
	25	1173	88,23	87,92	0,31
	23	1128	82,38	79,73	2,65
	24	953	80,25	80,69	-0,44
	22	766	70,77	70,97	-0,20
	22	612	66,83	69,18	-2,36
	33	1555	116,79	118,15	-1,36
	20	481	58,78	58,04	0,74
	23	1143	82,76	83,64	-0,88
	23	807	74,17	74,99	-0,83
	27	1015	88,88	88,60	0,28
	24	1296	89,03	89,81	-0,78
	26	1200	91,27	91,58	-0,31
	23	1063	80,72	82,72	-2,01
	27	1087	90,73	90,84	-0,11
	24	1333	89,97	87,39	2,58
	34	763	98,87	99,60	-0,73
	23	569	68,08	71,21	-3,13

En ella se comparan uno a uno los valores de variable endógena dadas por el modelo, tras introducir los valores de variable exógena de la muestra de predicción, con los valores correspondientes

de los costes de reparto de cada día de la muestra de predicción. La tabla que resume de esta prueba de validación es la tabla 149.

Tabla 149: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08196

	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	81,81	81,68	0,134	Límite superior I.C 0,655	$t_{0,025;42}$ 1,96
Varianza	387,32	384,96	3,10		
Varianza/n			0,071		
Máximo	121,92	118,81		Límite inferior I.C -0,386	
Mínimo	25,64	24,58			

El modelo de costes de reparto para este tipo de vehículo se considera válido ya que en la prueba de Welch, el intervalo creado contiene el cero.

2.3.5. Modelo de costes de recogida

Tras la consecución del modelo de costes de reparto, a continuación se crea el modelo de costes de recogida. Los pasos a seguir son análogos a los del modelo de costes de reparto. Los valores de máxima verosimilitud para cada uno de los supuestos son los siguientes.

Tabla 150: VMV modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
79,879					-187,6		0%	21,30
9,419	4,346				-159,7	27,825	73%	11,12
0,626	2,773	0,018			-103,5	56,205	98%	2,95
-0,667	2,960	0,018	-0,006		-103,4	0,076	98%	2,99
2,748	2,801	0,015		0	-102,8	0,652	98%	2,94

Tabla 151: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
4,381					-187,6		0%	21,30
3,515	0,052				-162,6	24,974	69%	11,90
3,398	0,033	0,0002			-126,8	35,773	94%	5,14
2,807	0,104	0,0002	-0,002		-113,7	13,063	97%	3,82
2,807	0,104	0,0002	-0,002	0	-113,7	0,000	97%	3,87

Tabla 152: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-2,000					-245,2		-1459%	
-5,146	-12,412				-245,0	0,249	-1480%	
-122,517		-8069,795			-245,0	0,004	-1479%	

Tabla 153: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-8,937					-187,6		0%	21,30
-4,980	-0,241				-161,2	26,377	71%	11,51
-4,391	-0,157	-0,001			-114,5	46,651	97%	3,84
-2,871	-0,348	-0,001	0,006		-105,0	9,550	98%	3,10
-2,871	-0,348	-0,001	0,006	0	-105,0	0,000	98%	3,14

Tabla 154: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
80,696					-186,1		0%	21,31
12,923	4,053				-163,5	22,628	72%	11,26
2,724	2,672	0,018			-102,8	60,993	98%	3,02
-2,167	3,353	0,018	-0,022		-100,6	1,873	98%	3,10
-1,840	2,736	0,021		-0,000001	-104,8	-4,161	98%	3,13

Tabla 155: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
4,392					-186,1		0%	21,32
3,626	0,045				-165,1	21,037	67%	12,15
3,455	0,031	0,0002			-114,9	50,179	94%	5,29
3,455	0,031	0,0002	0		-114,9	0,000	94%	5,36
3,455	0,031	0,0002		0	-114,9	0,000	94%	5,36

Tabla 156: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,012					-186,1		0%	21,31
0,022	-0,001				-165,9	20,181	64%	12,81
0,022	-0,001	0			-165,9	0,000	63%	12,97
0,022	-0,001		0		-165,9	0,000	63%	12,97

Tabla 157: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-8,987					-186,1		0%	21,32
-5,360	-0,215				-164,4	21,710	70%	11,71
-4,620	-0,145	-0,001			-108,0	56,392	97%	3,96
-4,628	-0,148	-0,001	0,00013		-107,7	0,303	96%	4,02
-4,628	-0,148	-0,001		0	-110,1	-2,362	96%	4,09

Tabla 158: VMV modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
16,720					-241,8		-901%	67,38
22,589	0,357				-233,6	8,187	-605%	56,53
22,621	0,358	0,039			-198,6	35,057	-58%	26,74
-1,507	3,164	0,018	-0,014		-102,7	95,894	98%	3,01
-1,507	3,164	0,018	-0,014	0,00000001	-102,7	0,017	98%	3,05

Se puede comprobar cómo bajo el supuesto de que los errores se distribuyen de acuerdo a una distribución logística, el valor de máxima verosimilitud es máximo, siendo prácticamente idéntico al de la distribución doble exponencial. Para la simulación de los costes de recogida del vehículo de 1500 kg. para

este código postal se empleará el modelo de regresión calculado bajo el supuesto de que los errores se distribuyen de acuerdo a una distribución de probabilidad logística.

2.3.5.1. Selección de variables para modelo de costes de recogida

El cuadro obtenido en el apartado anterior para la distribución logística se puede observar como el cociente de verosimilitud es mayor de dos unidades al introducir las variables número de recogidas, peso total recogido y número de recogidas elevada a la segunda potencia, al introducirlas en el modelo una a una. Sin embargo, al introducir la variable peso total recogido elevado a la segunda potencia, el cociente de verosimilitud sólo se diferencia en 0,017 unidades, por lo que la incorporación de esta variable al modelo no aporta un mejor ajuste a los datos. Por tanto, este modelo de recogidas incluirá únicamente las tres primeras variables mencionadas, con los parámetros de regresión 3,164, 0,018 y -0,014, respectivamente, con un valor umbral igual a -1,507.

2.3.5.2. Adecuación del modelo

Cuando los errores se distribuyen con una distribución de probabilidad no normal, si éstos son independientes y se distribuyen idénticamente, las inferencias bajo la teoría de normalidad son asintóticamente correctas. Por tanto, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. El análisis gráfico se puede ver en la figura 70.

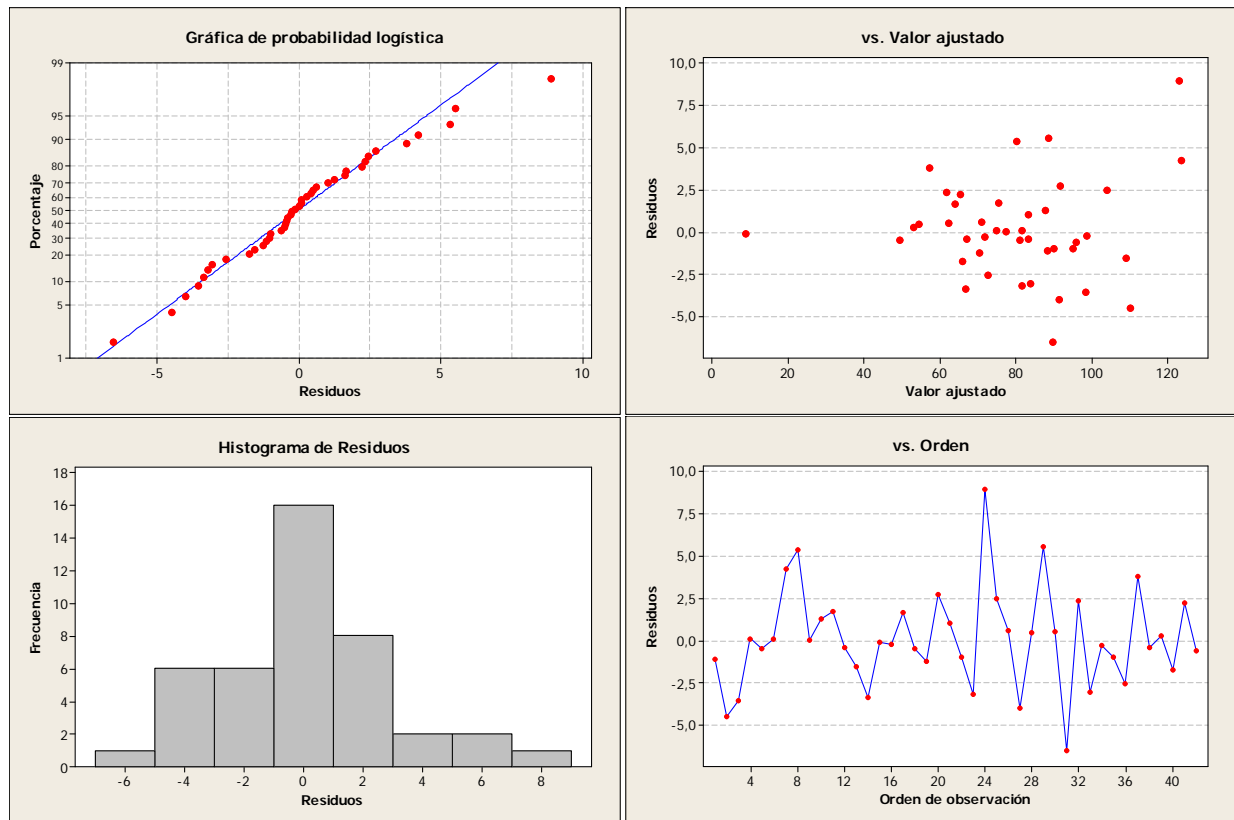


Figura 70: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08196

En la primera gráfica de la figura 70 se puede observar como los errores se distribuyen según una distribución logística ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.3.5.3. Validación del modelo de costes de recogida

Para la validación del modelo se va a emplear la muestra de predicción obtenida anteriormente en el apartado 2.3.1.3. Con esta muestra se va a realizar una prueba *t* pareada, comparando la respuesta dada por el modelo con los costes de recogida de la muestra de predicción. El cuadro que resume esta prueba *t* pareada es el de la tabla 159.

Tabla 159: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08196

	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	78,83	78,47	0,369	Límite superior I.C	$t_{0,025;39}$
Varianza	443,03	480,62	7,83	1,215	1,96
Varianza/n			0,186		
Máximo	128,42	128,89		Límite inferior I.C	
Mínimo	27,14	27,58		-0,478	

Como el intervalo de esta prueba incluye el cero, se concluye que el modelo predice los datos de manera adecuada, considerándose válido. Los resultados de la diferencia entre los valores estimados por el modelo y los de la muestra de predicción para confeccionar el cuadro anterior se obtuvieron de la tabla 160.

Tabla 160: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08196

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
-1,507	3,164	0,018	-0,014			
	15	2343	225	84,98	86,87	-1,90
	25	3310	625	128,42	128,89	-0,46
	16	1836	256	78,58	79,37	-0,79
	21	1469	441	85,21	86,21	-1,00
	17	1488	289	75,02	79,39	-4,38
	17	2951	289	101,35	104,22	-2,87
	17	1860	289	81,72	82,19	-0,48
	21	1873	441	92,48	94,14	-1,67
	12	2095	144	72,16	69,38	2,77
	21	1791	441	91,00	96,73	-5,73
	19	2839	361	104,66	104,93	-0,28
	20	2494	400	101,07	99,11	1,96
	13	1292	169	60,52	58,41	2,11
	18	2333	324	92,90	89,36	3,54
	8	996	64	40,84	44,70	-3,86
	11	1112	121	51,62	48,95	2,67
	20	2458	400	100,42	103,53	-3,11
	15	1347	225	67,05	60,92	6,13
	17	2136	289	86,68	86,14	0,54
	11	2280	121	72,64	68,31	4,34
	20	1690	400	86,59	87,00	-0,41
	4	901	16	27,14	27,58	-0,44
	16	2009	256	81,70	80,27	1,42
	14	2069	196	77,29	75,89	1,40
	20	2023	400	92,59	92,96	-0,37
	17	2072	289	85,53	85,09	0,44
	16	1974	256	81,07	79,68	1,38
	22	2826	484	112,19	115,52	-3,33
	17	2826	289	99,10	103,68	-4,57
	15	2264	225	83,56	83,68	-0,12
	9	1102	81	45,67	42,95	2,72
	18	2482	324	95,59	93,98	1,60
	16	1032	256	64,11	57,03	7,08
	13	883	169	53,15	51,50	1,65
	17	1376	289	73,00	70,26	2,74
	15	1478	225	69,41	72,00	-2,59
	14	1210	196	61,83	61,37	0,45
	17	3102	289	104,07	102,48	1,59
	14	1732	196	71,22	68,25	2,97
	7	725	49	33,01	33,79	-0,78
	14	2047	196	76,89	73,95	2,94
	11	1969	121	67,05	64,89	2,15

2.3.6. Modelo de carga de trabajo

En este modelo se van a estudiar la distribución de probabilidad que siguen las variables aleatorias que conforman la carga la de trabajo, así como sus parámetros. Al igual que para el resto de vehículos de los códigos postales anteriores estas variables son el número de repartos por día y el peso de cada uno de los repartos, en el caso de los repartos, y el número de recogidas por día y el peso de cada recogida, en el de las recogidas.

Las muestras que se emplean para las variables número de repartos, peso total repartido, número de recogidas y peso total recogido son las mismas muestras que las muestras de calibración que se emplearon para ajustar los modelos de los costes de reparto y recogida. Las muestras para las variables peso de cada reparto y peso de cada recogida, es una muestra con los pesos unitarios de cada uno de los repartos y de cada una de las recogidas.

2.3.6.1. Número de repartos por día

A través de Minitab se obtienen los parámetros para cada una de las distribuciones de probabilidad que hacen que éstas sean las que mejor se ajusten a los valores de la muestra. Esta variable aleatoria es un conteo de algún elemento, y por tanto, una variable discreta. Al ser una variable discreta, la primera distribución de probabilidad que se comprueba es la distribución de Poisson. El test no paramétrico con el que se comprueba si el ajuste para las distribuciones discretas es adecuado es el test Chi cuadrado, el cual se puede ver en la tabla 161.

Tabla 161: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08196

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
9-19	7	0,163	0,1685	7	0,0082	alfa=0,05 k=5 Par. Est. =1 n=5-1-1=3
20-22	9	0,209	0,1454	6	1,2086	
23-26	11	0,256	0,2329	10	0,0973	
27-30	10	0,233	0,1428	6	2,4256	
31-35	6	0,140	0,0745	3	2,4392	
	43				6,18	

En la tabla se puede observar como el valor límite para aceptar la hipótesis de que la una distribución de Poisson con lambda 24,23 ajusta adecuadamente la variable número de repartos por día es 7,85. Este valor corresponde a una distribución Chi cuadrado con 3 grados de libertad y nivel de significancia del 5%. Los grados de libertad es igual al número de intervalos de los que consta el test, al que hay que restar el número de parámetros estimados (para el caso de la distribución de Poisson es únicamente el parámetro lambda) y otra unidad. Como en este caso el valor del test es 6,18, menor al valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.3.6.2. Peso de cada reparto

La segunda variable a analizar para el modelo de carga de trabajo es el peso de cada reparto. Esta variable es una variable aleatoria continua, y al igual que para el resto de variables continuas, las distribuciones de probabilidad a analizar serán también continuas. Los valores dados por Minitab y los resultados del test Kolmogorov-Smirnov para cada una de las distribuciones analizadas se pueden ver en la tabla 162.

Tabla 162: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08196

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,239	alfa	0,090	alfa	0,065	Media	0,045	Lambda	0,245
37,95	5,250	0,6945	1,985	0,7692	1,432	2,766	0,999	27,29	5,391
Desv. Est.	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	Desv. Est.	$c_{\alpha} = 1,358$	Theta	$c_{\alpha} = 1,358$
52,47	$\alpha = 0,05$	54,57	$\alpha = 0,05$	32,02	$\alpha = 0,05$	1,417	$\alpha = 0,05$	23,59	$\alpha = 0,05$

En la tabla 162 se puede ver como las únicas distribuciones que ajustan satisfactoriamente los datos de la muestra de la variable son las distribuciones de probabilidad Weibull y lognormal, siendo la distribución lognormal la que menos discrepancia presenta. Por lo tanto, para la simulación se empleará esta distribución para estimar los pesos de cada uno de las recogidas.

2.3.6.3. Número de recogidas por día

Al igual que para la variable número de repartos por día, esta variable es un conteo de algún elemento. Por ello, la primera distribución de probabilidad que se va a comprobar si ajusta adecuadamente a los valores de la muestra de esta variable es la distribución de probabilidad discreta de Poisson. Al ser una distribución discreta, el test no paramétrico a realizar es el test Chi cuadrado. Los resultados de este test se pueden ver en la tabla 163.

Tabla 163: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08196

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
3-12	7	0,167	0,1793	8	0,0376	alfa=0,05 k=5 Est. par. =1 n=5-1-1=3
13-15	9	0,214	0,2663	11	0,4268	
16-17	10	0,238	0,1936	8	0,4301	
18-20	10	0,238	0,2167	9	0,0891	
21-25	6	0,143	0,0987	4	0,8314	
	42				1,81	7,85

El valor lambda dado por Minitab para una distribución de Poisson con los datos de la muestra de esta variable aleatoria es 16,21. En la tabla se puede observar como el valor límite para aceptar la hipótesis de que la una distribución de Poisson con lambda 16,21 ajusta adecuadamente la variable número de repartos por día es 7,85. Este valor corresponde a una distribución Chi cuadrado con 3 grados de libertad y nivel de significancia del 5%. Como en este caso el valor del test es 1,81, menor al valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.3.6.4. Peso de cada recogida

La última variable a analizar en la carga de trabajo es el peso de cada recogida. Esta variable aleatoria es continua, y se determinan los valores de cada distribución, así como la bondad de ajuste de las mismas de la misma manera que las variables continuas anteriores. Los resultados son los de la tabla 164.

Tabla 164: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08196

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,234	alfa	0,068	alfa	0,058	Media	0,029	Lambda	0,256
115,8	4,489	0,5108	1,300	0,6294	1,109	3,513	0,556	84,82	4,906
Dev. Est.	c _α = 1,358	beta	c _α = 1,358	beta	c _α = 1,358	Dev. Est.	c _α = 1,358	Theta	c _α = 1,358
160,3	α = 0,05	226,7	α = 0,05	81,86	α = 0,05	1,825	α = 0,05	79,63	α = 0,05

En la tabla 164 se puede ver como las únicas distribuciones que ajustan satisfactoriamente los datos de la muestra de la variable son las distribuciones de probabilidad gamma, Weibull y lognormal, siendo la distribución lognormal la que menos discrepancia presenta. Por lo tanto, para la simulación se empleará esta distribución para estimar los pesos de cada uno de las recogidas.

2.3.7. Simulación de los costes de reparto y recogida, y de la carga de trabajo

En este apartado se llevará a cabo la simulación de los modelos creados para el estudio del vehículo de 1500 kg. de peso máximo autorizado en el código postal 08198. En primer lugar se simulará la carga de trabajo. Para ello hay que determinar el número de repartos y de recogidas para cada uno de los días simulados, y el peso total repartido y el peso total recogido por día. Empleando los ajustes de las distribuciones de probabilidad de Poisson obtenidos en los apartados 2.3.6.1 y 2.3.6.3. se simularán el número de repartos y el número de recogidas para cada uno días que se desean simular.

Posteriormente, se simulará el peso correspondiente de cada uno de los repartos y las recogidas mediante las distribuciones de probabilidad calculadas en los apartados 2.3.6.2. y 2.3.6.4. Sumando el peso unitario de todos los repartos simulados en cada uno de los días, se calculará el peso total repartido por día. Con el peso unitario de las recogidas simulado para cada uno de los días, se calculará el peso total recogido por día.

Finalmente, se simularán los modelos de costes de reparto y costes de recogida. Como valores de entrada a estos modelos se emplearán los datos simulados para el modelo de carga de trabajo. El número de repartos y el peso total repartido por día serán los datos de entrada para el modelo de costes de reparto. Análogamente, el número de recogidas y el peso total recogido por día serán los valores de entrada para el modelo de costes de recogida.

2.3.7.1. Simulación de los costes de reparto

La carga de trabajo relativa a los repartos para el vehículo de 1500 kg. de peso máximo autorizado del código postal 08196 es la siguiente: el número de repartos puede ser descrito mediante una distribución de probabilidad Poisson con parámetro lambda igual a 24,23. Mientras que el de cada reparto sigue una distribución de probabilidad lognormal con valor medio igual a 2,766 y desviación estándar 1,417.

Un resumen de la carga de trabajo simulada es la de la tabla 165, en el cual se puede observar el promedio de repartos por día simulados, su desviación estándar, el número máximo de repartos a realizar en un día, así como el número mínimo. Y los mismos datos para la variable peso total repartido.

Tabla 165: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08196

	Repartos por día	Peso total repartido por día
Promedio	24	735
Desv. Estan.	5	251
Máximo	39	1742
Mínimo	13	241

La simulación de la carga de trabajo en 250 días bajo las condiciones previamente enunciadas, son las entradas para la simulación en el modelo de costes de reparto. Para cada uno de los días se obtiene un estimado puntual del coste de reparto para la carga de trabajo simulada correspondiente. A partir de cada estimado puntual se desarrolla un intervalo de predicción, teniendo un total de 250 intervalos de predicción. Estos intervalos de predicción se calculan tal y como se explicó en el apartado 4.1.1.3 de la metodología. Los resultados que aquí se exponen son el valor medio de límite inferior de las 250 simulaciones, la varianza y desviación estándar de este límite inferior, así como un intervalo de confianza de este límite inferior. Para el límite superior del intervalo de predicción los resultados que se presentan son análogos. Estos resultados son los de la tabla 166.

Tabla 166: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08196

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	70,09	79,42
Varianza	249	249
Desv. Estan.	15,78	15,79
I.C	68,44	77,78
	71,74	81,07

En la tabla 166 se puede observar como el límite inferior de los estimados puntuales de los costes de repartos se distribuyen con un valor medio de 70,09 euros por día y una desviación estándar de 15,78. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 79,42 y una desviación estándar de 15,79 euros por día.

2.3.7.2. Simulación de los costes de recogida

La carga de trabajo relativa a las recogidas para este código postal queda como sigue: el número de recogidas puede ser expresado mediante una distribución de probabilidad Poisson con parámetro lambda igual a 16,21. El de cada recogida, por otro lado, puede ser descrito mediante una distribución de probabilidad lognormal con valor medio igual a 3,513 y desviación estándar 1,825.

Un cuadro resumen de la carga de trabajo simulada es la tabla 167, en el cual se puede observar el promedio de recogidas por día simulados, su desviación estándar, el número máximo de recogidas a realizar en un día, así como el número mínimo. Y los datos análogos para la variable peso total recogido.

Tabla 167: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08196

	Recogidas por día	Peso total recogido por día
Promedio	16	1294
Desv. Estan.	4	588
Máximo	29	4157
Mínimo	6	279

Para la simulación de los costes de recogida, los resultados que se exponen son idénticos al caso de los costes de reparto. Se exponen los valores medios, varianza y desviación estándar, así como un intervalo de confianza de los límites superior e inferior de los límites de predicción de los 250 días simulados a partir de la carga de trabajo previamente simulada bajo las condiciones expuestas en el primer párrafo de este subapartado. Los resultados son los de la tabla 168.

Tabla 168: Resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08196

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	62,52	74,89
Varianza	349	351
Desv. Estan.	18,68	18,73
I.C	60,57	72,93
	64,47	76,84

En la tabla 168 se puede observar como el límite inferior de los estimados puntuales de los costes de repartos se distribuyen con un valor medio de 62,52 euros por día y una desviación estándar de 18,68. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 74,89 y una desviación estándar de 18,73 euros por día.

2.4. Código postal 08195

En este código postal no se dan repartos superiores a 300 kg. de peso, ni recogidas superiores a 700 kg., por lo que en él solo existe volumen de trabajo para vehículos de 1500 kg. de peso máximo autorizado.

Los pasos a seguir en la elaboración del modelo de costes de la red de distribución capilar de este código postal son análogos a los empleados en el código postal anterior. Tal y como se llevó a cabo en los tres tipos de vehículos del código postal anterior, el primer paso en la elaboración del modelo es la toma de muestras.

2.4.1. Toma de muestras

Se toman muestras de aquellas variables que se identificaron como las determinantes para el buen funcionamiento de la red de distribución capilar. Estas son el coste de reparto por día y el coste de recogida por día. Éstas a su vez dependen del número de repartos por día y del peso repartido por día, y del número de recogidas por día y del peso total repartido por día respectivamente. Esta toma de muestras se ha llevado a cabo mediante un estudio observacional, y se han obtenido 84 observaciones para cada una de ellas. Las muestras obtenidas para las variables relacionadas con los repartos se pueden observar en la tabla 169, mientras que en la tabla 170 se pueden ver las relacionadas con las recogidas.

Tabla 169: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08195

Día	REPARTOS			Día	REPARTOS		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	10	344	33	43	19	580	60
2	22	653	68	44	11	661	45
3	8	263	25	45	13	513	45
4	17	192	42	46	7	75	17
5	17	655	56	47	13	291	36
6	11	319	33	48	18	482	53
7	12	195	32	49	14	760	50
8	16	513	50	50	14	179	36
9	11	90	25	51	19	511	59
10	10	315	33	52	22	314	58
11	20	575	62	53	17	316	47
12	11	506	37	54	8	92	20
13	10	472	36	55	13	459	44
14	12	182	32	56	15	181	37
15	10	247	30	57	11	82	25
16	7	110	18	58	10	157	26
17	9	243	27	59	11	164	29
18	12	186	31	60	14	264	39
19	18	641	58	61	19	583	56
20	13	514	43	62	19	912	69
21	15	254	40	63	9	175	25
22	11	552	38	64	10	477	33
23	1	7	2	65	14	502	47
24	26	986	86	66	13	344	40
25	17	455	50	67	11	328	34
26	9	480	36	68	12	216	33
27	5	66	13	69	8	197	23
28	21	599	65	70	14	602	47
29	19	601	59	71	16	678	55
30	16	268	44	72	10	661	38
31	9	123	23	73	8	76	19
32	9	128	24	74	14	499	47
33	16	544	50	75	12	326	35
34	22	923	72	76	15	854	55
35	16	428	51	77	15	813	55
36	6	106	16	78	12	370	38
37	17	375	50	79	7	78	18
38	15	323	43	80	6	165	16
39	15	346	42	81	10	545	36
40	13	461	44	82	10	293	32
41	10	705	38	83	6	230	20
42	9	202	25	84	8	470	29

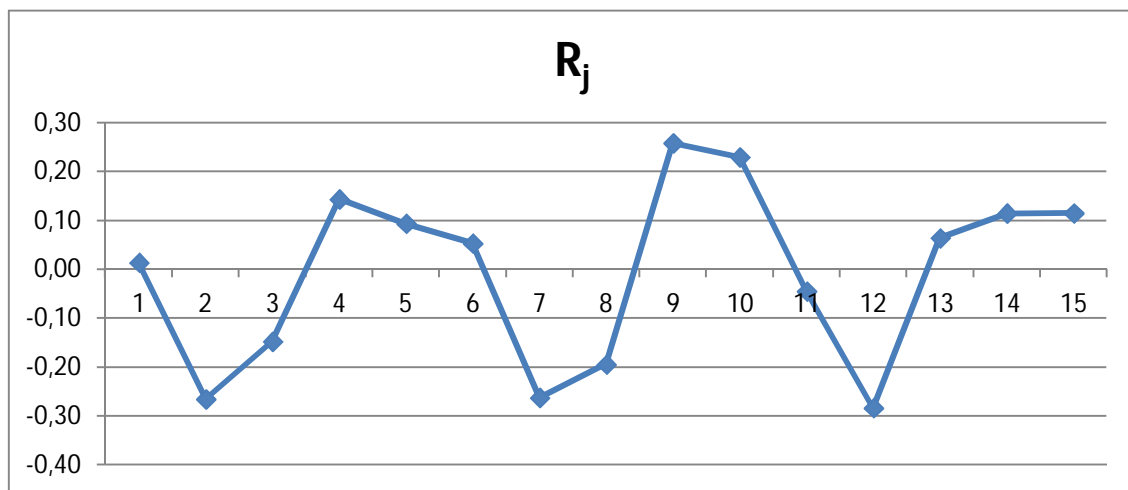
Tabla 170: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08195

Día	RECOGIDAS			Día	RECOGIDAS		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	5	533	23	43	6	792	31
2	4	732	26	44	0	0	0
3	2	133	10	45	4	525	21
4	2	298	10	46	3	126	12
5	5	307	20	47	1	331	9
6	2	339	13	48	2	341	11
7	4	455	20	49	6	713	32
8	5	239	19	50	4	267	19
9	4	338	18	51	6	257	22
10	1	13	3	52	7	664	33
11	4	1.402	36	53	4	261	18
12	4	381	17	54	5	359	22
13	3	248	15	55	4	275	17
14	3	154	12	56	4	374	18
15	3	105	12	57	3	268	14
16	5	354	21	58	5	154	17
17	4	1.004	29	59	7	1.041	41
18	5	444	23	60	5	715	27
19	4	1.290	35	61	3	563	20
20	3	306	16	62	4	224	17
21	6	860	31	63	7	526	31
22	3	580	20	64	3	258	15
23	3	309	14	65	6	939	36
24	3	212	13	66	3	60	9
25	4	193	17	67	3	711	22
26	3	230	15	68	4	61	11
27	2	139	8	69	4	848	28
28	2	207	10	70	5	369	23
29	2	124	9	71	3	195	14
30	6	518	26	72	3	391	18
31	6	785	32	73	4	772	25
32	4	353	20	74	3	336	16
33	2	737	19	75	3	747	21
34	2	161	9	76	5	522	23
35	4	602	23	77	3	242	15
36	4	279	17	78	3	598	20
37	4	552	21	79	2	181	10
38	6	635	28	80	3	243	15
39	4	540	22	81	1	39	4
40	3	265	14	82	2	323	11
41	5	291	20	83	4	239	17
42	3	115	12	84	2	191	10

2.4.2. Independencia de la muestras

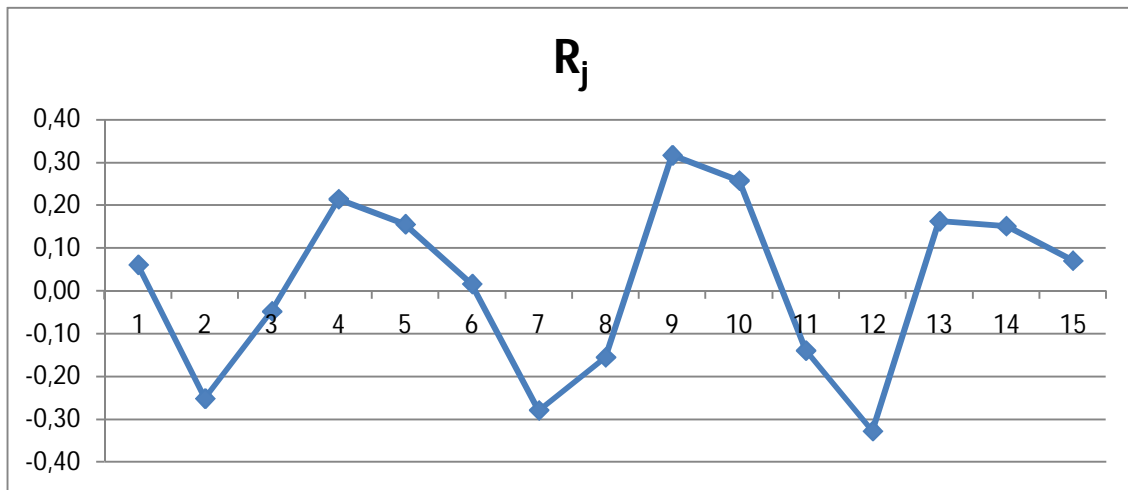
Tras la obtención de las muestras, hay que comprobar si las observaciones de de las mismas son independientes. Que las observaciones sean independientes es fundamental, ya que para emplear el método de máxima verosimilitud en la elaboración de los modelos de costes de reparto y recogida, las muestras utilizadas han de ser independientes. Por otro lado, los test no paramétricos de bondad de ajuste de las distribuciones de probabilidad también presuponen que las muestras empleadas son independientes.

En primer lugar se analiza la independencia de la muestras de los repartos. Como esta muestra se compone de tres variables, coste de reparto por día, número de repartos por día y peso total repartido por día, es necesario comprobar la independencia de cada una de ellas. En primer lugar se muestran las gráficas de autocorrelación.



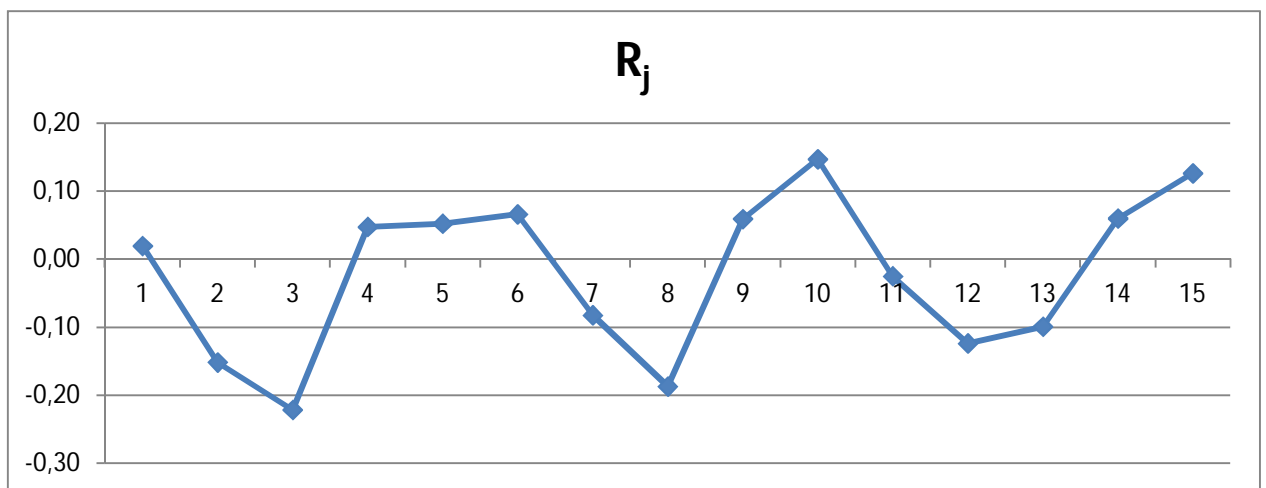
R_j	0,01	-0,27	-0,15	0,14	0,09	0,05	-0,26	-0,19	0,26	0,23	-0,04	-0,28	0,06	0,11	0,11
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	3	-62	-34	33	22	12	-61	-45	60	53	-10	-66	15	27	27

Figura 71: Grafica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08195



R_j	0,06	-0,25	-0,05	0,21	0,16	0,02	-0,28	-0,15	0,32	0,26	-0,14	-0,33	0,16	0,15	0,07
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	1,3	-5,2	-1,0	4,5	3,2	0,3	-5,8	-3,2	6,6	5,3	-2,9	-6,8	3,4	3,1	1,5

Figura 72: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08195



R_j	0,02	-0,15	-0,22	0,05	0,05	0,07	-0,08	-0,19	0,06	0,15	-0,03	-0,12	-0,10	0,06	0,13
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	1006	-7869	-11497	2454	2699	3425	-4276	-9712	3064	7619	-1303	-6423	-5149	3102	6554

Figura 73: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08195

En las tres gráficas se puede observar como en ninguna de ellas los valores de R_j se encuentran en el rango de -1 y 1, y que en ninguna de ellas se puede apreciar un patrón determinado. Por tanto, esta prueba indica la independencia de las observaciones de las muestras.

Tras las gráficas de autocorrelación, se muestran a continuación las gráficas de dispersión para cada una de las variables.

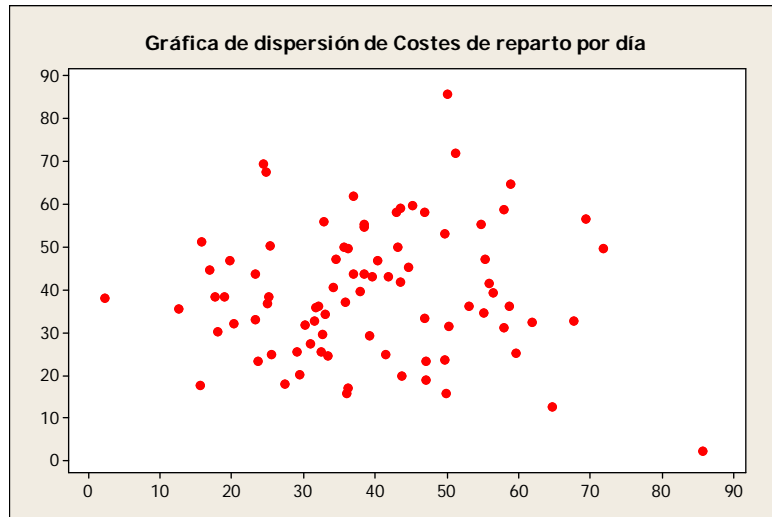


Figura 74: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08195

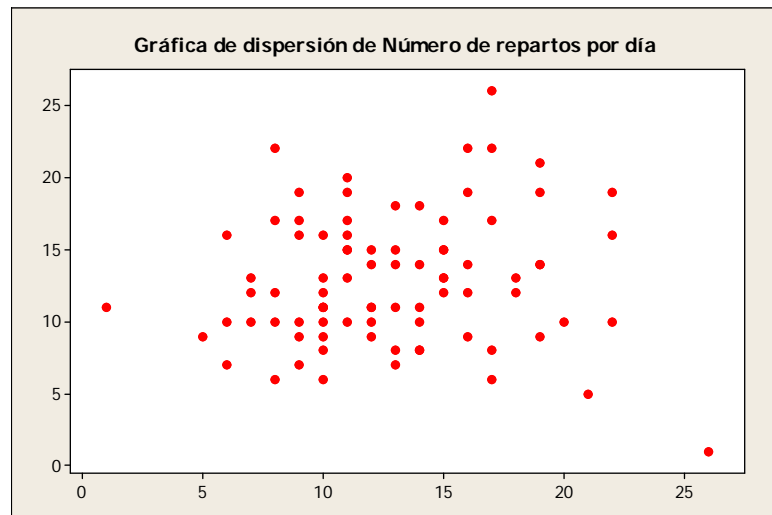


Figura 75: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08195

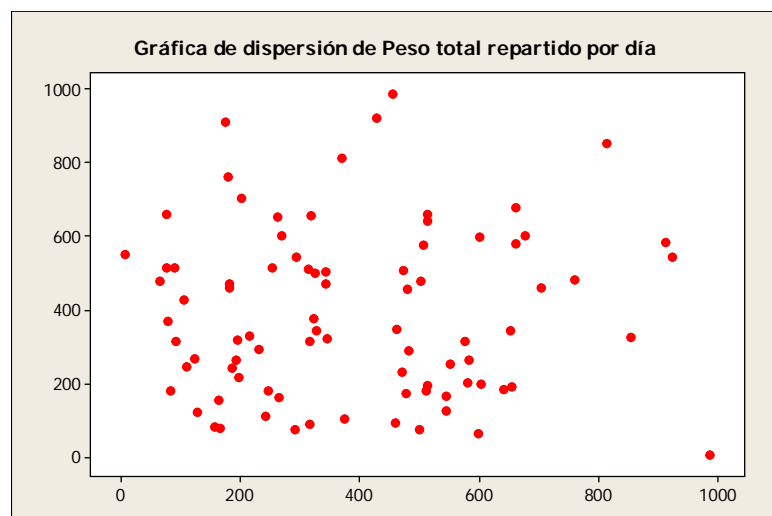
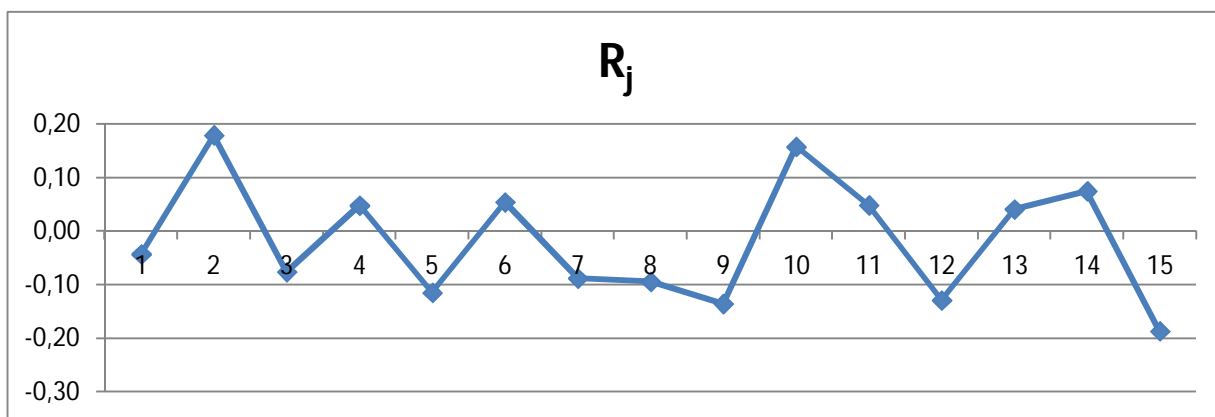


Figura 76: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08195

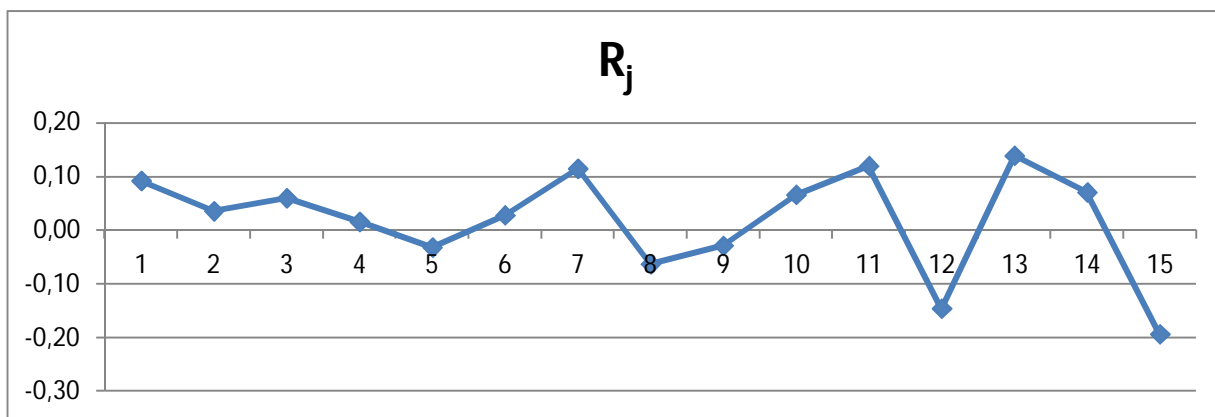
Como en las tres gráficas los puntos se distribuyen de manera aleatoria a lo largo de todo el plano, esta prueba también concluye que las observaciones son independientes.

Tras comprobar que las muestras relativas a los repartos son independientes, hay que analizar si las muestras correspondientes a las recogidas también son independientes. Al igual que para los repartos se comenzará mostrando las gráficas de autocorrelación.



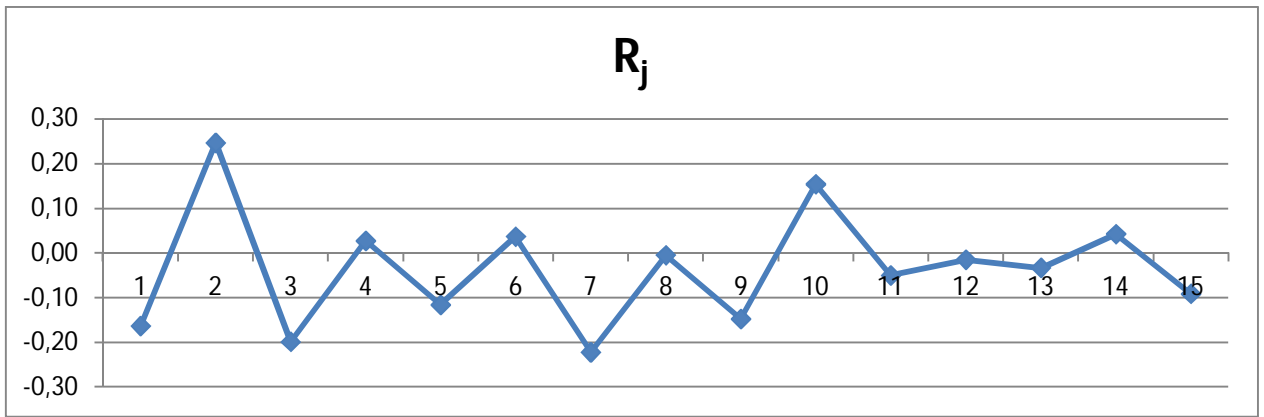
R_j	-0,04	0,18	-0,08	0,05	-0,11	0,05	-0,09	-0,09	-0,14	0,16	0,05	-0,13	0,04	0,07	-0,19
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-3	11	-5	3	-7	3	-6	-6	-9	10	3	-8	3	5	-12

Figura 77: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08195



R_j	0,09	0,04	0,06	0,02	-0,03	0,03	0,11	-0,06	-0,03	0,07	0,12	-0,15	0,14	0,07	-0,19
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	0,2	0,1	0,1	0,0	-0,1	0,1	0,2	-0,1	-0,1	0,1	0,3	-0,3	0,3	0,2	-0,4

Figura 78: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08195



R_j	-0,16	0,25	-0,20	0,03	-0,12	0,04	-0,22	0,00	-0,15	0,15	-0,05	-0,01	-0,03	0,04	-0,09
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-13167	19862	-16017	2185	-9340	2953	-17905	-400	-11884	12396	-4031	-1207	-2683	3413	-7361

Figura 79: Gráfica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08195

En las tres gráficas se puede ver como los valores de R_j no superan los valores 1 y -1, y que no se da un patrón determinado en ninguna de ellas por lo que se concluye que las observaciones de cada muestra son independientes. Tras estudiar las gráficas de autocorrelación, a continuación se verifica la independencia de las observaciones mediante las gráficas de dispersión.

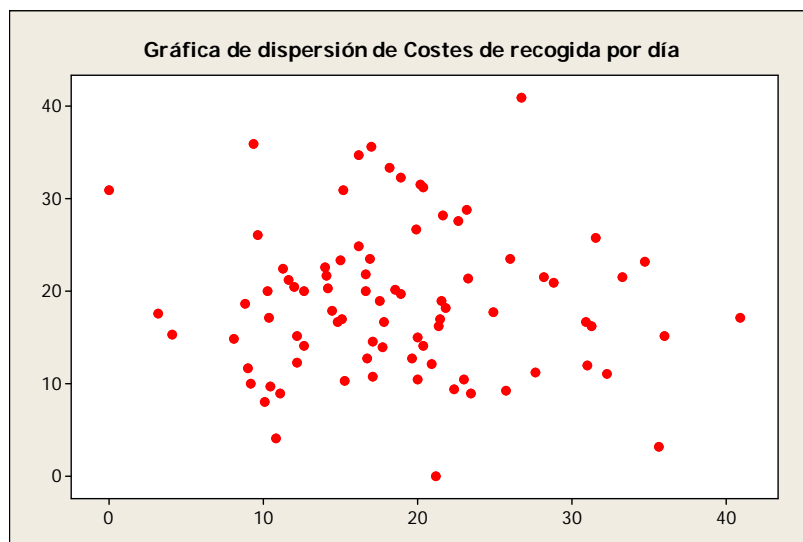


Figura 80: Gráfica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08195

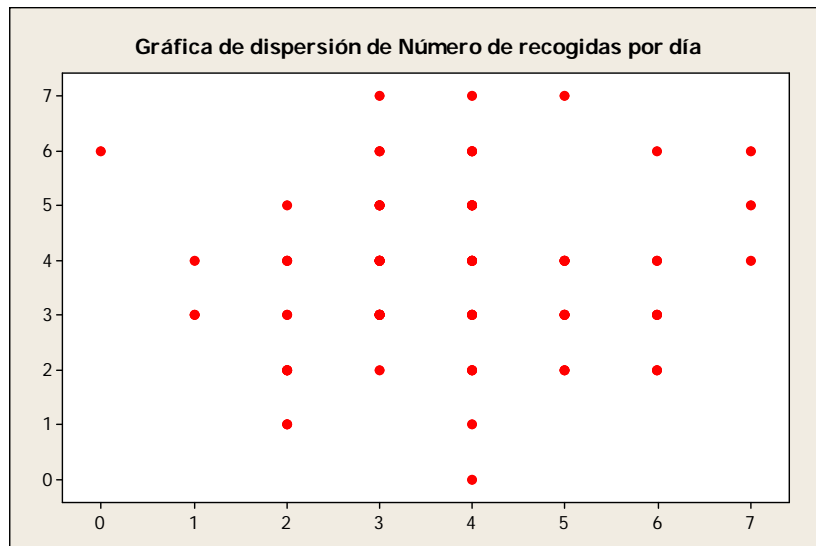


Figura 81: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08195

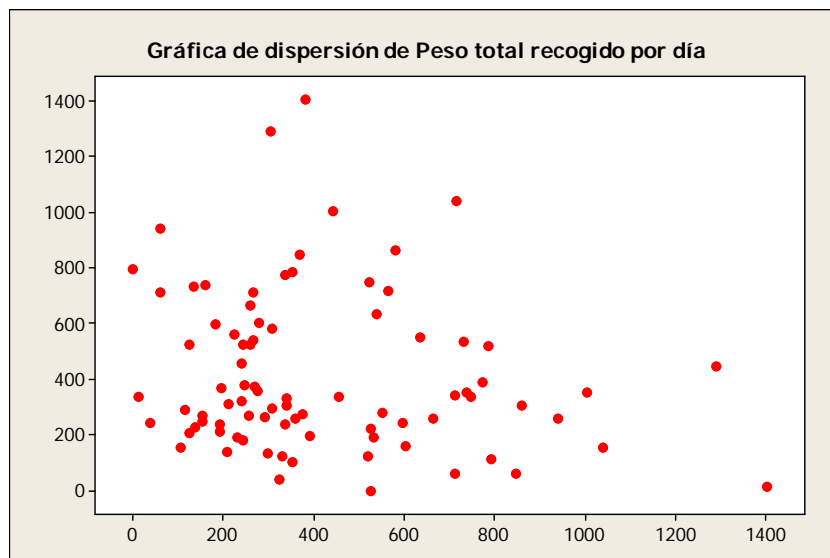


Figura 82: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08195

En las tres gráficas se puede observar como todos los puntos se distribuyen de manera aleatoria en los planos correspondientes, pudiéndose concluir que las observaciones de las muestras son independientes.

2.4.3. Separación de las muestras de calibración y predicción

Tras verificar que las 84 observaciones obtenidas del estudio observacional de cada una de las variables son independientes, hay que separar estas 84 observaciones en dos muestras diferentes, la de calibración y la de predicción. La de calibración se emplea para la creación y elaboración de los modelos necesarios para el estudio de los costes de la red de distribución, mientras que la de predicción se emplea para comprobar que los modelo elaborados mediante el análisis de regresión estiman los costes de reparto

y recogida de manera adecuada. Esta separación se lleva a cabo mediante una asignación de un número aleatorio a cada una de las observaciones, y si éste es mayor que el valor 0,5, se asigna a la muestra de calibración. Si por el contrario, este número aleatorio es menor que 0,5, se asigna a la muestra de predicción. El listado de números aleatorios asignados a cada una de las observaciones relativas al reparto es el de la tabla 171.

Tabla 171: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08195

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,229	22	0,903	43	0,622	64	0,556
2	0,449	23	0,205	44	0,587	65	0,812
3	0,316	24	0,157	45	0,959	66	0,339
4	0,697	25	0,287	46	0,376	67	0,298
5	0,488	26	0,062	47	0,185	68	0,599
6	0,994	27	0,179	48	0,360	69	0,130
7	0,754	28	0,304	49	0,811	70	0,816
8	0,846	29	0,508	50	0,175	71	0,816
9	0,273	30	0,333	51	0,765	72	0,665
10	0,595	31	0,400	52	0,760	73	0,863
11	0,849	32	0,046	53	0,773	74	0,462
12	0,340	33	0,266	54	0,918	75	0,416
13	0,625	34	0,378	55	0,398	76	0,237
14	0,006	35	0,657	56	0,942	77	0,908
15	0,879	36	0,833	57	0,197	78	0,219
16	0,681	37	0,715	58	0,548	79	0,286
17	0,737	38	0,840	59	0,930	80	0,331
18	0,649	39	0,901	60	0,764	81	0,656
19	0,413	40	0,087	61	0,027	82	0,010
20	0,365	41	0,930	62	0,556	83	0,081
21	0,827	42	0,632	63	0,292	84	0,697

Quedando las muestras de calibración y predicción de acuerdo a esta asignación como se muestra en la tabla 172.

Tabla 172: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08195

Día	REPARTOS. MUESTRA DE CALIBRACIÓN			Día	REPARTOS. MUESTRA DE PREDICCIÓN		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	17	192	42	1	10	344	33
2	11	319	33	2	22	653	68
3	12	195	32	3	8	263	25
4	16	513	50	4	17	655	56
5	10	315	33	5	11	90	25
6	20	575	62	6	11	506	37
7	10	472	36	7	12	182	32
8	10	247	30	8	18	641	58
9	7	110	18	9	13	514	43
10	9	243	27	10	1	7	2
11	12	186	31	11	26	986	86
12	15	254	40	12	17	455	50
13	11	552	38	13	9	480	36
14	19	601	59	14	5	66	13
15	16	428	51	15	21	599	65
16	6	106	16	16	16	268	44
17	17	375	50	17	9	123	23
18	15	323	43	18	9	128	24
19	15	346	42	19	16	544	50
20	10	705	38	20	22	923	72
21	9	202	25	21	13	461	44
22	19	580	60	22	7	75	17
23	11	661	45	23	13	291	36
24	13	513	45	24	18	482	53
25	14	760	50	25	14	179	36
26	19	511	59	26	13	459	44
27	22	314	58	27	11	82	25
28	17	316	47	28	19	583	56
29	8	92	20	29	9	175	25
30	15	181	37	30	13	344	40
31	10	157	26	31	11	328	34
32	11	164	29	32	8	197	23
33	14	264	39	33	14	499	47
34	19	912	69	34	12	326	35
35	10	477	33	35	15	854	55
36	14	502	47	36	12	370	38
37	12	216	33	37	7	78	18
38	14	602	47	38	6	165	16
39	16	678	55	39	10	293	32
40	10	661	38	40	6	230	20
41	8	76	19				
42	15	813	55				
43	10	545	36				
44	8	470	29				

En este caso, debido a la asignación aleatoria, se tienen 44 observaciones para la muestra de calibración y 40 observaciones para la muestra de predicción relativas a los repartos. Una vez elaboradas estas muestras para los repartos, se realiza un proceso análogo para las recogidas. La asignación de números aleatorios a cada una de las 84 observaciones relativas a las recogidas es la de la tabla 173.

Tabla 173: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08195

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,652	22	0,985	43	0,145	64	0,409
2	0,028	23	0,478	44	0,316	65	0,827
3	0,749	24	0,044	45	0,683	66	0,837
4	0,830	25	0,323	46	0,062	67	0,917
5	0,840	26	0,426	47	0,509	68	0,195
6	0,887	27	0,727	48	0,257	69	0,633
7	0,771	28	0,589	49	0,038	70	0,658
8	0,093	29	0,292	50	0,621	71	0,342
9	0,258	30	0,619	51	0,128	72	0,308
10	0,065	31	0,031	52	0,897	73	0,567
11	0,952	32	0,945	53	0,726	74	0,043
12	0,463	33	0,972	54	0,799	75	0,357
13	0,859	34	0,950	55	0,443	76	0,102
14	0,010	35	0,929	56	0,862	77	0,929
15	0,590	36	0,035	57	0,549	78	0,048
16	0,224	37	0,874	58	0,402	79	0,436
17	0,539	38	0,760	59	0,331	80	0,806
18	0,823	39	0,711	60	0,823	81	0,311
19	0,277	40	0,306	61	0,364	82	0,032
20	0,337	41	0,678	62	0,020	83	0,635
21	0,491	42	0,140	63	0,418	84	0,392

A partir de este listado de números aleatorios se confeccionan las muestras de calibración y predicción relativas a las recogidas. Estas muestras se pueden observar en la tabla 174.

Tabla 174: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08195

Día	RECOGIDAS. MUESTRA DE CALIBRACIÓN			Día	RECOGIDAS. MUESTRA DE PREDICCIÓN		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	5	533	23	1	4	732	26
2	2	133	10	2	5	239	19
3	2	298	10	3	4	338	18
4	5	307	20	4	1	13	3
5	2	339	13	5	4	381	17
6	4	455	20	6	3	154	12
7	4	1.402	36	7	5	354	21
8	3	248	15	8	4	1.290	35
9	3	105	12	9	3	306	16
10	4	1.004	29	10	6	860	31
11	5	444	23	11	3	309	14
12	3	580	20	12	3	212	13
13	2	139	8	13	4	193	17
14	2	207	10	14	3	230	15
15	6	518	26	15	2	124	9
16	4	353	20	16	6	785	32
17	2	737	19	17	4	279	17
18	2	161	9	18	3	265	14
19	4	602	23	19	3	115	12
20	4	552	21	20	6	792	31
21	6	635	28	21	0	0	0
22	4	540	22	22	3	126	12
23	5	291	20	23	2	341	11
24	4	525	21	24	6	713	32
25	1	331	9	25	6	257	22
26	4	267	19	26	4	275	17
27	7	664	33	27	5	154	17
28	4	261	18	28	7	1.041	41
29	5	359	22	29	3	563	20
30	4	374	18	30	4	224	17
31	3	268	14	31	7	526	31
32	5	715	27	32	3	258	15
33	6	939	36	33	4	61	11
34	3	60	9	34	3	195	14
35	3	711	22	35	3	391	18
36	4	848	28	36	3	336	16
37	5	369	23	37	3	747	21
38	4	772	25	38	5	522	23
39	3	242	15	39	3	598	20
40	3	243	15	40	2	181	10
41	4	239	17	41	1	39	4
				42	2	323	11
				43	2	191	10

2.4.4. Modelo de costes de reparto

Como para el resto de modelos de costes de reparto elaborados de este trabajo de investigación, los parámetros de regresión se calculan empleando el método de máxima verosimilitud. Se estudiará el valor de máxima verosimilitud obtenido bajo diferentes premisas, como diferentes distribuciones de probabilidad, ó transformaciones de las variables exógenas. Los parámetros de regresión del modelo que mayor valor de verosimilitud proporcionen, serán los que se empleen en la simulación del modelo para estimar los costes de reparto. Los supuestos que se van a analizar son los mismos que los modelos de reparto anteriores, y el estudio de máxima verosimilitud completo para cada uno de estos supuestos se pueden ver a continuación.

Tabla 175: VMV modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
40,276					-174,2		0%	12,82
1,449	2,966				-136,9	37,256	81%	5,56
-1,812	2,386	0,027			-75,1	61,803	99%	1,38
-1,314	2,305	0,027	0,003		-75,1	0,026	99%	1,40
-2,373	2,373	0,031		-0,000005	-74,5	0,542	99%	1,38

Tabla 176: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
3,696					-174,2		0%	12,82
2,785	0,067				-140,7	33,477	78%	6,06
2,668	0,056	0,001			-100,4	40,249	96%	2,46
2,197	0,126	0,001	-0,002		-90,1	10,389	98%	1,97
2,197	0,126	0,001	-0,002	0	-90,1	0,000	98%	1,99

Tabla 177: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,025					-174,2		0%	12,82
0,044	-0,001				-146,0	28,209	72%	6,83
0,047	-0,001	-0,00001			-122,9	23,101	90%	4,09
0,072	-0,005	-0,00001	0,00011		-111,9	10,942	94%	3,23
0,072	-0,005	-0,00001	0,00011	0	-111,9	0,000	93%	3,27

Tabla 178: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-6,346					-174,2		0%	12,82
-3,332	-0,226				-138,5	35,645	80%	5,77
-2,987	-0,186	-0,002			-83,9	54,671	98%	1,69
-2,215	-0,305	-0,002	0,004		-77,8	6,065	99%	1,49
-2,002	-0,296	-0,004	0,004	0,000002	-75,4	2,348	99%	1,36

Tabla 179: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
38,907					-177,4		-1%	12,90
-1,273	3,157				-141,0	36,381	81%	5,62
-2,303	2,491	0,025			-77,3	63,672	99%	1,46
-2,303	2,491	0,025	-0,0000001		-77,3	0,000	99%	1,48
-2,303	2,491	0,025		0	-77,3	0,000	99%	1,48

Tabla 180: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
3,669					-177,4		-1%	12,87
2,691	0,073				-144,2	33,173	77%	6,15
2,698	0,054	0,001			-99,8	44,374	96%	2,48
2,698	0,054	0,001	0		-99,8	0,000	96%	2,51
2,698	0,054	0,001		0	-99,8	0,000	96%	2,51

Tabla 181: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-222,230					-237,1		-1010%	42,72
0,046	-0,002				-147,8	89,353	70%	7,02
0,047	-0,001	-0,00001			-120,6	27,200	89%	4,22
0,047	-0,001	-0,00001	0,0000002		-120,5	0,077	89%	4,26
0,047	-0,001	-0,00001		0	-120,6	-0,062	89%	4,28

Tabla 182: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-6,204					-177,4		-2%	12,95
-3,083	-0,241				-142,6	34,737	79%	5,84
-2,957	-0,188	-0,002			-85,3	57,368	98%	1,72
-2,161	-0,309	-0,002	0,004		-78,7	6,554	99%	1,50
-2,161	-0,309	-0,002	0,004	0	-78,7	0,000	99%	1,52

Tabla 183: VMV modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
10,262					-223,6		-561%	32,96
0,746	3,010				-138,3	85,242	81%	5,57
-1,846	2,389	0,027			-75,6	62,708	99%	1,38
-1,029	2,253	0,027	0,005		-75,6	0,081	99%	1,40
-2,233	2,374	0,030		-0,000004	-75,3	0,295	99%	1,38

El supuesto que mayor valor de máxima verosimilitud ha proporcionado ha sido el que considera que los errores se distribuyen de acuerdo a una distribución normal. Por lo tanto, se empleará como modelo de costes de reparto el análisis de regresión bajo el supuesto de que los errores se distribuyen bajo una distribución normal.

2.4.4.1. Selección de variables del modelo de costes de reparto

En la tabla referente al estudio de máxima verosimilitud con los errores bajo la suposición de que se distribuyen de acuerdo a función normal, se puede observar como al introducir las variables exógenas número de repartos y peso total repartido por día, el cociente de verosimilitud es mayor que dos unidades cuando se introducen cada una de estas variables de una en una en el modelo. Sin embargo, al introducir la variable exógena número de repartos elevada a la segunda potencia, el cociente de verosimilitud es de 0,026 unidades, por lo que la adición de esta variable no aporta un mejor ajuste del modelo. Por último, al calcular el valor de máxima verosimilitud con las dos primeras variables exógenas más la variable peso total repartido elevada al cuadrado, el cociente de verosimilitud entre el modelo con dos variables y el modelo con estas tres variables es de 0,6 unidades, por lo que la incorporación de esta variable tampoco aporta un mejor ajuste del modelo, y no se empleará en el modelo de costes de reparto que se utilizará posteriormente en la simulación.

En definitiva, el modelo de costes de reparto para el código postal 08195 constará de las variables número de repartos por día y peso total repartido por día, con los siguientes parámetros de regresión respectivamente, 2,386 y -0,027. El valor umbral del modelo de regresión es -1,812.

2.4.4.2. Adecuación del modelo de costes de reparto

Como el mayor valor máximo de verosimilitud es el que aporta una distribución de probabilidad normal, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. Este análisis gráfico se puede ver en la figura 83.

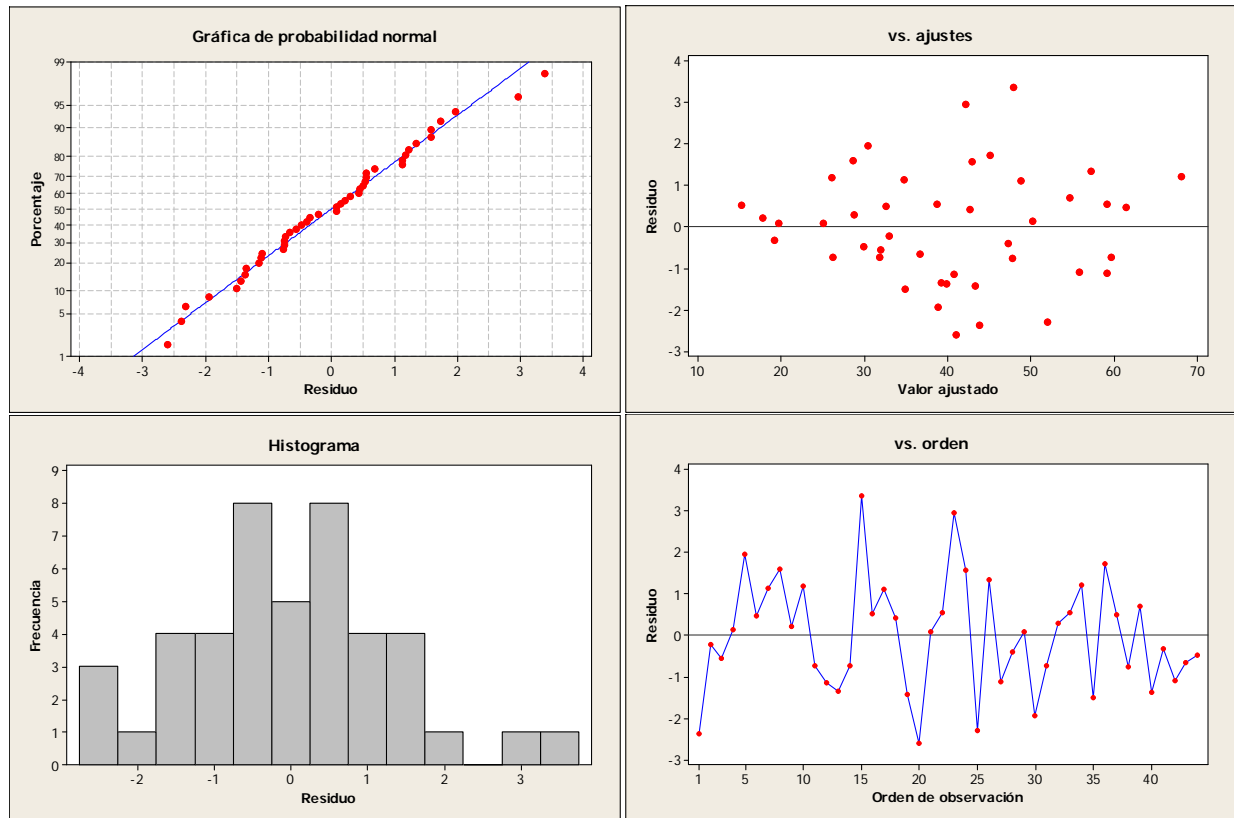


Figura 83: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08195

En la primera gráfica se puede observar como los errores se distribuyen de manera normal ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.4.4.3. Validación del modelo de costes de reparto

Para comprobar que el modelo predice de manera adecuada los valores de los costes de reparto ante nuevos valores de las variables exógenas, se realiza la prueba *t* pareada o prueba de Welch. Los nuevos valores de las variables exógenas son los valores de la muestra de predicción, y la validación del modelo se lleva a cabo comparando uno a uno los valores de la variable endógena estimados por el modelo tras introducir los nuevos valores de las variables exógenas, con los valores correspondientes de los costes de reparto de la muestra de predicción.

La tabla que resume los resultados de esta prueba de Welch es la tabla 184.

Tabla 184: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08195

	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	38,29	38,37	-0,080	Límite superior I.C	$t_{0,025;37}$
Varianza	333,13	310,76	2,11	0,370	1,96
Varianza/n			0,053		
Máximo	86,80	85,68		Límite inferior I.C	
Mínimo	0,76	2,31		-0,530	

El intervalo de confianza obtenido para esta prueba contiene el valor cero, por lo que se concluye que el modelo estima nuevos costes de reparto de manera adecuada. Para elaborarlo se ha empleado la distribución t de Student, con un nivel de significancia del 95% y 37 grados de libertad, cuyo valor es 1.96. Los valores Z que se emplean para realizar en esta prueba de Welch se obtienen de la tabla 185.

Tabla 185: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
-1,812	2,386	0,027			
	10	344	31,32	32,69	-1,37
	22	653	68,28	67,67	0,60
	8	263	24,36	24,85	-0,48
	17	655	56,40	55,99	0,41
	11	90	26,86	25,37	1,49
	11	506	38,07	37,07	1,00
	12	182	31,72	31,84	-0,12
	18	641	58,41	58,03	0,38
	13	514	43,06	42,94	0,11
	1	7	0,76	2,31	-1,55
	26	986	86,80	85,68	1,12
	17	455	51,01	50,01	1,00
	9	480	32,60	35,63	-3,03
	5	66	11,90	12,71	-0,82
	21	599	64,43	64,68	-0,25
	16	268	43,58	43,63	-0,05
	9	123	22,98	23,32	-0,34
	9	128	23,11	23,68	-0,57
	16	544	51,02	49,80	1,22
	22	923	75,55	71,84	3,72
	13	461	41,63	43,62	-1,99
	7	75	16,91	17,01	-0,10
	13	291	37,05	36,34	0,71
	18	482	54,12	53,13	1,00
	14	179	36,41	36,19	0,22
	13	459	41,58	43,70	-2,12
	11	82	26,64	25,03	1,61
	19	583	59,23	56,41	2,82
	9	175	24,38	24,52	-0,14
	13	344	38,48	40,43	-1,95
	11	328	33,27	34,15	-0,87
	8	197	22,58	23,38	-0,80
	14	499	45,04	47,12	-2,08
	12	326	35,60	34,65	0,96
	15	854	56,99	55,15	1,84
	12	370	36,79	38,43	-1,64
	7	78	16,99	17,69	-0,70
	6	165	16,95	15,72	1,23
	10	293	29,94	32,09	-2,15
	6	230	18,70	20,25	-1,55

2.4.5. Modelo de costes de recogida

Una vez finalizado el modelo de los costes de reparto, hay que realizar el modelo de costes de recogida. La creación este modelo es análogo al de los costes de reparto, por lo que los pasos a seguir son idénticos.

Tabla 186: VMV modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
19,731					-138,6		0%	7,20
3,356	4,331				-117,4	21,139	63%	4,35
0,943	2,984	0,016			-54,5	62,988	98%	0,95
0,088	3,487	0,016	-0,065		-53,9	0,511	98%	0,95
0,939	2,983	0,016		0	-54,5	-0,511	98%	0,96

Tabla 187: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,982					-138,6		0%	7,20
2,197	0,199				-119,2	19,357	60%	4,54
2,053	0,147	0,001			-76,9	42,282	95%	1,64
1,705	0,323	0,001	-0,020		-69,8	7,147	96%	1,40
1,705	0,323	0,001	-0,020	0	-69,8	0,000	96%	1,42

Tabla 188: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-918,876					-182,9		-771%	21,23
0,084	-0,008				-121,8	61,144	55%	4,84
0,091	-0,006	0,000			-94,5	27,295	88%	2,52
0,125	-0,022	0,000	0,002		-85,6	8,874	92%	2,06
0,125	-0,022	0,000	0,002	0	-85,6	0,000	92%	2,08

Tabla 189: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máxima veros.	Coci. de veros.	R ² adj.	σ
-4,442					-138,6		0%	7,20
-2,610	-0,474				-118,1	20,437	62%	4,43
-2,302	-0,340	-0,002			-64,2	53,923	97%	1,20
-1,898	-0,559	-0,002	0,026		-59,8	4,417	98%	1,10
-1,809	-0,544	-0,002	0,026	0,0000004	-56,6	3,190	98%	1,02

Tabla 190: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
20,163					-139,9		0%	7,21
1,982	4,363				-114,3	25,603	60%	4,54
0,584	3,099	0,016			-58,3	55,990	98%	0,97
0,183	3,563	0,016	-0,078		-57,4	0,852	98%	0,96
0,584	3,099	0,016		0,0000004	-59,4	-1,976	98%	0,97

Tabla 191: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
3,004					-139,9		0%	7,21
2,102	0,206				-116,3	23,528	57%	4,74
2,152	0,133	0,001			-77,9	38,405	94%	1,76
2,146	0,135	0,001	-0,001		-77,2	0,718	94%	1,77
2,146	0,135	0,001		0	-78,0	-0,744	94%	1,77

Tabla 192: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-6365					-191,7		-771%	21,23
0,093	-0,010				-121,0	70,653	52%	5,01
0,083	-0,005	-0,00003			-95,3	25,710	86%	2,71
0,084	-0,005	-0,00003	0,000003		-95,1	0,205	86%	2,71
0,084	-0,005	-0,00003		0	-95,2	-0,101	86%	2,71

Tabla 193: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-4,490					-139,9		0%	7,21
-2,471	-0,471				-114,4	25,480	59%	4,63
-2,336	-0,339	-0,002			-67,2	47,197	97%	1,21
-1,902	-0,568	-0,002	0,027		-63,1	4,062	98%	1,12
-1,902	-0,568	-0,002	0,027	0,0000001	-63,0	0,158	97%	1,15

Tabla 194: VMV modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
6,534					-183,5		-345%	15,17
8,723	0,219				-164,3	19,189	-206%	12,60
0,956	2,993	0,016			-56,0	108,344	98%	0,95
0,088	3,549	0,016	-0,075		-55,4	0,553	98%	0,95
0,089	3,045	0,019		-0,000002	-57,4	-2,029	98%	1,00

En las tablas se puede observar como el valor de máxima verosimilitud es mayor para la distribución normal. Por tanto, para la simulación se empleará el modelo con la suposición de que los errores se distribuyen de manera normal, y sin transformaciones lineales de las variables exógenas.

2.4.5.1. Selección de variables del modelo costes de recogida

Si se observa el cuadro en el que los errores se distribuyen según una distribución de probabilidad normal, se puede observar como al introducir las variables número de recogidas y peso total recogido, el valor de máxima verosimilitud al introducir cada una de estas variables una a una, difiere en más de dos unidades, por lo que el modelo que incluye estas variables ajusta mejor los datos que aquel que no las tiene. Sin embargo, al introducir la variable número de recogidas por día elevada al cuadrado el valor de máxima verosimilitud solo se incrementa en 0,511 unidades, por lo que la adición de esta variable no aporta un mejor ajuste. Lo mismo ocurre al introducir la variable peso total recogido por día elevada a la segunda potencia, junto con las dos primeras variables, ya que el cociente de verosimilitud es de cero

unidades al comparar el valor de máxima verosimilitud del modelo con las dos primeras variables, y el modelo de estas dos junto con el peso total recogido por día elevado al cuadrado.

Resumiendo, en este modelo de costes de recogida se incluirán las variables número de recogidas por día y peso total recogido por día con los parámetros de regresión 2,984 y 0,016 respectivamente. El valor umbral para el modelo es 0,943.

2.4.5.2. Adecuación del modelo de costes de recogida

Como el mayor valor máximo de verosimilitud es el que aporta una distribución de probabilidad normal, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. Este análisis gráfico se puede ver en la figura 84.

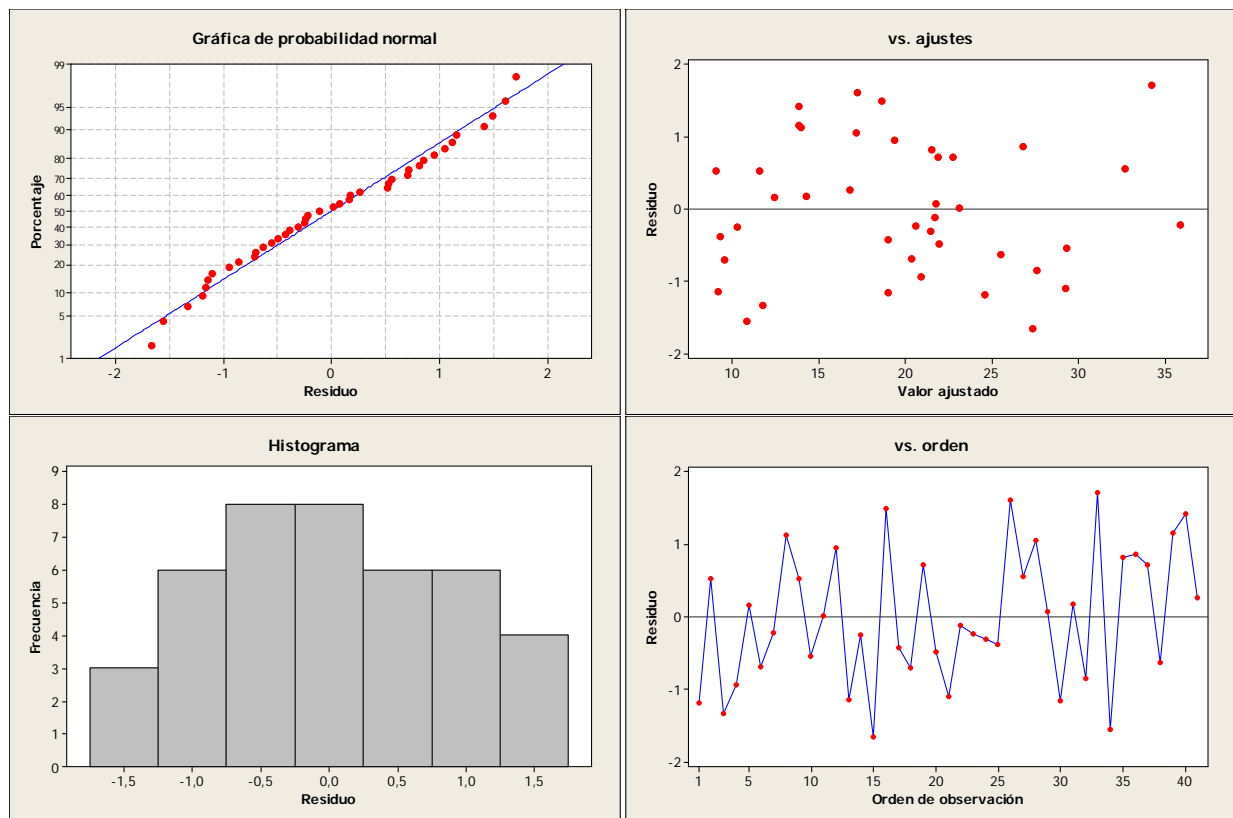


Figura 84: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08195

En la primera gráfica se puede observar como los errores se distribuyen de manera normal ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y

por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.4.5.3. Validación del modelo de costes de recogida

Para comprobar que el modelo predice de manera adecuada los valores de los costes de reparto ante nuevos valores de las variables exógenas, se realiza la prueba *t* pareada o prueba de Welch, con los datos de la muestra de predicción previamente obtenida. Y el cuadro que resume esta prueba de Welch es la tabla 195.

Tabla 195: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08195

	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch		
Media	18,91	17,77	0,184	Límite superior I.C	$t_{0,025;41}$ 2,04	
Varianza	77,74	74,53	1,14			0,563
Varianza/n			0,035	Límite inferior I.C		
Máximo	38,90	40,89				-0,195
Mínimo	0,94	0,00				

En él se puede observar como el modelo estima los costes de recogida de manera adecuada, ya que el intervalo de confianza creado incluye el cero. Para elaborarlo se ha empleado la distribución *t* de Student, con un nivel de significancia del 95% y 41 grados de libertad, cuyo valor es 1.96. Los valores Z con los que se elabora esta prueba provienen de la tabla 196.

Tabla 196: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08195

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
0,943	2,984	0,016			
	4	732	24,88	26,03	-1,15
	5	239	19,78	18,94	0,84
	4	338	18,42	17,56	0,86
	1	13	4,14	3,14	1,00
	4	381	19,13	17,00	2,12
	3	154	12,42	12,21	0,21
	5	354	21,67	20,91	0,76
	4	1290	34,03	34,74	-0,71
	3	306	14,91	16,13	-1,22
	6	860	32,95	31,24	1,71
	3	309	14,96	14,15	0,82
	3	212	13,37	12,66	0,72
	4	193	16,04	16,73	-0,68
	3	230	13,67	14,81	-1,15
	2	124	8,94	9,21	-0,26
	6	785	31,72	31,54	0,18
	4	279	17,45	16,89	0,57
	3	265	14,24	14,04	0,20
	3	115	11,78	12,00	-0,22
	6	792	31,83	30,96	0,87
	0	0	0,94	0,00	0,94
	3	126	11,96	11,64	0,32
	2	341	12,50	11,06	1,44
	6	713	30,54	32,27	-1,73
	6	257	23,06	21,53	1,53
	4	275	17,39	16,67	0,72
	5	154	18,39	17,08	1,31
	7	1041	38,90	40,89	-1,99
	3	563	19,13	19,93	-0,80
	4	224	16,55	16,65	-0,10
	7	526	30,46	30,92	-0,47
	3	258	14,13	15,19	-1,07
	4	61	13,88	11,23	2,65
	3	195	13,09	13,98	-0,89
	3	391	16,31	17,68	-1,38
	3	336	15,40	16,16	-0,76
	3	747	22,14	21,39	0,75
	5	522	24,42	23,25	1,17
	3	598	19,70	20,03	-0,33
	2	181	9,88	10,24	-0,36
	1	39	4,57	4,11	0,45
	2	323	12,21	10,81	1,40
	2	191	10,04	10,39	-0,35

2.4.6. *Modelo de carga de trabajo*

En este modelo se va a identificar que distribución de probabilidad siguen las variables aleatorias con las que se puede modelar la carga la de trabajo, así como los parámetros que las caracterizan. Estas variables son el número de repartos por día y el peso de cada uno de los repartos, en el caso de los repartos, y el número de recogidas por día y el peso de cada recogida, en el de las recogidas.

Las muestras que se emplean para las variables número de repartos, y número de recogidas son las mismas muestras que las muestras de calibración que se emplearon para ajustar los modelos de los costes de reparto y recogida. Las muestras para las variables peso de cada reparto y peso de cada

recogida, es una muestra con los pesos unitarios de cada uno de los repartos y de cada una de las recogidas.

2.4.6.1. Número de repartos por día

Minitab busca los parámetros para cada una de los distintos tipos de distribuciones de probabilidad, siendo estos parámetros tales que la distribución de probabilidad que caracterizan, es la que más se aproxima a los datos de la muestra. Una vez que se tienen estos datos hay que comprobar que la distribución de probabilidad que conforma se ajuste suficientemente bien a los datos. Para realizar esta comprobación se realizan los denominados test no paramétricos de bondad de ajuste. Debido a la naturaleza de esta variable aleatoria, la cual es un conteo de algún elemento, la primera función de distribución de probabilidad que se va a tratar de comprobar el ajuste es la distribución de Poisson. Como esta distribución es discreta se emplea el test no paramétrico de bondad de ajuste de Chi cuadrado. El resultado de este test se puede ver en la tabla 197.

Tabla 197: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08195

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
6-8	5	0,114	0,0957	4	0,1484	alfa=0,05 k=6 Par. Est.=1 n=6-1-1=4
9-10	10	0,227	0,1483	7	1,8513	
11-13	8	0,182	0,3191	14	2,5985	
14-15	9	0,205	0,1925	8	0,0334	
16-17	6	0,136	0,1299	6	0,0142	
18-22	6	0,136	0,1064	5	0,3720	
	44				5,02	9,488

El valor lambda dado por Minitab para una distribución de Poisson con los datos de la muestra de esta variable aleatoria es 13,09. En la tabla se puede observar como el valor límite para aceptar la hipótesis de que la una distribución de Poisson con lambda 13,09 ajusta adecuadamente la variable número de repartos por día es 9,488. Este valor corresponde a una distribución Chi cuadrado con 4 grados de libertad y nivel de significancia del 5%. Como en este caso el valor del test es 5,02, inferior al valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.4.6.2. Peso de cada reparto

La segunda variable a analizar en el modelo de carga de trabajo es el peso de cada reparto. Para esta variable continua se presentan los datos aportados por Minitab para cada uno de los parámetros de cada distribución, así como los resultados de los test de Kolmogorov-Smirnov en la tabla 198.

Tabla 198: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08195

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,258	alfa	0,099	alfa	0,084	Media	0,051	Lambda	0,259
30,77	6,218	0,6696	2,389	0,7473	2,023	2,518	1,224	21,45	6,227
Desv. Est.	$c_\alpha = 1,358$	beta	$c_\alpha = 1,358$	beta	$c_\alpha = 1,358$	Desv. Est.	$c_\alpha = 1,358$	Theta	$c_\alpha = 1,358$
46,17	$\alpha = 0,05$	45,95	$\alpha = 0,05$	25,14	$\alpha = 0,05$	1,411	$\alpha = 0,05$	19,64	$\alpha = 0,05$

Como se puede ver la única distribución que supera satisfactoriamente el test de Kolmogorov-Smirnov es la distribución lognormal. Por lo tanto, para la simulación se empleará esta distribución para estimar los pesos de cada uno de los repartos, y no será necesario emplear la distribución de probabilidad calculada anteriormente para la variable peso total repartido por día.

2.4.6.3. Número de recogidas por día

Esta variable al igual que el número de repartos por día, es un conteo de algún elemento, y por tanto, es una distribución discreta. Por ser una distribución discreta, la primera distribución que se va a comprobar si ajusta adecuadamente es la distribución de Poisson. El resultado de la prueba de bondad de ajuste Chi cuadrado es el que se puede ver en la tabla 199.

Tabla 199: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08195

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
1-2	8	0,195	0,2721	11	0,8921	alfa=0,05 k=4 Est. par. =1 n=4-1-1=2
3	8	0,195	0,2054	8	0,0212	
4	14	0,341	0,1941	8	4,5829	
5-7	11	0,268	0,2892	12	0,0622	
	41				5,56	

El valor lambda dado por Minitab para una distribución de Poisson con los datos de la muestra de esta variable aleatoria es 3,78. En la tabla se puede observar como el valor límite para aceptar la hipótesis de que la una distribución de Poisson con lambda 3,78 ajusta adecuadamente la variable número de repartos por día es 5,991. Este valor corresponde a una distribución Chi cuadrado con 2 grados de libertad y nivel de significancia del 5%. Como en este caso el valor del test es 5,56, menor al valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.4.6.4. Peso de cada recogida

Por último, para finalizar con la elaboración del modelo de carga de trabajo para este código postal se va a analizar la variable peso de cada recogida. Esta variable es una variable continua, y las distribuciones a analizar son también continuas. Los parámetros de las distintas distribuciones analizadas, así como los resultados de los test de Kolmogorov-Smirnov de cada una se pueden ver en la tabla 200.

Tabla 200: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08195

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,207	alfa	0,068	alfa	0,064	Media	0,076	Lambda	0,223
121,1	2,603	0,6383	0,854	0,7391	0,796	3,838	0,952	92,5	2,788
Desv. Est.	$c_\alpha = 1,358$	beta	$c_\alpha = 1,358$	beta	$c_\alpha = 1,358$	Desv. Est.	$c_\alpha = 1,358$	Theta	$c_\alpha = 1,358$
151,2	$\alpha = 0,05$	189,6	$\alpha = 0,05$	100,6	$\alpha = 0,05$	1,629	$\alpha = 0,05$	75,5	$\alpha = 0,05$

Como se puede observar las únicas distribuciones que han superado el test de Kolmogorov-Smirnov han sido las distribuciones gamma, Weibull y lognormal, siendo la distribución Weibull la que menos discrepancia presenta. Por lo tanto, para la simulación se empleará esta distribución para estimar los pesos de cada uno de los repartos, y no será necesario emplear la distribución de probabilidad calculada anteriormente para la variable peso total repartido por día.

2.4.7. *Simulación de los costes de reparto y recogida, y de la carga de trabajo*

En este apartado se llevará a cabo la simulación de los modelos creados para el estudio del vehículo de 1500 kg. de peso máximo autorizado en el código postal 08198. En primer lugar se simulará la carga de trabajo. Para ello hay que determinar el número de repartos y de recogidas para cada uno de los días simulados, y el peso total repartido y el peso total recogido por día. Empleando los ajustes de las distribuciones de probabilidad de Poisson obtenidos en los apartados 2.4.6.1 y 2.4.6.3. se simularán el número de repartos y el número de recogidas para cada uno días que se desean simular.

Posteriormente, se simulará el peso correspondiente de cada uno de los repartos y las recogidas mediante las distribuciones de probabilidad calculadas en los apartados 2.4.6.2. y 2.4.6.4. Sumando el peso unitario de todos los repartos simulados en cada uno de los días, se calculará el peso total repartido

por día. Con el peso unitario de las recogidas simulado para cada uno de los días, se calculará el peso total recogido por día.

Finalmente, se simularán los modelos de costes de reparto y costes de recogida. Como valores de entrada a estos modelos se emplearán los datos simulados para el modelo de carga de trabajo. El número de repartos y el peso total repartido por día serán los datos de entrada para el modelo de costes de reparto. Análogamente, el número de recogidas y el peso total recogido por día serán los valores de entrada para el modelo de costes de recogida.

2.4.7.1. Simulación de los costes de reparto

La carga de trabajo relativa a los repartos para este código postal es la siguiente: el número de repartos puede ser descrito mediante una distribución de probabilidad Poisson con parámetro lambda igual a 13,09. Mientras que el peso de cada reparto sigue una distribución de probabilidad lognormal con media 2,518 y desviación estándar 1,411.

Un resumen de la carga de trabajo simulada es el de la tabla 201, en el cual se puede observar el promedio de repartos por día simulados, su desviación estándar, el número máximo de repartos a realizar en un día, así como el número mínimo. Y los mismos datos para la variable peso total repartido, siendo este peso total repartido la suma de los pesos simulados para cada una de los repartos.

Tabla 201: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08195

	Repartos por día	Peso total repartido por día
Promedio	13	376
Desv. Estan.	4	185
Máximo	23	900
Mínimo	4	42

Esta simulación de la carga de trabajo se emplea como entrada para el modelo de costes de reparto. Como se han simulado 250 días de carga de trabajo, se obtendrán los estimados puntuales de costes de reparto para 250 días. Los resultados que aquí se exponen son el valor medio de límite inferior de las 250 simulaciones, la varianza y desviación estándar de este límite inferior, así como un intervalo de confianza de este límite inferior. Para el límite superior del intervalo de predicción los resultados que se presentan son análogos. Estos resultados son los de la tabla 202.

Tabla 202: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08195

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	37,29	42,85
Varianza	146	146
Desv. Estan.	12,07	12,09
I.C	36,03	41,59
	38,55	44,11

En la tabla se puede observar como el límite inferior de los estimados puntuales de los costes de repartos se distribuyen con un valor medio de 37,29 euros por día y una desviación estándar de 12,07. Mientras que el límite superior de los estimados puntuales de los costes de reparto tiene un valor medio de 42,85y una desviación estándar de 12,09 euros por día.

2.4.7.2. Simulación de los costes de recogida

La carga de trabajo relativa a las recogidas para este código postal queda como sigue: el número de recogidas puede ser expresado mediante una distribución de probabilidad Poisson con un parámetro lambda igual a 3,78. Por otro lado, el peso de cada recogida puede ser descrito mediante una distribución Weibull con un valor alfa igual a 0,7391 y valor beta igual a 100,6. Un cuadro resumen de la carga de trabajo simulada es el de la tabla 203, en el cual se puede observar el promedio de recogidas por día simulados, su desviación estándar, el número máximo de recogidas a realizar en un día, así como el número mínimo. Y los datos análogos para la variable peso total recogido, siendo este peso total recogido la suma de los pesos simulados para cada una de las recogidas.

Tabla 203: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08195

	Recogidas por día	Peso total recogido por día
Promedio	4	423
Desv. Estan.	2	330
Máximo	11	1785
Mínimo	0	0

Y la simulación de los costes de recogida correspondiente a esta carga de trabajo para 250 días se resume en la tabla 204.

Tabla 204: Resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08195

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	17,16	21,07
Varianza	103	104
Desv. Estan.	10,12	10,18
I.C	16,10	20,01
	18,22	22,13

Donde se puede comprobar como el límite inferior de los estimados puntuales de los costes de recogida se distribuyen con un valor medio de 17,16 euros por día y una desviación estándar de 10,12. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 21,07 y una desviación estándar de 10,18 euros por día.

2.5. Código postal 08174

En este código postal no se dan repartos superiores a 300 kg. de peso, ni recogidas superiores a 700 kg., por lo que en él solo existe volumen de trabajo para vehículos de 1500 kg. de peso máximo autorizado.

Los pasos a seguir en la elaboración del modelo de costes de la red de distribución capilar de este código postal son análogos a los empleados en los códigos postales anteriores, siendo el primer paso en la elaboración del modelo la toma de muestras.

2.5.1. Toma de muestras

Se toman muestras de aquellas variables que se identificaron como las determinantes para el buen funcionamiento de la red de distribución capilar. Estas son el coste de reparto por día y el coste de recogida por día. Éstas a su vez dependen del número de repartos por día y del peso repartido por día, y del número de recogidas por día y del peso total repartido por día respectivamente. Esta toma de muestras se ha llevado a cabo mediante un estudio observacional, y se han obtenido 84 observaciones para cada una de ellas. Las muestras obtenidas para las variables relacionadas con los repartos se pueden observar en la tabla 205, mientras que en la tabla 206 se pueden ver las relacionadas con las recogidas.

Tabla 205: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08174

Día	REPARTOS			Día	REPARTOS		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	4	778	25	43	2	42	6
2	4	108	14	44	7	155	19
3	2	47	6	45	3	46	9
4	2	49	6	46	4	100	11
5	2	103	8	47	4	120	14
6	5	317	19	48	6	155	18
7	4	67	11	49	3	8	6
8	5	100	14	50	3	228	14
9	2	24	5	51	9	245	28
10	2	36	6	52	1	16	3
11	2	24	5	53	3	29	7
12	1	8	2	54	4	37	10
13	2	29	5	55	6	61	14
14	3	90	11	56	3	17	7
15	6	137	18	57	1	4	2
16	8	315	26	58	7	149	20
17	3	94	10	59	6	106	16
18	2	11	4	60	2	54	7
19	2	43	6	61	1	3	2
20	6	69	14	62	1	8	2
21	3	77	9	63	1	8	2
22	2	26	5	64	3	43	9
23	3	236	14	65	1	4	2
24	6	351	24	66	4	334	17
25	5	68	14	67	6	65	15
26	4	47	10	68	4	223	14
27	1	1	2	69	3	69	9
28	5	121	14	70	4	187	13
29	3	17	7	71	6	137	18
30	0	0	0	72	5	171	14
31	8	103	20	73	3	26	7
32	4	96	11	74	3	97	10
33	1	9	2	75	5	197	17
34	4	81	11	76	2	50	6
35	4	13	8	77	5	58	13
36	1	4	2	78	4	404	19
37	3	37	8	79	1	14	3
38	1	12	3	80	4	141	14
39	2	73	7	81	5	58	13
40	4	50	11	82	3	893	22
41	3	27	7	83	3	14	6
42	7	71	17	84	3	15	6

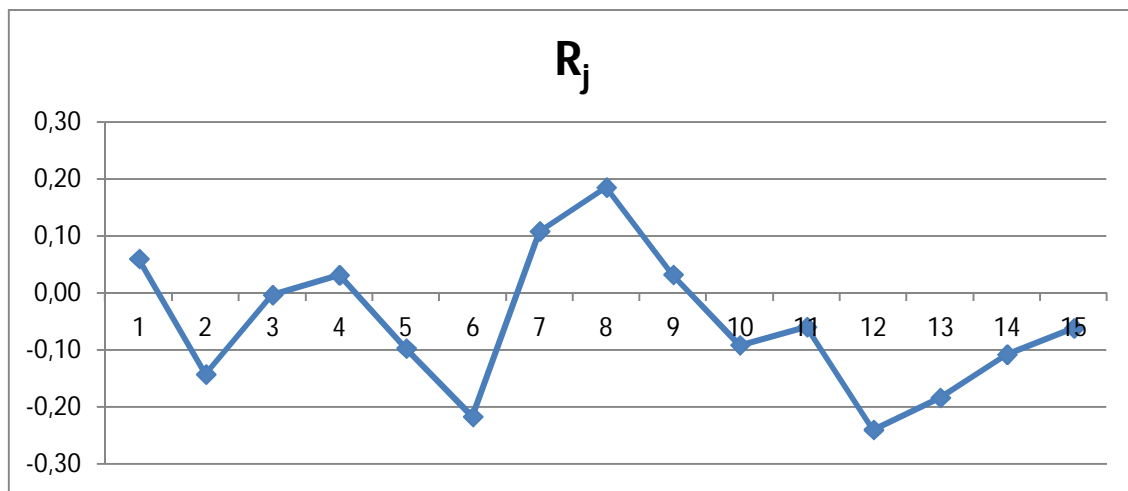
Tabla 206: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08174

Día	RECOGIDAS			Día	RECOGIDAS		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	1	1	2	43	1	28	4
2	0	0	0	44	1	11	3
3	2	106	8	45	2	290	10
4	1	32	4	46	1	5	2
5	0	0	0	47	0	0	0
6	4	69	11	48	1	33	4
7	3	20	8	49	0	0	0
8	2	39	6	50	0	0	0
9	2	34	6	51	0	0	0
10	2	104	8	52	3	91	11
11	3	57	9	53	0	0	0
12	3	26	8	54	0	0	0
13	4	173	15	55	2	30	6
14	2	69	8	56	1	49	4
15	3	175	11	57	1	4	2
16	0	0	0	58	0	0	0
17	2	314	13	59	2	50	8
18	2	42	7	60	3	45	9
19	2	13	5	61	2	16	5
20	1	11	3	62	2	3	4
21	2	46	7	63	2	38	6
22	1	2	2	64	1	284	8
23	2	96	7	65	0	0	0
24	4	289	16	66	1	11	3
25	1	1	2	67	5	99	16
26	2	2	4	68	3	79	11
27	1	1	2	69	0	0	0
28	2	15	5	70	4	82	12
29	2	4	4	71	2	35	6
30	2	67	8	72	1	26	4
31	1	1	2	73	0	0	0
32	1	34	4	74	2	6	4
33	0	0	0	75	2	165	9
34	5	85	15	76	1	8	3
35	1	28	4	77	1	70	5
36	0	0	0	78	0	0	0
37	2	26	6	79	1	12	3
38	1	247	8	80	1	88	5
39	2	78	7	81	2	92	9
40	0	0	0	82	2	6	4
41	4	146	15	83	0	0	0
42	0	0	0	84	2	44	6

2.5.2. Independencia de las muestras

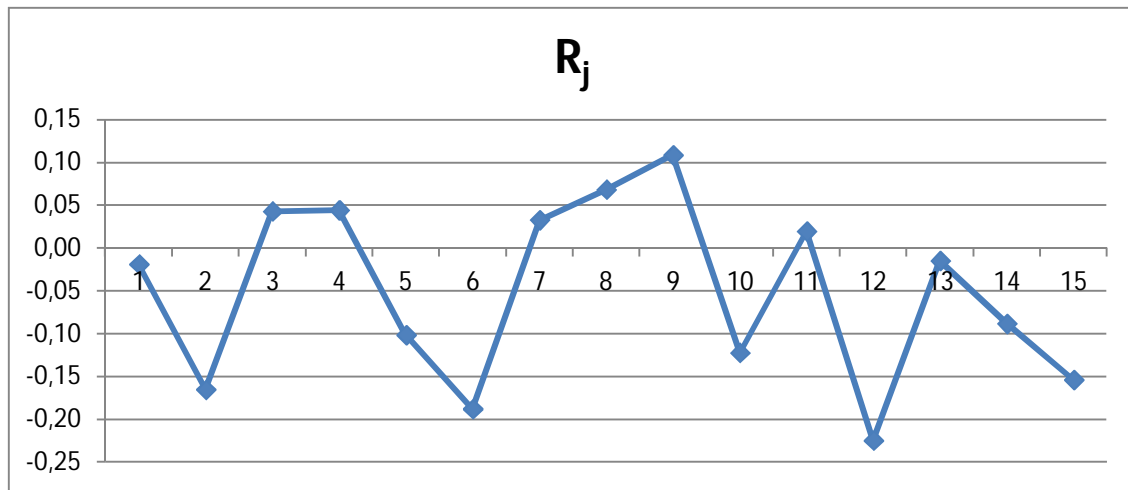
Tras la obtención de las muestras, hay que comprobar si las observaciones de las mismas son independientes. Que las observaciones sean independientes es fundamental, ya que para emplear el método de máxima verosimilitud en la elaboración de los modelos de costes de reparto y recogida, las muestras utilizadas han de ser independientes. Por otro lado, los test no paramétricos de bondad de ajuste de las distribuciones de probabilidad también presuponen que las muestras empleadas son independientes.

En primer lugar se analiza la independencia de la muestras de los repartos. Como esta muestra se compone de tres variables, coste de reparto por día, número de repartos por día y peso total repartido por día, es necesario comprobar la independencia de cada una de ellas. En primer lugar se muestran las gráficas de autocorrelación.



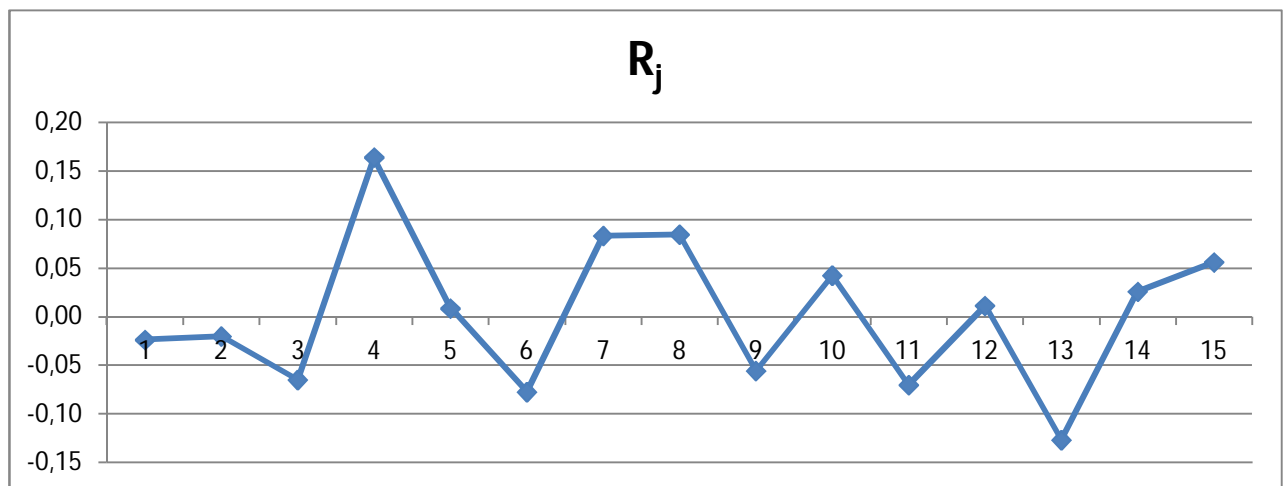
R_j	0,06	-0,14	0,00	0,03	-0,10	-0,22	0,11	0,19	0,03	-0,09	-0,06	-0,24	-0,18	-0,11	-0,06
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	2	-6	0	1	-4	-9	4	7	1	-4	-2	-10	-7	-4	-3

Figura 85: Gráfica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08174



R_j	-0,02	-0,17	0,04	0,04	-0,10	-0,19	0,03	0,07	0,11	-0,12	0,02	-0,22	-0,01	-0,09	-0,15
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-0,1	-0,6	0,2	0,2	-0,4	-0,7	0,1	0,2	0,4	-0,4	0,1	-0,8	-0,1	-0,3	-0,6

Figura 86: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08174



R_j	-0,02	-0,02	-0,06	0,16	0,01	-0,08	0,08	0,08	-0,06	0,04	-0,07	0,01	-0,13	0,03	0,06
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-500	-426	-1370	3463	182	-1637	1760	1787	-1176	895	-1480	240	-2679	547	1185

Figura 87: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08174

En las tres gráficas se puede observar como en ninguna de ellas los valores de R_j se encuentran en el rango de -1 y 1, y que en ninguna de ellas se puede apreciar un patrón determinado. Por tanto, esta prueba indica la independencia de las observaciones de las muestras.

Tras las gráficas de autocorrelación, se muestran a continuación las gráficas de dispersión para cada una de las variables, como segunda prueba para comprobar la independencia de las observaciones de cada muestra.

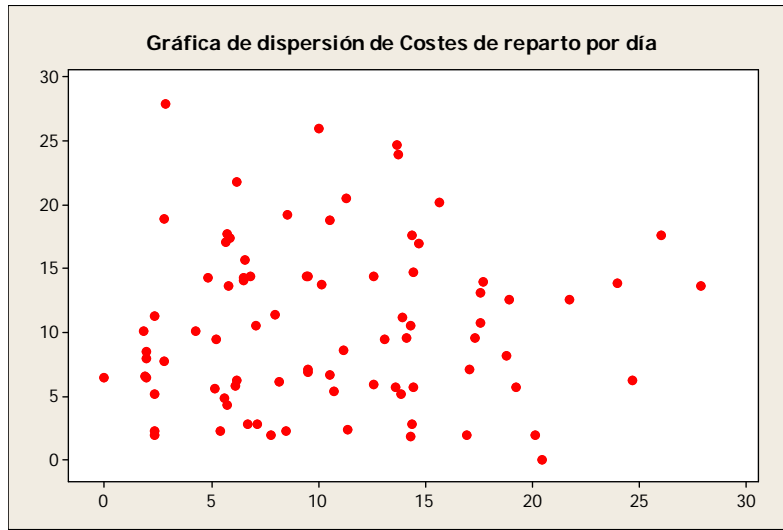


Figura 88: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08174

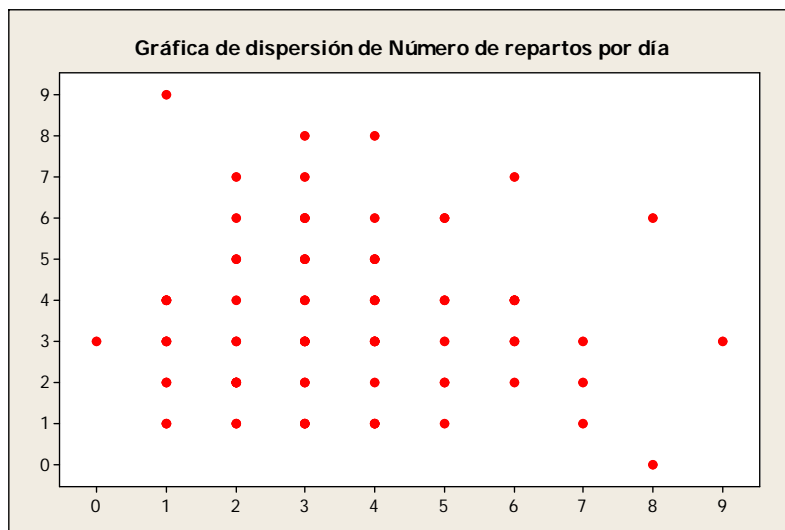


Figura 89: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08174

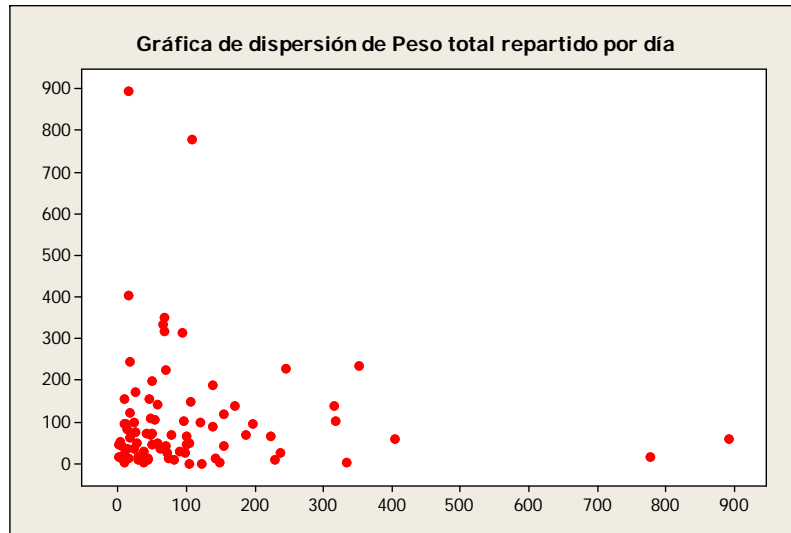
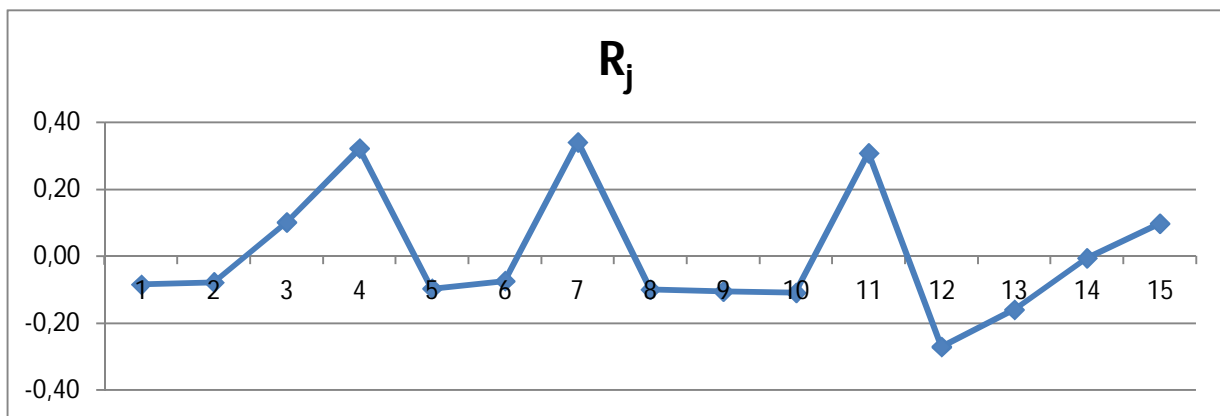


Figura 91: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08174

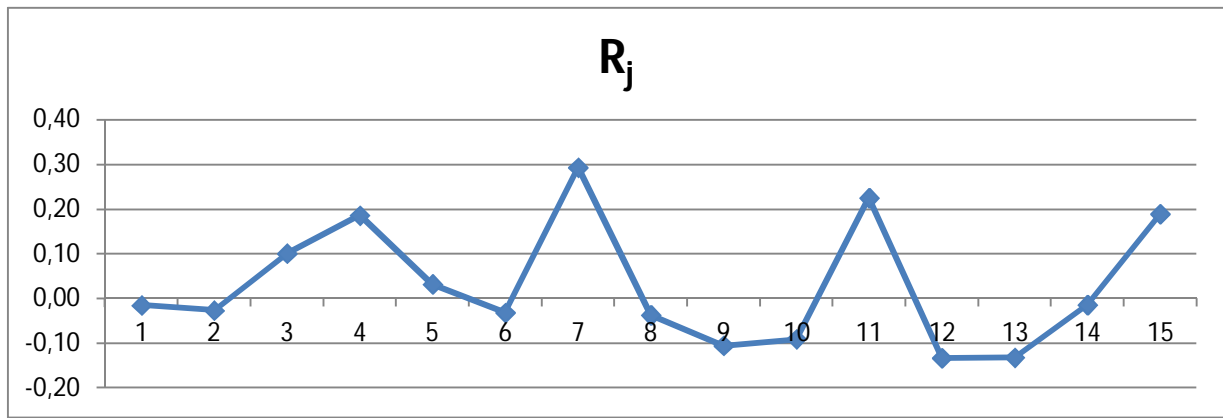
Como en las tres gráficas los puntos se distribuyen de manera aleatoria a lo largo de todo el plano, esta prueba también concluye que las observaciones son independientes.

Tras comprobar que las muestras relativas a los repartos son independientes, hay que analizar si las muestras correspondientes a las recogidas también son independientes. Al igual que para los repartos se comenzará mostrando las gráficas de autocorrelación.



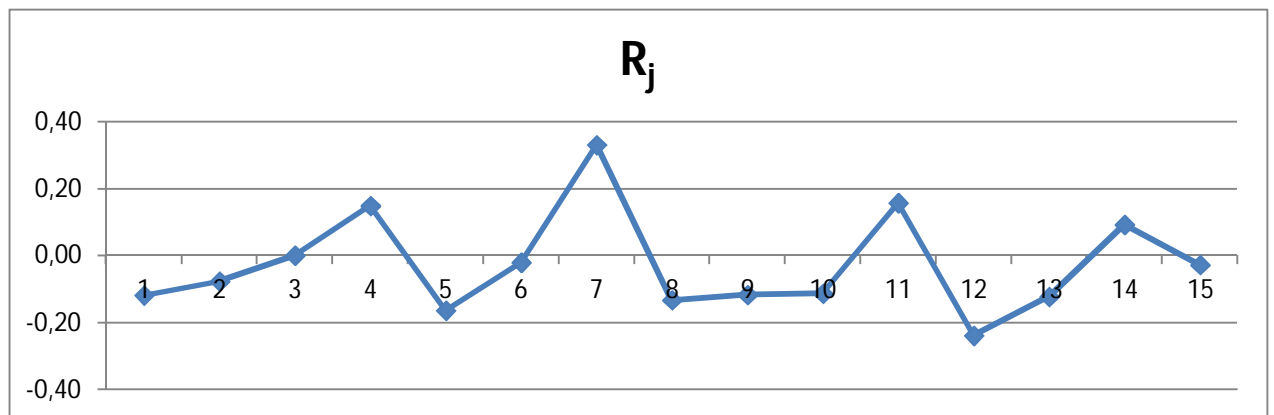
R_j	-0,08	-0,08	0,10	0,32	-0,10	-0,07	0,34	-0,10	-0,11	-0,11	0,31	-0,27	-0,16	-0,01	0,10
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-2	-1	2	6	-2	-1	6	-2	-2	-2	6	-5	-3	0	2

Figura 92: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08174



R_j	-0,02	-0,03	0,10	0,19	0,03	-0,03	0,29	-0,04	-0,11	-0,09	0,22	-0,13	-0,13	-0,01	0,19
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	0,0	0,0	0,2	0,3	0,0	0,0	0,4	-0,1	-0,2	-0,1	0,3	-0,2	-0,2	0,0	0,3

Figura 93: Grafica de autocorrelación número de recogidas por día vehículo 1500 kg. PMA código 08174



R_j	-0,12	-0,08	0,00	0,15	-0,16	-0,02	0,33	-0,13	-0,12	-0,11	0,16	-0,24	-0,12	0,09	-0,03
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-630	-408	5	795	-870	-108	1766	-707	-622	-601	841	-1273	-654	492	-144

Figura 94: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08174

Al igual que en el caso de los repartos, en las tres gráficas se puede ver como los valores de R_j no superan los valores 1 y -1, y que no se da un patrón determinado en ninguna de ellas por lo que se concluye que las observaciones de cada muestra son independientes. Tras estudiar las gráficas de autocorrelación, a continuación se verifica la independencia de las observaciones mediante las gráficas de dispersión.

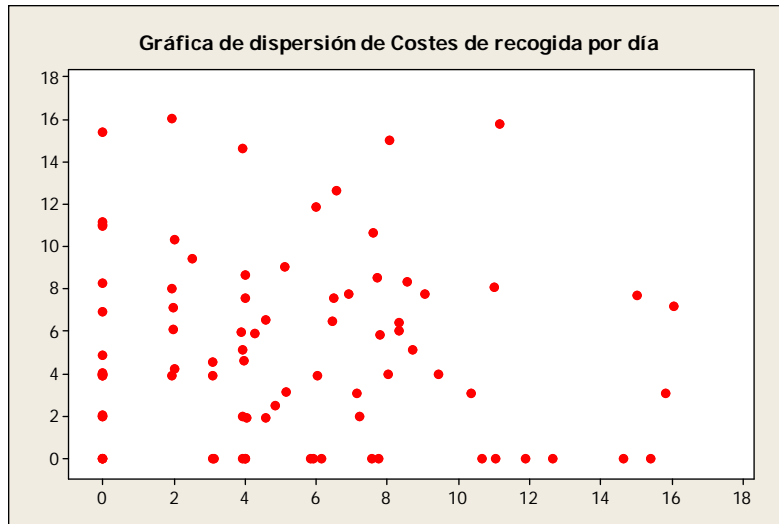


Figura 95: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08174

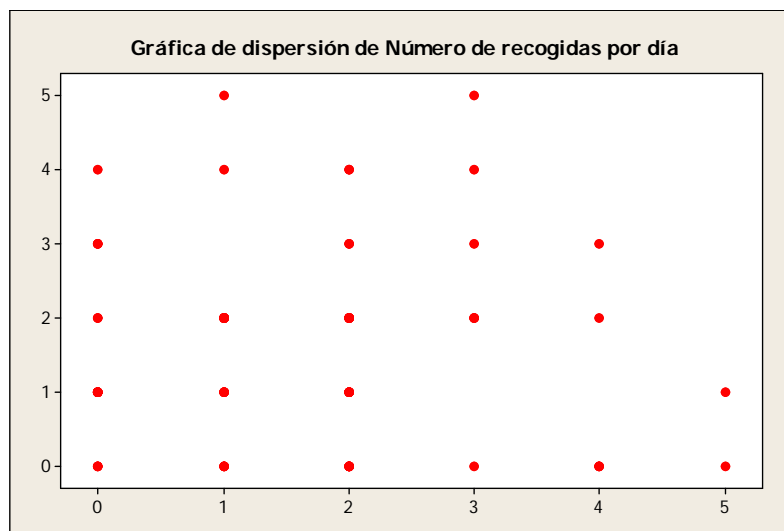


Figura 96: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08174

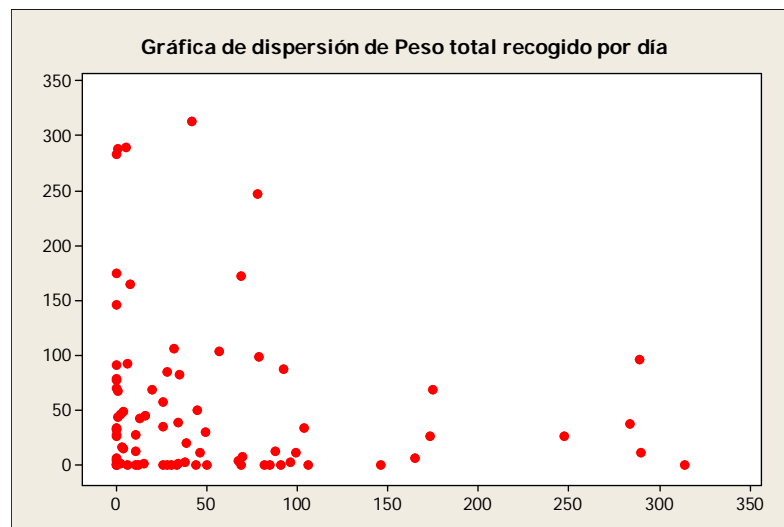


Figura 97: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08174

En las tres gráficas se puede observar como todos los puntos se distribuyen de manera aleatoria en los planos correspondientes, pudiéndose concluir que las observaciones de las muestras son independientes.

2.5.3. Separación de las muestras de calibración y predicción

Tras verificar que las 84 observaciones obtenidas del estudio observacional de cada una de las variables son independientes, hay que separar estas 84 observaciones en dos muestras diferentes, la de calibración y la de predicción. La de calibración se emplea para la creación y elaboración de los modelos necesarios para el estudio de los costes de la red de distribución, mientras que la de predicción se emplea para comprobar que los modelo elaborados mediante el análisis de regresión estiman los costes de reparto y recogida de manera adecuada. Esta separación se lleva a cabo mediante una asignación de un número aleatorio a cada una de las observaciones, y si éste es mayor que el valor 0,5, se asigna a la muestra de calibración. Si por el contrario, este número aleatorio es menor que 0,5, se asigna a la muestra de predicción. El listado de números aleatorios asignados a cada una de las observaciones relativas al reparto es el que se puede ver en la tabla 207.

Tabla 207: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08174

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,084	22	0,602	43	0,757	64	0,291
2	0,111	23	0,586	44	0,929	65	0,932
3	0,097	24	0,076	45	0,722	66	0,893
4	0,194	25	0,257	46	0,266	67	0,333
5	0,657	26	0,441	47	0,305	68	0,835
6	0,526	27	0,581	48	0,465	69	0,577
7	0,280	28	0,030	49	0,304	70	0,696
8	0,967	29	0,368	50	0,393	71	0,871
9	0,041	30	0,377	51	0,352	72	0,764
10	0,418	31	0,938	52	0,760	73	0,889
11	0,153	32	0,893	53	0,606	74	0,241
12	0,357	33	0,242	54	0,803	75	0,713
13	0,018	34	0,545	55	0,908	76	0,403
14	0,697	35	0,771	56	0,487	77	0,562
15	0,836	36	0,683	57	0,006	78	0,186
16	0,055	37	0,759	58	0,911	79	0,020
17	0,450	38	0,448	59	0,820	80	0,356
18	0,991	39	0,769	60	0,482	81	0,475
19	0,868	40	0,659	61	0,867	82	0,236
20	0,490	41	0,810	62	0,791	83	0,045
21	0,457	42	0,602	63	0,054	84	0,597

Quedando las muestras de calibración y predicción de acuerdo a esta asignación como se muestra en la tabla 208. En ella se puede observar cómo debido a la asignación aleatoria, se tienen 42 observaciones para la muestra de calibración y 42 observaciones para la muestra de predicción relativas a los repartos.

Tabla 208: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08174

Día	REPARTOS. MUESTRA DE CALIBRACIÓN			Día	REPARTOS. MUESTRA DE PREDICCIÓN		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	2	103	8	1	4	778	25
2	5	317	19	2	4	108	14
3	5	100	14	3	2	47	6
4	3	90	11	4	2	49	6
5	6	137	18	5	4	67	11
6	2	11	4	6	2	24	5
7	2	43	6	7	2	36	6
8	2	26	5	8	2	24	5
9	3	236	14	9	1	8	2
10	1	1	2	10	2	29	5
11	8	103	20	11	8	315	26
12	4	96	11	12	3	94	10
13	4	81	11	13	6	69	14
14	4	13	8	14	3	77	9
15	1	4	2	15	6	351	24
16	3	37	8	16	5	68	14
17	2	73	7	17	4	47	10
18	4	50	11	18	5	121	14
19	3	27	7	19	3	17	7
20	7	71	17	20	0	0	0
21	2	42	6	21	1	9	2
22	7	155	19	22	1	12	3
23	3	46	9	23	4	100	11
24	1	16	3	24	4	120	14
25	3	29	7	25	6	155	18
26	4	37	10	26	3	8	6
27	6	61	14	27	3	228	14
28	7	149	20	28	9	245	28
29	6	106	16	29	3	17	7
30	1	3	2	30	1	4	2
31	1	8	2	31	2	54	7
32	1	4	2	32	1	8	2
33	4	334	17	33	3	43	9
34	4	223	14	34	6	65	15
35	3	69	9	35	3	97	10
36	4	187	13	36	2	50	6
37	6	137	18	37	4	404	19
38	5	171	14	38	1	14	3
39	3	26	7	39	4	141	14
40	5	197	17	40	5	58	13
41	5	58	13	41	3	893	22
42	3	15	6	42	3	14	6

Una vez elaboradas estas muestras para los repartos, se realiza un proceso análogo para las recogidas. La asignación de números aleatorios a cada una de las 84 observaciones relativas a las recogidas es el de la tabla 209.

Tabla 209: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08174

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,178	22	0,494	43	0,652	64	0,247
2	0,579	23	0,676	44	0,889	65	0,497
3	0,576	24	0,591	45	0,559	66	0,344
4	0,546	25	0,083	46	0,412	67	0,952
5	0,554	26	0,297	47	0,366	68	0,996
6	0,084	27	0,912	48	0,596	69	0,429
7	0,204	28	0,242	49	0,160	70	0,660
8	0,269	29	0,088	50	0,570	71	0,571
9	0,369	30	0,378	51	0,563	72	0,437
10	0,470	31	0,955	52	0,938	73	0,930
11	0,801	32	0,131	53	0,434	74	0,803
12	0,675	33	0,722	54	0,196	75	0,743
13	0,485	34	0,676	55	0,725	76	0,873
14	0,074	35	0,749	56	0,417	77	0,148
15	0,335	36	0,675	57	0,329	78	0,213
16	0,695	37	0,596	58	0,979	79	0,221
17	0,837	38	0,284	59	0,642	80	0,801
18	0,574	39	0,154	60	0,091	81	0,642
19	0,326	40	0,781	61	0,250	82	0,826
20	0,066	41	0,204	62	0,096	83	0,989
21	0,205	42	0,950	63	0,081	84	0,781

A partir de este listado de números aleatorios se confeccionan las muestras de calibración y predicción relativas a las recogidas. Estas muestras se pueden observar en la tabla 210. En ellos se puede comprobar cómo debido a la asignación aleatoria, se tienen 43 observaciones para la muestra de calibración y 41 observaciones para la muestra de predicción relativas a las recogidas.

Tabla 210: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08174

Día	RECOGIDAS. MUESTRA DE CALIBRACIÓN			Día	RECOGIDAS. MUESTRA DE PREDICCIÓN		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	0	0	0	1	1	1	2
2	2	106	8	2	4	69	11
3	1	32	4	3	3	20	8
4	0	0	0	4	2	39	6
5	3	57	9	5	2	34	6
6	3	26	8	6	2	104	8
7	0	0	0	7	4	173	15
8	2	314	13	8	2	69	8
9	2	42	7	9	3	175	11
10	2	96	7	10	2	13	5
11	4	289	16	11	1	11	3
12	1	1	2	12	2	46	7
13	1	1	2	13	1	2	2
14	0	0	0	14	1	1	2
15	5	85	15	15	2	2	4
16	1	28	4	16	2	15	5
17	0	0	0	17	2	4	4
18	2	26	6	18	2	67	8
19	0	0	0	19	1	34	4
20	0	0	0	20	1	247	8
21	1	28	4	21	2	78	7
22	1	11	3	22	4	146	15
23	2	290	10	23	1	5	2
24	1	33	4	24	0	0	0
25	0	0	0	25	0	0	0
26	0	0	0	26	0	0	0
27	3	91	11	27	0	0	0
28	2	30	6	28	1	49	4
29	0	0	0	29	1	4	2
30	2	50	8	30	3	45	9
31	5	99	16	31	2	16	5
32	3	79	11	32	2	3	4
33	4	82	12	33	2	38	6
34	2	35	6	34	1	284	8
35	0	0	0	35	0	0	0
36	2	6	4	36	1	11	3
37	2	165	9	37	0	0	0
38	1	8	3	38	1	26	4
39	1	88	5	39	1	70	5
40	2	92	9	40	0	0	0
41	2	6	4	41	1	12	3
42	0	0	0				
43	2	44	6				

2.5.4. Modelo de costes de reparto

Como para el resto de modelos de costes de reparto elaborados de este trabajo de investigación, los parámetros de regresión se calculan empleando el método de máxima verosimilitud. Se estudiará el valor de máxima verosimilitud obtenido bajo diferentes premisas, como diferentes distribuciones de probabilidad, ó transformaciones de las variables exógenas. Los parámetros de regresión del modelo que mayor valor de verosimilitud proporcionen, serán los que se empleen en la simulación del modelo para estimar los costes de reparto. Los supuestos que se van a analizar son los mismos que los modelos de reparto anteriores, y el estudio de máxima verosimilitud completo para cada uno de estos supuestos se pueden ver a continuación.

Tabla 211: VMV modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
10,496					-131,7		0%	5,64
0,277	2,769				-89,8	41,935	86%	2,10
0,021	2,191	0,027			-41,0	48,847	99%	0,67
0,019	2,193	0,027	-0,0001		-41,0	0,000	99%	0,67
-0,112	2,032	0,044		-0,00005	-32,3	8,706	99%	0,55

Tabla 212: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,351					-131,7		0%	5,64
1,496	0,213				-100,5	31,279	77%	2,71
1,338	0,191	0,002			-74,4	26,047	93%	1,48
0,848	0,443	0,002	-0,025		-63,4	11,062	96%	1,15
0,859	0,390	0,005	-0,022	-0,00001	-58,2	5,116	97%	1,03

Tabla 213: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-1559					-163,6		-355%	12,03
0,155	-0,014				-108,8	54,747	66%	3,31
0,176	-0,014	0,000			-92,2	16,565	84%	2,26
0,254	-0,050	0,000	0,003		-80,1	12,134	91%	1,71
0,254	-0,050	0,000	0,003	0	-80,1	0,000	91%	1,74

Tabla 214: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-3,240					-131,7		0%	5,64
-1,709	-0,395				-95,1	36,627	82%	2,39
-1,552	-0,334	-0,004			-59,4	35,739	97%	1,03
-1,154	-0,560	-0,004	0,024		-52,4	6,972	98%	0,89
-1,155	-0,502	-0,007	0,021	0,00001	-43,9	8,471	98%	0,73

Tabla 215: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
9,927					-137,1		-1%	5,67
-0,641	2,967				-87,7	49,385	86%	2,15
-0,127	2,254	0,025			-43,3	44,395	98%	0,70
-0,127	2,254	0,025	0,0002		-43,3	0,002	98%	0,71
-0,113	2,213	0,027		-0,00001	-41,3	1,998	99%	0,69

Tabla 216: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,350					-137,1		0%	5,64
1,342	0,237				-102,0	35,096	76%	2,79
1,374	0,187	0,002			-77,3	24,711	93%	1,50
0,778	0,461	0,002	-0,027		-64,0	13,327	96%	1,17
0,781	0,462	0,002	-0,027	-0,0000004	-63,4	0,622	96%	1,16

Tabla 217: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-50796					-169,9		-355%	12,03
0,162	-0,016				-111,8	58,103	65%	3,35
0,164	-0,013	-0,0002			-95,5	16,287	83%	2,31
0,162	-0,013	-0,0002	0,00002		-95,3	0,186	83%	2,36
0,162	-0,013	-0,0002		0	-95,5	-0,225	82%	2,37

Tabla 218: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-3,231					-137,1		0%	5,64
-1,539	-0,421				-95,5	41,663	81%	2,43
-1,564	-0,326	-0,004			-62,2	33,254	97%	1,04
-1,072	-0,582	-0,004	0,026		-54,5	7,706	97%	0,90
-1,073	-0,583	-0,004	0,026	0,000001	-53,6	0,881	97%	0,90

Tabla 219: VMV modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
10,401					-133,5		0%	5,64
-0,005	2,779				-88,7	44,829	86%	2,12
-0,044	2,205	0,027			-41,5	47,174	99%	0,67
-0,064	2,219	0,027	-0,002		-41,5	0,002	99%	0,68
-0,170	2,010	0,048		-0,00006	-29,7	11,792	99%	0,56

El supuesto que mayor valor de máxima verosimilitud ha proporcionado ha sido el que considera que los errores se distribuyen de acuerdo a una distribución logística. Por lo tanto, se empleará como modelo de costes de reparto el análisis de regresión bajo el supuesto de que los errores se distribuyen bajo una distribución logística.

2.5.4.1. Selección de variables del modelo de costes de reparto

En la tabla referente al estudio de máxima verosimilitud con los errores bajo la suposición de que se distribuyen de acuerdo a función logística, se puede observar como al introducir las variables exógenas número de repartos y peso total repartido por día, el cociente de verosimilitud es mayor que dos unidades cuando se introducen cada una de estas variables de una en una en el modelo. Sin embargo, al introducir la variable exógena número de repartos elevada a la segunda potencia, el cociente de verosimilitud es de 0,002 unidades, por lo que la adición de esta variable no aporta un mejor ajuste del modelo. Por último, al calcular el valor de máxima verosimilitud con las dos primeras variables exógenas más la variable peso total repartido elevada al cuadrado, el cociente de verosimilitud entre el modelo con dos variables y el modelo con estas tres variables es de más de 11 unidades, por lo que esta variable sí que aporta un mejor ajuste del modelo, y se empleará en el modelo de costes de reparto que se empleará posteriormente en las simulación.

En definitiva, el modelo de costes de reparto para el código postal 08174 constará de las variables número de repartos por día, peso total repartido por día y peso total repartido por día elevado al cuadrado, con los siguientes parámetros de regresión respectivamente, 2,01, 0,048 y -0,00006. El valor umbral para este modelo de regresión es igual a -0,17.

2.5.4.2. Adecuación del modelo de costes de reparto

Cuando los errores se distribuyen con una distribución de probabilidad no normal, si éstos son independientes y se distribuyen idénticamente, las inferencias bajo la teoría de normalidad son asintóticamente correctas. Por tanto, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. Este análisis gráfico se puede ver en la figura 98.

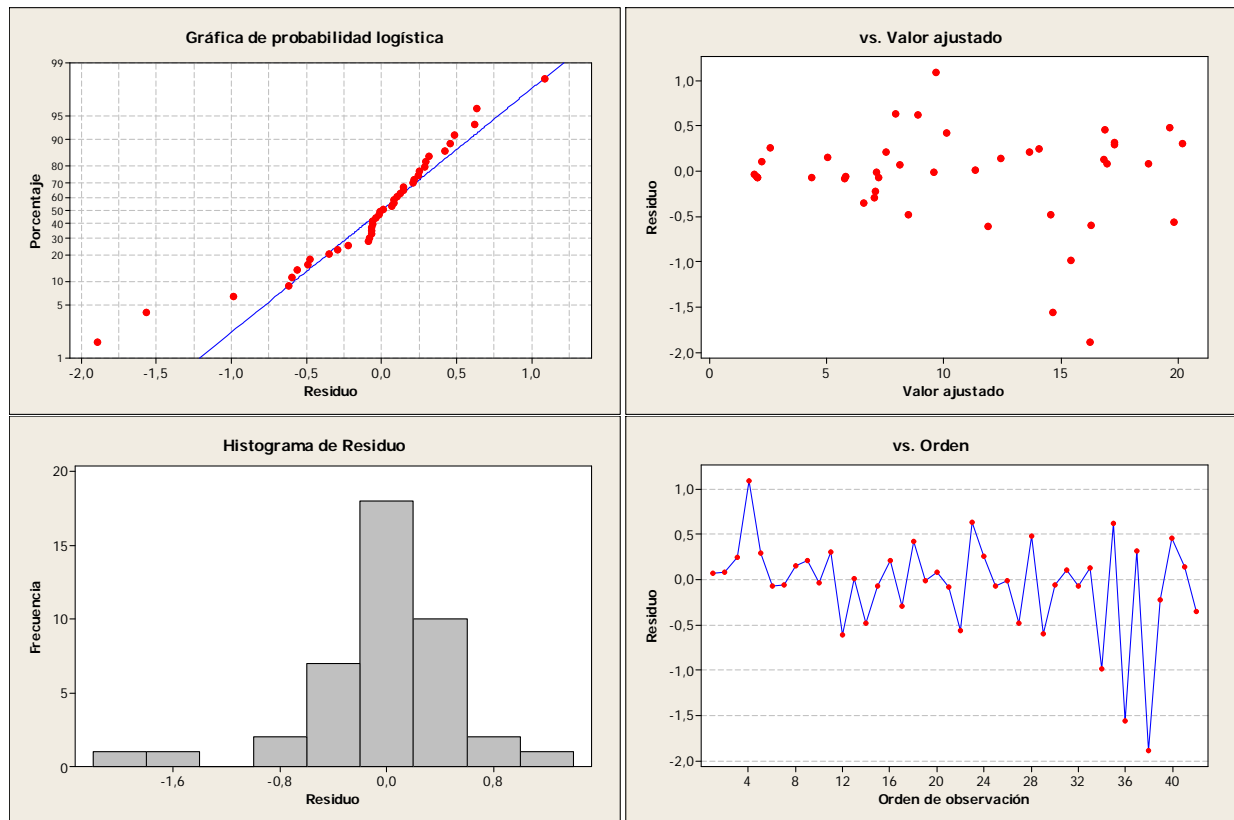


Figura 98: Análisis gráficos de los residuales para modelo de costes de reparto vehículo 1500 kg. PMA código postal 08174

En la primera gráfica se puede observar como los errores se distribuyen de manera normal ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.5.4.3. Validación del modelo de costes de reparto

Para comprobar que el modelo predice de manera adecuada los valores de los costes de reparto ante nuevos valores de las variables exógenas, se realiza la prueba t pareada o prueba de Welch.

Los nuevos valores de las variables exógenas son los valores de la muestra de predicción, y la validación del modelo se lleva a cabo comparando uno a uno los valores de la variable endógena estimados por el modelo tras introducir los nuevos valores de las variables exógenas, con los valores correspondientes de los costes de reparto de la muestra de predicción. Los valores Z que se emplean para realizar en esta prueba de Welch se obtienen de la tabla 220.

Tabla 220: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_4$	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
-0,170454	2,0097336	0,0480078	-0,00006			
	4	778	605284	6,91	24,66	-17,75
	4	108	11664	12,32	13,66	-1,34
	2	47	2209	5,97	5,80	0,16
	2	49	2401	6,05	6,13	-0,08
	4	67	4489	10,80	10,57	0,23
	2	24	576	4,96	4,88	0,09
	2	36	1296	5,50	5,59	-0,09
	2	24	576	4,96	5,17	-0,21
	1	8	64	2,22	2,33	-0,11
	2	29	841	5,19	5,44	-0,25
	8	315	99225	24,75	26,01	-1,26
	3	94	8836	9,81	10,06	-0,25
	6	69	4761	14,90	14,41	0,49
	3	77	5929	9,18	9,42	-0,24
	6	351	123201	20,94	23,94	-3,00
	5	68	4624	12,85	13,75	-0,90
	4	47	2209	9,99	10,15	-0,16
	5	121	14641	14,76	14,29	0,47
	3	17	289	6,66	6,50	0,15
	0	0	0	-0,17	0,00	-0,17
	1	9	81	2,27	2,34	-0,08
	1	12	144	2,41	2,77	-0,37
	4	100	10000	12,04	11,15	0,89
	4	120	14400	12,72	13,91	-1,19
	6	155	24025	17,81	17,68	0,13
	3	8	64	6,24	5,74	0,50
	3	228	51984	13,51	13,60	-0,09
	9	245	60025	25,88	27,89	-2,01
	3	17	289	6,66	6,50	0,15
	1	4	16	2,03	1,96	0,07
	2	54	2916	6,26	6,59	-0,33
	1	8	64	2,22	2,33	-0,11
	3	43	1849	7,81	8,52	-0,71
	6	65	4225	14,74	14,69	0,05
	3	97	9409	9,92	9,53	0,39
	2	50	2500	6,09	5,90	0,19
	4	404	163216	16,93	18,90	-1,96
	1	14	196	2,50	2,81	-0,31
	4	141	19881	13,38	14,37	-0,99
	5	58	3364	12,45	12,57	-0,12
	3	893	797449	-1,74	21,76	-23,50
	3	14	196	6,52	6,22	0,30

Y la tabla que resume los resultados de esta prueba de Welch es la tabla 221.

Tabla 221: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08174

	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	9,22	10,49	-1,269	Límite superior I.C	$t_{0,025;39}$
Varianza	39,42	50,01	20,16	0,089	1,96
Varianza/n			0,480		
Máximo	25,88	27,89		Límite inferior I.C	
Mínimo	-1,74	0,00		-2,627	

El intervalo de confianza obtenido para esta prueba contiene el valor cero, por lo que se concluye que el modelo estima nuevos costes de reparto de manera adecuada. Para elaborarlo se ha empleado la distribución *t* de Student, con un nivel de significancia del 95% y 39 grados de libertad, cuyo valor es 1.96.

2.5.5. Modelo de costes de recogida

Una vez finalizado el modelo de los costes de reparto, hay que realizar el modelo de costes de recogida. La creación este modelo es análogo al de los costes de reparto, por lo que los pasos a seguir son idénticos.

Tabla 222: VMV modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
7,410					-86,3		0%	3,98
0,719	3,096				-63,6	22,692	76%	1,95
0,225	2,550	0,022			-36,4	27,207	96%	0,82
0,364	2,418	0,022	0,023		-36,4	0,029	96%	0,84
-0,095	2,332	0,044		-0,00007	-29,8	6,619	97%	0,68

Tabla 223: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,003					-86,3		0%	3,98
1,275	0,310				-67,3	18,999	70%	2,20
1,166	0,272	0,002			-50,9	16,459	89%	1,31
0,781	0,584	0,002	-0,049		-47,8	3,110	91%	1,21
0,632	0,521	0,008	-0,046	-0,00002	-38,4	9,401	95%	0,91

Tabla 224: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-19884					-109,9		-357%	8,52
0,208	-0,030				-71,0	38,899	62%	2,47
0,221	-0,027	-0,0002			-61,3	9,671	79%	1,84
0,280	-0,074	-0,0001	0,007		-59,2	2,142	81%	1,75
0,277	-0,074	-0,0001	0,007	0	-59,1	0,089	80%	1,78

Tabla 225: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08174

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-2,722					-86,3		0%	3,98
-1,594	-0,498				-65,3	21,020	73%	2,06
-1,455	-0,428	-0,004			-43,1	22,207	93%	1,02
-1,150	-0,686	-0,003	0,042		-41,2	1,873	94%	0,98
-1,329	-0,376	-0,010		0,00002	-34,3	6,928	96%	0,78

2.5.5.1. Selección de variables del modelo de costes de recogida

En la tabla referente al estudio de máxima verosimilitud con los errores bajo la suposición de que se distribuyen de acuerdo a función normal, se puede observar como al introducir las variables exógenas número de recogidas y peso total recogido por día, el cociente de verosimilitud es mayor que dos unidades cuando se introducen cada una de estas variables de una en una en el modelo. Sin embargo, al introducir la variable exógena número de recogidas elevada a la segunda potencia, el cociente de verosimilitud es de 0,029 unidades, por lo que la adición de esta variable no aporta un mejor ajuste del modelo. Por último, al calcular el valor de máxima verosimilitud con las dos primeras variables exógenas más la variable peso total recogido por día elevada al cuadrado, el cociente de verosimilitud entre el modelo con dos variables y el modelo con estas tres variables es de más de 6 unidades, por lo que esta variable sí que aporta un mejor ajuste del modelo, y se empleará en el modelo de costes de reparto que se empleará posteriormente en las simulación.

En definitiva, el modelo de costes de recogida para el código postal 08174 constará de las variables número de recogidas por día, peso total recogido por día y peso total recogido por día elevado al cuadrado, con los siguientes parámetros de regresión respectivamente, 2,332, 0,044 y -0,00007. El valor umbral para este modelo de regresión es igual a -0,095.

2.5.5.2. Adecuación del modelo de costes de recogida

Como el mayor valor máximo de verosimilitud es el que aporta una distribución de probabilidad normal, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. Este análisis gráfico se puede ver en la figura 99.

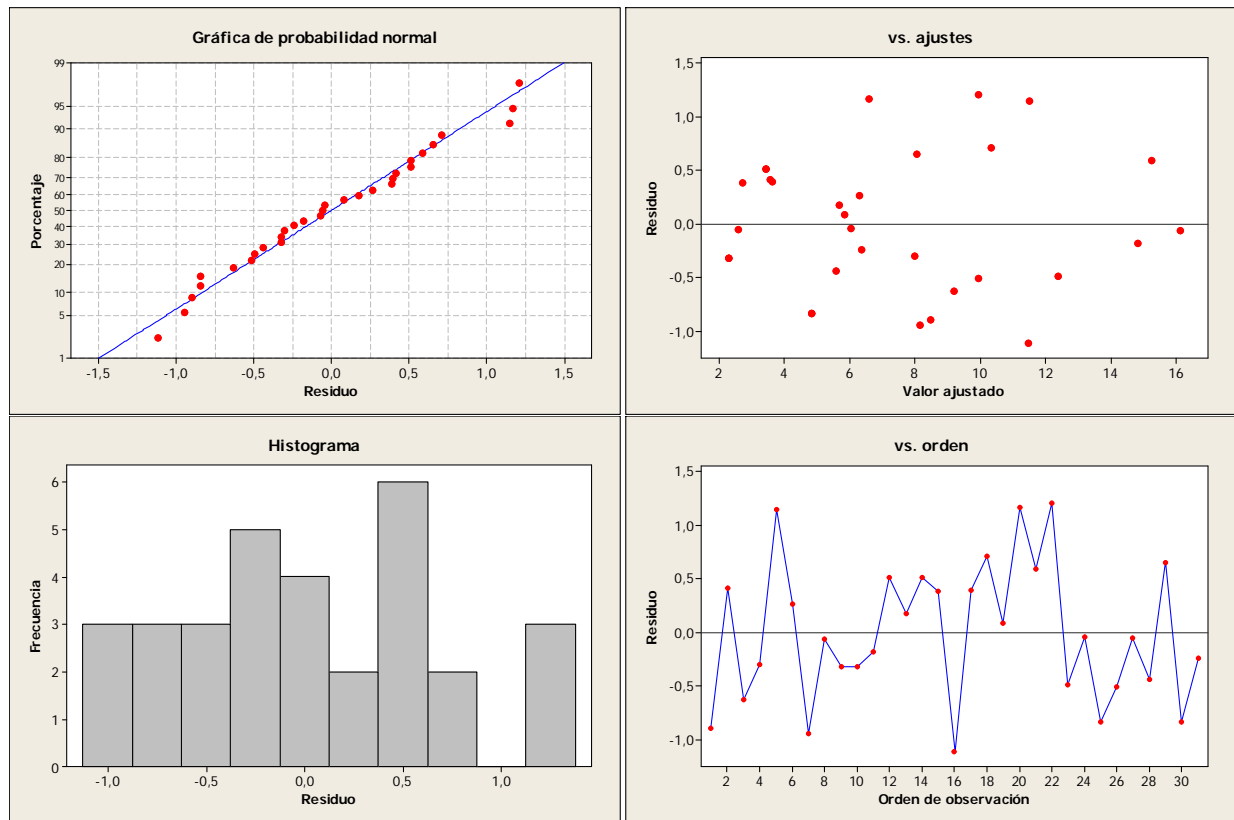


Figura 99: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08174

En la primera gráfica se puede observar como los errores se distribuyen de manera normal ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.5.5.3. Validación del modelo de costes de recogida

Para comprobar que el modelo predice de manera adecuada los valores de los costes de reparto ante nuevos valores de las variables exógenas, se realiza la prueba t pareada o prueba de Welch, con los datos de la muestra de predicción previamente obtenida. Los valores Z con los que se elabora este cuadro provienen de la tabla 231.

En él se puede observar como el modelo estima los costes de recogida de manera adecuada, ya que el intervalo de confianza creado incluye el cero. Para elaborarlo se ha empleado la distribución *t* de Student, con un nivel de significancia del 95% y 38 grados de libertad, cuyo valor es 1.96.

2.5.6. *Modelo de carga de trabajo*

En este modelo se va a identificar que distribución de probabilidad siguen las variables aleatorias con las que se puede modelar la carga la de trabajo, así como los parámetros que las caracterizan. Estas variables son el número de repartos por día y el peso de cada uno de los repartos, en el caso de los repartos, y el número de recogidas por día y el peso de cada recogida, en el de las recogidas.

Las muestras que se emplean para las variables número de repartos, peso total repartido, número de recogidas y peso total recogido son las mismas muestras que las muestras de calibración que se emplearon para ajustar los modelos de los costes de reparto y recogida. Las muestras para las variables peso de cada reparto y peso de cada recogida, es una muestra con los pesos unitarios de cada uno de los repartos y de cada una de las recogidas.

2.5.6.1. Número de repartos por día

Minitab busca los parámetros para cada una de los distintos tipos de distribuciones de probabilidad, siendo estos parámetros tales que la distribución de probabilidad que caracterizan, es la que más se aproxima a los datos de la muestra. Una vez que se tienen estos datos hay que comprobar que la distribución de probabilidad que conforma se ajuste suficientemente bien a los datos. Para realizar esta comprobación se realizan los denominados test no paramétricos de bondad de ajuste. Debido a la naturaleza de esta variable aleatoria, la cual es un conteo de algún elemento, la primera función de distribución de probabilidad que se va a tratar de comprobar el ajuste es la distribución de Poisson. Como esta distribución es discreta se emplea el test no paramétrico de bondad de ajuste de Chi cuadrado. El resultado de este test se puede ver en la tabla 233.

Tabla 233: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08174

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
1	6	0,143	0,1171	5	0,2385	alfa=0,05 k=6 Par. Est. =1 n=6-1-1=4
2	6	0,143	0,1700	7	0,1817	
3	9	0,214	0,2091	9	0,0054	
4	8	0,190	0,1929	8	0,0013	
5	5	0,119	0,1424	6	0,1607	
6-8	8	0,190	0,1551	7	0,3399	
	42				0,93	9,488

El valor lambda dado por Minitab para una distribución de Poisson con los datos de la muestra de esta variable aleatoria es 3,69. En la tabla se puede observar como el valor límite para aceptar la hipótesis

de que la una distribución de Poisson con lambda 3,69 ajusta adecuadamente la variable número de repartos por día es 9,488. Este valor corresponde a una distribución Chi cuadrado con 4 grados de libertad y nivel de significancia del 5%. Como en este caso el valor del test es 0,93, menor al valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.5.6.2. Peso de cada reparto

Para esta variable continua se presentan los datos aportados por Minitab para cada uno de los parámetros de cada distribución, así como los resultados de los test de Kolmogorov-Smirnov.

Tabla 234: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08174

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,276	alfa	0,146	alfa	0,110	Media	0,064	Lambda	0,263
23,82	3,473	0,7006	1,827	0,7617	1,380	2,308	0,803	27,48	3,285
Desv. Est.	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	Desv. Est.	$c_{\alpha} = 1,358$	Theta	$c_{\alpha} = 1,358$
39,71	$\alpha = 0,05$	34	$\alpha = 0,05$	19,57	$\alpha = 0,05$	1,311	$\alpha = 0,05$	24,7	$\alpha = 0,05$

Como se puede ver la única distribución que supera satisfactoriamente el test de Kolmogorov-Smirnov es la distribución lognormal. Por lo tanto, para la simulación se empleará esta distribución para estimar los pesos de cada uno de los repartos, y no será necesario emplear la distribución de probabilidad calculada anteriormente para la variable peso total repartido por día.

2.5.6.3. Número de recogidas por día

Esta variable al igual que el número de repartos por día, es un conteo de algún elemento, y por tanto, es una distribución discreta. Por ser una distribución discreta, la primera distribución que se va a comprobar si ajusta adecuadamente es la distribución de Poisson. El resultado de la prueba de bondad de ajuste Chi cuadrado es el de la tabla 235.

Tabla 235: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08174

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
0	12	0,279	0,2105	9	0,9596	alfa=0,05 k=4 Est. par. =1 n=4-1-1=2
1	9	0,209	0,3280	14	1,8478	
2	14	0,326	0,2556	11	0,8250	
3-5	8	0,186	0,2059	9	0,0822	
	43				3,71	

El valor lambda dado por Minitab para una distribución de Poisson con los datos de la muestra de esta variable aleatoria es 1,56. En la tabla se puede observar como el valor límite para aceptar la hipótesis de que la una distribución de Poisson con lambda 1,56 ajusta adecuadamente la variable número de repartos por día es 5,991. Este valor corresponde a una distribución Chi cuadrado con 2 grados de libertad y nivel de significancia del 5%. Como en este caso el valor del test es 3,71, menor al valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.5.6.4. Peso de cada recogida

Por último, para finalizar con la elaboración del modelo de carga de trabajo para este código postal se va a analizar la variable peso de cada recogida. Esta variable es una variable continua, y las distribuciones a analizar son también continuas. Los parámetros de las distintas distribuciones analizadas, así como los resultados de los test de Kolmogorov-Smirnov de cada una se pueden ver en la tabla 236.

Tabla 236: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08174

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,251	alfa	0,102	alfa	0,094	Media	0,089	Lambda	0,244
35,03	2,068	0,6223	0,835	0,7168	0,775	2,569	0,729	24,43	2,004
Desv. Est.	c _α = 1,358	beta	c _α = 1,358	beta	c _α = 1,358	Desv. Est.	c _α = 1,358	Theta	c _α = 1,358
53,42	α = 0,05	53,29	α = 0,05	27,67	α = 0,05	1,53	α = 0,05	22,29	α = 0,05

Como se puede observar las únicas distribuciones que han superado el test de Kolmogorov-Smirnov han sido las distribuciones gamma, Weibull y lognormal, siendo la distribución lognormal la que menos discrepancia presenta. Por lo tanto, para la simulación se empleará esta distribución para estimar los pesos de cada uno de los repartos, y no será necesario emplear la distribución de probabilidad calculada anteriormente para la variable peso total repartido por día.

2.5.7. Simulación de los costes de reparto y recogida, y de la carga de trabajo

En este apartado se llevará a cabo la simulación de los modelos creados para el estudio del vehículo de 1500 kg. de peso máximo autorizado en el código postal 08198. En primer lugar se simulará la carga de trabajo. Para ello hay que determinar el número de repartos y de recogidas para cada uno de los días simulados, y el peso total repartido y el peso total recogido por día. Empleando los ajustes de las distribuciones de probabilidad de Poisson obtenidos en los apartados 2.5.6.1 y 2.5.6.3. se simularán el número de repartos y el número de recogidas para cada uno días que se desean simular.

Posteriormente, se simulará el peso correspondiente de cada uno de los repartos y las recogidas mediante las distribuciones de probabilidad calculadas en los apartados 2.5.6.2. y 2.5..6.4. Sumando el peso unitario de todos los repartos simulados en cada uno de los días, se calculará el peso total repartido por día. Con el peso unitario de las recogidas simulado para cada uno de los días, se calculará el peso total recogido por día.

Finalmente, se simularán los modelos de costes de reparto y costes de recogida. Como valores de entrada a estos modelos se emplearán los datos simulados para el modelo de carga de trabajo. El número de repartos y el peso total repartido por día serán los datos de entrada para el modelo de costes de reparto. Análogamente, el número de recogidas y el peso total recogido por día serán los valores de entrada para el modelo de costes de recogida.

2.5.7.1. Simulación de los costes de reparto

La carga de trabajo relativa a los repartos para este código postal es la siguiente: el número de repartos puede ser descrito mediante una distribución de probabilidad Poisson con parámetro lambda igual a 3.69. Mientras que el peso de cada reparto sigue una distribución de probabilidad lognormal con media 2,308 y desviación estándar 1,311.

Un resumen de la carga de trabajo simulada es la tabla 237, en el cual se puede observar el promedio de repartos por día simulados, su desviación estándar, el número máximo de repartos a realizar en un día, así como el número mínimo. Y los mismos datos para la variable peso total repartido, siendo este peso total repartido la suma de los pesos simulados para cada una de los repartos.

Tabla 237: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08174

	Repartos por día	Peso total repartido por día
Promedio	4	76
Desv. Estan.	2	73
Máximo	13	408
Mínimo	0	0

Esta simulación de la carga de trabajo se emplea como entrada para el modelo de costes de reparto. Como se han simulado 250 días de carga de trabajo, se obtendrán los estimados puntuales de costes de reparto para 250 días. Los resultados que aquí se exponen son el valor medio de límite inferior de las 250 simulaciones, la varianza y desviación estándar de este límite inferior, así como un intervalo de confianza de este límite inferior. Para el límite superior del intervalo de predicción los resultados que se presentan son análogos. Estos resultados son los de la tabla 238.

Tabla 238: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08174

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	9,24	11,54
Varianza	34	35
Desv. Estan.	5,79	5,88
I.C	8,63	10,93
	9,84	12,15

En la tabla se puede observar como el límite inferior de los estimados puntuales de los costes de repartos se distribuyen con un valor medio de 9,24 euros por día y una desviación estándar de 5,79. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 11,54 y una desviación estándar de 5,88 euros por día.

2.5.7.2. Simulación de los costes de recogida

La carga de trabajo relativa a las recogidas para este código postal queda como sigue: el número de recogidas puede ser expresado mediante una distribución de probabilidad Poisson con un parámetro lambda igual a 1,56. Por otro lado, el peso de cada recogida puede ser descrito mediante una distribución lognormal con un valor medio igual a 2,569 y desviación estándar 1,53. Un cuadro resumen de la carga de trabajo simulada es la tabla 239, en el cual se puede observar el promedio de recogidas por día simulados, su desviación estándar, el número máximo de recogidas a realizar en un día, así como el número mínimo. Y los datos análogos para la variable peso total recogido, siendo este peso total recogido la suma de los pesos simulados para cada una de las recogidas.

Tabla 239: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08174

	Recogidas por día	Peso total recogido por día
Promedio	2	52
Desv. Estan.	1	86
Máximo	6	486
Mínimo	0	0

Y la simulación de los costes de recogida correspondiente a esta carga de trabajo para 250 días se resume en la tabla 240.

Tabla 240: resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08174

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	3,41	6,02
Varianza	20	24
Desv. Estan.	4,47	4,91
I.C	2,95	5,50
	3,88	6,53

Donde se puede comprobar como el límite inferior de los estimados puntuales de los costes de recogida se distribuyen con un valor medio de 3,41 euros por día y una desviación estándar de 4,47. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 6,02 y una desviación estándar de 4,91 euros por día.

2.6. Código postal 08173

Como en este código postal se dan solamente repartos de menos de 300 kg. de peso y recogidas de menos de 700 kg. existe únicamente volumen de trabajo para un vehículo de 1500 kg. de peso máximo autorizado. Hay que analizar por tanto los costes de reparto y recogida, así como las carga de trabajo para un vehículo de este tipo.

2.6.1. Toma de muestras

Como el resto de códigos postales estudiados hasta este momento el primer paso para la realización de los modelos, una vez identificadas las variables que se desean modelar, es la toma de muestras de cada una de estas variables. Las variables serán las mismas que para el resto de modelos que se realicen, y son los costes de reparto y los costes de recogida. Las variables exógenas que determinan estas variables endógenas son el número de repartos y el peso total repartido, para el modelo de los costes de reparto, y el número de recogidas y el peso total recogido para el caso de la variable costes de recogida. Las 84 observaciones recopiladas para el modelo de costes de reparto como para el modelo costes de recogida son las tablas 241 y 242 respectivamente.

Tabla 241: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08173

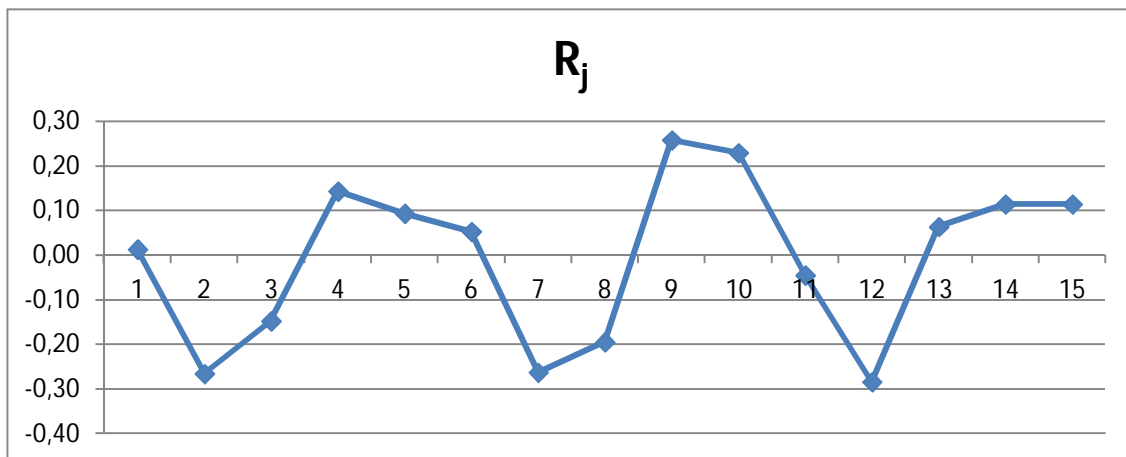
Día	REPARTOS			Día	REPARTOS		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	7	442	25	43	6	213	20
2	2	262	11	44	4	34	9
3	3	35	8	45	4	174	13
4	6	162	18	46	0	0	0
5	4	57	11	47	5	42	11
6	7	251	24	48	7	469	27
7	4	146	13	49	8	66	19
8	4	64	11	50	7	215	20
9	6	241	21	51	4	205	15
10	4	124	13	52	8	89	20
11	14	244	38	53	7	78	17
12	7	41	15	54	2	248	11
13	3	62	9	55	4	273	15
14	5	154	15	56	4	19	8
15	5	77	14	57	4	30	9
16	4	109	12	58	5	77	13
17	3	136	11	59	5	27	10
18	2	47	6	60	3	13	6
19	9	91	21	61	9	88	20
20	3	7	6	62	6	320	24
21	6	167	18	63	3	102	11
22	4	210	13	64	7	105	17
23	2	30	5	65	6	65	15
24	5	114	15	66	2	118	9
25	7	238	25	67	8	362	27
26	5	87	13	68	6	346	21
27	0	0	0	69	7	90	18
28	7	154	21	70	4	20	8
29	4	115	13	71	4	86	11
30	6	90	17	72	4	64	10
31	5	161	17	73	4	29	9
32	8	81	19	74	10	303	30
33	5	49	12	75	5	82	13
34	3	156	11	76	3	264	13
35	5	177	16	77	8	125	21
36	4	63	10	78	8	139	22
37	6	103	16	79	5	28	11
38	5	196	16	80	2	53	7
39	5	16	10	81	5	57	13
40	5	74	13	82	4	73	12
41	5	44	11	83	6	240	18
42	9	30	17	84	3	110	10

Tabla 242: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08173

Día	RECOGIDAS			Día	RECOGIDAS		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	2	59	7	43	1	85	5
2	2	60	7	44	0	0	0
3	1	83	5	45	4	32	10
4	1	33	4	46	3	398	17
5	2	171	10	47	3	188	12
6	3	375	17	48	2	9	4
7	3	28	8	49	1	1	2
8	1	7	3	50	2	66	7
9	1	166	6	51	2	93	8
10	5	47	12	52	1	379	10
11	2	5	4	53	2	52	7
12	7	219	24	54	2	17	5
13	3	159	13	55	3	62	9
14	1	74	5	56	3	72	9
15	1	22	4	57	2	35	6
16	0	0	0	58	1	11	3
17	2	443	12	59	2	9	4
18	2	70	8	60	3	156	11
19	2	84	8	61	3	68	9
20	2	559	15	62	2	171	9
21	1	32	4	63	4	197	15
22	2	142	8	64	1	50	4
23	1	231	8	65	3	81	10
24	1	294	8	66	1	41	4
25	3	27	8	67	1	208	7
26	1	75	5	68	1	221	7
27	2	59	7	69	1	176	7
28	2	91	9	70	1	34	4
29	3	338	14	71	2	167	9
30	1	68	5	72	1	45	4
31	7	211	24	73	2	42	6
32	3	68	9	74	4	429	20
33	4	82	11	75	3	688	20
34	1	31	4	76	2	44	6
35	2	56	7	77	2	22	6
36	1	28	4	78	5	162	19
37	1	298	9	79	1	13	3
38	0	0	0	80	1	27	4
39	2	72	7	81	4	455	19
40	1	20	3	82	2	13	5
41	2	47	8	83	3	122	11
42	1	437	10	84	3	14	7

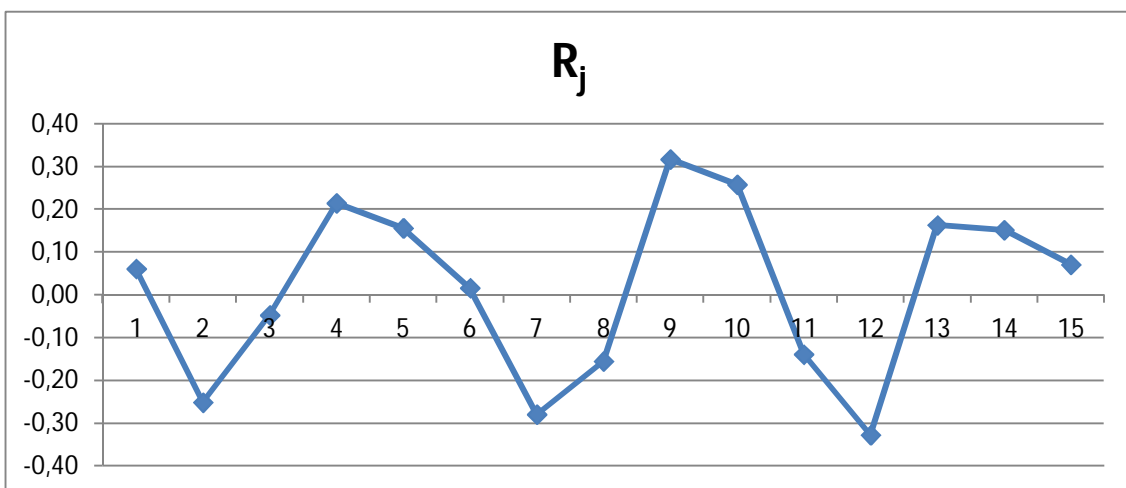
2.6.2. Independencia de las muestras

Tras la obtención de las muestras de las variables mediante el estudio observacional diseñado, el siguiente paso es comprobar la independencia de cada una de las observaciones de las muestras para poder aplicar con garantías el método de máxima verosimilitud y los test no paramétricos de bondad de ajuste. En primer lugar se comprobará la independencia de las observaciones de la muestra correspondiente a los repartos. Para cada una de las variables, una endógena y dos exógenas, se realizan las gráficas de autocorrelación correspondientes.



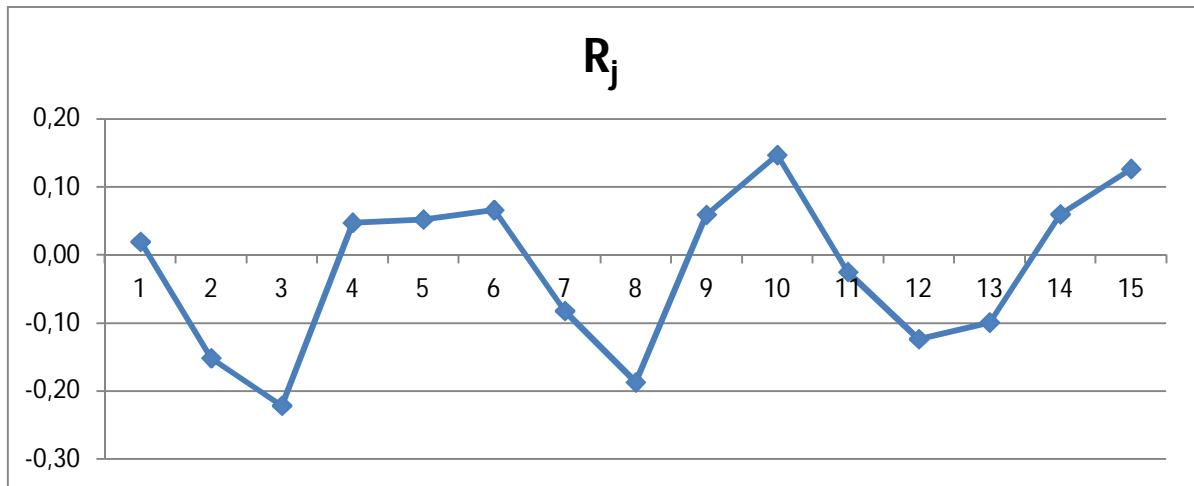
R_j	-0,05	-0,14	0,01	-0,14	0,08	-0,07	0,02	-0,21	-0,11	0,26	-0,09	-0,09	0,14	0,04	-0,05
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-2	-6	0	-6	3	-3	1	-8	-4	10	-4	-4	6	2	-2

Figura 100: Grafica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08173



R_j	-0,07	-0,17	-0,02	-0,10	0,02	-0,13	-0,03	-0,05	-0,02	0,20	-0,01	-0,08	0,10	-0,02	-0,17
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-0,3	-0,9	-0,1	-0,5	0,1	-0,6	-0,2	-0,2	-0,1	1,0	0,0	-0,4	0,5	-0,1	-0,8

Figura 101: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08173



R_j	-0,01	-0,12	0,01	-0,15	0,07	-0,03	0,13	-0,08	-0,08	0,04	-0,21	0,00	0,09	0,11	-0,08
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-71	-1265	67	-1537	664	-281	1294	-798	-866	458	-2108	33	944	1133	-811

Figura 102: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08173

En las tres gráficas se puede observar como en ninguna de ellas los valores de R_j se encuentran en el rango de -1 y 1, y que en ninguna de ellas se puede apreciar un patrón determinado. Por tanto, esta prueba muestra la independencia de las observaciones de esta muestra.

A continuación, una vez expuestas las gráficas de autocorrelación, se muestra las gráficas de dispersión para las mismas variables, como segunda comprobación de la independencia de las observaciones. Estas gráficas son las figuras 103, 104 y 105

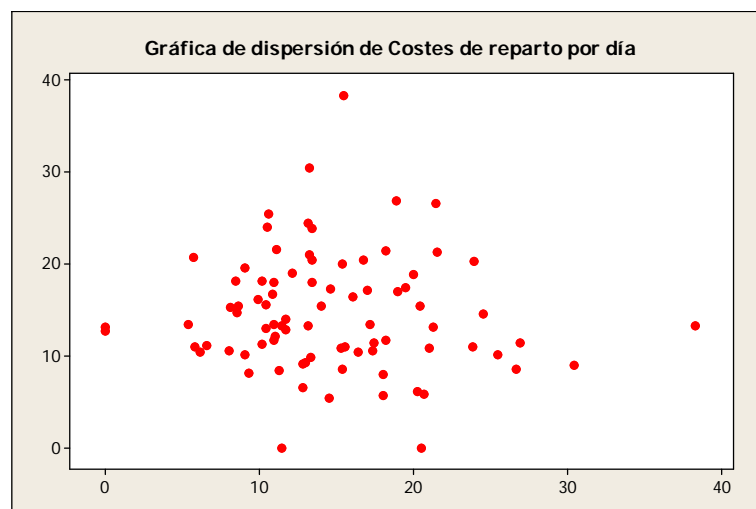


Figura 103: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08173

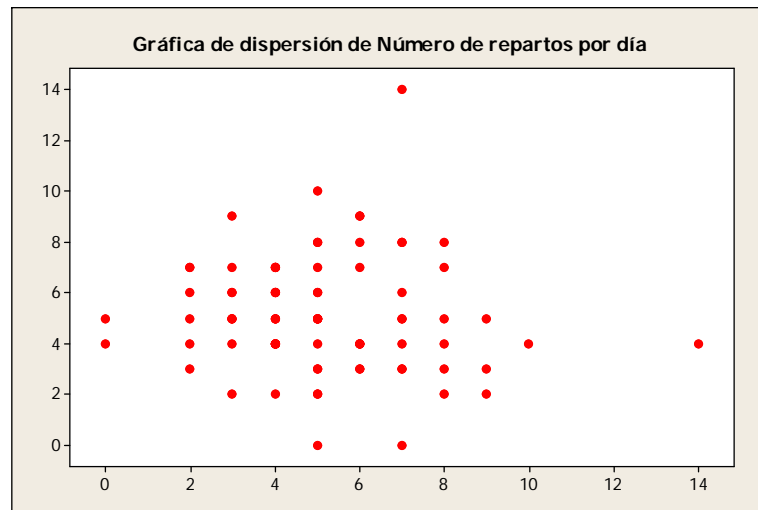


Figura 104: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08173

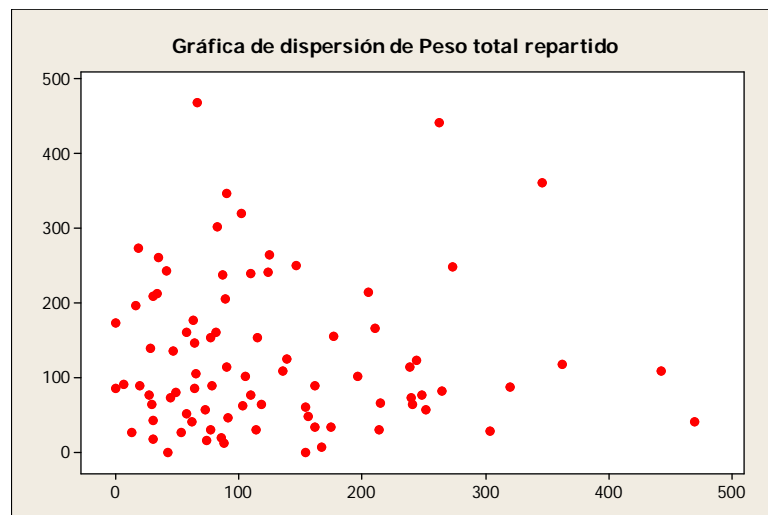
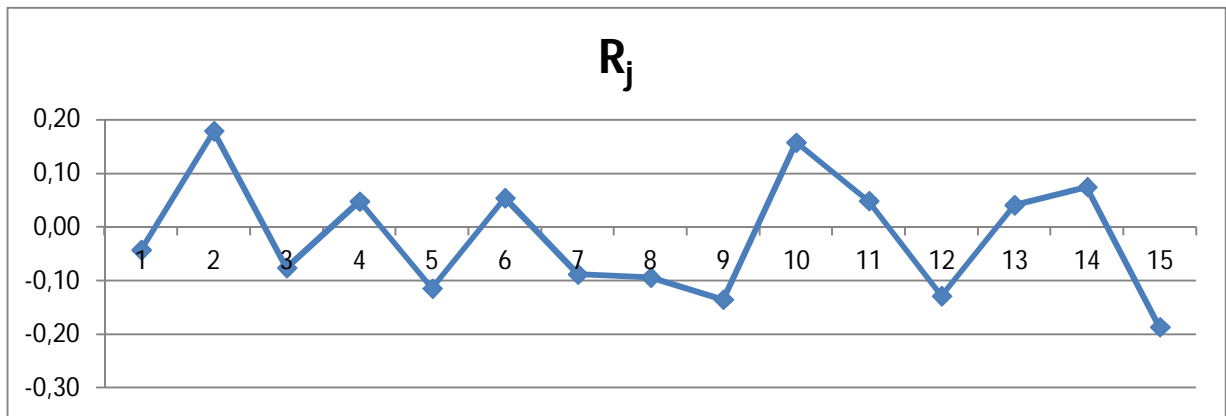


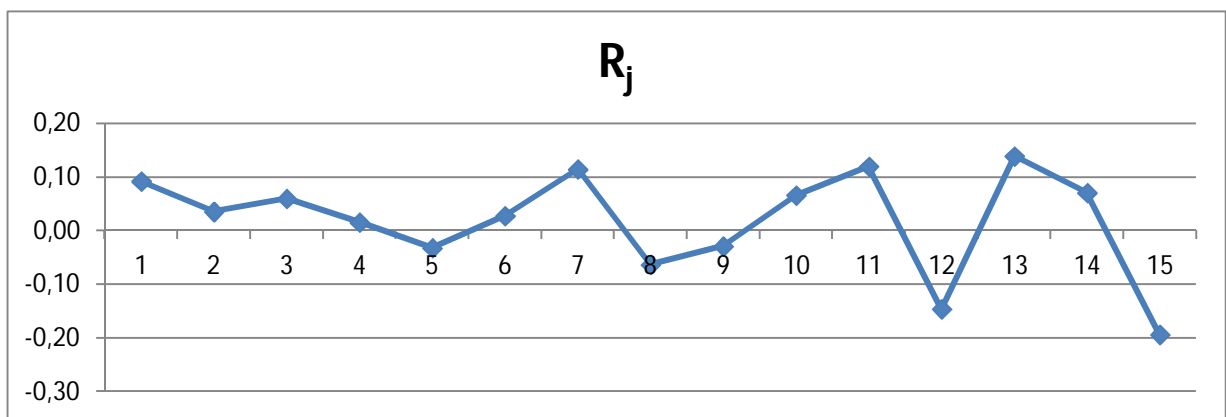
Figura 105: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08173

En estas gráficas se puede comprobar cómo los puntos están distribuidos aleatoriamente por los planos, lo cual indica que no existe correlación entre las observaciones y éstas son independientes. Una vez comprobado la independencia de la muestra correspondiente a los repartos, se comprobará la independencia de la muestra de las recogidas. De la misma manera que para los repartos, en primer lugar se realizarán las gráficas de autocorrelación.



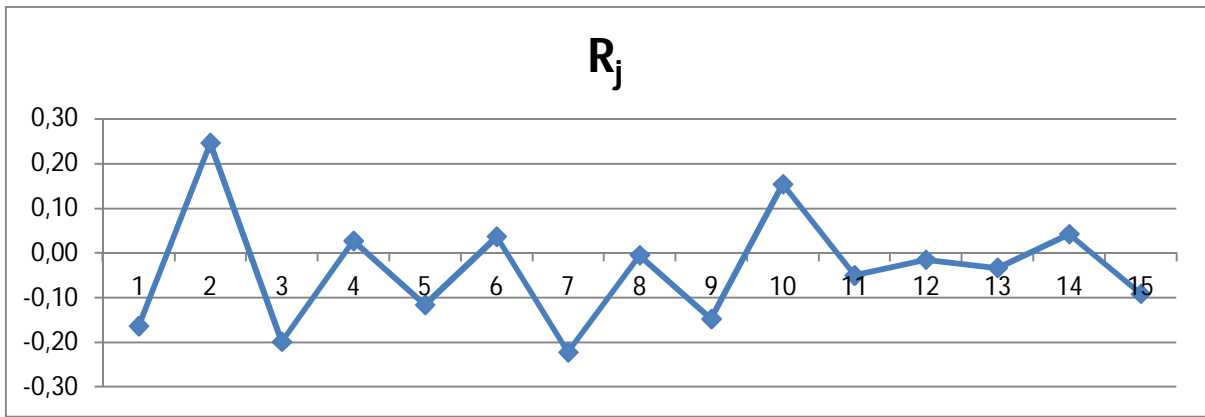
R_j	-0,02	0,11	-0,06	0,05	-0,12	0,05	-0,09	-0,10	-0,14	0,15	0,05	-0,13	0,04	0,07	-0,19
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	0	-3	-1	-2	-2	2	0	-1	-4	0	2	0	0	3	1

Figura 106: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08173



R_j	0,07	0,17	-0,07	-0,08	-0,11	0,01	-0,15	-0,07	-0,23	-0,03	-0,20	-0,06	-0,04	0,09	0,06
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	0,1	0,3	-0,1	-0,1	-0,2	0,0	-0,3	-0,1	-0,4	-0,1	-0,4	-0,1	-0,1	0,1	0,1

Figura 107: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08173



R_j	-0,04	0,25	-0,20	0,01	-0,03	0,18	-0,25	0,07	-0,10	-0,07	0,06	-0,05	-0,03	0,05	-0,02
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	-790	-5070	132	-528	-696	3602	1381	-1030	-1998	-1374	1279	-1017	-700	1036	-386

Figura 108: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08173

En este caso también las gráficas muestran como no existe ningún patrón determinado y los valores R_j no superan en ningún caso los valores -1 y 1, por lo que esta prueba determina que las observaciones de muestra correspondiente a las recogidas son independientes. Las gráficas de dispersión para estas observaciones se muestran a continuación:

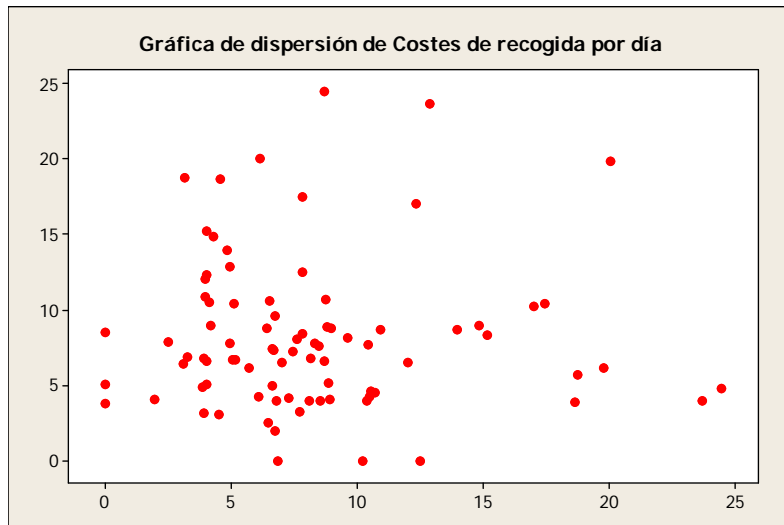


Figura 109: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08173

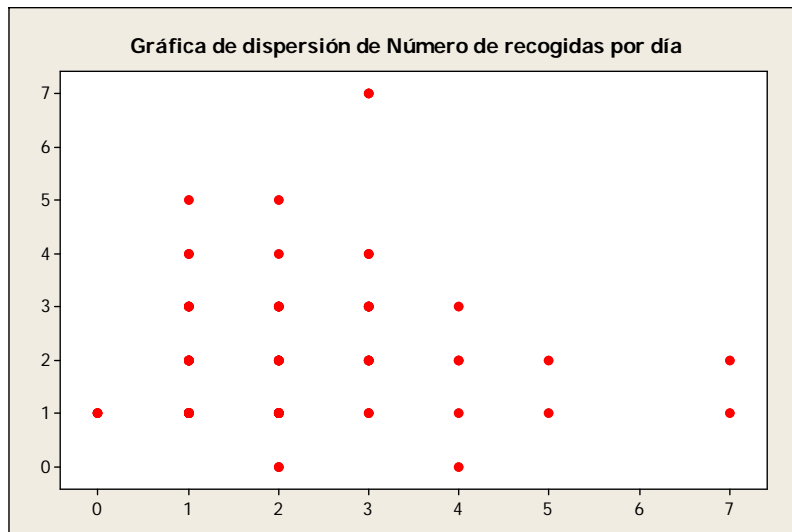


Figura 110: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08173

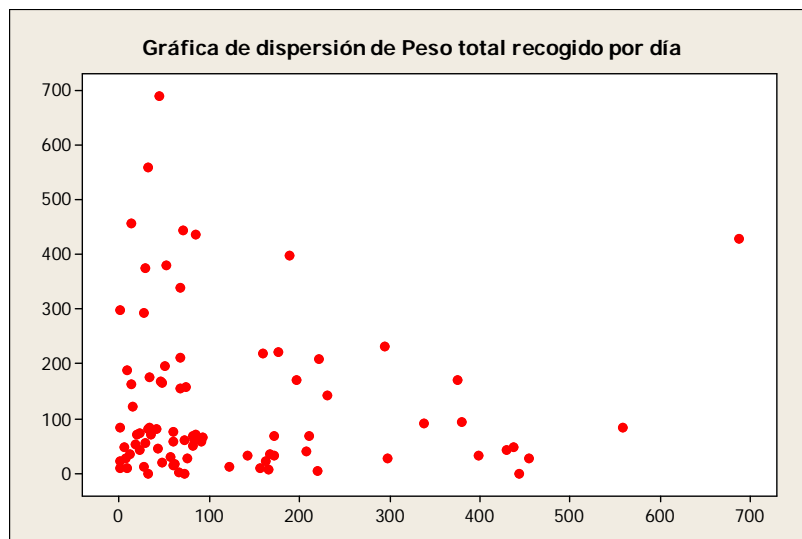


Figura 111: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08173

Como los puntos de todos los gráficos se distribuyen de manera aleatoria a lo largo de cada plano correspondiente, se concluye que las observaciones de la muestra son independientes.

2.6.3. Separación de las muestras de calibración y predicción

Para la elaboración de los modelos de costes de reparto y costes de recogida es necesario el separar las 84 observaciones en la muestra de calibración y la muestra de predicción. La primera se emplea para la creación del modelo y la segunda para la validación del mismo. Esta separación se realiza mediante asignación aleatoria. Se asigna un número aleatorio proveniente de una distribución uniforme a cada una de las observaciones, y si éste es mayor o igual que 0.5, esta observación se incluirá en la muestra de predicción. Por el contrario, si es menor de 0.5, se incluye en la muestra de calibración.

La asignación de números aleatorios a cada una de las observaciones de reparto se muestra en la tabla 243.

Tabla 243: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08173

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,144	22	0,261	43	0,494	64	0,217
2	0,169	23	0,029	44	0,275	65	0,156
3	0,112	24	0,928	45	0,816	66	0,348
4	0,372	25	0,182	46	0,061	67	0,237
5	0,049	26	0,883	47	0,019	68	0,784
6	0,280	27	0,098	48	0,737	69	0,779
7	0,366	28	0,509	49	0,769	70	0,842
8	0,382	29	0,048	50	0,415	71	0,239
9	0,356	30	0,338	51	0,524	72	0,735
10	0,669	31	0,858	52	0,344	73	0,723
11	0,945	32	0,003	53	0,969	74	0,686
12	0,831	33	0,106	54	0,816	75	0,665
13	0,067	34	0,926	55	0,961	76	0,803
14	0,199	35	0,339	56	0,817	77	0,325
15	0,820	36	0,952	57	0,349	78	0,355
16	0,025	37	0,091	58	0,947	79	0,630
17	0,214	38	0,995	59	0,379	80	0,817
18	0,927	39	0,736	60	0,235	81	0,376
19	0,964	40	0,829	61	0,936	82	0,420
20	0,817	41	0,216	62	0,348	83	0,627
21	0,541	42	0,814	63	0,542	84	0,485

Quedando la muestra de calibración y la de predicción de la que aparece en la tabla 244.

Tabla 244: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08173

Día	REPARTOS. MUESTRA DE CALIBRACIÓN			Día	REPARTOS. MUESTRA DE PREDICCIÓN		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	4	124	13	1	7	442	25
2	14	244	38	2	2	262	11
3	7	41	15	3	3	35	8
4	5	77	14	4	6	162	18
5	2	47	6	5	4	57	11
6	9	91	21	6	7	251	24
7	3	7	6	7	4	146	13
8	6	167	18	8	4	64	11
9	5	114	15	9	6	241	21
10	5	87	13	10	3	62	9
11	7	154	21	11	5	154	15
12	5	161	17	12	4	109	12
13	3	156	11	13	3	136	11
14	4	63	10	14	4	210	13
15	5	196	16	15	2	30	5
16	5	16	10	16	7	238	25
17	5	74	13	17	0	0	0
18	9	30	17	18	4	115	13
19	4	174	13	19	6	90	17
20	7	469	27	20	8	81	19
21	8	66	19	21	5	49	12
22	4	205	15	22	5	177	16
23	7	78	17	23	6	103	16
24	2	248	11	24	5	44	11
25	4	273	15	25	6	213	20
26	4	19	8	26	4	34	9
27	5	77	13	27	0	0	0
28	9	88	20	28	5	42	11
29	3	102	11	29	7	215	20
30	6	346	21	30	8	89	20
31	7	90	18	31	4	30	9
32	4	20	8	32	5	27	10
33	4	64	10	33	3	13	6
34	4	29	9	34	6	320	24
35	10	303	30	35	7	105	17
36	5	82	13	36	6	65	15
37	3	264	13	37	2	118	9
38	5	28	11	38	8	362	27
39	2	53	7	39	4	86	11
40	6	240	18	40	8	125	21
				41	8	139	22
				42	5	57	13
				43	4	73	12
				44	3	110	10

Se puede observar como para este caso, debido a la asignación aleatoria, la muestra de calibración a quedado conformada por 40 observaciones, mientras que la predicción a esta compuesta por 44.

Para el caso de la separación de las muestras de calibración y predicción relativas a las recogidas del vehículo del código postal 01895, la asignación de números aleatorios para cada una de las observaciones es la de la tabla 245.

Tabla 245: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08173

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,785	22	0,840	43	0,360	64	0,807
2	0,470	23	0,563	44	0,904	65	0,813
3	0,383	24	0,573	45	0,202	66	0,696
4	0,423	25	0,179	46	0,423	67	0,531
5	0,523	26	0,846	47	0,775	68	0,959
6	0,197	27	0,049	48	0,343	69	0,661
7	0,452	28	0,943	49	0,681	70	0,603
8	0,828	29	0,223	50	0,659	71	0,792
9	0,426	30	0,419	51	0,190	72	0,425
10	0,419	31	0,677	52	0,935	73	0,697
11	0,651	32	0,997	53	0,819	74	0,363
12	0,220	33	0,172	54	0,686	75	0,259
13	0,076	34	0,861	55	0,157	76	0,579
14	0,764	35	0,703	56	0,156	77	0,111
15	0,528	36	0,854	57	0,128	78	0,543
16	0,418	37	0,571	58	0,806	79	0,890
17	0,673	38	0,532	59	0,164	80	0,260
18	0,533	39	0,845	60	0,472	81	0,007
19	0,537	40	0,652	61	0,798	82	0,947
20	0,833	41	0,594	62	0,876	83	0,081
21	0,640	42	0,656	63	0,540	84	0,330

Y las correspondientes muestras de calibración y predicción son las que aparecen en la tabla 246

Tabla 246: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08173

Día	RECOGIDAS. MUESTRA DE CALIBRACIÓN			Día	RECOGIDAS. MUESTRA DE PREDICCIÓN		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	2	59	7	1	2	60	7
2	2	171	10	2	1	83	5
3	1	7	3	3	1	33	4
4	2	5	4	4	3	375	17
5	1	74	5	5	3	28	8
6	1	22	4	6	1	166	6
7	2	443	12	7	5	47	12
8	2	70	8	8	7	219	24
9	2	84	8	9	3	159	13
10	2	559	15	10	0	0	0
11	1	32	4	11	3	27	8
12	2	142	8	12	2	59	7
13	1	231	8	13	3	338	14
14	1	294	8	14	1	68	5
15	1	75	5	15	4	82	11
16	2	91	9	16	1	85	5
17	7	211	24	17	4	32	10
18	3	68	9	18	3	398	17
19	1	31	4	19	2	9	4
20	2	56	7	20	2	93	8
21	1	28	4	21	3	62	9
22	1	298	9	22	3	72	9
23	0	0	0	23	2	35	6
24	2	72	7	24	2	9	4
25	1	20	3	25	3	156	11
26	2	47	8	26	1	45	4
27	1	437	10	27	4	429	20
28	0	0	0	28	3	688	20
29	3	188	12	29	2	22	6
30	1	1	2	30	1	27	4
31	2	66	7	31	4	455	19
32	1	379	10	32	3	122	11
33	2	52	7	33	3	14	7
34	2	17	5				
35	1	11	3				
36	3	68	9				
37	2	171	9				
38	4	197	15				
39	1	50	4				
40	3	81	10				
41	1	41	4				
42	1	208	7				
43	1	221	7				
44	1	176	7				
45	1	34	4				
46	2	167	9				
47	2	42	6				
48	2	44	6				
49	5	162	19				
50	1	13	3				
51	2	13	5				

Teniendo para este caso, debido a la asignación aleatoria, la muestra de calibración conformada por 51 observaciones, mientras que la predicción a esta compuesta por 33.

2.6.4. Modelo de costes de reparto

Como para el resto de modelos de costes de reparto elaborados de este trabajo de investigación, los parámetros de regresión se calculan empleando el método de máxima verosimilitud. Se estudiará el valor de máxima verosimilitud obtenido bajo diferentes premisas, como diferentes distribuciones de probabilidad, ó transformaciones de las variables exógenas. Los parámetros de regresión del modelo que mayor valor de verosimilitud proporcionen, serán los que se empleen en la simulación del modelo para estimar los costes de reparto. Los supuestos que se van a analizar son los mismos que los modelos de reparto anteriores, y el estudio de máxima verosimilitud completo para cada uno de estos supuestos se pueden ver a continuación.

Tabla 247: VMV modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08173

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
15,095					-131,1		0%	6,49
2,388	2,353				-100,8	30,312	77%	3,08
0,065	2,106	0,028			-54,6	46,200	98%	0,98
1,197	1,721	0,028	0,028		-53,0	1,582	98%	0,96
-0,654	2,107	0,041		-0,00003	-49,5	3,473	98%	0,88

Tabla 248: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08173

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,714					-131,1		0%	6,49
2,027	0,119				-101,9	29,213	76%	3,17
1,924	0,102	0,001			-74,6	27,211	94%	1,63
1,637	0,184	0,001	-0,005		-65,9	8,735	96%	1,33
1,468	0,197	0,003	-0,006	-0,000005	-55,7	10,202	97%	1,04

Tabla 249: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08173

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,658					-167,8		-527%	16,26
0,097	-0,005				-107,5	60,274	68%	3,65
0,103	-0,004	-0,0001			-92,9	14,684	84%	2,56
0,136	-0,013	-0,0001	0,0005		-79,2	13,681	92%	1,85
0,136	-0,013	-0,0001	0,0005	0	-79,2	0,000	92%	1,85

Tabla 250: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08173

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-3,885					-131,1		0%	6,49
-2,350	-0,274				-100,3	30,724	78%	3,05
-2,100	-0,240	-0,003			-60,1	40,240	97%	1,13
-1,884	-0,307	-0,003	0,004		-57,7	2,367	97%	1,08
-1,592	-0,345	-0,006	0,007	0,00001	-49,3	9,154	98%	0,81

como el valor de máxima verosimilitud es mayor que en el supuesto de una función cadena para la misma distribución de probabilidad de los errores. El valor es un poco menor que en el caso de que los errores se distribuyen según una distribución logística, por lo que en la simulación se empleará este modelo, pero se puede comprobar como una transformación de las variables exógenas proporciona un mejor ajuste a los datos que el mismo modelo sin esa transformación.

2.6.4.1. Selección de variables para el modelo de costes de reparto

En el cuadro del estudio de máxima verosimilitud para los errores distribuidos de acuerdo a una distribución de probabilidad logística, se puede comprobar cómo al introducir una a una las cuatro variables en estudio (número de repartos por día, peso total repartido por día, número de repartos por día elevado a la segunda potencia, y peso total repartido por día elevado a la segunda potencia) el cociente de verosimilitud es superior a dos unidades en cada uno de los casos. Por tanto, se incluyen las cuatro variables en el modelo de costes de reparto, con sus correspondientes parámetros de regresión: 1,847, 0,039, 0,021 y -0,00002. El valor umbral para este modelo es 0,16.

2.6.4.2. Adecuación del modelo de costes de reparto

Cuando los errores se distribuyen con una distribución de probabilidad no normal, si éstos son independientes y se distribuyen idénticamente, las inferencias bajo la teoría de normalidad son asintóticamente correctas. Por tanto, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. Este análisis gráfico se puede ver en la figura 112.

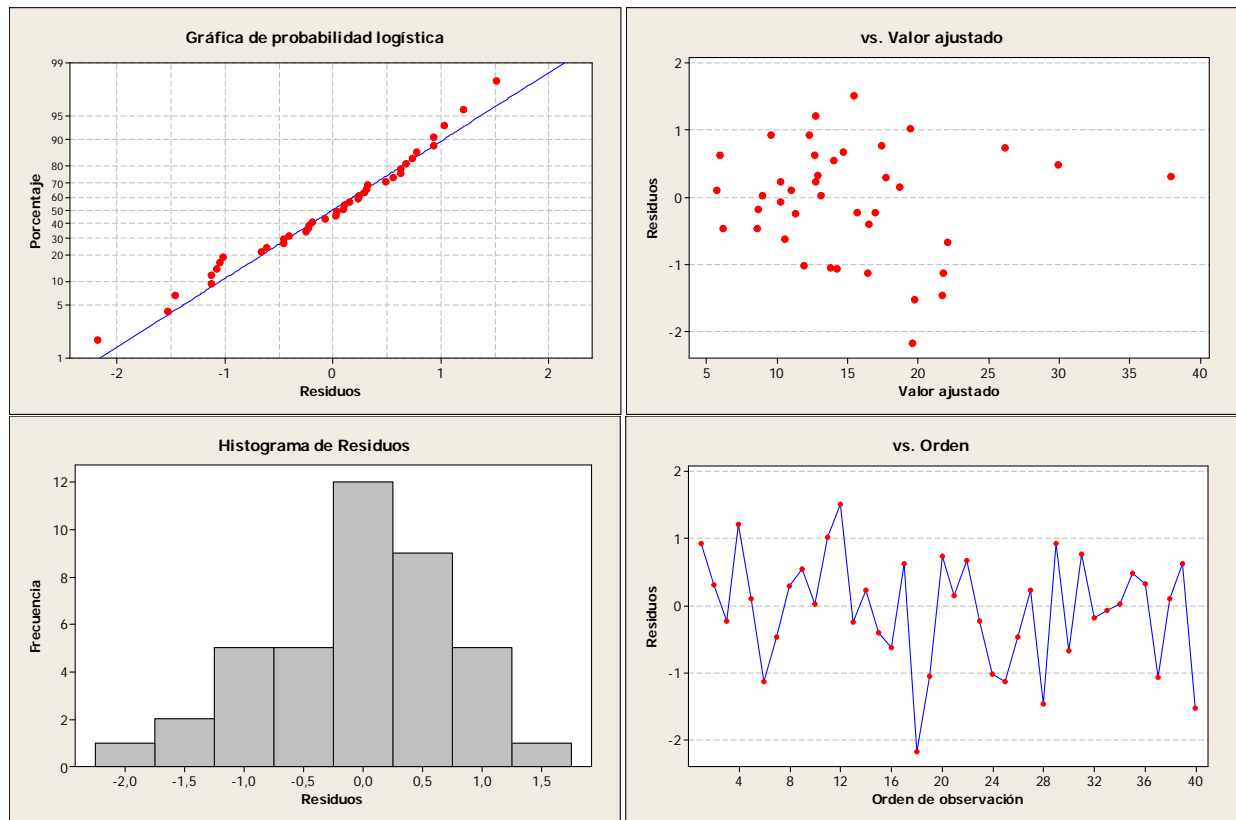


Figura 112: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08173

En la primera gráfica se puede observar como los errores se distribuyen de manera normal ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.6.4.3. Validación del modelo de costes de reparto

Para comprobar que el modelo predice de manera adecuada los valores de los costes de reparto ante nuevos valores de las variables exógenas, se realiza la prueba *t* pareada o prueba de Welch.

Los nuevos valores de las variables exógenas son los valores de la muestra de predicción, y la validación del modelo se lleva a cabo comparando uno a uno los valores de la variable endógena estimados por el modelo tras introducir los nuevos valores de las variables exógenas, con los valores correspondientes de los costes de reparto de la muestra de predicción.

La tabla que resume los resultados de esta prueba de Welch es la tabla 256.

Tabla 256: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08173

	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	14,07	14,19	-0,119	Límite superior I.C	$t_{0,025;40}$
Varianza	40,05	39,73	0,86	0,155	1,96
Varianza/n			0,020		
Máximo	27,88	26,66		Límite inferior I.C	
Mínimo	0,16	0,00		-0,393	

El intervalo de confianza obtenido para esta prueba contiene el valor cero, por lo que se concluye que el modelo estima nuevos costes de reparto de manera adecuada. Para elaborarlo se ha empleado la distribución t de Student, con un nivel de significancia del 95% y 40 grados de libertad, cuyo valor es 1.96.

Los valores Z que se emplean para realizar en esta prueba de Welch se obtienen de la tabla 257.

Tabla 257: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08173

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
0,160	1,846	0,0393	0,020	-0,00002			
	7	442	49	195364	27,58	25,50	2,08
	2	262	4	68644	12,86	10,62	2,25
	3	35	9	1225	7,24	8,04	-0,81
	6	162	36	26244	17,84	18,03	-0,20
	4	57	16	3249	10,06	10,95	-0,89
	7	251	49	63001	22,72	23,87	-1,16
	4	146	16	21316	13,19	13,44	-0,24
	4	64	16	4096	10,32	10,92	-0,61
	6	241	36	58081	20,30	21,02	-0,72
	3	62	9	3844	8,25	8,63	-0,38
	5	154	25	23716	15,49	15,39	0,10
	4	109	16	11881	11,93	11,71	0,22
	3	136	9	18496	10,86	10,96	-0,10
	4	210	16	44100	15,25	13,39	1,86
	2	30	4	900	5,10	5,41	-0,31
	7	238	49	56644	22,33	24,52	-2,19
	0	0	0	0	0,16	0,00	0,16
	4	115	16	13225	12,14	13,44	-1,31
	6	90	36	8100	15,37	17,15	-1,78
	8	81	64	6561	19,32	19,00	0,32
	5	49	25	2401	11,79	12,17	-0,38
	5	177	25	31329	16,25	15,57	0,67
	6	103	36	10609	15,83	16,39	-0,57
	5	44	25	1936	11,61	11,46	0,15
	6	213	36	45369	19,46	19,53	-0,07
	4	34	16	1156	9,19	9,09	0,10
	0	0	0	0	0,16	0,00	0,16
	5	42	25	1764	11,53	11,50	0,04
	7	215	49	46225	21,64	20,02	1,62
	8	89	64	7921	19,61	20,41	-0,80
	4	30	16	900	9,04	9,34	-0,30
	5	27	25	729	10,96	10,47	0,50
	3	13	9	169	6,40	6,18	0,22
	6	320	36	102400	22,52	23,98	-1,45
	7	105	49	11025	18,02	17,33	0,68
	6	65	36	4225	14,46	14,66	-0,20
	2	118	4	13924	8,30	8,59	-0,30
	8	362	64	131044	27,88	26,66	1,22
	4	86	16	7396	11,11	11,28	-0,16
	8	125	64	15625	20,87	21,31	-0,44
	8	139	64	19377	21,35	21,56	-0,21
	5	57	25	3249	12,09	12,83	-0,73
	4	73	16	5329	10,64	11,75	-1,10
	3	110	9	12100	9,97	10,17	-0,20

2.6.5. Modelo de costes de recogida

Una vez finalizado el modelo de los costes de reparto, hay que realizar el modelo de costes de recogida. La creación este modelo es análogo al de los costes de reparto, por lo que los pasos a seguir son idénticos.

Tabla 264: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08173

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,146					-143,6		-1%	4,43
0,174	-0,019				-128,8	14,858	55%	2,95
0,190	-0,016	0,000			-109,8	18,952	74%	2,23
0,190	-0,016	0,000	0		-109,8	0,000	74%	2,25
0,190	-0,016	0,000		0	-109,8	0,000	74%	2,25

Tabla 265: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08173

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-2,615					-143,6		-1%	4,43
-1,506	-0,565				-118,2	25,455	64%	2,65
-1,487	-0,437	-0,003			-72,6	45,515	94%	1,05
-1,317	-0,608	-0,003	0,026		-66,7	5,961	95%	0,97
-1,308	-0,580	-0,004	0,026	0,000003	-54,9	11,823	97%	0,79

Tabla 266: VMV modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08173

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
6,904					-144,0		-1%	4,43
1,294	3,267				-115,2	28,784	70%	2,42
0,202	2,879	0,018			-55,5	59,710	97%	0,75
0,513	2,564	0,018	0,049		-54,4	1,145	97%	0,74
-0,017	2,783	0,025		-0,00002	-49,2	5,202	98%	0,66

Se puede observar como los valores de máxima verosimilitud para los supuestos de que los errores se distribuyen según una distribución de probabilidad normal. Es por ello que se empleará como modelo de costes de reparto el análisis de regresión bajo el supuesto de que los errores se distribuyen bajo una distribución normal.

2.6.5.1. Selección de variables para el modelo de costes de recogida

Si se observa la tabla del apartado anterior correspondiente al supuesto de distribución logística de los errores, se puede observar como al añadir la variable número de repartos por día, el cociente de verosimilitud es superior a dos unidades, por lo que adición de esta variable aporta un mejor ajuste al modelo con su inclusión. Con la variable peso total repartido por día ocurre exactamente lo mismo, ya que al introducirla en el modelo el cociente de verosimilitud aumenta en 60 unidades. Sin embargo, al introducir la variable número de repartos por día elevada a la segunda potencia, el cociente de verosimilitud es de tan sólo 1.238, por lo que la adición de esta variable al modelo no aporta un mejor ajuste. Por tanto, como esta variable no aporta un mejor ajuste, al calcular el valor de máxima verosimilitud con las dos primeras variables más la variable peso total repartido elevada a la segunda potencia, no se debe incluir la variable número de repartos por día elevada al cuadrado. Al obtener el valor de máxima verosimilitud se observa como el cociente de verosimilitud es mayor de dos unidades,

por lo que el modelo al incluir esta variable aporta un mejor ajuste a los datos que con únicamente las dos primeras variables.

En definitiva, este modelo de costes de reparto se compone de tres variables, número de repartos por día, con un parámetro de regresión igual 2,773, peso total repartido por día, con su correspondiente parámetro de regresión igual a 0,026, y número de recogidas por día elevada a la segunda potencia, con un parámetro igual a -0,00002. El valor umbral para este modelo es igual a -0,107.

2.6.5.2. Adecuación del modelo de costes de recogida

Como el mayor valor máximo de verosimilitud es el que aporta una distribución de probabilidad normal, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. El análisis gráfico se puede observar en la figura 113.

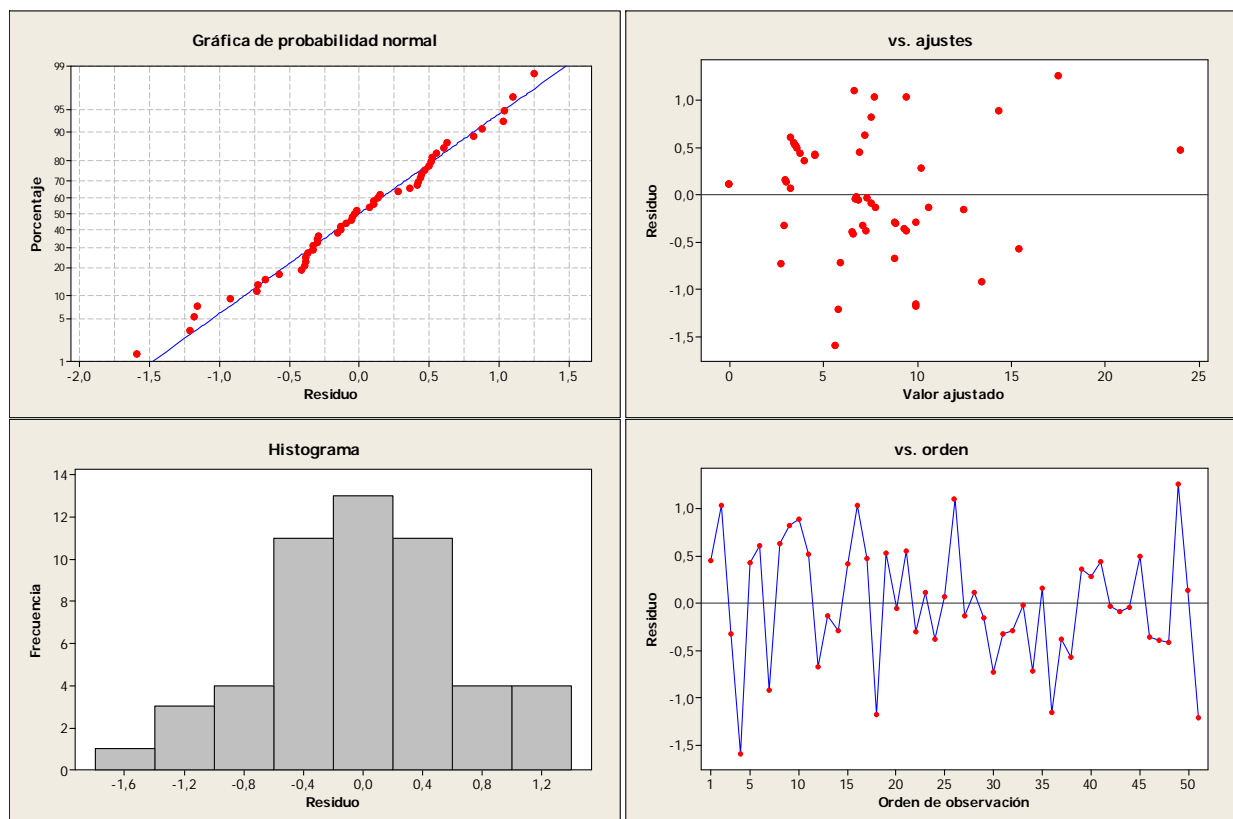


Figura 113: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08173

En la primera gráfica se puede observar como los errores se distribuyen de manera normal ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la

varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.6.5.3. Validación del modelo de costes de recogida

Para comprobar que el modelo predice de manera adecuada los valores de los costes de reparto ante nuevos valores de las variables exógenas, se realiza la prueba *t* pareada o prueba de Welch, con los datos de la muestra de predicción previamente obtenida. El cuadro que resume esta prueba de Welch es la tabla 267.

Tabla 267: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08173

	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	9,77	9,50	0,270	Límite superior I.C	$t_{0,025;30}$
Varianza	30,23	32,10	1,40	0,691	2,04
Varianza/n			0,043		
Máximo	24,14	23,70		Límite inferior I.C	
Mínimo	-0,11	0,00		-0,150	

En él se puede observar como el modelo estima los costes de recogida de manera adecuada, ya que el intervalo de confianza creado incluye el cero. Para elaborarlo se ha empleado la distribución *t* de Student, con un nivel de significancia del 95% y 30 grados de libertad, cuyo valor es 2,04. Los valores Z con los que se elabora este cuadro provienen de la tabla 268.

Tabla 268: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08173

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_4$	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
-0,107	2,773	0,026	-0,00002			
	2	60	3600	6,94	6,66	0,28
	1	83	6889	4,70	5,06	-0,36
	1	33	1089	3,51	4,01	-0,51
	3	375	140625	15,42	17,46	-2,04
	3	28	784	8,93	7,83	1,09
	1	166	27556	6,49	6,50	0,00
	5	47	2209	14,94	12,03	2,92
	7	219	47961	24,14	23,70	0,44
	3	159	25281	11,90	12,86	-0,96
	0	0	0	-0,11	0,00	-0,11
	3	27	729	8,90	7,81	1,09
	2	59	3481	6,91	6,65	0,27
	3	338	114244	14,94	13,97	0,96
	1	68	4624	4,35	4,83	-0,47
	4	82	6724	13,00	10,91	2,09
	1	85	7225	4,75	5,10	-0,35
	4	32	1024	11,80	10,20	1,60
	3	398	158404	15,69	17,01	-1,32
	2	9	81	5,67	4,05	1,62
	2	93	8649	7,71	8,16	-0,45
	3	62	3844	9,76	8,84	0,92
	3	72	5184	9,99	8,79	1,20
	2	35	1225	6,33	6,42	-0,09
	2	9	81	5,67	4,50	1,18
	3	156	24336	11,83	10,71	1,12
	1	45	2025	3,80	4,21	-0,41
	4	429	184041	18,81	19,81	-1,01
	3	688	473344	17,50	20,07	-2,57
	2	22	484	6,00	5,69	0,32
	1	27	729	3,36	3,92	-0,56
	4	455	207025	19,06	18,63	0,44
	3	122	14884	11,12	10,56	0,56
	3	14	196	8,57	6,53	2,04

2.6.6. Modelo de carga de trabajo

En este modelo se va a identificar que distribución de probabilidad siguen las variables aleatorias con las que se puede modelar la carga la de trabajo, así como los parámetros que las caracterizan. Estas variables son el número de repartos por día y el peso de cada uno de los repartos, en el caso de los repartos, y el número de recogidas por día y el peso de cada recogida, en el de las recogidas.

Las muestras que se emplean para las variables número de repartos, peso total repartido, número de recogidas y peso total recogido son las mismas muestras que las muestras de calibración que se emplearon para ajustar los modelos de los costes de reparto y recogida. Las muestras para las variables peso de cada reparto y peso de cada recogida, es una muestra con los pesos unitarios de cada uno de los repartos y de cada una de las recogidas.

2.6.6.1. Número de repartos por día

Minitab busca los parámetros para cada una de los distintos tipos de distribuciones de probabilidad, siendo estos parámetros tales que la distribución de probabilidad que caracterizan, es la que más se aproxima a los datos de la muestra. Una vez que se tienen estos datos hay que comprobar que la distribución de probabilidad que conforma se ajuste suficientemente bien a los datos. Para realizar esta comprobación se realizan los denominados test no paramétricos de bondad de ajuste. Debido a la naturaleza de esta variable aleatoria, la cual es un conteo de algún elemento, la primera función de distribución de probabilidad que se va a tratar de comprobar el ajuste es la distribución de Poisson. Como esta distribución es discreta se emplea el test no paramétrico de bondad de ajuste de Chi cuadrado. El resultado de este test se puede ver en la tabla 269.

Tabla 269: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08173

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
2-3	7	0,175	0,2133	9	0,2750	alfa=0,05 k=5 Par. Est. =1 n=5-1-1=3
4	9	0,225	0,1600	6	1,0555	
5	10	0,250	0,1728	7	1,3787	
6-7	8	0,200	0,2755	11	0,8281	
8-14	6	0,150	0,1783	7	0,1802	
	40				3,72	7,85

En la tabla se puede observar como el valor límite para aceptar la hipótesis de que la una distribución de Poisson con lambda 5,45 ajusta adecuadamente la variable número de repartos por día es 7,85. Este valor corresponde a una distribución Chi cuadrado con 3 grados de libertad y nivel de significancia del 5%. Los grados de libertad es igual al número de intervalos de los que consta el test, al que hay que restar el número de parámetros estimados (para el caso de la distribución de Poisson es únicamente el parámetro lambda) y otra unidad. Como en este caso el valor del test es 3,72, menor al valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.6.6.2. Peso de cada reparto

La segunda variable a analizar para el modelo de carga de trabajo es el peso de cada reparto. Esta variable es una variable aleatoria continua, y al igual que para el resto de variables continuas, las distribuciones de probabilidad a analizar serán también continuas. Los valores dados por Minitab y los

resultados del test Kolmogorov-Smirnov para cada una de las distribuciones analizadas se pueden ver en la tabla 270.

Tabla 270: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08173

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,296	alfa	0,147	alfa	0,127	Media	0,085	Lambda	0,293
23,92	4,388	0,5896	2,162	0,6859	1,865	2,124	1,249	14,99	4,320
Desv. Est.	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	Desv. Est.	$c_{\alpha} = 1,358$	Theta	$c_{\alpha} = 1,358$
43,87	$\alpha = 0,05$	40,57	$\alpha = 0,05$	17,47	$\alpha = 0,05$	1,433	$\alpha = 0,05$	16,29	$\alpha = 0,05$

En el cuadro se puede ver cómo la única distribución que supera satisfactoriamente el test de Kolmogorov-Smirnov es la distribución lognormal. Por lo tanto, para la simulación se empleará esta distribución para estimar los pesos de cada uno de los repartos.

2.6.6.3. Número de recogidas por día

Al igual que para la variable número de repartos por día, esta variable es un conteo de algún elemento. Por ello, la primera distribución de probabilidad que se va a comprobar si ajusta adecuadamente a los valores de la muestra de esta variable es la distribución de probabilidad discreta de Poisson. Al ser una distribución discreta, el test no paramétrico a realizar es el test Chi cuadrado. Los resultados de este test se puede ver en la tabla 271.

Tabla 271: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08173

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
0-1	32	0,381	0,3848	32	0,0032	alfa=0,05
2	27	0,321	0,2702	23	0,8142	k=4
3	16	0,190	0,1874	16	0,0043	Est. par. =1
4-7	9	0,107	0,1576	13	1,3567	n=4-1-1=2
	84				2,18	5,991

El valor lambda dado por Minitab para una distribución de Poisson con los datos de la muestra de esta variable aleatoria es 2,08. En la tabla se puede observar como el valor límite para aceptar la hipótesis de que una distribución de Poisson con lambda 2,08 ajusta adecuadamente la variable número de repartos por día es 5,991. Este valor corresponde a una distribución Chi cuadrado con 2 grados de libertad y nivel de significancia del 5%. Los grados de libertad es igual al número de intervalos de los que consta el test, al que hay que restar el número de parámetros estimados (para el caso de la distribución de Poisson es únicamente el parámetro lambda) y otra unidad. Como en este caso el valor del test es 2,18, es menor que el valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de recogidas por día.

2.6.6.4. Peso de cada recogida

Por último, para finalizar con la elaboración del modelo de carga de trabajo para este código postal se va a analizar la variable peso de cada recogida. Esta variable es una variable continua, y las distribuciones a analizar son también continuas. Los parámetros de las distintas distribuciones analizadas, así como los resultados de los test de Kolmogorov-Smirnov de cada una se pueden ver en la tabla 272.

Tabla 272: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08173

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,284	alfa	0,115	alfa	0,082	Media	0,059	Lambda	0,261
67,18	2,705	0,5889	1,090	0,6915	0,778	3,155	0,559	44,47	2,479
Desv. Est.	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	beta	$c_{\alpha} = 1,358$	Desv. Est.	$c_{\alpha} = 1,358$	Theta	$c_{\alpha} = 1,358$
105,8	$\alpha = 0,05$	114,1	$\alpha = 0,05$	51	$\alpha = 0,05$	1,586	$\alpha = 0,05$	44,12	$\alpha = 0,05$

En la tabla 272 se puede ver como las distribuciones de probabilidad gamma, Weibull y lognormal superan satisfactoriamente el test de Kolmogorov-Smirnov. Se podría emplear cualquiera de las tres para la simulación de la variable peso de cada recogida, pero como la que menor discrepancia presenta es la distribución lognormal, será ésta la que se utilice.

2.6.7. *Simulación de los costes de reparto y recogida, y de la carga de trabajo*

En este apartado se llevará a cabo la simulación de los modelos creados para el estudio del vehículo de 1500 kg. de peso máximo autorizado en el código postal 08198. En primer lugar se simulará la carga de trabajo. Para ello hay que determinar el número de repartos y de recogidas para cada uno de los días simulados, y el peso total repartido y el peso total recogido por día. Empleando los ajustes de las distribuciones de probabilidad de Poisson obtenidos en los apartados 2.6.6.1 y 2.6.6.3. se simularán el número de repartos y el número de recogidas para cada uno días que se desean simular.

Posteriormente, se simulará el peso correspondiente de cada uno de los repartos y las recogidas mediante las distribuciones de probabilidad calculadas en los apartados 2.6.6.2. y 2.6.6.4. Sumando el peso unitario de todos los repartos simulados en cada uno de los días, se calculará el peso total repartido por día. Con el peso unitario de las recogidas simulado para cada uno de los días, se calculará el peso total recogido por día.

Finalmente, se simularán los modelos de costes de reparto y costes de recogida. Como valores de entrada a estos modelos se emplearán los datos simulados para el modelo de carga de trabajo. El número de repartos y el peso total repartido por día serán los datos de entrada para el modelo de costes de reparto.

Análogamente, el número de recogidas y el peso total recogido por día serán los valores de entrada para el modelo de costes de recogida.

2.6.7.1. Simulación de los costes de reparto

La carga de trabajo relativa a los repartos para este código postal es la siguiente: el número de repartos puede ser descrito mediante una distribución de probabilidad Poisson con parámetro lambda igual a 5,45. Mientras que el peso de cada reparto sigue una distribución de probabilidad lognormal con media 2,124 y desviación estándar 1,433.

Un resumen de la carga de trabajo simulada es el siguiente, en el cual se puede observar el promedio de repartos por día simulados, su desviación estándar, el número máximo de repartos a realizar en un día, así como el número mínimo. Y los mismos datos para la variable peso total repartido, siendo este peso total repartido la suma de los pesos simulados para cada una de los repartos.

Tabla 273: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08173

	Repartos por día	Peso total repartido por día
Promedio	5	106
Desv. Estan.	2	81
Máximo	11	443
Mínimo	1	1

Esta simulación de la carga de trabajo se emplea como entrada para el modelo de costes de reparto. Como se han simulado 250 días de carga de trabajo, se obtendrán los estimados puntuales de costes de reparto para 250 días. Los resultados que aquí se exponen son el valor medio de límite inferior de las 250 simulaciones, la varianza y desviación estándar de este límite inferior, así como un intervalo de confianza de este límite inferior. Para el límite superior del intervalo de predicción los resultados que se presentan son análogos. Estos resultados son los siguientes:

Tabla 274: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08173

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	12,73	16,30
Varianza	34	34
Desv. Estan.	5,84	5,84
I.C	12,12	15,69
	13,34	16,90

En la tabla se puede observar como el límite inferior de los estimados puntuales de los costes de repartos se distribuyen con un valor medio de 12,73 euros por día y una desviación estándar de 5,84. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 16,30 y una desviación estándar de 5,84 euros por día.

2.6.7.2. Simulación de los costes de recogida

La carga de trabajo relativa a los repartos para este código postal es la siguiente: el número de repartos puede ser descrito mediante una distribución de probabilidad Poisson con parámetro lambda igual a 2,18, y el peso de cada una de las recogidas mediante una distribución de probabilidad lognormal con media igual a 3,155 y desviación estándar 1,586.

Un resumen de la carga de trabajo simulada es el siguiente, en el cual se puede observar el promedio de repartos por día simulados, su desviación estándar, el número máximo de repartos a realizar en un día, así como el número mínimo. Y los mismos datos para la variable peso total repartido.

Tabla 275: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08173

	Recogidas por día	Peso total recogido por día
Promedio	2	117
Desv. Estan.	2	140
Máximo	7	720
Mínimo	0	0

Y la simulación de los costes de recogida correspondiente a esta carga de trabajo para 250 días se resume en el siguiente cuadro.

Tabla 276: Resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08173

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	7,12	9,84
Varianza	37	38
Desv. Estan.	6,05	6,18
I.C	6,49	9,20
	7,76	10,49

Donde se puede comprobar como el límite inferior de los estimados puntuales de los costes de recogida se distribuyen con un valor medio de 7,12 euros por día y una desviación estándar de 6,05.

Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 9,84 y una desviación estándar de 6,18 euros por día.

2.7. Código postal 08172

En este cuarto código postal analizado, no se dan repartos superiores a 300 kg. de peso, ni recogidas superiores a 700 kg., por lo que en él solo existe volumen de trabajo para vehículos de 1500 kg. de peso máximo autorizado.

Los pasos a seguir en la elaboración del modelo de costes de la red de distribución capilar de este código postal son análogos a los empleados en los códigos postales anteriores. Tal y como se llevó a cabo en los tres tipos de vehículos del código postal anterior, el primer paso en la elaboración del modelo es la toma de muestras.

2.7.1. Toma de muestras

Se toman muestras de aquellas variables que se identificaron como las determinantes para el buen funcionamiento de la red de distribución capilar. Estas son el coste de reparto por día y el coste de recogida por día. Éstas a su vez dependen del número de repartos por día y del peso repartido por día, y del número de recogidas por día y del peso total repartido por día respectivamente. Esta toma de muestras se ha llevado a cabo mediante un estudio observacional, y se han obtenido 84 observaciones para cada una de ellas. Las muestras obtenidas para las variables relacionadas con los repartos se pueden observar en la tabla 277, mientras que en la tabla 278 se pueden ver las relacionadas con las recogidas.

Tabla 277: Toma de muestras de los repartos del vehículo de 1500 kg. PMA para el código postal 08172

Día	REPARTOS			Día	REPARTOS		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	11	310	34	43	8	254	25
2	12	281	33	44	4	60	11
3	14	270	38	45	9	513	35
4	9	82	21	46	8	167	24
5	5	71	12	47	6	287	20
6	3	151	12	48	7	138	22
7	5	72	12	49	14	348	43
8	5	16	10	50	7	240	22
9	9	165	24	51	13	167	34
10	4	32	9	52	7	141	22
11	11	158	29	53	8	143	23
12	12	297	35	54	6	89	18
13	14	195	37	55	7	178	21
14	11	323	36	56	8	535	33
15	9	363	30	57	10	140	28
16	2	58	6	58	10	143	28
17	6	22	12	59	11	488	39
18	13	278	36	60	9	275	32
19	6	78	16	61	12	122	32
20	13	426	39	62	5	27	12
21	10	586	36	63	7	296	25
22	6	775	34	64	9	132	26
23	3	114	11	65	7	383	26
24	6	170	17	66	14	782	49
25	11	94	24	67	8	331	24
26	6	108	17	68	8	166	23
27	4	137	12	69	5	155	17
28	7	134	20	70	9	170	26
29	8	136	23	71	8	78	20
30	8	300	25	72	8	142	21
31	13	290	37	73	7	172	21
32	9	276	28	74	3	29	8
33	6	334	22	75	7	75	19
34	8	322	27	76	5	135	17
35	9	116	23	77	8	70	20
36	3	71	8	78	4	33	11
37	11	503	40	79	7	915	37
38	7	125	19	80	11	199	33
39	5	81	13	81	9	273	30
40	11	137	28	82	8	180	26
41	8	68	19	83	2	7	4
42	8	114	20	84	1	5	2

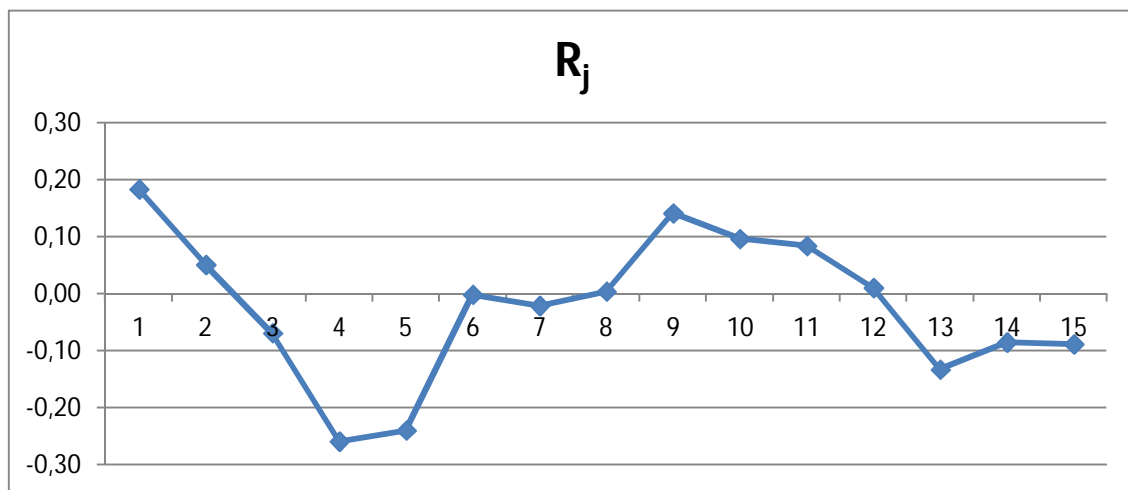
Tabla 278: Toma de muestras de las recogidas del vehículo de 1500 kg. PMA para el código postal 08172

Día	RECOGIDAS			Día	RECOGIDAS		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	3	276	14	43	0	0	0
2	5	606	25	44	2	719	19
3	2	344	13	45	3	794	24
4	4	147	15	46	5	652	27
5	5	159	19	47	6	448	24
6	2	46	7	48	3	265	15
7	1	22	4	49	4	477	21
8	3	428	16	50	4	417	20
9	5	411	24	51	3	404	18
10	3	496	17	52	2	461	15
11	4	854	27	53	5	231	19
12	3	545	19	54	5	98	15
13	8	552	33	55	5	796	30
14	3	286	16	56	3	465	18
15	4	344	20	57	4	294	18
16	2	223	11	58	4	299	18
17	6	178	21	59	1	86	5
18	3	422	16	60	3	40	9
19	6	453	26	61	2	261	12
20	2	60	7	62	3	66	9
21	3	374	15	63	1	437	10
22	4	295	18	64	3	73	10
23	3	378	16	65	3	89	11
24	6	586	30	66	2	605	17
25	4	161	15	67	1	36	4
26	5	158	17	68	2	779	19
27	2	303	11	69	1	40	4
28	4	644	24	70	3	257	12
29	4	267	17	71	2	105	9
30	5	921	32	72	2	186	9
31	6	881	34	73	4	140	14
32	4	556	21	74	2	128	8
33	4	169	15	75	1	212	7
34	4	386	21	76	2	39	7
35	5	610	27	77	1	226	8
36	4	232	17	78	1	10	3
37	4	304	19	79	3	197	14
38	7	1.520	46	80	0	0	0
39	4	778	25	81	3	49	8
40	5	804	28	82	1	41	4
41	5	76	14	83	3	91	12
42	6	161	21	84	2	198	11

2.7.2. Independencia de la muestras

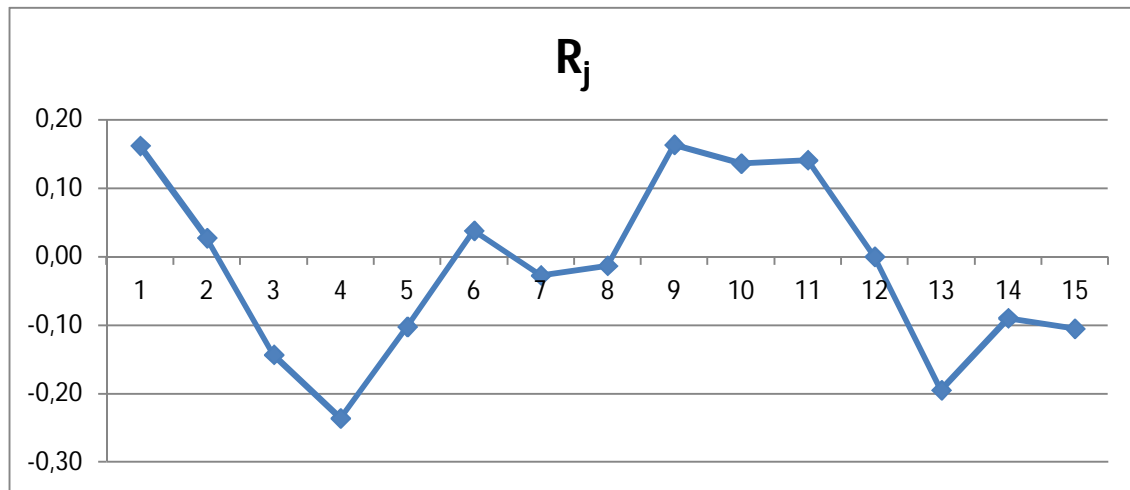
Tras la obtención de las muestras, hay que comprobar si las observaciones de de las mismas son independientes. Que las observaciones sean independientes es fundamental, ya que para emplear el método de máxima verosimilitud en la elaboración de los modelos de costes de reparto y recogida, las muestras utilizadas han de ser independientes. Por otro lado, los test no paramétricos de bondad de ajuste de las distribuciones de probabilidad también presuponen que las muestras empleadas son independientes.

En primer lugar se analiza la independencia de la muestras de los repartos. Como esta muestra se compone de tres variables, coste de reparto por día, número de repartos por día y peso total repartido por día, es necesario comprobar la independencia de cada una de ellas. En primer lugar se muestran las gráficas de autocorrelación.



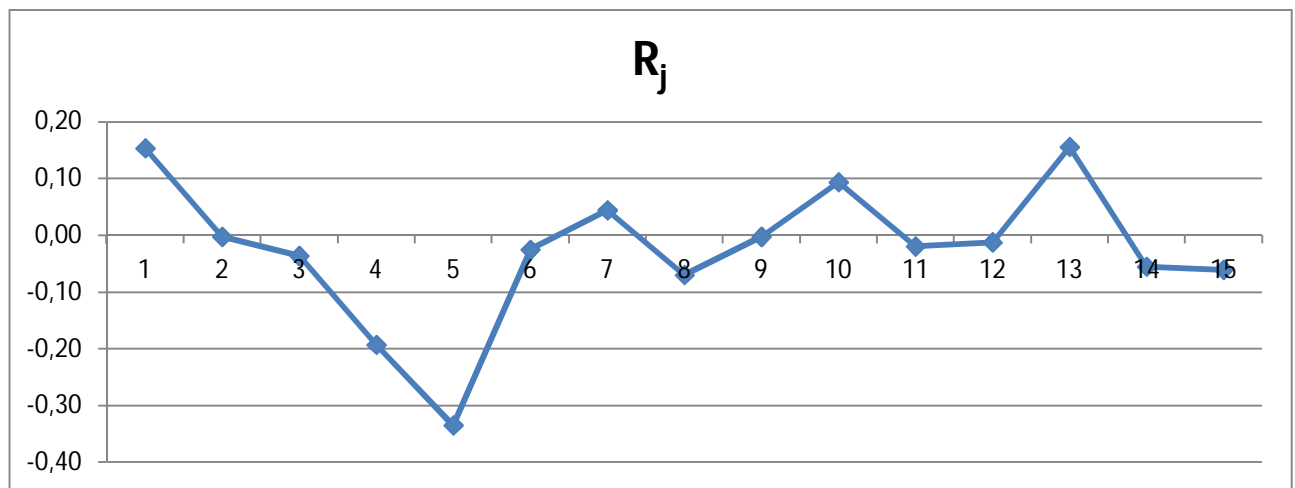
R_j	0,18	0,05	-0,07	-0,26	-0,24	0,00	-0,02	0,00	0,14	0,10	0,08	0,01	-0,13	-0,08	-0,09
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	18	5	-7	-25	-23	0	-2	0	14	9	8	1	-13	-8	-9

Figura 114: Grafica de autocorrelación de coste de reparto por día vehículo 1500 kg. PMA código 08172



R_j	0,16	0,03	-0,14	-0,24	-0,10	0,04	-0,03	-0,01	0,16	0,14	0,14	0,00	-0,19	-0,09	-0,11
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	1,5	0,3	-1,3	-2,2	-1,0	0,4	-0,3	-0,1	1,5	1,3	1,3	0,0	-1,8	-0,8	-1,0

Figura 115: Grafica de autocorrelación de número de repartos por día vehículo 1500 kg. PMA código 08172



R_j	0,15	0,00	-0,04	-0,19	-0,33	-0,03	0,04	-0,07	0,00	0,09	-0,02	-0,01	0,16	-0,06	-0,06
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	4767	-81	-1124	-5989	-10388	-786	1371	-2169	-78	2910	-609	-386	4832	-1717	-1873

Figura 116: Grafica de autocorrelación del peso total repartido por día vehículo 1500 kg. PMA código 08172

En las tres gráficas se puede observar como en ninguna de ellas los valores de R_j se encuentran en el rango de -1 y 1, y que en ninguna de ellas se puede apreciar un patrón determinado. Por tanto, esta prueba muestra la independencia de las observaciones de esta muestra.

Tras las gráficas de autocorrelación, se muestran a continuación las gráficas de dispersión para cada una de las variables.

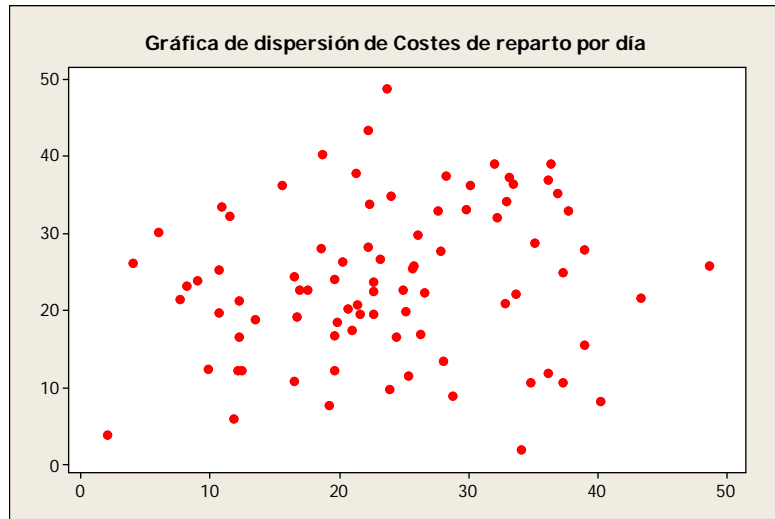


Figura 117: Grafica de dispersión de coste de reparto por día vehículo 1500 kg. PMA código 08172

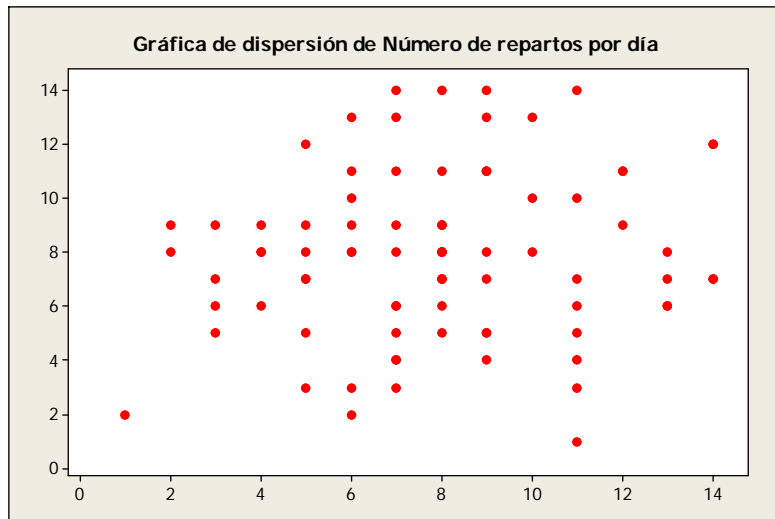


Figura 118: Grafica de dispersión de número de repartos por día vehículo 1500 kg. PMA código 08172

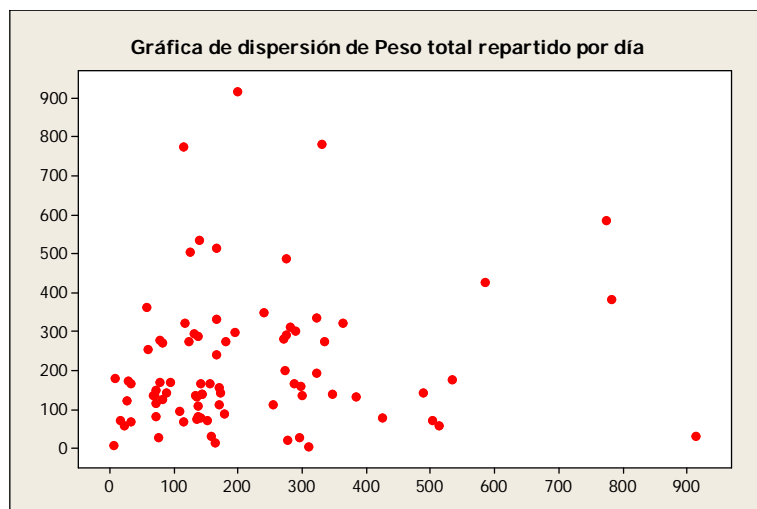
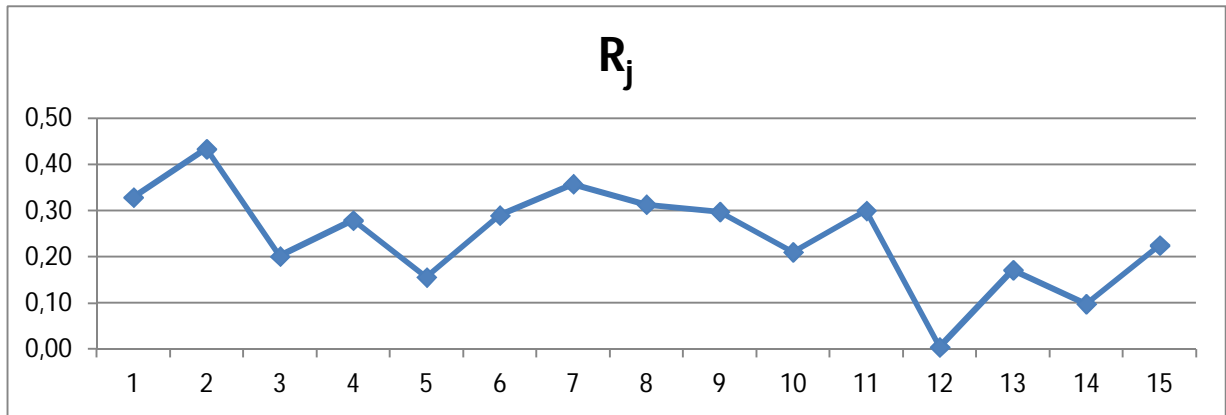


Figura 119: Grafica de dispersión del peso total repartido por día vehículo 1500 kg. PMA código 08172

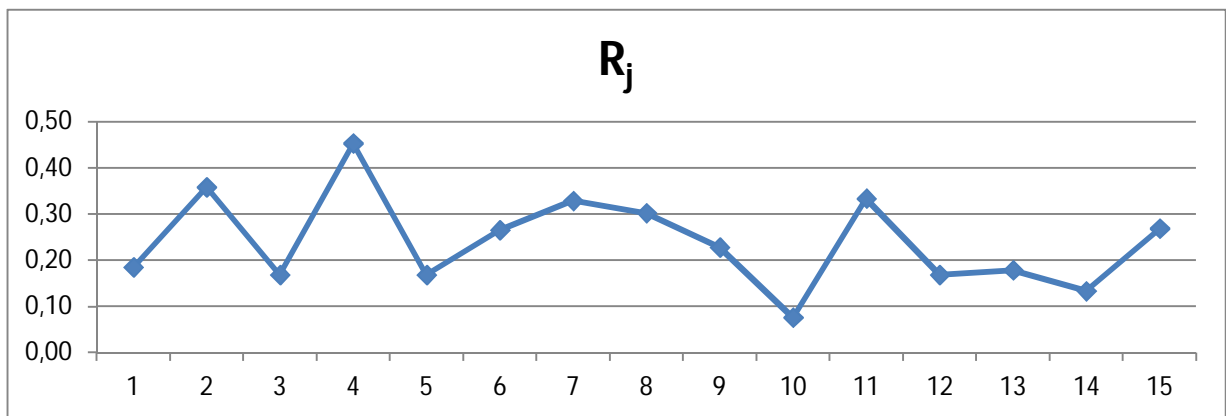
Como en las tres gráficas los puntos se distribuyen de manera aleatoria a lo largo de todo el plano, esta prueba también concluye que las observaciones son independientes.

Tras comprobar que las muestras relativas a los repartos son independientes, hay que analizar si las muestras correspondientes a las recogidas también lo son. Al igual que para los repartos se comenzará mostrando las gráficas de autocorrelación.



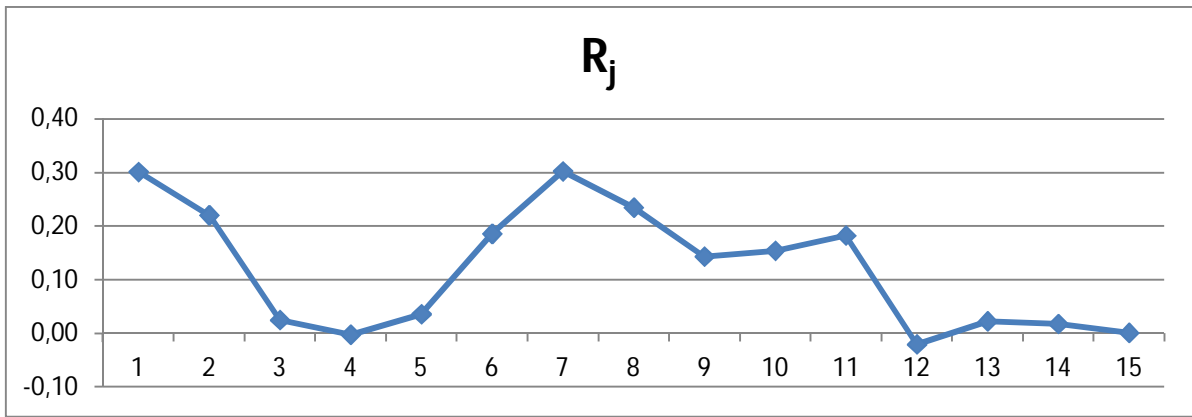
R_j	0,33	0,43	0,20	0,28	0,16	0,29	0,36	0,31	0,30	0,21	0,30	0,00	0,17	0,10	0,23
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	23	30	14	19	11	20	25	22	21	15	21	0	12	7	16

Figura 120: Grafica de autocorrelación de coste de recogida por día vehículo 1500 kg. PMA código 08172



R_j	0,18	0,36	0,17	0,45	0,17	0,27	0,33	0,30	0,23	0,08	0,33	0,17	0,18	0,13	0,27
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	0,5	0,9	0,4	1,2	0,4	0,7	0,8	0,8	0,6	0,2	0,9	0,4	0,5	0,3	0,7

Figura 121: Grafica de autocorrelación de número de recogidas por día vehículo 1500 kg. PMA código 08172



R_j	0,30	0,22	0,02	0,00	0,04	0,19	0,30	0,23	0,14	0,15	0,18	-0,02	0,02	0,02	0,00
j	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
C_j	22934	16783	1899	-169	2736	14155	23030	17864	10912	11756	13898	-1547	1714	1348	89

Figura 122: Grafica de autocorrelación del peso total recogido por día vehículo 1500 kg. PMA código 08172

En las tres gráficas se puede ver como los valores de R_j no superan los valores 1 y -1, y que no se da un patrón determinado en ninguna de ellas por lo que se concluye que las observaciones de cada muestra son independientes. Tras estudiar las gráficas de autocorrelación, a continuación se verifica la independencia de las observaciones mediante las gráficas de dispersión.

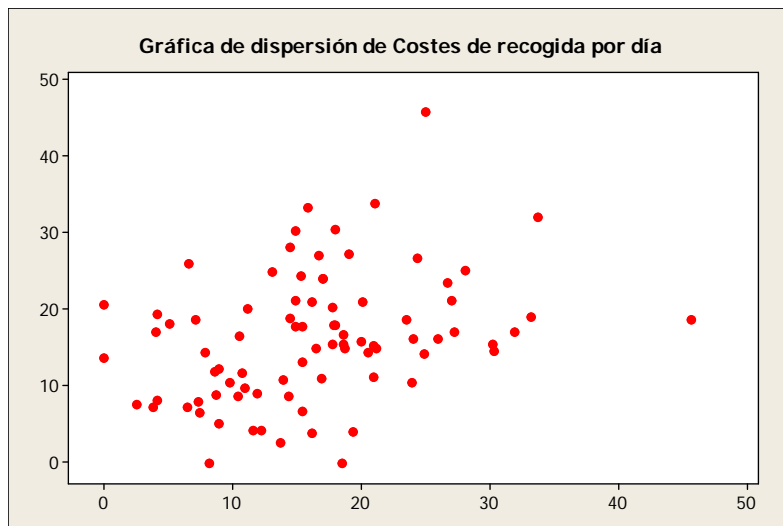


Figura 123: Grafica de dispersión de coste de recogida por día vehículo 1500 kg. PMA código 08172

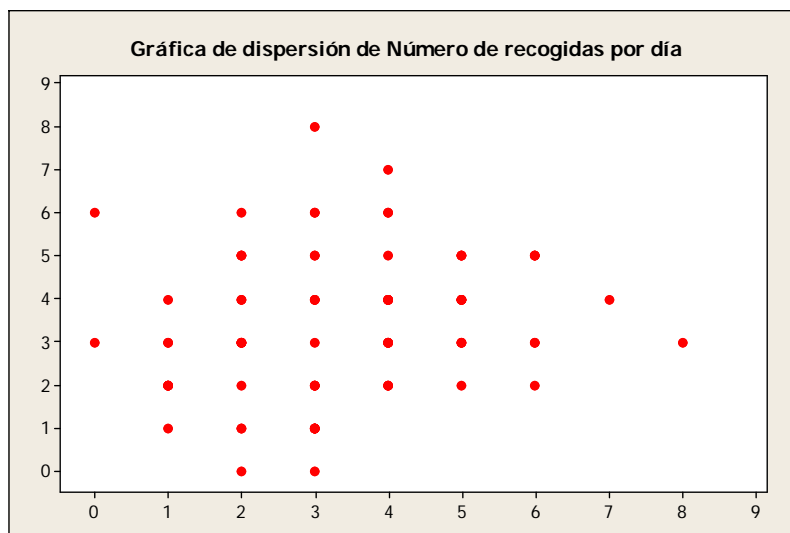


Figura 124: Grafica de dispersión de número de recogidas por día vehículo 1500 kg. PMA código 08172

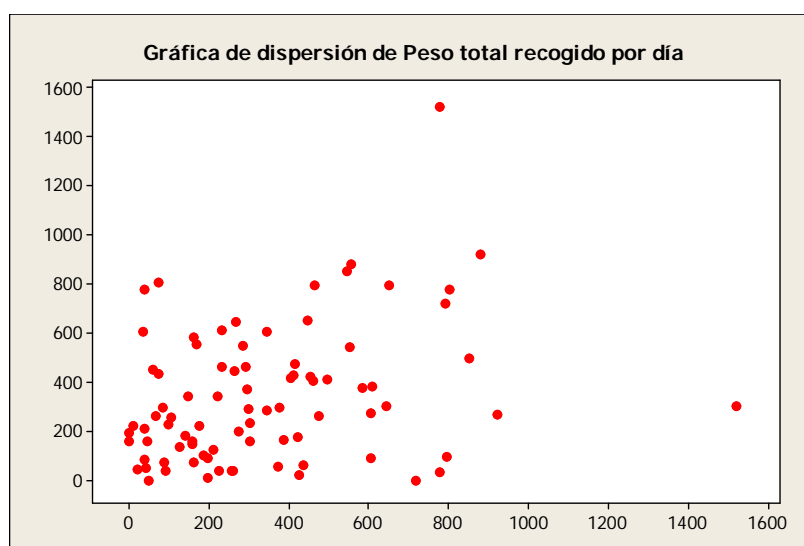


Figura 125: Grafica de dispersión de peso total recogido por día vehículo 1500 kg. PMA código 08172

En las tres gráficas se puede observar como todos los puntos se distribuyen de manera aleatoria en los planos correspondientes, pudiéndose concluir que las observaciones de las muestras son independientes.

2.7.3. Separación de las muestras de calibración y predicción

Tras verificar que las 84 observaciones obtenidas del estudio observacional de cada una de las variables son independientes, hay que separar estas 84 observaciones en dos muestras diferentes, la de calibración y la de predicción. La de calibración se emplea para la creación y elaboración de los modelos necesarios para el estudio de los costes de la red de distribución, mientras que la de predicción se emplea para comprobar que los modelos elaborados mediante el análisis de regresión estiman los costes de reparto y recogida de manera adecuada. Esta separación se lleva a cabo mediante una asignación de un número aleatorio a cada una de las observaciones, y si éste es mayor o igual que el valor 0,5, se asigna a la

muestra de calibración. Si por el contrario, este número aleatorio es menor que 0,5, se asigna a la muestra de predicción. El listado de números aleatorios asignados a cada una de las observaciones relativas al reparto es el de la tabla 279.

Tabla 279: Asignación de números aleatorios a las observaciones de la muestra de repartos vehículo 1500 kg. PMA código postal 08172

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,669	22	0,369	43	0,373	64	0,002
2	0,457	23	0,329	44	0,574	65	0,365
3	0,922	24	0,015	45	0,451	66	0,896
4	0,254	25	0,146	46	0,182	67	0,575
5	0,751	26	0,892	47	0,678	68	0,995
6	0,664	27	0,655	48	0,600	69	0,541
7	0,168	28	0,083	49	0,986	70	0,776
8	0,870	29	0,533	50	0,885	71	0,381
9	0,610	30	0,846	51	0,960	72	0,741
10	0,178	31	0,845	52	0,672	73	0,928
11	0,578	32	0,636	53	0,178	74	0,552
12	0,086	33	0,661	54	0,881	75	0,088
13	0,633	34	0,328	55	0,809	76	0,716
14	0,911	35	0,941	56	0,111	77	0,134
15	0,710	36	0,990	57	0,827	78	0,795
16	0,285	37	0,630	58	0,021	79	0,440
17	0,372	38	0,009	59	0,340	80	0,001
18	0,520	39	0,006	60	0,019	81	0,126
19	0,555	40	0,418	61	0,272	82	0,026
20	0,792	41	0,328	62	0,116	83	0,157
21	0,556	42	0,042	63	0,342	84	0,343

Quedando las muestras de calibración y predicción de acuerdo a esta asignación como se muestra en la tabla 280.

Tabla 280: Muestra de calibración y de predicción de repartos del vehículo de 1500 kg. PMA código postal 08172

Día	REPARTOS. MUESTRA DE CALIBRACIÓN			Día	REPARTOS. MUESTRA DE PREDICCIÓN		
	Repartos por día	Peso total repartido por día	Costes de reparto por día		Repartos por día	Peso total repartido por día	Costes de reparto por día
1	11	310	34	1	12	281	33
2	14	270	38	2	9	82	21
3	5	71	12	3	5	72	12
4	3	151	12	4	4	32	9
5	5	16	10	5	12	297	35
6	9	165	24	6	2	58	6
7	11	158	29	7	6	22	12
8	14	195	37	8	6	775	34
9	11	323	36	9	3	114	11
10	9	363	30	10	6	170	17
11	13	278	36	11	11	94	24
12	6	78	16	12	7	134	20
13	13	426	39	13	8	322	27
14	10	586	36	14	7	125	19
15	6	108	17	15	5	81	13
16	4	137	12	16	11	137	28
17	8	136	23	17	8	68	19
18	8	300	25	18	8	114	20
19	13	290	37	19	8	254	25
20	9	276	28	20	9	513	35
21	6	334	22	21	8	167	24
22	9	116	23	22	8	143	23
23	3	71	8	23	8	535	33
24	11	503	40	24	10	143	28
25	4	60	11	25	11	488	39
26	6	287	20	26	9	275	32
27	7	138	22	27	12	122	32
28	14	348	43	28	5	27	12
29	7	240	22	29	7	296	25
30	13	167	34	30	9	132	26
31	7	141	22	31	7	383	26
32	6	89	18	32	8	78	20
33	7	178	21	33	7	75	19
34	10	140	28	34	8	70	20
35	14	782	49	35	7	915	37
36	8	331	24	36	11	199	33
37	8	166	23	37	9	273	30
38	5	155	17	38	8	180	26
39	9	170	26	39	2	7	4
40	8	142	21	40	1	5	2
41	7	172	21				
42	3	29	8				
43	5	135	17				
44	4	33	11				

En este caso, debido a la asignación aleatoria, se tienen 44 observaciones para la muestra de calibración y 40 observaciones para la muestra de predicción relativas a los repartos. Una vez elaboradas estas muestras para los repartos, se realiza un proceso análogo para las recogidas. La asignación de números aleatorios a cada una de las 84 observaciones relativas a las recogidas es el de la tabla 281.

Tabla 281: Asignación de números aleatorios a las observaciones de la muestra de recogidas vehículo 1500 kg. PMA código postal 08172

Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio	Día	Número aleatorio
1	0,462	22	0,190	43	0,617	64	0,935
2	0,875	23	0,595	44	0,626	65	0,550
3	0,337	24	0,331	45	0,950	66	0,466
4	0,191	25	0,797	46	0,595	67	0,430
5	0,362	26	0,127	47	0,962	68	0,422
6	0,588	27	0,922	48	0,576	69	0,471
7	0,195	28	0,080	49	0,919	70	0,117
8	0,541	29	0,295	50	0,576	71	0,630
9	0,057	30	0,323	51	0,552	72	0,630
10	0,339	31	0,698	52	0,638	73	0,535
11	0,526	32	0,400	53	0,762	74	0,704
12	0,468	33	0,167	54	0,027	75	0,787
13	0,969	34	0,365	55	0,926	76	0,042
14	0,172	35	0,341	56	0,182	77	0,505
15	0,914	36	0,158	57	0,772	78	0,416
16	0,504	37	0,750	58	0,570	79	0,947
17	0,789	38	0,345	59	0,680	80	0,862
18	0,612	39	0,745	60	0,023	81	0,896
19	0,115	40	0,340	61	0,831	82	0,849
20	0,977	41	0,088	62	0,439	83	0,659
21	0,550	42	0,888	63	0,867	84	0,182

A partir de este listado de números aleatorios se confeccionan las muestras de calibración y predicción relativas a las recogidas. Estas muestras se pueden observar en la tabla 282.

Tabla 282: Muestra de calibración y de predicción de recogidas del vehículo de 1500 kg. PMA código postal 08172

Día	RECOGIDAS. MUESTRA DE CALIBRACIÓN			Día	RECOGIDAS. MUESTRA DE PREDICCIÓN		
	Recogidas por día	Peso total recogido por día	Costes de recogida por día		Recogidas por día	Peso total recogido por día	Costes de recogida por día
1	5	606	25	1	3	276	14
2	2	46	7	2	2	344	13
3	3	428	16	3	4	147	15
4	4	854	27	4	5	159	19
5	8	552	33	5	1	22	4
6	4	344	20	6	5	411	24
7	2	223	11	7	3	496	17
8	6	178	21	8	3	545	19
9	3	422	16	9	3	286	16
10	2	60	7	10	6	453	26
11	3	374	15	11	4	295	18
12	3	378	16	12	6	586	30
13	4	161	15	13	5	158	17
14	2	303	11	14	4	644	24
15	6	881	34	15	4	267	17
16	4	304	19	16	5	921	32
17	4	778	25	17	4	556	21
18	6	161	21	18	4	169	15
19	0	0	0	19	4	386	21
20	2	719	19	20	5	610	27
21	3	794	24	21	4	232	17
22	5	652	27	22	7	1.520	46
23	6	448	24	23	5	804	28
24	3	265	15	24	5	76	14
25	4	477	21	25	5	98	15
26	4	417	20	26	3	465	18
27	3	404	18	27	3	40	9
28	2	461	15	28	3	66	9
29	5	231	19	29	2	605	17
30	5	796	30	30	1	36	4
31	4	294	18	31	2	779	19
32	4	299	18	32	1	40	4
33	1	86	5	33	3	257	12
34	2	261	12	34	2	39	7
35	1	437	10	35	1	10	3
36	3	73	10	36	2	198	11
37	3	89	11				
38	2	105	9				
39	2	186	9				
40	4	140	14				
41	2	128	8				
42	1	212	7				
43	1	226	8				
44	3	197	14				
45	0	0	0				
46	3	49	8				
47	1	41	4				
48	3	91	12				

2.7.4. Modelo de costes de reparto

Como para el resto de modelos de costes de reparto elaborados de este trabajo de investigación, los parámetros de regresión se calculan empleando el método de máxima verosimilitud. Se estudiará el valor de máxima verosimilitud obtenido bajo diferentes premisas, como diferentes distribuciones de probabilidad, ó transformaciones de las variables exógenas. Los parámetros de regresión del modelo que mayor valor de verosimilitud proporcionen, serán los que se empleen en la simulación del modelo para estimar los costes de reparto. Los supuestos que se van a analizar son los mismos que los modelos de reparto anteriores, y el estudio de máxima verosimilitud completo para cada uno de estos supuestos se pueden ver a continuación.

Tabla 283: VMV modelo de costes de reparto con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
24,564					-164,5		0%	10,29
0,137	2,961				-110,8	53,683	91%	3,07
0,510	2,318	0,023			-75,4	35,396	98%	1,39
-0,555	2,606	0,023	-0,016		-75,0	0,381	98%	1,40
-0,083	2,263	0,031		-0,00001	-73,2	1,805	98%	1,34

Tabla 284: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
3,201					-164,5		0%	10,29
2,239	0,109				-118,7	45,808	87%	3,68
2,255	0,092	0,001			-103,8	14,875	93%	2,65
1,527	0,258	0,001	-0,009		-83,6	20,211	97%	1,70
1,506	0,239	0,001	-0,007	-0,000001	-80,7	2,895	98%	1,61

Tabla 285: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,041					-164,5		0%	10,29
0,075	-0,004				-127,4	37,141	81%	4,48
0,074	-0,003	-0,00001			-121,7	5,713	85%	3,98
0,131	-0,015	-0,00001	0,001		-94,0	27,686	96%	2,15
0,131	-0,015	-0,00001	0,001	0	-94,0	0,000	96%	2,17

Tabla 286: VMV modelo de costes de reparto con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-4,956					-164,5		0%	10,29
-2,481	-0,289				-113,9	50,572	90%	3,30
-2,512	-0,235	-0,002			-89,4	24,534	97%	1,91
-1,621	-0,453	-0,002	0,012		-78,7	10,722	98%	1,52
-1,704	-0,374	-0,004	0,008	0,000003	-73,8	4,866	98%	1,38

Tabla 287: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
22,645					-166,8		-4%	10,47
-0,890	3,067				-112,3	54,501	91%	3,10
0,620	2,258	0,024			-77,6	34,700	98%	1,41
0,620	2,258	0,024	0		-77,6	0,000	98%	1,43
0,629	2,202	0,028		-0,00001	-74,6	2,984	98%	1,40

Tabla 290: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
3,120					-166,8		-4%	10,47
2,255	0,108				-119,8	46,953	87%	3,68
2,287	0,085	0,001			-103,1	16,725	92%	2,88
1,796	0,202	0,001	-0,006		-88,3	14,771	97%	1,84
1,486	0,260	0,001	-0,009	-0,0000004	-82,2	6,095	98%	1,63

Tabla 291: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,044					-166,8		-4%	10,47
0,072	-0,004				-127,4	39,398	81%	4,51
0,072	-0,003	-0,00001			-120,9	6,487	85%	4,02
0,072	-0,003	-0,00001	0		-120,9	0,019	84%	4,07
0,072	-0,003	-0,00001		0	-120,9	-0,026	84%	4,07

Tabla 292: VMV modelo de costes de reparto con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-4,759					-166,8		-4%	10,47
-2,423	-0,294				-115,9	50,840	90%	3,31
-2,574	-0,223	-0,002			-91,5	24,476	96%	1,94
-1,442	-0,497	-0,002	0,014		-79,4	12,028	98%	1,56
-1,442	-0,497	-0,002	0,014	0	-79,4	0,000	98%	1,58

Tabla 293: VMV modelo de costes de reparto con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
24,123					-165,7		0%	10,30
0,414	2,904				-110,7	55,002	91%	3,09
0,662	2,294	0,023			-75,9	34,811	98%	1,40
-0,734	2,685	0,022	-0,022		-75,2	0,733	98%	1,40
-0,299	2,346	0,031	-0,006	-0,00001	-73,2	1,952	98%	1,36

En las tablas se puede observar como el valor de máxima verosimilitud es mayor para la distribución normal y para la distribución logística. El valor es idéntico para ambas, pero en el caso de la distribución normal el valor R² ajustado es un poco mayor que para la distribución logística. Por otro lado, la varianza es menor para la distribución normal, algo que es lógico ya que los estimados de los parámetros de regresión cuando los errores se distribuyen de manera normal son los de menor varianza, de acuerdo al teorema de Gauss-Markov. Por tanto, el modelo que se empleará para estimar los costes de

reparto de este vehículo en este código postal será el elaborado considerando que los errores se distribuyen normalmente.

2.7.4.1. Selección de variables para el modelo de costes de reparto

Tras la inclusión una a una de las variables número de repartos y peso total repartido, se puede comprobar en el cuadro relativo al supuesto de que los errores se distribuyen de manera normal, como el cociente de verosimilitud es mayor de dos unidades. No ocurre lo mismo al introducir, la variable número de repartos elevado a la segunda potencia, siendo en esta ocasión el cociente de verosimilitud de 0,381 unidades. Sin embargo, al introducir la cuarta variable, peso total repartido a la segunda potencia junto con las dos primeras variables, y sin tener en cuenta los valores de la tercera variable, se puede comprobar cómo en esta ocasión el cociente de verosimilitud es 2,186 comparándolo con el modelo de solo dos variables. Por tanto, la adición de esta cuarta variable aporta un mejor ajuste a los datos.

Resumiendo, el modelo de costes de reparto consta de tres variables, número de repartos, peso total repartido por día, y peso total repartido por día elevado al cuadrado. Sus correspondientes parámetros de regresión son 2,263, 0,031 y -0,00001, con un valor umbral igual a -0,083.

2.7.4.2. Adecuación del modelo de costes de reparto

Como el mayor valor máximo de verosimilitud es el que aporta una distribución de probabilidad normal, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores.

En la primera gráfica se puede observar como los errores se distribuyen de manera normal ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón. Este análisis gráfico se puede ver en la figura 126.

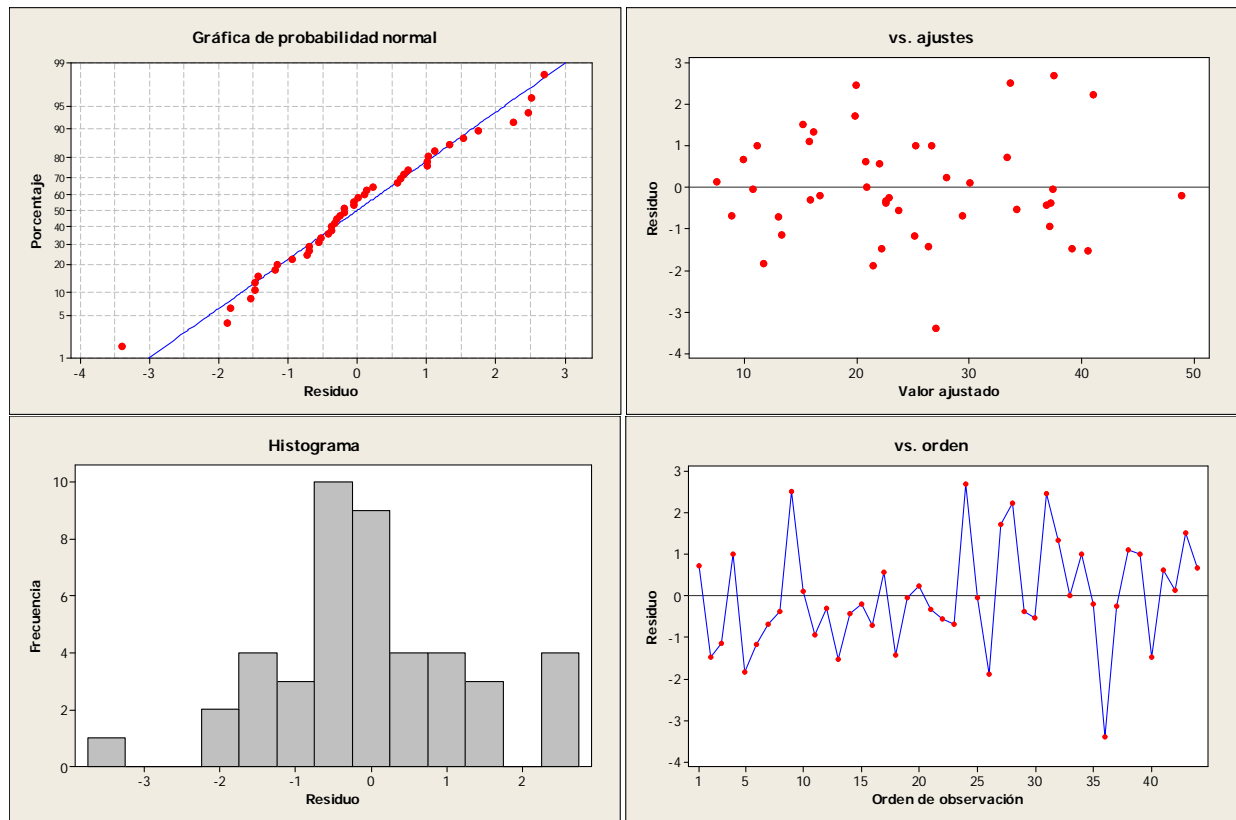


Figura 126: Análisis gráficos de los residuales para modelo de costes de repartos vehículo 1500 kg. PMA código postal 08172

2.7.4.3. Validación del modelo de costes de reparto

Para comprobar que el modelo predice de manera adecuada los valores de los costes de reparto ante nuevos valores de las variables exógenas, se realiza la prueba *t* pareada o prueba de Welch, con los datos de la muestra de predicción previamente obtenida. El cuadro que resume esta prueba de Welch es el de la tabla 294.

Tabla 294: Resultado prueba Welch para el modelo de costes de reparto del vehículo 1500 kg. PMA para el código postal 08172

	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	23,50	22,75	-0,258	Límite superior I.C	$t_{0,025;37}$
Varianza	59,77	89,54	2,59	0,241	1,96
Varianza/n			0,065		
Máximo	37,27	39,02		Límite inferior I.C	
Mínimo	6,22	2,03		-0,757	

En él se puede observar como el modelo estima los costes de recogida de manera adecuada, ya que el intervalo de confianza creado incluye el cero. Para elaborarlo se ha empleado la distribución *t* de Student, con un nivel de significancia del 95% y 37 grados de libertad, cuyo valor es 1.96. Los valores Z con los que se elabora este cuadro provienen de la tabla 295.

Tabla 295: Valores de la prueba de Welch para el modelo de costes de reparto del vehículo de 1500 kg. para el código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_4$	Costes de recogida estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
-0,082683	2,2632716	0,0312657	-0,00001			
	12	281	78961	34,93	32,92	2,01
	9	82	6724	22,77	21,30	1,47
	5	72	5184	13,42	12,49	0,93
	4	32	1024	9,96	9,04	0,92
	12	297	88209	35,33	35,20	0,13
	2	58	3364	6,22	6,04	0,18
	6	22	484	14,18	11,90	2,28
	6	775	600625	30,67	33,51	-2,84
	3	114	12996	10,12	10,95	-0,83
	6	170	28900	18,47	16,53	1,94
	11	94	8836	27,65	24,49	3,16
	7	134	17956	19,74	19,62	0,12
	8	322	103684	26,87	26,63	0,25
	7	125	15625	19,48	18,77	0,71
	5	81	6561	13,69	13,47	0,21
	11	137	18769	28,88	28,07	0,81
	8	68	4624	20,10	18,58	1,52
	8	114	12996	21,44	19,85	1,59
	8	254	64516	25,21	25,19	0,02
	9	513	263169	33,23	34,85	-1,62
	8	167	27889	22,92	24,00	-1,08
	8	143	20449	22,25	22,64	-0,39
	8	535	286225	31,39	32,90	-1,51
	10	143	20449	26,78	27,91	-1,13
	11	488	238144	37,27	39,02	-1,75
	9	275	75625	28,00	32,05	-4,05
	12	122	14884	30,72	32,20	-1,48
	5	27	729	12,07	11,54	0,53
	7	296	87616	23,99	25,40	-1,42
	9	132	17424	24,21	25,71	-1,50
	7	383	146689	26,01	25,76	0,26
	8	78	6084	20,39	20,27	0,12
	7	75	5625	18,04	19,22	-1,18
	8	70	4900	20,15	19,68	0,47
	7	915	837225	34,53	37,29	-2,76
	11	199	39601	30,57	33,14	-2,57
	9	273	74529	27,95	29,82	-1,88
	8	180	32400	23,27	26,15	-2,88
	2	7	49	4,66	4,01	0,65
	1	5	25	2,34	2,03	0,30

2.7.5. Modelo de costes de recogida

Una vez finalizado el modelo de los costes de reparto, hay que realizar el modelo de costes de recogida. La creación este modelo es análogo al de los costes de reparto, por lo que los pasos a seguir son idénticos.

Tabla 296: VMV modelo de costes de recogida con distribución de probabilidad normal vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
15,552					-167,2		0%	7,97
2,430	4,117				-133,3	33,930	75%	3,97
0,438	2,909	0,018			-56,4	76,945	99%	0,81
0,035	3,220	0,018	-0,043		-55,3	1,108	99%	0,80
0,238	2,893	0,020		-0,000002	-55,7	-0,441	99%	0,81

Tabla 297: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena logarítmica vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,744					-167,2		0%	7,97
2,042	0,205				-140,5	26,737	66%	4,62
1,875	0,151	0,001			-107,4	33,146	91%	2,34
1,495	0,364	0,001	-0,023		-94,1	13,230	95%	1,80
1,482	0,370	0,001	-0,023	-0,0000002	-92,7	1,423	95%	1,76

Tabla 298: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena recíproca vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,064					-167,2		0%	7,97
0,095	-0,009				-147,7	19,560	55%	5,36
0,105	-0,006	-0,00005			-128,1	19,597	80%	3,60
0,138	-0,024	-0,00003	0,002		-118,4	9,661	86%	2,98
0,138	-0,024	-0,00003	0,002	0	-118,4	0,000	86%	3,01

Tabla 299: VMV modelo de costes de recogida con distribución de probabilidad normal con función cadena raíz vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-3,944					-167,2		0%	7,97
-2,354	-0,477				-136,5	30,709	72%	4,25
-2,042	-0,349	-0,002			-87,3	49,223	96%	1,54
-1,599	-0,621	-0,002	0,032		-74,3	13,042	98%	1,19
-1,420	-0,615	-0,003	0,032	0,000002	-67,1	7,144	98%	1,03

Tabla 300: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
15,359					-169,8		0%	7,97
2,353	4,088				-135,7	34,133	75%	3,98
0,556	2,861	0,018			-58,6	77,103	99%	0,81
0,393	2,999	0,018	-0,016		-57,9	0,691	99%	0,81
0,431	2,855	0,019		-0,000002	-58,4	-0,488	99%	0,81

Tabla 301: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena logarítmica vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
2,732					-169,8		0%	7,97
2,043	0,192				-144,5	25,363	65%	4,73
1,939	0,154	0,001			-106,4	38,016	90%	2,47
1,490	0,389	0,001	-0,026		-95,5	10,926	95%	1,84
1,490	0,389	0,001	-0,026	0	-95,5	0,000	95%	1,86

Tabla 302: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena recíproca vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
0,065					-169,8		0%	7,97
0,098	-0,009				-150,8	19,036	54%	5,41
0,100	-0,006	-0,00005			-128,5	22,311	77%	3,81
0,104	-0,009	-0,00005	0,0003		-124,7	3,801	80%	3,56
0,104	-0,009	-0,00005	0,0003	0	-124,7	0,000	80%	3,60

Tabla 303: VMV modelo de costes de recogida con distribución de probabilidad doble exponencial con función cadena raíz vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
-3,924					-169,8		0%	7,97
-2,241	-0,502				-139,7	30,092	71%	4,27
-2,073	-0,356	-0,002			-87,0	52,692	96%	1,57
-1,762	-0,534	-0,002	0,022		-76,2	10,840	98%	1,24
-1,762	-0,534	-0,002	0,022	0	-76,2	0,000	98%	1,25

Tabla 304: VMV modelo de costes de recogida con distribución de probabilidad logística vehículo 1500 kg. PMA código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	Valor máx. veros.	Coci. de veros.	R ² adj.	σ
15,294					-168,0		0%	7,97
2,194	4,114				-133,8	34,109	75%	3,98
0,420	2,916	0,018			-57,1	76,730	99%	0,81
0,080	3,203	0,018	-0,039		-56,0	1,153	99%	0,80
0,081	2,912	0,020		-0,000003	-56,9	-0,916	99%	0,81

En las tablas anteriores se puede observar como el valor de máxima verosimilitud es mayor para la distribución normal. Por tanto, el modelo que se empleará para estimar los costes de reparto de este vehículo en este código postal será el elaborado considerando que los errores se distribuyen normalmente.

2.7.5.1. Selección de las variables del modelo de costes de recogida

Si se observa el cuadro en el que los errores se distribuyen según una distribución de probabilidad normal, se puede observar como al introducir las variables número de recogidas y peso total recogido, el valor de máxima verosimilitud al introducir cada una de estas variables una a una, difiere en más de dos unidades, por lo que el modelo que incluye estas variables ajusta mejor a los datos que aquel que no las tiene. Sin embargo, al introducir la variable número de recogidas por día elevada al cuadrado el valor de máxima verosimilitud sólo se incrementa en 1,108 unidades, por lo que la adición de esta variable no aporta un mejor ajuste. Lo mismo ocurre al incluir la variable peso total recogido elevada al segundo cuadrado. El cociente de verosimilitud también difiere en menos de dos unidades, por lo que no se incorpora esta variable al modelo de costes de recogida.

En definitiva, este modelo constará únicamente con las variables número de recogidas por día y peso total recogido por día. Los parámetros de regresión para estas variables son 2,909 y 0,018, con un valor umbral igual a 0,438.

2.7.5.2. Adecuación del modelo

Como el mayor valor máximo de verosimilitud es el que aporta una distribución de probabilidad normal, los pasos para comprobar la adecuación del modelo serán los del método de mínimos cuadrados, es decir, hay que realizar un análisis gráfico para comprobar que se cumplen los supuestos de los errores. Este análisis gráfico se puede comprobar en la figura 127.

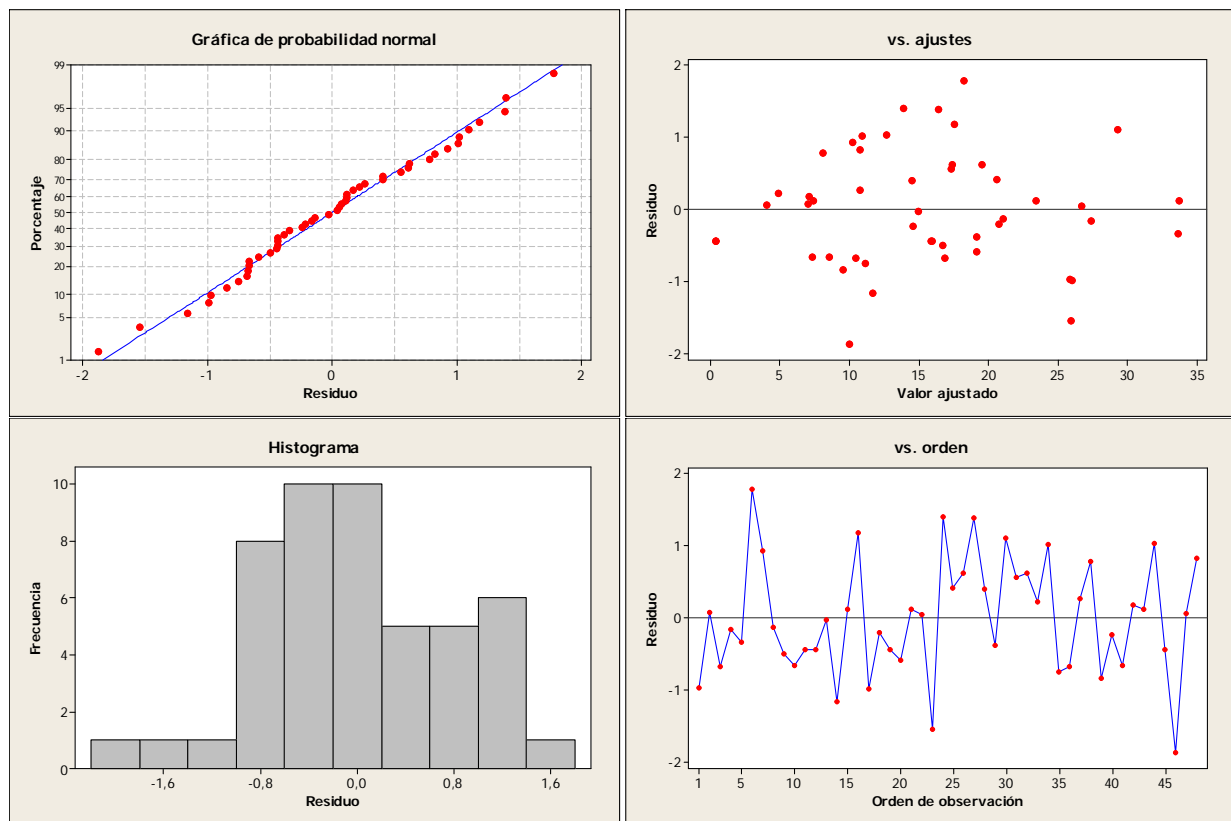


Figura 127: Análisis gráficos de los residuales para modelo de costes de recogidas vehículo 1500 kg. PMA código postal 08172

En la primera gráfica se puede observar como los errores se distribuyen de manera normal ya que los puntos caen más o menos sobre una línea recta. En la segunda, en la que aparecen los residuos frente a sus valores ajustados, los puntos se distribuyen de manera aleatoria en todo el eje y, lo que indica que la varianza es constante para los diferentes valores de los residuos, y por tanto se cumple el principio de homocedasticidad. En la tercera, se puede observar un histograma de la distribución de los residuales. Y por último, en la cuarta se puede observar como los residuales son independientes, ya que en la gráfica de éstos en función del tiempo no se aprecia ningún tipo de patrón.

2.7.5.3. Validación del modelo de costes de recogida

Para comprobar que el modelo predice de manera adecuada los valores de los costes de reparto ante nuevos valores de las variables exógenas, se realiza la prueba t pareada o prueba de Welch, con los datos de la muestra de predicción previamente obtenida. Los valores Z con los que se elabora este cuadro provienen de la tabla 305.

Tabla 305: Valores de la prueba de Welch para el modelo de costes de recogida del vehículo de 1500 kg. para el código postal 08172

$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.
0,438	2,909	0,018			
	3	276	14,12	14,18	-0,06
	2	344	12,43	13,07	-0,65
	4	147	14,71	15,46	-0,75
	5	159	17,84	18,71	-0,87
	1	22	3,74	3,84	-0,10
	5	411	22,36	24,06	-1,70
	3	496	18,06	17,08	0,98
	3	545	18,94	19,05	-0,11
	3	286	14,30	15,85	-1,56
	6	453	26,02	26,03	-0,01
	4	295	17,37	17,81	-0,44
	6	586	28,40	30,23	-1,83
	5	158	17,82	16,51	1,31
	4	644	23,63	24,03	-0,40
	4	267	16,86	17,01	-0,14
	5	921	31,50	31,99	-0,48
	4	556	22,05	21,06	0,99
	4	169	15,11	14,98	0,13
	4	386	19,00	21,19	-2,19
	5	610	25,93	27,09	-1,17
	4	232	16,24	16,78	-0,54
	7	1520	48,07	45,70	2,36
	5	804	29,41	28,18	1,22
	5	76	16,35	14,46	1,89
	5	98	16,74	14,54	2,20
	3	465	17,51	17,97	-0,47
	3	40	9,88	9,00	0,88
	3	66	10,35	8,69	1,66
	2	605	17,11	16,97	0,13
	1	36	3,99	4,06	-0,07
	2	779	20,23	19,40	0,83
	1	40	4,06	4,13	-0,06
	3	257	13,78	12,32	1,46
	2	39	6,96	6,52	0,44
	1	10	3,53	2,57	0,96
	2	198	9,81	10,81	-1,00

El cuadro que resume esta prueba de Welch es la tabla 306.

Tabla 306: Resultado prueba Welch para el modelo de costes de recogida del vehículo 1500 kg. PMA para el código postal 08172

	Costes de reparto estimados	Costes de la muestra de predicción	Z=Costes Est. - Costes Predic.	Prueba de Welch	
Media	17,34	17,26	0,079	Límite superior I.C	$t_{0,025;34}$
Varianza	78,11	78,25	1,29	0,450	1,96
Varianza/n			0,036		
Máximo	48,07	45,70		Límite inferior I.C	
Mínimo	3,53	2,57		-0,292	

En la tabla 306 se puede observar como el modelo estima los costes de recogida de manera adecuada, ya que el intervalo de confianza creado incluye el cero. Para elaborarlo se ha empleado la distribución t de Student, con un nivel de significancia del 95% y 35 grados de libertad, cuyo valor es 1.96. Los valores Z con los que se elabora este cuadro provienen de la siguiente tabla.

2.7.6. Modelo de carga de trabajo

En este modelo se van a estudiar la distribución de probabilidad que siguen las variables aleatorias que conforman la carga la de trabajo, así como sus parámetros. Al igual que en casos anteriores estas variables son el número de repartos por día y el peso de cada uno de los repartos, en el caso de los repartos, y el número de recogidas por día y el peso de cada recogida, en el de las recogidas.

Las muestras que se emplean para las variables número de repartos y número de recogidas y peso total recogido son las mismas muestras que las muestras de calibración que se emplearon para ajustar los modelos de los costes de reparto y recogida. Las muestras para las variables peso de cada reparto y peso de cada recogida, es una muestra con los pesos unitarios de cada uno de los repartos y de cada una de las recogidas.

2.7.6.1. Número de repartos por día

A través de Minitab se obtienen los parámetros para cada una de las distribuciones de probabilidad que hacen que éstas sean las que mejor se ajusten a los valores de la muestra. Esta variable aleatoria es un conteo de algún elemento, y por tanto, una variable discreta. Al ser una variable discreta, la primera distribución de probabilidad que se comprueba es la distribución de Poisson. El test no paramétrico con el que se comprueba si el ajuste para las distribuciones discretas es adecuado es el test Chi cuadrado. Para los datos de la muestra de calibración de esta variable, Minitab indica que la distribución de Poisson a comprobar es aquella con parámetro lambda igual a 8,25. Los resultados del test paramétrico para esta distribución de Poisson son los de la tabla 307.

Tabla 307: Prueba Chi Cuadrado número de repartos para el vehículo 1500 kg. PMA código postal 08172

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
3-4	6	0,136	0,0862	4	1,2854	alfa=0,05 k=5 Par. Est. =1 n=5-1-1=3
5-6	9	0,205	0,1976	9	0,0107	
7-8	10	0,227	0,2739	12	0,3492	
9-11	11	0,250	0,3115	14	0,5342	
13-14	8	0,182	0,1308	6	0,8751	
	44				3,05	7,85

En la tabla se puede observar como el valor límite para aceptar la hipótesis de que la una distribución de Poisson con lambda 8,25 ajusta adecuadamente la variable número de repartos por día es 7,85. Este valor corresponde a una distribución Chi cuadrado con 3 grados de libertad y nivel de significancia del 5%. Los grados de libertad es igual al número de intervalos de los que consta el test, al que hay que restar el número de parámetros estimados (para el caso de la distribución de Poisson es únicamente el parámetro lambda) y otra unidad. Como en este caso el valor del test es 3,05, menor al valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.7.6.2. Peso de cada reparto

La segunda variable a analizar para el modelo de carga de trabajo es el peso de cada reparto. Esta variable es una variable aleatoria continua, y al igual que para el resto de variables continuas, las distribuciones de probabilidad a analizar serán también continuas. Los valores dados por Minitab y los resultados del test Kolmogorov-Smirnov para cada una de las distribuciones analizadas se pueden ver en la tabla 308.

Tabla 308: Test Kolmogorov-Smirnov peso de cada reparto para el vehículo 1500 kg. PMA código postal 08172

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,259	alfa	0,136	alfa	0,104	Media	0,064	Lambda	0,253
26,26	4,958	0,7379	2,587	0,7886	1,987	2,455	1,224	18,43	4,828
Desv. Est.	c _α = 1,358	beta	c _α = 1,358	beta	c _α = 1,358	Desv. Est.	c _α = 1,358	Theta	c _α = 1,358
39,59	α = 0,05	35,58	α = 0,05	22,39	α = 0,05	1,291	α = 0,05	16,34	α = 0,05

Como se puede ver la única distribución que supera satisfactoriamente el test de Kolmogorov-Smirnov es la distribución lognormal. Por lo tanto, para la simulación se empleará esta distribución para estimar los pesos de cada uno de los repartos.

2.7.7.3. Número de recogidas por día

Al igual que para la variable número de repartos por día, esta variable es un conteo de algún elemento. Por ello, la primera distribución de probabilidad que se va a comprobar si ajusta adecuadamente a los valores de la muestra de esta variable es la distribución de probabilidad discreta de Poisson. Al ser una distribución discreta, el test no paramétrico a realizar es el test Chi cuadrado. El resultado de este test se puede ver en la tabla 309.

Tabla 309: Prueba Chi Cuadrado número de recogidas para el vehículo 1500 kg. PMA código postal 08172

Observaciones	O _i	f(O _i)	F(E _i)	E _i	Chi-cuadrado	
0-1	7	0,146	0,1728	8	0,2025	alfa=0,05 k=5 Est. par. =1 n=5-1-1=3
2	10	0,208	0,2097	10	0,0004	
3	12	0,250	0,2228	11	0,1596	
4	10	0,208	0,1775	9	0,2565	
5-8	9	0,188	0,2172	10	0,1945	
	48				0,81	

El valor lambda dado por Minitab para una distribución de Poisson con los datos de la muestra de esta variable aleatoria es 3,19. En la tabla se puede observar como el valor límite para aceptar la hipótesis de que la una distribución de Poisson con lambda 3,19 ajusta adecuadamente la variable número de repartos por día es 7,85. Este valor corresponde a una distribución Chi cuadrado con 3 grados de libertad y nivel de significancia del 5%. Como en este caso el valor del test es 0,81, menor al valor límite, se acepta la hipótesis de que esta distribución ajusta adecuadamente la variable número de repartos por día.

Tras comprobar que mediante una distribución de Poisson se puede ajustar esta variable, no es necesario comprobar si existen distribuciones continuas que ajusten adecuadamente los valores de la muestra, ya que al ser esta variable discreta, es preferible ajustarla mediante distribuciones de probabilidad discretas.

2.7.7.4. Peso de cada recogida

Esta variable aleatoria es la última a analizar para el modelo de carga de trabajo, y es una variable continua. Se determinan los valores de cada distribución, así como la bondad de ajuste de las mismas de igual manera que las variables continuas anteriores. Los resultados son los de la tabla 309.

Tabla 310: Test Kolmogorov-Smirnov peso de cada recogida para el vehículo 1500 kg. PMA código postal 08172

Normal	Test K - S	Gamma	Test K - S	Weibull	Test K - S	Lognormal	Test K - S	Logística	Test K - S
Media	0,224	alfa	0,056	alfa	0,046	Media	0,080	Lambda	0,231
102	2,774	0,6143	0,692	0,718	0,564	3,623	0,985	75,23	2,854
Desv. Est.	c _α = 1,358	beta	c _α = 1,358	beta	c _α = 1,358	Desv. Est.	c _α = 1,358	Theta	c _α = 1,358
136,9	α = 0,05	166,1	α = 0,05	82,24	α = 0,05	1,65	α = 0,05	64,15	α = 0,05

En la tabla 310 se puede observar como las distribuciones gamma, Weibull y lognormal superan satisfactoriamente el test de Kolmogorov-Smirnov, siendo la distribución Weibull la que menor discrepancia presenta.

2.7.8. Simulación de los costes de reparto y recogida, y de la carga de trabajo

En este apartado se llevará a cabo la simulación de los modelos creados para el estudio del vehículo de 1500 kg. de peso máximo autorizado en el código postal 08198. En primer lugar se simulará la carga de trabajo. Para ello hay que determinar el número de repartos y de recogidas para cada uno de los días simulados, y el peso total repartido y el peso total recogido por día. Empleando los ajustes de las distribuciones de probabilidad de Poisson obtenidos en los apartados 2.7.6.1 y 2.7.6.3. se simularán el número de repartos y el número de recogidas para cada uno días que se desean simular.

Posteriormente, se simulará el peso correspondiente de cada uno de los repartos y las recogidas mediante las distribuciones de probabilidad calculadas en los apartados 2.7.6.2. y 2.7.6.4. Sumando el peso unitario de todos los repartos simulados en cada uno de los días, se calculará el peso total repartido por día. Con el peso unitario de las recogidas simulado para cada uno de los días, se calculará el peso total recogido por día.

Finalmente, se simularán los modelos de costes de reparto y costes de recogida. Como valores de entrada a estos modelos se emplearán los datos simulados para el modelo de carga de trabajo. El número de repartos y el peso total repartido por día serán los datos de entrada para el modelo de costes de reparto. Análogamente, el número de recogidas y el peso total recogido por día serán los valores de entrada para el modelo de costes de recogida.

2.7.14.3. Simulación de los costes de reparto

La carga de trabajo relativa a los repartos para este código postal es la siguiente: el número de repartos puede ser descrito mediante una distribución de probabilidad Poisson con parámetro lambda igual a 8.25. Mientras que el peso de cada reparto sigue una distribución de probabilidad lognormal con media 2,455 y desviación estándar 1,291.

Un resumen de la carga de trabajo simulada es de la tabla 311, en el cual se puede observar el promedio de repartos por día simulados, su desviación estándar, el número máximo de repartos a realizar en un día, así como el número mínimo. Y los mismos datos para la variable peso total repartido, siendo este peso total repartido la suma de los pesos simulados para cada una de los repartos.

Tabla 311: Resumen de la simulación de la carga de trabajo relativa a los repartos del vehículo 1500 kg. PMA en el código postal 08172

	Repartos por día	Peso total repartido por día
Promedio	8	211
Desv. Estan.	3	137
Máximo	16	791
Mínimo	2	18

La simulación de la carga de trabajo en 250 días bajo las condiciones previamente enunciadas, son las entradas para la simulación en el modelo de costes de reparto. Para cada uno de los días se obtiene un estimado puntual del coste de reparto para la carga de trabajo simulada correspondiente. A partir de cada estimado puntual se desarrolla un intervalo de predicción, teniendo un total de 250 intervalos de predicción. Estos intervalos de predicción se calculan tal y como se explicó en el apartado 4.1.1.3. Los resultados que aquí se exponen son el valor medio de límite inferior de las 250 simulaciones, la varianza y desviación estándar de este límite inferior, así como un intervalo de confianza de este límite inferior. Para el límite superior del intervalo de predicción los resultados que se presentan son análogos. Estos resultados son los de la tabla 312.

Tabla 312: Resumen de la simulación de los costes de reparto del vehículo de 1500 kg. de PMA en el código postal 08172

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	21,08	26,55
Varianza	73	74
Desv. Estan.	8,54	8,60
I.C	20,19	25,65
	21,97	27,44

En el cuadro se puede ver como el límite inferior oscila entre los valores 20,19 y 21,97 euros por día, mientras que el límite superior va desde 25,65 hasta 27,44 euros.

2.7.14.4. Simulación de costes de recogida

La carga de trabajo relativa a las recogidas para este código postal queda como sigue: el número de recogidas puede ser expresado mediante una distribución de probabilidad Poisson con parámetro lambda 3,19. El peso de cada recogida, por otro lado, puede ser descrito mediante una distribución de probabilidad Weibull con parámetro alfa igual a 0,718, y una parámetro beta igual a 82,24. Un cuadro resumen de la carga de trabajo simulada es la tabla 313, en el cual se puede observar el promedio de recogidas por día simulados, su desviación estándar, el número máximo de recogidas a realizar en un día, así como el número mínimo. Y los datos análogos para la variable peso total recogido.

Tabla 313: Resumen de la simulación de la carga de trabajo relativo a las recogidas del vehículo 1500 kg. PMA en el código postal 08172

	Recogidas por día	Peso total recogido por día
Promedio	3	365
Desv. Estan.	2	329
Máximo	8	1744
Mínimo	0	0

Para la simulación de los costes de recogida, los resultados que se exponen son idénticos al caso de los costes de reparto. Se exponen los valores medios, varianza y desviación estándar, así como un intervalo de confianza de los límites superior e inferior de los límites de predicción de los 250 días simulados a partir de la carga de trabajo previamente simulada bajo las condiciones expuestas en el párrafo anterior. Los resultados son los que se puede observar en la tabla 314.

Tabla 314: Resumen de la simulación de los costes de recogida del vehículo de 1500 kg. de PMA en el código postal 08172

	Límite inferior interv. Predic.	Límite superior interv. Predic.
Promedio	14,85	18,15
Varianza	94	95
Desv. Estan.	9,69	9,77
I.C	13,83	17,13
	15,86	19,17

En la tabla se puede observar como el límite inferior de los estimados puntuales de los costes de recogidas se distribuyen con un valor medio de 14,85 euros por día y una desviación estándar de 9,69. Mientras que el límite superior de los estimados puntuales de los costes de reparto tienen un valor medio de 18,15 y una desviación estándar de 9,77 euros por día.

3. RESUMEN DE LOS RESULTADOS

En este apartado se va a presentar un resumen del apartado 2 de este capítulo, para poder observar de una manera más clara los resultados obtenidos. En primer lugar se presenta en la tabla 315 el resumen de los diferentes modelos elaborados.

Tabla 315: Resumen de los modelos elaborados para cada uno de los códigos postales

		Modelo costes de Reparto	Modelo costes de Recogida	Distr. Prob. Núm. Repartos	Distr. Prob. Núm. Recogidas	Distr. Prob. Peso de cada Reparto	Distr. Prob. Peso de cada Recogida
8198	1500 kg. PMA	Errores distribuidos según distr. Normal	Errores distribuidos según distr. Doble Exp.	Distr. Prob. Poisson	Distr. Prob. Lognormal	Distr. Prob. Poisson	Distr. Prob. Lognormal
8198	3500 kg. PMA	Errores distribuidos según distr. Logística	Errores distribuidos según distr. Logística	Distr. Prob. Poisson	Distr. Prob. Lognormal	Distr. Prob. Poisson	Distr. Prob. Gamma
8198	Más de 3500 kg. PMA		Errores distribuidos según distr. Doble Exp.			Distr. Prob. Empírica	Distr. Prob. Gamma
8197	1500 kg. PMA	Errores distribuidos según distr. Logística	Errores distribuidos según distr. Normal	Distr. Prob. Poisson	Distr. Prob. Lognormal	Distr. Prob. Poisson	Distr. Prob. Lognormal
8196	1500 kg. PMA	Errores distribuidos según distr. Normal	Errores distribuidos según distr. Logística	Distr. Prob. Poisson	Distr. Prob. Lognormal	Distr. Prob. Poisson	Distr. Prob. Lognormal
8195	1500 kg. PMA	Errores distribuidos según distr. Normal	Errores distribuidos según distr. Normal	Distr. Prob. Poisson	Distr. Prob. Lognormal	Distr. Prob. Poisson	Distr. Prob. Weibull
8174	1500 kg. PMA	Errores distribuidos según distr. Logística	Errores distribuidos según distr. Normal	Distr. Prob. Poisson	Distr. Prob. Lognormal	Distr. Prob. Poisson	Distr. Prob. Lognormal
8173	1500 kg. PMA	Errores distribuidos según distr. Logística	Errores distribuidos según distr. Normal	Distr. Prob. Poisson	Distr. Prob. Lognormal	Distr. Prob. Poisson	Distr. Prob. Lognormal
8172	1500 kg. PMA	Errores distribuidos según distr. Normal	Errores distribuidos según distr. Normal	Distr. Prob. Poisson	Distr. Prob. Lognormal	Distr. Prob. Poisson	Distr. Prob. Weibull

En la tabla 315 se puede comprobar cómo se han llevado a cabo el modelado de 9 vehículos en total en los 7 códigos postales analizados: 7 vehículos de 1500 kg. de PMA autorizado, uno por cada uno

de los códigos postales, 1 vehículo de 3500 kg. de PMA para el código postal 08198, y otro de más de 3500 kg. de PMA, también para el código postal 08198.

En los 9 vehículos modelados se han llevado a cabo un total de 17 análisis de regresión: 8 análisis de regresión para modelar los costes relativos a los repartos, y 9 análisis de regresión para modelar los costes relativos a las recogidas.

En la tabla 315 se puede comprobar cómo los análisis de regresión elaborados bajo el supuesto de que los errores se distribuyen bajo diferentes funciones de distribución de probabilidad han dado mejores resultados que aquellos modelos en los que las variables exógenas son transformadas bajo el criterio de los modelos lineales generalizados. De los 17 análisis de regresión, 10 de los modelos de costes han sido elaborados bajo el supuesto que los errores se distribuyen de acuerdo a distribución de probabilidad normal. 2 de ellos se han elaborado bajo el supuesto de que los errores se distribuyen de acuerdo a una distribución de probabilidad doble exponencial. Por último, los 5 modelos restantes han sido elaborados de acuerdo al supuesto de que los errores se distribuyen de acuerdo a una distribución de probabilidad logística.

Por otro lado, para cada uno de los 9 vehículos estudiados, se ha analizado la carga de trabajo para cada uno de ellos. En la tabla 315 se puede comprobar cómo el número de repartos de todos los vehículos se ha podido modelar mediante la distribución de probabilidad de Poisson. Respecto a la variable número de recogidas, en ocho de los nueve vehículos también se ha podido modelar esta variable mediante la distribución de probabilidad de Poisson. Únicamente, para el vehículo de más de 3500 kg. de PMA en el código postal 08198 no se ha podido ajustar la variable número de recogidas mediante ninguna función de distribución de probabilidad, y se ha tenido que modelar mediante una probabilidad de distribución empírica.

Para la variable peso de cada reparto, en los ocho casos modelados, esta variable ha sido modelada mediante una distribución de probabilidad lognormal. Para la variable peso de cada recogida, en 5 de los 9 vehículos modelados también se ha podido ajustar mediante una distribución de probabilidad lognormal. En otros dos vehículos la variable se ha ajustado mediante una distribución de probabilidad gamma, y en los dos últimos vehículos se ha empleado una distribución de probabilidad Weibull para el ajuste de la variable. Sin embargo, tal y como se puede comprobar en las tablas 94, 72, 200 y 310, la variable peso de cada recogida para los vehículos de 3500 kg. de PMA y más de 3500 kg. de PMA del código postal 08198, y los vehículo de 1500 kg. de PMA de los códigos postales 08195 y 08197, se puede modelar mediante una distribución de probabilidad lognormal. Por lo que se puede concluir que tanto la variable peso de cada reparto como la variable peso de cada recogida se pueden modelar mediante distribuciones de probabilidad lognormales.

CONCLUSIONS

In this thesis, an overnight parcel logistics company's capillary distribution network has been modelled by stochastic techniques. In particular, the capillary distribution network an overnight parcel logistics company has in Sant Cugat del Vallés, belonging to the hub located in Barberá del Vallés, in Barcelona's metropolitan area.

This stochastic modelling has been created to develop a tool which can help in the decision making related to the design of capillary distribution networks, both new ones or existing ones that need to be improved.

The tool developed in this research provides a stochastic modelling of this capillary distribution network. More specifically, the way the costs of each distribution areas are obtained for the design of the capillary network is by regression analysis. To predict the workload of each distribution area, the number of deliveries and collections, and the weight of each one, a model will be created by probability distribution functions.

With this tool seven distribution areas have been modelled, creating nine different models for nine different vehicles: seven vehicles of 1500 kg. GVW, one vehicle per each distribution area, another two ones, one of 3500 kg. GVW and another one of more than 3500 kg. GVW for the same distribution area.

This tool enables a parcel logistics company's capillary distribution network designers to create or redesign a network based on the incomes of the drivers, being these incomes crucial to ensure their continuity in the company. This continuity maximizes driver familiarity within their distribution area. With increased familiarity, driver performance enhances due to ease in finding addresses and locations, and efficiency in organizing daily routes. In this way, their ability to make deliveries and collections increases, being their production higher.

The first hypothesis made in the introduction, that the distribution costs can be predicted through a regression analysis, has been modeled successfully using the maximum likelihood technique. These distribution costs can be separated into delivery costs and collection costs. These two costs have been proved as independent, so a model can be developed for each cost in each area. In the case of the delivery costs, the outcome for the model has been identified as the delivery cost per day. The exogenous variables used to predict the value of this outcome, have been recognised as the number of deliveries per day and the total delivered weight per day. As

for the collection costs, the outcome for the model is the collection cost per day, and the exogenous variables are the number of collections per day and the total collected weight per day.

Due to the methodology based on the use of the maximum likelihood technique it has been possible to elaborate each delivery and collection cost model, based on regression analysis, under different assumptions, as the residuals of the regression model are distributed by different probability distribution functions, and the exogenous variables are transformed under the generalized linear model criteria. The different regression models have been compared through the maximum likelihood values, and for each vehicle in the different distribution areas the one with best results has been chosen. This methodology based in maximum likelihood, has given the possibility also to identify from all the exogenous variables used for the model adjustment, which ones have a significant impact in the behavior of the endogenous variable.

For the cases studied in this research, the regression models made under the assumption that the residuals are distributed by different probability distribution functions have had better results than the ones under the assumption of the exogenous variables' transformation. The first models have provided a better fit to the data of the samples, being their maximum likelihood values higher than the ones of the models made under the generalized linear models criteria.

These models have been developed with around 42 observations each sample, what it is approximately the number of working days in two months. The possibility of developing these models with only 42 observations per sample, allows the managers of the parcel logistics companies to check every two months if there are significant changes in the environment that these kinds of companies have to deal with, anticipating to its dynamism, and to any reduction in the income of the drivers.

In addition to the model for predicting the distribution costs of each area in a capillary distribution network, the second hypothesis made in the introduction, to create a model to estimate the workload of each distribution area, has been worked out satisfactorily.

The methodology followed to create this model has been adjusted by probability distribution functions. The variables that have been fitted are the number of deliveries per day, the weight of each delivery, the number of collections, and the weight of each collection. For the variable number of deliveries, in the eight vehicles that have been modelled, it has been

possible to adjust this variable with a Poisson probability distribution function. As for the variable number of collections, it has been possible to adjust the sample values with a Poisson probability distribution function in eight of the nine vehicles that have been modelled. It has been impossible to model for the vehicle of more than 3500 kg. of GVW in 08198 post code with any probability distribution function, and an empirical distribution function has been used instead.

For the variable weight of each delivery, all vehicles have been modelled with a lognormal probability distribution function. However, as for the variable weight of each collection, gamma, Weibull and lognormal probability distribution functions have been used to model the different vehicles. It is remarkable to say that, despite not using the lognormal probability distribution function, in all of them it could be used, because in these cases the lognormal probability distribution function successfully overcomes the validation tests.

Once the models are developed, according to the third hypothesis, a simulation has been run, and the distribution costs and the workload of each distribution area have been estimated. This simulation, made under the criteria of the Monte Carlo method, was of 250 days, more or less the number of working days in a year. After the simulation it is at disposal the number of deliveries, the delivered weight, the number of collections, and the collected weight, as well as, the delivery costs and the collection costs for each distribution area, and it is possible to study how the system is going to behave. With all this information, a new design of the capillary distribution network can be undertaken. This new design and optimization of the capillary distribution network can be hold in future researches. This new design can take into account the estimated costs to guarantee the continuity of the drivers, ensuring for them a minimum income, with a reasonable workload.

This kind of simulations can also contribute in the research of any modification of the charges for services in the capillary distribution network. The managers can predict the savings for the parcel logistics company in case they decide to reduce them; or they can advance how much the costs can increase if they have to pay more for them.

In summary, the hypotheses made in the introduction of this work have been successfully proven. A stochastic tool, based on regression analysis, has been developed and has been proved that it is valid technique to help in the decision making related to the design of a capillary distribution network. With the stochastic tool developed in this research, it has been

proved that the number of collections and deliveries can be estimated by Poisson probability distribution functions and the weight of each one of them by a lognormal probability distribution function. Moreover, the regression models made under the assumption that the residuals are distributed by different probability distribution functions have had better results than the ones under the assumption of the exogenous variables' transformation.

BIBLIOGRAFÍA

- Ashish, J. (2004). *Protocol design for a public logistics network*. North Carolina State University, USA.
- Ayyub, B. M., Mc Cuen, R. H. (2003). *Probability statistics and reliability for engineers and scientists*. Florida, USA: Chapman & Hall.
- Baldarcci, R., Bartolini E. y Aristide, M. (2011). An Exact Algorithm for the Pickup and Delivery Problem with Time Windows. *Operations Research*, 59 no. 2, 414-426.
- Banks, J. (1998). *Handbook of simulation: principles, methodology, advances, applications, and practice*. Nueva Jersey, USA: John Wiley & Sons.
- Batarra, M., Gunes, E., Laporte, G., Vigo, D., (2010). The Traveling Salesman Problem with Pickups, Deliveries, and Handling Costs. *Journal Transportation Science*, 44 no. 3, 383-399.
- Benton, W.C. y Rosseti, M. D. (1991). *The vehicle scheduling problem with intermittent customer demands*. The Ohio State University, USA.
- Berbeglia, G., Cordeau, J.F., Laporte, G. (2009). Dynamic pickup and delivery problems. *European Journal of Operational Research*, 202 no. 1, 8-15.
- Bertsimas, D. J. (1991). *A vehicle routing problem with stochastic demand*. Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
- Bertsimas, D. J., Karthik N. y Chung-Piaw T. (2004). Probabilistic combinatorial optimization: moments, semidefinite programming, and asymptotic bounds. *Society for Industrial and Applied mathematics*, 15, no 1, 185-209.
- Bowers Riego, D. (2005). *Optimization of Vehicle Routes for a Third Party logistics provider*. North Carolina State University, USA.
- Brown, A. M. (2003). *Improving the efficiency of hub operations in a less-than-truckload distribution network*. Virginia Polytechnic Institute, Virginia, USA.
- Caliskan Demirag, O. (2007). *Demand management in decentralized logistics*. Georgia Institute of Technology, USA.
- Çınlar, E. (1975). *Introduction to stochastic processes*. New Jersey, USA: Prentice Hall.
- Cordeau, J.F., Gendreau M., Laporte, G., Potvin J.Y. y Semet, F. (2002). A guide to vehicle routing heuristics. *Journal of the Operational Research Society*, 53 no 5, 512-522.
- Coss Bu, R. (1993). *Simulación: Un enfoque práctico*. México D.F., Mexico: Noriega Editores.
- Crainic, T.C. (2003). Long Haul Freight Transportation. Capítulo en: R.W. Hall, *Handbook of Transportation Science*, second edition. Kluwer Academic Publishers.

- D'Agostino R. y Stephens, M. A. (1986). *Goodness of fit techniques*. New York, USA: Marcel Dekker.
- Banhart, C. y Daeki, K. (1995). Routing models and solution procedures for regional less-than-truckload operations problems. *Annals of operational research*, 61, 67-90.
- Davison, A.C. (2003). *Statistical models*. Cambridge, Reino Unido: Cambridge University Press.
- Dennis, W. T. (2011). *Parcel and small package delivery industry*. North Charleston, USA: CreateSpace.
- Devang Bhalchandra, D. (2004). *Patterns of Freight Flow and Design of a less-than-truckload distribution network*. Georgia Institute of Technology, USA.
- Dong, G., Tang, J., Lai, K.K. y Kong, Y. (2009). An exact algorithm for vehicle routing and scheduling problem of free pickup and delivery service in flight ticket sales companies based on set-partitioning model. *Springer Science and Business*, 22, no.5, 789-799.
- Draper, N. y Smith H. (1998). *Applied regression analysis*. New Jersey, USA: John Wiley & Sons.
- Eksioglu, B., Volkan Vural, A. y Reisnan, A. (2009). The vehicle routing problem: A taxonomic review. *Journal of Computers & Industrial Engineering*, 57, no 4, 1472-1483.
- Estrada Romeu, M. A. (2007). *Análisis de estrategias eficientes en la logística de distribución de paquetería*. Univeristat Politecnica de Catalunya, Barcelona, España.
- Fishburn, P.C. (1978). *Operations research. An introduction to modern applications*. Princeton, USA: Auerbach Publisher
- Frosini, P. (2005). *Using innovative ICT tools to develop an agency for flexible transport and logistics services in urban and rural areas*. Università degli studi di Firenze.
- Fu, H. (2004). *Development of dynamic travel demand models for hurricane evacuation*. Louisiana State University, USA.
- Garthwaite, P.H., Jolliffe, I.T. y Jones, B. (2002). *Statistical inference*. Oxford, Reino Unido: Oxford University Press.
- Gendreau, M., Laporte, g., Segini, R. (1996). Stochastic vehicle routing. *European Journal Of Operational Research*, 88, no 1, 3-12.
- Gnanasekaran, V. (2004). *Improved opportunity cost algorithm for carrier selection in combinatorial auctions*. Louisiana State University, USA.
- Gonzales King-Kee, K. C. (2001). *Método de mínimos cuadrados ponderados para la estimación de los modelos lineales generalizados*. UNMSM, Lima, Perú.
- Gordon, G. (1978). *System simulation*. New Jersey, USA: Prentice Hall.

- Guasch, A., Piera, M. A., Casanovas, J. y Figueras, J. (2002). *Modelado y simulación. Aplicación a procesos logísticos de fabricación y servicios*. Barcelona, España: Edicions UPC.
- Gültekin, K. (2007). *Procurement in truckload transportation*. Georgia Institute of Technology, USA.
- Graybill, F.A. (1961). *An introduction to linear statistical models*. Nueva York, USA: Mac Graw Hill
- Hammersley, J.M. y Handscomb, D.C. (1968). *Monte Carlo Methods*. Londres, Reino Unido: Mehteun.
- Halls, P. (1990). Pseudo-likelihood theory for empirical likelihood. *Annals of Statistics*, 18, no 1, 121-140.
- Halton, J.H. (1971). A Retrospective and prospective survey of the Monte Carlo method. *SIAM review*, 12.
- Ho, C. (2005). *Public Logistics Network Cost Analysis*. North Carolina State University, USA.
- Hocking, R. R. (2003). *Methods and applications of linear models*. New Jersey, USA: John Willey and Sons.
- Hoff, A., Gribskovskaia, I., Laporte, G. y Lokketangen, A. (2009). Lasso solution strategies for the vehicle routing problem with pickups and deliveries. *European Journal of Operational Research*, 192, no 3, 755-766.
- House, W.C. (1972). *Operations research. An introduction to modern applications*. Princeton, USA: Auerbach Publisher.
- Hillier, F. y Lieberman, G. (1967). *Introduction to operations research*. Londres, Reino Unido: Holden Day.
- Hu, T.C. (1969). *Integer programming and network flows*. Massachusetts, USA: Addison-Wesley.
- Hyungtae, K. (2004). *Reliability Modelling with Load-Shared Data and Product-Ordering Decisions Considering Uncertainty in logistics operation*. Georgia Institute of Technology, USA.
- Holst, H. y Ingo Singe, I. (2011). Precariousness and Informality. Work and Employment in the German Parcel and Delivery Industry. *Work Package 2. SODIPER Research Report, from European Commission's DG Employment*
- Ingalls, R., Rosseti, M., Smith J. y Peters, B. (2004). *Proceedings of 2004 Winter Simulation Conference*. Washington D.C, USA.
- Jain, A. (2004). *Protocol design for a public logistics network*. North Carolina State University, USA.
- Johansson R. (1993). *System modelling & identification*. New Jersey, USA: Prentice Hall.

- Karacik, B. (2007). *Driver management for less-than-truckload carriers*. Georgia Institute of Technology, USA.
- Keeney, R.L. (1978) *Operations research. An introduction to modern applications*. Princeton, USA: Auerbach Publisher.
- Kennedy, J. F. (2008). *Estimating pedestrian volumes and crashes at urban signalized intersections*. Virginia Polytechnic Institute, USA.
- Kleijnen, J. P.C. (1974). *Statistical techniques in simulation*. New York, USA: Maral Dekker inc.
- Kleinbaum, D.G., Kupper, L.L. y Muller, K.E. (1988). *Applied regression analysis and other multivariable methods*. Belmont, California, USA: Duxbury Press.
- Kulkarni, A. N. (1972). *Analysis of the airline systems performance by varying inputs*. California Institute of Technology, USA.
- Kunkel, M. y Schwind, M. (2011). Cost and Market-based Pricing in the Courier Express and Parcel Service Industry. *Proceeding CEC '11 Proceedings of the 2011 IEEE 13th Conference on Commerce and Enterprise Computing*, 58-68.
- Kutner, M. H., Nachtsheim C. J. y Neter, J. (2004). *Applied linear regression models*. New York, USA: Mc Graw Hill.
- Kuyzu, G. (2007). *Procurement in Truckload Transportation*. Georgia Institute of Technology, USA.
- Laporte, G. (1992). The vehicle routing problem: an overview of exact and approximate algorithms. *European journal of operational research* 59, 345-358.
- Laporte, G. y Louveaux F.V. (1993). The integer L-shaped method for stochastic integer programs with complete recourse. *Operation Research Letters*, 13, no 3, 133-142.
- Lasala Calleja, P. (1994). *Introducción a la inteligencia artificial y los sistemas expertos*. Zaragoza, España: Prentice Hall.
- Law, A. M. y Kelton, W. D. (2007). *Simulation modelling and analysis*. New York, USA: Mc Graw Hill.
- Lei, H., Laporte, G. y Bo, G. (2011). The capacitated vehicle routing problem with stochastic demands and time windows. *Computers & Operations Research*, 38 no 12, 1775-1783.
- Infante, J. C. (2007). *Propuesta de rediseño de la estrategia de despacho a domicilio a regiones de una empresa de retail*. Universidad de Santiago de Chile, Chile.
- Lewis, T.G. y Smith, B.J. (1979). *Computer principles of modeling and simulation*. Boston, USA: Houghton Mifflin Company.
- Lieberman, H. (2006). *Introducción a la investigación de operaciones*. Mexico D.F., Mexico: Mc Graw Hill Interamericana.

- Limpitton, T. (2008). *Real time multi-period truckload routing problems*. Massachusetts Institute of Technology, USA.
- Madariaga, J. (2004). *Profiling Best Practices: An Explanatory Analysis of BoxPlant Trucking Logistics in the Paper Industry*. Georgia Institute of Technology, USA.
- Manivannan, M.S. (1998). *Handbook of simulation: principles, methodology, advances, applications and practice*. New Jersey, USA: John Willey & Sons.
- Marasco, A. (2008). Third-party logistics: A literature review. *International Journal of Production Economics*, 113, 127-147.
- Martinez Caro, L. y Martinez García, J.A. (2007). Measuring perceived service quality in urgent transport service. *Journal of Retailing and Customer Services*, 14, 60-72.
- Masiel, H. y Gnugnoli, G. (1972). *Simulation of discrete stochastic systems*. Chicago, USA: Henley-on-Thames Science Research Associates, Inc.
- Mc Cullagh, P. y Nelder, J. A. (1983). *Generalized Linear Models*. Londres, Reino Unido: Chapman & Hall,.
- Milton, J.S. y Arnold, J.C. (2004). *Probabilidad y estadística con aplicaciones para ingeniería y ciencias computacionales*. México D.F., México: Mc Graw Hill.
- Minsky, M. (1965). *Matter, Mind and Models*. Washington D.C., USA: Spartan Books.
- Moder, J. J. y Elmaghraby, S. E. (1978). *Handbook of operations research*. New York, USA: Van Nostrand Reinhold Compan.
- Montgomery, D., Peck, E. A., Vining, G.G. (2002). *Introducción al análisis de regresión lineal*. México D.F., México: Compañía editorial Continental.
- Moore D., Mc Cabo G. y Craig B. (2009). *Introduction to the practice of statistics*. New York, USA: Freeman and Company.
- Myers, R.H. (1990). *Classical and modern regression with applications*. Belmont, California, USA: Duxbury Press.
- Namboothiri, R. (2006). *Planning Container Drayage Operations at Congested Seaport*. Georgia Institute of Technology, USA.
- Navidi, W. (2006). *Estadística para ingenieros y científicos*. México D.F., México: Mac Graw Hill.
- Naylor, T. H. (1977). *Experimentos de simulación en computadoras con modelos de sistemas económicos*. México D.F., México: Editorial Limusa.
- Novoa, C. y Storer, R. (2009). An approximate dynamic programming approach for the vehicle routing problem with stochastic demands. *European Journal of Operational Research*, 196, no 2, 509-515.
- Nune, R. (2007). *Path prediction and path diversion identifying methodologies for hazardous materials transported by malicious entities*. Virginia Polytechnic Institute, USA.

- Örsan Özener, O. (2008). *Collaboration in transport*. Georgia Institute of Technology, USA.
- Oxford Economics (2009). *The impact of the express delivery industry on the global economy*. Oxford University, Reino Unido.
- Ozgun, C. D. (2007). *Demand management in decentralized logistics*. Georgia Institute of Technology, USA.
- Palma Méndez, J. T. y Marín Morales, R. (2008). *Inteligencia artificial*. Madrid, España: Mc Graw Hill.
- Palmgren, M. (2005). *Optimal Truck Scheduling*. Linköpings Universitet Institute of Technology, Suecia.
- Patel, J., Kapadia, C.H.; Owen, D.B. (1976). *Handbook of statistical distribution*. New York, USA: Maral Dekker inc.
- Peña, D. (2002). *Regresión y diseño de experimentos*. Madrid, España: Alianza editorial.
- Pino, R., Gómez Gómez, A. y De Abajo Martínez, N. (2001). *Introducción a la inteligencia artificial: sistemas expertos, redes neuronales artificiales y computación evolutiva*. Oviedo, España: Servicio de publicaciones Universidad de Oviedo.
- Ravula, P. P. (2007). *Design, Simulation, Analysis and Optimization of transportation system for a biomass to ethanol conversion plant*. Virginia Polytechnic Institute, USA.
- Ribeiro, R. (2004). *Integrated distribution management problems*. Barcelona, España: Universitat Pompeu Fabra.
- Ríos Insua, D., Ríos Insua, S., Martín Jiménez, J. (2004). *Simulación. Métodos y aplicaciones*. Madrid, España: RA-MA Editorial.
- Ríos Insua, S. (2004). *Investigación operativa. Modelos determinísticos y estocásticos*. Madrid, España: Editorial Centro de Estudios Ramón Areces.
- Rubenstein, R. Y. y Melamed, B. (1998). *Modern simulation and modeling*. New York, USA: John Willey & Sons.
- Seber, G.A.F. y Lee, A. (2003). *Linear regression analysis*. New Jersey, USA: John Willey & Sons.
- Snee, R.D. (1997). Validation of regression models: Methods and examples. *Technometrics*, 19, no 4, 415-428.
- Shannon, R. E. (1988). *Simulación de sistemas. Diseño, desarrollo e implantación*. México D.F., México: Prentice Hall.
- Shao, J. (2003). *Mathematical statistics*. New York, USA: Springer.
- Shapiro, S. S. y Groos, A. (1981). *Statistical modeling technique*. New York, USA: M. Dekker.

- Smith, C. A. (2005). *Simulation Model Used As Design Improvement Decision Tool for Warehouse Material Flow*. Massachusetts Institute of Technology, USA.
- Smith, W.L. (1978). *Operations research. An introduction to modern applications*. Princeton, USA: Auerbach Publisher.
- Sobol, I.M. (1976). *Lecciones populares de matemáticas. Método de Montecarlo*. Moscú, Rusia: Editorial MIR.
- Subramanian, A., Drummond, L.M.A., Bentes C., Chi L.S. y FARIAS R. (2010). A parallel heuristic for the Vehicle Routing Problem with Simultaneous Pickup and Delivery. *Computer & Operations Research*, 37, no 11, 1899-1911.
- Szandtner, T.A.B. (1978). *Operations research. An introduction to modern applications*. Princeton, USA: Auerbach Publisher.
- Taha, H. A. (2004). *Investigación de operaciones*. México D.F., México: Pearson Educación.
- Taha, H. A. (1978). *Operations research. An introduction to modern applications*. Princeton, USA: Auerbach Publisher.
- Tang Montané, F.A. y Dieguez Galvao, R. (2006). A tabu search algorithm for the vehicle routing problem with simultaneous pick-up and delivery service. *Computers and Operations Research*, 33, 595-619.
- Val Blasco, S., Pardo Rodrigo, M. J. y Larrodé Pellicer, E. (2007). Desarrollo de modelos de transporte urbano de mercancías para la mejora medioambiental y sostenibilidad de una ciudad. *GITEL*, Zaragoza, España.
- Wang, N. (2006). *Statistical Learning in logistics and manufacturing systems*. Georgia Institute of Technology, USA.
- Webster, A. L. (2000). *Applied statistics for business and economics*. New York, USA: Mac Graw Hill.
- Winston, W. L. (1994). *Investigación de operaciones*. México D.F., México: Grupo editorial Iberoamericana.
- Xuelei, N. (2006). *New Results in Detection, Estimation, and Model Selection*. Georgia Institute of Technology, USA.
- Zhang, Y. (2002). *Scalability of car-following and lane-changing models in microscopic traffic simulation systems*. B.S. Southeast University, USA.
- Zhong, J., Hall, R.W. y Dessouky, M. (2004). *Territory planning and vehicle dispatching*. University of Southern California, Los Angeles, USA.

Entrevista nº1: DIRECTOR DE TERMINAL

¿Qué es necesario en cada uno de estos campos para que cada zona de reparto/recogida funcione correctamente?

CONDUCTOR:

- El chofer de la zona debe tener una buena presencia personal. Él es una parte de la empresa y la imagen más visible de la misma de cara a los clientes, ya que casi todos los días es a él a quien ve los clientes.
- El chofer ha de tener perfil comercial. Los clientes de su zona de reparto son también sus clientes. Debe conocer sus necesidades, los problemas que tienen y tratar de solucionarlos en la medida de sus posibilidades, y en su defecto, comunicarlos a la persona responsable de la terminal lo más antes posible. En definitiva, ha de mimarlos como un si de su comercial se tratara.

VEHÍCULO

- Al igual que se le exige buena presencia personal, también se le exige a su vehículo. Es una prolongación de conductor, y por tanto de la empresa en la que trabaja.
- La capacidad del vehículo es muy importante. No se pueden quedar sin hacer recogidas por problemas de capacidad, ya que son ingresos que no vuelven.

TERMINAL Y PERSONAL DE LA TERMINAL

- Hay que conseguir que los vehículos salgan lo antes posible de la terminal al reparto, para maximizar el tiempo que están en reparto.
- La terminal debe asignar una zona de reparto y recogida que aporte unos ingresos mínimos tales como para que garanticen la continuidad del chofer en la empresa. Los clientes se acostumbran a los chóferes, y si el chofer es bueno y mimas a sus clientes, los clientes son recelosos del cambio de chofer.

ZONA DE REPARTO Y RECOGIDA

- Hay que considerar el tiempo que el vehículo tarda en llegar a la zona de reparto y/o recogida. Cuanto mayor sea la distancia, y por tanto el tiempo en llegar, menor será el tiempo efectivo que tiene para repartir y recoger mercancía.
- Al igual que la distancia a la zona de reparto, la extensión de la misma es importante. Una zona de reparto demasiado extensa repercute en la calidad del servicio de reparto y recogida.

POLITICA COMERCIAL

- La política comercial es a nivel de terminal, no por zonas de repartos. No considero que se deba hacer políticas comerciales diferenciadas por zonas de reparto, sino por tamaño de cliente, independientemente de la zona en la que se encuentre.

Puntúe de mayor a menor importancia cada una de estas variables

CHOFER	BUENA PRESENCIA PERSONAL	10
	PERFIL COMERCIAL	11
	CONOCIMIENTO MEDIOS TECNICOS	3
	MOTIVACION, REGIMEN DE TRABAJO	18
	COMUNICACIÓN CON DEPARTAMENTOS	4
	CONTINUIDAD	20
VEHICULO	LIMPIEZA	9
	CAPACIDAD	15
	PLATAFORMA	1
	MANTENIMIENTO	2
TERMINAL	HORA DE SALIDA	12
	INGRESOS	19
	PERSONAL ADMINISTRATIVO/ALMACEN	8
ZONA DE REPARTO	DISTANCIA A LA ZONA	16
	EXTENSION ZONA	17
	CLIENTES ESPECIALES	6
	NUMERO DE RECOGIDAS/REPARTOS	14
POLITICAS COMERCIALES	TIPOLOGIA DE MERCANCIA	7
	CAPTACION NUEVOS CLIENTES	5
	SOLUCION A LOS PROBLEMAS :FIDELIZACIÓN	13

Entrevista nº2: SUBDIRECTOR DE TERMINAL

¿Qué es necesario en cada uno de estos campos para que cada zona de reparto/recogida funcione correctamente?

CONDUCTOR:

- Principal variable a considerar. Se ha de asegurar su continuidad, los clientes empatizan con él y les aporta seguridad.
- El chofer ha de tener perfil comercial. Ha de tener don de gentes e instinto a la hora de poder captar nuevos clientes, o nuevos destinos de un cliente actual.
- El chofer de la zona debe tener una buena presencia personal.
- Conocedor de los medios técnicos con los que va dotado. Ir provisto de todo el material necesario.

VEHÍCULO

- Limpio y bien rotulado.
- Con la capacidad adecuada.
- Si es posible disponer de trampilla móvil.

TERMINAL Y PERSONAL DE LA TERMINAL

- Llegar a zonas de reparto antes de las 9 horas de la mañana.

ZONA DE REPARTO Y RECOGIDA

- El vehículo ha de tener un número adecuado de repartos y recogidas. Si el número es excesivo se ha de tener previsto refuerzos.
- Zonas alejadas de la terminal. Hay que considerar el tiempo de tránsito hasta ellas y de retorno de las mismas.
- Equilibrio extensión/producción de la zona. Si para que llegue a unos niveles aceptables de ingresos ha de cubrir una zona muy amplia, ante cualquier imprevisto se perderá eficacia, quedando repartos y/o recogidas sin realizar.

POLITICA COMERCIAL

- Se ha de valorar los clientes con tipología especial y/o con horarios de entrega y recogidas cerrados, instando a los comerciales a no aceptar este tipo de clientes.

Puntúe de mayor a menor importancia cada una de estas variables

CHOFER	BUENA PRESENCIA PERSONAL	9
	PERFIL COMERCIAL	10
	CONOCIMIENTO MEDIOS TECNICOS	6
	MOTIVACION, REGIMEN DE TRABAJO	14
	COMUNICACIÓN CON DEPARTAMENTOS	5
	CONTINUIDAD	20
VEHICULO	LIMPIEZA	8
	CAPACIDAD	17
	PLATAFORMA	7
	MANTENIMIENTO	4
TERMINAL	HORA DE SALIDA	13
	INGRESOS	19
	PERSONAL ADMINISTRATIVO/ALMACEN	3
ZONA DE REPARTO	DISTANCIA A LA ZONA	15
	EXTENSION ZONA	16
	CLIENTES ESPECIALES	11
	NUMERO DE RECOGIDAS/REPARTOS	18
POLITICAS COMERCIALES	TIPOLOGIA DE MERCANCIA	12
	CAPTACION NUEVOS CLIENTES	2
	SOLUCION A LOS PROBLEMAS :FIDELIZACIÓN	1

Entrevista nº3: JEFE DE TRÁFICO

¿Qué es necesario en cada uno de estos campos para que cada zona de reparto/recogida funcione correctamente?

CONDUCTOR:

El carácter del conductor sería una de las partes más importantes a tener en cuenta, ya que son las personas que están cara al público (clientes) cada día. El buen trato con el cliente, el conocer a los responsables o conocer las necesidades o complicaciones de cada uno de los clientes de su zona, son una gran baza para la gestión de la ruta.

Si le añadimos que cuiden el aspecto personal (una buena imagen), así como que mantengan en buen estado el vehículo (limpieza y reparaciones al día), y que a la vez demuestren a los clientes que tratan y se preocupan de la mercancía de manera individualizada, como si solo llevaran la mercancía de ese cliente en concreto, tendremos un aumento de calidad y evitaremos muchos de los problemas que surgen a lo largo del día, consiguiendo que los clientes se encuentren satisfechos y sumando más valor a la ruta.

VEHÍCULO

Según las características de la mercancía a repartir o recoger en la zona, el vehículo necesario para la zona puede variar, ya que puede ser necesario un tipo de vehículo pequeño (de 1500 kg. de peso máximo autorizado), o bien un vehículo carrozado, o un vehículo con lona, etc. Por lo general, es bueno tener algún vehículo plataforma para cubrir clientes puntuales o servicios más especiales.

TERMINAL Y PERSONAL DE LA TERMINAL

La zona de reparto no funcionará correctamente si detrás de los conductores el personal administrativo no gestiona las incidencias diarias de las rutas, que mejoren, busquen soluciones a los problemas, y entrelacen las necesidades de los clientes con los problemas operativos de las zonas.

ZONA DE REPARTO Y RECOGIDA

Hay que considerar la distancia a la ruta desde la terminal. Cuanto más alejada se encuentre la zona a cubrir, menor es el tiempo que se puede trabajar en la zona, con lo cual el servicio tenderá a ser de menor calidad.

Si la zona a cubrir por el vehículo (sobre todo si tiene que cubrir varias poblaciones) influirá en el sentido de que cuanto mayor sea el área a cubrir, menor será el nivel de calidad que se obtendrá, principalmente por las pérdidas de tiempo entre desplazamientos. Los horarios a cubrir suelen ser los mismos de 9.00 a 13.00 y 14.00 a 17.00 horas.

Un número excesivo repartos/recogidas influirá en la calidad de las mismas.

POLITICA COMERCIAL

Si nos están acordes con la operativa diaria de la zona, o bien con las características reales de dicha ruta, peor calidad de reparto/recogida hacia los clientes de esa zona.

Se debe tener en cuenta si la mercancía de esa ruta se trata de paquetería, palets o mercancía irregular. Todo esto influirá en el servicio que se podría realizar, y el tipo de vehículo que tendrá que operar en la zona, por ejemplo si necesitamos un vehículo con plataforma.

Puntúe de mayor a menor importancia cada una de estas variables

CHOFER	BUENA PRESENCIA PERSONAL	8
	PERFIL COMERCIAL	5
	CONOCIMIENTO MEDIOS TECNICOS	1
	MOTIVACION, REGIMEN DE TRABAJO	4
	COMUNICACIÓN CON DEPARTAMENTOS	9
	CONTINUIDAD	16
VEHICULO	LIMPIEZA	7
	CAPACIDAD	17
	PLATAFORMA	10
	MANTENIMIENTO	6
TERMINAL	HORA DE SALIDA	14
	INGRESOS	15
	PERSONAL ADMINISTRATIVO/ALMACEN	13
ZONA DE REPARTO	DISTANCIA A LA ZONA	19
	EXTENSION ZONA	20
	CLIENTES ESPECIALES	11
	NUMERO DE RECOGIDAS/REPARTOS	18
POLITICAS COMERCIALES	TIPOLOGIA DE MERCANCIA	12
	CAPTACION NUEVOS CLIENTES	3
	SOLUCION A LOS PROBLEMAS :FIDELIZACIÓN	2

Entrevista nº4: RESPONSABLE DE RECOGIDAS

¿Qué es necesario en cada uno de estos campos para que cada zona de reparto/recogida funcione correctamente?

CONDUCTOR:

- El no cambiar un chofer en la zona de reparto/recogida posibilita que el número total repartos y recogidas a realizar por día aumente considerablemente. Por otro lado, un amplio conocimiento de su zona incrementa la capacidad de reacción ante cualquier imprevisto.
- El chofer ha de tener una buena relación con los departamentos de la terminal, así como comunicación fluida.

VEHÍCULO

- Capacidades máximas de carga.
- Buen estado y conservación del vehículo.

TERMINAL Y PERSONAL DE LA TERMINAL

- Que la mercancía llegue a la terminal en los horarios establecidos. Si es posible adelantarlos, mejor.
- Que la descarga y ubicación de las mercancías se haga en el menor tiempo posible para que el horario de salida de los choferes de la terminal no se demore. Prioridad en la salida de los choferes que tengan zonas de reparto más alejadas.

ZONA DE REPARTO Y RECOGIDA

- Valorar el número de repartos/recogidas diarios.
- Realizar un estudio previo de cada zona de reparto en cuanto a poblaciones a visitar, extensión de la zona de influencia, distancia y tiempos de tránsito. Conocer si la zona principalmente es industrial, o por el contrario, más comercial. Analizar los costes de cada zona de manera individualizada.

POLITICA COMERCIAL

- Estudio previo de los clientes en cada zona y de la tipología de mercancía de cada uno.

Puntúe de mayor a menor importancia cada una de estas variables

CHOFER	BUENA PRESENCIA PERSONAL	7
	PERFIL COMERCIAL	1
	CONOCIMIENTO MEDIOS TECNICOS	2
	MOTIVACION, REGIMEN DE TRABAJO	18
	COMUNICACIÓN CON DEPARTAMENTOS	11
	CONTINUIDAD	19
VEHICULO	LIMPIEZA	6
	CAPACIDAD	14
	PLATAFORMA	5
	MANTENIMIENTO	10
TERMINAL	HORA DE SALIDA	13
	INGRESOS	20
	PERSONAL ADMINISTRATIVO/ALMACEN	12
ZONA DE REPARTO	DISTANCIA A LA ZONA	17
	EXTENSION ZONA	16
	CLIENTES ESPECIALES	8
	NUMERO DE RECOGIDAS/REPARTOS	15
POLITICAS COMERCIALES	TIPOLOGIA DE MERCANCIA	9
	CAPTACION NUEVOS CLIENTES	3
	SOLUCION A LOS PROBLEMAS :FIDELIZACIÓN	4

Entrevista nº5: CONDUCTOR nº1

¿Qué es necesario en cada uno de estos campos para que cada zona de reparto/recogida funcione correctamente?

CONDUCTOR:

El contacto diario con el cliente es mediante el autónomo, el trato personal y profesional con el cliente debe ser correcto, dar una buena imagen y prestar un buen servicio contribuye a la continuidad del mismo.

VEHÍCULO

La capacidad del vehículo ha de ser adecuada a la mercancía a recoger.

TERMINAL Y PERSONAL DE LA TERMINAL

Una buena comunicación entre todos los departamentos es esencial: la oficina administrativa para la información de cualquier incidencia ante el cliente y el departamento comercial para el seguimiento de envíos y entregas en destino.

El personal de almacén realice cuanto antes la descarga de los vehículos de ruta, para agilizar la salida de los vehículos a reparto.

ZONA DE REPARTO Y RECOGIDA

El número de vehículos por población ha de ser acorde al número de repartos y recogidas a realizar.

POLITICA COMERCIAL

La captación y fidelización de los clientes en cada una de las zonas, para asegurar unos ingresos mínimos para los conductores.

Puntúe de mayor a menor importancia cada una de estas variables

CHOFER	BUENA PRESENCIA PERSONAL	10
	PERFIL COMERCIAL	9
	CONOCIMIENTO MEDIOS TECNICOS	3
	MOTIVACION, REGIMEN DE TRABAJO	19
	COMUNICACIÓN CON DEPARTAMENTOS	11
	CONTINUIDAD	18
VEHICULO	LIMPIEZA	2
	CAPACIDAD	14
	PLATAFORMA	1
	MANTENIMIENTO	5
TERMINAL	HORA DE SALIDA	8
	INGRESOS	20
	PERSONAL ADMINISTRATIVO/ALMACEN	12
ZONA DE REPARTO	DISTANCIA A LA ZONA	7
	EXTENSION ZONA	13
	CLIENTES ESPECIALES	4
	NUMERO DE RECOGIDAS/REPARTOS	15
POLITICAS COMERCIALES	TIPOLOGIA DE MERCANCIA	6
	CAPTACION NUEVOS CLIENTES	17
	SOLUCION A LOS PROBLEMAS :FIDELIZACIÓN	16

Entrevista nº6: CONDUCTOR nº2

¿Qué es necesario en cada uno de estos campos para que cada zona de reparto/recogida funcione correctamente?

CONDUCTOR:

Un trato adecuado con el cliente es fundamental, adaptándose a las necesidades del mismo. Ser diestro a la hora de realizar cuanto más repartos y recogidas mejor.

VEHÍCULO

Un mantenimiento del vehículo ha de ser adecuado para evitar averías inoportunas para y no dejar repartos y o recogidas sin realizar.

TERMINAL Y PERSONAL DE LA TERMINAL

Que el departamento administrativo colabore los más posible en agilizar todos los repartos y/o recogidas.

Disponer cuanto antes la mercancía para realizar la carga a los vehículos de reparto y abandonar cuanto antes la terminal

ZONA DE REPARTO Y RECOGIDA

El número de poblaciones a visitar y la distancia entre ellos ha de ser racional para poder satisfacer de manera adecuada, y con la calidad requerida, todos los repartos y recogidas asignados

POLITICA COMERCIAL

Garantizar la continuidad de los clientes para asegurar unos ingresos mínimos para los conductores.

Puntúe de mayor a menor importancia cada una de estas variables

CHOFER	BUENA PRESENCIA PERSONAL	5
	PERFIL COMERCIAL	4
	CONOCIMIENTO MEDIOS TECNICOS	3
	MOTIVACION, REGIMEN DE TRABAJO	11
	COMUNICACIÓN CON DEPARTAMENTOS	8
	CONTINUIDAD	15
VEHICULO	LIMPIEZA	2
	CAPACIDAD	6
	PLATAFORMA	1
	MANTENIMIENTO	12
TERMINAL	HORA DE SALIDA	17
	INGRESOS	19
	PERSONAL ADMINISTRATIVO/ALMACEN	13
ZONA DE REPARTO	DISTANCIA A LA ZONA	18
	EXTENSION ZONA	20
	CLIENTES ESPECIALES	9
	NUMERO DE RECOGIDAS/REPARTOS	10
POLITICAS COMERCIALES	TIPOLOGIA DE MERCANCIA	7
	CAPTACION NUEVOS CLIENTES	14
	SOLUCION A LOS PROBLEMAS :FIDELIZACIÓN	16