CAPÍTULO 2:

ALGUNOS MODELOS DE REGRESIÓN BORROSA

2.1. CONSIDERACIONES PREVIAS

A continuación analizaremos algunos modelos de regresión borrosa que, como toda técnica de regresión, su objetivo es determinar una relación funcional entre una variable dependiente con varias variables explicativas y la podemos considerar alternativa a la más conocida de mínimos cuadrados, aunque ésta última es una técnica estadística. Asimismo, observaremos que es más versátil que los convencionales modelos de regresión estadísticos, ya que permite hallar relaciones funcionales cuando la variable dependiente, las variables independientes o ambas no se manifiestan como valores únicos, sino como intervalos de confianza.

A grosso modo, la técnica de regresión de mínimos cuadrados parte de que la relación existente entre un variable explicada, y las explicativas, que es un vector m-dimensional $x=(x_1, x_2,\ldots,x_m)$, viene dada por un hiperplano, es decir:

$$y= a_0 + a_1x_1 + a_2x_2 +\ldots+ a_mx_m$$

Así, para un conjunto de n observaciones –una muestra- $\{(Y_1,X_1), (Y_2,X_2),\ldots,(Y_j,X_j),\ldots,(Y_n,X_n)\}$:

donde:

Y_j: Es la observación número j de la variable dependiente y, j=1,2,…,n.

X_j: Es la observación número j con j=1,2,…,n, de la variable independiente x. X_j es una variable m-dimensional $X_j = (X_{0j}, X_{1j}, X_{2j},..., X_{ij},..., X_{mj})$, donde $X_{0j} = 1$ $\forall j$, y X_{ij} es el valor observado sobre la variable i-ésima en el j-ésimo componente de la muestra.

dado que difícilmente podremos ajustar un hiperplano que sea compatible con todos los elementos de la muestra, se supone que la relación entre Y_j y el vector X_j es lineal $\forall j$, pero que existe un elemento de perturbación estocástico, que se comporta como una variable aleatoria de media 0, de forma que para la j-ésima observación se cumple:

$$Y_j= a_0 + a_1X_{1j} + a_2X_{2j} +\ldots+ a_mX_{mj} + \varepsilon_j$$

siendo ε_j el término de perturbación estocástico. Si, como se supone habitualmente, la varianza es igual para los errores y que éstos están incorrelacionados, $E[\varepsilon_k \cdot \varepsilon_j]=0\ \forall k \neq j$, el método de estimación utilizado es el conocido de mínimos cuadrados ordinarios. En este caso, para estimar los parámetros a_i, $i=1,1,2,\ldots,m$, se resuelve un programa cuadrático cuya función objetivo es: $z = \sum_{j=1}^{n}\left(Y_j - \hat{Y}_j\right)^2$, donde $\hat{Y}_j$ es el valor estimado de Y_j como función de las variables X_{1j}, X_{2j},..X_{mj}, tras haber hallado los valores de $a_1, a_2,\ldots,a_m$.

Utilizando regresión borrosa también asumimos que la relación entre la variable explicada y las explicativas es lineal, pero en este caso, si disponemos de una muestra $\{(Y_1,X_1), (Y_2,X_2),\ldots,(Y_j,X_j),\ldots,(Y_n,X_n)\}$, las posibles divergencias que pudieran surgir entre la j-ésima observación de la variable independiente Y_j y su estimación $\hat{Y}_j$, viene dada porque entre x e y existe una relación borrosa del tipo:

$$y = \widetilde{A}_0 + \widetilde{A}_1 x_1 + \widetilde{A}_2 x_2 + \ldots + \widetilde{A}_m x_m$$

donde los coeficientes $\widetilde{A}_0$, $\widetilde{A}_1,\ldots,\widetilde{A}_m$, son números borrosos, por lo que el i-ésimo queda caracterizado por:

$$\widetilde{A}_i = \left\{x, \mu_{\widetilde{A}_i}(x)\right\} = \left\{A_{i\alpha} = \left[A_i^1(\alpha), A_i^2(\alpha)\right] \mid 0 \leq \alpha \leq 1\right\}$$

Es decir, las divergencias que se producen respecto a la teórica relación lineal no tienen naturaleza aleatoria, sino borrosa. Asimismo, podemos comprobar que el término de error no queda introducido como sumando en el hiperplano, sino que es incorporado en los coeficientes $\widetilde{A}_i$, $i=0,2,\ldots,m$, al asumirse que son números borrosos. Por supuesto, de forma análoga a la técnica de mínimos cuadrados, una vez se disponga de la muestra, nuestro objetivo debe ser ajustar los coeficientes $\widetilde{A}_i$.

Respecto a esta forma de modelización, creemos que de alguna forma, ofrece ciertas ventajas sobre la tradicional técnica de regresión. En primer lugar, porque las estimaciones que obtengamos después de ajustar los coeficientes borroso, no serán variables aleatorias, y por tanto, en muchas ocasiones de difícil tratamiento numérico, sino números borrosos, cuyo tratamiento es más sencillo. Por otra parte, si el fenómeno de estudio es de carácter económico o social, las observaciones que del mismo se obtienen son consecuencia de la interacción entre las creencias, expectativas, etc. de los agentes que participan en dicho fenómeno, y por tanto, ya hemos señalado que en nuestra opinión, no es del todo adecuado modelizar dicho fenómeno utilizando la teoría de la probabilidad. Por ejemplo, el precio de los activos que se negocian en los mercados financieros es la consecuencia de las expectativas que tienen los participantes sobre el devenir de la economía, la confianza que a los operadores les generan los emisores de dichos activos etc. Posiblemente en este caso sea excesivamente simplificadora la existencia de linealidad entre la variable explicada y las variables explicativas lo cual se asume utilizando tanto la regresión convencional como la regresión borrosa, pero creemos que es más realista modelizar el sesgo que puede darse entre las realizaciones de la variable dependiente y el valor que teóricamente éstas pueden tomar asumiendo que la relación entre variable dependiente y variables explicativas es borrosa, que si damos una naturaleza aleatoria a dicho sesgo. Respecto a los precios de los activos financieros, estaremos asumiendo, como mínimo, el fuerte componente subjetivo que implica su determinación.

Por otra parte, en muchas circunstancias las observaciones de la variable dependiente, de la variable independiente o de ambas no vienen dadas por un número cierto, sino por un intervalo de confianza. Por ejemplo, el precio que se negocia en los mercados financieros durante una sesión para un determinado activo difícilmente es único, sino que éste suele negociarse dentro de una horquilla delimitada por un precio máximo y por un precio mínimo. Para utilizar las técnicas de mínimos cuadrados –o la más sofisticada de máximo verosímilitud- deben cuantificarse las observaciones de la variable explicada (y explicativa) a través de un único número, utilizándose por ejemplo, el precio medio negociado o el último precio en el modelo que se vaya a implementar. Es evidente que este proceder implica una importante pérdida de información. Para implementar los métodos de regresión borrosa no hace falta reducir el valor de las variables observadas a un número real, cuando son observadas como intervalos, así, podremos ajustar la relación funcional que busquemos trabajando con todos los valores observados siendo posible entonces utilizar toda la información disponible.

A continuación analizamos los dos modelos que habitualmente son utilizados en aplicaciones empíricas, propuestos por Tanaka e Ishibuchi (1992) y Tanaka (1987) y Sakawa y Yano (1992) respectivamente.

2.2. EL MODELO DE TANAKA E ISHIBUCHI DE REGRESIÓN BORROSA

2.2.1. Regresión con intervalos de confianza

Estos autores parten de que para un determinado fenómeno, el observador dispone de una muestra que representamos como: $\{(Y_1,X_1), (Y_2,X_2),\ldots,(Y_j,X_j),\ldots,(Y_n,X_n)\}$, donde:

a) Y_j es la observación j-ésima de la variable dependiente, $j=1,2,\ldots,n$, y suponemos que viene dada a través de un intervalo de confianza:

$$Y_j = \left[Y_j^1, Y_j^2\right]$$

así, podemos representar el intervalo de confianza Y_j a través de su centro y de su radio como:

$$Y_j = < Y_{jC}, Y_{jR} >$$

donde:

$$Y_{jC} = \frac{Y_j^1 + Y_j^2}{2} \text{ y } Y_{jR} = \frac{Y_j^2 - Y_j^1}{2}$$

si Y_j, $j=1,2,\ldots,n$ son números ciertos, $Y_{jR} = 0 \; \forall j$.

b) X_j: es el vector observado en la j-ésima observación sobre las variables independientes, con $j=1,2,\ldots,n$. Así, X_j es una variable m-dimensional $X_j = \left(X_{0j}, X_{1j}, X_{2j}, \ldots, X_{ij}, \ldots, X_{mj}\right)$, donde $X_{0j} = 1 \; \forall j$, y X_{ij} es el valor en la j-ésima observación de la muestra para la variable i-ésima. En cualquier caso asumimos que se tratan de observaciones crisp.

En el modelo de Tanaka se asume que la relación existente entre la variable dependiente, que es un intervalo de confianza, y la variable independiente $x = (x_1, x_2,\ldots,x_m)$, que es un vector con componentes ciertos es lineal, de forma que:

$$y = A_0 + A_1x_1 + A_2x_2 + \ldots + A_mx_m$$

donde A_i, $i = 0,1,\ldots,m$ son intervalos de confianza:

$$A_i = <a_{iC}, a_{iR}>, \; i=0,1,\ldots,m.$$

Evidentemente, el objetivo final es determinar los centros y radios de A_i de forma que sean compatibles con las observaciones de que se dispone. De esta forma, si denominamos como $\hat{Y}_j =$

$\left\langle \hat{Y}_{jC}, \hat{Y}_{jR} \right\rangle$ al intervalo de confianza correspondiente al valor que estimamos para la j-ésima variable independiente, Y_j, después de hallar los parámetros $A_1, A_2, \ldots, A_m$, como dicha estimación se realizaría mediante la suma:

$$\hat{Y}_j = \sum_{i=0}^{m} A_i X_{ij} \text{ , j=1,2,}\ldots\text{,n.}$$

podemos expresarla a través de sus centros y sus radios, que serán funciones de los centros y de los radios de los parámetros A_i, i=1,2,...,m:

$$\left\langle \hat{Y}_{jC}, \hat{Y}_{jR} \right\rangle = \sum_{i=0}^{m} \left\langle a_{iC}, a_{iR} \right\rangle X_{ij} = \left\langle \sum_{i=0}^{m} a_{iC} X_{ij}, \sum_{i=0}^{m} a_{iR} \left| X_{ij} \right| \right\rangle$$

En Tanaka e Ishibuchi (1992) se propone dos modelos de regresión borrosa. El primero de ellos está asociado al concepto de posibilidad. El objetivo buscado en este modelo es que las estimaciones que obtengamos de los valores de la variable independiente observados tras ajustar los parámetros $A_1, A_2, \ldots, A_m$ contengan a Y_j $\forall$j, es decir, que $\hat{Y}_j$ sea posiblemente Y_j, minimizando la incertidumbre de las estimaciones realizadas sobre Y_j, $\hat{Y}_j$ tras ajustar los parámetros A_i. Asimismo, dicha incertidumbre puede se medida por la amplitud de dicho intervalo de confianza, $\hat{Y}_j$, más concretamente, por su radio, y así, como el radio de la estimación de la j-ésima variable independiente es:

$$\hat{Y}_{jR} = \sum_{i=0}^{m} a_{iR} \left| X_{ij} \right|$$

la incertidumbre total que genera la muestra tras ajustarse los parámetros A_i, i=0,2,…,m es la suma de los radios de las estimaciones que realicemos sobre las variables dependientes observadas, es decir:

$$z = \sum_{j=1}^{n} \hat{Y}_{jR} = \sum_{j=1}^{n} \sum_{i=0}^{m} a_{iR} \left| X_{ij} \right|$$

Así, como buscamos minimizar la función z de forma que las estimaciones $\hat{Y}_j$ contengan las variables independientes observadas, Y_j, es decir: $Y_j \subseteq \hat{Y}_j$ $\forall$j, para la determinación de los parámetros A_i deberemos resolver el siguiente programa lineal, siendo las variables decisión del programa los centros y los radios de los parámetros A_i:

$$\text{Minimizar } z= \sum_{j=1}^{n} \hat{Y}_{jR} = \sum_{j=1}^{n}\sum_{i=0}^{m} a_{iR} \left|X_{ij}\right|$$

sujeto a:

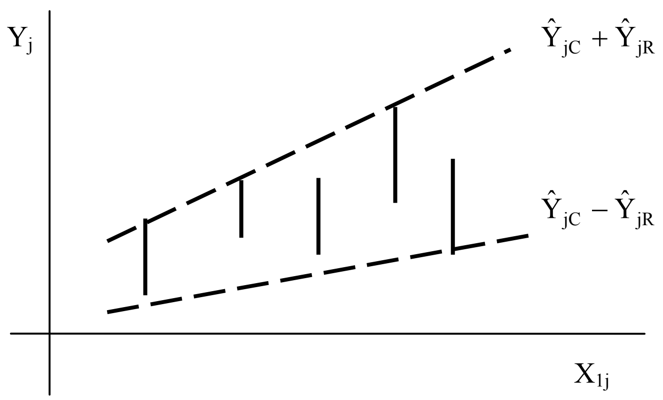
$$\hat{Y}_{jC} - \hat{Y}_{jR} = \sum_{i=0}^{m} a_{iC} X_{ij} - \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right| \leq Y_{jC} - Y_{jR} \qquad j=1,2,...,n$$

$$\hat{Y}_{jC} + \hat{Y}_{jR} = \sum_{i=0}^{m} a_{iC} X_{ij} + \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right| \geq Y_{jC} + Y_{jR} \qquad j=1,2,...,n$$

$$a_{iR} \geq 0 \qquad i=0,1,...,m$$

Los dos primeros bloques de restricciones aseguran que las estimaciones sobre la variable dependiente realizadas contengan a las realmente observadas. Así, con el primero aseguramos que el extremo inferior de las estimaciones sean inferiores a los extremos inferiores de las observaciones mientras que con el segundo bloque de restricciones indicamos que los extremos superiores de las estimaciones deben ser superiores a los de las observaciones. El tercer bloque de restricciones se plantea porque el radio de un intervalo de confianza no puede ser negativo.

Representamos a continuación el modelo de regresión borroso basado en el concepto de posibilidad, para un modelo del tipo $Y_j = A_0 + A_1 X_{1j}$.



El segundo modelo de regresión borrosa propuesto en Tanaka e Ishibuchi (1992), está asociado al concepto de necesidad. En este caso se exigen que las estimaciones que se obtienen de las variables independientes $\hat{Y}_j^N$, estén contenidas en las realmente observadas, Y_j, j=1,2,…,n, es decir, $\hat{Y}_j^N \subseteq Y_j$, lo cual puede interpretarse como que $\hat{Y}_j^N$ es necesario en Y_j.

Si notamos a los coeficientes a estimar como $A_i^N = \left\langle a_{iC}^N, a_{iR}^N \right\rangle$, i=0,1,…,m., la incertidumbre de las estimaciones realizadas sobre las observaciones de las variables independientes es la suma de los radios de las estimaciones $\hat{Y}_j^N$, j=1,2,...,n, es decir:

$$z^N = \sum_{j=1}^{n} \hat{Y}_{jR}^N = \sum_{j=1}^{n}\sum_{i=0}^{m} a_{iR}^N \left|X_{ij}\right|$$

En este modelo los autores proponen maximizar la incertidumbre de las estimaciones teniéndose en cuenta que éstas deben estar contenidas en los intervalos de confianza realmente observados para dichas variables dependientes a_{iC}^N y a_{iR}^N, i=0,1,2,…,m. Para determinar los centros y los radios de los coeficientes A_i, debemos, pues, plantear el programa lineal:

$$\text{Maximizar } z^N = \sum_{j=1}^{n} \hat{Y}_{jR}^N = \sum_{j=1}^{n}\sum_{i=0}^{m} a_{iR}^N \left|X_{ij}\right|$$

sujeto a:

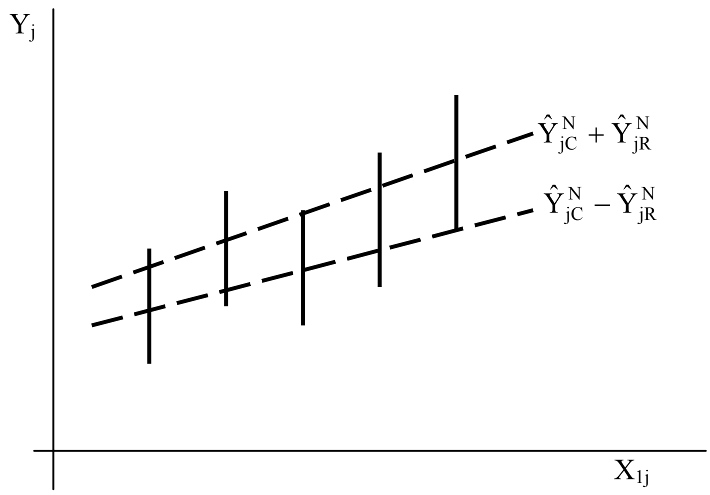
$$\hat{Y}_{jC}^N - \hat{Y}_{jR}^N = \sum_{i=0}^{m} a_{iC}^N X_{ij} - \sum_{i=0}^{m} a_{iR}^N \left|X_{ij}\right| \geq Y_{jC} - Y_{jR} \qquad j=1,2,...,n$$

$$\hat{Y}_{jC}^N + \hat{Y}_{jR}^N = \sum_{i=0}^{m} a_{iC}^N X_{ij} + \sum_{i=0}^{m} a_{iR}^N \left|X_{ij}\right| \leq Y_{jC} + Y_{jR} \qquad j=1,2,...,n$$

$$a_{iR}^N \geq 0 \qquad i=0,1,...,m$$

Los dos primeros bloques de restricciones aseguran que las estimaciones sobre la variable dependiente realizadas estén contenidas en las realmente observadas. Con el primero de los dos conseguimos que el extremo inferior de las estimaciones sea superior al extremo inferior de las observaciones; y con el segundo exigimos que los extremos superiores de los intervalos de confianza correspondientes a las estimaciones sean inferiores a los de las observaciones. El tercer bloque de restricciones asegura que el radio de los parámetros A_0, A_1,...,A_n no sea negativo.

Representamos a continuación, gráficamente y suponiendo una única variable independiente, el modelo de regresión borrosa asociado al concepto de necesidad propuesto por Tanaka:



Tanaka (1987) demuestra que el programa lineal asociado al modelo de regresión posibilística siempre tiene solución, mientras que el planteado para la regresión borrosa construida a partir del concepto de necesidad no puede asegurarse su existencia. Es fácil intuir que éste último, por ejemplo, no tendrá solución si las observaciones de la variable independiente Y_j. j=1,2,…n son observaciones ciertas y la muestra no proviene de un hiperplano con coeficientes ciertos. Por esta razón, en aplicaciones prácticas, se suele únicamente utilizar el modelo de regresión posibilística, y no el asociado al concepto de necesidad, pues para éste, en muchas ocasiones, no existe solución.

2.2.2. Regresión con números borrosos

En este caso partiremos de que, para un determinado fenómeno, el observador dispone de una muestra que representamos como: $\{(Y_1,X_1), (Y_2,X_2),\ldots,(Y_j,X_j),\ldots,(Y_n,X_n)\}$, donde las variables independientes son vectores crisp m+1-dimensionales y las n observaciones de la variable dependiente son intervalos de confianza. La j-ésima observación es pues un intervalo de confianza que puede representarse a través de su centro y su radio como $< Y'_{jC}, Y'_{jR} >$.

Asimismo suponemos que la j-ésima observación de la variable dependiente es un α^*-corte arbitrario del número borroso del que proviene donde α^* viene prefijada por el decisor de forma arbitraria. El número borroso que cuantifica el valor de la j-ésima observación sobre la variable dependiente es, asimismo, un número borroso L-L de Dubois y Prade simétrico que notamos como $\tilde{Y}_j = (Y_{jC}, Y_{jR})_L$. La función forma de $\tilde{Y}'_j$, L(x), al igual que el nivel α^* observado viene elegida por el decisor, y su forma dependerá del problema que tratemos.

De esta forma, como el α^*-corte de $\tilde{Y}_j$, $Y_{j\alpha^*}$ es:

$$Y_{j\alpha^*} = \left[Y_{jC} - Y_{jR} L^{-1}\left(\alpha^*\right), Y_{jC} + Y_{jR} L^{-1}\left(\alpha^*\right)\right], \text{ j=1,2,…,n.}$$

podemos inferir el centro y el radio de $\tilde{Y}_j$ a través de su α-corte observado, que lo es a un nivel α^*, es decir, es $Y_{j\alpha^*}$. En concreto:

$$Y'_{jC} = Y_{jC}$$

$$Y'_{jR} = Y_{jR} L^{-1}\left(\alpha^*\right) \Rightarrow Y_{jR} = \frac{Y'_{jR}}{L^{-1}\left(\alpha^*\right)}$$

Finalmente, el modelo de regresión a estimar será:

$$\tilde{y} = \tilde{A}_0 + \tilde{A}_1 x_1 + ... + \tilde{A}_m x_m$$

Respecto a los parámetros $\tilde{A}_i$, i=0,1,2,…,m serán números borrosos L-L simétricos de Dubois y Prade, siendo la función forma L(x) a tomar idéntica a la elegida para las observaciones de la variable explicada. Así, los parámetros A_i a estimar pueden ser expresados como $\tilde{A}_i = (a_{iC}, a_{iR})_L$, i=0,1,…,m.

Como siempre, el objetivo final es estimar el centro y el radio de $\tilde{A}_i$ $\forall$i, siendo para ello necesario tener en cuenta que tras estimar los centros y los radios de los parámetros $\tilde{A}_i$, obtendremos unas estimaciones del valor de las observaciones de las variables independientes $\hat{\tilde{Y}}_j = \left(\hat{Y}_{jC}, \hat{Y}_{jR}\right)_L$. Más concretamente, como:

$$\hat{\tilde{Y}}_j = \sum_{i=0}^{m} \tilde{A}_i X_{ij}$$

los centros y los radios de $\hat{\tilde{Y}}_j$ vendrán dados por:

$$\hat{\tilde{Y}}_j = \left(\hat{Y}_{jC}, \hat{Y}_{jR}\right)_L = \sum_{i=0}^{m} \left(a_{iC}, a_{iR}\right)_L X_{ij} = \left(\sum_{i=0}^{m} a_{iC} X_{ij}, \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right|\right)_L$$

apuntamos que debe ser tenido en cuenta que $X_{0j} = 1$ $\forall$j.

Así, como $\hat{\tilde{Y}}_j$ es un número borroso de Dubois y Prade simétrico, su función de pertenencia será:

$$\mu_{\hat{\hat{Y}}_j}(x)=\begin{cases} L\left(\dfrac{\left|\sum_{i=0}^{m} a_{iC}X_{ij}-x\right|}{\sum_{i=0}^{m} a_{iR}\left|X_{ij}\right|}\right) & \sum_{i=0}^{m} a_{iC}X_{ij}-\sum_{i=0}^{m} a_{iR}\left|X_{ij}\right|\le x\le \sum_{i=0}^{m} a_{iC}X_{ij}+\sum_{i=0}^{m} a_{iR}\left|X_{ij}\right| \\ 0 & \text{en otro caso} \end{cases}$$

La determinación de los parámetros a_{iC} y a_{iR}, i=0,1,…,m se realizará a través del modelo de regresión posibilística de intervalos de confianza planteado anteriormente. Apuntamos que en este caso los intervalos de confianza que entran en juego, son los α^*-cortes de $\hat{\tilde{Y}}_j$, j=1,2,…n que es desconocido, ya que desconocemos los de $\tilde{A}_i$, i=0,1,…,m y los α^*-cortes de $\tilde{Y}_j$, j=1,2,…,n. Respecto al α^*-corte de $\hat{\tilde{Y}}_j$, $\hat{Y}_{j\alpha^*}=\left\langle \hat{Y}'_{jC}, \hat{Y}'_{jR}\right\rangle$, podemos observar que este se halla a través de los α-cortes a dicho nivel α^* de los parámetros $\tilde{A}_i$, los cuales notamos como $A_{i\alpha^*}=\left\langle a'_{iC}, a'_{iR}\right\rangle$, i=0,1,…,m. De esta forma, $\hat{Y}_{j\alpha^*}$ se obtiene como:

$$\left\langle \hat{Y}'_{jC}, \hat{Y}'_{jR}\right\rangle=\sum_{i=0}^{m}\left\langle a'_{iC}, a'_{iR}\right\rangle X_{ij}=\left\langle \sum_{i=0}^{m} a'_{iC}X_{ij}, \sum_{i=0}^{m} a'_{iR}\left|X_{ij}\right|\right\rangle$$

Entonces, para hallar a'_{iC} y a'_{iR}, i=0,1,…,m el programa lineal que plantearemos es el siguiente:

Minimizar $z=\sum_{j=1}^{n}\sum_{i=0}^{m} a'_{iR}\left|X_{ij}\right|$

sujeto a:

$$\hat{Y}'_{jC}-\hat{Y}'_{jR}=\sum_{i=0}^{m} a'_{iC}X_{ij}-\sum_{i=0}^{m} a'_{iR}\left|X_{ij}\right|\le Y'_{jC}-Y'_{jR} \qquad j=1,2,...,n$$

$$\hat{Y}'_{jC}+\hat{Y}'_{jR}=\sum_{i=0}^{m} a'_{iC}X_{ij}+\sum_{i=0}^{m} a'_{iR}\left|X_{ij}\right|\ge Y'_{jC}+Y'_{jR} \qquad j=1,2,...,n$$

$$a'_{iR}\ge 0 \qquad i=0,1,...,m$$

Finalmente, ya hallados los centros y radios del α^*-corte de los coeficientes $\tilde{A}_i$, a'_{iC} y a'_{iR}, debemos calcular el centro y el radio de $\tilde{A}_i=(a_{iC}, a_{iR})_L$ $\forall$i. Así, teniendo en cuenta que:

$$A_{i\alpha^*}=\left[a_{iC}-a_{iR}L^{-1}\left(\alpha^*\right), a_{iC}+a_{iR}L^{-1}\left(\alpha^*\right)\right]$$

podemos determinar a_{iC} y a_{iR}, i=0,1,…,m como:

$$a'_{iC} = a_{iC}$$

$$a'_{iR} = a_{iR} L^{-1}(\alpha^*) \Rightarrow a_{iR} = \frac{a'_{iR}}{L^{-1}(\alpha^*)}$$

2.3. MODELO DE SAKAWA Y YANO

2.3.1. Regresión posibilística con inputs ciertos y outputs borrosos

Como en el modelo de Tanaka, Sakawa y Yano (1992) suponen que el conjunto de observaciones sobre las variables dependientes e independientes $\{(Y_1,X_1), (Y_2,X_2),\ldots,(Y_j,X_j),\ldots,(Y_n,X_n)\}$, las variables independientes son vectores crisp m-dimensionales y las n observaciones de la variable independiente son intervalos de confianza. De esta forma, la j-ésima observación de la variable dependiente Y_j queda cuantificada por su centro y su radio como $< Y'_{jC}, Y'_{jR} >$ con $j=1,2,\ldots,n$.

Asimismo suponemos que la j-ésima observación de la variable dependiente es un α^*-corte arbitrario de un número borroso L-L de Dubois y Prade simétrico, siendo dicho número borroso denotado por $\tilde{Y}_j = (Y_{jC}, Y_{jR})_L$. Así, si determinamos, la función forma de $\tilde{Y}_j$, a través de su α^*-corte observado $Y_{j\alpha^*} = < Y'_{jC}, Y'_{jR} >$ podemos hallar el centro y el radio del número borroso del que proviene de forma análoga a lo comentado en el modelo de regresión anterior para outputs borrosos. En concreto:

$$Y'_{jC} = Y_{jC}$$

$$Y'_{jR} = Y_{jR} L^{-1}(\alpha^*) \Rightarrow Y_{jR} = \frac{Y'_{jR}}{L^{-1}(\alpha^*)}$$

Por otra parte, y como siempre, el modelo de regresión a estimar es:

$$\tilde{y} = \tilde{A}_0 + \tilde{A}_1 x_1 + \ldots + \tilde{A}_m x_m$$

donde los parámetros $\tilde{A}_i$, $i=0,1,2,\ldots,m$ serán números borrosos L-L simétricos de Dubois y Prade, y por tanto, estos vienen representados por $\tilde{A}_i = (a_{iC}, a_{iR})_L$, $i=0,1,\ldots,m$. Asimismo, el objetivo final es la determinación del centro y el radio de $\tilde{A}_i$ $\forall i$. Tras estimar dichos parámetros, obtendremos unas estimaciones del valor de las observaciones de las variables independientes $\hat{\tilde{Y}}_j = (\hat{Y}_{jC}, \hat{Y}_{jR})_L$ donde:

$$\hat{\tilde{Y}}_j = \left(\hat{Y}_{jC}, \hat{Y}_{jR}\right)_L = \sum_{i=0}^{m} \left(a_{iC}, a_{iR}\right)_L X_{ij} = \left(\sum_{i=0}^{m} a_{iC} X_{ij}, \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right|\right)_L$$

Sakawa y Yano, para medir la congruencia de la relación lineal borrosa de las variables independientes con la explicada respecto a las observaciones obtenidas, utilizan el concepto de posibilidad de que dos números borrosos sean iguales. Según dicha medida, para dos números borrosos $\tilde{A}$ y $\tilde{B}$, donde:

$$\tilde{A} = \left\{x, \mu_{\tilde{A}}(x)\right\} = \left\{A_\alpha = \left[A^1(\alpha), A^2(\alpha)\right] \mid 0 \le \alpha \le 1\right\}$$

$$\tilde{B} = \left\{x, \mu_{\tilde{B}}(x)\right\} = \left\{B_\alpha = \left[B^1(\alpha), B^2(\alpha)\right] \mid 0 \le \alpha \le 1\right\}$$

la posibilidad de que $\tilde{A} = \tilde{B}$ es igual o mayor que α, $\mathrm{Poss}\left(\tilde{A} = \tilde{B}\right) \ge \alpha$, si se cumple que:

$A^2(\alpha) \ge B^1(\alpha)$ y $A^1(\alpha) \le B^2(\alpha)$

De esta forma, diremos que la posibilidad de que la j-ésima observación de la variable dependiente, $\tilde{Y}_j$ es igual en un grado igual o superior a α a su estimación $\hat{\tilde{Y}}_j$, si se cumple:

$$\sum_{i=0}^{m} a_{iC} X_{ij} - \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right| L^{-1}(\alpha) \le Y_{jC} + Y_{jR} L^{-1}(\alpha)$$

$$\sum_{i=0}^{m} a_{iC} X_{ij} + \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right| L^{-1}(\alpha) \ge Y_{jC} - Y_{jR} L^{-1}(\alpha)$$

Finalmente, el programa matemático que proponen Sakawa y Yano para determinar los parámetros a_{iC} y a_{iR}, consiste en resolver un programa donde se minimice la incertidumbre de las estimaciones de $\tilde{Y}_j$, $\hat{\tilde{Y}}_j$, para todas la observaciones realizadas y se maximice conjuntamente el grado de cumplimiento de $\tilde{Y}_j = \hat{\tilde{Y}}_j$ $\forall$j. Entonces, para hallar los centros y los radios de los coeficientes $\tilde{A}_i$, i=0,1,...,m debemos resolver el programa multiobjetivo no lineal:

Minimizar z= $\displaystyle\sum_{j=1}^{n} \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right|$

Maximizar α

sujeto a:

$$\sum_{i=0}^{m} a_{iC} X_{ij} - \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right| L^{-1}(\alpha) \le Y_{jC} + Y_{jR} L^{-1}(\alpha) \qquad j=1,2,...,n$$

$$\sum_{i=0}^{m} a_{iC} X_{ij} + \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right| L^{-1}(\alpha) \geq Y_{jC} - Y_{jR} L^{-1}(\alpha) \qquad j=1,2,...,n$$

$$a_{iR} \geq 0 \qquad i=0,1,...,m$$

Si fijamos para la segunda función objetivo un valor numérico α^*, es decir, $\alpha \geq \alpha^*$, que bien puede ser el nivel de presunción que el decisor considera que es al que se obtienen las observaciones de la variable dependiente, el programa matemático a resolver se reduce a uno lineal que expresamos como:

$$\text{Minimizar } z= \sum_{j=1}^{n} \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right|$$

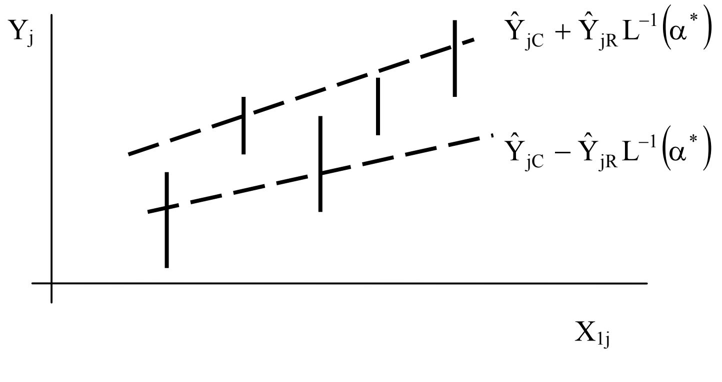
sujeto a:

$$\sum_{i=0}^{m} a_{iC} X_{ij} - \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right| L^{-1}(\alpha^*) \leq Y_{jC} + Y_{jR} L^{-1}(\alpha^*) \qquad j=1,2,...,n$$

$$\sum_{i=0}^{m} a_{iC} X_{ij} + \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right| L^{-1}(\alpha^*) \geq Y_{jC} - Y_{jR} L^{-1}(\alpha^*) \qquad j=1,2,...,n$$

$$a_{iR} \geq 0 \qquad i=0,1,...,m$$

Para un modelo de regresión con una variable independiente, $\tilde{Y} = \tilde{A}_0 + \tilde{A}_1 X_1$, y tomándose el nivel de presunción α^* para el cumplimiento de $\tilde{Y}_j = \hat{\tilde{Y}}_j$, el funcionamiento del modelo de Sakawa y Yano quedaría representado en la siguiente figura:



Es decir, con este modelo exigimos una menor congruencia entre las observaciones de la variable explicada $\tilde{Y}_j$ y sus estimaciones, $\hat{\tilde{Y}}_j$, que en el de Tanaka e Ishibuchi. En el modelo de Sakawa y Yano se exige, para un nivel α^* prefijado por el decisor, que exista únicamente cierto solapamiento entre el α-corte de la variable independiente observada, $Y_{j\alpha*}$ y su valor estimado tras hallar los parámetros a_{iC}, a_{iR}, i=0,1,…,m, $\hat{Y}_{j\alpha*}$, es decir: $Y_{j\alpha*} \cap \hat{Y}_{j\alpha*} \neq \emptyset$, mientras que con el de Tanaka exigimos que $\hat{Y}_{j\alpha*} \cap Y_{j\alpha*} = Y_{j\alpha*}$.

Queremos señalar que el modelo de regresión posibilístico de Tanaka puede ser adaptado al planteamiento realizado por Sakawa y Yano, pero teniendo en cuenta que no exigimos la igualdad entre observaciones y estimaciones, sino que exigimos que el hiperplano ajustado incluya las observaciones realizadas. En este caso para dos números borrosos $\tilde{A}$ y $\tilde{B}$, diremos que $\tilde{A}$ contiene a $\tilde{B}$ con una posibilidad igual o superior a α $\text{Poss}\left(\tilde{B} \subseteq \tilde{A}\right) \geq \alpha$ si $B_\alpha \subseteq A_\alpha$, es decir, si:

$A^1(\alpha) \leq B^1(\alpha)$ y $A^2(\alpha) \geq B^2(\alpha)$

De esta forma, diremos que la posibilidad de que la j-ésima observación de la variable dependiente, $\tilde{Y}_j$ está contenida con un grado igual o superior a α en $\hat{\tilde{Y}}_j$, si se cumple:

$$\sum_{i=0}^{m} a_{iC} X_{ij} - \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right| L^{-1}(\alpha) \leq Y_{jC} - Y_{jR} L^{-1}(\alpha)$$

$$\sum_{i=0}^{m} a_{iC} X_{ij} + \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right| L^{-1}(\alpha) \geq Y_{jC} + Y_{jR} L^{-1}(\alpha)$$

Así, podríamos plantear un el programa matemático para determinar los parámetros a_{iC} y a_{iR}, de forma que se minimice la incertidumbre de las estimaciones de $\tilde{Y}_j$, $\hat{\tilde{Y}}_j$, y conjuntamente, se maximice Poss($\tilde{Y}_j \subseteq \hat{\tilde{Y}}_j$) $\forall$j. Para ello debemos plantear el programa multiobjetivo no lineal:

Minimizar z= $\sum_{j=1}^{n} \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right|$

Maximizar α

sujeto a:

$$\sum_{i=0}^{m} a_{iC} X_{ij} - \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right| L^{-1}(\alpha) \leq Y_{jC} - Y_{jR} L^{-1}(\alpha) \qquad j=1,2,...,n$$

$$\sum_{i=0}^{m} a_{iC} X_{ij} + \sum_{i=0}^{m} a_{iR} \left|X_{ij}\right| L^{-1}(\alpha) \geq Y_{jC} + Y_{jR} L^{-1}(\alpha) \qquad j=1,2,...,n$$

$$a_{iR} \geq 0 \qquad i=0,1,...,m$$

Si fijamos para la segunda función objetivo un cumplimiento mínimo de α^* que es el nivel al cual el decisor considera que se obtienen los intervalos de confianza $<Y'_{jC}, Y'_{jR}>$, j=1,2,…,n, el programa matemático que debemos plantear queda transformado en un programa lineal convencional cuyo planteamiento será:

$$\text{Minimizar } z=\sum_{j=1}^{n}\sum_{i=0}^{m} a_{iR}\left|X_{ij}\right|$$

sujeto a:

$$\sum_{i=0}^{m} a_{iC}X_{ij} - \sum_{i=0}^{m} a_{iR}\left|X_{ij}\right|L^{-1}\left(\alpha^*\right) \leq Y_{jC} - Y_{jR}L^{-1}\left(\alpha^*\right) \qquad j=1,2,...,n$$

$$\sum_{i=0}^{m} a_{iC}X_{ij} + \sum_{i=0}^{m} a_{iR}\left|X_{ij}\right|L^{-1}\left(\alpha^*\right) \geq Y_{jC} + Y_{jR}L^{-1}\left(\alpha^*\right) \qquad j=1,2,...,n$$

$$a_{iR} \geq 0 \qquad i=0,1,...,m$$

2.3.2. Regresión posibilista con inputs y outputs borrosos

En este caso partimos de un conjunto de observaciones $\{(Y_1,X_1), (Y_2,X_2),\ldots,(Y_j,X_j),\ldots,(Y_n,X_n)\}$, donde como siempre, las n observaciones de la variable independiente son intervalos de confianza. Para la j-ésima observación representamos a dicho intervalo a través de su centro y su radio como $< Y'_{jC}, Y'_{jR} >$ con j=1,2,…,n. Si suponemos que estas observaciones son un α^*-corte arbitrario de un número borroso L-L de Dubois y Prade simétrico que finalmente cuantifica a la j-ésima observación de la variable dependiente, siendo el correspondiente a Y_j, $\tilde{Y}_j = \left(Y_{jC}, Y_{jR}\right)_L$ podemos obtener sus centros y sus radios como:

$$Y'_{jC} = Y_{jC}$$

$$Y'_{jR} = Y_{jR}L^{-1}\left(\alpha^*\right) \Rightarrow Y_{jR} = \frac{Y'_{jR}}{L^{-1}\left(\alpha^*\right)}$$

Por otra parte, los inputs, X_j, j=1,2,…,n son vectores m-dimensionales, para las cuales, el i-ésimo componente es un intervalo de confianza. Así, X_j puede ser representado como:

$$X_j = (<1,0>, < X'_{1jC}, X'_{1jR} >, < X'_{2jC}, X'_{2jR} >,\ldots,< X'_{ijC}, X'_{ijR} >,\ldots,< X'_{mjC}, X'_{mjR} >)$$

Asimismo, el intervalo de confianza que cuantifica a la observación j-ésima de la variable i-ésima, con i≥1, también provienen de un número L-L de Dubois y Prade simétrico, es decir: $\tilde{X}_{ij} = \left(X_{ijC}, X_{ijR}\right)_L$, que también es observado a un nivel α^*. Por tanto, se obtienen las siguientes relaciones entre los centros y radios de las observaciones y los números borrosos de los que provienen:

$$X'_{ijC} = X_{ijC}$$

$$X'_{ijR} = X_{ijR} L^{-1}\left(\alpha^*\right) \Rightarrow X_{ijR} = \frac{X'_{ijR}}{L^{-1}\left(\alpha^*\right)}$$

Así pues, el modelo de regresión que debemos utilizar en este caso es:

$$\tilde{y} = \tilde{A}_0 + \tilde{A}_1 \tilde{x}_1 + ... + \tilde{A}_m \tilde{x}_m$$

donde los parámetros $\tilde{A}_i$, i=0,1,2,…,m serán como siempre L-L simétricos, por tanto, el i-ésimo coeficiente lo representaremos como $\tilde{A}_i = (a_{iC}, a_{iR})_L$, i=0,1,…,n.. Tras estimar dichos parámetros, más concretamente sus centros y sus radios, obtendremos unas estimaciones del valor de las observaciones de las variables independientes $\hat{\tilde{Y}}_j$ que no serán L-L de Dubois y Prade –aunque con los instrumentos propuestos en el apartado 1.4. podríamos aproximarlo por uno de la misma naturaleza-. En concreto, los α-cortes de $\hat{\tilde{Y}}_{j_\alpha} = \left[\hat{Y}_j^1(\alpha), \hat{Y}_j^2(\alpha)\right]$, se hallarán de la siguiente forma:

$$\hat{Y}_j^1(\alpha) = \sum_{i=0}^{m} \mathrm{Min}\left\{\left(a_{iC} - a_{iR} L^{-1}(\alpha)\right)\cdot\left(X_{ijC} - X_{ijR} L^{-1}(\alpha)\right); \left(a_{iC} - a_{iR} L^{-1}(\alpha)\right)\cdot\left(X_{ijC} + X_{ijR} L^{-1}(\alpha)\right);\right.$$
$$\left.\left(a_{iC} + a_{iR} L^{-1}(\alpha)\right)\cdot\left(X_{ijC} - X_{ijR} L^{-1}(\alpha)\right); \left(a_{iC} + a_{iR} L^{-1}(\alpha)\right)\cdot\left(X_{ijC} + X_{ijR} L^{-1}(\alpha)\right)\right\}$$

y

$$\hat{Y}_j^2(\alpha) = \sum_{i=0}^{m} \mathrm{Max}\left\{\left(a_{iC} - a_{iR} L^{-1}(\alpha)\right)\cdot\left(X_{ijC} - X_{ijR} L^{-1}(\alpha)\right); \left(a_{iC} - a_{iR} L^{-1}(\alpha)\right)\cdot\left(X_{ijC} + X_{ijR} L^{-1}(\alpha)\right);\right.$$
$$\left.\left(a_{iC} + a_{iR} L^{-1}(\alpha)\right)\cdot\left(X_{ijC} - X_{ijR} L^{-1}(\alpha)\right); \left(a_{iC} + a_{iR} L^{-1}(\alpha)\right)\cdot\left(X_{ijC} + X_{ijR} L^{-1}(\alpha)\right)\right\}$$

Al igual que en el caso de inputs ciertos, en este modelo se propone minimizar la amplitud conjunta de todas las estimaciones que obtenemos sobre las variables dependientes realizadas, $\hat{\tilde{Y}}_j$, j=1,2,…n, y simultáneamente maximizar el grado de presunción con que $\mathrm{Poss}\left(\hat{\tilde{Y}}_j = \tilde{Y}_j\right) \geq \alpha$, j=1,2,…,n. El programa matemático que debemos plantear y resolver para hallar en este caso a_{iC} y a_{iR}, i=0,1,...,m es análogo al del apartado 2.3.1.:

Minimizar z= $\sum_{j=1}^{n}\left[\hat{Y}_j^2(0) - \hat{Y}_j^1(0)\right]$

Maximizar α

sujeto a:

$$\hat{Y}_j^1(\alpha) \leq Y_{jC} + Y_{jR} L^{-1}(\alpha) \qquad j=1,2,...,n$$

$\hat{Y}_j^2(\alpha) \geq Y_{jC} - Y_{jR} L^{-1}(\alpha)$ j=1,2,...,n

$a_{iR} \geq 0$ i=0,1,...,m

De esta forma, si exigimos para la segunda función objetivo un nivel mínimo α^* de cumplimiento, que será al que el decisor considera que ha obtenido los intervalos de confianza que cuantifican a las observaciones obtenidas de la variable dependiente y las variables independientes, el programa anterior se transforma en el siguiente, con función objetivo única:

$$\text{Minimizar } z = \sum_{j=1}^{n} \left[\hat{Y}_j^2(0) - \hat{Y}_j^1(0)\right]$$

sujeto a:

$\hat{Y}_j^1(\alpha^*) \leq Y_{jC} + Y_{jR} L^{-1}(\alpha^*)$ j=1,2,...,n

$\hat{Y}_j^2(\alpha^*) \geq Y_{jC} - Y_{jR} L^{-1}(\alpha^*)$ j=1,2,...,n

$a_{iR} \geq 0$ i=0,1,...,m

Por otra parte, podemos adaptar el planteamiento de Sakawa y Yano al modelo de Tanaka, teniendo en cuenta que la adecuación entre $\hat{\tilde{Y}}_j$ e $\tilde{Y}_j$ viene dada por el grado α con el que $\text{Poss}\left(\tilde{Y}_j \subseteq \hat{\tilde{Y}}_j\right) \geq \alpha$. Si minimizamos la incertidumbre de las estimaciones sobre la variable explicada y conjuntamente maximizamos el nivel de presunción con que dichas estimaciones de la variable dependiente incluyen sus valores realmente observados, el programa multiobjetivo que debemos plantear es:

$$\text{Minimizar } z = \sum_{j=1}^{n} \left[\hat{Y}_j^2(0) - \hat{Y}_j^1(0)\right]$$

Maximizar α

sujeto a:

$\hat{Y}_j^1(\alpha) \leq Y_{jC} - Y_{jR} L^{-1}(\alpha)$ j=1,2,...,n

$\hat{Y}_j^2(\alpha) \geq Y_{jC} + Y_{jR} L^{-1}(\alpha)$ j=1,2,...,n

$a_{iR} \geq 0$ i=0,1,...,m

De esta forma, si exigimos para la segunda función objetivo un nivel mínimo de cumplimiento α^*, el programa anterior se transforma en un programa matemático con un único objetivo:

$$\text{Minimizar } z = \sum_{j=1}^{n} \left[\hat{Y}_j^2(0) - \hat{Y}_j^1(0)\right]$$

sujeto a:

$$\hat{Y}_j^1\left(\alpha^*\right) \le Y_{jC} - Y_{jR}L^{-1}\left(\alpha^*\right) \qquad j=1,2,...,n$$

$$\hat{Y}_j^2\left(\alpha^*\right) \ge Y_{jC} + Y_{jR}L^{-1}\left(\alpha^*\right) \qquad j=1,2,...,n$$

$$a_{iR} \ge 0 \qquad i=0,1,...,m$$