

UNIVERSIDAD DE OVIEDO

Departamento de Estadística e I.O. y D.M.



**Estadística con datos imprecisos
basada en una métrica generalizada**

Tesis Doctoral

Manuel Montenegro Hermida

Don Pedro Gil Álvarez, Director del Departamento de Estadística e Investigación Operativa y Didáctica de la Matemática de la Universidad de Oviedo, autoriza la presentación de la Tesis Doctoral realizada por D. Manuel F. Montenegro Hermida, bajo la dirección de las Doctoras Dña. María Rosa Casals Varela, Dña. Ana Colubi Cervero y Dña María Ángeles Gil Álvarez, con el título “Estadística con datos imprecisos basada en una métrica generalizada”.

Oviedo, 4 de Febrero de 2003

Dña. María Rosa Casals Varela, Dña. Ana Colubi Cervero y Dña. María Ángeles Gil Álvarez, Profesoras del Departamento de Estadística e Investigación Operativa y Didáctica de la Matemática de la Universidad de Oviedo,

C E R T I F I C A N:

Que la memoria titulada “Estadística con datos imprecisos basada en una métrica generalizada” para optar al Grado de Doctor por la Universidad de Oviedo, ha sido realizada por D. Manuel F. Montenegro Hermida bajo su dirección, dando su conformidad para que sea presentada para su defensa en la Universidad de Oviedo.

Oviedo, 4 de Febrero de 2003

*A mis padres,
a Julio y a Dani*

Agradecimientos

Deseo destacar el agradecimiento debido a algunas personas que me ayudaron a llegar a buen puerto en la tarea de realizar una tesis doctoral. Sobre todo quisiera agradecer muy profundamente el apoyo recibido de mis tres directoras, Rosa, Ana y M^a Ángeles, que supieron compaginar la labor de dirección con su amistad.

También quiero expresar mi gratitud, muy especialmente, a Nacho Martínez López por su apoyo científico, pero también por esas ‘discusiones’ que sirven para que uno ejercite su mente y, en muchos casos, comprenda (que no comparta) planteamientos distintos del propio. Por supuesto del grupo de personas que me dieron apoyo científico y personal no puedo olvidarme de Gil González y Miguel López. Además, en Miguel se da un tercer argumento para el agradecimiento ya que, al compartir despacho con él, tuvo que soportarme en muchas ocasiones, lo que espero sinceramente sepa perdonarme. Espero que todos sean capaces de llegar a sentir el sincero agradecimiento que quiero hacerles llegar mediante estas líneas. Tampoco quisiera dejar de nombrar explícitamente a Hortensia López, Teresa López y Asun Lubiano por sus comentarios y también por su apoyo personal. Uno al final no sabe muy bien si lo que se llega a agradecer más a lo largo de una tarea como ésta es el apoyo científico que, formalmente, influye más en el resultado, o el personal, sin el que, a lo mejor, no es posible llegar al final de la misma.

Quisiera, también, dar las gracias a todos los compañeros del Departamento de Estadística e Investigación Operativa y Didáctica de la Matemática de la Universidad de Oviedo y, en especial, a su director D. Pedro Gil.

Aunque algunas personas duden de la conveniencia de situarlo aquí quisiera usar este espacio para agradecer el esfuerzo de todas aquellas personas que, de un modo u otro, hacen que much@s de mis paisan@s no se sientan sol@s ahora que ven peligrar su modo de vida, que también es el de mis padres y, por tanto, forma parte de mis raíces. Espero que el grito de *NUNCA MÁIS* no llegue tarde para mi tierra. Esto también me da pie para resaltar a l@s amig@s que desde allí siempre me hicieron llegar su calor a lo largo de estos años: Senén, Pili, Genucha, Xose, Carlos ...

Finalmente, aunque no en último lugar, quisiera dar las gracias, por todo, a mis padres, a Julio y a Dani.

Durante este tiempo he sido miembro de los equipos de investigación de los Proyectos de Investigación del MCYT Ref. BFM2001-3494 y BFM2002-01057. Quiero agradecer la ayuda económica recibida.

Prólogo

Podemos encontrar aplicaciones de la Estadística prácticamente en todos los campos (Biología, Ciencias de la Educación, Economía, Medicina, Sociología, etc.). Buena parte de los problemas reales con experimentos aleatorios, admiten modelos apropiados dentro del Cálculo de Probabilidades para los que la Estadística dispone de procedimientos de resolución bien consolidados. La mayoría de estos modelos involucran variables aleatorias (reales o vectoriales) asociadas al experimento considerado, y ciertas características de su distribución.

Sin embargo, hay otros problemas reales para los que los modelos tradicionales son insuficientes, porque no recogen adecuadamente toda la incertidumbre subyacente. Entre estos, cabe mencionar los que consideran datos experimentales imprecisos. En ellos, a la incertidumbre debida a la aleatoriedad del resultado experimental se une la imprecisión en la definición de los datos. A menudo, los datos imprecisos se someten a una codificación real, lo que conlleva que, en las medidas de error que intervienen en las distintas técnicas estadísticas aplicables a los problemas, se ignore parte de la incertidumbre (ya que, por la codificación considerada, pueden identificarse datos diferentes).

Los conjuntos aleatorios (elementos aleatorios con valores de conjunto clásico) y las variables aleatorias difusas (que, en el sentido de Puri y Rales-

cu, representan elementos aleatorios con valores de conjunto difuso), se han introducido como un modelo útil y bien establecido para un buen número de elementos aleatorios con valores imprecisos.

En las últimas décadas, se han desarrollado diversos estudios probabilísticos sobre conjuntos aleatorios y variables aleatorias difusas. En este sentido cabe destacar, entre otros, los relativos a la formalización de la medibilidad de las variables aleatorias difusas, las leyes fuertes de los grandes números según diferentes métricas para las dos clases de elementos aleatorios, la definición de medidas de centralización y dispersión, etc.

Por el contrario, el tratamiento de estos elementos aleatorios desde una perspectiva estadística apenas ha recibido atención.

En esta memoria van a presentarse estudios de introducción a una metodología estadística para los conjuntos aleatorios con valores de intervalo compacto real, y para las variables aleatorias difusas con valores de conjunto difuso real. Estos estudios tienen en común que se basan en una *métrica generalizada* definida sobre los espacios de valores de los dos tipos de elementos aleatorios.

En el **Capítulo 1** se analiza en detalle el *problema de regresión y correlación lineal entre dos conjuntos aleatorios con valores de intervalo compacto real*.

En primer lugar, se obtienen de forma descriptiva las soluciones “mínimo-cuadráticas” de dos tipos de relaciones lineales (según que la “ordenada en el origen” tome valor real o de intervalo), y se discuten las analogías y diferencias con respecto al caso de variables reales. Merece mención especial la caracterización de los casos en los que la solución óptima no es única. Para la determinación práctica de las relaciones óptimas se ha desarrollado un algoritmo, ilustrándose su aplicación con un ejemplo real.

Estos estudios sirven de apoyo para la definición de un coeficiente para la proporción de variabilidad del conjunto aleatorio “dependiente” que se explica por la variabilidad del “independiente”, y que extiende el conocido coeficiente de determinación en la correlación lineal entre dos variables reales. Las propiedades del nuevo coeficiente corroboran su utilidad como medida del grado de dependencia funcional (más concretamente, “lineal” en el sentido de la clase de relación considerada) entre los dos elementos aleatorios. En esta parte cabe destacar las discrepancias detectadas entre los dos tipos de relaciones lineales, ya que el problema de regresión para el primer tipo es una particularización del correspondiente al segundo, pero no ocurre lo mismo con el problema de correlación. El cálculo del coeficiente se presenta también de forma algorítmica, y la aplicación se ilustra completando el ejemplo visto anteriormente.

Los resultados obtenidos sobre ciertas relaciones funcionales entre elementos aleatorios imprecisos, pueden emplearse para la *realización de predicciones*. Para contrastar si la consideración de una relación lineal entre los dos elementos aleatorios implica una mejora en la predicción de los valores del elemento dependiente en función de los del independiente, se desarrollan diversos tests de hipótesis: basados en condiciones de normalidad, en la teoría para grandes muestras, o en técnicas bootstrap.

En el **Capítulo 2** se desarrollan *procedimientos de contraste de hipótesis a partir de los datos proporcionados por la observación de una variable aleatoria difusa*. A diferencia de los tests de la última parte del Capítulo 1, en el presente el contraste no se refiere a un parámetro con valores reales, sino difusos: el *valor esperado (difuso) de la variable aleatoria difusa* considerada. Sobre la base de la métrica adoptada, la hipótesis nula del contraste bilateral para una muestra puede expresarse como la de anulación de cierta función real del valor esperado difuso.

Teniendo en cuenta esta equivalencia se van a desarrollar tests de hipótesis para: el valor esperado de una variable aleatoria difusa, la igualdad de valores esperados de dos variables aleatorias difusas, y la igualdad de valores esperados de un número finito cualquiera de variables.

Para el contraste de todas esas hipótesis se construyen *métodos exactos para variables aleatorias difusas normales* (en el sentido de Puri y Ralescu). La normalidad para estos elementos aleatorios constituye un modelo poco realista y sólo adaptable a un reducido número de situaciones prácticas.

Por el contrario, la suposición de que la(s) variable(s) aleatoria(s) difusa(s) en estudio toma(n) un número finito de valores distintos es aplicable a la mayoría de las situaciones prácticas. Para este caso, se establecen *métodos aproximados basados en técnicas asintóticas* (más concretamente, las que aprovechan las buenas propiedades de los estimadores consistentes y asintóticamente normales), y *métodos basados en técnicas bootstrap*.

A través de la *simulación de varios tipos de variables aleatorias difusas*, y considerando distintas métricas particulares y diversos tamaños muestrales, se realiza una *comparación de la utilidad de los diferentes métodos de contraste propuestos*, obteniéndose conclusiones de interés.

La memoria se completa con el planteamiento de algunas nuevas líneas de investigación que han surgido a lo largo de la elaboración de la misma, así como con el de problemas iniciados en cuyo tratamiento y análisis podría abundarse más en el futuro.

Contenido

Agradecimientos	iii
Prólogo	v
Contenido	ix
Índice de Figuras	xv
Índice de Tablas	xvii
Preliminares	1
0.1 Valores imprecisos; definiciones y aritmética	1
0.1.1 Valores de intervalo	2
0.1.2 Valores difusos	2
0.2 Elementos aleatorios con valores imprecisos; formalización y valor esperado	3
0.2.1 Conjuntos aleatorios	3
0.2.2 Variables aleatorias difusas	5

0.3	Métricas generalizadas entre valores imprecisos	7
0.3.1	Métrica sobre $\mathcal{K}_c(\mathbb{R})$	7
0.3.2	Métrica sobre $\mathcal{F}_c(\mathbb{R})$	10
1	Estudios de regresión y correlación para conjuntos aleatorios compactos y convexos de \mathbb{R}	13
1.1	Estudios descriptivos de regresión lineal entre dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$	15
1.1.1	Relación lineal de Tipo 1: coeficiente b con valores en \mathbb{R}	18
1.1.2	Relación lineal de Tipo 2: coeficiente B con valores en $\mathcal{K}_c(\mathbb{R})$	26
1.1.3	Algoritmo para la determinación de la relación lineal óptima de Tipo 2	40
1.1.4	Aplicación sobre un ejemplo real	42
1.2	Estudios descriptivos de correlación lineal entre dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$	45
1.2.1	Análisis de correlación para relaciones lineales de Tipo 1	46
1.2.2	Análisis de correlación para relaciones lineales de Tipo 2	50
1.2.3	Algoritmo para el cálculo del coeficiente de determinación en relaciones de Tipo 2	57
1.2.4	Aplicación sobre un ejemplo real	59

1.3	Estudios inferenciales en problemas de relaciones lineales entre dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$	60
1.3.1	Contraste sobre R_{XY}^2 cuando (mid X , mid Y , spr X , spr Y) tiene distribución normal	64
1.3.2	Contraste sobre R_{XY}^2 cuando X e Y son conjuntos aleatorios normales	74
1.3.3	Contrastes asintóticos sobre R_{XY}^2 cuando X e Y son conjuntos aleatorios simples	76
1.3.4	Contrastes sobre el coeficiente a en modelos lineales	96
1.3.5	Aplicación sobre un ejemplo real	100
1.4	Valoración final y problemas abiertos	101
2	Contraste de hipótesis sobre el valor esperado difuso de variables aleatorias difusas	105
2.1	Contrastes sobre la media poblacional de una variable aleatoria difusa con valores en $\mathcal{F}_c(\mathbb{R})$	107
2.1.1	Test bilateral para el valor esperado de una variable aleatoria difusa normal	108
2.1.2	Test bilateral asintótico para el valor esperado de una variable aleatoria difusa simple	112
2.1.3	Test bilateral bootstrap para el valor esperado de una variable aleatoria difusa simple	124
2.1.4	Estudios de simulación: comparación de técnicas de contraste sobre el valor esperado de una variable aleatoria difusa	130

2.2	Contrastes sobre la igualdad de medias poblacionales de dos variables aleatorias difusas con valores en $\mathcal{F}_c(\mathbb{R})$	133
2.2.1	Test bilateral para la igualdad de valores esperados poblacionales de dos variables aleatorias difusas normales	134
2.2.2	Test bilateral asintótico para la igualdad de valores esperados poblacionales de dos variables aleatorias difusas simples	137
2.2.3	Test bilateral bootstrap para la igualdad de valores esperados poblacionales de dos variables aleatorias difusas simples	145
2.2.4	Estudios de simulación: comparación de técnicas de contraste sobre la igualdad de valores esperados de dos variables aleatorias difusas	150
2.3	Contrastes tipo ANOVA unifactorial para una variable aleatoria difusa con valores en $\mathcal{F}_c(\mathbb{R})$	152
2.3.1	Test ANOVA unifactorial para variables aleatorias difusas normales	153
2.3.2	Test ANOVA unifactorial asintótico para variables aleatorias difusas simples	155
2.3.3	Test ANOVA unifactorial bootstrap para variables aleatorias difusas simples	164
2.3.4	Estudios de simulación: comparación de técnicas de contraste ANOVA para variables aleatorias difusas	165
2.4	Valoración final y problemas abiertos	166

<i>Contenido</i>	xiii
Epílogo	169
Referencias y Bibliografía	173
Índice Temático	182

Índice de Figuras

1.1	Representación gráfica de $\hat{Y}(\omega) = aX(\omega) + b$	18
1.2	Representación gráfica de $\hat{Y}(\omega) = aX(\omega) + B$	27
1.3	Diagrama de “rectángulos” del Ejemplo 1.1.2 y soluciones óptimas del tipo $\hat{Y}(\omega) = aX(\omega) + B$	37
1.4	Diagrama de los datos rectangulares en el Ejemplo 1.1.3	44
2.1	Descripción, según los expertos, de la observación ‘ m minutos’ en el Ejemplo 2.1.1	111
2.2	Conjuntos difusos correspondientes al nivel de conocimiento en informática	123
2.3	Valores de la variable nubosidad	144
2.4	Valores de la variable tiempo de TV	163

Índice de Tablas

1.1	Datos sobre los rangos de la tensión arterial sistólica (X) y diastólica (Y)	43
2.1	Nivel de conocimientos adquirido	123
2.2	Simulación para los tests asintóticos (una muestra) .	125
2.3	Simulación para muestras de tamaño pequeño/medio	132
2.4	Simulación para grandes muestras	133
2.5	Datos de nubosidad en ciertos días de Ω_1 y Ω_2	144
2.6	Simulación de los tests asintóticos. Dos muestras . .	151
2.7	Datos muestrales del tiempo dedicado a la TV	163
2.8	Simulación de los tests asintóticos. ANOVA	166

Preliminares

Este breve capítulo recoge los conceptos y resultados ya establecidos que son fundamentales en el desarrollo de esta memoria.

En primer lugar, se recuerda la definición de los espacios de valores de los elementos aleatorios imprecisos con los que va a trabajarse, y la aritmética para esos valores.

En segundo término se presentan las formalizaciones para los elementos aleatorios con valores de conjunto y con valores de conjunto difuso, y sus correspondientes valores esperados. Se particularizará a los casos que van a tratarse en la memoria, destacándose las propiedades de interés que tengan relevancia en los estudios de los siguientes capítulos.

Por último, se exponen las métricas generalizadas en las que se basan las técnicas estadísticas propuestas y que representan la contribución principal de este trabajo.

0.1 Valores imprecisos; definiciones y aritmética

En esta sección se indican los valores que pueden tomar los elementos aleatorios de cuyo análisis estadístico nos vamos a ocupar, así como algunas de las

operaciones entre ellos, y se muestran varias de las notaciones adoptadas a lo largo de la memoria.

0.1.1 Valores de intervalo

Se denotará por $\mathcal{K}_c(\mathbb{R})$ la clase de los intervalos reales compactos no vacíos (a los que haremos referencia habitualmente como *valores de intervalo*).

El espacio $\mathcal{K}_c(\mathbb{R})$ puede dotarse de una estructura semilineal, inducida por el *producto por un escalar* y la *suma de Minkowski*, definidos de forma que:

$$\lambda A = \{\lambda a \mid a \in A\},$$

$$A + B = \{a + b \mid a \in A, b \in B\},$$

para $A, B \in \mathcal{K}_c(\mathbb{R})$, $\lambda \in \mathbb{R}$. Como $\mathcal{K}_c(\mathbb{R})$ no es un grupo respecto a la suma, al no existir un elemento simétrico para cada intervalo compacto no vacío, $\mathcal{K}_c(\mathbb{R})$ no tiene estructura de espacio vectorial.

0.1.2 Valores difusos

Por $\mathcal{F}_c(\mathbb{R})$ representaremos la clase de las funciones de \mathbb{R} en $[0, 1]$ semi-continuas superiormente, cuasi-convexas, normalizadas y con clausura del soporte compacta (a los que haremos referencia habitualmente como *valores o números difusos*). En otras palabras, $\mathcal{F}_c(\mathbb{R})$ denota el espacio de los subconjuntos difusos $\tilde{V} : \mathbb{R} \rightarrow [0, 1]$ cuyos α -niveles \tilde{V}_α son elementos de $\mathcal{K}_c(\mathbb{R})$, cualquiera que sea $\alpha \in [0, 1]$ (donde $\tilde{V}_\alpha = \{x \in \mathbb{R} \mid \tilde{V}(x) \geq \alpha\}$ si $\alpha \in (0, 1]$, $\tilde{V}_0 = \text{cl}\{x \in \mathbb{R} \mid \tilde{V}(x) > 0\}$).

El espacio $\mathcal{F}_c(\mathbb{R})$ puede dotarse de una estructura semilineal, inducida por el *producto de un conjunto difuso por un escalar* y la *suma de conjuntos difusos*, extendidas siguiendo el *principio de extensión de Zadeh* (1975).

Ambas leyes de composición (la externa y la interna) resultan compatibles con las obtenidas aplicando la aritmética de conjuntos anterior para cada α -nivel, de modo que para todo $\tilde{U}, \tilde{V} \in \mathcal{F}_c(\mathbb{R})$ y $\lambda \in \mathbb{R}$ se definen $\lambda\tilde{U}$ y $\tilde{U} \oplus \tilde{V}$ de manera que:

$$(\lambda\tilde{U})_\alpha = \lambda\tilde{U}_\alpha = \{\lambda u \mid u \in \tilde{U}_\alpha\},$$

$$(\tilde{U} \oplus \tilde{V})_\alpha = \tilde{U}_\alpha + \tilde{V}_\alpha = \{u + v \mid u \in \tilde{U}_\alpha, v \in \tilde{V}_\alpha\},$$

para todo $\alpha \in [0, 1]$.

0.2 Elementos aleatorios con valores imprecisos; formalización y valor esperado

En esta sección se incluyen los elementos aleatorios con los que vamos a trabajar: los elementos aleatorios con valores de conjunto clásico (más concretamente, de intervalo compacto no vacío), y los elementos aleatorios cuyos valores son números difusos.

0.2.1 Conjuntos aleatorios

En la clase $\mathcal{K}(\mathbb{R}^p)$ de los conjuntos compactos y no vacíos de \mathbb{R}^p , se define la métrica de Hausdorff como sigue:

$$d_H(A, B) = \max \left\{ \sup_{a \in A} \inf_{b \in B} \|a - b\|, \sup_{b \in B} \inf_{a \in A} \|a - b\| \right\},$$

donde $\|\cdot\|$ representa la norma euclídea correspondiente y $A, B \in \mathcal{K}(\mathbb{R}^p)$. Sobre $\mathcal{K}(\mathbb{R})$ se reemplazará $\|\cdot\|$ por $|\cdot|$.

Si $A, B \in \mathcal{K}_c(\mathbb{R})$, la métrica de Hausdorff admite la expresión alternativa siguiente:

$$d_H(A, B) = \max\{ |\inf A - \inf B|, |\sup A - \sup B| \}.$$

Dado un espacio de probabilidad (Ω, \mathcal{A}, P) , se dice que una aplicación $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ es un *conjunto aleatorio con valores en $\mathcal{K}_c(\mathbb{R})$* asociado con el espacio medible (Ω, \mathcal{A}) (en lo que sigue, simplemente, conjunto aleatorio), si X es $(\mathcal{A}, \mathcal{B}_{d_H})$ -medible, donde \mathcal{B}_{d_H} es la σ -álgebra generada por la topología inducida por d_H en $\mathcal{K}_c(\mathbb{R})$. El concepto de conjunto aleatorio puede establecerse de forma más general, admitiendo como espacio de llegada la clase $\mathcal{K}(\mathbb{R}^p)$, o incluso $\mathcal{K}(\mathbb{B})$ para un espacio de Banach \mathbb{B} .

Si $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ es un conjunto aleatorio asociado al espacio (Ω, \mathcal{A}, P) , y $E[|X| | P] < \infty$ (propiedad a la que suele referirse como de *integrabilidad acotada de X*), con $|X|(\omega) = \sup\{|x| \mid x \in X(\omega)\}$ para todo $\omega \in \Omega$, entonces se define el *valor esperado de X en el sentido de Aumann (1965)* como el conjunto

$$E^A[X|P] = \{E(\xi|P) \mid \xi : \Omega \rightarrow \mathbb{R}, \xi \in L^1(\Omega, \mathcal{A}, P), \xi \in X \text{ c.s. } [P]\},$$

que en el caso en que X toma valores en $\mathcal{K}_c(\mathbb{R})$ corresponde al intervalo compacto

$$E^A[X|P] = [E(\inf X|P), E(\sup X|P)].$$

En la memoria consideraremos a veces dos tipos de conjuntos aleatorios especiales: normales y simples.

Dado un espacio de probabilidad (Ω, \mathcal{A}, P) , y un conjunto aleatorio X asociado, se dice que X es un *conjunto aleatorio normal* en el sentido de Lyashenko (1980) si, para cada $\omega \in \Omega$, el intervalo $X(\omega)$ puede expresarse como la suma de Minkowski

$$X(\omega) = \xi(\omega) + E^A[X|P],$$

donde $\xi : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria real asociada al espacio (Ω, \mathcal{A}, P) con distribución normal de media 0 y varianza $\sigma^2 = \text{Var}(\xi|P)$.

Diremos que X es un *conjunto aleatorio simple* si toma un número finito de valores distintos sobre el espacio muestral Ω .

Los conjuntos aleatorios se han examinado y utilizado ampliamente en muchos estudios (ver, por ejemplo, Kendall, 1974, Matheron, 1975, Diamond, 1990, Cressie, 1993, Molchanov, 1993, Stoyan *et al.*, 1995, Molchanov, 1998, Stoyan, 1998, Pick, 2000, Wang, 2001).

0.2.2 Variables aleatorias difusas

El concepto de variable aleatoria difusa que vamos a emplear, es una generalización de la noción de conjunto aleatorio con valores en $\mathcal{K}_c(\mathbb{R})$, y se debe a Puri and Ralescu (1986).

Dado un espacio de probabilidad (Ω, \mathcal{A}, P) , se dice que una aplicación $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c(\mathbb{R})$ es una *variable aleatoria difusa con valores en $\mathcal{F}_c(\mathbb{R})$* (también llamada conjunto difuso aleatorio o variable difusa aleatoria) asociada con el espacio medible (Ω, \mathcal{A}) si la aplicación α -nivel, $\mathcal{X}_\alpha : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ tal que $\mathcal{X}_\alpha(\omega) = (\mathcal{X}(\omega))_\alpha$ para todo $\omega \in \Omega$, es un conjunto aleatorio cualquiera que sea $\alpha \in [0, 1]$. Puri y Ralescu definieron el concepto de variable aleatoria difusa de modo más general, admitiendo como espacio de llegada la clase $\mathcal{F}(\mathbb{R}^p)$, o incluso $\mathcal{F}(\mathbb{B})$ para un espacio de Banach \mathbb{B} , donde $\mathcal{F}(\mathbb{B})$ es la clase de las funciones de \mathbb{B} en $[0, 1]$ semicontinuas superiormente, normalizadas y con aplicación 0-nivel con valores compactos.

Si \mathcal{X} es una variable aleatoria difusa con valores en $\mathcal{F}_c(\mathbb{R})$ (en lo que sigue, simplemente, variable aleatoria difusa), entonces para todo $\alpha \in [0, 1]$, las aplicaciones $\inf \mathcal{X}_\alpha : \Omega \rightarrow \mathbb{R}$ y $\sup \mathcal{X}_\alpha : \Omega \rightarrow \mathbb{R}$ son variables aleatorias reales.

La noción de valor esperado que va a considerarse generaliza la esperanza de un conjunto aleatorio en el sentido de Aumann, y se debe a Puri and Ralescu (1986).

Se dice que una variable aleatoria difusa \mathcal{X} es *acotada integrablemente* si su magnitud $|\mathcal{X}_0| \in L^1(\Omega, \mathcal{A}, P)$, donde $|\mathcal{X}_0|(\cdot) = \sup_{x \in \mathcal{X}_0(\cdot)} |x|$. Si \mathcal{X} es una variable aleatoria difusa acotada integrablemente, se define el *valor esperado de \mathcal{X}* como el único conjunto difuso $\tilde{E}(\mathcal{X}|P)$, tal que

$$(\tilde{E}(\mathcal{X}|P))_\alpha = E^A[\mathcal{X}_\alpha|P]$$

para todo $\alpha \in [0, 1]$.

En esta memoria hay dos tipos especiales de variables aleatorias difusas con los que trabajaremos a menudo: normales y simples.

Dado un espacio de probabilidad (Ω, \mathcal{A}, P) , y una variable aleatoria difusa \mathcal{X} asociada, se dice que \mathcal{X} es una *variable aleatoria difusa normal* en el sentido de Puri and Ralescu (1985) si, para cada $\alpha \in [0, 1]$ y $\omega \in \Omega$, el intervalo compacto $\mathcal{X}_\alpha(\omega)$ puede expresarse como la suma de Minkowski del valor real $\xi(\omega)$ y el α -nivel del valor esperado $\tilde{E}(\mathcal{X}|P)$, de forma que para cada $\omega \in \Omega$ se cumple que:

$$\mathcal{X}(\omega) = \xi(\omega) \oplus \tilde{E}(\mathcal{X}|P),$$

donde $\xi : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria real asociada al espacio (Ω, \mathcal{A}, P) con distribución normal de media 0 y varianza $\sigma^2 = \text{Var}(\xi|P)$.

Se dice que \mathcal{X} es una *variable aleatoria difusa simple* si $\mathcal{X}(\Omega)$ tiene cardinal finito.

Las variables aleatorias difusas han cobrado bastante interés en la última década, destacando los estudios de tipo probabilístico realizados (ver, por ejemplo, Klement *et al.*, 1986, López-Díaz and Gil, 1997, 1998, Näther, 1997, Körner, 1997b, Colubi *et al.*, 1999, 2001, 2002, Molchanov, 1999, entre otros).

0.3 Métricas generalizadas entre valores imprecisos

Sobre las clases $\mathcal{K}_c(\mathbb{R})$ y $\mathcal{F}_c(\mathbb{R})$ pueden definirse varias métricas (ver, por ejemplo, Diamond and Kloeden, 1994). La conveniencia de unas u otras depende de su finalidad. En este sentido, la métrica de Hausdorff es la base de la medibilidad de los conjuntos aleatorios, y para formalizar la medibilidad de las variables aleatorias difusas se han considerado distancias basadas en la de Hausdorff (ver Colubi *et al.*, 2002).

El propósito común de los estudios que se presentan en esta memoria es el desarrollo de métodos estadísticos con datos experimentales que toman valores de conjunto o valores de conjunto difuso. Estos métodos involucran explícita o implícitamente medidas de error. Con el objeto de que estas medidas se adapten fácilmente al cálculo práctico, y den lugar a procedimientos operativos en un contexto estadístico, las métricas que van a emplearse no se basan en la métrica de Hausdorff sino en distancias de tipo cuadrático. Asimismo, se trata de métricas generalizadas con una interpretación intuitiva sencilla.

0.3.1 Métrica sobre $\mathcal{K}_c(\mathbb{R})$

Sobre el espacio $\mathcal{K}_c(\mathbb{R})$ va a considerarse la W -distancia, que para $A, B \in \mathcal{K}_c(\mathbb{R})$ viene dada por:

$$d_W(A, B) = \sqrt{\int_{[0,1]} [f_A(\lambda) - f_B(\lambda)]^2 dW(\lambda)}$$

con

$$f_A(\lambda) = \lambda \sup A + (1 - \lambda) \inf A$$

para todo $\lambda \in [0, 1]$, y donde W es una medida de ponderación normalizada que se formaliza mediante una medida de probabilidad sobre el espacio medible $([0, 1], \mathcal{B}_{[0,1]})$ que corresponde a una distribución no degenerada (que en el Capítulo 1 supondremos además simétrica) en $[0, 1]$ (denotándose por $\mathcal{B}_{[0,1]}$ la σ -álgebra de Borel en $[0, 1]$).

La W -distancia es una particularización de una métrica introducida recientemente en un marco más general por Körner and Näther (2002). Sobre la base del Teorema de Radström (ver, por ejemplo, Diamond and Kloeden, 1994), cualquier conjunto $A \in \mathcal{K}_c(\mathbb{R})$ puede encajarse isométricamente mediante su función soporte s_A en un cono de un espacio de Hilbert de funciones. En consecuencia, un conjunto aleatorio puede contemplarse como una función aleatoria que toma valores en un espacio de Hilbert, y cualquier distancia L_2 entre las funciones soporte de dos conjuntos $A, B \in \mathcal{K}_c(\mathbb{R})$ podría expresarse en la forma:

$$D_K(s_A, s_B) = \sqrt{\sum_{(u,v) \in S^0 \times S^0} (s_A(u) - s_B(u))(s_A(v) - s_B(v)) K(u, v)}$$

para alguna función $K : S^0 \times S^0 \rightarrow \mathbb{R}$ (con S^0 la esfera unidad en \mathbb{R} , i.e., $S^0 = \{-1, 1\}$), donde K representa un núcleo definido positivo y simétrico, es decir, $K(1, 1) > 0$, $K(1, -1) = K(-1, 1)$ y $K(1, 1)K(-1, -1) > K(1, -1)K(-1, 1)$.

El cuadrado de la distancia precedente puede expresarse alternativamente como:

$$\begin{aligned} [D_K(s_A, s_B)]^2 &= (K(1, 1) - K(1, -1))[\sup A - \sup B]^2 \\ &+ (K(-1, -1) - K(1, -1))[\inf A - \inf B]^2 \\ &+ 4 K(1, -1)[\text{mid } A - \text{mid } B]^2, \end{aligned}$$

donde $\text{mid } A = [\sup A + \inf A]/2$ denota el punto medio del intervalo A . Si se exige para el núcleo K que asigne “peso” no negativo al cuadrado de la distancia euclídea entre los puntos medios (i.e., $K(1, -1) \geq 0$), entonces, para $A, B \in \mathcal{K}_c(\mathbb{R})$ se tiene que $D_K(s_A, s_B) = d_W(A, B)$.

Observación 0.3.1 Conviene señalar que la métrica utilizada por Diamond (1990), en sus estudios de regresión con intervalos aleatorios, corresponde a d_W con $W(0) = W(1) = .5$, $W(\lambda) = 0$ en el resto. Es evidente que d_W es equivalente a la métrica generalizada $d_{\vec{\lambda}}$ introducida por Bertoluzza *et al.* (1995), donde $\vec{\lambda} = (W(0), W(.5), W(1))$, pero la elección de la medida W sobre $[0, 1]$ es más fácil e intuitiva en la práctica que la elección de $\vec{\lambda}$.

Observación 0.3.2 Aunque la medida W no tiene realmente una interpretación estocástica, se ha formalizado en un contexto probabilístico para poder así trabajar con el espacio de probabilidad $(\Omega \times [0, 1], \mathcal{A} \otimes \mathcal{B}_{[0,1]}, P \otimes W)$, que resulta muy manejable y permite aplicar algunos resultados de interés conocidos de ese contexto.

Observación 0.3.3 La aplicación $f_{[0,1]} : \Omega \times [0, 1] \rightarrow \mathbb{R}$ puede tratarse como una variable aleatoria real constante respecto a P y, si W se supone asociada a una distribución simétrica en $[0, 1]$, $\sigma_{f_{[0,1]}}^2 = \text{Var}[f_{[0,1]}|P \otimes W] = \int_{[0,1]} \lambda^2 dW(\lambda) - .25 > 0$. Además, puede probarse que, si $A, B \in \mathcal{K}_c(\mathbb{R})$, la simetría de W garantiza que

$$[d_W(A, B)]^2 = [\text{mid } A - \text{mid } B]^2 + 4 \sigma_{f_{[0,1]}}^2 [\text{spr } A - \text{spr } B]^2$$

(donde “spr” representa la semiamplitud, es decir, $\text{spr } A = [\sup A - \inf A]/2$). En consecuencia, cuanto mayor es $\sigma_{f_{[0,1]}}^2$ mayor es la influencia de la distancia euclídea entre las semiamplitudes de A y B sobre $d_W(A, B)$, alcanzándose la influencia máxima para la métrica empleada por Diamond (1990).

0.3.2 Métrica sobre $\mathcal{F}_c(\mathbb{R})$

La métrica con la que va a trabajarse sobre $\mathcal{F}_c(\mathbb{R})$ parte de la consideración de dos medidas ponderadoras normalizadas, W y φ , que pueden formalizarse matemáticamente mediante sendas medidas de probabilidad sobre el espacio medible $([0, 1], \mathcal{B}_{[0,1]})$, W asociada con una distribución no degenerada y φ asociada con una función de distribución creciente (en sentido estricto) sobre $[0, 1]$. De nuevo, se puntualiza que W y φ no tienen un significado estocástico, aunque el empleo del modelo probabilístico sea muy útil.

La (W, φ) -distancia entre $\tilde{U}, \tilde{V} \in \mathcal{F}_c(\mathbb{R})$ se define por:

$$D_W^\varphi(\tilde{U}, \tilde{V}) = \sqrt{\int_{[0,1]} [d_W(\tilde{U}_\alpha, \tilde{V}_\alpha)]^2 d\varphi(\alpha)},$$

donde ahora $d_W : \mathcal{K}_c(\mathbb{R}) \times \mathcal{K}_c(\mathbb{R}) \rightarrow [0, +\infty)$ se define por:

$$d_W(\tilde{U}_\alpha, \tilde{V}_\alpha) = \sqrt{\int_{[0,1]} [f_{\tilde{U}}(\alpha, \lambda) - f_{\tilde{V}}(\alpha, \lambda)]^2 dW(\lambda)}$$

con $f_{\tilde{V}}(\alpha, \lambda) = \lambda \sup \tilde{V}_\alpha + (1 - \lambda) \inf \tilde{V}_\alpha$.

Observación 0.3.4 D_W^φ puede verse como una particularización de la métrica introducida por Körner and Nather (2002) en $\mathcal{F}_c(\mathbb{R}^p)$, de modo que para $\tilde{U}, \tilde{V} \in \mathcal{F}_c(\mathbb{R}^p)$:

$$\begin{aligned} & \left[D_K(\tilde{U}, \tilde{V}) \right]^2 \\ &= \int_{(S^{p-1})^2 \times [0,1]^2} (s_{\tilde{U}}(u, \alpha) - s_{\tilde{V}}(u, \alpha))(s_{\tilde{U}}(v, \beta) - s_{\tilde{V}}(v, \beta)) dK(u, \alpha, v, \beta), \end{aligned}$$

donde s denota la función soporte en $\mathcal{F}_c(\mathbb{R}^p)$, S^{p-1} es la esfera unidad en \mathbb{R}^p , y K es un núcleo definido positivo y simétrico. D_K corresponde a la distancia genérica L_2 en el espacio de Banach $\mathcal{L}(S^{p-1} \times [0, 1])$. Como indican Körner y Näther, si $p = 1$ la métrica D_W^φ coincide con D_K para el núcleo K tal que:

$$dK(u, \alpha, v, \beta) = \delta_{\alpha\beta} d\varphi(\alpha) K_1(u, v)$$

con

$$K_1(1, 1) = \int_{[0,1]} \lambda^2 dW(\lambda), \quad K_1(-1, -1) = \int_{[0,1]} (1 - \lambda)^2 dW(\lambda),$$

$$K_1(1, -1) = K_1(-1, 1) = \int_{[0,1]} \lambda(1 - \lambda) dW(\lambda).$$

Por último, debe señalarse que la independencia y la identidad de distribuciones se entenderán en la memoria en el sentido usual (ver, por ejemplo, Billingsley, 1995).

Capítulo 1

Estudios de regresión y correlación para conjuntos aleatorios compactos y convexos de \mathbb{R}

En este capítulo se presentan en primer lugar los estudios descriptivos desarrollados sobre dos modelos de regresión lineal entre dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$. La diferencia entre ambos modelos radica en la naturaleza de la ordenada en el origen de la relación lineal, que para el primero es de un elemento en \mathbb{R} , mientras que para el segundo es de un elemento en $\mathcal{K}_c(\mathbb{R})$.

Sobre la base de la W -distancia introducida en el Capítulo 0, se aplica el principio de mínimos cuadrados, obteniéndose las relaciones óptimas para los dos modelos considerados. Se concluye que la relación mínimo-cuadrática es prácticamente única para el primer modelo, y que en dicho

caso la pendiente no se anula salvo en situaciones muy especiales; para el segundo modelo, la unicidad se cumple mayoritariamente, aunque existen situaciones (caracterizadas en la memoria) en las que hay dos posibles relaciones óptimas, y la anulación de la pendiente se produce en situaciones esperadas y bien justificadas. El cálculo práctico para ese segundo modelo se presenta en forma algorítmica.

Las conclusiones derivadas de los estudios de regresión, conducen de forma inmediata a los de correlación lineal entre dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$. Así, a partir de la expresión de la variación residual para cada uno de los modelos, se define un “coeficiente de determinación” como medida adimensional del grado de dependencia lineal (o, mejor aún, del grado de bondad de las relaciones óptimas según la regresión lineal considerada) entre los dos conjuntos aleatorios. El coeficiente extendido está bien definido, aunque la expresión para el segundo modelo es bastante compleja. Para evitar las dificultades que puede implicar su cálculo práctico para ese segundo modelo, se presenta un algoritmo.

Los resultados de regresión y correlación lineales se ilustran con un ejemplo real.

Obviamente, los estudios descriptivos sólo suelen aplicarse sobre poblaciones pequeñas o, más asiduamente, sobre muestras extraídas de la población. Cuando se trabaja con poblaciones grandes y la información disponible es la de los datos proporcionados por una muestra de esa población, con el fin de tener cierta garantía de la idoneidad de las predicciones pertinentes se procederá habitualmente a contrastar la hipótesis de anulación del coeficiente de determinación extendido (o, alternativamente, la anulación de la pendiente de la relación óptima por regresión). En la última parte del Capítulo 1, se establecen varios procedimientos de contraste, dependiendo de ciertas suposiciones previas sobre los conjuntos aleatorios involucrados

en la relación (simples, normales, caracterizados por ciertas componentes reales normales, etc.) Estos procedimientos hacen uso de técnicas exactas, técnicas para grandes muestras y técnicas bootstrap, entre otras. Se añaden algunas conclusiones y problemas abiertos.

1.1 Estudios descriptivos de regresión lineal entre dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$

En el análisis de regresión lineal simple entre dos variables aleatorias reales, entendido en sentido descriptivo, el objetivo es buscar la relación lineal “más adecuada” entre una *variable dependiente* o *variable respuesta* Y y una *variable independiente* o *variable de predicción* X sobre una muestra o población dada. En otras palabras, se trata de determinar la recta $\hat{y} = ax + b$ que proporciona “mejor ajuste” o “estimación” a los datos. La adecuación, ajuste o estimación “mejor” se interpreta habitualmente en términos del *criterio de mínimos cuadrados* con la distancia euclídea en \mathbb{R} , que permite hallar fácilmente y de forma general la solución óptima, salvo que la variable de predicción sea degenerada (en cuyo caso habría una infinidad de posibles soluciones óptimas y carecería de interés).

El propósito de la presente sección es extender el análisis de regresión anterior cuando X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$. Diamond (1990) realizó algunos estudios sobre el problema que vamos a abordar en esta sección, si bien adoptó una métrica particular (y no siempre conveniente) y sólo halló las soluciones óptimas para situaciones muy especiales, en las que la obtención de conclusiones es inmediata. En consecuencia, la discusión y los resultados que se recogen en esta sección generali-

zan el trabajo de Diamond en varios aspectos: la métrica considerada, d_W , es más general e idónea; las soluciones óptimas se determinan para todos los casos en que existen; estas conclusiones sirven de soporte al desarrollo de los estudios de la Sección 1.2.

La extensión del análisis va a llevarse a cabo de acuerdo con dos tipos de relaciones: lineal con “ordenada en el origen” real; lineal con “ordenada en el origen” con valor de intervalo. Nos ocuparemos de estas extensiones en las dos subsecciones siguientes. Para formalizar la idea de relación “menos errónea”, el “error” va a cuantificarse en términos de la W -distancia introducida en la Subsección 0.3.1, y supondremos que la medida W corresponde a una distribución de probabilidad simétrica con respecto al valor $\lambda = 0.5$. Esta suposición no entraña pérdida de generalidad, ya que no existen razones realmente importantes para discriminar entre $[f_Y(\lambda) - f_{aX+b}(\lambda)]^2$ y $[f_Y(1 - \lambda) - f_{aX+b}(1 - \lambda)]^2$ a través de la ponderación asignada al valorar el error asociado con la estimación.

Antes de analizar la solución de los problemas, van a describirse algunas de las notaciones que se emplearán en dicho análisis, indicándose inicialmente su expresión en función de ciertos momentos de las variables aleatorias reales f_X , f_Y y $f_{[0,1]}$, definidas sobre el espacio de probabilidad $(\Omega \times [0, 1], \mathcal{A} \times \mathcal{B}_{[0,1]}, P \otimes W)$. Esta expresión permitirá aplicar resultados del Análisis de Regresión Lineal simple y múltiple con variables aleatorias reales.

Por otra parte, tanto en la búsqueda de las soluciones de los problemas, como en la discusión sobre la unicidad/duplicidad de esas soluciones, o a la hora de realizar los cálculos prácticos relativos a los resultados en este capítulo (para los que se podrá recurrir subsidiariamente a programas estadísticos bien conocidos), jugará un papel crucial la expresión de las notaciones en términos de ciertos momentos de las variables aleatorias reales

mid X , mid Y , spr X y spr Y (ver Subsección 0.3.1 para la definición de ‘mid’ y ‘spr’) asociadas al espacio de probabilidad (Ω, \mathcal{A}, P) , que también se incluye a continuación.

De este modo, si se denota por $\bar{f}_A(\lambda) = f_A(1 - \lambda)$ cualquiera que sea $\lambda \in [0, 1]$ y $A \in \mathcal{K}_c(\mathbb{R})$:

$$\mu_{f_{[0,1]}} = E(f_{[0,1]}|P \otimes W) = 0.5,$$

$$\mu_{f_X} = E(f_X|P \otimes W) = E(\bar{f}_X|P \otimes W) = E(\text{mid } X|P) = \text{mid } E^A[X|P],$$

$$\sigma_{f_X}^2 = \text{Var}(f_X|P \otimes W) = \text{Var}(\bar{f}_X|P \otimes W)$$

$$= 4 \sigma_{f_{[0,1]}}^2 E((\text{spr } X)^2|P) + \text{Var}(\text{mid } X|P),$$

$$\rho_{f_X f_Y} = \frac{\sigma_{f_X f_Y}}{\sqrt{\sigma_{f_X}^2 \sigma_{f_Y}^2}} = \frac{\text{Cov}(f_X, f_Y|P \otimes W)}{\sqrt{\sigma_{f_X}^2 \sigma_{f_Y}^2}}$$

$$= \frac{\text{Cov}(\text{mid } X, \text{mid } Y|P) + 4 \sigma_{f_{[0,1]}}^2 E(\text{spr } X \cdot \text{spr } Y|P)}{\sqrt{\sigma_{f_X}^2 \sigma_{f_Y}^2}},$$

$$\rho_{\bar{f}_X f_Y} = \frac{\text{Cov}(\text{mid } X, \text{mid } Y|P) - 4 \sigma_{f_{[0,1]}}^2 E(\text{spr } X \cdot \text{spr } Y|P)}{\sqrt{\sigma_{f_X}^2 \sigma_{f_Y}^2}} \leq \rho_{f_X f_Y},$$

$$\rho_{f_X f_{[0,1]}} = -\rho_{\bar{f}_X f_{[0,1]}} = 2 \sqrt{\frac{\sigma_{f_{[0,1]}}^2}{\sigma_{f_X}^2}} E(\text{spr } X|P).$$

Conviene reseñar que si X e Y son conjuntos aleatorios no degenerados, entonces

$$\sigma_{f_X}^2 > 0, \sigma_{f_Y}^2 > 0, \rho_{f_X f_{[0,1]}} \neq 1, \rho_{f_Y f_{[0,1]}} \neq 1.$$

1.1.1 Relación lineal de Tipo 1: coeficiente b con valores en \mathbb{R}

Sea (Ω, \mathcal{A}, P) un espacio de probabilidad, y sean $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ e $Y : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ dos conjuntos aleatorios no degenerados asociados a ese espacio. Si se observa el conjunto $X(\omega)$ y se quiere “estimar” el valor correspondiente $Y(\omega)$, puede tratarse de aproximarse Y mediante una función lineal de X (o, de forma más general, de $g(X)$, donde $g : \mathbb{R} \rightarrow \mathbb{R}$ es una función medible bien definida). La Figura 1.1 muestra la representación gráfica de la estimación $\hat{Y}(\omega) = aX(\omega) + b$ (con $a, b \in \mathbb{R}$):

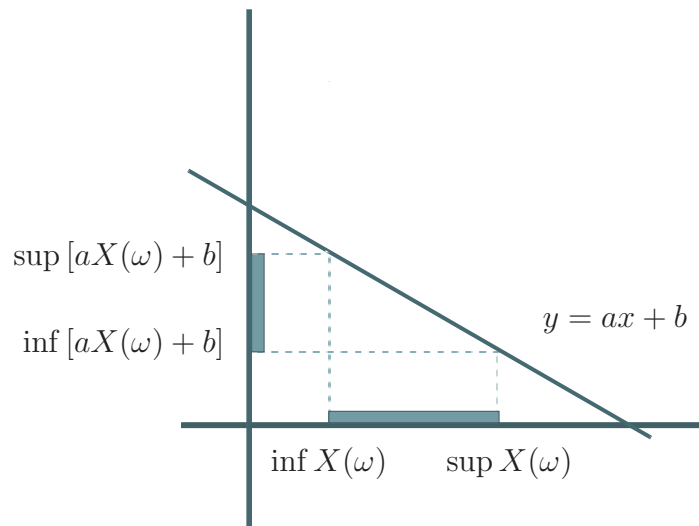


Fig. 1.1: Representación gráfica de $\hat{Y}(\omega) = aX(\omega) + b$

La búsqueda de la mejor relación de Y como función lineal de X , se va a llevar a cabo mediante la aplicación del *método de mínimos-cuadrados* basada en la W -distancia introducida en la Subsección 0.3.1. De este modo, la suma (o, precisando más, la media) de los cuadrados de los residuos se extiende por medio de la suma (la media) de los cuadrados de las W -

distancias entre los valores de Y observados y los estimados por la relación lineal.

Por lo tanto, el problema consiste en hallar el mínimo de la **función objetivo** $\phi : \mathbb{R} \times \mathbb{R} \rightarrow [0, +\infty)$ tal que

$$\phi(a, b) = E\left([d_W(Y, aX + b)]^2 \mid P\right)$$

para $a, b \in \mathbb{R}$.

En condiciones muy generales, el problema planteado admite solución (única salvo en situaciones muy especiales). Más concretamente:

Teorema 1.1.1 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c^*(\mathbb{R}) = \mathcal{K}_c(\mathbb{R}) \setminus \{\{x\} \mid x \in \mathbb{R}\}$ c.s. $[P]$ y $E(|X| \mid P) < \infty$, $E(|Y| \mid P) < \infty$, entonces existe solución para el problema, y la función objetivo ϕ alcanza mínimo absoluto en $(a_0^*, b_0^*) \in (\mathbb{R} \setminus \{0\}) \times \mathbb{R}$, con:*

$$(a_0^*, b_0^*) = \begin{cases} \left(\sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2}} \rho_{f_X f_Y}, \mu_{f_Y} - \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2}} \rho_{f_X f_Y} \cdot \mu_{f_X} \right) & \text{si } \rho_{f_X f_Y}^2 > \rho_{\bar{f}_X f_Y}^2, \\ \left(\sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2}} \rho_{\bar{f}_X f_Y}, \mu_{f_Y} - \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2}} \rho_{\bar{f}_X f_Y} \cdot \mu_{f_X} \right) & \text{en caso contrario.} \end{cases}$$

(La condición $\rho_{f_X f_Y}^2 > \rho_{\bar{f}_X f_Y}^2$ equivale a $\text{Cov}(\text{mid } X, \text{mid } Y \mid P) > 0$) La solución será única salvo en el caso en el que $\text{Cov}(\text{mid } X, \text{mid } Y \mid P) = 0$.

Demostración. Una de las diferencias más destacables con respecto al problema de regresión lineal simple con variables reales, estriba en el hecho de que el producto de un elemento de $\mathcal{K}_c(\mathbb{R})$ por un número real adopta expresiones distintas de acuerdo con el signo del escalar (es decir, según que $a \geq 0$ ó $a \leq 0$). Precizando más:

$$\phi(a, b) = \begin{cases} E \left([f_Y - (af_X + b)]^2 \mid P \otimes W \right) & \text{si } a \geq 0, \\ E \left([f_Y - (a\bar{f}_X + b)]^2 \mid P \otimes W \right) & \text{si } a \leq 0. \end{cases}$$

En particular, si consideramos una muestra aleatoria de n datos bidimensionales, $(X_1, Y_1), \dots, (X_n, Y_n)$, con $X_i, Y_i \in \mathcal{K}_c(\mathbb{R})$ ($i = 1, \dots, n$), se cumple que:

$$\phi(a, b) = \begin{cases} \frac{1}{n} \sum_{i=1}^n \int_{[0,1]} \left(f_{Y_i}(\lambda) - af_{X_i}(\lambda) - b \right)^2 dW(\lambda) & \text{si } a \geq 0, \\ \frac{1}{n} \sum_{i=1}^n \int_{[0,1]} \left(f_{Y_i}(\lambda) - af_{X_i}(1-\lambda) - b \right)^2 dW(\lambda) & \text{si } a \leq 0. \end{cases}$$

En consecuencia, el problema que se desea resolver puede contemplarse como una combinación de dos problemas de regresión lineal simple clásicos, el primero de los cuales involucra las variables aleatorias reales f_X y f_Y (si $a \geq 0$), y el segundo referido a las variables \bar{f}_X y f_Y (si $a \leq 0$).

Los estudios de regresión lineal simple entre variables aleatorias reales nos permiten asegurar que la función

$$\phi_0(a, b) = E \left([f_Y - (af_X + b)]^2 \mid P \otimes W \right)$$

alcanza su mínimo absoluto en \mathbb{R}^2 para los valores, si existen, dados por:

$$a_0 = \frac{\text{Cov}(f_X, f_Y \mid P \otimes W)}{\text{Var}(f_X \mid P \otimes W)} = \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2}} \rho_{f_X f_Y},$$

$$b_0 = E(f_Y \mid P \otimes W) - a_0 E(f_X \mid P \otimes W) = \mu_{f_Y} - \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2}} \rho_{f_X f_Y} \mu_{f_X},$$

y la función

$$\phi'_0(a, b) = E \left([f_Y - (a \bar{f}_X + b)]^2 \middle| P \otimes W \right)$$

alcanza su mínimo absoluto en \mathbb{R}^2 para los valores, si existen, dados por:

$$a'_0 = \frac{\text{Cov}(\bar{f}_X, f_Y | P \otimes W)}{\text{Var}(f_X | P \otimes W)} = \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2}} \rho_{\bar{f}_X f_Y},$$

$$b'_0 = E(f_Y | P \otimes W) - a'_0 E(f_X | P \otimes W) = \mu_{f_Y} - \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2}} \rho_{\bar{f}_X f_Y} \mu_{f_X}.$$

Además, se satisface que:

$$\text{Cov}(f_X, f_Y | P \otimes W) - \text{Cov}(\bar{f}_X, f_Y | P \otimes W) = 8 \sigma_{f_{[0,1]}}^2 E(\text{spr } X \cdot \text{spr } Y | P),$$

que en el caso de que X e Y tomen valores en $\mathcal{K}_c^*(\mathbb{R})$ c.s. $[P]$ (de hecho, en condiciones más generales) es positiva.

Por lo tanto, al ser $\text{Var}(f_X | P \otimes W)$ positiva salvo si X es un conjunto aleatorio degenerado en un valor real (lo que no es posible en las condiciones supuestas en el teorema), se concluye que $a_0 > a'_0$, de modo que $\phi(a, b)$ alcanza el mínimo absoluto en \mathbb{R}^2 en

- (a_0, b_0) si $a_0 > a'_0 \geq 0$,
- (a'_0, b'_0) si $0 \geq a_0 > a'_0$.

En caso contrario (i.e., si $a_0 > 0$ y $a'_0 < 0$) el mínimo absoluto de $\phi(a, b)$ en \mathbb{R}^2 corresponde al valor $\min\{\phi(a_0, b_0), \phi(a'_0, b'_0)\}$.

Además, la igualdad $\phi(a_0, b_0) = \phi(a'_0, b'_0)$ se cumplirá si, y sólo si, se anula el producto $\text{Cov}(\text{mid } X, \text{mid } Y | P) E(\text{spr } X \cdot \text{spr } Y | P)$, lo que al no reducirse X e Y a variables aleatorias reales casi seguro, equivale a que $\text{Cov}(\text{mid } X, \text{mid } Y | P) = 0$. \square

Como un ejemplo de caso especial en el que la solución del problema de Tipo 1 no es única puede considerarse el siguiente:

Ejemplo 1.1.1 Sea (Ω, \mathcal{A}, P) el espacio de probabilidad definido como sigue:

$$\begin{aligned}\Omega &= \{\omega_1, \omega_2, \omega_3\}, \mathcal{A} = \mathcal{P}(\Omega), \\ P(\omega_1) &= P(\omega_2) = P(\omega_3) = 1/3,\end{aligned}$$

y sean X e Y dos conjuntos aleatorios asociados, dados por:

$$\begin{aligned}X(\omega_1) &= [-1, 1], X(\omega_2) = [-2, 2], X(\omega_3) = [-3, 3], \\ Y(\omega_1) &= [-1, 1], Y(\omega_2) = [-2, 2], Y(\omega_3) = [-3, 3].\end{aligned}$$

Cualquiera que sea la medida W , se cumple que $a_1 = -a_2 = 1$, $b_1 = b_2 = 0$, de modo que las relaciones $\widehat{Y} = X$ y $\widehat{Y} = -X$ son las óptimas en esta situación.

Observación 1.1.1 El problema anterior podría presentarse alternativa-mente considerando la función objetivo definida directamente en términos de valores esperados de las variables $\text{mid } X$, $\text{mid } Y$, $\text{spr } X$ y $\text{spr } Y$ asociadas al espacio de probabilidad (Ω, \mathcal{A}, P) (a excepción de $\sigma_{f_{[0,1]}}^2$ que sólomente depende de W), como sigue:

$$\zeta(a, b) = E \left((\text{mid } Y - a \text{mid } X - b)^2 + 4 \sigma_{f_{[0,1]}}^2 (\text{spr } Y - |a| \text{spr } X)^2 \mid P \right)$$

con $a, b \in \mathbb{R}$. La determinación del mínimo absoluto de ζ en $\mathbb{R} \times \mathbb{R}$ puede realizarse ajustando dos veces por el método de mínimos-cuadrados dos líneas paralelas: una vez para $(\text{mid } X, \text{mid } Y)$ y $(2 \sigma_{f_{[0,1]}} \text{spr } X, 2 \sigma_{f_{[0,1]}} \text{spr } Y)$ (que será válida cuando la pendiente a sea no negativa), y otra vez para $(\text{mid } X, \text{mid } Y)$ y $(2 \sigma_{f_{[0,1]}} [-\text{spr } X], 2 \sigma_{f_{[0,1]}} \text{spr } Y)$ (que será válida cuando la pendiente a sea negativa). En esta aproximación del problema, se utilizarían las técnicas características del Álgebra Lineal para hallar las soluciones óptimas y calcular las sumas de cuadrados de los residuos.

En contraste con la solución óptima para el problema de regresión con variables aleatorias reales, merece destacarse la existencia de un rasgo distintivo: en las condiciones supuestas en el Teorema 1.1.1, la “pendiente” de la relación lineal nunca puede anularse. De hecho, se tiene que:

Proposición 1.1.2 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, y X no es degenerado en un valor real, entonces,*

$$a^* = 0 \Rightarrow \text{para casi todo } \omega : X(\omega) \in \mathbb{R} \text{ ó } Y(\omega) \in \mathbb{R}.$$

La condición necesaria que se acaba de establecer en la Proposición 1.1.2 no es suficiente para que $a_0^* = 0$. De este modo:

Contraejemplo 1.1.1 Sea (Ω, \mathcal{A}, P) el espacio de probabilidad definido de forma que:

$$\Omega = \{\omega_1, \omega_2\}, \mathcal{A} = \mathcal{P}(\Omega), P(\omega_1) = P(\omega_2) = 0.5$$

y sean X e Y dos conjuntos aleatorios asociados a ese espacio y dados por:

$$X(\omega_1) = [1, 2], X(\omega_2) = \{3\},$$

$$Y(\omega_1) = \{1\}, Y(\omega_2) = [2, 3].$$

Los conjuntos X e Y no toman simultáneamente valores en $\mathcal{K}_c^*(\mathbb{R})$ c.s. $[P]$ y, sin embargo, supuesto que W es la medida de Lebesgue en $[0, 1]$, se cumple que $a_0^* = a_0 = a'_0 = 0.9310$.

Otro rasgo diferencial con el caso real es el relativo a las conexiones entre la independencia estocástica y la anulación de la “pendiente” de la relación óptima. Así, la *independencia estocástica* de dos conjuntos aleatorios X e Y (es decir, $P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$ cualesquiera que sean $A, B \in \mathcal{B}_{d_H}$) no implica necesariamente la anulación de a_0^* . Ello se debe básicamente al hecho de que $E(f_X|P) = f_{E(X|P)}$, por lo que

$$\text{Cov}(f_X, f_Y|P \otimes W) = E\left(\text{Cov}(f_X, f_Y|P)\middle|W\right) + \text{Cov}(f_{E(X|P)}, f_{E(Y|P)}|W)$$

(y también $\text{Var}(f_X|P \otimes W) = E\left(\text{Var}(f_X|P)\middle|W\right) + \text{Var}(f_{E(X|P)}|W)$), por lo que la independencia de X e Y implica que $\text{Cov}(f_X(\lambda), f_Y(\lambda)|P) = 0$ para todo $\lambda \in [0, 1]$, pero puede subyacer una dependencia entre las variables aleatorias reales $f_{E(X|P)}$ y $f_{E(Y|P)}$ asociadas al espacio de probabilidad $([0, 1], \mathcal{B}_{[0,1]}, W)$.

El siguiente resultado establece una condición necesaria y suficiente para que la independencia de X e Y obligue a que $a_0^* = 0$.

Proposición 1.1.3 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$ y estocásticamente independientes, $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, y X no es degenerado en un valor real, entonces,*

$$a^* = 0 \Leftrightarrow \text{para casi todo } \omega : X(\omega) \in \mathbb{R} \text{ ó } Y(\omega) \in \mathbb{R}.$$

Demostración. En efecto, si X e Y son estocásticamente independientes, entonces, para cualquier $\lambda \in [0, 1]$ se tiene que $\text{mid } X$ y $\text{mid } Y$ son variables aleatorias reales asociadas con el mismo espacio de probabilidad y estocásticamente independientes, por lo cual $\text{Cov}(\text{mid } X, \text{mid } Y|P) = 0$. En consecuencia, sobre la base de la Proposición 1.1.2 puede afirmarse que $a_0 = a'_0 = 0$ si, y sólo si, $E(\text{spr } X \cdot \text{spr } Y|P) = 0$ para lo cual es condición necesaria y suficiente que o bien X o bien Y sea variable aleatoria real *c.s.* $[P]$. \square

Al igual que en el caso real, la independencia estocástica de X e Y no es condición necesaria para asegurar que $a_0^* = 0$, incluso si X e Y no se reducen a variables reales, como se comprueba a continuación:

Contraejemplo 1.1.2 Sea (Ω, \mathcal{A}, P) el espacio de probabilidad definido como sigue:

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}, \mathcal{A} = \mathcal{P}(\Omega),$$

$$P(\omega_1) = P(\omega_2) = P(\omega_3) = P(\omega_4) = 0.25,$$

y sean X e Y dos conjuntos aleatorios asociados y dados por:

$$X(\omega_1) = \{-1\}, X(\omega_2) = X(\omega_3) = [0, 2], X(\omega_4) = \{3\},$$

$$Y(\omega_1) = Y(\omega_4) = [2, 4], Y(\omega_2) = \{1\}, Y(\omega_3) = \{5\}.$$

En este caso, supuesto que W es la medida de Lebesgue en $[0, 1]$, se satisface que $a_0^* = 0$, pero X e Y no son independientes.

Observación 1.1.2 Como en el modelo de relación lineal que se ha considerado entre X e Y los parámetros a y b son reales, si Y es no degenerado el valor paramétrico $a = 0$ es “inadmisible”, mientras que si se permite que Y sea degenerado, entonces sólo podría llegarse a que $a = 0$ en el caso en que Y fuera un conjunto aleatorio degenerado en el valor real b . Por ello, no resulta sorprendente que la “pendiente” a_0^* no pueda nunca anularse si Y no es degenerado, incluso en el caso en que X e Y sean independientes. Se verificará en la subsección siguiente que cuando la “ordenada en el origen” toma valores en $\mathcal{K}_c(\mathbb{R})$ la independencia garantizará la anulación de la “pendiente”.

Para concluir esta subsección, se señala un último elemento diferencial con respecto al caso en que X e Y son variables aleatorias reales: si X e Y

son degeneradas y X no es variable aleatoria real *c.s.* $[P]$, entonces, existe una solución única (a_0^*, b_0^*) con $a_0^* = 0$ si, y sólo si, Y es degenerada en un valor real.

1.1.2 Relación lineal de Tipo 2: coeficiente B con valores en $\mathcal{K}_c(\mathbb{R})$

Sea (Ω, \mathcal{A}, P) un espacio de probabilidad, y sean $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ e $Y : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ dos conjuntos aleatorios no degenerados asociados a ese espacio.

Si se observa el conjunto $X(\omega)$ y se quiere “estimar” el valor correspondiente $Y(\omega)$, puede tratar de aproximarse Y mediante una función afín de X (o, de forma más general, de $g(X)$, donde $g : \mathbb{R} \rightarrow \mathbb{R}$ es una función medible bien definida).

La Figura 1.2 muestra la representación gráfica de la estimación $\hat{Y}(\omega) = aX(\omega) + B$ (con $a \in \mathbb{R}$, $B \in \mathcal{K}_c(\mathbb{R})$).

Al igual que en el problema anterior, la búsqueda de la mejor relación de Y como función afín de X , se va a llevar a cabo mediante la aplicación del *método de mínimos-cuadrados* basada en la W -distancia, de modo que la cuestión que ahora se aborda es la determinación del mínimo absoluto de la función objetivo $\phi : \mathbb{R} \times \mathcal{K}_c(\mathbb{R}) \rightarrow [0, +\infty)$ tal que

$$\phi(a, B) = E\left([d_W(Y, aX + B)]^2 \mid P\right)$$

para $a \in \mathbb{R}$, $B \in \mathcal{K}_c(\mathbb{R})$.

En virtud de la definición del producto de elementos de $\mathcal{K}_c(\mathbb{R})$ por un escalar, hallar el mínimo absoluto de ϕ equivale a hallar el mínimo absoluto de una función objetivo alternativa expresada en términos de valores esperados sobre el espacio de probabilidad $(\Omega \times [0, 1], \mathcal{A} \otimes \mathcal{B}_{[0,1]}, P \otimes W)$.

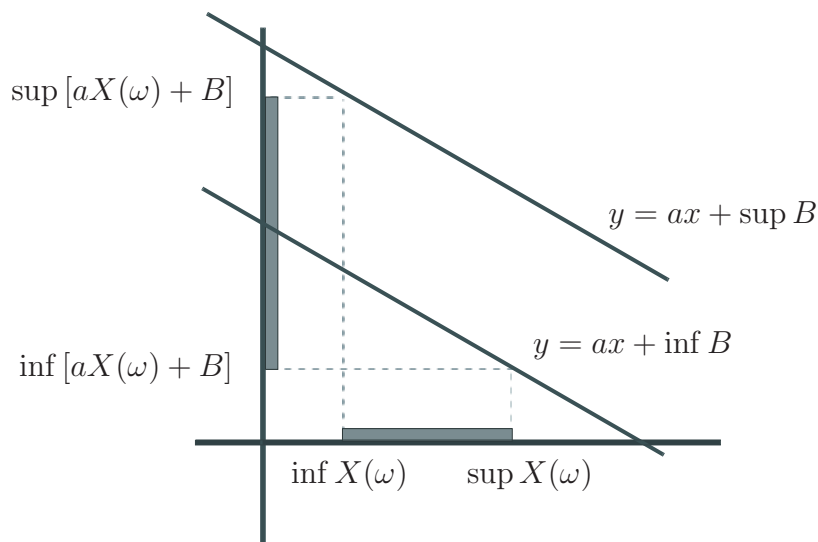


Fig. 1.2: Representación gráfica de $\hat{Y}(\omega) = aX(\omega) + B$

La Figura 1.2 muestra la representación gráfica de la estimación $\hat{Y}(\omega) = aX(\omega) + B$ (con $a \in \mathbb{R}$, $B \in \mathcal{K}_c(\mathbb{R})$).

Al igual que en el problema anterior, la búsqueda de la mejor relación de Y como función afín de X , se va a llevar a cabo mediante la aplicación del *método de mínimos-cuadrados* basada en la W -distancia, de modo que la cuestión que ahora se aborda es la determinación del mínimo absoluto de la función objetivo $\phi : \mathbb{R} \times \mathcal{K}_c(\mathbb{R}) \rightarrow [0, +\infty)$ tal que

$$\phi(a, B) = E\left([d_W(Y, aX + B)]^2 \mid P\right)$$

para $a \in \mathbb{R}$, $B \in \mathcal{K}_c(\mathbb{R})$.

En virtud de la definición del producto de elementos de $\mathcal{K}_c(\mathbb{R})$ por un escalar, hallar el mínimo absoluto de ϕ equivale a hallar el mínimo absoluto de una función objetivo alternativa expresada en términos de valores esperados sobre el espacio de probabilidad $(\Omega \times [0, 1], \mathcal{A} \otimes \mathcal{B}_{[0,1]}, P \otimes W)$.

Esta **función objetivo** alternativa viene dada por $\psi : \mathbb{R} \times [0, +\infty) \times \mathbb{R} \rightarrow [0, +\infty)$ tal que

$$\psi(a, b, c) = \begin{cases} E \left([f_Y - (af_X + bf_{[0,1]} + c)]^2 \mid P \otimes W \right) & \text{si } a \geq 0 \\ E \left([f_Y - (a\bar{f}_X + bf_{[0,1]} + c)]^2 \mid P \otimes W \right) & \text{si } a \leq 0 \end{cases}$$

con $b = 2 \operatorname{spr} B \in [0, +\infty)$ y $c = \inf B \in \mathbb{R}$.

En particular, si consideramos una muestra aleatoria de n datos bidimensionales, $(X_1, Y_1), \dots, (X_n, Y_n)$, con $X_i, Y_i \in \mathcal{K}_c(\mathbb{R})$ ($i = 1, \dots, n$), se cumple que:

$$\psi(a, b, c) = \begin{cases} \frac{1}{n} \sum_{i=1}^n \int_{[0,1]} \left(f_{Y_i}(\lambda) - af_{X_i}(\lambda) - b\lambda - c \right)^2 dW(\lambda) & \text{si } a \geq 0 \\ \frac{1}{n} \sum_{i=1}^n \int_{[0,1]} \left(f_{Y_i}(\lambda) - af_{X_i}(1-\lambda) - b\lambda - c \right)^2 dW(\lambda) & \text{si } a \leq 0. \end{cases}$$

En condiciones muy generales, el problema planteado admite solución (como se verá más tarde, no siempre única). Más concretamente:

Teorema 1.1.4 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$ y $E(|X| \mid P) < \infty$, $E(|Y| \mid P) < \infty$, entonces, existe solución para el problema, y la función objetivo ψ alcanza mínimo absoluto en $(a^*, b^*, c^*) \in \mathbb{R} \times [0, +\infty) \times \mathbb{R}$, con:*

$$(a^*, b^*, c^*) = \begin{cases} (a_1, b_1, c_1) & \text{en el CASO 1} \\ (a_2, b_2, c_2) & \text{en el CASO 2} \\ (a'_1, 0, c'_1) & \text{en el CASO 3} \\ (a'_2, 0, c'_2) & \text{en el CASO 4} \\ (0, b'', c'') & \text{en el CASO 5,} \end{cases}$$

es decir, ϕ alcanza mínimo absoluto en $(a^*, B^*) \in \mathbb{R} \times \mathcal{K}_c(\mathbb{R})$ tal que

$$(a^*, B^*) = \begin{cases} (a_1, [c_1, c_1 + b_1]) & \text{en el CASO 1} \\ (a_2, [c_2, c_2 + b_2]) & \text{en el CASO 2} \\ (a'_1, \{c'_1\}) & \text{en el CASO 3} \\ (a'_2, \{c'_2\}) & \text{en el CASO 4} \\ (0, E^A[Y|P]) & \text{en el CASO 5,} \end{cases}$$

donde:

$$\begin{aligned} a_1 &= \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2}} \cdot \frac{\rho_{f_X f_Y} - \rho_{f_X f_{[0,1]}} \rho_{f_Y f_{[0,1]}}}{1 - \rho_{f_X f_{[0,1]}}^2} \\ &= \frac{\text{Cov}(\text{mid } X, \text{mid } Y|P) + 4 \sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P)}{\sigma_{f_X}^2 (1 - \rho_{f_X f_{[0,1]}}^2)}, \\ a_2 &= \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2}} \cdot \frac{\rho_{\bar{f}_X f_Y} + \rho_{f_X f_{[0,1]}} \rho_{f_Y f_{[0,1]}}}{1 - \rho_{f_X f_{[0,1]}}^2} \\ &= \frac{\text{Cov}(\text{mid } X, \text{mid } Y|P) - 4 \sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P)}{\sigma_{f_X}^2 (1 - \rho_{f_X f_{[0,1]}}^2)}, \\ b_1 &= \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_{[0,1]}}^2}} \cdot \frac{\rho_{f_Y f_{[0,1]}} - \rho_{f_X f_{[0,1]}} \rho_{f_X f_Y}}{1 - \rho_{f_X f_{[0,1]}}^2}, \\ b_2 &= \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_{[0,1]}}^2}} \cdot \frac{\rho_{f_Y f_{[0,1]}} + \rho_{f_X f_{[0,1]}} \rho_{\bar{f}_X f_Y}}{1 - \rho_{f_X f_{[0,1]}}^2}, \end{aligned}$$

$$c_1 = \mu_{f_Y} - a_1 \mu_{f_X} - .5 b_1, \quad c_2 = \mu_{f_Y} - a_2 \mu_{f_X} - .5 b_2,$$

$$a'_1 = \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2}} \cdot \rho_{f_X f_Y}, \quad a'_2 = \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2}} \cdot \rho_{\bar{f}_X f_Y},$$

$$c'_1 = \mu_{f_Y} - a'_1 \mu_{f_X}, \quad c'_2 = \mu_{f_Y} - a'_2 \mu_{f_X},$$

$$b'' = \sqrt{\frac{\sigma_{f_Y}^2}{\sigma_{f_{[0,1]}}^2}} \cdot \rho_{f_Y f_{[0,1]}} = E(\text{spr } Y | P),$$

$$c'' = \mu_{f_Y} - .5 b'' = E(\text{inf } Y | P),$$

y los CASOS 1-5 corresponden a:

- CASO 1 tiene lugar si, y sólo si, ocurre una de las situaciones siguientes:
 - $a_1 > 0, a_2 \geq 0, b_1 > 0,$
 - $a_1^2 \geq a_2^2, a_1 > 0, a_2 < 0, b_1 > 0, b_2 > 0,$
 - $a_1^2 [1 - \rho_{f_X f_{[0,1]}}^2] \geq a_2'^2 - \sigma_{f_Y}^2 \rho_{f_Y f_{[0,1]}}^2 / \sigma_{f_X}^2,$
 - y $a_1 > 0, a_2 < 0, b_1 > 0, b_2 \leq 0;$
- CASO 2 tiene lugar si, y sólo si, ocurre una de las situaciones siguientes:
 - $a_1 \leq 0, a_2 < 0, b_2 > 0,$
 - $a_1^2 \leq a_2^2, a_1 > 0, a_2 < 0, b_1 > 0, b_2 > 0,$
 - $a_2^2 [1 - \rho_{f_X f_{[0,1]}}^2] \geq a_1'^2 - \sigma_{f_Y}^2 \rho_{f_Y f_{[0,1]}}^2 / \sigma_{f_X}^2,$
 - y $a_1 > 0, a_2 < 0, b_1 \leq 0, b_2 \leq 0;$

- CASO 3 *tiene lugar si, y sólo si, ocurre una de las situaciones siguientes:*

$$\begin{aligned}
 & - a_1 > 0, a_2 \geq 0, b_1 \leq 0, \\
 & - a_1'^2 \geq a_2'^2, a_1 > 0, a_2 < 0, b_1 \leq 0, b_2 > 0, \\
 & - a_2^2[1 - \rho_{f_X f_{[0,1]}}^2] \leq a_1'^2 - \sigma_{f_Y}^2 \rho_{f_Y f_{[0,1]}}^2 / \sigma_{f_X}^2, \\
 & \quad y a_1 > 0, a_2 < 0, b_1 \leq 0, b_2 \leq 0;
 \end{aligned}$$

- CASO 4 *tiene lugar si, y sólo si, ocurre una de las situaciones siguientes:*

$$\begin{aligned}
 & - a_1 \leq 0, a_2 < 0, b_2 \leq 0, \\
 & - a_1'^2 \leq a_2'^2, a_1 > 0, a_2 < 0, b_1 \leq 0, b_2 > 0, \\
 & - a_1^2[1 - \rho_{f_X f_{[0,1]}}^2] \leq a_2'^2 - \sigma_{f_Y}^2 \rho_{f_Y f_{[0,1]}}^2 / \sigma_{f_X}^2, \\
 & \quad y a_1 > 0, a_2 < 0, b_1 > 0, b_2 \leq 0;
 \end{aligned}$$

- CASO 5 *tiene lugar si, y sólo si, $a_1 \leq 0 \leq a_2$.*

Demostración. Si se emplea la función objetivo ψ , el problema que se desea resolver puede contemplarse como una combinación de dos problemas de regresión lineal múltiple clásicos, el primero de los cuales involucra las variables aleatorias reales f_X , $f_{[0,1]}$ y f_Y (si $a \geq 0$), y el segundo referido a las variables \bar{f}_X , $f_{[0,1]}$ y f_Y (si $a \leq 0$).

Los estudios de regresión lineal múltiple entre tres variables aleatorias reales nos permiten asegurar que la función

$$\psi_1(a, b, c) = E \left([f_Y - (af_X + bf_{[0,1]} + c)]^2 \mid P \otimes W \right)$$

alcanza su mínimo absoluto en \mathbb{R}^3 para (a_1, b_1, c_1) , en $\mathbb{R} \times \{0\} \times \mathbb{R}$ (i.e., supuesto $b = 0$) para $(a'_1, 0, c'_1)$, y en $\{0\} \times \mathbb{R} \times \mathbb{R}$ (i.e., supuesto $a = 0$)

para $(0, b'', c'')$. En consecuencia, denotando por (a_1^*, b_1^*, c_1^*) la solución que proporciona el óptimo de ψ_1 , se concluye que:

- a_1 y b_1 no pueden ser ambos negativos, puesto que si fuera así se tendría que $\rho_{f_Y f_{[0,1]}} < \rho_{f_X f_{[0,1]}} \rho_{f_X f_Y} < \rho_{f_X f_{[0,1]}}^2 \rho_{f_Y f_{[0,1]}} \leq \rho_{f_Y f_{[0,1]}}$, lo que conduce a una contradicción;
- si $a_1 \geq 0$ y $b_1 \geq 0$, entonces $(a_1^*, b_1^*, c_1^*) = (a_1, b_1, c_1)$;
- si $a_1 \geq 0$ y $b_1 < 0$, entonces $(a_1^*, b_1^*, c_1^*) = (a_1', 0, c_1')$ ya que $a_1 \geq 0$ implica que $\rho_{f_X f_Y} \geq \rho_{f_X f_{[0,1]}} \rho_{f_Y f_{[0,1]}} \geq 0$, de donde $a_1' \geq 0$;
- si $a_1 < 0$ y $b_1 \geq 0$, entonces $(a_1^*, b_1^*, c_1^*) = (0, b'', c'')$ ya que $b'' \geq 0$.

Por otro lado, también puede asegurarse que la función

$$\psi_2(a, b, c) = E \left([f_Y - (a \bar{f}_X + b f_{[0,1]} + c)]^2 \mid P \otimes W \right)$$

alcanza su mínimo absoluto en \mathbb{R}^3 para (a_2, b_2, c_2) , en $\mathbb{R} \times \{0\} \times \mathbb{R}$ para $(a_2', 0, c_2')$, y en $\{0\} \times \mathbb{R} \times \mathbb{R}$ para $(0, b'', c'')$. En consecuencia, denotando por (a_2^*, b_2^*, c_2^*) la solución que proporciona el óptimo de ψ_2 , se concluye que:

- no puede ocurrir que $a_2 > 0$ y $b_2 < 0$, puesto que si fuera así se tendría que $\rho_{f_Y f_{[0,1]}} < -\rho_{f_X f_{[0,1]}} \rho_{\bar{f}_X f_Y} < \rho_{f_X f_{[0,1]}}^2 \rho_{f_Y f_{[0,1]}} \leq \rho_{f_Y f_{[0,1]}}$, lo que conduce a una contradicción;
- si $a_2 \leq 0$ y $b_2 \geq 0$, entonces $(a_2^*, b_2^*, c_2^*) = (a_2, b_2, c_2)$;
- si $a_2 \leq 0$ y $b_2 < 0$, entonces $(a_2^*, b_2^*, c_2^*) = (a_2', 0, c_2')$ ya que $a_2 \leq 0$ implica que $\rho_{\bar{f}_X f_Y} \leq -\rho_{f_X f_{[0,1]}} \rho_{f_Y f_{[0,1]}} \leq 0$, de donde $a_2' \leq 0$;
- si $a_2 > 0$ y $b_2 \geq 0$, entonces $(a_2^*, b_2^*, c_2^*) = (0, b'', c'')$.

En resumen, el valor de $\psi(a, b, c)$ en los “posibles valores” en los que puede alcanzarse el óptimo será:

- Si $a_1 \geq 0$:

$$\begin{aligned}\psi(a_1, b_1, c_1) &= \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2] - a_1^2 \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2], \\ \psi(a'_1, 0, c'_1) &= \sigma_{f_Y}^2 \left[1 - a_1'^2 \frac{\sigma_{f_X}^2}{\sigma_{f_Y}^2} \right].\end{aligned}$$

- Si $a_2 \leq 0$:

$$\begin{aligned}\psi(a_2, b_2, c_2) &= \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2] - a_2^2 \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2], \\ \psi(a'_2, 0, c'_2) &= \sigma_{f_Y}^2 \left[1 - a_2'^2 \frac{\sigma_{f_X}^2}{\sigma_{f_Y}^2} \right].\end{aligned}$$

- Si $a_1 < 0$ y $a_2 > 0$:

$$\psi(0, b'', c'') = \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2].$$

Se satisface también que:

$$\begin{aligned}a_1 - a_2 &= 2 \frac{\sigma_{f_{[0,1]}}^2}{\sigma_{f_{[0,1]}}^2 (1 - \rho_{f_X f_{[0,1]}}^2)} \cdot \text{Cov}(\text{spr } X, \text{spr } Y|P), \\ b_1 - b_2 &= -2 \frac{\rho_{f_X f_{[0,1]}}}{\sigma_{f_{[0,1]}}^2 (1 - \rho_{f_X f_{[0,1]}}^2)} \cdot \text{Cov}(\text{mid } X, \text{mid } Y|P).\end{aligned}$$

Obsérvese que si $a_1 = a_2 = 0$, entonces $b_1 = b_2 = b'' \geq 0$ y $c_1 = c_2 = c''$, y que $b_1 = b_2 = 0$ implicaría que $a_1 = a'_1$, $a_2 = a'_2$, $c_1 = c'_1$ y $c_2 = c'_2$.

En consecuencia, el valor mínimo de $\psi(a, b, c)$ o (*variación residual*), $\psi(a^*, b^*, c^*) = \min \{ \psi(a_1^*, b_1^*, c_1^*), \psi(a_2^*, b_2^*, c_2^*) \}$, viene dado por:

- $\psi(a^*, b^*, c^*) = \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2] - a_1^2 \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2]$
si $a_1 \geq 0, a_2 \geq 0, b_1 > 0$;
- $\psi(a^*, b^*, c^*) = \sigma_{f_Y}^2 \left[1 - a_1^2 \frac{\sigma_{f_X}^2}{\sigma_{f_Y}^2} \right]$
si $a_1 \geq 0, a_2 \geq 0, b_1 \leq 0$;
- $\psi(a^*, b^*, c^*) = \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2] - a_2^2 \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2]$
si $a_1 \leq 0, a_2 \leq 0, b_2 > 0$;
- $\psi(a^*, b^*, c^*) = \sigma_{f_Y}^2 \left[1 - a_2^2 \frac{\sigma_{f_X}^2}{\sigma_{f_Y}^2} \right]$
si $a_1 \leq 0, a_2 \leq 0, b_2 \leq 0$;
- $\psi(a^*, b^*, c^*) = \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2]$
si $a_1 < 0, a_2 > 0$;
- $\psi(a^*, b^*, c^*) = \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2] - \max\{a_1^2, a_2^2\} \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2]$
si $a_1 > 0 > a_2, b_1 > 0, b_2 > 0$;
- $\psi(a^*, b^*, c^*) = \sigma_{f_Y}^2 \left[1 - \max\{a_1^2, a_2^2\} \frac{\sigma_{f_X}^2}{\sigma_{f_Y}^2} \right]$
si $a_1 > 0 > a_2, b_1 \leq 0, b_2 \leq 0$;
- $\psi(a^*, b^*, c^*) = \sigma_{f_Y}^2 \left[1 - \max \left\{ \rho_{f_Y f_{[0,1]}}^2 + a_1^2 \frac{\sigma_{f_X}^2}{\sigma_{f_Y}^2} [1 - \rho_{f_X f_{[0,1]}}^2], a_2^2 \frac{\sigma_{f_X}^2}{\sigma_{f_Y}^2} \right\} \right]$
si $a_1 > 0 > a_2, b_1 > 0, b_2 \leq 0$;
- $\psi(a^*, b^*, c^*) = \sigma_{f_Y}^2 \left[1 - \max \left\{ a_1^2 \frac{\sigma_{f_X}^2}{\sigma_{f_Y}^2}, \rho_{f_Y f_{[0,1]}}^2 + a_2^2 \frac{\sigma_{f_X}^2}{\sigma_{f_Y}^2} [1 - \rho_{f_X f_{[0,1]}}^2] \right\} \right]$
si $a_1 > 0 > a_2, b_1 \leq 0, b_2 > 0$.

La comparación dos a dos de estas variaciones residuales nos lleva a la determinación de la solución (a^*, b^*, c^*) en la que ψ alcanza el mínimo absoluto, y que se detalla en el enunciado del presente teorema. \square

El estudio de las variaciones residuales en la demostración del Teorema 1.1.4, nos permite obtener las conclusiones siguientes acerca de la unicidad de la solución óptima:

Teorema 1.1.5 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$ y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, entonces, existe solución única $(a^*, b^*, c^*) \in \mathbb{R} \times [0, +\infty) \times \mathbb{R}$ para el problema, siempre que no se cumpla que $a_1 > 0 > a_2$, y viene dada por:*

$$(a^*, b^*, c^*) = \begin{cases} (a_1, b_1, c_1) & \text{si } a_1 \geq 0, a_2 \geq 0, b_1 > 0, \\ (a'_1, 0, c'_1) & \text{si } a_1 \geq 0, a_2 \geq 0, b_1 \leq 0, \\ (a_2, b_2, c_2) & \text{si } a_1 \leq 0, a_2 \leq 0, b_2 > 0, \\ (a'_2, 0, c'_2) & \text{si } a_1 \leq 0, a_2 \leq 0, b_2 \leq 0, \\ (0, b'', c'') & \text{si } a_2 > 0 > a_1. \end{cases}$$

Cuando $a_1 > 0 > a_2$, la solución óptima no es necesariamente única, pero para los casos en los que la solución no es única se obtienen dos posibles (una en el cono correspondiente a $a > 0$ y otra en el cono correspondiente a $a < 0$). El siguiente resultado incluye una caracterización de tales casos.

Teorema 1.1.6 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$ y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, entonces, es condición necesaria y suficiente para que existan dos soluciones distintas para el problema que ocurra una de las situaciones siguientes:*

- $a_1 + a_2 = 0$, $b_1 = b_2$, $c_2 = c_1 + 2a_1\mu_{f_X}$ si $a_1 > 0, b_1 > 0$ and $a_2 < 0, b_2 > 0$;
- $a'_1 + a'_2 = 0$, $c'_2 = c'_1 + 2a'_1\mu_{f_X}$ (en cuyo caso $b'_1 = b'_2 = 0$) si $a_1 > 0, b_1 \leq 0$ y $a_2 < 0, b_2 \leq 0$;
- $a_2'^2 - a_1^2[1 - \rho_{f_X f_{[0,1]}}^2] = \frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2} \rho_{f_Y f_{[0,1]}}^2$ si $a_1 > 0, b_1 > 0$ y $a_2 < 0, b_2 \leq 0$;
- $a_1'^2 - a_2^2[1 - \rho_{f_X f_{[0,1]}}^2] = \frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2} \rho_{f_Y f_{[0,1]}}^2$ si $a_1 > 0, b_1 \leq 0$ y $a_2 < 0, b_2 > 0$.

Para ilustrar la existencia de situaciones en las que puede aparecer duplicidad de soluciones óptimas, se considera el siguiente:

Ejemplo 1.1.2 Sea (Ω, \mathcal{A}, P) el espacio de probabilidad definido como sigue:

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}, \mathcal{A} = \mathcal{P}(\Omega),$$

$$P(\omega_1) = P(\omega_2) = P(\omega_3) = P(\omega_4) = 0.25,$$

y sean X e Y dos conjuntos aleatorios asociados, dados por:

$$X(\omega_1) = [0, 2], X(\omega_2) = X(\omega_3) = [2, 6], X(\omega_4) = [6, 8],$$

$$Y(\omega_1) = Y(\omega_4) = [4, 6], Y(\omega_2) = [-1, 3], Y(\omega_3) = [7, 11].$$

Si W es la medida de Lebesgue en $[0, 1]$, se cumple que $a_1 = -a_2 = 1/46$, $b_1 = b_2 = 135/46$ y $c_1 = 317/92$, $c_2 = 333/92$, lo que conduce a la primera de las condiciones del Teorema 1.1.6.

La Figura 1.3 muestra el diagrama de los datos bidimensionales (rectangulares), junto con las correspondientes relaciones óptimas del Tipo 2:

$$\widehat{Y} = \frac{1}{46}X + \left[\frac{317}{92}, \frac{587}{92} \right], \quad \widehat{Y} = -\frac{1}{46}X + \left[\frac{333}{92}, \frac{603}{92} \right].$$

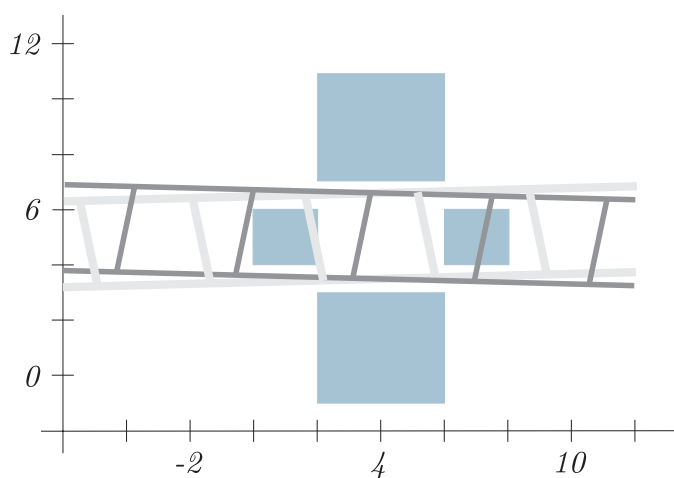


Fig. 1.3: Diagrama de “rectángulos” del Ejemplo 1.1.2 y soluciones óptimas del tipo $\widehat{Y}(\omega) = aX(\omega) + B$

En la práctica la falta de unicidad no es usualmente conveniente, por lo que en la mayoría de las situaciones que se han caracterizado en el Teorema 1.1.6, se recurriría a cambiar el tipo de relación funcional considerada (por ejemplo, reemplazando X por $g(X)$ para alguna función g adecuada).

Observación 1.1.3 De acuerdo con las notaciones utilizadas en esta sección, el caso examinado por Diamond (1990) y para el que probó la existencia de solución única y al que se refirió como caso en el que se involucran datos con valor de intervalo *coherentes* corresponde a $a_1 \geq a_2 \geq 0$ y $a_1 \leq a_2 \leq 0$.

Observación 1.1.4 El problema planteado en esta subsección podría presentarse también considerando la función objetivo:

$$\zeta(a, b, c) = E \left((\text{mid } Y - a \text{ mid } X - c)^2 + 4 \sigma_{f_{[0,1]}}^2 (\text{spr } Y - |a| \text{ spr } X - b)^2 \mid P \right)$$

con $a \in \mathbb{R}$, $b = \text{spr } B \in [0, +\infty)$, y $c = \text{mid } B \in \mathbb{R}$. La determinación del mínimo absoluto de ζ en $\mathbb{R} \times [0, +\infty) \times \mathbb{R}$ puede realizarse ajustando dos veces por el método de mínimos-cuadrados dos líneas paralelas: una vez para $(\text{mid } X, \text{mid } Y)$ y $(2 \sigma_{f_{[0,1]}} \text{ spr } X, 2 \sigma_{f_{[0,1]}} \text{ spr } Y)$, y otra vez para $(\text{mid } X, \text{mid } Y)$ y $(2 \sigma_{f_{[0,1]}} [-\text{spr } X], 2 \sigma_{f_{[0,1]}} \text{ spr } Y)$. Se recurriría después a las técnicas características del Álgebra Lineal para hallar las soluciones óptimas y los valores de ζ mínimos.

A diferencia del modelo de relación de la subsección precedente, en las condiciones supuestas en el Teorema 1.1.4, la “pendiente” de la relación lineal puede anularse. De hecho, se tiene que:

Proposición 1.1.8 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$ y $E(|X| \mid P) < \infty$, $E(|Y| \mid P) < \infty$, entonces, es condición necesaria y suficiente para que $a^* = 0$ (i.e., para que la predicción óptima de Y corresponda a un valor de intervalo constante) que $a_1 \leq 0 \leq a_2$ (es decir,*

$$\begin{aligned} & 4 \sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y \mid P) \\ & \leq \text{Cov}(\text{mid } X, \text{mid } Y \mid P) \\ & \leq -4 \sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y \mid P). \end{aligned}$$

Otro rasgo diferencial con respecto al problema de la subsección anterior, aunque similar al caso real, es el de la relación entre la independencia estocástica y la anulación de a^* de la relación óptima. Así, como la independencia de X e Y garantiza que $\text{mid } X$ y $\text{mid } Y$ son independientes, y también que lo son $\text{spr } X$ y $\text{spr } Y$, se deduce que:

Proposición 1.1.8 Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$ y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, entonces, si X e Y son estocásticamente independientes se cumple que $a^* = 0$.

Aunque la anulación de a^* es condición necesaria para la independencia, no se trata de una condición suficiente aun cuando X e Y no sean casi seguro reales. Así:

Contraejemplo 1.1.3 Sea (Ω, \mathcal{A}, P) el espacio de probabilidad definido como sigue:

$$\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}, \mathcal{A} = \mathcal{P}(\Omega),$$

$$P(\omega_1) = P(\omega_2) = P(\omega_3) = P(\omega_4) = 0.25,$$

y sean X e Y dos conjuntos aleatorios asociados y dados por:

$$X(\omega_1) = [-3, -1], X(\omega_2) = X(\omega_3) = [-1, 1], X(\omega_4) = [1, 3],$$

$$Y(\omega_1) = Y(\omega_4) = [-1, 1], Y(\omega_2) = [-3, -1], Y(\omega_3) = [1, 3].$$

Se cumple entonces, supuesto que W es la medida de Lebesgue en $[0, 1]$, que $a_1 = a_2 = 0$ y, por tanto, $a^* = 0$. Sin embargo X e Y no son independientes, ya que, por ejemplo:

$$P(X \in [-3, -1], Y \in [-3, -1]) = P(\emptyset) = 0$$

$$\neq 0.25 \cdot 0.25 = P(X \in [-3, -1]) \cdot P(Y \in [-3, -1]).$$

Observación 1.1.5 En algunos casos particulares las relaciones óptimas admiten expresiones muy sencillas, como sigue:

- Si $a_1 = a_2 = 0$ ó $a_2 > 0 > a_1$, entonces la relación óptima es:

$$\hat{Y} = [E(\inf Y | P), E(\sup Y | P)] = E^A[Y | P].$$

- Si $a_1 = a'_1 = 0$, $b_1 \leq 0$, debe ocurrir que $\rho_{f_X f_Y} = 0 = \rho_{f_X f_{[0,1]}} \rho_{f_Y f_{[0,1]}}$ y $\rho_{f_Y f_{[0,1]}} = 0$, lo que (ya que se ha supuesto que $\sigma_{f_Y}^2 \neq 0$) equivale a que Y sea real *c.s* $[P]$ y las variables aleatorias f_X y f_Y definidas sobre $(\Omega \times [0, 1], \mathcal{A} \otimes \mathcal{B}_{[0,1]}, P \otimes W)$ sean incorreladas. En este caso, la relación óptima corresponde a:

$$\hat{Y} = \text{mid } E^A[Y|P] = E(\text{mid } Y).$$

- Si $a_2 = a'_2 = 0$, $b_2 \leq 0$, debe ocurrir que $\rho_{\bar{f}_X f_Y} = 0 = -\rho_{f_X f_{[0,1]}} \rho_{f_Y f_{[0,1]}}$ y $\rho_{f_Y f_{[0,1]}} = 0$, lo que equivale a que Y sea real *c.s* $[P]$ y las variables aleatorias \bar{f}_X y f_Y sean incorreladas. La relación óptima coincide con la del caso precedente.

1.1.3 Algoritmo para la determinación de la relación lineal óptima de Tipo 2

Como apoyo para el cálculo práctico de la solución óptima de la relación de Tipo 2, más compleja que la del tipo 1, puede considerarse el algoritmo:

Paso 1. Calcular los coeficientes $a_1, a_2, a'_1, a'_2, b_1, b_2, c_1, c_2, c'_1, c'_2, b'', c'', \sigma_{f_X}^2, \sigma_{f_Y}^2, \rho_{f_X f_{[0,1]}}^2$ y $\rho_{f_Y f_{[0,1]}}^2$, e ir al Paso 2.

Paso 2. SI $a_1 \leq 0$ ENTONCES ir al Paso 3, EN CASO CONTRARIO ir al Paso 5.

Paso 3. SI $a_2 \geq 0$ ENTONCES la solución óptima corresponde a la del CASO 5, EN CASO CONTRARIO ir al Paso 4.

Paso 4. SI $b_2 > 0$ ENTONCES la solución óptima corresponde a la del CASO 2, EN CASO CONTRARIO la solución óptima corresponde a la del CASO 4.

Paso 5. SI $a_2 \geq 0$ ENTONCES ir al Paso 6, EN CASO CONTRARIO ir al Paso 7.

Paso 6. SI $b_1 > 0$ ENTONCES la solución óptima corresponde a la del CASO 1, EN CASO CONTRARIO la solución óptima corresponde a la del CASO 3.

Paso 7. SI $b_1 > 0$ ENTONCES ir al Paso 8, EN CASO CONTRARIO ir al Paso 11.

Paso 8. SI $b_2 > 0$ ENTONCES ir al Paso 9, EN CASO CONTRARIO ir al Paso 10.

Paso 9. SI $a_1^2 > a_2^2$ ENTONCES la solución óptima corresponde a la del CASO 1, EN CASO CONTRARIO

SI $a_1^2 < a_2^2$ ENTONCES la solución óptima corresponde a la del CASO 2,

EN CASO CONTRARIO existen dos soluciones óptimas: las que corresponden a los CASOS 1 y 2.

Paso 10. SI $a_1^2[1 - \rho_{f_X f_{[0,1]}}^2] > a_2'^2 - \frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2} \rho_{f_Y f_{[0,1]}}^2$ ENTONCES la solución óptima corresponde a la del CASO 1, EN CASO CONTRARIO

SI $a_1^2[1 - \rho_{f_X f_{[0,1]}}^2] < a_2'^2 - \frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2} \rho_{f_Y f_{[0,1]}}^2$ ENTONCES la solución óptima corresponde a la del CASO 4,

EN CASO CONTRARIO existen dos soluciones óptimas: la que corresponde al CASO 1 y la que corresponde al CASO 4.

Paso 11. SI $b_2 > 0$ ENTONCES ir al Paso 12, EN CASO CONTRARIO ir al Paso 13.

Paso 12. SI $a_1'^2 > a_2'^2$ ENTONCES la solución óptima corresponde a la del CASO 3, EN CASO CONTRARIO

SI $a_1'^2 < a_2'^2$ ENTONCES la solución óptima corresponde a la del CASO 4,

EN CASO CONTRARIO existen dos soluciones óptimas: la que corresponde al CASO 3 y la que corresponde al CASO 4.

Paso 13. SI $a_2^2[1 - \rho_{f_X f_{[0,1]}}^2] > a_1'^2 - \frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2} \rho_{f_Y f_{[0,1]}}^2$ ENTONCES la solución óptima corresponde a la del CASO 2, EN CASO CONTRARIO

SI $a_2^2[1 - \rho_{f_X f_{[0,1]}}^2] < a_1'^2 - \frac{\sigma_{f_Y}^2}{\sigma_{f_X}^2} \rho_{f_Y f_{[0,1]}}^2$ ENTONCES la solución óptima corresponde a la del CASO 3,

EN CASO CONTRARIO existen dos soluciones óptimas: la que corresponde al CASO 2 y la que corresponde al CASO 3.

1.1.4 Aplicación sobre un ejemplo real

El siguiente ejemplo ilustra en una situación real la aplicación de las conclusiones en las Subsecciones 1.1.1 y 1.1.2, sobre datos proporcionados por el Servicio de Nefrología del Hospital Valle del Nalón de Langreo (Asturias).

Ejemplo 1.1.3 Los datos ligados que aparecen en la Tabla 1.1 corresponden al “rango de la tensión arterial sistólica a lo largo de un día”, X , y “rango de la tensión arterial diastólica a lo largo del mismo día”, Y , en una población Ω de 59 pacientes hospitalizados.

Tabla 1.1: Datos sobre los rangos de la tensión arterial sistólica (X) y diastólica (Y)

X	Y	X	Y	X	Y
11.8-17.3	6.3-10.2	11.9-21.2	4.7-9.3	9.8-16.0	4.7-10.8
10.4-16.1	7.1-10.8	12.2-17.8	7.3-10.5	9.7-15.4	6.0-10.7
13.1-18.6	5.8-11.3	12.7-18.9	7.4-12.5	8.7-15.0	4.7-8.6
10.5-15.7	6.2-11.8	11.3-21.3	5.2-11.2	14.1-25.6	7.7-15.8
12.0-17.9	5.9-9.4	14.1-20.5	6.9-13.3	10.8-14.7	6.2-10.7
10.1-19.4	4.8-11.6	9.9-16.9	5.3-10.9	11.5-19.6	6.5-11.7
10.9-17.4	6.0-11.9	12.6-19.7	6.0-9.8	9.9-17.2	4.2-8.6
12.8-21.0	7.6-12.5	9.9-20.1	5.5-12.1	11.3-17.6	5.7-9.5
9.4-14.5	4.7-10.4	8.8-22.1	3.7-9.4	11.4-18.6	4.6-10.3
14.8-20.1	8.8-13.0	11.3-18.3	5.5-8.5	14.5-21.0	10.0-13.6
11.1-19.2	5.2-9.6	9.4-17.6	5.6-12.1	12.0-18.0	5.9-9.0
11.6-20.1	7.4-13.3	10.2-15.6	5.0-9.4	10.0-16.1	5.4-10.4
10.2-16.7	3.9-8.4	10.3-15.9	5.2-9.5	15.9-21.4	9.9-12.7
10.4-16.1	5.5-9.8	10.2-18.5	6.3-11.8	13.8-22.1	7.0-11.8
10.6-16.7	4.5-9.5	11.1-19.9	5.7-11.3	8.7-15.2	5.0-9.5
11.2-16.2	6.2-11.6	13.0-18.0	6.4-12.1	12.0-18.8	5.3-10.5
13.6-20.1	6.7-12.2	10.3-16.1	5.5-9.7	9.5-16.6	5.4-10.0
9.0-17.7	5.2-10.4	12.5-19.2	5.9-10.1	9.2-17.3	4.5-10.7
11.6-16.8	5.8-10.9	9.7-18.2	5.4-10.4	8.3-14.0	4.5-9.1
9.8-15.7	5.0-11.1	12.7-22.6	5.7-10.1		

Cada valor de X e Y se obtiene a partir de los distintos registros (habitualmente entre 60 y 70) de la tensión arterial realizados sobre cada paciente en distintos instantes de un mismo día.

La Figura 1.4 representa el diagrama de los “datos rectangulares” en el Ejemplo 1.1.3.

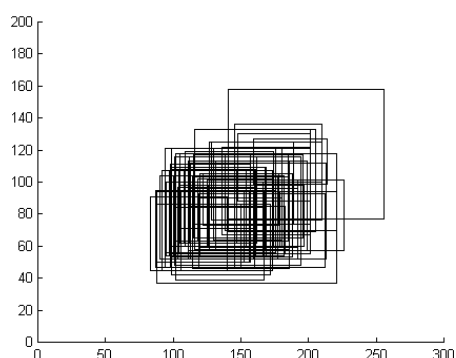


Fig. 1.4: Diagrama de los datos rectangulares en el Ejemplo 1.1.3

La relación lineal óptima del Tipo 1 de Y respecto de X , según el Teorema 1.1.1, cuando W es la medida de Lebesgue en $[0, 1]$, será $Y = 0.582 X - 0.2096$. Si esta relación se usara con fines de predicción, a un paciente con rango de tensión sistólica 11.5-16.1 le correspondería una predicción del rango de tensión diastólica igual a 6.5-9.2.

Por otra parte, la relación lineal óptima del Tipo 2 de Y respecto de X , según el Teorema 1.1.4, cuando W es la medida de Lebesgue en $[0, 1]$, será $Y = 0.4393 X + [0.9494, 2.8272]$. Si esta relación se usara con fines de predicción, a un paciente con rango de tensión sistólica 11.5-16.1 le correspondería una predicción del rango de tensión diastólica igual a 6.0-9.8.

1.2 Estudios descriptivos de correlación lineal entre dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$

Los estudios de *correlación* entre dos conjuntos aleatorios, van a referirse a la cuantificación del grado de “relación estadística” existente entre ellos. Se entiende que existe una “relación positiva” cuando los individuos que muestran “valores altos (bajos)” para uno de los conjuntos aleatorios tienden a mostrar “valores altos (bajos)” para el otro. Se entiende que existe una “relación negativa” cuando los individuos que muestran “valores altos (bajos)” para uno de los conjuntos aleatorios tienden a mostrar “valores bajos (altos)” para el otro. Alternativamente, se puede interpretar el objetivo de la correlación como de la valoración de la idoneidad con la que la función (del conjunto aleatorio independiente) óptima del tipo considerado (lineal, afín, etc.) explica la variación del conjunto aleatorio dependiente.

La definición de medidas del grado de idoneidad de una función lineal y de una función afín entre dos conjuntos aleatorios X e Y en la población o muestra considerada, va a inspirarse en la del coeficiente de determinación entre dos variables aleatorias reales, es decir, en la proporción de variación total que puede explicarse por la relación funcional.

En este sentido, sobre la base de las conclusiones obtenidas en la Sección 1.1, se determinará en primer lugar la *variación total* para los Tipos 1 y 2, que corresponde al valor máximo que podría alcanzar la *variación no explicada* (variación residual) en cada caso, y se calculará posteriormente la proporción de variación total que puede explicarse para los datos (pobla-cionales o muestrales) disponibles concretos, en cada uno de los dos tipos de relaciones.

1.2.1 Análisis de correlación para relaciones lineales de Tipo 1

Sea (Ω, \mathcal{A}, P) un espacio de probabilidad, y sean $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ e $Y : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ dos conjuntos aleatorios no degenerados asociados a ese espacio.

La variación residual en la relación de Tipo 1 es el mínimo absoluto de la función objetivo ϕ en \mathbb{R}^2 que, por aplicación del Teorema 1.1.1, viene dado por:

$$\begin{aligned} \phi(a_0^*, b_0^*) &= \begin{cases} (1 - \rho_{f_X f_Y}^2) \sigma_{f_Y}^2 & \text{si } \rho_{f_X f_Y}^2 > \rho_{\bar{f}_X \bar{f}_Y}^2, \\ (1 - \rho_{\bar{f}_X \bar{f}_Y}^2) \sigma_{f_Y}^2 & \text{en caso contrario} \end{cases} \\ &= \left[1 - \max \{ \rho_{f_X f_Y}^2, \rho_{\bar{f}_X \bar{f}_Y}^2 \} \right] \sigma_{f_Y}^2. \end{aligned}$$

Obviamente, cuanto mayor es el valor del $\max \{ \rho_{f_X f_Y}^2, \rho_{\bar{f}_X \bar{f}_Y}^2 \}$, menor es el valor de la variación residual (y, por tanto, mayor el de la variación explicada) por la relación lineal óptima de Tipo 1. De este modo, siguiendo las ideas del caso real, puede definirse:

Definición 1.2.1 Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, no degenerados y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, el **coeficiente de determinación en el sentido de la relación funcional de Tipo 1** de X e Y viene dado por el valor real:

$$\begin{aligned} r_{XY}^2 &= \max \{ \rho_{f_X f_Y}^2, \rho_{\bar{f}_X \bar{f}_Y}^2 \} \\ &= \frac{1}{\sqrt{\sigma_{f_X}^2 \sigma_{f_Y}^2}} \left\{ \left(\text{Cov}(\text{mid } X, \text{mid } Y | P) \right)^2 + 16 \sigma_{f_{[0,1]}}^4 \left(E(\text{spr } X \cdot \text{spr } Y | P) \right)^2 \right. \\ &\quad \left. + 8 \sigma_{f_{[0,1]}}^2 \left| \text{Cov}(\text{mid } X, \text{mid } Y | P) E(\text{spr } X \cdot \text{spr } Y | P) \right| \right\}. \end{aligned}$$

El rango de valores para el coeficiente adimensional r_{XY}^2 y la corroboración formal de que se trata de un índice para la proporción de variación total que puede explicarse por la relación óptima de Tipo 1, se recogen en el resultado siguiente:

Teorema 1.2.1 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, no degenerados y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, entonces:*

$$\begin{aligned} i) \quad r_{XY}^2 &= \frac{\text{variación total de } Y - \text{variación residual de Tipo 1 para } Y}{\text{variación total}} \\ &= \frac{\text{variación explicada por relación de Tipo 1}}{\text{variación total}}; \end{aligned}$$

$$ii) \quad 0 \leq r_{XY}^2 \leq 1.$$

Demostración. En efecto:

$$\text{variación residual de Tipo 1 para } Y = \phi(a_0^*, b_0^*) = [1 - r_{XY}^2] \sigma_{f_Y}^2,$$

por lo que

$$0 \leq \text{variación residual de Tipo 1 para } Y \leq \sigma_{f_Y}^2 = \text{variación total de } Y,$$

lo que prueba *i*).

Por otro lado, $0 \leq \phi(a_0^*, b_0^*) = [1 - r_{XY}^2] \sigma_{f_Y}^2$ garantiza que $1 - r_{XY}^2 \geq 0$, lo que demuestra *ii*). \square

Por lo que se refiere a los valores extremos del rango de r_{XY}^2 , en la proposición siguiente se exponen condiciones para alcanzarlos, si bien cabe reseñar que la variación total sólo sería alcanzable por la variación residual para conjuntos aleatorios y relaciones entre ellos muy especiales.

Proposición 1.2.2 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, no degenerados y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, entonces:*

- i) $r_{XY}^2 = 1$ si, y sólo si, existe $(a, b) \in \mathbb{R}^2$ tal que para casi todo $\omega \in \Omega$ se cumple que $Y = aX + b$. Más concretamente, $a \geq 0$ si, y sólo si, $\rho_{f_X f_Y}^2 = 1$ y $a \leq 0$ si, y sólo si, $\rho_{f_X f_Y}^2 = 1$;*
- ii) $r_{XY}^2 = 0$ si, y sólo si, las variables reales f_X y f_Y son incorreladas. Más concretamente:*
 - si X e Y toman valores c.s. $[P]$ en la clase $\mathcal{K}_c^*(\mathbb{R})$, entonces $r_{XY}^2 \neq 0$;*
 - cualesquiera que sean X e Y , es condición necesaria para que $r_{XY}^2 = 0$ que X e Y no tomen simultáneamente valores en $\mathcal{K}_c^*(\mathbb{R})$ c.s. $[P]$ (es decir, que para casi todo $\omega : X(\omega) \in \mathbb{R}$ ó $Y(\omega) \in \mathbb{R}$).*

Demostración.

- i) Si para cierto $(a, b) \in \mathbb{R}^2$ se cumple que $Y = aX + b$ casi seguro $[P]$, entonces, $\phi(a, b) = 0$, de modo que $[1 - r_{XY}^2] \sigma_{f_Y}^2 = 0$, y al ser $\sigma_{f_Y}^2 > 0$ en las condiciones supuestas, $r_{XY}^2 = 1$. Recíprocamente, $r_{XY}^2 = 1$ implica que $\phi(a_0^*, b_0^*) = [1 - r_{XY}^2] \sigma_{f_Y}^2 = 0$, es decir, que $d_W(Y, a_0^* X + b_0^*) = 0$ c.s. $[P]$, y por tanto $Y = a_0^* X + b_0^*$ casi seguro $[P]$.*

Además, si $Y = aX + b$ c.s. $[P]$, entonces, si $a \geq 0$ se cumple que:

$$\begin{aligned} \rho_{f_X f_Y} &= \frac{\text{Cov}(\text{mid } X, a \text{ mid } X + b | P) + 4 \sigma_{f_{[0,1]}}^2 E(\text{spr } X \cdot a \text{ spr } X | P)}{\sqrt{\sigma_{f_X}^2 a^2 \sigma_{f_X}^2}} \\ &= \frac{a \text{Var}(\text{mid } X | P) + 4a \sigma_{f_{[0,1]}}^2 E((\text{spr } X)^2 | P)}{a \sigma_{f_X}^2} = 1, \end{aligned}$$

mientras que si $a \leq 0$ se cumple que:

$$\begin{aligned} \rho_{\bar{f}_X f_Y} &= \frac{\text{Cov}(\text{mid } X, a \text{ mid } X + b | P) - 4 \sigma_{f_{[0,1]}}^2 E(\text{spr } X \cdot (-a \text{ spr } X) | P)}{\sqrt{\sigma_{f_X}^2 a^2 \sigma_{f_X}^2}} \\ &= \frac{a \text{Var}(\text{mid } X | P) + 4a \sigma_{f_{[0,1]}}^2 E((\text{spr } X)^2 | P)}{-a \sigma_{f_X}^2} = -1. \end{aligned}$$

- ii) Como $r_{XY}^2 = 0$ si, y sólo si, con las notaciones del Teorema 1.1.1 $a_0 = a'_0 = 0$ (lo que equivale a $a_0^* = 0$), las conclusiones de este apartado son triviales en virtud de la Proposición 1.1.2. \square

A partir de la Proposición 1.2.3 puede concluirse que la independencia estocástica no garantiza, en general, la anulación de r_{XY}^2 . Precizando más:

Proposición 1.2.3 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, no degenerados, estocásticamente independientes y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, entonces,*

$$r_{XY}^2 = 0 \Leftrightarrow \text{para casi todo } \omega : X(\omega) \in \mathbb{R} \text{ ó } Y(\omega) \in \mathbb{R}$$

El coeficiente de determinación en el sentido de la relación funcional de Tipo 1, satisface también las propiedades siguientes:

Proposición 1.2.4 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, no degenerados y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, entonces:*

- i) $r_{XY}^2 = r_{YX}^2$;
- ii) $r_{(aX+b)(cY+d)}^2 = r_{YX}^2$, cualesquiera que sean $a, b, c, d \in \mathbb{R}$.

1.2.2 Análisis de correlación para relaciones lineales de Tipo 2

Sea (Ω, \mathcal{A}, P) un espacio de probabilidad, y sean $X : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ e $Y : \Omega \rightarrow \mathcal{K}_c(\mathbb{R})$ dos conjuntos aleatorios no degenerados asociados a ese espacio.

La variación residual en la relación de Tipo 2 es el mínimo absoluto de la función objetivo ψ en $\mathbb{R} \times [0, +\infty) \times \mathbb{R}$ que, de acuerdo con la demostración del Teorema 1.1.4, y agrupando según los CASOS 1-5, viene dado por:

$$\psi(a^*, b^*, c^*) = \begin{cases} \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2] - a_1^2 \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2] & \text{en el CASO 1} \\ \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2] - a_2^2 \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2] & \text{en el CASO 2} \\ \sigma_{f_Y}^2 [1 - \rho_{f_X f_Y}^2] & \text{en el CASO 3} \\ \sigma_{f_Y}^2 [1 - \rho_{\bar{f}_X f_Y}^2] & \text{en el CASO 4} \\ 0 & \text{en el CASO 5.} \end{cases}$$

La “pendiente” óptima en las relaciones de Tipo 2 sólo se anula en unas pocas situaciones (ver Teorema 1.1.2), para las que la relación óptima corresponde a $\widehat{Y} = E^A[Y|P]$. En estas situaciones, la variación residual alcanza su valor máximo, que es igual a $\sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2]$, y que será la “variación total” de Y para este tipo de relaciones funcionales. Siguiendo las ideas del caso real y de la subsección precedente, puede definirse:

Definición 1.2.2 Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, no degenerados y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, el **coeficiente de determinación en el sentido de la relación funcional de Tipo 2** de X e Y viene dado por el valor real:

$$R_{XY}^2 = \begin{cases} \frac{a_1^2 \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2]}{\sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2]} & \text{en el CASO 1} \\ \frac{a_2^2 \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2]}{\sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2]} & \text{en el CASO 2} \\ \frac{\rho_{f_X f_Y}^2 - \rho_{f_Y f_{[0,1]}}^2}{1 - \rho_{f_Y f_{[0,1]}}^2} & \text{en el CASO 3} \\ \frac{\rho_{f_X f_Y}^2 - \rho_{f_Y f_{[0,1]}}^2}{1 - \rho_{f_Y f_{[0,1]}}^2} & \text{en el CASO 4} \\ 0 & \text{en el CASO 5.} \end{cases}$$

El rango de valores para el coeficiente adimensional R_{XY}^2 y la corroboración formal de que se trata de un índice para la proporción de variación total que puede explicarse por la relación óptima de Tipo 2, se recogen en el resultado siguiente:

Teorema 1.2.5 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, no degenerados y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, entonces:*

$$\begin{aligned} i) \quad R_{XY}^2 &= \frac{\text{variación total de } Y - \text{variación residual de Tipo 2 para } Y}{\text{variación total}} \\ &= \frac{\text{variación explicada por relación de Tipo 2}}{\text{variación total}}; \end{aligned}$$

$$ii) \quad 0 \leq R_{XY}^2 \leq 1.$$

Demostración. En efecto:

$$\begin{aligned} \text{variación residual de Tipo 2 para } Y &= \psi(a^*, b^*, c^*) \\ &= [1 - R_{XY}^2] \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2], \end{aligned}$$

por lo que:

$$\begin{aligned} 0 &\leq \text{variación residual de Tipo 2 para } Y \\ &\leq \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2] = \text{variación total de } Y, \end{aligned}$$

lo que prueba *i*).

Por otro lado, $0 \leq \psi(a^*, b^*, c^*) = [1 - R_{XY}^2] \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2]$ garantiza que $1 - R_{XY}^2 \geq 0$, lo que demuestra *ii*). \square

Por lo que concierne a los valores extremos del rango de R_{XY}^2 , en la proposición siguiente se exponen condiciones para alcanzarlos.

Proposición 1.2.6 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, no degenerados y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, entonces:*

i) $R_{XY}^2 = 1$ si, y sólo si, existe $(a, B) \in \mathbb{R} \times \mathcal{K}_c(\mathbb{R})$ tal que para casi todo $\omega \in \Omega$ se cumple que $Y = aX + B$;

ii) $R_{XY}^2 = 0$ si, y sólo si:

$$\begin{aligned} &4 \sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y | P) \\ &\leq \text{Cov}(\text{mid } X, \text{mid } Y | P) \\ &\leq -4 \sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y | P). \end{aligned}$$

Demostración.

i) Si para cierto $(a, B) \in \mathbb{R} \times \mathcal{K}_c(\mathbb{R})$ se cumple que $Y = aX + B$ casi seguro $[P]$, entonces, $\psi(a, 2 \text{spr } B, \text{inf } B) = 0$, lo que obliga a que $[1 - R_{XY}^2] \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2] = 0$, y al ser $\sigma_{f_Y}^2 > 0$ y $\rho_{f_Y f_{[0,1]}}^2 < 1$ en las condiciones supuestas, $R_{XY}^2 = 1$. Recíprocamente, $R_{XY}^2 = 1$ implica que $\psi(a^*, b^*, c^*) = [1 - R_{XY}^2] \sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2] = 0$, es decir, que $d_W(Y, a^* X + [c^*, c^* + b^*]) = 0$ *c.s.* $[P]$, y por tanto que $Y = a^* X + [c^*, c^* + b^*]$ casi seguro $[P]$.

ii) Como $R_{XY}^2 = 0$ si, y sólo si, $a^* = 0$, $R_{XY}^2 = 0$ se verificará si, y sólo si, $a_1 \leq 0$ y $a_2 \geq 0$. Como $a_1 \leq 0$ si, y sólo si, $\rho_{f_X f_Y} \leq \rho_{f_X f_{[0,1]}} \cdot \rho_{f_Y f_{[0,1]}}$ lo que equivale a que

$$\frac{\text{Cov}(\text{mid } X, \text{mid } Y|P) + 4\sigma_{f_{[0,1]}}^2 E(\text{spr } X \cdot \text{spr } Y|P)}{\sqrt{\sigma_{f_X}^2 \sigma_{f_Y}^2}} \leq \frac{4\sigma_{f_{[0,1]}}^2 E(\text{spr } X|P) \cdot E(\text{spr } Y|P)}{\sqrt{\sigma_{f_X}^2 \sigma_{f_Y}^2}},$$

de donde se puede deducir que la primera condición para que se verifique $R_{XY}^2 = 0$ es equivalente a la condición:

$$[\text{C.1}] \quad \text{Cov}(\text{mid } X, \text{mid } Y|P) + 4\sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P) \leq 0.$$

De forma análoga, se puede ver que la condición $a_2 \geq 0$ ocurre si, y sólo si, $-\rho_{\bar{f}_X f_Y} \leq \rho_{f_{[0,1]} f_X} \cdot \rho_{f_{[0,1]} f_Y}$, lo que equivale a que:

$$\begin{aligned} & \frac{4\sigma_{f_{[0,1]}}^2 E(\text{spr } X|P) \cdot E(\text{spr } Y|P)}{\sqrt{\sigma_{f_X}^2 \sigma_{f_Y}^2}} \\ & \geq \frac{4\sigma_{f_{[0,1]}}^2 E(\text{spr } X \cdot \text{spr } Y|P) - \text{Cov}(\text{mid } X, \text{mid } Y|P)}{\sqrt{\sigma_{f_X}^2 \sigma_{f_Y}^2}}, \end{aligned}$$

de donde se deduce que $a_2 \leq 0$ es equivalente a la condición:

$$[\text{C.2}] \quad -\text{Cov}(\text{mid } X, \text{mid } Y|P) + 4\sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P) \leq 0.$$

La combinación de las condiciones [C.1] y [C.2] da lugar a:

$$\begin{aligned} & 4\sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P) \\ & \leq \text{Cov}(\text{mid } X, \text{mid } Y|P) \\ & \leq -4\sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P). \end{aligned}$$

□

A partir de la Proposición 1.1.8 puede concluirse que la independencia estocástica conlleva la anulación de R_{XY}^2 . Así:

Proposición 1.2.7 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, no degenerados y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, entonces, si X e Y son estocásticamente independientes se cumple que $R_{XY}^2 = 0$.*

El coeficiente de determinación en el sentido de la relación funcional de Tipo 2, satisface también la propiedad de simetría siguiente:

Proposición 1.2.8 *Si X e Y son dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, no degenerados y $E(|X| | P) < \infty$, $E(|Y| | P) < \infty$, entonces $R_{XY}^2 = R_{YX}^2$.*

A diferencia de lo que ocurre en el caso de variables reales, y en la regresión de Tipo 1 (ver Teorema 1.2.4), no es cierto que $R_{(aX+B)(cY+D)}^2 = R_{XY}^2$, cualesquiera que sean $a, c \in \mathbb{R}$, $B, D \in \mathcal{K}_c(\mathbb{R})$. De este modo:

Contraejemplo 1.2.1 Sea (Ω, \mathcal{A}, P) el espacio de probabilidad definido de forma que:

$$\Omega = \{\omega_1, \omega_2\}, \mathcal{A} = \mathcal{P}(\Omega), P(\omega_1) = P(\omega_2) = 0.5$$

y sean X e Y dos conjuntos aleatorios asociados a ese espacio y dados por:

$$X(\omega_1) = [0, 2], X(\omega_2) = [1, 3],$$

$$Y(\omega_1) = [0, 2], Y(\omega_2) = [1, 5].$$

El coeficiente de determinación R_{XY}^2 es igual a 0.8901 (se calcula a través del CASO 3).

Si consideramos dos nuevos conjuntos aleatorios definidos por:

$$X' = -2X + [0, 2], \quad Y' = -3Y,$$

el coeficiente de determinación, $R_{X'Y'}^2$, también se calcula a través del CASO 3 y su valor es igual a 0.7500.

Observación 1.2.1 Aunque la igualdad $R_{(aX+B)(cY+D)}^2 = R_{XY}^2$ no sea válida, en general, puede garantizarse que es cierta en algunas condiciones. Concretamente, puede verificarse que:

- cuando X e Y corresponden al CASO 5, la igualdad es siempre válida;
- cuando X e Y corresponden al CASO 1, entonces si $a.c > 0$ y X' e Y' corresponden al CASO 1 o bien $a.c < 0$ y X' e Y' corresponden al CASO 2, la igualdad es cierta;
- cuando X e Y corresponden al CASO 2, entonces si $a.c > 0$ y X' e Y' corresponden al CASO 2 o bien $a.c < 0$ y X' e Y' corresponden al CASO 1, la igualdad es cierta.

Observación 1.2.2 Es importante señalar que, así como el análisis de la regresión de Tipo 1 es una particularización del de la regresión de Tipo 2, en el sentido de que las soluciones para el primero son las que obtendríamos para el segundo cuando $b^* = 0$ (es decir, cuando $B \in \mathbb{R}$), no ocurre lo mismo con el análisis de la correlación. Ello obedece al hecho de que la variación total para el problema de Tipo 1 no coincide con la correspondiente al Tipo 2, sino que esta última es inferior (es igual a la primera multiplicada por $[1 - \rho_{f_X f_Y}^2]$, que es un valor en $(0, 1)$ en las condiciones consideradas en la sección).

En consecuencia, la variación residual en la relación de Tipo 2 siempre será a lo sumo igual a la de Tipo 1 (es decir, para una misma elección

de W la relación óptima del segundo tipo será “más adecuada” que la del primero), pero el coeficiente de determinación para el primero no necesariamente es inferior al del segundo. En otras palabras, para comparar tipos de relaciones diferentes, la herramienta apropiada es la variación residual y no los coeficientes de determinación (que sí son un elemento idóneo para la comparación entre relaciones del mismo tipo).

Es también interesante observar que aunque un conjunto aleatorio X con valores en $\mathcal{K}_c(\mathbb{R})$ está caracterizado por las variables aleatorias reales $\text{mid } X$ y $\text{spr } X$, dados dos conjuntos aleatorios X e Y , el hecho de que $\text{mid } X$ y $\text{mid } Y$ satisfagan una relación lineal sobre toda la población y lo mismo ocurra con $\text{spr } X$ y $\text{spr } Y$ no obliga a que exista una relación del Tipo 2 sobre toda la población. De hecho, puede llegar a ocurrir que X e Y tengan una relación nula, como puede verse en el ejemplo siguiente:

Ejemplo 1.2.1 Sea (Ω, \mathcal{A}, P) el espacio de probabilidad definido por:

$$\Omega = \{\omega_1, \omega_2\}, \mathcal{A} = \mathcal{P}(\Omega), P(\omega_1) = P(\omega_2) = 0.5,$$

y sean X e Y dos conjuntos aleatorios asociados, dados por:

$$X(\omega_1) = [-1, 1], \quad X(\omega_2) = [-4, 6], \quad Y(\omega_1) = [-11, 15], \quad Y(\omega_2) = [0, 2].$$

En este caso, se satisfacen las relaciones:

$$\text{mid } Y(\omega_i) = 2 - \text{mid } X(\omega_i), \quad \text{spr } Y(\omega_i) = 16 - 3 \text{spr } X(\omega_i),$$

para todo $i \in \{1, 2\}$, de modo que los coeficientes de determinación para ambos pares de variables aleatorias reales son iguales a 1.

Sin embargo, si W es la medida de Lebesgue en $[0, 1]$, puede comprobarse fácilmente que:

$$a_1 = -17 < 0 < 15 = a_2,$$

que es condición necesaria y suficiente para la anulación de R_{XY}^2 y para que la relación óptima de Y respecto a X venga dada por:

$$\widehat{Y} = E^A[Y|P] = [-5.5, 8.5],$$

el valor estimado de Y no dependa del de X .

Por otro lado, podría comprobarse que tampoco la relación óptima de Tipo 1 conduce a un coeficiente de determinación máximo, puesto que $r_{XY}^2 = 0.0062$.

1.2.3 Algoritmo para el cálculo del coeficiente de determinación en relaciones de Tipo 2

Como apoyo para el cálculo práctico del coeficiente R_{XY}^2 , puede considerarse el algoritmo siguiente:

Paso 1. Calcular los valores $a_1, a_2, b_1, b_2, \sigma_{f_X}^2, \sigma_{f_Y}^2, \rho_{f_X f_{[0,1]}}^2, \rho_{f_Y f_{[0,1]}}^2, \rho_{f_X f_Y}^2$ y $\rho_{f_X f_{[0,1]}}^2$, e ir al Paso 2.

Paso 2. SI $a_1 \leq 0$ ENTONCES ir al Paso 3, EN CASO CONTRARIO ir al Paso 5.

Paso 3. SI $a_2 \geq 0$ ENTONCES el coeficiente de determinación viene dado por:

$$R_{XY}^2 = 0,$$

EN CASO CONTRARIO ir al Paso 4.

Paso 4. SI $b_2 > 0$ ENTONCES el coeficiente de determinación viene dado por:

$$R_{XY}^2 = \frac{a_2^2 \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2]}{\sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2]},$$

EN CASO CONTRARIO el coeficiente de determinación viene dado por:

$$R_{XY}^2 = \frac{\rho_{\bar{f}_X f_Y}^2 - \rho_{f_Y f_{[0,1]}}^2}{1 - \rho_{f_Y f_{[0,1]}}^2}.$$

Paso 5. SI $a_2 \geq 0$ ENTONCES ir al Paso 6, EN CASO CONTRARIO ir al Paso 7.

Paso 6. SI $b_1 > 0$ ENTONCES el coeficiente de determinación viene dado por:

$$R_{XY}^2 = \frac{a_1^2 \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2]}{\sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2]},$$

EN CASO CONTRARIO el coeficiente de determinación viene dado por:

$$R_{XY}^2 = \frac{\rho_{f_X f_Y}^2 - \rho_{f_Y f_{[0,1]}}^2}{1 - \rho_{f_Y f_{[0,1]}}^2}.$$

Paso 7. SI $b_1 > 0$ ENTONCES ir al Paso 8, EN CASO CONTRARIO ir al Paso 9.

Paso 8. SI $b_2 > 0$ ENTONCES el coeficiente de determinación viene dado por:

$$R_{XY}^2 = \frac{\max \{a_1^2, a_2^2\} \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2]}{\sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2]},$$

EN CASO CONTRARIO el coeficiente de determinación viene dado por:

$$R_{XY}^2 = \frac{\max \left\{ a_1^2 \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2], \sigma_{f_Y}^2 [\rho_{\bar{f}_X f_Y}^2 - \rho_{f_Y f_{[0,1]}}^2] \right\}}{\sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2]}.$$

Paso 9. SI $b_2 > 0$ ENTONCES el coeficiente de determinación viene dado por:

$$R_{XY}^2 = \frac{\max \left\{ \sigma_{f_Y}^2 [\rho_{f_X f_Y}^2 - \rho_{f_Y f_{[0,1]}}^2], a_2^2 \sigma_{f_X}^2 [1 - \rho_{f_X f_{[0,1]}}^2] \right\}}{\sigma_{f_Y}^2 [1 - \rho_{f_Y f_{[0,1]}}^2]},$$

EN CASO CONTRARIO el coeficiente de determinación viene dado por:

$$R_{XY}^2 = \frac{\max \left\{ \rho_{f_X f_Y}^2, \rho_{f_X f_{[0,1]}}^2 \right\} - \rho_{f_Y f_{[0,1]}}^2}{1 - \rho_{f_Y f_{[0,1]}}^2}.$$

1.2.4 Aplicación sobre un ejemplo real

La aplicación de los estudios de esta sección va a llevarse a cabo para los datos del Ejemplo 1.1.3 proporcionados por el Servicio de Nefrología del Hospital Valle del Nalón de Langreo (Asturias).

Ejemplo 1.2.2 Para los datos ligados que aparecen en la Tabla 1.1 (con $X =$ “rango de la tensión arterial sistólica a lo largo de un día”, e $Y =$ “rango de la tensión arterial diastólica a lo largo del mismo día”, en una población de 59 pacientes hospitalizados en el Hospital Valle del Nalón, el valor del coeficiente de determinación para la relación de Tipo 1 es:

$$r_{XY}^2 = 0.6999,$$

mientras que para la relación de Tipo 2 es

$$R_{XY}^2 = 0.4193.$$

Este ejemplo confirma las puntualizaciones hechas en la Observación 1.2.2.

1.3 Estudios inferenciales en problemas de relaciones lineales entre dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$

En las dos secciones anteriores se han determinado las relaciones óptimas de dos tipos entre dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, y se ha cuantificado el grado de dependencia entre esos conjuntos por la relación considerada.

A menudo los cálculos que implican estos estudios sólo pueden realizarse para una muestra de la población con la que se trabaja, por lo que se desea examinar hasta qué punto son extrapolables las conclusiones muestrales.

En los modelos lineales abordados en este capítulo hay muchos contrastes importantes que podrían llevarse a cabo, si bien la hipótesis cuyo contraste parece más relevante y viable es la de la anulación del coeficiente de determinación (que indicaría la no existencia de dependencia entre los conjuntos en el sentido de la relación funcional del tipo considerado o, desde otra perspectiva, significaría la ausencia de base para la predicción de Y a partir de X a través de la relación). Como la condición $r_{XY}^2 = 0$ ocurre sólo para conjuntos X e Y muy especiales, el contraste de la hipótesis $H_0 : r_{XY}^2 = 0$ no reviste mayor interés.

Esta sección va a centrarse mayoritariamente en la elaboración y discusión de métodos de contraste de la hipótesis nula:

$$H_0 : R_{XY}^2 = 0$$

frente a la alternativa:

$$H_a : R_{XY}^2 > 0,$$

supuesto que R_{XY}^2 es el coeficiente de determinación poblacional en el sentido de la relación funcional de Tipo 2 entre X e Y .

La metodología de alguna de las técnicas es claramente aplicable al contraste de otras hipótesis, como $H_0^\circ : a = 0$, en el supuesto de que se trabaje con un modelo lineal, del que también se ocupará esta sección.

Los contrastes se suponen realizados a partir de una muestra de tamaño n extraída al azar de la población, y sobre la que se observan los valores de intervalo que toman los conjuntos aleatorios X e Y . Para los diferentes métodos que se presentan en esta sección, se suponen condiciones distintas; más concretamente, se distinguirán los casos siguientes:

- Dados los conjuntos aleatorios X e Y con valores en $\mathcal{K}_c(\mathbb{R})$, se supondrá que el vector aleatorio

$$Z = (\text{mid } X, \text{mid } Y, \text{spr } X, \text{spr } Y)$$

sigue distribución normal 4-dimensional, $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, donde

$$\boldsymbol{\mu} = (E(\text{mid } X|P), E(\text{mid } Y|P), E(\text{spr } X|P), E(\text{spr } Y|P))$$

y

$$\boldsymbol{\Sigma} = \left(\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array} \right)$$

con $\boldsymbol{\Sigma}_{ij}$ cierta matriz de varianzas-covarianzas 2×2 para todo $i, j \in \{1, 2\}$.

- Los conjuntos aleatorios X e Y son normales.
- Los conjuntos aleatorios X e Y son simples.
- Los conjuntos aleatorios X e Y no deben cumplir ningún requisito especial.

Para el primer caso, conviene recordar (ver Muirhead, 1982) que si Z es un vector aleatorio con distribución 4-dimensional $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, entonces,

el vector $Z^1 = (\text{mid } X, \text{mid } Y)$ tiene distribución $\mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$ y el $Z^2 = (\text{spr } X, \text{spr } Y)$ tiene distribución $\mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$, con:

$$\boldsymbol{\mu}_1 = (E(\text{mid } X|P), E(\text{mid } Y|P)), \quad \boldsymbol{\mu}_2 = (E(\text{spr } X|P), E(\text{spr } Y|P))$$

y

$$\boldsymbol{\Sigma}_{11} = \begin{pmatrix} \text{Var}(\text{mid } X|P) & \text{Cov}(\text{mid } X, \text{mid } Y|P) \\ \text{Cov}(\text{mid } X, \text{mid } Y|P) & \text{Var}(\text{mid } Y|P), \end{pmatrix}$$

$$\boldsymbol{\Sigma}_{22} = \begin{pmatrix} \text{Var}(\text{spr } X|P) & \text{Cov}(\text{spr } X, \text{spr } Y|P) \\ \text{Cov}(\text{spr } X, \text{spr } Y|P) & \text{Var}(\text{spr } Y|P). \end{pmatrix}$$

En el estudio descriptivo de R_{XY}^2 de la sección anterior, se ha comprobado que la hipótesis nula $H_0 : R_{XY}^2 = 0$ se verifica si, y sólo si:

$$\begin{aligned} & 4\sigma_{f[0,1]}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P) \\ & \leq \text{Cov}(\text{mid } X, \text{mid } Y|P) \\ & \leq -4\sigma_{f[0,1]}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P), \end{aligned}$$

de modo que $R_{XY}^2 = 0$ si, y sólo si:

$$|\text{Cov}(\text{mid } X, \text{mid } Y|P)| + 4\sigma_{f[0,1]}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P) \leq 0.$$

Para contrastar la hipótesis nula:

$$H_0 : |\text{Cov}(\text{mid } X, \text{mid } Y|P)| + 4\sigma_{f[0,1]}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P) \leq 0$$

frente a la alternativa

$$H_a : |\text{Cov}(\text{mid } X, \text{mid } Y|P)| + 4\sigma_{f[0,1]}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P) > 0,$$

vamos a proceder en los tres primeros métodos de la forma siguiente:

- Se realiza en primer lugar el contraste entre las hipótesis:

$$\begin{cases} H_0^{(1)} : \text{Cov}(\text{mid } X, \text{mid } Y|P) = 0 \\ H_a^{(1)} : \text{Cov}(\text{mid } X, \text{mid } Y|P) \neq 0. \end{cases}$$

- Si se concluye que puede aceptarse la hipótesis $H_0^{(1)}$ en el contraste anterior para un nivel de significación prefijado α_1 , se desarrollará un segundo contraste, ahora entre las hipótesis:

$$\begin{cases} H_0^{(2)} : \text{Cov}(\text{spr } X, \text{spr } Y|P) \leq 0 \\ H_a^{(2)} : \text{Cov}(\text{spr } X, \text{spr } Y|P) > 0. \end{cases}$$

- Si puede aceptarse la hipótesis $H_0^{(2)}$ en este contraste para un nivel de significación prefijado α_2 , entonces podrá aceptarse la hipótesis nula inicial $H_0 : R_{XY}^2 = 0$ para un nivel de significación (por la desigualdad de Bonferroni) a lo sumo igual a $\alpha_1 + \alpha_2$;
- Si debe rechazarse la hipótesis $H_0^{(2)}$ en este contraste para el nivel prefijado α_2 , se rechazará H_0 en el contraste inicial.
- Si se concluye que debe rechazarse la hipótesis $H_0^{(1)}$ en el primer contraste, se presentan varias posibilidades:
 - Si tras realizar el segundo contraste debe rechazarse la hipótesis $H_0^{(2)}$, se optará por el rechazo de la hipótesis nula del contraste inicial $H_0 : R_{XY}^2 = 0$.
 - Si tras realizar el segundo contraste puede aceptarse la hipótesis $H_0^{(2)}$, podría plantearse un tercer contraste, esta vez entre las hipótesis:

$$\begin{cases} H_0^{(3)} : \text{Cov}(\text{spr } X, \text{spr } Y|P) = 0 \\ H_a^{(3)} : \text{Cov}(\text{spr } X, \text{spr } Y|P) \neq 0 (< 0). \end{cases}$$

- Si pudiera aceptarse la hipótesis $H_0^{(3)}$ en este tercer contraste para el nivel de significación α_3 , entonces debería rechazarse H_0 para un nivel de significación (por la desigualdad de Bonferroni) a lo sumo igual a $\alpha_1 + \alpha_2 + \alpha_3$.
- Si debe rechazarse la hipótesis $H_0^{(3)}$ en este tercer contraste, la situación no nos llevaría a una conclusión determinada y habría que estudiar el problema recurriendo a otras técnicas.

1.3.1 Contraste sobre R_{XY}^2 cuando (mid X , mid Y , spr X , spr Y) tiene distribución normal

Sean X e Y dos conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$, no degenerados y asociados con el espacio de probabilidad (Ω, \mathcal{A}, P) . Sea Z el vector aleatorio

$$Z = (\text{mid } X, \text{mid } Y, \text{spr } X, \text{spr } Y).$$

Sea (X_i, Y_i) , $i = 1, \dots, n$ ($n > 2$) una muestra aleatoria simple obtenida a partir del vector bidimensional de conjuntos aleatorios (X, Y) . A partir de esa muestra aleatoria simple se obtendrían de forma inmediata otras tres:

$$(Z_1, \dots, Z_n),$$

con $Z_i = (\text{mid } X_i, \text{mid } Y_i, \text{spr } X_i, \text{spr } Y_i)$, $i = 1, \dots, n$,

$$(Z_1^1, \dots, Z_n^1),$$

con $Z_i^1 = (\text{mid } X_i, \text{mid } Y_i)$, $i = 1, \dots, n$, y

$$(Z_1^2, \dots, Z_n^2),$$

con $Z_i^2 = (\text{spr } X_i, \text{spr } Y_i)$, $i = 1, \dots, n$.

Como condiciones de trabajo, en este caso, se supone que Z sigue una distribución normal 4-dimensional $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, donde:

$$\boldsymbol{\mu} = (E(\text{mid } X|P), E(\text{mid } Y|P), E(\text{spr } X|P), E(\text{spr } Y|P))$$

y

$$\boldsymbol{\Sigma} = \left(\begin{array}{c|c} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{array} \right),$$

con

$$\boldsymbol{\Sigma}_{11} = \begin{pmatrix} \text{Var}(\text{mid } X|P) & \text{Cov}(\text{mid } X, \text{mid } Y|P) \\ \text{Cov}(\text{mid } X, \text{mid } Y|P) & \text{Var}(\text{mid } Y|P) \end{pmatrix},$$

$$\boldsymbol{\Sigma}_{12} = \begin{pmatrix} \text{Cov}(\text{mid } X, \text{spr } X|P) & \text{Cov}(\text{mid } X, \text{spr } Y|P) \\ \text{Cov}(\text{mid } Y, \text{spr } X|P) & \text{Cov}(\text{mid } Y, \text{spr } Y|P) \end{pmatrix},$$

$$\boldsymbol{\Sigma}_{21} = \begin{pmatrix} \text{Cov}(\text{mid } X, \text{spr } X|P) & \text{Cov}(\text{mid } Y, \text{spr } X|P) \\ \text{Cov}(\text{mid } X, \text{spr } Y|P) & \text{Cov}(\text{mid } Y, \text{spr } Y|P) \end{pmatrix},$$

$$\boldsymbol{\Sigma}_{22} = \begin{pmatrix} \text{Var}(\text{spr } X|P) & \text{Cov}(\text{spr } X, \text{spr } Y|P) \\ \text{Cov}(\text{spr } X, \text{spr } Y|P) & \text{Var}(\text{spr } Y|P) \end{pmatrix}.$$

1.3.1.1 Contraste de $H_0^{(1)}$ frente a $H_a^{(1)}$

Como se señaló anteriormente, en las condiciones supuestas los vectores aleatorios bidimensionales Z^1 y Z^2 siguen distribución normal (ver Muirhead, 1982). Concretamente: $Z^1 = (\text{mid } X, \text{mid } Y)$ tiene distribución $\mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$ con:

$$\boldsymbol{\mu}_1 = (E(\text{mid } X|P), E(\text{mid } Y|P)).$$

Para contrastar la hipótesis $H_0^{(1)} : \text{Cov}(\text{mid } X, \text{mid } Y|P) = 0$ frente a la correspondiente alternativa $H_a^{(1)}$, se considera la muestra (Z_1^1, \dots, Z_n^1) .

Bajo la hipótesis $H_0^{(1)}$, que en las condiciones de normalidad supuestas (que obligan a $\text{Var}(\text{mid } X|P) > 0$ y $\text{Var}(\text{mid } Y|P) > 0$) equivale a que en el espacio (Ω, \mathcal{A}, P) el coeficiente de correlación poblacional entre las dos variables componentes de Z^1 se anule, es decir, $\rho_{\text{mid } X \text{ mid } Y} = 0$, se verifica que el estadístico:

$$T = \sqrt{n-2} \frac{r_{\text{mid } X \text{ mid } Y}}{\sqrt{1 - r_{\text{mid } X \text{ mid } Y}^2}}$$

sigue una distribución t_{n-2} (ver Muirhead, 1982), donde $r_{\text{mid } X \text{ mid } Y}$ representa el coeficiente de correlación lineal muestral de las dos componentes de Z^1 , esto es:

$$r_{\text{mid } X \text{ mid } Y} = \frac{\sum_{i=1}^n (\text{mid } X_i - \overline{\text{mid } X}) (\text{mid } Y_i - \overline{\text{mid } Y})}{\sqrt{\sum_{i=1}^n (\text{mid } X_i - \overline{\text{mid } X})^2 \sum_{i=1}^n (\text{mid } Y_i - \overline{\text{mid } Y})^2}}$$

con

$$\overline{\text{mid } X} = \frac{1}{n} \sum_{i=1}^n \text{mid } X_i.$$

De hecho, el resultado que se está aplicando proviene de uno más general, según el cual el estadístico T se distribuye (cualquiera que sea el valor de $\rho_{\text{mid } X \text{ mid } Y}$) de acuerdo con una t de Student no centrada, con $n-2$ grados de libertad y con parámetro de desviación igual a:

$$\sqrt{\sum_{i=1}^n (\text{mid } X_i - \overline{\text{mid } X})^2} \cdot \frac{\rho_{\text{mid } X \text{ mid } Y}}{\sqrt{1 - \rho_{\text{mid } X \text{ mid } Y}^2}}$$

(ver Johnson *et al.*, 1994).

En consecuencia, el siguiente teorema ofrece un procedimiento para el contraste de la hipótesis $H_0^{(1)}$:

Teorema 1.3.1 *Para contrastar al nivel de significación α_1 la hipótesis nula:*

$$H_0^{(1)} : \text{Cov}(\text{mid } X, \text{mid } Y | P) = 0$$

frente a la alternativa:

$$H_a^{(1)} : \text{Cov}(\text{mid } X, \text{mid } Y | P) \neq 0,$$

$H_0^{(1)}$ debe rechazarse siempre que

$$\frac{(n-2) r_{\text{mid } X \text{ mid } Y}^2}{1 - r_{\text{mid } X \text{ mid } Y}^2} > F_{1, n-2, \alpha_1},$$

donde $F_{1, n-2, \alpha_1}$ denota el percentil de orden $100(1 - \alpha_1)$ de la distribución F de Snedecor con 1 y $n - 2$ grados de libertad.

Demostración. Un camino alternativo para probar este resultado, y que se corresponde con la forma equivalente de expresar la condición de rechazo en el enunciado del teorema, consiste en utilizar directamente la técnica de la razón de verosimilitudes. Para ello, se considera la muestra (Z_1^1, \dots, Z_n^1) , y se determinan los estimadores máximo-verosímiles de $\boldsymbol{\mu}_1$ y $\boldsymbol{\Sigma}_{11}$, que para esa muestra vienen dados por:

$$\widehat{\boldsymbol{\mu}}_1 = (\overline{\text{mid } X}, \overline{\text{mid } Y}),$$

independientemente de que $H_0^{(1)}$ sea o no cierta, y:

$$\widehat{\boldsymbol{\Sigma}}_{11} = \begin{pmatrix} S_{\text{mid } X}^2 & 0 \\ 0 & S_{\text{mid } Y}^2 \end{pmatrix}$$

si $H_0^{(1)}$ es cierta, y

$$\widehat{S}_{11} = \begin{pmatrix} S_{\text{mid } X}^2 & S_{(\text{mid } X)(\text{mid } Y)} \\ S_{(\text{mid } X)(\text{mid } Y)} & S_{\text{mid } Y}^2 \end{pmatrix}$$

en el caso general, con:

$$S_{\text{mid } X}^2 = \frac{1}{n} \sum_{i=1}^n (\text{mid } X_i - \overline{\text{mid } X})^2,$$

$$S_{(\text{mid } X)(\text{mid } Y)} = \frac{1}{n} \sum_{i=1}^n (\text{mid } X_i - \overline{\text{mid } X}) (\text{mid } Y_i - \overline{\text{mid } Y}).$$

Como es sabido, el test de la razón de verosimilitudes para contrastar $H_0^{(1)}$ frente a $H_a^{(1)}$, está basado en el estadístico:

$$-2 \log \Lambda = 2n(M_a - \log M_g - 1)$$

donde M_a y M_g son las medias aritmética y geométrica, respectivamente, de los autovalores de la matriz $\widehat{\Sigma}_{11}^{-1} \widehat{S}_{11}$ (ver Mardia *et al.*, 1997). Por lo tanto:

$$-2 \log \Lambda = 2n \left(-\log \sqrt{1 - \frac{S_{\text{mid } X \text{ mid } Y}^2}{S_{\text{mid } X}^2 S_{\text{mid } Y}^2}} \right),$$

de donde:

$$\Lambda^{2/n} = 1 - \frac{S_{\text{mid } X \text{ mid } Y}^2}{S_{\text{mid } X}^2 S_{\text{mid } Y}^2}.$$

Si la hipótesis $H_0^{(1)}$ es cierta, $S_{\text{mid } X \text{ mid } Y}^2$ debe ser próxima a cero y, en consecuencia, $H_0^{(1)}$ debería rechazarse si $\Lambda^{2/n}$ toma valores ‘pequeños’.

El estadístico matricial

$$n \widehat{S}_{11} = \begin{pmatrix} n S_{\text{mid } X}^2 & n S_{\text{mid } X \text{ mid } Y} \\ n S_{\text{mid } X \text{ mid } Y} & n S_{\text{mid } Y}^2 \end{pmatrix}$$

sigue distribución de Wishart $\mathcal{W}(\Sigma, n-1)$ (ver Mardia *et al.*, 1997). Como $n > 2$, se cumple que bajo la hipótesis $H_0^{(1)}$

$$T_1 = n \left(S_{\text{mid } Y}^2 - \frac{S_{\text{mid } X \text{ mid } Y}^2}{S_{\text{mid } X}^2} \right)$$

sigue distribución $\text{Var}(\text{mid } Y|P) \cdot \chi_{n-2}^2 \equiv \mathcal{W}(\text{Var}(\text{mid } Y|P), n-2)$, y

$$T_2 = n \frac{S_{\text{mid } X \text{ mid } Y}^2}{S_{\text{mid } X}^2}$$

sigue distribución $\text{Var}(\text{mid } Y|P) \cdot \chi_1^2 \equiv \mathcal{W}(\text{Var}(\text{mid } Y|P), 1)$ y, además, ambos estadísticos son independientes (ver Mardia, 1997).

En consecuencia, resulta que:

$$\Lambda^{2/n} = \frac{T_1}{T_1 + T_2}$$

y el estadístico

$$\frac{(1 - \Lambda^{2/n})(n-2)}{\Lambda^{2/n}} = \frac{(n-2)T_2}{T_1} = T^2$$

sigue distribución F de Snedecor con 1 y $n-2$ grados de libertad, lo que demuestra el resultado. \square

Observación 1.3.1 La potencia del test anterior podría estudiarse sobre la base de la distribución t de Student no centrada del estadístico T cuando $\rho_{\text{mid } X \text{ mid } Y}$ tome valores no nulos (ver Johnson *et al.*, 1994).

1.3.1.2 Contraste de $H_0^{(2)}$ frente a $H_a^{(2)}$

En las condiciones supuestas el vector aleatorio $Z^2 = (\text{spr } X, \text{spr } Y)$ tiene distribución $\mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$ con:

$$\boldsymbol{\mu}_2 = (E(\text{spr } X|P), E(\text{spr } Y|P)).$$

Para contrastar la hipótesis unilateral $H_0^{(2)} : \text{Cov}(\text{spr } X, \text{spr } Y) \leq 0$ frente a la alternativa $H_a^{(2)} : \text{Cov}(\text{spr } X, \text{spr } Y) > 0$, se considera la muestra (Z_1^2, \dots, Z_n^2) .

La hipótesis $H_0^{(2)}$ equivale, en las condiciones supuestas, a que en el espacio (Ω, \mathcal{A}, P) el coeficiente de correlación poblacional entre las dos variables componentes de Z^2 no sea positivo, es decir, $\rho_{\text{spr } X \text{ spr } Y} \leq 0$. Para contrastar esa hipótesis equivalente se va a considerar el estadístico $r_{\text{spr } X \text{ spr } Y}$ que representa el coeficiente de correlación lineal muestral de las dos componentes de Z^2 , esto es:

$$r_{\text{spr } X \text{ spr } Y} = \frac{\sum_{i=1}^n (\text{spr } X_i - \overline{\text{spr } X}) (\text{spr } Y_i - \overline{\text{spr } Y})}{\sqrt{\sum_{i=1}^n (\text{spr } X_i - \overline{\text{spr } X})^2 \sum_{i=1}^n (\text{spr } Y_i - \overline{\text{spr } Y})^2}}$$

con

$$\overline{\text{spr } X} = \frac{1}{n} \sum_{i=1}^n \text{spr } X_i.$$

Para formalizar el contraste se determina la distribución de $r_{\text{spr } X \text{ spr } Y}$. Esta lleva asociada la función de densidad:

$$f(r) = \frac{(n-2) \Gamma(n-2)}{\Gamma(n-1/2) \sqrt{2\pi}} \cdot (1-r^2)^{(n-4)/2} \cdot {}_2F_1(1/2; 1/2; n-1/2; 1/2)$$

si $r \in (-1, 1)$, $f(r) = 0$ en caso contrario, donde:

$${}_2F_1(1/2; 1/2; n-1/2; 1/2) = \sum_{k=0}^{\infty} \frac{(1/2)_k (1/2)_k}{(n-1/2)_k} \cdot \frac{(1/2)^k}{k!}$$

y

$$\begin{aligned} \left(\frac{1}{2}\right)_k &= \frac{1}{2} \cdot \frac{3}{2} \cdot \dots \cdot \left(\frac{1}{2} + k - 1\right) \\ \left(n - \frac{1}{2}\right)_k &= \left(n - \frac{1}{2}\right) \cdot \left(n + \frac{1}{2}\right) \cdot \dots \cdot \left(n - \frac{1}{2} + k - 1\right). \end{aligned}$$

La función ${}_2F_1$ es conocida como función hipergeométrica (ver Muirhead, 1982). Existen tablas para la función de distribución correspondiente a esta distribución (ver David, 1938).

En el caso concreto en que $\rho_{\text{spr } X \text{ spr } Y} = 0$, la función de densidad adopta la expresión:

$$f_0(r) = \begin{cases} \frac{\Gamma((n-1)/2)}{\Gamma(1/2) \Gamma((n-2)/2)} \cdot (1-r^2)^{(n-4)/2} & \text{si } r \in (-1, 1) \\ 0 & \text{en caso contrario,} \end{cases}$$

que es simétrica respecto al valor $r = 0$ (ver Johnson *et al.*, 1994).

En consecuencia, el siguiente teorema ofrece un procedimiento para el contraste de la hipótesis $H_0^{(2)}$:

Teorema 1.3.2 *Para contrastar al nivel de significación α_2 la hipótesis nula:*

$$H_0^{(2)} : \text{Cov}(\text{spr } X, \text{spr } Y) \leq 0$$

frente a la alternativa:

$$H_a^{(2)} : \text{Cov}(\text{spr } X, \text{spr } Y) > 0,$$

$H_0^{(2)}$ debe rechazarse siempre que

$$r_{\text{spr } X \text{ spr } Y} > k_{\alpha_2},$$

donde k_{α_2} denota el percentil de orden $100(1 - \alpha_2)$ de la distribución cuya densidad es f_0 .

1.3.1.3 Contraste de $H_0^{(3)}$ frente a $H_a^{(3)}$

Para contrastar la hipótesis $H_0^{(3)} : \text{Cov}(\text{spr } X, \text{spr } Y|P) = 0$ frente a la alternativa $H_a^{(3)}$, se consideraría de nuevo la muestra (Z_1^2, \dots, Z_n^2) , y se seguiría un proceso análogo al del contraste de $H_0^{(1)}$ frente a $H_a^{(1)}$, reemplazando $\text{mid } X$ por $\text{spr } X$ y $\text{mid } Y$ por $\text{spr } Y$, es decir:

Teorema 1.3.3 *Para contrastar al nivel de significación α_3 la hipótesis nula:*

$$H_0^{(3)} : \text{Cov}(\text{spr } X, \text{spr } Y|P) = 0$$

frente a la alternativa:

$$H_a^{(3)} : \text{Cov}(\text{spr } X, \text{spr } Y|P) \neq 0,$$

$H_0^{(3)}$ debe rechazarse siempre que

$$\frac{(n-2) r_{\text{spr } X \text{ spr } Y}^2}{1 - r_{\text{spr } X \text{ spr } Y}^2} > F_{1, n-2, \alpha_3}.$$

1.3.1.4 Algoritmo para el contraste de H_0 frente a H_a cuando $(\text{mid } X, \text{mid } Y, \text{spr } X, \text{spr } Y)$ tiene distribución normal

Para la resolución práctica del contraste inicial planteado a través de las diferentes etapas en que se ha desglosado el mismo, puede recurrirse al algoritmo siguiente, que resume lo examinado en la Subsección 1.3.1.

Paso 1. Para la muestra considerada, y prefijado el nivel de significación α_1 , calcular el valor del estadístico

$$T^2 = \frac{(n-2) r_{\text{mid } X \text{ mid } Y}^2}{1 - r_{\text{mid } X \text{ mid } Y}^2}$$

y el valor crítico $F_{1, n-2, \alpha_1}$, e ir al Paso 2.

Paso 2. SI el valor de T^2 en la muestra es inferior o igual a $F_{1,n-2,\alpha_1}$,
ENTONCES ir al Paso 3, EN CASO CONTRARIO al Paso 4.

Paso 3. Para la muestra considerada, y prefijado el nivel de significación α_2 , calcular el valor del estadístico

$$T' = r_{\text{spr } X \text{ spr } Y}$$

y el valor crítico k_{α_2} (al que se refiere el Teorema 1.3.2). SI el valor de T' en la muestra es inferior o igual a k_{α_2} , ENTONCES puede aceptarse H_0 a un nivel de significación a lo sumo igual a $\alpha_1 + \alpha_2$, EN CASO CONTRARIO debe rechazarse H_0 a un nivel de significación a lo sumo igual a $\alpha_1 + \alpha_2$.

Paso 4. Para la muestra considerada, y prefijado el nivel de significación α_2 , calcular el valor del estadístico T' y el valor k_{α_2} del Paso 3. SI el valor de T' en la muestra es superior a k_{α_2} , ENTONCES debe rechazarse H_0 a un nivel de significación a lo sumo igual a $\alpha_1 + \alpha_2$, EN CASO CONTRARIO ir al Paso 5.

Paso 5. Para la muestra considerada, y prefijado el nivel de significación α_3 , calcular el valor del estadístico

$$(T'')^2 = \frac{(n-2)r_{\text{spr } X \text{ spr } Y}^2}{1 - r_{\text{spr } X \text{ spr } Y}^2}$$

y el valor crítico $F_{1,n-2,\alpha_3}$. SI el valor de $(T'')^2$ en la muestra es inferior o igual a $F_{1,n-2,\alpha_3}$, ENTONCES debe rechazarse H_0 a un nivel de significación a lo sumo igual a $\alpha_1 + \alpha_2 + \alpha_3$, EN CASO CONTRARIO recurrir a otra técnica de contraste para H_0 y H_a .

1.3.2 Contraste sobre R_{XY}^2 cuando X e Y son conjuntos aleatorios normales

En este caso va a suponerse que los conjuntos aleatorios X e Y asociados al espacio de probabilidad (Ω, \mathcal{A}, P) son normales de acuerdo con la definición de Lyashenko (1980), recopilada en la Subsección 0.2.1, lo que significa que para cada $\omega \in \Omega$ se cumple que:

$$X(\omega) = \xi(\omega) + E^A[X|P],$$

donde $\xi : \Omega \rightarrow \mathbb{R}$ es una variable aleatoria real asociada al espacio (Ω, \mathcal{A}, P) con distribución normal de media 0 y varianza $\sigma_\xi^2 = \text{Var}(\xi|P)$ e

$$Y(\omega) = \eta(\omega) + E^A[Y|P],$$

donde $\eta : \Omega \rightarrow \mathbb{R}$ es otra variable aleatoria real asociada a (Ω, \mathcal{A}, P) con distribución normal de media 0 y varianza $\sigma_\eta^2 = \text{Var}(\eta|P)$.

Por lo tanto:

$$\text{mid } X(\omega) = \xi(\omega) + \text{mid } E^A[X|P]$$

para todo $\omega \in \Omega$, de modo que $\text{mid } X$ es una variable aleatoria real con distribución $\mathcal{N}(\text{mid } E^A[X|P], \sigma_\xi^2)$.

Siguiendo un razonamiento análogo, para todo $\omega \in \Omega$ se cumple que:

$$\text{mid } Y(\omega) = \eta(\omega) + \text{mid } E^A[Y|P],$$

de modo que $\text{mid } Y$ es una variable con distribución $\mathcal{N}(\text{mid } E^A[Y|P], \sigma_\eta^2)$.

Por otro lado, para todo $\omega \in \Omega$ se cumple que:

$$\text{spr } X(\omega) = \text{spr } E^A[X|P],$$

$$\text{spr } Y(\omega) = \text{spr } E^A[Y|P],$$

por lo que $\text{spr } X$ y $\text{spr } Y$ son variables aleatorias reales degeneradas y, como consecuencia:

$$\text{Cov}(\text{spr } X, \text{spr } Y|P) = 0.$$

De esta última afirmación se deduce que en la situación considerada, la condición $|\text{Cov}(\text{mid } X, \text{mid } Y|P)| + 4\sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P) \leq 0$ se reduciría a $\text{Cov}(\text{mid } X, \text{mid } Y|P) = 0$ y la de la hipótesis alternativa se reduciría a $\text{Cov}(\text{mid } X, \text{mid } Y|P) \neq 0$.

Pero, al ser X e Y normales, un contraste entre las dos últimas hipótesis equivale a un contraste entre:

$$\begin{cases} H_0 : \text{Cov}(\xi, \eta|P) = 0 \\ H_a : \text{Cov}(\xi, \eta|P) \neq 0. \end{cases}$$

Por todo ello, en este caso no se precisaría contrastar $H_0^{(1)}$, $H_0^{(2)}$ y $H_0^{(3)}$; y, como si la hipótesis H_0 es cierta, el vector aleatorio (ξ, η) asociado al espacio de probabilidad (Ω, \mathcal{A}, P) tiene distribución normal bidimensional $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, donde:

$$\boldsymbol{\mu} = (0, 0)$$

y:

$$\boldsymbol{\Sigma} = \begin{pmatrix} \text{Var}(\xi|P) & 0 \\ 0 & \text{Var}(\eta|P) \end{pmatrix},$$

el contraste sobre la base de una muestra aleatoria simple de tamaño n a partir de (X, Y) se formalizaría como sigue:

Teorema 1.3.4 *Para contrastar al nivel de significación α la hipótesis nula:*

$$H_0 : R_{XY}^2 = 0$$

frente a la alternativa:

$$H_a : R_{XY}^2 \neq 0,$$

H_0 debe rechazarse siempre que

$$\frac{(n-2)r_{\text{mid } X \text{ mid } Y}^2}{1-r_{\text{mid } X \text{ mid } Y}^2} > F_{1,n-2,\alpha},$$

donde $F_{1,n-2,\alpha}$ denota el percentil de orden $100(1-\alpha)$ de la distribución F de Snedecor con 1 y $n-2$ grados de libertad.

Observación 1.3.2 Al igual que en la subsección precedente, la potencia de este último test se podría analizar a partir de la distribución t de Student no centrada que se indicaba en la Observación 1.3.1.

1.3.3 Contrastes asintóticos sobre R_{XY}^2 cuando X e Y son conjuntos aleatorios simples

En este caso va a suponerse que los conjuntos aleatorios X e Y asociados al espacio de probabilidad (Ω, \mathcal{A}, P) son simples y no degenerados.

Sean $A_1, \dots, A_M \in \mathcal{K}_c(\mathbb{R})$ los valores distintos de X en Ω , y sean $B_1, \dots, B_{M'} \in \mathcal{K}_c(\mathbb{R})$ los valores distintos de Y en Ω .

Sea L el número de pares distintos $(A_m, B_{m'}) \in \mathcal{K}_c(\mathbb{R}) \times \mathcal{K}_c(\mathbb{R})$ que aparecen en la población con probabilidad positiva ($\max\{M, M'\} \leq L \leq M \cdot M'$).

En lo que sigue estos L pares distintos van a disponerse según una ordenación elegida arbitrariamente, denotándose por $(A, B)_l$ el l -ésimo par en esa ordenación.

Sean:

$$p_l = P(\{\omega \in \Omega \mid (X(\omega), Y(\omega)) = (A, B)_l\}) > 0, \quad l = 1, \dots, L-1,$$

$$p_L = 1 - \sum_{l=1}^{L-1} p_l.$$

En esta situación, la “distribución conjunta” de X e Y está caracterizada por el valor (vectorial) poblacional del parámetro vectorial $\mathbf{p} = (p_1, \dots, p_{L-1})$, para el que se cumplen las condiciones de regularidad siguientes:

- El espacio paramétrico correspondiente es $\mathbb{P} = [0, 1]^{L-1}$, y el verdadero valor de \mathbf{p} pertenece al interior de \mathbb{P} , es decir, a $(0, 1)^{L-1}$. Además, como se cumple que $\mathbf{p} \in (0, 1)^{L-1}$, para valores distintos de \mathbf{p} las medidas de probabilidad correspondientes deben diferir.
- El conjunto de “valores” (pares de conjuntos) distintos de (X, Y) en la población, $\{(A, B)_1, \dots, (A, B)_L\}$, no depende de \mathbf{p} .
- Para cada valor $\mathbf{p} \in (0, 1)^{L-1}$, se cumple que si $\Omega_l = \{\omega \in \Omega \mid (X(\omega), Y(\omega)) = (A, B)_l\}$ y $P(\Omega_l \mid \mathbf{p}) = p_l$ para cada $l \in \{1, \dots, L\}$, entonces $\log P(\Omega_l \mid \mathbf{p})$ es derivable parcialmente hasta el tercer orden con respecto a las componentes p_i de \mathbf{p} y en un entorno de \mathbf{p} se satisface además que:

$$\sum_{l=1}^L p_l \left| \frac{\partial^3}{\partial p_i \partial p_j \partial p_k} \log P(\Omega_l \mid \mathbf{p}) \right| < \infty$$

para cualesquiera $i, j, k \in \{1, \dots, L-1\}$.

- La matriz de información de Fisher asociada a la familia \mathcal{P} (familia de las medidas de probabilidad correspondientes a los distintos valores de $\mathbf{p} \in \mathbb{P}$) para el valor \mathbf{p} viene dada por:

$$\begin{aligned} \mathbf{I}_{\mathcal{X}}^F(\mathbf{p}) &= [\mathbf{I}_{ij}^F(\mathbf{p})] = \left[\sum_{l=1}^L p_l \frac{\partial \log p_l}{\partial p_i} \cdot \frac{\partial \log p_l}{\partial p_j} \right] \\ &= \left[\frac{\delta_{ij}}{p_i} + \frac{1}{p_L} \right] \quad \text{con } \delta_{ij} = \text{delta de Kronecker} \end{aligned}$$

$(i, j \in \{1, \dots, L-1\})$. $I_{\mathcal{X}}^F(\mathbf{p})$ está bien definida (ya que $I_{ij}^F(\mathbf{p})$ es finito para cualquier par (i, j)) y es una matriz definida positiva (es decir, la forma cuadrática asociada es definida positiva) para cada $\mathbf{p} \in (0, 1)^{L-1}$, ya que para cualquier $i \in \{1, \dots, L-1\}$ se cumple que:

$$\begin{vmatrix} I_{11}^F(\mathbf{p}) & \cdots & I_{1i}^F(\mathbf{p}) \\ \vdots & \ddots & \vdots \\ I_{i1}^F(\mathbf{p}) & \cdots & I_{ii}^F(\mathbf{p}) \end{vmatrix} = \frac{p_1 + \cdots + p_i + p_L}{p_1 \cdots p_i \cdots p_L} > 0$$

(lo que garantiza - ver, por ejemplo, Rao, 1973 - que la matriz de información de Fisher es definida positiva).

Sea (X_i, Y_i) , $i = 1, \dots, n$ una muestra aleatoria simple obtenida a partir del vector bidimensional de conjuntos aleatorios (X, Y) . Se denotará por f_{ni} el estadístico correspondiente a la frecuencia relativa de $(A, B)_i$ en la muestra aleatoria, y $\mathbf{f}_n = (f_{n1}, \dots, f_{n(L-1)})$.

Para definir estadísticos que resuelvan los contrastes de $H_0^{(1)}$ frente a $H_a^{(1)}$, $H_0^{(2)}$ frente a $H_a^{(2)}$, y $H_0^{(3)}$ frente a $H_a^{(3)}$, aprovechando los resultados conocidos de la Teoría de Grandes Muestras para \mathbf{p} , va a tenerse en cuenta que la sucesión $\{\mathbf{f}_n\}_n$ es solución del sistema de ecuaciones de verosimilitud, puesto que son los estimadores máximo-verosímiles de \mathbf{p} . En consecuencia, la sucesión de estimadores $\{\mathbf{f}_n\}_n$ es fuertemente consistente y se distribuye asintóticamente según una distribución normal $(L-1)$ -dimensional $\mathcal{N}(\mathbf{p}, [I_{\mathcal{X}}^F(\mathbf{p})]^{-1}/n)$ cuando $n \rightarrow \infty$, con:

$$[I_{ij}^F(\mathbf{p})]^{-1} = [p_i(\delta_{ij} - p_j)],$$

ya que $|[I_{ij}^F(\mathbf{p})]| = \left(\prod_{l=1}^L p_l\right)^{-1}$ y, para $i, j \in \{1, \dots, L-1\}$, el adjunto de $I_{ij}^F(\mathbf{p})$ es ${}^{\alpha}I_{ij}^F(\mathbf{p}) = p_i(\delta_{ij} - p_j) \left(\prod_{l=1}^L p_l\right)^{-1}$.

Además, $n(\mathbf{f}_n - \mathbf{p})\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})(\mathbf{f}_n - \mathbf{p})^t$ converge en distribución a una chi-cuadrado χ_{L-1}^2 cuando $n \rightarrow \infty$ (ver, por ejemplo, Rao, 1973, Serfling, 1980, Lehmann, 1983).

1.3.3.1 Contraste asintótico de $H_0^{(1)}$ frente a $H_a^{(1)}$

A partir de las condiciones anteriores se establece un resultado de apoyo para resolver el primer contraste.

Proposición 1.3.5 *Para cada $n \in \mathbb{N}$, se consideran n vectores bidimensionales de conjuntos aleatorios $(X_1, Y_1), \dots, (X_n, Y_n)$, independientes e idénticamente distribuidos que (X, Y) (es decir, una muestra aleatoria simple de tamaño n a partir de (X, Y)), definidos sobre el espacio de probabilidad (Ω, \mathcal{A}, P) de forma que $P(\{\omega \in \Omega \mid (X(\omega), Y(\omega)) = (A, B)_l\}) = p_l$ con $p_l \in (0, 1)$, $l = 1, \dots, L$, $\sum_{l=1}^L p_l = 1$. Se cumple entonces que:*

- i) Si $\mathbf{f}_n = (f_{n1}, \dots, f_{n(L-1)}) \in [0, 1]^{L-1}$, con $f_{nl} =$ frecuencia relativa de $(A, B)_l$ en la correspondiente realización de la muestra aleatoria simple de tamaño n ($l = 1, \dots, L-1$, $f_{nL} = 1 - \sum_{l=1}^{L-1} f_{nl}$), y $\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{f}_n)$ representa la covarianza muestral entre $\text{mid } X$ y $\text{mid } Y$ correspondiente, entonces, $\{\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{f}_n)\}_n$ es una sucesión de estimadores de $\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p})$ que es fuertemente consistente, es decir, cuando $n \rightarrow \infty$ se tiene que:

$$\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{f}_n) \xrightarrow{c.s.} \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}),$$

cualquiera que sea $\mathbf{p} = (p_1, \dots, p_{L-1})$ (con $p_1, \dots, p_{L-1} \in (0, 1)$ y $\sum_{l=1}^{L-1} p_l < 1$).

ii) La sucesión de variables aleatorias reales:

$$\{\sqrt{n} [\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{f}_n) - \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})]\}_n$$

converge en ley hacia una normal unidimensional $\mathcal{N}(0, \sigma^2(\mathbf{p}))$, con:

$$\sigma^2(\mathbf{p}) = \text{Var}([\text{mid } X - E(\text{mid } X | \mathbf{p})] [\text{mid } Y - E(\text{mid } Y | \mathbf{p})] | \mathbf{p}),$$

siempre que $\sigma^2(\mathbf{p}) > 0$.

iii) Si $\sigma^2(\mathbf{p}) = 0$ y para algún par $(i, j) \in \{1, \dots, L-1\} \times \{1, \dots, L-1\}$ se cumple que:

$$h_{ij} = \frac{\partial^2 \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})}{\partial p_i \partial p_j}$$

$$= [\text{mid}_1(A, B)_i - \text{mid}_1(A, B)_L] [\text{mid}_2(A, B)_j - \text{mid}_2(A, B)_L]$$

$$+ [\text{mid}_1(A, B)_j - \text{mid}_1(A, B)_L] [\text{mid}_2(A, B)_i - \text{mid}_2(A, B)_L] > 0$$

(con $\text{mid}_1(A, B)_l = \text{punto medio del primer componente de } (A, B)_l$ y $\text{mid}_2(A, B)_l = \text{punto medio del segundo componente de } (A, B)_l$), entonces:

$$\{2n [\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{f}_n) - \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})]\}_n$$

es una sucesión de variables aleatorias reales que converge en ley hacia una combinación lineal de, a lo sumo, $L-1$ variables chi-cuadrado con 1 grado de libertad e independientes.

Demostración.

- i) En las condiciones que se acaban de señalar se garantiza que $\mathbf{f}_n \xrightarrow{c.s.} \mathbf{p}$, por lo que de acuerdo con los resultados de la Teoría Asintótica en Inferencia Paramétrica (ver, por ejemplo, Serfling, 1980), al ser $\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})$ continua en un entorno de \mathbf{p} , se concluye que:

$$\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{f}_n) \xrightarrow{c.s.} \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})$$

cuando $n \rightarrow \infty$, es decir, $\{\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{f}_n)\}_n$ es una sucesión estimadora de $\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})$ fuertemente consistente.

- ii) Por las condiciones supuestas, para n suficientemente grande es posible desarrollar $\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{f}_n)$ en un entorno de \mathbf{p} . De este modo, el desarrollo de Taylor de primer orden de $\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{f}_n)$ será:

$$\begin{aligned} \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{f}_n) &= \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p}) \\ &+ \nabla \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})(\mathbf{f}_n - \mathbf{p})^t \\ &+ \frac{1}{2}(\mathbf{f}_n - \mathbf{p}) \mathbf{H} (\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p}_n^*)) (\mathbf{f}_n - \mathbf{p})^t, \end{aligned}$$

con ∇ el correspondiente vector gradiente, $\mathbf{H} = [h_{ij}]$ la matriz hessiana $(L - 1) \times (L - 1)$, y $\mathbf{p}_n^* \in \mathbb{P}$ tal que $\|\mathbf{p}_n^* - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$.

Por lo tanto:

$$\begin{aligned} &\sqrt{n} [\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{f}_n) - \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})] \\ &= \nabla \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p}) (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t \\ &+ \frac{1}{2}(\mathbf{f}_n - \mathbf{p}) \mathbf{H} (\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p}_n^*)) (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t. \end{aligned}$$

El vector gradiente vendrá dado por la matriz fila:

$$\begin{aligned} & \nabla \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}) \\ &= \left(\frac{\partial}{\partial p_1} \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}) \cdots \frac{\partial}{\partial p_{L-1}} \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}) \right) \end{aligned}$$

donde, para $i \in \{1, \dots, L-1\}$, se cumple que:

$$\begin{aligned} & \frac{\partial}{\partial p_i} \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}) \\ &= [\text{mid}_1(A, B)_i \text{mid}_2(A, B)_i - \text{mid}_1(A, B)_L \text{mid}_2(A, B)_L] \\ & \quad - [\text{mid}_1(A, B)_i - \text{mid}_1(A, B)_L] E(\text{mid } Y \mid P) \\ & \quad - [\text{mid}_2(A, B)_i - \text{mid}_2(A, B)_L] E(\text{mid } X \mid P). \end{aligned}$$

Como $\sqrt{n}(\mathbf{f}_n - \mathbf{p})$ se distribuye asintóticamente como una normal $(L-1)$ -dimensional $\mathcal{N}(\mathbf{0}, [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1})$, según las propiedades de la convergencia en ley (ver, por ejemplo, Serfling 1980), se tiene que cuando $n \rightarrow \infty$ la sucesión

$$\left\{ \nabla \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}) (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t \right\}_n$$

converge en ley a una variable con distribución:

$$\mathcal{N}\left(0, \nabla \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}) [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1} (\nabla \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}))^t\right),$$

siempre que sea positiva la varianza:

$$\begin{aligned} \sigma^2(\mathbf{p}) &= \nabla \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}) [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1} (\nabla \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}))^t \\ &= \sum_{i=1}^{L-1} \sum_{j=1}^{L-1} p_i (\delta_{ij} - p_j) \left(\frac{\partial}{\partial p_i} \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}) \right) \\ & \quad \cdot \left(\frac{\partial}{\partial p_j} \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}) \right) = \mathbf{W}^* \boldsymbol{\Sigma} \mathbf{W}^{*t}, \end{aligned}$$

donde \mathbf{W}^* es la matriz línea $(W_1^* \dots W_L^*)$, con:

$$W_l^* = [\text{mid}_1(A, B)_l - E(\text{mid } X | \mathbf{p})] [\text{mid}_2(A, B)_l - E(\text{mid } Y | \mathbf{p})],$$

y:

$$\Sigma = \begin{pmatrix} p_1(1-p_1) & \cdots & -p_1 p_L \\ \vdots & \ddots & \vdots \\ -p_L p_1 & \cdots & p_L(1-p_L) \end{pmatrix},$$

es decir:

$$\sigma^2(\mathbf{p}) = \text{Var}([\text{mid } X - E(\text{mid } X | \mathbf{p})] [\text{mid } Y - E(\text{mid } Y | \mathbf{p})] | \mathbf{p}).$$

En virtud de las propiedades relativas a las convergencias en ley y en probabilidad (ver, por ejemplo, Serfling, 1980), al ser $\|\mathbf{p}_n^* - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$ se tiene que cuando $n \rightarrow \infty$:

$$\left(\frac{\partial^2 \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})}{\partial p_i \partial p_j} \right)_{\mathbf{p}=\mathbf{p}_n^*} \xrightarrow{p} \frac{\partial^2 \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})}{\partial p_i \partial p_j},$$

para cualesquiera $i, j \in \{1, \dots, L-1\}$. Por otro lado, $\mathbf{f}_n \xrightarrow{c.s.} \mathbf{p}$ y, por lo tanto, $\mathbf{f}_n \xrightarrow{p} \mathbf{p}$, cuando $n \rightarrow \infty$, de donde:

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p}) \mathbf{H}(\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p}_n^*)) \xrightarrow{p} \mathbf{0}$$

y, como $\sqrt{n}(\mathbf{f}_n - \mathbf{p}) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1})$, se tiene que:

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p}) \mathbf{H}(\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p}_n^*)) (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t \xrightarrow{\mathcal{L}} 0$$

y, en consecuencia:

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p}) \mathbf{H}(\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p}_n^*)) (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t \xrightarrow{p} 0.$$

Los resultados anteriores garantizan que la sucesión:

$$\left\{ \sqrt{n} [\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{f}_n) - \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p})] \right\}_n$$

converge en ley hacia una variable con distribución

$$\mathcal{N} \left(0, \nabla \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}) [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1} (\nabla \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}))^t \right)$$

cuando $n \rightarrow \infty$ y, por lo tanto, se verifica *ii*) siempre que $\sigma^2(\mathbf{p}) > 0$.

iii) La matriz $\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})$ es definida positiva, por lo que la forma cuadrática asociada a $\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})$ y a $[\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1}$ es definida positiva. En consecuencia, si $\sigma^2(\mathbf{p}) = 0$ se debe cumplir que $\nabla \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}) = 0$.

El desarrollo de Taylor de segundo orden de $\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{f}_n)$, permite asegurar que para n suficientemente grande:

$$\begin{aligned} \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{f}_n) &= \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p}) \\ &+ \frac{1}{2} (\mathbf{f}_n - \mathbf{p}) \mathbf{H} (\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p})) (\mathbf{f}_n - \mathbf{p})^t \\ &+ \frac{1}{6} \sum_{i=1}^{L-1} \sum_{j=1}^{L-1} \sum_{k=1}^{L-1} \left(\frac{\partial^3 \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p})}{\partial p_i \partial p_j \partial p_k} \right)_{\mathbf{p}=\mathbf{p}_n^{**}} \\ &\quad \cdot (f_{ni} - p_i)(f_{nj} - p_j)(f_{nk} - p_k), \end{aligned}$$

con $\mathbf{p}_n^{**} \in \mathbb{P}$ tal que $\|\mathbf{p}_n^{**} - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$.

Por lo tanto:

$$\begin{aligned} &2n [\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{f}_n) - \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p})] \\ &= (\sqrt{n}(\mathbf{f}_n - \mathbf{p})) \mathbf{H} (\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p})) (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t \\ &+ \frac{1}{3} \sum_{i=1}^{L-1} \sum_{j=1}^{L-1} \sum_{k=1}^{L-1} \left(\frac{\partial^3 \text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p})}{\partial p_i \partial p_j \partial p_k} \right)_{\mathbf{p}=\mathbf{p}_n^{**}} \\ &\quad \cdot (f_{ni} - p_i) (\sqrt{n}(f_{nj} - p_j)) (\sqrt{n}(f_{nk} - p_k)). \end{aligned}$$

Razonando como en el apartado *ii*), puede concluirse que:

$$(f_{ni} - p_i) (\sqrt{n}(f_{nj} - p_j)) (\sqrt{n}(f_{nk} - p_k)) \xrightarrow{p} 0.$$

Además (ver, por ejemplo, Serfling, 1980), se tiene que la sucesión:

$$\left\{ (\sqrt{n}(\mathbf{f}_n - \mathbf{p})) \mathbf{H} (\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})) (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t \right\}_n$$

converge en ley hacia:

$$Y \mathbf{H} (\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})) Y^t,$$

donde Y es un vector aleatorio con distribución normal $(L-1)$ -dimensional $\mathcal{N}(\mathbf{0}, [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1})$.

Como se supone que alguna de las derivadas segundas

$$h_{ij} = \partial^2 \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p}) / \partial p_i \partial p_j$$

es positiva y, por lo tanto, que $\mathbf{H} (\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p}))$ no se reduce a la matriz nula de orden $(L-1) \times (L-1)$, y ya que el rango de $[\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1}$ es $L-1$, debe cumplirse que el vector Y puede expresarse como $Y = Z \mathbf{B}$ con Z vector aleatorio $(L-1)$ -dimensional cuyas componentes son $L-1$ variables aleatorias independientes e idénticamente distribuidas, con distribución $\mathcal{N}(0, 1)$, y \mathbf{B} es una matriz $(L-1) \times (L-1)$ tal que $\mathbf{B} \mathbf{B}^t = [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1}$. Además, existe una transformación $Z = U \mathbf{C}$ en la que \mathbf{C} es una matriz ortogonal, de forma que:

$$\begin{aligned} & Y \mathbf{H} (\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})) Y^t \\ &= Z \mathbf{B} \mathbf{H} (\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})) \mathbf{B}^t Z^t \\ &= U \mathbf{C} \mathbf{B} \mathbf{H} (\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})) \mathbf{B}^t \mathbf{C}^t U^t \\ &= \lambda_1 U_1^2 + \cdots + \lambda_q U_q^2, \end{aligned}$$

con $\lambda_1, \dots, \lambda_q$ ($q \leq L - 1$) los autovalores no nulos de la matriz $(L - 1) \times (L - 1)$ dada por $\mathbf{B} \mathbf{H} (\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p})) \mathbf{B}^t$, y donde U_1, \dots, U_q son variables aleatorias independientes e idénticamente distribuidas con distribución $\mathcal{N}(0, 1)$.

En consecuencia, el estudio de la forma cuadrática

$$Y \mathbf{H} (\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p})) Y^t$$

se reduce en primer lugar al de una forma cuadrática de un vector normal $(L - 1)$ -dimensional con componentes independientes y $\mathcal{N}(0, 1)$, y este último se reduce a su vez al de una combinación lineal de variables aleatorias independientes e idénticamente distribuidas, con distribución chi-cuadrado χ_1^2 , es decir, una combinación lineal de cuadrados de variables aleatorias independientes e idénticamente distribuidas, con distribución $\mathcal{N}(0, 1)$ (ver, por ejemplo, Rao, 1973, y Serfling, 1980, para una revisión de los resultados básicos de las afirmaciones precedentes). \square

Sobre la base de los resultados en la proposición anterior, y de las propiedades de la convergencia en ley y en probabilidad, se obtiene de forma prácticamente inmediata el siguiente resultado:

Teorema 1.3.6 *Para cada $n \in \mathbb{N}$, se consideran n vectores bidimensionales de conjuntos aleatorios $(X_1, Y_1), \dots, (X_n, Y_n)$, independientes e idénticamente distribuidos que (X, Y) (es decir, una muestra aleatoria simple de tamaño n a partir de (X, Y)), definidos sobre el espacio de probabilidad (Ω, \mathcal{A}, P) de forma que $P(\{\omega \in \Omega \mid (X(\omega), Y(\omega)) = (A, B)_l^*\}) = p_l$ con $p_l \in (0, 1)$, $l = 1, \dots, L$, $\sum_{l=1}^L p_l = 1$. Si $\mathbf{f}_n = (f_{n1}, \dots, f_{n(L-1)}) \in [0, 1]^{L-1}$, con $f_{nl} =$ frecuencia relativa de $(A, B)_l$ en la correspondiente realización de la*

muestra aleatoria simple de tamaño n ($l = 1, \dots, L-1$, $f_{nL} = 1 - \sum_{l=1}^{L-1} f_{nl}$), y $\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{f}_n)$ representa la covarianza muestral entre $\text{mid } X$ y $\text{mid } Y$ correspondiente, entonces, se cumple que la sucesión:

$$\left\{ \frac{\sqrt{n} [\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{f}_n) - \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})]}{\sqrt{\sigma^2(\mathbf{f}_n)}} \right\}_n$$

converge en ley hacia una distribución normal $\mathcal{N}(0, 1)$ cuando $n \rightarrow \infty$, siempre que $\sigma^2(\mathbf{p}) > 0$ y su estimador analógico $\sigma^2(\mathbf{f}_n) > 0$.

Bajo la hipótesis $H_0^{(1)}$ se cumple que $\sigma^2(\mathbf{p}) = 0$ si, y sólo si,

$$[\text{mid } X - E(\text{mid } X | \mathbf{p})] [\text{mid } Y - E(\text{mid } Y | \mathbf{p})]$$

es una variable aleatoria real degenerada en 0 (caso verdaderamente especial, carente de interés práctico y en el que, trivialmente, se cumple que $\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p}) = 0$). Además:

Observación 1.3.3 Aunque el enunciado del apartado *iii*) en la Proposición 1.3.5 es formalmente válido, hay que observar que $\sigma^2(\mathbf{p}) = 0$ implica en condiciones generales que $[\text{mid } X - E(\text{mid } X | \mathbf{p})] [\text{mid } Y - E(\text{mid } Y | \mathbf{p})]$ será una variable aleatoria real degenerada y, por lo tanto, se produce la igualdad $\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{f}_n) = \text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{p})$, de modo que el estadístico en el apartado *iii*) de la Proposición 1.3.5 tiene distribución (exacta) degenerada en el 0.

Como consecuencia del teorema y observación anteriores, se obtiene el siguiente procedimiento de contraste:

Teorema 1.3.7 *Para contrastar al nivel de significación nominal α_1 la hipótesis nula:*

$$H_0^{(1)} : \text{Cov}(\text{mid } X, \text{mid } Y | P) = 0$$

frente a la alternativa:

$$H_a^{(1)} : \text{Cov}(\text{mid } X, \text{mid } Y | P) \neq 0,$$

en el supuesto de que X e Y son conjuntos aleatorios tales que

$$[\text{mid } X - E(\text{mid } X | \mathbf{p})] [\text{mid } Y - E(\text{mid } Y | \mathbf{p})]$$

no sea casi seguro $[P]$ igual a 0, $H_0^{(1)}$ debe rechazarse siempre que

$$\frac{n [\text{Cov}(\text{mid } X, \text{mid } Y | \mathbf{f}_n)]^2}{\sigma^2(\mathbf{f}_n)} > \chi_{1, \alpha_1}^2,$$

donde χ_{1, α_1}^2 es el percentil de orden $100(1 - \alpha_1)$ de la distribución chi-cuadrado con 1 grado de libertad.

Además, la probabilidad de rechazar $H_0^{(1)}$ bajo la hipótesis alternativa converge a 1 cuando $n \rightarrow \infty$.

1.3.3.2 Contraste asintótico de $H_0^{(2)}$ frente a $H_a^{(2)}$

Recurriendo a argumentos análogos a los empleados en la Proposición 1.3.5, pero reemplazando los ‘mid’ por los correspondientes ‘spr’, se puede llegar al resultado siguiente:

Teorema 1.3.8 *Para cada $n \in \mathbb{N}$, se consideran n vectores bidimensionales de conjuntos aleatorios $(X_1, Y_1), \dots, (X_n, Y_n)$, independientes e idénticamente distribuidos que (X, Y) (es decir, una muestra aleatoria simple de tamaño n a partir de (X, Y)), definidos sobre el espacio de probabilidad (Ω, \mathcal{A}, P) de forma que $P(\{\omega \in \Omega \mid (X(\omega), Y(\omega)) = (A, B)_l^*\}) = p_l$ con $p_l \in (0, 1)$, $l =$*

$1, \dots, L$, $\sum_{l=1}^L p_l = 1$. Si $\mathbf{f}_n = (f_{n1}, \dots, f_{n(L-1)}) \in [0, 1]^{L-1}$, con f_{nl} = frecuencia relativa de $(A, B)_l$ en la correspondiente realización de la muestra aleatoria simple de tamaño n ($l = 1, \dots, L - 1$, $f_{nL} = 1 - \sum_{l=1}^{L-1} f_{nl}$), y $\text{Cov}(\text{spr } X, \text{spr } Y | \mathbf{f}_n)$ representa la covarianza muestral entre $\text{spr } X$ y $\text{spr } Y$ correspondientes, entonces, se cumple que la sucesión:

$$\left\{ \frac{\sqrt{n} [\text{Cov}(\text{spr } X, \text{spr } Y | \mathbf{f}_n) - \text{Cov}(\text{spr } X, \text{spr } Y | \mathbf{p})]}{\sqrt{\sigma_*^2(\mathbf{f}_n)}} \right\}_n$$

converge en ley hacia una distribución normal $\mathcal{N}(0, 1)$ cuando $n \rightarrow \infty$, siempre que la varianza:

$$\sigma_*^2(\mathbf{p}) = \text{Var}([\text{spr } X - E(\text{spr } X | \mathbf{p})] [\text{spr } Y - E(\text{spr } Y | \mathbf{p})] | \mathbf{p})$$

sea positiva y su estimador analógico $\sigma_*^2(\mathbf{f}_n) > 0$.

Cuando la hipótesis $H_0^{(2)} : \text{Cov}(\text{spr } X, \text{spr } Y) \leq 0$ es cierta, se cumple que $\sigma_*^2(\mathbf{p}) = 0$ si, y sólo si, $[\text{spr } X - E(\text{spr } X | \mathbf{p})] [\text{spr } Y - E(\text{spr } Y | \mathbf{p})]$ es una variable aleatoria real degenerada en un valor no positivo (situación poco general, pero mucho menos restrictiva que la de la anulación de $\sigma^2(\mathbf{p})$ bajo $H_0^{(1)}$) (véase, por ejemplo, el caso de valores de intervalo con la amplitud común para uno cualquiera de los conjuntos aleatorios).

Por otra parte, en condiciones generales la anulación de $\sigma_*^2(\mathbf{p}) = 0$ implica que $[\text{spr } X - E(\text{spr } X | \mathbf{p})] [\text{spr } Y - E(\text{spr } Y | \mathbf{p})]$ debe ser una variable aleatoria real degenerada y, por lo tanto, $\text{Cov}(\text{spr } X, \text{spr } Y | \mathbf{f}_n) = \text{Cov}(\text{spr } X, \text{spr } Y | \mathbf{p})$, de modo que carecería de sentido el estudio de la distribución del estadístico $2n [\text{Cov}(\text{spr } X, \text{spr } Y | \mathbf{f}_n) - \text{Cov}(\text{spr } X, \text{spr } Y | \mathbf{p})]$ según el apartado *iii*) de la Proposición 1.3.5. Por ello, se considera el contraste siguiente:

Teorema 1.3.9 *Para contrastar al nivel de significación nominal α_2 la hipótesis nula:*

$$H_0^{(2)} : \text{Cov}(\text{spr } X, \text{spr } Y) \leq 0$$

frente a la alternativa:

$$H_a^{(2)} : \text{Cov}(\text{spr } X, \text{spr } Y) > 0,$$

entonces, si $\sigma_^2(\mathbf{p}) > 0$ y $\sigma_*^2(\mathbf{f}_n) > 0$, la hipótesis $H_0^{(2)}$ debe rechazarse siempre que:*

$$\frac{\sqrt{n} \text{Cov}(\text{spr } X, \text{spr } Y | \mathbf{f}_n)}{\sqrt{\sigma_*^2(\mathbf{f}_n)}} > \gamma_{\alpha_2},$$

donde γ_{α_2} representa el percentil de orden $100(1 - \alpha_2)$ de la distribución normal típica.

Además, la probabilidad de rechazar $H_0^{(2)}$ bajo la hipótesis alternativa converge a 1 cuando $n \rightarrow \infty$.

1.3.3.3 Contraste asintótico de $H_0^{(3)}$ frente a $H_a^{(3)}$

Para contrastar la hipótesis $H_0^{(3)} : \text{Cov}(\text{spr } X, \text{spr } Y | P) = 0$ frente a la alternativa $H_a^{(3)}$, se consideraría de nuevo la muestra (Z_1^2, \dots, Z_n^2) , y se seguiría un proceso análogo al del contraste de $H_0^{(1)}$ frente a $H_a^{(1)}$, reemplazando $\text{mid } X$ por $\text{spr } X$ y $\text{mid } Y$ por $\text{spr } Y$, es decir:

Teorema 1.3.10 *Para contrastar al nivel de significación nominal α_3 la hipótesis nula:*

$$H_0^{(3)} : \text{Cov}(\text{spr } X, \text{spr } Y | P) = 0$$

frente a la alternativa:

$$H_a^{(3)} : \text{Cov}(\text{spr } X, \text{spr } Y | P) \neq 0,$$

entonces, si $\sigma_*^2(\mathbf{p}) > 0$ y $\sigma_*^2(\mathbf{f}_n) > 0$, la hipótesis $H_0^{(3)}$ debe rechazarse siempre que:

$$\frac{n [\text{Cov}(\text{spr } X, \text{spr } Y | \mathbf{f}_n)]^2}{\sigma_*^2(\mathbf{f}_n)} > \chi_{1, \alpha_3}^2.$$

Además, la probabilidad de rechazar $H_0^{(3)}$ bajo la hipótesis alternativa converge a 1 cuando $n \rightarrow \infty$.

1.3.3.4 Algoritmo para el contraste asintótico de H_0 frente a H_a cuando X e Y son conjuntos aleatorios simples

Para la resolución práctica del contraste (asintótico) a través de las diferentes etapas en que se ha desglosado el mismo, puede recurrirse al algoritmo siguiente, que resume lo examinado en la Subsección 1.3.3.

Paso 1. Para la muestra considerada, y prefijado el nivel de significación nominal α_1 , calcular el valor del estadístico

$$G = \frac{n [S_{\text{mid } X \text{ mid } Y}]^2}{\sigma^2(\mathbf{f}_n)}$$

y el valor crítico χ_{1, α_1}^2 , e ir al Paso 2.

Paso 2. SI el valor de G en la muestra es inferior o igual a χ_{1, α_1}^2 , ENTONCES ir al Paso 3, EN CASO CONTRARIO ir al Paso 4.

Paso 3. Para la muestra considerada, y prefijado el nivel de significación nominal α_2 , calcular el valor del estadístico

$$G' = \frac{\sqrt{n} S_{\text{spr } X \text{ spr } Y}}{\sqrt{\sigma_*^2(\mathbf{f}_n)}}$$

y el valor crítico γ_{2, α_2} . SI el valor de G' en la muestra es inferior o igual a γ_{2, α_2} , ENTONCES puede aceptarse H_0 a un nivel de

significación a lo sumo igual a $\alpha_1 + \alpha_2$, EN CASO CONTRARIO debe rechazarse H_0 a un nivel de significación a lo sumo igual a $\alpha_1 + \alpha_2$.

Paso 4. Para la muestra considerada, y prefijado el nivel de significación nominal α_2 , calcular el valor del estadístico G' y el valor $\gamma_{2\alpha_2}$ del Paso 3. SI el valor de G' en la muestra es superior a $\gamma_{2\alpha_2}$, ENTONCES debe rechazarse H_0 a un nivel de significación a lo sumo igual a $\alpha_1 + \alpha_2$, EN CASO CONTRARIO ir al Paso 5.

Paso 5. Para la muestra considerada, y prefijado el nivel de significación nominal α_3 , calcular el valor del estadístico

$$G'' = \frac{n [S_{\text{spr } X \text{ spr } Y}]^2}{\sigma_*^2(\mathbf{f}_n)}$$

y el valor crítico χ_{1, α_3}^2 . SI el valor de G'' en la muestra es inferior o igual a χ_{1, α_3}^2 , ENTONCES debe rechazarse H_0 a un nivel de significación a lo sumo igual a $\alpha_1 + \alpha_2 + \alpha_3$, EN CASO CONTRARIO recurrir a otra técnica de contraste para H_0 y H_a .

1.3.3.5 Contraste asintótico alternativo sobre R_{XY}^2 cuando X e Y son conjuntos aleatorios simples

En el caso en que X e Y sean conjuntos aleatorios simples, el contraste asintótico podría desarrollarse alternativamente siguiendo unas etapas diferentes a las de las subsecciones anteriores. El contraste resultaría globalmente algo más complejo, si bien siempre permitiría llegar a alguna conclusión (es decir, no se presentarían situaciones como la que aparece al final del Paso 5 en el último algoritmo). No obstante, algunos de los pasos de este procedimiento no podrían haberse realizado en general para conjuntos aleatorios con las características de las dos subsecciones precedentes.

En las condiciones y con las notaciones de esta subsección, para contrastar la hipótesis nula:

$$H_0 : |\text{Cov}(\text{mid } X, \text{mid } Y|P)| + 4 \sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P) \leq 0$$

frente a la alternativa

$$H_a : |\text{Cov}(\text{mid } X, \text{mid } Y|P)| + 4 \sigma_{f_{[0,1]}}^2 \text{Cov}(\text{spr } X, \text{spr } Y|P) > 0,$$

se puede proceder también de la forma siguiente:

- Se realiza en primer lugar el contraste entre las hipótesis:

$$\begin{cases} H_0^{(1*)} : \text{Cov}(\text{spr } X, \text{spr } Y|P) \leq 0 \equiv H_0^{(2)} \\ H_a^{(1*)} : \text{Cov}(\text{spr } X, \text{spr } Y|P) > 0 \equiv H_a^{(2)}. \end{cases}$$

- Si se concluye que debe rechazarse la hipótesis $H_0^{(1*)}$ en el contraste anterior para un nivel de significación prefijado nominal α_1 , entonces H_0 debe rechazarse a ese nivel.
- Si se concluye que puede aceptarse la hipótesis $H_0^{(1*)}$ en ese primer contraste para un nivel de significación prefijado nominal α_1 , entonces se desarrollará un segundo contraste, ahora entre las hipótesis:

$$\begin{cases} H_0^{(2*)} : [\text{Cov}(\text{mid } X, \text{mid } Y|P)]^2 - 16 \sigma_{f_{[0,1]}}^4 [\text{Cov}(\text{spr } X, \text{spr } Y|P)]^2 \leq 0 \\ H_a^{(2*)} : [\text{Cov}(\text{mid } X, \text{mid } Y|P)]^2 - 16 \sigma_{f_{[0,1]}}^4 [\text{Cov}(\text{spr } X, \text{spr } Y|P)]^2 > 0. \end{cases}$$

- Si puede aceptarse la hipótesis $H_0^{(2*)}$ en este contraste para un nivel de significación nominal prefijado α_2 , entonces podrá aceptarse la hipótesis nula inicial $H_0 : R_{XY}^2 = 0$ para un nivel de significación nominal (por la desigualdad de Bonferroni) a lo sumo igual a $\alpha_1 + \alpha_2$;

- Si debe rechazarse la hipótesis $H_0^{(2^*)}$ en este contraste para el nivel nominal prefijado α_2 , entonces se rechazará H_0 en el contraste inicial para un nivel de significación nominal (por la desigualdad de Bonferroni) a lo sumo igual a $\alpha_1 + \alpha_2$.

El primero de estos dos contrastes está recogido en el apartado 1.3.3.2 de esta subsección (concretamente, en el Teorema 1.3.9).

Por lo que se refiere al segundo contraste, y con argumentos análogos a los de los Teoremas 1.3.8 y 1.3.9, se concluye que:

Teorema 1.3.11 *Para cada $n \in \mathbb{N}$, se consideran n vectores bidimensionales de conjuntos aleatorios $(X_1, Y_1), \dots, (X_n, Y_n)$, independientes e idénticamente distribuidos que (X, Y) (es decir, una muestra aleatoria simple de tamaño n a partir de (X, Y)), definidos sobre el espacio de probabilidad (Ω, \mathcal{A}, P) de forma que $P(\{\omega \in \Omega \mid (X(\omega), Y(\omega)) = (A, B)_l^*\}) = p_l$ con $p_l \in (0, 1)$, $l = 1, \dots, L$, $\sum_{l=1}^L p_l = 1$. Si $\mathbf{f}_n = (f_{n1}, \dots, f_{n(L-1)}) \in [0, 1]^{L-1}$, con $f_{nl} =$ frecuencia relativa de $(A, B)_l$ en la correspondiente realización de la muestra aleatoria simple de tamaño n ($l = 1, \dots, L-1$, $f_{nL} = 1 - \sum_{l=1}^{L-1} f_{nl}$), y $\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{f}_n)$, $\text{Cov}(\text{spr } X, \text{spr } Y \mid \mathbf{f}_n)$ representan las covarianzas muestrales entre $\text{mid } X$ y $\text{mid } Y$, $\text{spr } X$ y $\text{spr } Y$ correspondientes, respectivamente, entonces, se cumple que la sucesión:*

$$\left\{ \sqrt{\frac{n}{\sigma_{**}^2(\mathbf{f}_n)}} \left[\left([\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{f}_n)]^2 - 16 \sigma_{f_{[0,1]}}^4 [\text{Cov}(\text{spr } X, \text{spr } Y \mid \mathbf{f}_n)]^2 \right) - \left([\text{Cov}(\text{mid } X, \text{mid } Y \mid \mathbf{p})]^2 - 16 \sigma_{f_{[0,1]}}^4 [\text{Cov}(\text{spr } X, \text{spr } Y \mid \mathbf{p})]^2 \right) \right] \right\}_n$$

converge en ley hacia una distribución normal $\mathcal{N}(0, 1)$ cuando $n \rightarrow \infty$, siempre que la varianza:

$$\begin{aligned} & \sigma_{**}^2(\mathbf{p}) \\ = & \text{Var}\left(\text{Cov}(\text{mid } X, \text{mid } Y|\mathbf{p}) [\text{mid } X - E(\text{mid } X|\mathbf{p})] [\text{mid } Y - E(\text{mid } Y|\mathbf{p})] \right. \\ & \left. - 16 \sigma_{f[0,1]}^4 \text{Cov}(\text{spr } X, \text{spr } Y|\mathbf{p}) [\text{spr } X - E(\text{spr } X|\mathbf{p})] [\text{spr } Y - E(\text{spr } Y|\mathbf{p})] \mid \mathbf{p}\right) \\ & \text{sea positiva y su estimador análogo } \sigma_{**}^2(\mathbf{f}_n) > 0. \end{aligned}$$

Si $\sigma_{**}^2(\mathbf{p}) = 0$ y $\sigma_{**}^2(\mathbf{f}_n) = 0$, lo cual sólo ocurrirá ocasionalmente, el resultado anterior podría ampliarse recurriendo al desarrollo de Taylor de segundo orden y razonando como en el *iii*) de la Proposición 1.3.5. En condiciones de positividad de $\sigma_{**}^2(\mathbf{p})$ y $\sigma_{**}^2(\mathbf{f}_n)$ se tiene, como consecuencia, que:

Teorema 1.3.12 *Para contrastar al nivel de significación nominal α_2 la hipótesis nula:*

$$H_0^{(2*)} : [\text{Cov}(\text{mid } X, \text{mid } Y|P)]^2 - 16 \sigma_{f[0,1]}^4 [\text{Cov}(\text{spr } X, \text{spr } Y|P)]^2 \leq 0$$

frente a la alternativa:

$$H_a^{(2*)} : [\text{Cov}(\text{mid } X, \text{mid } Y|P)]^2 - 16 \sigma_{f[0,1]}^4 [\text{Cov}(\text{spr } X, \text{spr } Y|P)]^2 > 0,$$

*entonces, si $\sigma_{**}^2(\mathbf{p}) > 0$ y $\sigma_{**}^2(\mathbf{f}_n) > 0$, la hipótesis $H_0^{(2*)}$ debe rechazarse siempre que:*

$$\begin{aligned} & \sqrt{\frac{n}{\sigma_{**}^2(\mathbf{f}_n)}} [(\text{Cov}(\text{mid } X, \text{mid } Y|\mathbf{f}_n))^2 \\ & - 16 \sigma_{f[0,1]}^4 (\text{Cov}(\text{spr } X, \text{spr } Y|\mathbf{f}_n))^2] > \gamma_{\alpha_2}. \end{aligned}$$

1.3.4 Contrastes sobre el coeficiente a en modelos lineales

En cualquier estudio científico cuyo objetivo sea determinar si un elemento aleatorio proporciona información lineal útil acerca de otra característica aleatoria medida sobre la misma población parece natural suponer que, en principio, la relación entre ambas variables se puede expresar a partir de una recta más un error aleatorio y discernir si la pendiente de esa recta es cero (es decir, la variable dependiente no aporta información lineal).

Así, si (Ω, \mathcal{A}, P) es un espacio de probabilidad y (X, Y) es un conjunto aleatorio bidimensional asociado a él, la admisión del modelo lineal implica la existencia de $a \in \mathbb{R}$ y $B \in \mathcal{K}_c(\mathbb{R})$, desconocidos pero fijos, tales que $Y|X = aX + \epsilon_X$ donde ϵ_X es un conjunto aleatorio cuyo valor esperado $E^A[\epsilon_X|P] = B$. En estas condiciones se puede probar el siguiente resultado:

Proposición 1.3.13 *Si (X, Y) es un conjunto aleatorio bidimensional con valores en $\mathcal{K}_c(\mathbb{R}) \times \mathcal{K}_c(\mathbb{R})$ para el que existen $a \in \mathbb{R}$, $B \in \mathcal{K}_c(\mathbb{R})$ y un conjunto aleatorio ϵ_X con $E^A[\epsilon_X|P] = B$, que verifican que:*

$$(Y|X) = aX + \epsilon_X$$

y X no es degenerado, las condiciones siguientes son equivalentes:

- i) $a = 0$;*
- ii) $\text{Cov}(\text{mid } X, \text{mid } Y|P) = 0$ y $\text{Cov}(\text{spr } X, \text{spr } Y|P) = 0$.*

Demostración. Nótese, en primer lugar, que $(Y|X) = aX + \epsilon_X$ implica que $E^A[Y|X] = aX + B$, denotando abreviadamente por $E^A[Y|X]$ el valor esperado de Y condicionado por cada valor de conjunto de X . Se cumple, por tanto, que:

$$\text{mid } E^A[Y|X] = a \text{mid } X + \text{mid } B, \quad \text{spr } E^A[Y|X] = |a| \text{spr } X + \text{spr } B.$$

En consecuencia:

$$E(\text{mid } Y | X) = a \text{ mid } X + \text{mid } B, \quad E(\text{spr } Y | X) = |a| \text{ spr } X + \text{spr } B.$$

Por otro lado, al ser ‘mid’ una función medible, $\sigma(\text{mid } X) \subset \sigma(X)$, y las propiedades de la esperanza condicionada aseguran que:

$$\begin{aligned} E(\text{mid } Y | \text{mid } X) &= E(E(\text{mid } Y | X) | \text{mid } X) \\ &= E(a \text{ mid } X + \text{mid } B | \text{mid } X) = a \text{ mid } X + \text{mid } B, \end{aligned}$$

es decir, se verifica el modelo de regresión lineal real entre las variables aleatorias $\text{mid } X$ y $\text{mid } Y$ determinado por $a \text{ mid } X + \text{mid } B$, y, por tanto, se tiene que:

$$a = \frac{\text{Cov}(\text{mid } X, \text{mid } Y | P)}{S_{\text{mid } X}^2}.$$

De la misma manera se comprueba que $E(\text{spr } Y | \text{spr } X) = |a| \text{ spr } X + \text{spr } B$, y entonces:

$$|a| = \frac{\text{Cov}(\text{spr } X, \text{spr } Y | P)}{S_{\text{spr } X}^2}.$$

Así, si $a = 0$ se tiene que $\text{Cov}(\text{mid } X, \text{mid } Y | P) = 0$ y $\text{Cov}(\text{spr } X, \text{spr } Y | P) = 0$. Recíprocamente, dado que X es no degenerado, o bien la variable $\text{mid } X$ es no degenerada (en cuyo caso $\text{Cov}(\text{mid } X, \text{mid } Y | P) = 0$ implica que $a = 0$), o bien la variable $\text{spr } X$ es no degenerada (en cuyo caso $\text{Cov}(\text{spr } X, \text{spr } Y | P) = 0$ implica que $a = 0$). \square

Si se dispone únicamente de información muestral para determinar si la relación supuesta tiene o no “pendiente” nula (es decir, que no existe relación lineal de Y en función de X), la Proposición 1.3.13 asegura que contrastar $H_0^\circ : a = 0$ frente a $H_a^\circ : a \neq 0$ es equivalente a contrastar:

$$H_0^\circ : \text{Cov}(\text{mid } X, \text{mid } Y | P) = 0 \quad \text{y} \quad \text{Cov}(\text{spr } X, \text{spr } Y | P) = 0$$

o, equivalentemente, $[\text{Cov}(\text{mid } X, \text{mid } Y|P)]^2 + [\text{Cov}(\text{spr } X, \text{spr } Y|P)]^2 = 0$, frente a la alternativa:

$$H_a^\circ : \text{Cov}(\text{mid } X, \text{mid } Y|P) \neq 0 \quad \text{o} \quad \text{Cov}(\text{spr } X, \text{spr } Y|P) \neq 0.$$

Así, si se dispone de una una muestra aleatoria $(X_1, Y_1), \dots, (X_n, Y_n)$, se tiene que $Y_i = aX_i + \epsilon_{X_i}$ con ϵ_{X_i} conjuntos aleatorios independientes. La equivalencia anterior hace que el problema se reduzca a realizar un test multivariante acerca de una afirmación sobre la matriz Σ de varianzas-covarianzas del vector aleatorio $Z = (\text{mid } X, \text{mid } Y, \text{spr } X, \text{spr } Y)$ a partir de una muestra aleatoria $Z_i = (\text{mid } X_i, \text{mid } Y_i, \text{spr } X_i, \text{spr } Y_i)$ con $i = 1 \dots, n$.

Para ello se puede recurrir a técnicas similares a las desarrolladas en las subsecciones precedentes, e incluso la suposición adicional de independencia entre los vectores aleatorios $Z^1 = (\text{mid } X, \text{mid } Y)$ y $Z^2 = (\text{spr } X, \text{spr } Y)$ en condiciones de normalidad daría lugar a nuevos procedimientos.

Por otro lado, resulta especialmente interesante el empleo del *método bootstrap* desarrollado en Beran and Srivastava (1985) ya que, como allí se señala, los test multivariantes clásicos suelen requerir la hipótesis de normalidad y, en cambio, al aplicar las técnicas bootstrap se consigue un test asintótico válido cualquiera que sea la distribución de partida (sin necesidad de exigir que el número de valores distintos de cada conjunto aleatorio sea finito). De esa manera, contrastar:

$$H_0^\circ : \text{Cov}(\text{mid } X, \text{mid } Y|P) = 0 \quad \text{y} \quad \text{Cov}(\text{spr } X, \text{spr } Y|P) = 0$$

frente a

$$H_a^\circ : \text{Cov}(\text{mid } X, \text{mid } Y|P) \neq 0 \quad \text{o} \quad \text{Cov}(\text{spr } X, \text{spr } Y|P) \neq 0,$$

es equivalente a comprobar si los elementos de posiciones (1,2) y (3,4) (y sus simétricos) de la matriz Σ son o no cero.

Los resultados de Beran and Srivastava (1985) se pueden particularizar a este caso y el método bootstrap se expresaría como sigue:

- En primer lugar se calcula la población bootstrap

$$V_i = \left(\pi(\widehat{S})^{(1/2)} \widehat{S}^{(-1/2)} Z_i^T \right)^T$$

para todo $i = 1, \dots, n$, donde para una matriz 4×4 , \mathfrak{M} , $\pi(\mathfrak{M})$ es la matriz que deja invariantes todos los elementos de \mathfrak{M} salvo \mathbf{m}_{12} , \mathbf{m}_{34} , \mathbf{m}_{21} y \mathbf{m}_{43} a los que les asocia el valor 0 y \widehat{S} denota la matriz de varianzas-covarianzas muestral, esto es:

$$\widehat{S} = \left(\begin{array}{c|c} \widehat{S}_{11} & \widehat{S}_{12} \\ \hline \widehat{S}_{21} & \widehat{S}_{22} \end{array} \right).$$

- Se remuestra de la población (V_1, \dots, V_n) para conseguir una nueva muestra (V_1^*, \dots, V_n^*) . El estadístico del contraste bootstrap es $T_n^* = n h(\widehat{S}^*)$ donde $h(\mathfrak{M}) = \mathbf{m}_{12}^2 + \mathbf{m}_{34}^2$ y \widehat{S}^* denota la matriz de varianzas-covarianzas asociada a la muestra (V_1^*, \dots, V_n^*) . La distribución de este estadístico se puede aproximar por el método de Monte Carlo.

A partir de estas etapas, se concluye que:

Teorema 1.3.14 *Para contrastar al nivel de significación nominal α la hipótesis nula:*

$$H_0^o : a = 0$$

frente a

$$H_0^o : a \neq 0,$$

bajo condiciones de linealidad (es decir, $Y_i = aX_i + \epsilon_{X_i}$, $i = 1, \dots, n$, con ϵ_{X_i} conjuntos aleatorios independientes), entonces, H_0^o debe rechazarse siempre que:

$$T_n = n h(\widehat{S}) > c_\alpha,$$

donde c_α denota el $100(1 - \alpha)$ percentil de la distribución de T_n^* .

Observación 1.3.4 La condición de linealidad entre X e Y que se ha supuesto en esta subsección, tiene un valor añadido: las hipótesis y conclusiones estadísticas no dependen de la métrica considerada sobre $\mathcal{K}_c(\mathbb{R})$ (a diferencia de los contrastes sobre R_{XY}^2 en la subsección precedente).

Observación 1.3.5 El Teorema 1.3.14 es una particularización de los resultados de Beran and Srivastava (1985), por lo que no es preciso insistir en sus ventajas operacionales y en su idoneidad con muestras de cualquier tamaño frente a los demás procedimientos. Por ello, en este caso no van a desarrollarse simulaciones que comparen las diferentes técnicas de contraste.

No obstante, para los procedimientos de contraste en el Capítulo 2, sí se llevarán a cabo simulaciones con fines comparativos.

1.3.5 Aplicación sobre un ejemplo real

Si los datos proporcionados por el Servicio de Nefrología del Hospital Valle del Nalón de Langreo (Asturias) en el Ejemplo 1.1.3 son datos muestrales a partir de una población de 3000 individuos, resultaría que:

Ejemplo 1.3.1 Suponiendo que existe una relación lineal del tipo $Y|X = aX + \epsilon_X$ donde ϵ_X es un conjunto aleatorio cuyo valor esperado es B , al aplicar el método bootstrap indicado en el Teorema 1.3.14 para contrastar $H_0^\circ : a = 0$ frente a $H_a^\circ : a \neq 0$, se obtendría un p -valor igual a 0.0068. En consecuencia, a los niveles de significación usuales se rechazaría la hipótesis nula y, por tanto, se podría concluir (a dichos niveles de significación) que la variable X aporta información acerca de la variable Y según el modelo lineal propuesto.

1.4 Valoración final y problemas abiertos

Como valoración final del Capítulo 1 cabe señalar las contribuciones realizadas en cuanto a:

- la determinación de las soluciones generales de los problemas de regresión lineal entre dos conjuntos aleatorios compactos y convexos de \mathbb{R} (hasta el momento sólo se disponía de soluciones para casos particulares en los que la discusión era bastante inmediata);
- la definición de índices del grado de dependencia a través de las relaciones óptimas obtenidas en el análisis de la regresión lineal (cuestión absolutamente novedosa para la que no existía metodología previa);
- la construcción de técnicas estadísticas inferenciales relativas al nuevo parámetro real “grado de dependencia entre los dos conjuntos aleatorios” considerados (tema para el que tampoco existían antecedentes en la literatura).

Aunque la suposición de que W es simétrica no resulta restrictiva en el contexto del tratamiento de mínimos cuadrados para el problema de regresión en este capítulo, el análisis del problema en el caso en el que se prescindiera de esa suposición podría llevarse a cabo siguiendo la ideas en la Observación 1.1.4 y considerando la función objetivo definida por:

$$\begin{aligned} \zeta(a, b, c) = E & \left[\left((\text{mid } Y - a \text{ mid } X - c) \right. \right. \\ & \left. \left. + (2\mu_{f_{[0,1]}} - 1)(\text{spr } Y - |a| \text{ spr } X - b) \right)^2 \right. \\ & \left. + 4\sigma_{f_{[0,1]}}^2 (\text{spr } Y - |a| \text{ spr } X - b)^2 \middle| P \right] \end{aligned}$$

con $a \in \mathbb{R}$, $b = \text{spr } B \in [0, +\infty)$ y $c = \text{mid } B \in \mathbb{R}$. La solución puede obtenerse empleando técnicas tradicionales para los problemas de regresión/correlación.

Por otro lado, cuando el conjunto aleatorio Y toma valores de intervalo y el X es casi seguro real, Diamond (1990) ha señalado que una relación del tipo $Y = AX + B$ con $A, B \in \mathcal{K}_c(\mathbb{R})$ resulta mucho más natural que las que se examinan en este capítulo, y ha desarrollado una discusión sobre la solución del problema. Además, Körner and Näther (1998) (ver también Näther y Körner, 2002, como referencia más reciente y amplia) han estudiado la regresión para esta relación y sus soluciones para el caso más general en el que Y toma valores en el espacio $\mathcal{F}_c(\mathbb{R})$.

En relación con la investigación de este capítulo, se plantean varios problemas de interés para un futuro próximo. Entre ellos cabe citar:

- Los estudios y discusiones en este capítulo pueden extenderse a situaciones más complejas, como la de la regresión múltiple del tipo:

$$Y = a_1X_1 + \dots + a_kX_k + B,$$

con $a_1, \dots, a_k \in \mathbb{R}$ y $B \in \mathcal{K}_c(\mathbb{R})$, donde X_1, \dots, X_k e Y son conjuntos aleatorios con valores en $\mathcal{K}_c(\mathbb{R})$. En este caso, la determinación de las soluciones óptimas y de índices sobre la idoneidad de la relación lineal implicarán la consideración de 2^k conos, lo que habitualmente derivará en una complejidad enorme que tal vez habrá que abordar desde otra metodología.

- Es también interesante extender los análisis del capítulo reemplazando los conjuntos aleatorios por variables aleatorias difusas, si bien deberá comenzarse tal extensión limitando la forma de los valores de dichas variables de modo que dependiendo de esa forma los valores puedan caracterizarse por dos o más parámetros.

- Otra cuestión asociada que sería útil examinar es la de la búsqueda de relaciones funcionales no lineales que tengan interpretación interesante y se adapten adecuadamente al método de mínimos cuadrados basado en la métrica D_W .

Capítulo 2

Contraste de hipótesis sobre el valor esperado difuso de variables aleatorias difusas

En el capítulo anterior se han presentado algunos contrastes de hipótesis sobre una característica relativa a conjuntos aleatorios y que toma valor real. En este capítulo, el objetivo general es el desarrollo de algunas técnicas de contraste de hipótesis sobre una característica relativa a variables aleatorias difusas y que a su vez toma valor difuso: el valor esperado en el sentido de Puri and Ralescu (1986).

En primer lugar, se llevarán a cabo estudios sobre el contraste de la igualdad del valor esperado de una variable aleatoria difusa con un elemento prefijado de $\mathcal{F}_c(\mathbb{R})$, en segundo lugar sobre el contraste de la igualdad de los valores esperados de dos variables aleatorias difusas, y en tercer lugar sobre el contraste tipo ANOVA (es decir, sobre la igualdad de valores esperados de una variable aleatoria difusa en varias poblaciones).

En estos estudios, sobre la base de diferentes metodologías se construirán procedimientos de contraste para dos casos: el de las variables aleatorias difusas normales (operativo, pero difícilmente un modelo ajustable a los problemas reales), y el de las variables aleatorias difusas simples (menos operativo, pero ajustable a prácticamente todos los problemas reales). Estos procedimientos tienen un fundamento común, que es el empleo de la métrica D_W^φ introducida en la Subsección 0.3.2.

Para las variables normales, se define un estadístico que se reduce al del contraste clásico de la media de una población normal (tanto si la varianza es conocida, como si es desconocida), de forma que las conclusiones inferenciales serán inmediatas.

Para las variables simples se examinarán dos tipos de técnicas: asintóticas y bootstrap. Las primeras están inspiradas en la misma idea de los desarrollos en la Subsección 1.3.3, es decir, en aprovechar las propiedades de ciertos estimadores consistentes y asintóticamente normales. No obstante, a diferencia de lo que ocurría para los contrastes asintóticos del coeficiente de determinación en el capítulo anterior, los del presente utilizan la aproximación de la combinación lineal de chi-cuadrado en vez de la normal. Las técnicas bootstrap permitirán aproximar la distribución de un estadístico basado en el remuestreo a partir de una muestra de referencia. Estos dos tipos de técnicas se compararán a través de estudios de simulación de variables aleatorias difusas.

Los resultados obtenidos se ilustrarán con algunos ejemplos.

Finalmente, se incluirán conclusiones y problemas abiertos relacionados.

2.1 Contrastes sobre la media poblacional de una variable aleatoria difusa con valores en $\mathcal{F}_c(\mathbb{R})$

Las inferencias tradicionales para contrastar si el valor de la media poblacional de una variable aleatoria real puede admitirse que es igual a (o bien mayor o igual que, o menor o igual que) un número real dado, se basan habitualmente en la diferencia o en la distancia euclídea entre la correspondiente media muestral y la poblacional. La distribución exacta del estadístico que involucra dicha diferencia o distancia suele determinarse bajo condiciones de normalidad de la variable.

En la presente sección va a analizarse en primer término el contraste bilateral para el valor esperado (difuso) de una variable aleatoria difusa, en el caso en el que esta variable tiene distribución normal en el sentido definido por Puri and Ralescu (1986). Se comprobará que el procedimiento en el que la distancia euclídea en \mathbb{R} se reemplaza por la métrica D_W^c en $\mathcal{F}_c(\mathbb{R})$, se reduce a un contraste de hipótesis con variables aleatorias reales normales, lo que redundará en la simplicidad y operatividad del método.

Sin embargo, como ya se ha señalado, la suposición de normalidad para una variable aleatoria difusa representa una exigencia que no se ajusta a la mayoría de los problemas reales, especialmente al obligar a que todos los valores de la variable adopten la misma forma aunque con diferente localización.

Para evitar los inconvenientes prácticos que plantea este hecho, se desarrollan en esta sección otras dos aproximaciones del problema de contraste: la basada en la Teoría de Grandes Muestras y la basada en las técnicas bootstrap.

Se ilustrará la aplicación de los procedimientos construidos con algunos ejemplos reales, y finalmente se compararán los mismos mediante estudios de simulación basados en las ideas de Colubi *et al.* (2002).

2.1.1 Test bilateral para el valor esperado de una variable aleatoria difusa normal

Sean (Ω, \mathcal{A}, P) un espacio de probabilidad, $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c(\mathbb{R})$ una variable aleatoria difusa normal asociada a ese espacio con varianza σ^2 y $\tilde{E}(\mathcal{X}|P)$ su media poblacional.

El primer problema que se aborda es el de contrastar la hipótesis nula $H_0 : \tilde{E}(\mathcal{X}|P) = \tilde{V} \in \mathcal{F}_c(\mathbb{R})$ (es decir, $[D_W^\varphi(\tilde{E}(\mathcal{X}|P), \tilde{V})]^2 = 0$) frente a la alternativa $H_a : \tilde{E}(\mathcal{X}|P) \neq \tilde{V}$ (es decir, $[D_W^\varphi(\tilde{E}(\mathcal{X}|P), \tilde{V})]^2 > 0$), cuando se supone que σ^2 es desconocida. Para ello, se considera una muestra aleatoria simple, $\mathcal{X}_1, \dots, \mathcal{X}_n$ a partir de \mathcal{X} , y se denota por $\bar{\mathcal{X}}$ la media muestral, es decir:

$$\bar{\mathcal{X}} = \frac{1}{n} (\mathcal{X}_1 \oplus \dots \oplus \mathcal{X}_n).$$

Para este propósito va a emplearse el estadístico:

$$Z = \frac{[D_W^\varphi(\bar{\mathcal{X}}, \tilde{V})]^2 - [D_W^\varphi(\tilde{E}(\mathcal{X}|P), \tilde{V})]^2}{\frac{S^2}{n-1}},$$

donde $S^2 = \frac{1}{n} \sum_{i=1}^n [D_W^\varphi(\mathcal{X}_i, \bar{\mathcal{X}})]^2$.

Por las propiedades de la función $f(\cdot, \cdot)$ y de la métrica D_W^φ en la Subsección 0.3.2 (ver Lubiano Gómez, 1999), si se expresa $\mathcal{X}_i = \xi_i \oplus \tilde{E}(\mathcal{X}|P)$ (y, por tanto, $\bar{\mathcal{X}} = \bar{\xi} \oplus \tilde{E}(\mathcal{X}|P)$), puede concluirse que cuando H_0 es cierta:

$$\left[D_W^\varphi(\bar{\mathcal{X}}, \tilde{V}) \right]^2 - \left[D_W^\varphi(\tilde{E}(\mathcal{X}|P), \tilde{V}) \right]^2 = \bar{\xi}^2$$

y:

$$S^2 = \frac{1}{n} \sum_{i=1}^n [\xi_i - \bar{\xi}]^2,$$

por lo que Z coincide con el estadístico clásico para poblaciones normales no difusas con media 0, es decir:

$$\frac{n(n-1)\bar{\xi}^2}{\sum_{i=1}^n [\xi_i - \bar{\xi}]^2}$$

y, en consecuencia, bajo H_0 (que implica $E(\xi|P) = 0$) el estadístico Z tiene distribución F de Snedecor con 1 y $n-1$ grados de libertad. Por todo ello, puede afirmarse que:

Teorema 2.1.1 *Para contrastar al nivel de significación α la hipótesis nula:*

$$H_0 : \tilde{E}(\mathcal{X}|P) = \tilde{V}$$

frente a la alternativa:

$$H_a : \tilde{E}(\mathcal{X}|P) \neq \tilde{V},$$

H_0 debe rechazarse siempre que:

$$\frac{n(n-1) \left[D_W^\varphi(\bar{\mathcal{X}}, \tilde{V}) \right]^2}{\sum_{i=1}^n \left[D_W^\varphi(\mathcal{X}_i, \bar{\mathcal{X}}) \right]^2} > F_{1, n-1, \alpha},$$

donde $F_{1, n-1, \alpha}$ denota el percentil de orden $100(1-\alpha)$ de la distribución F de Snedecor con 1 y $n-1$ grados de libertad.

El segundo problema que va a tratarse es el de contrastar la hipótesis nula H_0 frente a la alternativa H_a del problema anterior, cuando se supone que la varianza σ^2 es conocida y se dispone de una muestra aleatoria simple, $\mathcal{X}_1, \dots, \mathcal{X}_n$ a partir de \mathcal{X} . Razonando en forma análoga al problema precedente, puede concluirse que:

Teorema 2.1.2 *Para contrastar al nivel de significación α la hipótesis nula:*

$$H_0 : \tilde{E}(\mathcal{X}|P) = \tilde{V}$$

frente a la alternativa:

$$H_a : \tilde{E}(\mathcal{X}|P) \neq \tilde{V},$$

H_0 debe rechazarse siempre que:

$$\frac{n \left[D_W^\varphi(\bar{\mathcal{X}}, \tilde{V}) \right]^2}{\sigma^2} > \chi_{1, \alpha}^2,$$

donde $\chi_{1, \alpha}^2$ es el percentil de orden $100(1-\alpha)$ de la distribución chi-cuadrado con 1 grado de libertad.

En resumen, e independientemente de las medidas de ponderación W y φ que se hayan elegido, si H_0 es cierta pueden utilizarse los tests para las variables aleatorias reales normales correspondientes a las *diferencias de Hukuhara* $\xi_i = \mathcal{X}_i -_h \tilde{E}(\mathcal{X}|P)$, que en la situación supuesta son sencillas de determinar. Por esta razón, la distribución de los estadísticos de contraste cuando $\tilde{E}(\mathcal{X}|P)$ corresponde a una traslación real a partir de \tilde{V} (suposición adecuada en el contexto de las variables aleatorias difusas normales) coincide con la de los contrastes gaussianos clásicos (téngase en cuenta que $D_W^\varphi(\tilde{V} + c, \tilde{V}) = |c|$ para cualquier $c \in \mathbb{R}$), de modo que podría determinarse la potencia de los tests anteriores en forma análoga a la de los correspondientes con variables aleatorias reales normales.

Los resultados que acaban de establecerse se ilustran a continuación:

Ejemplo 2.1.1 En una empresa se ha desarrollado un estudio sobre el tiempo que dedican sus trabajadores a la comida del mediodía, con el objeto de ajustar en lo posible el horario común. Para este fin, y ya que la empresa es de gran tamaño, se selecciona al azar una muestra de 20 empleados. Un supervisor se responsabiliza de medir el tiempo \mathcal{X} que destina a esa comida cada uno de los trabajadores de la muestra en un día concreto. Aunque las medidas suelen expresarse en términos de valores exactos, los expertos en este tipo de problemas indican que la observación ‘ m minutos’ estaría descrita de forma más apropiada mediante el número difuso triangular simétrico $\tilde{m} = \text{Tri}(m - 1, m, m + 1)$ que se representa en la Figura 2.1.

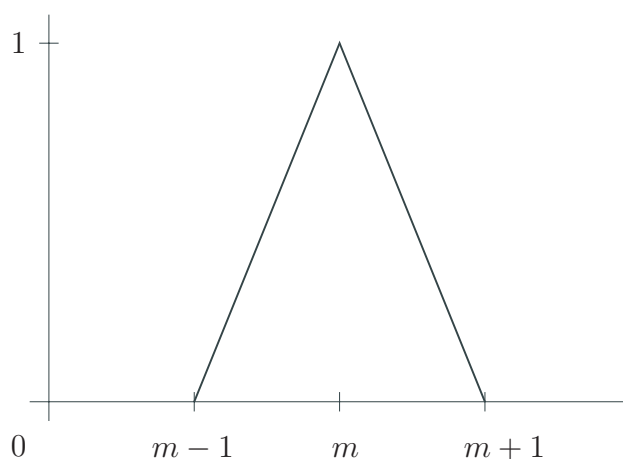


Fig. 2.1: Descripción, según los expertos, de la observación ‘ m minutos’ en el Ejemplo 2.1.1

Por otro lado, experiencias anteriores apuntan a que la distribución de m es normal.

La empresa se plantea ajustar el horario admitiendo que el tiempo medio dedicado a la comida del mediodía es de ‘20 minutos’ (más concretamente, $\tilde{20} = \text{Tri}(19, 20, 21)$).

Si la realización muestral disponible es $(\tilde{9}, \tilde{12}, \tilde{14}, \tilde{15}, \tilde{16}, \tilde{16}, \tilde{17}, \tilde{20}, \tilde{21}, \tilde{22}, \tilde{23}, \tilde{27}, \tilde{29}, \tilde{29}, \tilde{31}, \tilde{32}, \tilde{33}, \tilde{34}, \tilde{38}, \tilde{41})$, y para contrastar $H_0 : \tilde{E}(\mathcal{X}|P) = \tilde{20}$ se emplea el resultado del Teorema 2.1.1, se obtendrá que el p -valor es 0.069. Como consecuencia, al nivel de significación $\alpha = 0.05$ no se rechazaría la posibilidad de ajustar el tiempo medio dedicado a la comida del mediodía en ‘20 minutos’.

2.1.2 Test bilateral asintótico para el valor esperado de una variable aleatoria difusa simple

Cuando no parece admisible la normalidad de la variable aleatoria real, los procedimientos tradicionales de contraste sobre la media, aplicables en condiciones bastante generales, están basados a menudo en la Teoría de Grandes Muestras.

En el caso de variables aleatorias difusas simples (condición que no es totalmente general desde un punto de vista teórico, pero sí lo es en la práctica), puede seguirse un proceso análogo al de la Subsección 1.3.3, para lo cual se considera un estadístico en el que figura el valor esperado (difuso) muestral y se aprovechan sus propiedades como estimador consistente y asintóticamente normal.

De este modo, cuando sobre una población cualquiera se observa una variable aleatoria difusa simple, puede definirse un parámetro vectorial para el que su estimador analógico es consistente y asintóticamente normal, y el valor esperado difuso poblacional es una función de dicho parámetro cuyo estimador analógico va a heredar las propiedades fundamentales para grandes muestras.

Sean (Ω, \mathcal{A}, P) un espacio de probabilidad y $\mathcal{X} : \Omega \rightarrow \mathcal{F}_c(\mathbb{R})$ una variable aleatoria difusa simple asociada a ese espacio. Sean $\tilde{x}_1, \dots, \tilde{x}_L \in \mathcal{F}_c(\mathbb{R})$

los distintos valores de \mathcal{X} en la población, y p_1, \dots, p_L las probabilidades respectivas ($P(\{\omega \in \Omega | \mathcal{X}(\omega) = \tilde{x}_l\}) = p_l > 0, l = 1, \dots, L, \sum_{l=1}^L p_l = 1$).

Si se denota la media poblacional de \mathcal{X} por $\tilde{E}(\mathcal{X}|\mathbf{p})$, con $\mathbf{p} = (p_1, \dots, p_{L-1}) \in (0, 1)^{L-1}$, se tiene que:

$$\tilde{E}(\mathcal{X}|\mathbf{p}) = p_1 \tilde{x}_1 \oplus \dots \oplus p_L \tilde{x}_L.$$

Al igual que en la Subsección 1.3.3, el parámetro vectorial \mathbf{p} cumple las condiciones de regularidad allí indicadas.

Sea $\mathcal{X}_1, \dots, \mathcal{X}_n$ una muestra aleatoria simple obtenida a partir de la variable aleatoria difusa \mathcal{X} . Se denotará por f_{nl} el estadístico correspondiente a la frecuencia relativa de \tilde{x}_l en la muestra aleatoria, y por $\mathbf{f}_n = (f_{n1}, \dots, f_{n(L-1)})$. Obviamente,

$$\bar{\mathcal{X}} = \tilde{E}(\mathcal{X}|\mathbf{f}_n) = f_{n1} \tilde{x}_1 \oplus \dots \oplus f_{nL} \tilde{x}_L.$$

Como se recordó en el capítulo anterior, la sucesión de estimadores $\{\mathbf{f}_n\}_n$ es fuertemente consistente y se distribuye asintóticamente según una distribución normal $(L-1)$ -dimensional $\mathcal{N}\left(\mathbf{p}, [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1}/n\right)$ cuando $n \rightarrow \infty$, y $n(\mathbf{f}_n - \mathbf{p})\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})(\mathbf{f}_n - \mathbf{p})^t$ converge en distribución a una chi-cuadrado χ_{L-1}^2 cuando $n \rightarrow \infty$. En consecuencia:

Proposición 2.1.3 *Para cada $n \in \mathbb{N}$, se consideran n variables aleatorias difusas independientes e igualmente distribuidas que \mathcal{X} , definidas sobre el espacio de probabilidad (Ω, \mathcal{A}, P) de forma que $P(\{\omega \in \Omega | \mathcal{X}(\omega) = \tilde{x}_l\}) = p_l \in (0, 1), l = 1, \dots, L, \sum_{l=1}^L p_l = 1$. Se cumple entonces que:*

- i) Si $\mathbf{f}_n = (f_{n1}, \dots, f_{n(L-1)}) \in [0, 1]^{L-1}$, con $f_{nl} =$ frecuencia relativa de \tilde{x}_l en la correspondiente realización de la muestra aleatoria simple de tamaño n ($l = 1, \dots, L-1, f_{nL} = 1 - \sum_{l=1}^{L-1} f_{nl}$), y $\tilde{E}(\mathcal{X}|\mathbf{f}_n)$ representa

la media muestral de \mathcal{X} , entonces, cualquiera que sea $\tilde{V} \in \mathcal{F}_c(\mathbb{R})$, se satisface que $\left\{ \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 \right\}_n$ es una sucesión de estimadores de $\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2$ que es fuertemente consistente, es decir, cuando $n \rightarrow \infty$ se tiene que:

$$\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 \xrightarrow{c.s.} \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2,$$

cualquiera que sea $\mathbf{p} = (p_1, \dots, p_{L-1}) \in (0, 1)^{L-1}$.

ii) La sucesión de variables aleatorias reales:

$$\left\{ \sqrt{n} \left[\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 - \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right] \right\}_n$$

converge en ley hacia una normal unidimensional $\mathcal{N}(0, \sigma^2(\mathbf{p}))$, con:

$$\sigma^2(\mathbf{p}) = \text{Var} \left(E \left(f_{\mathcal{X}} \left[f_{\tilde{E}(\mathcal{X}|\mathbf{p})} - f_{\tilde{V}} \right] \middle| W \otimes \varphi \right) \middle| \mathbf{p} \right),$$

siempre que $\sigma^2(\mathbf{p}) > 0$.

iii) Si $\sigma^2(\mathbf{p}) = 0$ y para algún par $(i, j) \in \{1, \dots, L-1\} \times \{1, \dots, L-1\}$ se cumple que:

$$\begin{aligned} h_{ij} &= \frac{\partial^2 \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2}{\partial p_i \partial p_j} \\ &= 2 E \left([f_{\tilde{x}_i} - f_{\tilde{x}_L}] [f_{\tilde{x}_j} - f_{\tilde{x}_L}] \middle| W \otimes \varphi \right) > 0, \end{aligned}$$

entonces:

$$\left\{ 2n \left[\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 - \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right] \right\}_n$$

es una sucesión de variables aleatorias reales que converge en ley hacia una combinación lineal de, a lo sumo, $L-1$ variables chi-cuadrado con 1 grado de libertad e independientes.

Demostración.

- i) En las condiciones de la proposición puede asegurarse que $\mathbf{f}_n \xrightarrow{c.s.} \mathbf{p}$, por lo que de acuerdo con los resultados de la Teoría Asintótica en Inferencia Paramétrica, al ser $\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2$ continua en un entorno de \mathbf{p} , se concluye que:

$$\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 \xrightarrow{c.s.} \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2$$

cuando $n \rightarrow \infty$, es decir, $\left\{ \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 \right\}_n$ es una sucesión estimadora de $\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2$ fuertemente consistente.

- ii) Por las condiciones supuestas, para n suficientemente grande es posible desarrollar $\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2$ en un entorno de \mathbf{p} . De este modo, el desarrollo de Taylor de primer orden de $\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2$ será:

$$\begin{aligned} \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 &= \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \\ &+ \nabla \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 (\mathbf{f}_n - \mathbf{p})^t \\ &+ \frac{1}{2} (\mathbf{f}_n - \mathbf{p}) \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}_n^*), \tilde{V}) \right]^2 \right) (\mathbf{f}_n - \mathbf{p})^t, \end{aligned}$$

con ∇ el correspondiente vector gradiente, $\mathbf{H} = [h_{ij}]$ la matriz hessiana $(L - 1) \times (L - 1)$, y $\mathbf{p}_n^* \in \mathbb{P}$ tal que $\|\mathbf{p}_n^* - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$.

Por lo tanto:

$$\sqrt{n} \left[\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 - \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right]$$

$$\begin{aligned}
&= \nabla \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t \\
&+ \frac{1}{2}(\mathbf{f}_n - \mathbf{p}) \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}_n^*), \tilde{V}) \right]^2 \right) (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t.
\end{aligned}$$

El vector gradiente vendrá dado por la matriz fila:

$$\begin{aligned}
&\nabla \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \\
&= \left(\frac{\partial}{\partial p_1} \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2; \dots; \frac{\partial}{\partial p_{L-1}} \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right)
\end{aligned}$$

donde, para $i \in \{1, \dots, L-1\}$, se cumple que:

$$\begin{aligned}
&\frac{\partial}{\partial p_i} \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \\
&= E \left(2 \left[f_{\tilde{x}_i} - f_{\tilde{x}_L} \right] \left[f_{\tilde{E}(\mathcal{X}|\mathbf{p})} - f_{\tilde{V}} \right] \middle| W \otimes \varphi \right).
\end{aligned}$$

Como $\sqrt{n}(\mathbf{f}_n - \mathbf{p})$ se distribuye asintóticamente como una normal $(L-1)$ -dimensional $\mathcal{N}(\mathbf{0}, [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1})$, según las propiedades de la convergencia en ley se tiene que cuando $n \rightarrow \infty$ la sucesión

$$\left\{ \nabla \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t \right\}_n$$

converge en ley a una variable con distribución:

$$\mathcal{N} \left(0, \nabla \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1} \left(\nabla \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right)^t \right),$$

siempre que sea positiva la varianza:

$$\begin{aligned}\sigma^2(\mathbf{p}) &= \nabla \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1} \left(\nabla \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right)^t \\ &= \sum_{i=1}^{L-1} \sum_{j=1}^{L-1} p_i(\delta_{ij} - p_j) \left(\frac{\partial}{\partial p_i} \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right) \\ &\quad \cdot \left(\frac{\partial}{\partial p_j} \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right),\end{aligned}$$

es decir:

$$\sigma^2(\mathbf{p}) = \text{Var} \left(E \left(f_{\mathcal{X}} \left[f_{\tilde{E}(\mathcal{X}|\mathbf{p})} - f_{\tilde{V}} \right] \middle| W \otimes \varphi \right) \middle| \mathbf{p} \right).$$

En virtud de las propiedades relativas a las convergencias en ley y en probabilidad, al ser $\|\mathbf{p}_n^* - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$ se tiene que cuando $n \rightarrow \infty$:

$$\left(\frac{\partial^2 \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2}{\partial p_i \partial p_j} \right)_{\mathbf{p}=\mathbf{p}_n^*} \xrightarrow{p} \frac{\partial^2 \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2}{\partial p_i \partial p_j},$$

para cualesquiera $i, j \in \{1, \dots, L-1\}$. Por otro lado, $\mathbf{f}_n \xrightarrow{c.s.} \mathbf{p}$, y por lo tanto $\mathbf{f}_n \xrightarrow{p} \mathbf{p}$, cuando $n \rightarrow \infty$, de donde:

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p}) \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}_n^*), \tilde{V}) \right]^2 \right) \xrightarrow{p} \mathbf{0}$$

y, como $\sqrt{n}(\mathbf{f}_n - \mathbf{p}) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1})$, se tiene que:

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p}) \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}_n^*), \tilde{V}) \right]^2 \right) (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t \xrightarrow{\mathcal{L}} \mathbf{0}$$

y, en consecuencia:

$$\frac{1}{2}(\mathbf{f}_n - \mathbf{p}) \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}_n^*), \tilde{V}) \right]^2 \right) (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t \xrightarrow{p} \mathbf{0}.$$

Los resultados anteriores garantizan que la sucesión:

$$\left\{ \sqrt{n} \left[\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 - \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right] \right\}_n$$

converge en ley hacia una variable con distribución

$$\mathcal{N} \left(0, \nabla \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \left[\mathbf{I}_{\mathcal{X}}^F(\mathbf{p}) \right]^{-1} \left(\nabla \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right)^t \right)$$

cuando $n \rightarrow \infty$, por lo tanto se verifica *ii*) siempre que $\sigma^2(\mathbf{p}) > 0$.

iii) La matriz $\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})$ es definida positiva, por lo que la forma cuadrática asociada a $\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})$ y a $[\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1}$ es definida positiva. En consecuencia, si $\sigma^2(\mathbf{p}) = 0$ se debe cumplir que $\nabla \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 = 0$.

Si se considera ahora el desarrollo de Taylor de segundo orden de $\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2$, puede asegurarse que para n suficientemente grande:

$$\begin{aligned} & \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 = \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \\ & + \frac{1}{2} (\mathbf{f}_n - \mathbf{p}) \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right) (\mathbf{f}_n - \mathbf{p})^t \\ & + \frac{1}{6} \sum_{i=1}^{L-1} \sum_{j=1}^{L-1} \sum_{k=1}^{L-1} \left(\frac{\partial^3 \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2}{\partial p_i \partial p_j \partial p_k} \right)_{\mathbf{p}=\mathbf{p}_n^{**}} \\ & \quad \cdot (f_{ni} - p_i)(f_{nj} - p_j)(f_{nk} - p_k), \end{aligned}$$

con $\mathbf{p}_n^{**} \in \mathbb{P}$ tal que $\|\mathbf{p}_n^{**} - \mathbf{p}\| \leq \|\mathbf{f}_n - \mathbf{p}\|$.

Por lo tanto:

$$\begin{aligned}
 & 2n \left[\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 - \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right] \\
 &= (\sqrt{n}(\mathbf{f}_n - \mathbf{p})) \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right) (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t \\
 &+ \frac{1}{3} \sum_{i=1}^{L-1} \sum_{j=1}^{L-1} \sum_{k=1}^{L-1} \left(\frac{\partial^3 \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2}{\partial p_i \partial p_j \partial p_k} \right)_{\mathbf{p}=\mathbf{p}_n^{**}} \\
 &\quad \cdot (f_{ni} - p_i) (\sqrt{n}(f_{nj} - p_j)) (\sqrt{n}(f_{nk} - p_k)).
 \end{aligned}$$

Razonando como en el apartado *ii*), puede concluirse que:

$$(f_{ni} - p_i) (\sqrt{n}(f_{nj} - p_j)) (\sqrt{n}(f_{nk} - p_k)) \xrightarrow{p} 0.$$

Además, se tiene que la sucesión:

$$\left\{ (\sqrt{n}(\mathbf{f}_n - \mathbf{p})) \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right) (\sqrt{n}(\mathbf{f}_n - \mathbf{p}))^t \right\}_n$$

converge en ley hacia:

$$Y \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right) Y^t,$$

donde Y es un vector aleatorio con distribución normal $(L-1)$ -dimensional $\mathcal{N}(\mathbf{0}, [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1})$.

Como se supone que alguna de las derivadas segundas

$$h_{ij} = \partial^2 \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 / \partial p_i \partial p_j$$

es positiva y, por lo tanto, que $\mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right)$ no se reduce a la matriz nula de orden $(L-1) \times (L-1)$, y ya que el rango de $[\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1}$

es $L - 1$, debe cumplirse que el vector Y puede expresarse como $Y = Z \mathbf{B}$ con Z vector aleatorio $(L - 1)$ -dimensional cuyas componentes son $L - 1$ variables aleatorias independientes e idénticamente distribuidas, con distribución $\mathcal{N}(0, 1)$, y \mathbf{B} es una matriz $(L - 1) \times (L - 1)$ tal que $\mathbf{B} \mathbf{B}^t = [\mathbf{I}_{\mathcal{X}}^F(\mathbf{p})]^{-1}$. Además, existe una transformación $Z = U \mathbf{C}$ en la que \mathbf{C} es una matriz ortogonal, de forma que:

$$\begin{aligned} & Y \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right) Y^t \\ &= Z \mathbf{B} \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right) \mathbf{B}^t Z^t \\ &= U \mathbf{C} \mathbf{B} \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right) \mathbf{B}^t \mathbf{C}^t U^t \\ &= \lambda_1 U_1^2 + \cdots + \lambda_q U_q^2, \end{aligned}$$

con $\lambda_1, \dots, \lambda_q$ ($q \leq L - 1$) los autovalores no nulos de la matriz $(L - 1) \times (L - 1)$ dada por $\mathbf{B} \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right) \mathbf{B}^t$, y donde U_1, \dots, U_q son variables aleatorias independientes e idénticamente distribuidas con distribución $\mathcal{N}(0, 1)$.

En consecuencia, el estudio de la forma cuadrática

$$Y \mathbf{H} \left(\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right) Y^t$$

se reduce en primer lugar al de una forma cuadrática de un vector normal $(L - 1)$ -dimensional con componentes independientes y $\mathcal{N}(0, 1)$, y este último se reduce a su vez al de una combinación lineal de variables aleatorias independientes e idénticamente distribuidas, con distribución chi-cuadrado χ_1^2 , es decir, una combinación lineal de cuadrados de variables aleatorias independientes e idénticamente distribuidas, con distribución $\mathcal{N}(0, 1)$. \square

Cuando se cumple que $\tilde{E}(\mathcal{X}|\mathbf{p}) = \tilde{V}$, que equivale a $\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 = 0$, entonces la varianza $\sigma^2(\mathbf{p})$ se anula. Como implicación, sobre la base de la proposición precedente y de las propiedades de las convergencias en ley y en probabilidad (especialmente, del Teorema de Slutsky), se tiene que:

Teorema 2.1.4 *Para cada $n \in \mathbb{N}$, se consideran n variables aleatorias difusas, $\mathcal{X}_1 \dots, \mathcal{X}_n$, independientes e idénticamente distribuidas que \mathcal{X} (es decir, una muestra aleatoria simple de tamaño n a partir de \mathcal{X}), definidas sobre el espacio de probabilidad (Ω, \mathcal{A}, P) de forma que $P(\{\omega \in \Omega \mid \mathcal{X} = \tilde{x}_l\}) = p_l$ con $p_l \in (0, 1)$, $l = 1, \dots, L$, $\sum_{l=1}^L p_l = 1$.*

Si $\sigma^2(\mathbf{p})$ se anula, $\mathbf{f}_n = (f_{n1}, \dots, f_{n(L-1)}) \in [0, 1]^{L-1}$, con f_{nl} = frecuencia relativa de \tilde{x}_l en la correspondiente realización de la muestra aleatoria simple de tamaño n ($l = 1, \dots, L-1$, $f_{nL} = 1 - \sum_{l=1}^{L-1} f_{nl}$), y $\bar{\mathcal{X}}$ representa el valor esperado muestral de \mathcal{X} en el espacio de probabilidad, entonces se cumple que la sucesión de variables aleatorias reales:

$$\left\{ 2n \left[\left[D_W^\varphi(\bar{\mathcal{X}}, \tilde{V}) \right]^2 - \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V}) \right]^2 \right] \right\}_n$$

converge en ley cuando $n \rightarrow \infty$ hacia una combinación lineal de variables chi-cuadrado con 1 grado de libertad e independientes:

$$\hat{\lambda}_1 \chi_{1,1}^2 + \dots + \hat{\lambda}_q \chi_{1,q}^2$$

(con $q \leq L-1$), en la que $\hat{\lambda}_1, \dots, \hat{\lambda}_q$ son los autovalores muestrales no nulos de la matriz $(L-1) \times (L-1)$ dada por:

$$\mathbf{B}(\mathbf{f}_n) \mathbf{H} \left(\left[D_W^\varphi(\bar{\mathcal{X}}, \tilde{V}) \right]^2 \right) \mathbf{B}(\mathbf{f}_n)^t,$$

con:

$$\mathbf{B}(\mathbf{f}_n) \mathbf{B}(\mathbf{f}_n)^t = [I_{\mathcal{X}}^F(\mathbf{f}_n)]^{-1} = [f_{ni}(\delta_{ij} - f_{nj})],$$

y la matriz hessiana muestral dada por:

$$\begin{aligned} \mathbf{H} \left(\left[D_W^\varphi(\bar{\mathcal{X}}, \tilde{V}) \right]^2 \right) &= \left[\partial^2 \left[D_W^\varphi(\bar{\mathcal{X}}, \tilde{V}) \right]^2 / \partial f_{ni} \partial f_{nj} \right] \\ &= 2 E \left([f_{\tilde{x}_i} - f_{\tilde{x}_L}] [f_{\tilde{x}_j} - f_{\tilde{x}_L}] \mid W \otimes \varphi \right) \end{aligned}$$

(que es independiente de la muestra).

Como consecuencia del teorema anterior, se obtiene el siguiente procedimiento de contraste:

Teorema 2.1.5 Para contrastar al nivel de significación nominal α la hipótesis nula:

$$H_0 : \tilde{E}(\mathcal{X}|P) = \tilde{V}$$

frente a la alternativa:

$$H_a : \tilde{E}(\mathcal{X}|P) \neq \tilde{V},$$

H_0 debe rechazarse siempre que:

$$2n \left[D_W^\varphi(\bar{\mathcal{X}}, \tilde{V}) \right]^2 > \tau_\alpha,$$

donde τ_α es el percentil de orden $100(1 - \alpha)$ de la distribución de $\hat{\lambda}_1 \chi_{1,1}^2 + \dots + \hat{\lambda}_q \chi_{1,q}^2$.

Además, la probabilidad de rechazar H_0 bajo la hipótesis alternativa converge a 1 cuando $n \rightarrow \infty$.

Observación 2.1.1 La combinación lineal de variables chi-cuadrado en el Teorema 2.1.5 puede aproximarse siguiendo criterios que proporcionen aproximaciones apropiadas (como el sugerido por Pardo and Zografos, 2000). De hecho, en las simulaciones posteriores que involucran dicha distribución asintótica se ha seguido ese criterio.

La conclusión del Teorema 2.1.5 se aplica a continuación sobre un ejemplo real.

Ejemplo 2.1.2 Para analizar el porcentaje de conocimiento de informática a nivel de usuario adquirido por los estudiantes del primer curso de Gestión y Administración Pública de la Universidad de Oviedo (población Ω), se seleccionó una muestra aleatoria de 80 estudiantes y se les preguntó por su percepción personal sobre este tema. Las respuestas están recopiladas en la Tabla 2.1.

Tabla 2.1: Nivel de conocimientos adquirido

Respuestas	Nulo	Bajo	Medio	Medio-alto	Alto
Frecuencias	1	22	27	23	7

$$\begin{aligned}
 Nulo &= \text{Tra}(0, 0, 0, 10), \\
 Bajo &= \text{Tra}(0, 10, 30, 40), \\
 Medio &= \text{Tra}(30, 40, 60, 70), \\
 Medio-alto &= \text{Tra}(50, 60, 70, 80), \\
 Alto &= \text{Tra}(70, 80, 100, 100).
 \end{aligned}$$

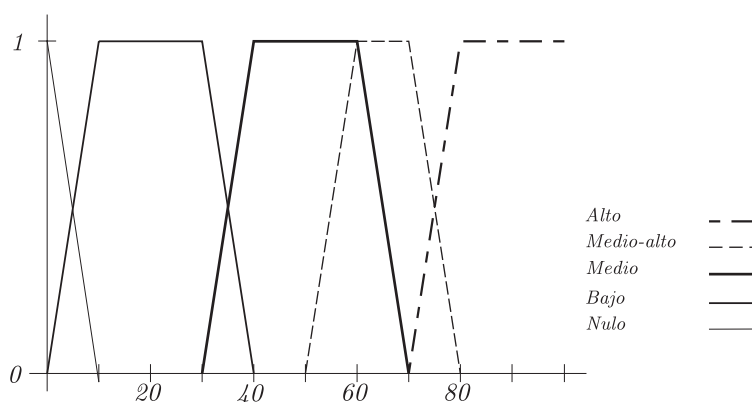


Fig. 2.2: Conjuntos difusos correspondientes al nivel de conocimiento en informática

Los profesores de esos alumnos han identificado las respuestas con los números difusos trapezoidales de la Figura 2.2, denotando por \tilde{V} el conjunto difuso trapezoidal $\text{Tra}(\inf \tilde{V}_0, \inf \tilde{V}_1, \sup \tilde{V}_1, \sup \tilde{V}_0)$.

Los profesores creen que el porcentaje de conocimiento adquirido por sus estudiantes es, en media, “en torno a un 50% o quizás un poco más” y expresan esa percepción mediante el número difuso $\text{Tra}(45, 50, 60, 65)$.

Si para contrastar

$$H_0 : \tilde{E}(\mathcal{X}|P) = \text{Tra}(45, 50, 60, 65)$$

se utiliza el test asintótico empírico del Teorema 2.1.5, se obtiene un p -valor igual a 0.003, por lo que H_0 debería ser rechazada a los niveles de significación nominales usuales ($\alpha = 0.05, 0.01, 0.005$). Sin embargo, al contrastar $H_0 : \tilde{E}(\mathcal{X}|P) = \text{Medio}$ se obtendría un p -valor igual a 0.141 y, en este caso, no se rechazaría la hipótesis a esos niveles.

2.1.3 Test bilateral bootstrap para el valor esperado de una variable aleatoria difusa simple

Con el fin de justificar el uso práctico, en referencia a la significación, del test asintótico empírico propuesto en la subsección precedente, e investigar a partir de qué tamaños muestrales pueden considerarse aceptables las aproximaciones obtenidas, se han realizado estudios de simulación siguiendo las pautas sugeridas en Colubi *et al.* (2002).

De la misma manera, para evaluar la precisión perdida por la estimación de los autovalores necesaria en la aplicación del test del Teorema 2.1.5 y con el objetivo añadido de disponer de un punto de referencia para posteriores comparaciones, se ha simulado el comportamiento del test asintótico teórico (en el que se suponen conocidos los autovalores no nulos poblacionales).

En la Tabla 2.2 se recoge un resumen de los resultados obtenidos en 10000 realizaciones de los tests para ciertas poblaciones finitas simuladas y distintos tamaños muestrales. Para realizar tales simulaciones, se ha elegido la medida de Lebesgue para formalizar tanto W como φ , aunque más adelante se mostrará un estudio más amplio en este sentido.

Tabla 2.2: **Simulación para los tests asintóticos (una muestra)**

	n=5	n=10	n=30	n=100	n=300	n=500
Test asintótico	4.74	4.41	4.38	5.12	5.09	4.95
Test asintótico empírico	16.09	12	6.78	6.01	5.24	5.01

La tabla presenta el porcentaje de rechazos a nivel $\alpha = 0.05$ en esas 10000 realizaciones cuando la hipótesis nula es cierta. Así, el error muestral asintótico para un nivel de confianza del 95% que subyace en relación con el valor nominal 0.05 del test es de 0.004272.

A la vista de los resultados de la Tabla 2.2, puede concluirse que la estimación de los autovalores hace perder mucha precisión respecto al valor nominal de la significación y, de hecho, se comprueba que los tamaños muestrales necesarios para obtener resultados razonables son ciertamente elevados. Por este motivo, y con el fin de establecer analogías con los métodos usuales en el caso real, parece interesante recurrir a técnicas bootstrap en este contexto.

Como punto de partida pueden consultarse los resultados clásicos referentes al bootstrap en el test t de una muestra que aparecen en Efron and Tibshirani (1993). Una de las particularidades del espacio $\mathcal{F}_c(\mathbb{R})$, que habrá que tener en cuenta en cualquier situación en la que se pretendan utilizar estas técnicas, es la falta de linealidad. En el bootstrap clásico, para conseguir que la distribución de la que remuestrear esté bajo la hipó-

tesis nula, se suelen calcular residuos por diferencia respecto a la media muestral. En el presente caso, la falta de linealidad obliga a conseguir colocar las muestras bajo la hipótesis nula sin utilizar en ningún momento el operador diferencia. Como se señala en Shao and Tu (1995) existen diversos métodos bootstrap válidos para llevar a cabo un mismo test.

Para el problema que se considera en esta sección, el propósito será introducir únicamente técnicas básicas que podrán ser mejoradas en el futuro siguiendo las ideas que llevan a perfeccionar, bajo ciertas condiciones, las técnicas clásicas (ver, por ejemplo, Sutton, 1993, para tests de una muestra, o Hall and Martin, 1988, para tests de dos muestras). La primera técnica que sugiere la forma del estadístico con el que se soluciona el test asintótico se podría plantear como sigue:

Sea (Ω, \mathcal{A}, P) un espacio de probabilidad y \mathcal{X} una variable aleatoria difusa asociada a él que toma L valores diferentes $\tilde{x}_1, \dots, \tilde{x}_L$ con probabilidades respectivas p_1, \dots, p_L . Para cada $n \in \mathbb{N}$ se considera una muestra aleatoria de n observaciones de \mathcal{X} independientes, denotándose por $\mathbf{f}_n = (f_{n1}, \dots, f_{nL})$ el vector aleatorio de las frecuencias muestrales.

Se considera ahora una nueva población dada por una realización de la muestra aleatoria anterior. Se remuestrea a partir de esta última población, obteniendo de nuevo una muestra de tamaño n . Si se denota el vector aleatorio asociado a esa nueva muestra por $\mathbf{f}_n^* = (f_{n1}^*, \dots, f_{nL}^*)$, el estadístico del contraste será:

$$T_n^1 = \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{E}(\mathcal{X}|\mathbf{f}_n^*)) \right]^2$$

y su distribución puede aproximarse por el método de Monte Carlo. Se concluye que:

Teorema 2.1.6 *Para contrastar al nivel de significación nominal α la hipótesis nula:*

$$H_0 : \tilde{E}(\mathcal{X}|P) = \tilde{V}$$

frente a la alternativa:

$$H_a : \tilde{E}(\mathcal{X}|P) \neq \tilde{V},$$

H_0 debe rechazarse (de acuerdo con una primera aproximación bootstrap) siempre que:

$$\left[D_W^\varphi(\bar{\mathcal{X}}, \tilde{V}) \right]^2 > z_\alpha,$$

donde $\bar{\mathcal{X}} = \tilde{E}(\mathcal{X}|\mathbf{f}_n^*)$ y z_α es el $100(1 - \alpha)$ percentil de la distribución de T_n^1 .

Demostración. Es equivalente considerar el estadístico y el valor crítico multiplicados por $2n$.

Siguiendo las indicaciones en Shao and Tu (1995), se probará que la sucesión:

$$\left\{ 2n \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n^*), \tilde{E}(\mathcal{X}|\mathbf{f}_n)) \right]^2 \right\}_n$$

es fuertemente consistente con la sucesión:

$$\left\{ 2n \left[D_W^\varphi(\bar{\mathcal{X}}, \tilde{V}) \right]^2 \right\}_n$$

bajo la hipótesis H_0 .

Sea \mathbf{f}_m^* una distribución de frecuencias obtenida en un remuestreo de tamaño m , y sea $\tilde{V} \in \mathcal{F}_c(\mathbb{R})$ un número difuso cualquiera. Al igual que en la Proposición 2.1.3, se puede obtener el desarrollo de Taylor de segundo orden de $\left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_m^*), \tilde{V}) \right]^2$ en un entorno de \mathbf{f}_n como sigue:

$$\begin{aligned} \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_m^*), \tilde{V}) \right]^2 &= \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 \\ &+ \nabla \left[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V}) \right]^2 (\mathbf{f}_m^* - \mathbf{f}_n)^t \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2} (\mathbf{f}_m^* - \mathbf{f}_n) \mathbf{H} \left(\left[D_W^\varphi \left(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{V} \right) \right]^2 \right) (\mathbf{f}_m^* - \mathbf{f}_n)^t \\
& + \frac{1}{6} \sum_i \sum_j \sum_k \left(\frac{\partial^3 \left[D_W^\varphi \left(\tilde{E}(\mathcal{X}|\mathbf{p}), \tilde{V} \right) \right]^2}{\partial p_i \partial p_j \partial p_k} \right)_{\mathbf{p}=\mathbf{p}_{n,m}^{**}} \\
& \quad \cdot (f_{mi}^* - f_{ni}) (f_{mj}^* - f_{nj}) (f_{mk}^* - f_{nk}),
\end{aligned}$$

con $\|\mathbf{p}_{n,m}^{**} - \mathbf{f}_n\| \leq \|\mathbf{f}_m^* - \mathbf{f}_n\|$.

Siguiendo un razonamiento análogo al aplicado en la Proposición 2.1.3, y teniendo en cuenta que la matriz hessiana no depende de \tilde{V} ni del valor del parámetro vectorial \mathbf{f}_n , se puede evaluar la expresión en $\tilde{V} = \tilde{E}(\mathcal{X}|\mathbf{f}_n)$ y probar entonces que los límites reiterados

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} 2m \left[D_W^\varphi \left(\tilde{E}(\mathcal{X}|\mathbf{f}_m^*), \tilde{E}(\mathcal{X}|\mathbf{f}_n) \right) \right]^2, \\
& \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} 2m \left[D_W^\varphi \left(\tilde{E}(\mathcal{X}|\mathbf{f}_m^*), \tilde{E}(\mathcal{X}|\mathbf{f}_n) \right) \right]^2
\end{aligned}$$

coinciden en ley.

De esta manera, eligiendo $m = n$, se obtiene que:

$$\left\{ 2n \left[D_W^\varphi \left(\tilde{E}(\mathcal{X}|\mathbf{f}_n^*), \tilde{E}(\mathcal{X}|\mathbf{f}_n) \right) \right]^2 \right\}_n$$

converge en ley cuando $n \rightarrow \infty$ hacia la distribución de una combinación lineal

$$\lambda_1 Z_1^2 + \dots + \lambda_q Z_q^2,$$

donde $\lambda_1, \dots, \lambda_q$ son los autovalores no nulos en *iii*) de la Proposición 2.1.3, de lo que se deduce la consistencia fuerte buscada.

Así se asegura que el test es asintóticamente correcto en el sentido de alcanzar, en el límite, la significación nominal y, de la misma manera, bajo la hipótesis alternativa la probabilidad de rechazo es 1 en el límite. \square

Como suele ocurrir en el caso real (ver, por ejemplo, Efron and Tibshirani, 1993) se puede mejorar notablemente la precisión del test bootstrap incorporando la variabilidad. En el siguiente teorema se prueba que incluyendo esa variabilidad el bootstrap es consistente y, más adelante, se ilustrará la mejora que conlleva mediante simulaciones.

Teorema 2.1.7 *Para contrastar al nivel de significación nominal α la hipótesis nula*

$$H_0 : \tilde{E}(\mathcal{X}|\mathbf{p}) = \tilde{V}$$

frente a la alternativa

$$H_a : \tilde{E}(\mathcal{X}|\mathbf{p}) \neq \tilde{V},$$

H_0 debe rechazarse (según una segunda aproximación bootstrap) siempre que:

$$\frac{[D_W^\varphi(\bar{\mathcal{X}}, \tilde{V})]^2}{\hat{S}(\mathcal{X}|\mathbf{f}_n)} > z_\alpha,$$

donde z_α es el $100(1 - \alpha)$ percentil de la distribución de

$$T_n^2 = \frac{[D_W^\varphi(\tilde{E}(\mathcal{X}|\mathbf{f}_n), \tilde{E}(\mathcal{X}|\mathbf{f}_n^*))]^2}{\hat{S}^2(\mathcal{X}|\mathbf{f}_n^*)},$$

y:

$$\hat{S}^2(\mathcal{X}|\mathbf{f}_n) = \frac{n}{n-1} S^2(\mathcal{X}|\mathbf{f}_n) = \frac{1}{n-1} \sum_{l=1}^L [D_W^\varphi(\tilde{x}_l, \tilde{E}(\mathcal{X}|\mathbf{f}_n))]^2 f_{nl}.$$

Demostración. En el Teorema 2.1.6 se probó la convergencia en ley del numerador de T_n^2 hacia una variable aleatoria real que se denotará por T . Por otro lado, los resultados de Lubiano *et al.* (1999) permiten asegurar que $\{\hat{S}^2(\mathcal{X}|\mathbf{f}_n^*)\}_n$ converge casi seguro a

$$S^2(\mathcal{X}|\mathbf{p}) = \frac{1}{n} \sum_{l=1}^L [D_W^\varphi(\tilde{x}_l, \tilde{E}(\mathcal{X}|\mathbf{p}))]^2 p_l$$

cuando $n \rightarrow \infty$. En consecuencia, aplicando el Teorema de Slutsky se concluye que la sucesión $\{T_n^2\}_n$ converge en ley hacia la distribución de

$$\frac{(n-1)T}{nS^2(\mathcal{X}|\mathbf{p})} = (n-1)T / \sum_{l=1}^L \left[D_W^\varphi(\tilde{x}_l, \tilde{E}(\mathcal{X}|\mathbf{p})) \right]^2 p_l$$

en ley cuando $n \rightarrow \infty$. Esta distribución límite se puede aproximar por Monte Carlo, y resolver así el test mediante bootstrap. \square

Observación 2.1.2 Siguiendo las ideas del teorema precedente, puede desarrollarse una versión bootstrap de la Proposición 2.1.3-ii) que sería útil para el estudio de la potencia del test del Teorema 2.1.7. Este estudio podría llevarse a cabo en términos del parámetro real $\theta_W^\varphi = D_W^\varphi(\tilde{E}(\mathcal{X}|P), \tilde{V})$.

Los métodos que acaban de establecerse se aplican a continuación para los datos del Ejemplo 2.1.2.

Ejemplo 2.1.3 Se considera la situación en el Ejemplo 2.1.2.

Si para contrastar $H_0 : \tilde{E}(\mathcal{X}|P) = \text{Tra}(45, 50, 60, 65)$ se utiliza la segunda técnica bootstrap (Teorema 2.1.7), se obtiene un p -valor igual a 0.001, de forma que H_0 debería ser rechazada a los niveles de significación usuales ($\alpha = 0.05, 0.01, 0.005$). Sin embargo, al contrastar $H_0 : \tilde{E}(\mathcal{X}|P) = \text{Medio}$ se obtendría un p -valor igual a 0.135 y, en este caso, no se rechazaría la hipótesis a esos niveles.

2.1.4 Estudios de simulación: comparación de técnicas de contraste sobre el valor esperado de una variable aleatoria difusa

Ya se han mostrado algunas conclusiones parciales obtenidas en estudios comparativos de simulación.

Los estudios que se presentan en esta subsección se han llevado a cabo en un marco más amplio con el fin de analizar el comportamiento de los tests en distintas situaciones. En concreto, se han simulado poblaciones con diferente número de modalidades $L = 5$, $L = 10$ y $L = 20$ (en realidad se han simulado también poblaciones con 50 y 100 modalidades pero las conclusiones que se obtenían no difieren sustancialmente de las que se expondrán en la memoria). Asimismo, se han elegido distintas distribuciones sobre esos valores. En primer lugar, se han hecho simulaciones considerando tamaños muestrales pequeños y medios. Las conclusiones que se extraen de estos estudios nos permitirán simplificar los realizados para grandes muestras, ya que éstos implican un alto coste computacional.

Para muestras pequeñas y de tamaño medio se han elegido diferentes medidas φ para ponderar la información de los α -niveles del número difuso y la medida W elegida ha sido la de Lebesgue. A continuación φ_1 denotará la medida de Lebesgue, φ_2 representará una distribución $\beta(2, 10)$ (i.e., a menor nivel mayor ponderación) y φ_3 será una $\beta(10, 2)$ (i.e., a mayor nivel mayor ponderación). Hemos observado que, como cabía esperar, no hay una elección uniformemente buena de φ , ya que depende de la forma de los conjuntos difusos. Por esta razón, para grandes muestras hemos optado únicamente por φ_1 .

En estos estudios se han considerado cuatro tests. T_1 representará el test asintótico teórico (para el que se suponen conocidos los autovalores poblacionales), T_2 denotará el test asintótico empírico (en el que se estiman los autovalores poblacionales por sus análogos muestrales), T_3 representará el primer test bootstrap y T_4 representará el segundo.

En la Tabla 2.3 se muestran las conclusiones obtenidas, en media, para muestras pequeñas y de tamaño medio en la simulación de tres tipos de variables aleatorias difusas considerando para cada una dos distribuciones

sobre sus valores poblacionales. Cada simulación corresponde a 10000 iteraciones y el número de réplicas bootstrap ha sido 1000. Se muestran los porcentajes de rechazo al nivel de significación 0.05 cuando la hipótesis nula es cierta.

Tabla 2.3: Simulación para muestras de tamaño pequeño/medio

		T_1			T_2			T_3			T_4		
		φ_1	φ_2	φ_3	φ_1	φ_2	φ_3	φ_1	φ_2	φ_3	φ_1	φ_2	φ_3
$L = 5$	$n = 5$	4.74	4.14	3.77	16.09	16.99	23.72	15.78	16.93	26.6	2.29	2.44	2.34
	$n = 10$	4.41	4.68	4.38	12	12.29	12.42	12.43	13.07	12.82	5.77	5.31	5.32
	$n = 30$	4.38	4.98	4.88	6.78	6.85	7.12	7.14	6.94	7.83	4.96	4.66	4.65
	$n = 100$	5.12	5.44	5.32	6.01	5.9	6.04	6.3	5.97	5.99	5.02	5.11	4.95
$L = 10$	$n = 5$	5.09	4.49	4.5	15.6	18.45	18.99	16.85	17.63	19.28	5.13	4.44	8.47
	$n = 10$	4.53	4.52	4.4	10.42	11.34	10.79	11.52	11.58	21.01	3.85	4.73	16.21
	$n = 30$	5.15	4.91	5	6.63	6.24	6.78	6.71	6.72	6.85	4.65	4.6	4.69
	$n = 100$	4.89	4.64	5.01	5.59	5.61	5.74	5.69	5.61	5.71	4.94	4.92	4.97
$L = 20$	$n = 5$	4.56	4.6	4.61	23.21	20.96	20.29	21.13	19.2	19.59	6.29	5.56	6.62
	$n = 10$	4.31	4.19	3.97	12.78	11.78	14.7	12.95	12.92	14.94	7.71	6.69	9.34
	$n = 30$	5.18	5.05	5.12	6.59	6.61	7.7	7.17	6.72	7.92	4.61	4.6	4.94
	$n = 100$	4.89	4.85	4.91	5.51	5.73	5.88	5.7	5.78	5.94	4.81	5	4.9

El test T_1 se ha simulado únicamente como referencia, puesto que no es aplicable en la práctica (pues no se conocen los autovalores poblacionales). En la Tabla 2.3 se aprecia que, para estos tamaños muestrales tanto el test T_2 como el T_3 proporcionan resultados inaceptables. Parece que el test T_2 funciona ligeramente mejor que el T_3 pero peor que el T_4 . Este último test proporciona mejores resultados a partir de $n = 30$.

Convendría remarcar que, debido al coste computacional de estos estudios, se han realizado únicamente 10000 repeticiones de cada test, lo que implica un error muestral asintótico cercano a 5 centésimas, que explicaría en parte las variaciones que se producen en torno al valor nominal.

En las simulaciones de grandes muestras hemos tenido en cuenta únicamente la medida φ_1 y los tests T_2 y T_4 . Esta simplificación ha permitido

elevar el número de repeticiones a 40000, con lo que se consigue reducir el error muestral a 0.002136. Los resultados se presentan en la Tabla 2.4

Tabla 2.4: **Simulación para grandes muestras**

	$L = 5$		$L = 10$		$L = 20$	
	T_2	T_4	T_2	T_4	T_2	T_4
$n = 300$	4.88	4.83	5.3	5.06	5.16	5.09
$n = 500$	5.22	5.08	4.91	4.87	5.19	5.07

2.2 Contrastes sobre la igualdad de medias poblacionales de dos variables aleatorias difusas con valores en $\mathcal{F}_c(\mathbb{R})$

Para contrastar la igualdad de medias poblacionales de dos variables aleatorias reales se recurre usualmente a un estadístico que involucra la diferencia o distancia euclídea entre las correspondientes medias muestrales. La distribución exacta de ese estadístico suele determinarse bajo condiciones de normalidad de las variables.

En esta sección se examina en primer lugar el contraste bilateral para la igualdad de valores esperados poblacionales de dos variables aleatorias difusas, cuando estas variables tengan distribución normal y se disponga de dos muestras independientes. Se comprobará que, al igual que en el caso de una muestra el procedimiento en el que la distancia euclídea en \mathbb{R} se reemplaza por la métrica D_W^φ en $\mathcal{F}_c(\mathbb{R})$, se reduce a un contraste de hipótesis con variables aleatorias reales normales.

Por otro lado, se estudiarán otras dos aproximaciones del problema de contraste: la basada en la Teoría de Grandes Muestras y la basada en las técnicas bootstrap. Finalmente, se compararán los diferentes procedimientos mediante simulaciones y se ilustrarán con algunos ejemplos.

2.2.1 Test bilateral para la igualdad de valores esperados poblacionales de dos variables aleatorias difusas normales

Sean $(\Omega_1, \mathcal{A}_1, P_1)$ y $(\Omega_2, \mathcal{A}_2, P_2)$ dos espacios de probabilidad, $\mathcal{X}_1 : \Omega_1 \rightarrow \mathcal{F}_c(\mathbb{R})$ y $\mathcal{X}_2 : \Omega_2 \rightarrow \mathcal{F}_c(\mathbb{R})$ dos variables aleatorias difusas normales asociadas cada una a su respectivo espacio, y con varianzas σ_1^2, σ_2^2 , respectivamente, y $\tilde{E}(\mathcal{X}_1|P_1), \tilde{E}(\mathcal{X}_2|P_2)$ sus medias poblacionales.

El primer problema es contrastar la hipótesis nula $H_0 : \tilde{E}(\mathcal{X}_1|P_1) = \tilde{E}(\mathcal{X}_2|P_2)$ (es decir, $[D_W^\varphi(\tilde{E}(\mathcal{X}_1|P_1), \tilde{E}(\mathcal{X}_2|P_2))]^2 = 0$) frente a la alternativa $H_a : \tilde{E}(\mathcal{X}_1|P_1) \neq \tilde{E}(\mathcal{X}_2|P_2)$ (es decir, $[D_W^\varphi(\tilde{E}(\mathcal{X}_1|P_1), \tilde{E}(\mathcal{X}_2|P_2))]^2 > 0$), cuando se supone que $\sigma_1^2 = \sigma_2^2 = \sigma^2$ es desconocida. Para ello, se consideran una primera muestra aleatoria simple, $\mathcal{X}_{11}, \dots, \mathcal{X}_{1n_1}$ a partir de \mathcal{X}_1 y otra segunda muestra aleatoria simple independiente de la anterior, $\mathcal{X}_{21}, \dots, \mathcal{X}_{2n_2}$ a partir de \mathcal{X}_2 . Se denota por $\bar{\mathcal{X}}_1$ la media muestral de \mathcal{X}_1 y por $\bar{\mathcal{X}}_2$ la media muestral de \mathcal{X}_2 , es decir:

$$\bar{\mathcal{X}}_1 = \frac{1}{n_1} (\mathcal{X}_{11} \oplus \dots \oplus \mathcal{X}_{1n_1}), \quad \bar{\mathcal{X}}_2 = \frac{1}{n_2} (\mathcal{X}_{21} \oplus \dots \oplus \mathcal{X}_{2n_2}).$$

Para este propósito va a emplearse el estadístico:

$$Z = \frac{[D_W^\varphi(\bar{\mathcal{X}}_1, \bar{\mathcal{X}}_2)]^2 - [D_W^\varphi(\tilde{E}(\mathcal{X}_1|P_1), \tilde{E}(\mathcal{X}_2|P_2))]^2}{S^2 \left[\frac{1}{n_1} + \frac{1}{n_2} \right]},$$

donde

$$S^2 = \frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2 - 2}$$

y

$$S_j^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} [D_W^\varphi(\mathcal{X}_{ji}, \bar{\mathcal{X}}_j)]^2, \quad j = 1, 2.$$

Por las propiedades de la función $f(\cdot, \cdot)$ y de la métrica D_W^φ , si se expresa $\mathcal{X}_{ji} = \xi_{ji} \oplus \tilde{E}(\mathcal{X}_j|P_j)$ y, por tanto, $\bar{\mathcal{X}}_j = \bar{\xi}_j \oplus \tilde{E}(\mathcal{X}_j|P_j)$ ($j = 1, 2$), puede concluirse que cuando H_0 es cierta:

$$[D_W^\varphi(\bar{\mathcal{X}}_1, \bar{\mathcal{X}}_2)]^2 = [\bar{\xi}_1 - \bar{\xi}_2]^2,$$

$$\left[D_W^\varphi(\tilde{E}(\mathcal{X}_1|P_1), \tilde{E}(\mathcal{X}_2|P_2)) \right]^2 = [E(\xi_1|P_1) - E(\xi_2|P_2)]^2,$$

y

$$S_j^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} [\xi_{ji} - \bar{\xi}_j]^2$$

($j = 1, 2$), por lo que Z coincide con el estadístico clásico para la igualdad de medias de dos poblaciones normales con varianzas desconocidas coincidentes, es decir:

$$\frac{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2} [\bar{\xi}_1 - \bar{\xi}_2]^2}{\sum_{i=1}^{n_1} [\xi_{1i} - \bar{\xi}_1]^2 + \sum_{i=1}^{n_2} [\xi_{2i} - \bar{\xi}_2]^2}$$

y, en consecuencia, bajo H_0 (que implica $E(\xi_1|P_1) = E(\xi_2|P_2)$) el estadístico Z tiene distribución F de Snedecor con 1 y $n_1 + n_2 - 2$ grados de libertad. Por todo ello, puede afirmarse que:

Teorema 2.2.1 *Para contrastar al nivel de significación α la hipótesis nula:*

$$H_0 : \tilde{E}(\mathcal{X}_1|P_1) = \tilde{E}(\mathcal{X}_2|P_2)$$

frente a la alternativa:

$$H_a : \tilde{E}(\mathcal{X}_1|P_1) \neq \tilde{E}(\mathcal{X}_2|P_2),$$

H_0 debe rechazarse siempre que:

$$\frac{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2} [D_W^\varphi(\bar{\mathcal{X}}_1, \bar{\mathcal{X}}_2)]^2}{\sum_{i=1}^{n_1} [D_W^\varphi(\mathcal{X}_{1i}, \bar{\mathcal{X}}_1)]^2 + \sum_{i=1}^{n_2} [D_W^\varphi(\mathcal{X}_{2i}, \bar{\mathcal{X}}_2)]^2} > F_{1, n_1 + n_2 - 2, \alpha},$$

donde $F_{1, n_1 + n_2 - 2, \alpha}$ denota el percentil de orden $100(1 - \alpha)$ de la distribución F de Snedecor con 1 y $n_1 + n_2 - 2$ grados de libertad.

El segundo problema que va a tratarse es el de contrastar la hipótesis nula H_0 frente a la alternativa H_a del problema anterior, cuando se supone que las varianzas σ_1^2 y σ_2^2 son conocidas y se dispone de sendas muestras aleatorias simples independientes, $\mathcal{X}_{11}, \dots, \mathcal{X}_{1n_1}$ a partir de \mathcal{X}_1 y $\mathcal{X}_{21}, \dots, \mathcal{X}_{2n_2}$ a partir de \mathcal{X}_2 . Razonando en forma análoga al problema precedente, puede concluirse que:

Teorema 2.2.2 *Para contrastar al nivel de significación α la hipótesis nula:*

$$H_0 : \tilde{E}(\mathcal{X}_1|P_1) = \tilde{E}(\mathcal{X}_2|P_2)$$

frente a la alternativa:

$$H_a : \tilde{E}(\mathcal{X}_1|P_1) \neq \tilde{E}(\mathcal{X}_2|P_2),$$

H_0 debe rechazarse siempre que:

$$\frac{[D_W^\varphi(\bar{\mathcal{X}}_1, \bar{\mathcal{X}}_2)]^2}{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} > \chi_{1,\alpha}^2,$$

donde $\chi_{1,\alpha}^2$ es el percentil de orden $100(1-\alpha)$ de la distribución chi-cuadrado con 1 grado de libertad.

De este modo, con independencia de la elección de W y φ , si H_0 es cierta pueden utilizarse los tests para las variables aleatorias reales normales correspondientes a las *diferencias de Hukuhara* $\xi_{ji} = \mathcal{X}_{ji} -_h \tilde{E}(\mathcal{X}_j|P_j)$ ($j = 1, 2$), que en ambos casos son sencillas de determinar. La potencia de los tests precedentes podría estudiarse en forma análoga a la de los tests tradicionales para variables aleatorias reales normales.

2.2.2 Test bilateral asintótico para la igualdad de valores esperados poblacionales de dos variables aleatorias difusas simples

En el caso de variables aleatorias difusas simples, puede seguirse un proceso análogo al de la Subsección 2.1.2, para lo cual se considera un estadístico en el que figuran los valores esperados muestrales que son función de sendos estimadores consistentes y asintóticamente normales de los estimadores analógicos de dos parámetros vectoriales poblacionales.

Sean $(\Omega_1, \mathcal{A}_1, P_1)$ y $(\Omega_2, \mathcal{A}_2, P_2)$ dos espacios de probabilidad, $\mathcal{X}_1 : \Omega_1 \rightarrow \mathcal{F}_c(\mathbb{R})$ y $\mathcal{X}_2 : \Omega_2 \rightarrow \mathcal{F}_c(\mathbb{R})$ dos variables aleatorias difusas simples asociadas cada una a su respectivo espacio.

Sean $\tilde{x}_{11}, \dots, \tilde{x}_{1L_1} \in \mathcal{F}_c(\mathbb{R})$ los distintos valores de \mathcal{X}_1 en la población, y p_{11}, \dots, p_{1L_1} las probabilidades respectivas ($P_1(\{\omega_1 \in \Omega_1 | \mathcal{X}_1(\omega_1) = \tilde{x}_{1l_1}\}) = p_{1l_1} > 0, l_1 = 1, \dots, L_1, \sum_{l_1=1}^{L_1} p_{1l_1} = 1$).

Sean $\tilde{x}_{21}, \dots, \tilde{x}_{2L_2} \in \mathcal{F}_c(\mathbb{R})$ los distintos valores de \mathcal{X}_2 en la población, y p_{21}, \dots, p_{2L_2} las probabilidades respectivas ($P_2(\{\omega_2 \in \Omega_2 | \mathcal{X}_2(\omega_2) = \tilde{x}_{2l_2}\}) = p_{2l_2} > 0, l_2 = 1, \dots, L_2, \sum_{l_2=1}^{L_2} p_{2l_2} = 1$).

Si se denota la media poblacional de \mathcal{X}_1 por $\tilde{E}(\mathcal{X}_1 | \mathbf{p}_1)$, con

$$\mathbf{p}_1 = (p_{11}, \dots, p_{1(L_1-1)}) \in (0, 1)^{L_1-1},$$

se tiene que:

$$\tilde{E}(\mathcal{X}_1 | \mathbf{p}_1) = p_{11} \tilde{x}_{11} \oplus \dots \oplus p_{1L_1} \tilde{x}_{1L_1}.$$

Análogamente, si se denota la media poblacional de \mathcal{X}_2 por $\tilde{E}(\mathcal{X}_2 | \mathbf{p}_2)$, con

$$\mathbf{p}_2 = (p_{21}, \dots, p_{2(L_2-1)}) \in (0, 1)^{L_2-1},$$

se tiene que:

$$\tilde{E}(\mathcal{X}_2 | \mathbf{p}_2) = p_{21} \tilde{x}_{21} \oplus \dots \oplus p_{2L_2} \tilde{x}_{2L_2}.$$

Razonando de acuerdo con el comienzo de la Subsección 1.3.3, los parámetros vectoriales \mathbf{p}_1 y \mathbf{p}_2 cumplen las condiciones de regularidad allí indicadas.

Sea $\mathcal{X}_{11}, \dots, \mathcal{X}_{1n_1}$ una muestra aleatoria simple a partir de la variable \mathcal{X}_1 , y $\mathcal{X}_{21}, \dots, \mathcal{X}_{2n_2}$ una muestra aleatoria simple a partir de \mathcal{X}_2 independiente de la muestra anterior. Se denotará por $f_{n_1 l_1}$ el estadístico correspondiente a la frecuencia relativa de \tilde{x}_{1l_1} en la primera muestra aleatoria, y por $\mathbf{f}_{n_1} = (f_{n_1 1}, \dots, f_{n_1(L_1-1)})$. De forma similar, se denotará por $f_{n_2 l_2}$ el estadístico correspondiente a la frecuencia relativa de \tilde{x}_{2l_2} en la segunda muestra aleatoria, y por $\mathbf{f}_{n_2} = (f_{n_2 1}, \dots, f_{n_2(L_2-1)})$. Obviamente:

$$\bar{\mathcal{X}}_1 = \frac{1}{n_1} (\mathcal{X}_{11} \oplus \dots \oplus \mathcal{X}_{1n_1}) = \tilde{E}(\mathcal{X}|\mathbf{f}_{n_1}) = f_{n_1 1} \tilde{x}_{11} \oplus \dots \oplus f_{n_1 L_1} \tilde{x}_{1L_1},$$

$$\bar{\mathcal{X}}_2 = \frac{1}{n_2} (\mathcal{X}_{21} \oplus \dots \oplus \mathcal{X}_{2n_2}) = \tilde{E}(\mathcal{X}|\mathbf{f}_{n_2}) = f_{n_2 1} \tilde{x}_{21} \oplus \dots \oplus f_{n_2 L_2} \tilde{x}_{2L_2}.$$

La sucesión de estimadores $\{\mathbf{f}_{n_j}\}_n$ es fuertemente consistente y se distribuye asintóticamente según una distribución normal $(L_j - 1)$ -dimensional $\mathcal{N}\left(\mathbf{p}_j, \left[\mathbf{I}_{\mathcal{X}_j}^F(\mathbf{p}_j)\right]^{-1} / n_j\right)$ cuando $n_j \rightarrow \infty$, y $n_j(\mathbf{f}_{n_j} - \mathbf{p}_j)\mathbf{I}_{\mathcal{X}_j}^F(\mathbf{p}_j)(\mathbf{f}_{n_j} - \mathbf{p}_j)^t$ converge en distribución a una chi-cuadrado $\chi_{L_j-1}^2$ cuando $n_j \rightarrow \infty$, para todo $j \in \{1, 2\}$. En consecuencia, razonando como en la Proposición 2.1.3, se llega al resultado siguiente:

Proposición 2.2.3 *Para cada par $(n_1, n_2) \in \mathbb{N} \times \mathbb{N}$, se consideran n_j variables aleatorias difusas independientes e igualmente distribuidas que \mathcal{X}_j , definidas sobre el espacio de probabilidad $(\Omega_j, \mathcal{A}_j, P_j)$ de forma que $P_j(\{\omega_j \in \Omega_j \mid \mathcal{X}_j(\omega_j) = \tilde{x}_{jl_j}\}) = p_{jl_j} \in (0, 1)$, $l_j = 1, \dots, L_j$, $\sum_{l_j=1}^{L_j} p_{jl_j} = 1$, $j = 1, 2$. Supongamos ambas muestras independientes. Se cumple entonces que:*

- i) Si $\mathbf{f}_{n_j} = (f_{n_j 1}, \dots, f_{n_j(L_j-1)}) \in [0, 1]^{L_j-1}$, con $f_{n_j l_j} =$ frecuencia relativa de \tilde{x}_{jl_j} en la correspondiente realización de la j -ésima muestra aleatoria simple de tamaño n_j ($l_j = 1, \dots, L_j - 1$, $f_{n_j L_j} = 1 - \sum_{l_j=1}^{L_j-1} f_{n_j l_j}$), y $\tilde{E}(\mathcal{X}_j|\mathbf{f}_{n_j})$ representa la media muestral de \mathcal{X}_j , entonces, se satisface que $\left\{ \left[D_W^\varphi(\tilde{E}(\mathcal{X}_1|\mathbf{f}_{n_1}), \tilde{E}(\mathcal{X}_2|\mathbf{f}_{n_2})) \right]_{n_1, n_2}^2 \right\}$ es una sucesión de estimadores de $\left[D_W^\varphi(\tilde{E}(\mathcal{X}_1|\mathbf{p}_1), \tilde{E}(\mathcal{X}_2|\mathbf{p}_2)) \right]^2$ que es fuertemente consistente, es decir, cuando $n_1 \rightarrow \infty$ y $n_2 \rightarrow \infty$ se tiene que:

$$\left[D_W^\varphi(\tilde{E}(\mathcal{X}_1|\mathbf{f}_{n_1}), \tilde{E}(\mathcal{X}_2|\mathbf{f}_{n_2})) \right]^2 \xrightarrow{c.s.} \left[D_W^\varphi(\tilde{E}(\mathcal{X}_1|\mathbf{p}_1), \tilde{E}(\mathcal{X}_2|\mathbf{p}_2)) \right]^2,$$
 cualesquiera que sean $\mathbf{p}_1 = (p_{11}, \dots, p_{1L_1-1}) \in (0, 1)^{L_1-1}$ y $\mathbf{p}_2 = (p_{21}, \dots, p_{2L_2-1}) \in (0, 1)^{L_2-1}$.

ii) La sucesión de variables aleatorias reales:

$$\left\{ \sqrt{n_1 + n_2} \left[\left[D_W^\varphi(\tilde{E}(\mathcal{X}_1|\mathbf{f}_{n_1}), \tilde{E}(\mathcal{X}_2|\mathbf{f}_{n_2})) \right]^2 - \left[D_W^\varphi(\tilde{E}(\mathcal{X}_1|\mathbf{p}_1), \tilde{E}(\mathcal{X}_2|\mathbf{p}_2)) \right]^2 \right] \right\}_{n_1, n_2}$$

converge en ley hacia una normal unidimensional $\mathcal{N}(0, \sigma^2(\mathbf{p}_1, \mathbf{p}_2))$, con:

$$\begin{aligned} \sigma^2(\mathbf{p}_1, \mathbf{p}_2) = & 4 \left[\text{Var} \left(E \left(f_{\mathcal{X}_1} \left[f_{\tilde{E}(\mathcal{X}_1|\mathbf{p}_1)} - f_{\tilde{E}(\mathcal{X}_2|\mathbf{p}_2)} \right] \mid W \otimes \varphi \right) \mid \mathbf{p}_1 \right) \right. \\ & \left. + \text{Var} \left(E \left(f_{\mathcal{X}_2} \left[f_{\tilde{E}(\mathcal{X}_2|\mathbf{p}_2)} - f_{\tilde{E}(\mathcal{X}_1|\mathbf{p}_1)} \right] \mid W \otimes \varphi \right) \mid \mathbf{p}_2 \right) \right], \end{aligned}$$

siempre que $\sigma^2(\mathbf{p}_1, \mathbf{p}_2) > 0$.

iii) Si $\sigma^2(\mathbf{p}_1, \mathbf{p}_2) = 0$ y se cumple que :

$$\begin{aligned} h_{i_1 k_1} &= \frac{\partial^2 \left[D_W^\varphi(\tilde{E}(\mathcal{X}_1|\mathbf{f}_{n_1}), \tilde{E}(\mathcal{X}_2|\mathbf{f}_{n_2})) \right]^2}{\partial p_{1i_1} \partial p_{1k_1}} \\ &= 2 E \left([f_{\tilde{x}_{1i_1}} - f_{\tilde{x}_{1L_1}}][f_{\tilde{x}_{1k_1}} - f_{\tilde{x}_{1L_1}}] \mid W \otimes \varphi \right) > 0, \end{aligned}$$

para algún $(i_1, k_1) \in \{1, \dots, L_1 - 1\} \times \{1, \dots, L_1 - 1\}$, o bien:

$$\begin{aligned} h_{i_1 k_2} &= \frac{\partial^2 \left[D_W^\varphi(\tilde{E}(\mathcal{X}_1|\mathbf{f}_{n_1}), \tilde{E}(\mathcal{X}_2|\mathbf{f}_{n_2})) \right]^2}{\partial p_{1i_1} \partial p_{2k_2}} \\ &= -2 E \left([f_{\tilde{x}_{1i_1}} - f_{\tilde{x}_{1L_1}}][f_{\tilde{x}_{2k_2}} - f_{\tilde{x}_{2L_2}}] \mid W \otimes \varphi \right) > 0, \end{aligned}$$

para algún $(i_1, k_2) \in \{1, \dots, L_1 - 1\} \times \{1, \dots, L_2 - 1\}$, o bien:

$$h_{i_2 k_1} = \frac{\partial^2 \left[D_W^\varphi(\tilde{E}(\mathcal{X}_1 | \mathbf{f}_{n_1}), \tilde{E}(\mathcal{X}_2 | \mathbf{f}_{n_2})) \right]^2}{\partial p_{2i_2} \partial p_{1k_1}}$$

$$= -2 E \left([f_{\tilde{x}_{2i_1}} - f_{\tilde{x}_{2L_2}}][f_{\tilde{x}_{1k_1}} - f_{\tilde{x}_{1L_1}}] | W \otimes \varphi \right) > 0,$$

para algún $(i_2, k_1) \in \{1, \dots, L_2 - 1\} \times \{1, \dots, L_1 - 1\}$, o bien:

$$h_{i_2 k_2} = \frac{\partial^2 \left[D_W^\varphi(\tilde{E}(\mathcal{X}_1 | \mathbf{f}_{n_1}), \tilde{E}(\mathcal{X}_2 | \mathbf{f}_{n_2})) \right]^2}{\partial p_{2i_2} \partial p_{2k_2}}$$

$$= 2 E \left([f_{\tilde{x}_{2i_2}} - f_{\tilde{x}_{2L_2}}][f_{\tilde{x}_{2k_2}} - f_{\tilde{x}_{2L_2}}] | W \otimes \varphi \right) > 0,$$

para algún $(i_2, k_2) \in \{1, \dots, L_2 - 1\} \times \{1, \dots, L_2 - 1\}$, entonces:

$$\left\{ 2(n_1 + n_2) \left[\left[D_W^\varphi(\tilde{E}(\mathcal{X}_1 | \mathbf{f}_{n_1}), \tilde{E}(\mathcal{X}_2 | \mathbf{f}_{n_2})) \right]^2 - \left[D_W^\varphi(\tilde{E}(\mathcal{X}_1 | \mathbf{p}_1), \tilde{E}(\mathcal{X}_2 | \mathbf{p}_2)) \right]^2 \right] \right\}_{n_1, n_2}$$

es una sucesión de variables aleatorias reales que cuando $n_1 \rightarrow \infty$ y $n_2 \rightarrow \infty$ converge en ley hacia una combinación lineal de, a lo sumo, $L_1 + L_2 - 2$ variables chi-cuadrado con 1 grado de libertad e independientes.

Cuando se cumple la igualdad $\tilde{E}(\mathcal{X}_1 | \mathbf{p}_1) = \tilde{E}(\mathcal{X}_2 | \mathbf{p}_2)$, que equivale a que $\left[D_W^\varphi(\tilde{E}(\mathcal{X}_1 | \mathbf{p}_1), \tilde{E}(\mathcal{X}_2 | \mathbf{p}_2)) \right]^2 = 0$, entonces la varianza $\sigma^2(\mathbf{p}_1, \mathbf{p}_2)$ se anula. Como implicación, sobre la base de la proposición precedente y de las propiedades de las convergencias en ley y en probabilidad (y, especialmente, del Teorema de Slutsky), se obtiene el siguiente procedimiento de contraste:

Teorema 2.2.4 Para contrastar al nivel de significación nominal α la hipótesis nula:

$$H_0 : \tilde{E}(\mathcal{X}_1|\mathbf{p}_1) = \tilde{E}(\mathcal{X}_2|\mathbf{p}_2)$$

frente a la alternativa:

$$H_a : \tilde{E}(\mathcal{X}_1|\mathbf{p}_1) \neq \tilde{E}(\mathcal{X}_2|\mathbf{p}_2),$$

H_0 debe rechazarse siempre que:

$$2(n_1 + n_2) [D_W^\varphi(\bar{\mathcal{X}}_1, \bar{\mathcal{X}}_2)]^2 > \tau_\alpha,$$

donde τ_α es el percentil de orden $100(1 - \alpha)$ de la distribución de $\hat{\lambda}_1 \chi_{1,1}^2 + \dots + \hat{\lambda}_q \chi_{1,q}^2$ en la que $\hat{\lambda}_1, \dots, \hat{\lambda}_q$ son los autovalores muestrales no nulos de la matriz $(L_1 + L_2 - 2) \times (L_1 + L_2 - 2)$ dada por:

$$\mathbf{B}(\mathbf{f}_{n_1}, \mathbf{f}_{n_2}) \mathbf{H} \left([D_W^\varphi(\bar{\mathcal{X}}_1, \bar{\mathcal{X}}_2)]^2 \right) \mathbf{B}(\mathbf{f}_{n_1}, \mathbf{f}_{n_2})^t,$$

con:

$$\begin{aligned} & \mathbf{B}(\mathbf{f}_{n_1}, \mathbf{f}_{n_2}) \mathbf{B}(\mathbf{f}_{n_1}, \mathbf{f}_{n_2})^t = [I_{\mathcal{X}}^F(\mathbf{f}_{n_1}, \mathbf{f}_{n_2})]^{-1} \\ & = \left(\begin{array}{c|c} \frac{n_1 + n_2}{n_1} [I_{\mathcal{X}_1}^F(\mathbf{f}_{n_1})]^{-1} & \mathbf{0} \\ \hline \mathbf{0} & \frac{n_1 + n_2}{n_2} [I_{\mathcal{X}_2}^F(\mathbf{f}_{n_2})]^{-1} \end{array} \right), \end{aligned}$$

y la matriz hessiana muestral dada por:

$$\begin{aligned} & \mathbf{H} \left([D_W^\varphi(\bar{\mathcal{X}}_1, \bar{\mathcal{X}}_2)]^2 \right) \\ & = \left(\begin{array}{c|c} [h_{i_1 k_1}] & [h_{i_1 k_2}] \\ \hline [h_{i_2 k_1}] & [h_{i_2 k_2}] \end{array} \right), \end{aligned}$$

(que es independiente de la muestra).

Además, la probabilidad de rechazar H_0 bajo H_a converge a 1 cuando $n \rightarrow \infty$.

La conclusión del Teorema 2.2.4 se aplica a continuación sobre un ejemplo real cuyos datos fueron proporcionados por miembros del Servicio de Conservación de la Naturaleza, Dirección Regional de Montes y Medio Ambiente de la Consejería de Agricultura, Medio Rural y Pesca del Principado de Asturias.

Ejemplo 2.2.1 Se consideran las poblaciones Ω_1 y Ω_2 de los días de dos determinados años, en los cuales se ha observado la variables nubosidad que, por seguir fielmente la notación del Teorema 2.2.4 se denotará por $(\mathcal{X})_1$ en la primera población y por $(\mathcal{X})_2$ en la segunda. Supongamos que se han seleccionado al azar y de forma independiente 50 días de Ω_1 y 45 de Ω_2 .

Las variables $(\mathcal{X})_1$ y $(\mathcal{X})_2$ toman los valores comunes *soleado* (\tilde{x}_1), *despejado* (\tilde{x}_2), *nubes y claros* (\tilde{x}_3), *nuboso* (\tilde{x}_4) y *totalmente cubierto* (\tilde{x}_5). Expertos en la medición de estos valores los han descrito en términos de los siguientes números difusos (entendidos como porcentajes) con soporte contenido en $[0, 100]$ y basados en S -curvas (ver Figura 2.3):

$$\begin{aligned} \tilde{x}_1 &= \tilde{x}_{11} = \tilde{x}_{21} = S(0, 20), \\ \tilde{x}_2 &= \tilde{x}_{12} = \tilde{x}_{22} = \begin{cases} S(10, 20) & \text{en } (-\infty, 20] \\ 1 - S(50, 70) & \text{en otro caso,} \end{cases} \\ \tilde{x}_3 &= \tilde{x}_{13} = \tilde{x}_{23} = \begin{cases} S(20, 40) & \text{en } [20, 40] \\ 1 & \text{en } [40, 50] \\ 1 - S(50, 70) & \text{en } [50, 70] \\ 0 & \text{en otro caso} \end{cases} \\ \tilde{x}_4 &= \tilde{x}_{14} = \tilde{x}_{24} = \begin{cases} S(60, 70) & \text{en } [60, 70] \\ 1 & \text{en } [70, 80] \\ 1 - S(80, 90) & \text{en } [80, 90] \\ 0 & \text{en otro caso,} \end{cases} \end{aligned}$$

$\tilde{x}_5 = \tilde{x}_{15} = \tilde{x}_{25} = 1 - S(80, 100)$, donde:

$$S(a, b)(t) = \begin{cases} 0 & \text{if } t \leq a \\ 2 \left(\frac{t-a}{b-a} \right)^2 & \text{si } t \in \left[a, \frac{a+b}{2} \right] \\ 1 - 2 \left(\frac{t-b}{b-a} \right)^2 & \text{si } t \in \left[\frac{a+b}{2}, b \right] \\ 1 & \text{en otro caso,} \end{cases}$$

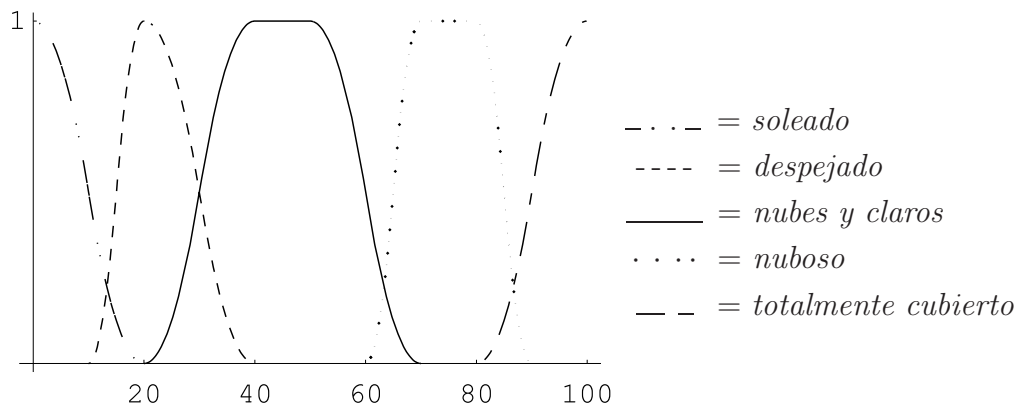


Fig. 2.3: Valores de la variable nubosidad

Para las muestras de días consideradas, los datos difusos observados se han recogido en la Tabla 2.5.

Tabla 2.5: Datos de nubosidad en ciertos días de Ω_1 y Ω_2

$\Omega_j \setminus \mathcal{X}$	\tilde{x}_1	\tilde{x}_2	\tilde{x}_3	\tilde{x}_4	\tilde{x}_5
Ω_1	5	21	12	9	3
Ω_2	2	19	7	12	5

Para contrastar la hipótesis nula H_0 de igualdad de los dos valores esperados difusos de la nubosidad en los años Ω_1 y Ω_2 , cuando W y φ se identifican con la medida de Lebesgue en $[0, 1]$, y aplicando las ideas de Pardo and Zografos (2000) para la aproximación de los autovalores muestrales, se concluye que el p -valor vendrá dado (aproximadamente) por $p = 0.126$, de modo que podría aceptarse H_0 para los niveles de significación nominales más usuales.

2.2.3 Test bilateral bootstrap para la igualdad de valores esperados poblacionales de dos variables aleatorias difusas simples

Los mismos motivos que llevan a introducir técnicas bootstrap en el test de una muestra hacen interesante realizar estudios análogos para el contraste con dos muestras. Al igual que anteriormente, se comenzará probando la consistencia del bootstrap elemental.

Sean $(\Omega_1, \mathcal{A}_1, P_1)$ y $(\Omega_2, \mathcal{A}_2, P_2)$ dos espacios de probabilidad. Se consideran $\mathcal{X}_1 : \Omega_1 \rightarrow \mathcal{F}_c(\mathbb{R})$ y $\mathcal{X}_2 : \Omega_2 \rightarrow \mathcal{F}_c(\mathbb{R})$ dos variables aleatorias difusas simples asociadas cada una a su respectivo espacio, que toman L_1 y L_2 valores diferentes, respectivamente.

Los valores de \mathcal{X}_1 se denotarán por $\tilde{x}_{11}, \dots, \tilde{x}_{1L_1}$ y sus probabilidades por p_{11}, \dots, p_{1L_1} . Igualmente los valores de \mathcal{X}_2 se denotarán por $\tilde{x}_{21}, \dots, \tilde{x}_{2L_2}$ y sus probabilidades por p_{21}, \dots, p_{2L_2} . Para cada $n_1 \in \mathbb{N}$ se considera una muestra aleatoria de n_1 observaciones independientes de \mathcal{X}_1 denotándose por $\mathbf{f}_{n_1} = (f_{n_11}, \dots, f_{n_1L_1})$ el vector aleatorio de frecuencias muestrales y, de la misma manera, para cada $n_2 \in \mathbb{N}$ se considera una muestra aleatoria de n_2 observaciones independientes de \mathcal{X}_2 y se denota por $\mathbf{f}_{n_2} = (f_{n_21}, \dots, f_{n_2L_2})$ su vector aleatorio de frecuencias muestrales.

Para especificar el estadístico del test bootstrap se consideran ahora nuevas poblaciones que verifiquen la hipótesis nula a partir de las realizaciones de las muestras aleatorias obtenidas. Como ya se ha comentado, para conseguir las nuevas poblaciones en estas condiciones en el caso real, es habitual calcular los residuos de los valores muestrales a la media común. En las condiciones de esta subsección, se conseguirá que las nuevas poblaciones tengan la misma media sumando a cada muestra la media de la otra, esto es, \mathcal{X}_1^* tomará los valores $\tilde{x}_{11} \oplus \bar{\mathcal{X}}_2, \dots, \tilde{x}_{1L_1} \oplus \bar{\mathcal{X}}_2$ con vector de distribución de frecuencias relativas muestrales $\mathbf{f}_{n_1} = (f_{n_11}, \dots, f_{n_1L_1})$ y \mathcal{X}_2^* tomará los valores $\tilde{x}_{21} \oplus \bar{\mathcal{X}}_1, \dots, \tilde{x}_{2L_2} \oplus \bar{\mathcal{X}}_1$ con vector de distribución de frecuencias relativas muestrales $\mathbf{f}_{n_2} = (f_{n_21}, \dots, f_{n_2L_2})$.

Así, se remuestreará de esas poblaciones obteniendo de nuevo muestras de tamaños n_1 y n_2 , respectivamente. Se denotan los vectores aleatorios asociados a las nuevas muestras por $\mathbf{f}_{n_1}^* = (f_{n_11}^*, \dots, f_{n_1L_1}^*)$ y $\mathbf{f}_{n_2}^* = (f_{n_21}^*, \dots, f_{n_2L_2}^*)$. El estadístico del contraste será

$$T_{n_1, n_2}^1 = \left[D_W^\varphi(\tilde{E}(\mathcal{X}_1^* | \mathbf{f}_{n_1}^*), \tilde{E}(\mathcal{X}_2^* | \mathbf{f}_{n_2}^*)) \right]^2$$

y su distribución puede aproximarse por el método de Monte Carlo. Se concluye a partir de ello que:

Teorema 2.2.5 *Para contrastar al nivel de significación nominal $\alpha \in [0, 1]$ la hipótesis nula:*

$$H_0 : \tilde{E}(\mathcal{X}_1 | \mathbf{p}_1) = \tilde{E}(\mathcal{X}_2 | \mathbf{p}_2)$$

frente a la alternativa

$$H_a : \tilde{E}(\mathcal{X}_1 | \mathbf{p}_1) \neq \tilde{E}(\mathcal{X}_2 | \mathbf{p}_2),$$

H_0 debe rechazarse (según una primera aproximación bootstrap) cuando

$$\left[D_W^\varphi(\bar{\mathcal{X}}_1, \bar{\mathcal{X}}_2) \right]^2 > z_\alpha,$$

donde z_α es el $100(1 - \alpha)$ percentil de la distribución de T_{n_1, n_2}^1 .

Demostración. Se utilizará la misma técnica que en el Teorema 2.1.6 aunque la notación será ligeramente diferente debido a la complicación que conlleva el manejo de dos muestras. En primer lugar, se debe notar que es equivalente multiplicar por $2(n_1 + n_2)$ el estadístico y el valor crítico. Sean \mathbf{f}_s^* y \mathbf{g}_t^* dos muestras aleatorias de tamaños s y t extraídas respectivamente de las poblaciones consideradas para \mathcal{X}_1^* y \mathcal{X}_2^* . Sean \mathbf{f} y \mathbf{g} dos distribuciones cualesquiera sobre esas poblaciones. El desarrollo de Taylor de $\left[D_W^\varphi \left(\tilde{E}(\mathcal{X}_1^* | \mathbf{f}_s^*), \tilde{E}(\mathcal{X}_2^* | \mathbf{g}_t^*) \right) \right]^2$ en un entorno de (\mathbf{f}, \mathbf{g}) se puede expresar como sigue:

$$\begin{aligned} & \left[D_W^\varphi \left(\tilde{E}(\mathcal{X}_1^* | \mathbf{f}_s^*), \tilde{E}(\mathcal{X}_2^* | \mathbf{g}_t^*) \right) \right]^2 = \left[D_W^\varphi \left(\tilde{E}(\mathcal{X}_1^* | \mathbf{f}), \tilde{E}(\mathcal{X}_2^* | \mathbf{g}) \right) \right]^2 \\ & \quad + \nabla \left[D_W^\varphi \left(\tilde{E}(\mathcal{X}_1^* | \mathbf{f}), \tilde{E}(\mathcal{X}_2^* | \mathbf{g}) \right) \right]^2 ((\mathbf{f}, \mathbf{g}) - (\mathbf{f}_s^*, \mathbf{g}_t^*))^t \\ & \quad + \frac{1}{2} ((\mathbf{f}, \mathbf{g}) - (\mathbf{f}_s^*, \mathbf{g}_t^*)) \mathbf{H} \left(\left[D_W^\varphi \left(\tilde{E}(\mathcal{X}_1^* | \mathbf{f}), \tilde{E}(\mathcal{X}_2^* | \mathbf{g}) \right) \right]^2 \right) ((\mathbf{f}, \mathbf{g}) - (\mathbf{f}_s^*, \mathbf{g}_t^*))^t \\ & \quad + R(\mathbf{f}, \mathbf{g}, \mathbf{f}_s, \mathbf{g}_t), \end{aligned}$$

donde el resto $R(\mathbf{f}, \mathbf{g}, \mathbf{f}_s, \mathbf{g}_t)$ es análogo al que aparece en el Teorema 2.1.3.

Evaluando la expresión anterior en el punto $(\mathbf{f}, \mathbf{g}) = (\mathbf{f}_{n_1}, \mathbf{f}_{n_2})$, los razonamientos seguidos en la Proposición 2.2.3 y en el Teorema 2.1.6 implican la consistencia.

Conviene notar que de nuevo la matriz hessiana no cambia aunque a las modalidades de las poblaciones originales se les haya sumado una constante y se calcule respecto a otra distribución de probabilidad ya que se tiene que:

$$\begin{aligned} & \mathbf{H} \left(\left[D_W^\varphi \left(\tilde{E}(\mathcal{X}_1 | P_1), \tilde{E}(\mathcal{X}_2 | P_2) \right) \right]^2 \right) \\ & = \mathbf{H} \left(\left[D_W^\varphi \left(\tilde{E}(\mathcal{X}_1 + \tilde{V}_1 | P'_1), \tilde{E}(\mathcal{X}_2 + \tilde{V}_2 | P'_2) \right) \right]^2 \right) \end{aligned}$$

cualesquiera que sean $\tilde{V}_1, \tilde{V}_2 \in \mathcal{F}_c(\mathbb{R})$ y P_1, P_2, P'_1, P'_2 distribuciones de probabilidad en las respectivas poblaciones. \square

De nuevo, con el fin de mejorar la precisión, parece conveniente incluir la variabilidad en el estadístico del test bootstrap.

Si se supone que las varianzas poblacionales son iguales, se puede considerar el estadístico análogo al propuesto en Efron and Tibshirani (1993), es decir:

$$T_{n_1, n_2}^2 = \frac{\left[D_W^\varphi(\tilde{E}(\mathcal{X}_1^* | \mathbf{f}_{n_1}^*), \tilde{E}(\mathcal{X}_2^* | \mathbf{f}_{n_2}^*)) \right]^2}{\hat{S}^2 \left(\mathcal{X}^* \left| \frac{n_1 \mathbf{f}_{n_1}^*}{n_1 + n_2}, \frac{n_2 \mathbf{f}_{n_2}^*}{n_1 + n_2} \right. \right)}$$

donde el denominador representa la estimación de la varianza de la población global \mathcal{X}^* , que viene dada por:

$$\hat{S}^2(\mathcal{X}_1^* | \mathbf{f}_{n_1}^*) \frac{n_1 - 1}{n_1 + n_2 - 2} + \hat{S}^2(\mathcal{X}_2^* | \mathbf{f}_{n_2}^*) \frac{n_2 - 1}{n_1 + n_2 - 2}.$$

De esta manera:

Teorema 2.2.6 *Para contrastar al nivel de significación nominal $\alpha \in [0, 1]$ la hipótesis nula:*

$$H_0 : \tilde{E}(\mathcal{X}_1 | P_1) = \tilde{E}(\mathcal{X}_2 | P_2)$$

frente a la alternativa

$$H_a : \tilde{E}(\mathcal{X}_1 | P_1) \neq \tilde{E}(\mathcal{X}_2 | P_2),$$

H_0 debe rechazarse (según una segunda aproximación bootstrap) siempre que:

$$\frac{\left[D_W^\varphi(\bar{\mathcal{X}}_1, \bar{\mathcal{X}}_2) \right]^2}{\hat{S}^2(\mathcal{X}_1 | \mathbf{f}_{n_1}) \frac{n_1 - 1}{n_1 + n_2 - 2} + \hat{S}^2(\mathcal{X}_2 | \mathbf{f}_{n_2}) \frac{n_2 - 1}{n_1 + n_2 - 2}} > z_\alpha,$$

donde z_α es el $100(1 - \alpha)$ percentil de la distribución de T_{n_1, n_2}^2 .

Nótese que es equivalente multiplicar tanto el estadístico como el valor crítico por $n_1 n_2 / (n_1 + n_2)$ y calcular su raíz cuadrada y conseguir así la forma exacta propuesta en Efron and Tibshirani (1993).

Siguiendo las pautas de Efron and Tibshirani (1993) se podría ampliar el ámbito de aplicación del test eliminando la suposición de varianzas iguales. Sin embargo, se debe recordar que el test de dos muestras sin la hipótesis de igualdad de varianzas se ha solucionado de manera aproximada de diversas formas en la literatura y es un problema que estaba en estudio y lo sigue estando en la actualidad (ver, por ejemplo, Scheffé, 1959 y Akahira, 2002). En Babu and Singh (1993) se prueba la adecuabilidad de esta técnica en el caso real, si bien se señala que los tamaños muestrales deben tender a infinito a la misma velocidad lo que se traduce, en la práctica, en que dichos tamaños muestrales sean 'comparables'. Este comentario se debe tener en cuenta a la hora de aplicar el test (tanto en el contexto real como en el difuso). El estadístico que se consideraría en esta situación sería:

$$T_{n_1, n_2}^3 = \frac{\left[D_W^\varphi(\tilde{E}(\mathcal{X}_1^* | \mathbf{f}_{n_1}^*), \tilde{E}(\mathcal{X}_2^* | \mathbf{f}_{n_2}^*)) \right]^2}{\frac{\hat{S}^2(\mathcal{X}_1^* | \mathbf{f}_{n_1}^*)}{n_1} + \frac{\hat{S}^2(\mathcal{X}_2^* | \mathbf{f}_{n_2}^*)}{n_2}},$$

de modo que:

Teorema 2.2.7 *Para contrastar al nivel de significación nominal $\alpha \in [0, 1]$ la hipótesis nula:*

$$H_0 : \tilde{E}(\mathcal{X}_1 | P_1) = \tilde{E}(\mathcal{X}_2 | P_2)$$

frente a la alternativa

$$H_a : \tilde{E}(\mathcal{X}_1 | P_1) \neq \tilde{E}(\mathcal{X}_2 | P_2),$$

H_0 debe rechazarse (según la tercera aproximación bootstrap) siempre que:

$$\frac{\left[D_W^\varphi(\bar{\mathcal{X}}_{1n}, \bar{\mathcal{X}}_{2m}) \right]^2}{\frac{\widehat{S}^2(\mathcal{X}_1 | \mathbf{f}_{n_1})}{n_1} + \frac{\widehat{S}^2(\mathcal{X}_2 | \mathbf{f}_{n_2})}{n_2}} > z_\alpha,$$

donde z_α es el $100(1 - \alpha)$ percentil de la distribución de T_{n_1, n_2}^3 .

Observación 2.2.1 Siguiendo las ideas del teorema anterior, puede desarrollarse una versión bootstrap de la Proposición 2.2.3-ii) que sería útil para el estudio de la potencia del test en el Teorema 2.2.7 (que podría llevarse a cabo en términos del parámetro $\vartheta_W^\varphi = D_W^\varphi(\tilde{E}(\mathcal{X}_1 | P_1), \tilde{E}(\mathcal{X}_2 | P_2))$).

2.2.4 Estudios de simulación: comparación de técnicas de contraste sobre la igualdad de valores esperados de dos variables aleatorias difusas

En la práctica aplicaremos únicamente el tercer bootstrap por ser el más general y, previsiblemente y según los resultados obtenidos para una muestra, el más preciso. Para comparar el comportamiento de los tests propuestos se han desarrollado de nuevo algunas simulaciones.

En este caso se realizaron 40000 repeticiones, 1000 repeticiones bootstrap en cada una y se simuló únicamente algunas poblaciones elegidas al azar considerando ciertos tamaños muestrales. Se incluyen como novedad los resultados del test bootstrap para variables aleatorias reales recogido en Efron and Tibshirani (1993) con el fin efectuar comparaciones en relación a los tamaños muestrales necesarios para obtener conclusiones aceptables.

Tabla 2.6: Simulación de los tests asintóticos. Dos muestras

	$n_1 = 30, n_2 = 30$	$n_1 = 30, n_1 = 100$	$n_1 = 100, n_2 = 100$	$n_1 = 100, n_2 = 500$
M_1	5.05	4.97	5.09	4.91
M_2	6.02	7.21	5.35	5.56
M_3	5	5.85	5.04	4.97
M_4	5.02	7.24	5.11	5.13

Se denotará por M_1 el test asintótico, en el que se suponen conocidos los autovalores poblacionales. M_2 será el test asintótico empírico, en el que se estiman dichos autovalores. M_3 será el tercer test bootstrap y M_4 será el test bootstrap real aplicado al λ -promedio de las variables aleatorias difusas consideradas en cada caso. Las probabilidades φ y W serán la medida de Lebesgue en $[0, 1]$. Los resultados se presentan en la Tabla 2.6

En la Tabla 2.6 se puede apreciar que, al estimar los autovalores se pierde precisión y por ello, aunque el test M_1 presenta un buen comportamiento, M_2 necesita tamaños muestrales superiores. El test bootstrap en ambiente difuso M_3 mejora el comportamiento del test M_2 con tamaños muestrales medios y tiene un comportamiento muy similar a su análogo real M_4 , apreciándose en ambos los problemas derivados de la estimación de la varianza.

Ejemplo 2.2.2 Supongamos que se consideran las poblaciones Ω_1 y Ω_2 del Ejemplo 2.2.1 y los datos muestrales de la Tabla 2.5.

Al aplicar la técnica bootstrap para el test M_3 se obtiene un p -valor igual a 0.130 y, en consecuencia, no se rechazaría la hipótesis de igualdad de los valores esperados en ambas poblaciones.

2.3 Contrastes tipo ANOVA unifactorial para una variable aleatoria difusa con valores en $\mathcal{F}_c(\mathbb{R})$

La presente sección se ocupa del estudio inferencial sobre las medias de una variable aleatoria difusa en varias poblaciones independientes. Más concretamente, está dedicada a los problemas de contraste de hipótesis en el modelo del Análisis de Varianza unifactorial con efectos fijos, cuando el factor corresponde a una variable aleatoria difusa.

En general, se considerará un factor que puede actuar con J niveles distintos, y una variable respuesta con valores imprecisos y asimilable con una variable aleatoria difusa, \mathcal{X} , determinando J poblaciones que se supondrán independientes. A partir de la j -ésima población, que se denotará por \mathcal{X}_j , se generará una muestra aleatoria simple, $\mathcal{X}_{j1}, \dots, \mathcal{X}_{jn_j}$, de modo que:

$$\mathcal{X}_{ji} = \tilde{\mu}_j \oplus \mathcal{E}_{ji},$$

con $j = 1, \dots, J$, $i = 1, \dots, n_j$, lo que indica que el valor de la i -ésima observación (aleatoria) en la muestra de la j -ésima población se obtiene de acuerdo con un modelo lineal que viene dado por la suma de dos componentes: el *valor esperado en el j -ésimo nivel*, $\tilde{\mu}_j$ (o valor esperado de \mathcal{X}_j , que suele interpretarse como la suma de la media global de \mathcal{X} y el efecto asociado al j -ésimo nivel del factor), y un término de *error aleatorio difuso*, \mathcal{E}_{ji} , que matemáticamente se identificará con una variable aleatoria difusa.

Por las propiedades del valor esperado de una variable aleatoria difusa (ver, por ejemplo, Puri and Ralescu, 1986), si $(\Omega_j, \mathcal{A}_j, P_j)$ es el espacio probabilístico sobre el que está definida \mathcal{X}_j , se satisface que:

$$\tilde{E}(\mathcal{X}_{ji}|P_j) = \tilde{\mu}_j \oplus \tilde{E}(\mathcal{E}_{ji}|P_j),$$

para $j = 1, \dots, J$, $i = 1, \dots, n_j$, de donde se concluye que:

$$\tilde{E}(\mathcal{E}_{ji}|P_j) = \mathbf{1}_{\{0\}}.$$

El propósito común en esta sección es contrastar la hipótesis nula

$$H_0 : \tilde{\mu}_1 = \dots = \tilde{\mu}_J$$

a partir de las muestras aleatorias simples disponibles.

Para ello, va a realizarse en primer lugar el estudio cuando las variables aleatorias difusas \mathcal{X}_j son normales, comprobándose que el procedimiento en el que la distancia euclídea en \mathbb{R} se reemplaza por la métrica D_W^φ en $\mathcal{F}_c(\mathbb{R})$, se reduce al contraste ANOVA unifactorial tradicional.

Posteriormente, para el caso de variables aleatorias difusas simples, se desarrollarán procedimientos asintóticos y bootstrap.

Se ilustrará la aplicación de los procedimientos construidos con algunos ejemplos reales, y finalmente se compararán mediante estudios de simulación.

2.3.1 Test ANOVA unifactorial para variables aleatorias difusas normales

Sean $(\Omega_1, \mathcal{A}_1, P_1), \dots, (\Omega_J, \mathcal{A}_J, P_J)$, J espacios de probabilidad, $\mathcal{X}_1 : \Omega_1 \rightarrow \mathcal{F}_c(\mathbb{R}), \dots, \mathcal{X}_J : \Omega_J \rightarrow \mathcal{F}_c(\mathbb{R})$ variables aleatorias difusas normales asociadas cada una a su respectivo espacio, con varianza desconocida común σ^2 , y $\tilde{\mu}_1 = \tilde{E}(\mathcal{X}_1|P_1), \dots, \tilde{\mu}_J = \tilde{E}(\mathcal{X}_J|P_J)$ sus medias poblacionales respectivas.

El problema es contrastar la hipótesis nula

$$H_0 : \tilde{E}(\mathcal{X}_1|P_1) = \dots = \tilde{E}(\mathcal{X}_J|P_J)$$

(es decir,

$$\sum_{j=1}^J n_j \left[D_W^\varphi(\tilde{E}(\mathcal{X}_j|P_j), \tilde{E}(\mathcal{X}|P)) \right]^2 = 0,$$

con

$$\tilde{E}(\mathcal{X}|P) = \frac{n_1}{n} \tilde{E}(\mathcal{X}_1|P_1) \oplus \dots \oplus \frac{n_J}{n} \tilde{E}(\mathcal{X}_J|P_J).$$

Para ello, se consideran muestras aleatorias simples, $\mathcal{X}_{j1}, \dots, \mathcal{X}_{jn_j}$ a partir de \mathcal{X}_j ($j = 1, \dots, J$), y se supone que las J muestras son independientes. Se denota por $\bar{\mathcal{X}}_j$ la media muestral de \mathcal{X}_j , es decir:

$$\bar{\mathcal{X}}_j = \frac{1}{n_j} (\mathcal{X}_{j1} \oplus \dots \oplus \mathcal{X}_{jn_j}),$$

y por $\bar{\mathcal{X}}$ la media muestral global:

$$\bar{\mathcal{X}} = \frac{n_1}{n} \bar{\mathcal{X}}_1 \oplus \dots \oplus \frac{n_J}{n} \bar{\mathcal{X}}_J,$$

donde $n = n_1 + \dots + n_J$.

Para el fin planteado va a emplearse el estadístico:

$$Z = \frac{\sum_{j=1}^J n_j [D_W^\varphi(\bar{\mathcal{X}}_j, \bar{\mathcal{X}})]^2}{\sum_{j=1}^J \sum_{i=1}^{n_j} [D_W^\varphi(\mathcal{X}_{ji}, \bar{\mathcal{X}}_j)]^2}.$$

Por las propiedades de la función $f(\cdot, \cdot)$ y de la métrica D_W^φ , y para el modelo lineal considerado, si \mathcal{X}_j es una variable aleatoria difusa normal, el error aleatorio \mathcal{E}_{ji} se reduce a una variable aleatoria real con media nula y varianza σ^2 , y puede concluirse que cuando H_0 es cierta:

$$[D_W^\varphi(\bar{\mathcal{X}}_j, \bar{\mathcal{X}})]^2 = \left[\frac{1}{n_j} \sum_{i=1}^{n_j} \mathcal{E}_{ji} - \frac{1}{n} \sum_{j=1}^J \sum_{i=1}^{n_j} \mathcal{E}_{ji} \right]^2,$$

$$\sum_{j=1}^J \sum_{i=1}^{n_j} [D_W^\varphi(\mathcal{X}_{ji}, \bar{\mathcal{X}}_j)]^2 = \sum_{j=1}^J \sum_{i=1}^{n_j} \left[\mathcal{E}_{ji} - \frac{1}{n_j} \sum_{i=1}^{n_j} \mathcal{E}_{ji} \right]^2$$

para $j = 1, \dots, J$, por lo que Z coincide con el estadístico clásico para el contraste ANOVA unifactorial con variables reales normales en condiciones de homocedasticidad e independencia entre muestras y, en consecuencia, bajo H_0 el estadístico Z tiene distribución F de Snedecor con $J - 1$ y $n - J$ grados de libertad. Por todo ello, puede afirmarse que:

Teorema 2.3.1 *Para contrastar al nivel de significación α la hipótesis nula:*

$$H_0 : \tilde{\mu}_1 = \dots = \tilde{\mu}_J$$

frente a la alternativa de que no todas esas medias difusas sean coincidentes, H_0 debe rechazarse siempre que:

$$\frac{\sum_{j=1}^J n_j [D_W^\varphi(\bar{\mathcal{X}}_j, \bar{\mathcal{X}})]^2}{\sum_{j=1}^J \sum_{i=1}^{n_j} [D_W^\varphi(\mathcal{X}_{ji}, \bar{\mathcal{X}}_j)]^2} > F_{J-1, n-J, \alpha},$$

donde $F_{J-1, n-J, \alpha}$ denota el percentil de orden $100(1 - \alpha)$ de la distribución F de Snedecor con $J - 1$ y $n - J$ grados de libertad.

2.3.2 Test ANOVA unifactorial asintótico para variables aleatorias difusas simples

El resultado precedente es una extensión del problema estudiado en la Subsección 2.2.1. De modo análogo, para el caso de variables aleatorias difusas simples va a desarrollarse una extensión de los resultados de la Subsección 2.2.2.

Sean $(\Omega_1, \mathcal{A}_1, P_1), \dots, (\Omega_J, \mathcal{A}_J, P_J)$, J espacios de probabilidad, $\mathcal{X}_1 : \Omega_1 \rightarrow \mathcal{F}_c(\mathbb{R}), \dots, \mathcal{X}_J : \Omega_J \rightarrow \mathcal{F}_c(\mathbb{R})$ variables aleatorias difusas simples asociadas cada una a su respectivo espacio, con $\tilde{\mu}_1 = \tilde{E}(\mathcal{X}_1|P_1), \dots, \tilde{\mu}_J = \tilde{E}(\mathcal{X}_J|P_J)$ sus medias poblacionales respectivas.

Sean $\tilde{x}_{j1}, \dots, \tilde{x}_{jL_j} \in \mathcal{F}_c(\mathbb{R})$ los distintos valores de \mathcal{X}_j en la j -ésima población y p_{j1}, \dots, p_{jL_j} las probabilidades respectivas ($P_j(\{\omega_j \in \Omega_j | \mathcal{X}_j(\omega_j) = \tilde{x}_{jl_j}\}) = p_{jl_j} > 0, l_j = 1, \dots, L_j, \sum_{l_j=1}^{L_j} p_{jl_j} = 1$) para $j = 1, \dots, J$.

Si se denota la media poblacional de \mathcal{X}_j por $\tilde{E}(\mathcal{X}_j|\mathbf{p}_j)$, con

$$\mathbf{p}_j = (p_{j1}, \dots, p_{j(L_j-1)}) \in (0, 1)^{L_j-1},$$

se tiene que:

$$\tilde{E}(\mathcal{X}_j|\mathbf{p}_j) = p_{j1} \tilde{x}_{j1} \oplus \dots \oplus p_{jL_j} \tilde{x}_{jL_j}$$

y los parámetros vectoriales \mathbf{p}_j cumplen las condiciones de regularidad indicadas en la Subsección 1.3.3.

Sea $\mathcal{X}_{j1}, \dots, \mathcal{X}_{jn_j}$ una muestra aleatoria simple a partir de la variable $\mathcal{X}_j, j = 1, \dots, J$. Se denotará por f_{njl_j} el estadístico correspondiente a la frecuencia relativa de \tilde{x}_{jl_j} en la j -ésima muestra aleatoria, y por $\mathbf{f}_{n_j} = (f_{nj1}, \dots, f_{nj(L_j-1)})$. Obviamente:

$$\bar{\mathcal{X}}_j = \frac{1}{n_j} (\mathcal{X}_{j1} \oplus \dots \oplus \mathcal{X}_{jn_j}) = \tilde{E}(\mathcal{X}|\mathbf{f}_{n_j}) = f_{nj1} \tilde{x}_{j1} \oplus \dots \oplus f_{njL_j} \tilde{x}_{jL_j}.$$

Es fundamental señalar que la descomposición de la variación total en el ANOVA unifactorial con variables aleatorias reales, tiene su análogo en el caso de variables aleatorias difusas, pudiendo verificarse a través del empleo de la función auxiliar $f(\cdot, \cdot)$ por las propiedades de esta función y las de la métrica D_W^φ que:

$$\sum_{j=1}^J \sum_{i=1}^{n_j} [D_W^\varphi(\mathcal{X}_{ji}, \bar{\mathcal{X}})]^2 = \sum_{j=1}^J n_j [D_W^\varphi(\bar{\mathcal{X}}_j, \bar{\mathcal{X}})]^2 + \sum_{j=1}^J \sum_{i=1}^{n_j} [D_W^\varphi(\mathcal{X}_{ji}, \bar{\mathcal{X}}_j)]^2.$$

En consecuencia:

$$\sum_{j=1}^J \sum_{i=1}^{n_j} [D_W^\varphi(\mathcal{X}_{ji}, \bar{\mathcal{X}}_j)]^2$$

representa una medida de la *variación dentro de las poblaciones*, mientras que

$$\sum_{j=1}^J n_j [D_W^\varphi(\bar{\mathcal{X}}_j, \bar{\mathcal{X}})]^2$$

es una medida de la *variación entre poblaciones*.

Por otro lado, la sucesión de estimadores $\{\mathbf{f}_{n_j}\}_n$ es fuertemente consistente y se distribuye asintóticamente según una distribución normal (L_j-1) -dimensional $\mathcal{N}\left(\mathbf{p}_j, \left[\mathbf{I}_{\mathcal{X}_j}^F(\mathbf{p}_j)\right]^{-1}/n_j\right)$ cuando $n_j \rightarrow \infty$, y $n_j(\mathbf{f}_{n_j} - \mathbf{p}_j)\mathbf{I}_{\mathcal{X}_j}^F(\mathbf{p}_j)$ $(\mathbf{f}_{n_j} - \mathbf{p}_j)^t$ converge en distribución a una chi-cuadrado $\chi_{L_j-1}^2$ cuando $n_j \rightarrow \infty$ para todo $j \in \{1, \dots, J\}$. En consecuencia, razonando como en la Proposición 2.1.3 se llega al resultado siguiente:

Proposición 2.3.2 *Para cada J -upla $(n_1, \dots, n_J) \in \mathbb{N}^J$, se consideran n_j variables aleatorias difusas independientes e igualmente distribuidas que \mathcal{X}_j , definidas sobre el espacio de probabilidad $(\Omega_j, \mathcal{A}_j, P_j)$ de forma que $P_j(\{\omega_j \in \Omega_j \mid \mathcal{X}_j(\omega_j) = \tilde{x}_{jl_j}\}) = p_{jl_j} \in (0, 1)$, $l_j = 1, \dots, L_j$, $\sum_{l_j=1}^{L_j} p_{jl_j} = 1$, $j = 1, \dots, J$. Supongamos las J muestras independientes. Se cumple entonces que:*

- i) Si $\mathbf{f}_{n_j} = (f_{n_j 1}, \dots, f_{n_j(L_j-1)}) \in [0, 1]^{L_j-1}$, con $f_{n_j l_j} =$ frecuencia relativa de \tilde{x}_{jl_j} en la correspondiente realización de la j -ésima muestra aleatoria simple de tamaño n_j ($l_j = 1, \dots, L_j - 1$, $f_{n_j L_j} = 1 - \sum_{l_j=1}^{L_j-1} f_{n_j l_j}$), $\tilde{E}(\mathcal{X}_j | \mathbf{f}_{n_j})$ representa la media muestral de \mathcal{X}_j , y para $\mathbf{p} = (\mathbf{p}_1, \dots, \mathbf{p}_J)$ se denota la media ponderada global por:*

$$\tilde{E}(\mathcal{X} | \mathbf{p}) = \frac{n_1}{n} \tilde{E}(\mathcal{X}_1 | \mathbf{p}_1) \oplus \dots \oplus \frac{n_J}{n} \tilde{E}(\mathcal{X}_J | \mathbf{p}_J),$$

entonces, cuando $n_j \rightarrow \infty$ para todo $j \in \{1, \dots, J\}$ se satisface que:

$$\left[D_W^\varphi(\bar{\mathcal{X}}_j, \bar{\mathcal{X}}) \right]^2 \xrightarrow{c.s.} \left[D_W^\varphi(\tilde{E}(\mathcal{X}_j|\mathbf{p}_j), \tilde{E}(\mathcal{X}|\mathbf{p})) \right]^2,$$

$$\left[D_W^\varphi(\mathcal{X}_{ji}, \bar{\mathcal{X}}_j) \right]^2 \xrightarrow{c.s.} \left[D_W^\varphi(\mathcal{X}_{ji}, \tilde{E}(\mathcal{X}_j|\mathbf{p}_j)) \right]^2,$$

cualesquiera que sean $\mathbf{p}_j = (p_{j1}, \dots, p_{jL_j-1}) \in (0, 1)^{L_j-1}$, $j = 1, \dots, J$.

ii) La sucesión de variables aleatorias reales:

$$\left\{ \frac{\sqrt{n_1 + \dots + n_J} \left[\frac{\sum_{j=1}^J n_j \left[D_W^\varphi(\bar{\mathcal{X}}_j, \bar{\mathcal{X}}) \right]^2}{\sum_{j=1}^J n_j} - \frac{\sum_{j=1}^J n_j \left[D_W^\varphi(\tilde{E}(\mathcal{X}_j|\mathbf{p}_j), \tilde{E}(\mathcal{X}|\mathbf{p})) \right]^2}{\sum_{j=1}^J \sum_{i=1}^{n_j} \left[D_W^\varphi(\mathcal{X}_{ji}, \bar{\mathcal{X}}_j) \right]^2} \right]}{\sum_{j=1}^J \sum_{i=1}^{n_j} \left[D_W^\varphi(\mathcal{X}_{ji}, \tilde{E}(\mathcal{X}_j|\mathbf{p}_j)) \right]^2} \right\}_{n_1, \dots, n_J}$$

converge en ley, cuando $n_j \rightarrow \infty$ para todo $j \in \{1, \dots, J\}$, hacia una normal unidimensional $\mathcal{N}(0, \sigma^2(\mathbf{p}_1, \dots, \mathbf{p}_J))$, siempre que $\sigma^2(\mathbf{p}_1, \dots, \mathbf{p}_J) > 0$, con:

$$\sigma^2(\mathbf{p}_1, \dots, \mathbf{p}_J) = \nabla \frac{A(\mathbf{p})}{B(\mathbf{p})} \left[I_{\mathcal{X}}^F(\mathbf{p}_1, \dots, \mathbf{p}_J) \right]^{-1} \left(\nabla \frac{A(\mathbf{p})}{B(\mathbf{p})} \right)^t$$

con:

$$A(\mathbf{p}) = \sum_{j=1}^J n_j \left[D_W^\varphi(\tilde{E}(\mathcal{X}_j|\mathbf{p}_j), \tilde{E}(\mathcal{X}|\mathbf{p})) \right]^2,$$

$$B(\mathbf{p}) = \sum_{j=1}^J \sum_{i=1}^{n_j} \left[D_W^\varphi(\mathcal{X}_{ji}, \tilde{E}(\mathcal{X}_j|\mathbf{p}_j)) \right]^2,$$

$$\begin{aligned} \nabla \frac{A(\mathbf{p})}{B(\mathbf{p})} &= \left(\frac{\partial}{\partial p_{11}} \frac{A(\mathbf{p})}{B(\mathbf{p})}, \dots, \frac{\partial}{\partial p_{1(L_1-1)}} \frac{A(\mathbf{p})}{B(\mathbf{p})}, \dots, \right. \\ &\quad \left. \frac{\partial}{\partial p_{J1}} \frac{A(\mathbf{p})}{B(\mathbf{p})}, \dots, \frac{\partial}{\partial p_{J(L_J-1)}} \frac{A(\mathbf{p})}{B(\mathbf{p})} \right), \\ &\quad \frac{\partial}{\partial p_{kl_k}} A(\mathbf{p}) \\ &= 2 \sum_{j=1}^J n_j \left[E \left((f_{\tilde{x}_{kl_k}} - f_{\tilde{x}_{kL_k}}) \left(\delta_{jk} - \frac{n_k}{n} \right) (f_{\tilde{E}(\mathcal{X}_j|\mathbf{p}_j)} - f_{\tilde{E}(\mathcal{X}|\mathbf{p})}) \mid W \otimes \varphi \right) \right], \\ \frac{\partial}{\partial p_{kl_k}} B(\mathbf{p}) &= -2 \sum_{i_k=1}^{L_k} n_{ki_k} E \left((f_{\tilde{x}_{ki_k}} - f_{\tilde{E}(\mathcal{X}_k|p_k)}) (f_{\tilde{x}_{kl_k}} - f_{\tilde{x}_{kL_k}}) \mid W \otimes \varphi \right), \\ &\quad [I_{\mathcal{X}}^F(\mathbf{f}_{n_1}, \dots, \mathbf{f}_{n_J})]^{-1} \\ &= \begin{pmatrix} \frac{n_1 + \dots + n_J}{n_1} [I_{\mathcal{X}_1}^F(\mathbf{f}_{n_1})]^{-1} & \dots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \dots & \frac{n_1 + \dots + n_J}{n_J} [I_{\mathcal{X}_2}^F(\mathbf{f}_{n_J})]^{-1} \end{pmatrix}. \end{aligned}$$

iii) Si $\sigma^2(\mathbf{p}_1, \dots, \mathbf{p}_J) = 0$ y se cumple que :

$$h_{i_j k_{j'}}(\mathbf{p}) = \frac{\partial^2}{\partial p_{j i_j} \partial p_{j' k_{j'}}} \frac{A(\mathbf{p})}{B(\mathbf{p})} > 0,$$

para algún $(i_j, k_{j'}) \in \{1, \dots, L_j - 1\} \times \{1, \dots, L_{j'} - 1\}; j, j' \in \{1, \dots, J\}$,
con

$$\begin{aligned} &\frac{\partial^2}{\partial p_{j i_j} \partial p_{j' k_{j'}}} A(\mathbf{p}) \\ &= n_j \left(\delta_{j j'} - \frac{n_{j'}}{n} \right) E \left((f_{\tilde{x}_{j' k_{j'}}} - f_{\tilde{x}_{j' L_{j'}}}) (f_{\tilde{x}_{j i_j}} - f_{\tilde{x}_{j L_j}}) \mid W \otimes \varphi \right) \end{aligned}$$

$$\frac{\partial^2}{\partial p_{j i_j} \partial p_{j' k_{j'}}} B(\mathbf{p})$$

$$= -2\delta_{j j'} n_j E\left(\left(f_{\tilde{x}_{j' k_{j'}}} - f_{\tilde{x}_{j' L_{j'}}}\right) \left(f_{\tilde{x}_{j i_j}} - f_{\tilde{x}_{j L_j}}\right) \mid W \otimes \varphi\right),$$

entonces:

$$\left\{ 2(n_1 + \dots + n_J) \left[\frac{\sum_{j=1}^J n_j [D_W^\varphi(\bar{\mathcal{X}}_j, \bar{\mathcal{X}})]^2}{\sum_{j=1}^J \sum_{i=1}^{n_j} [D_W^\varphi(\mathcal{X}_{ji}, \bar{\mathcal{X}}_j)]^2} \right. \right. \\ \left. \left. - \frac{\sum_{j=1}^J n_j [D_W^\varphi(\tilde{E}(\mathcal{X}_j | \mathbf{p}_j), \tilde{E}(\mathcal{X} | \mathbf{p}))]^2}{\sum_{j=1}^J \sum_{i=1}^{n_j} [D_W^\varphi(\mathcal{X}_{ji}, \tilde{E}(\mathcal{X}_j | \mathbf{p}_j))]^2} \right] \right\}_{n_1, \dots, n_J}$$

es una sucesión de variables aleatorias reales que converge en ley, cuando $n_j \rightarrow \infty$ para todo $j \in \{1, \dots, J\}$, hacia una combinación lineal de, a lo sumo, $L_1 + \dots + L_J - J$ variables chi-cuadrado con 1 grado de libertad e independientes.

Cuando se cumple la igualdad $\tilde{E}(\mathcal{X}_1 | \mathbf{p}_1) = \dots = \tilde{E}(\mathcal{X}_J | \mathbf{p}_J)$, que equivale a que

$$\frac{\sum_{j=1}^J n_j [D_W^\varphi(\tilde{E}(\mathcal{X}_j | \mathbf{p}_j), \tilde{E}(\mathcal{X} | \mathbf{p}))]^2}{\sum_{j=1}^J \sum_{i=1}^{n_j} [D_W^\varphi(\mathcal{X}_{ji}, \tilde{E}(\mathcal{X}_j | \mathbf{p}_j))]^2} = 0,$$

entonces la varianza $\sigma^2(\mathbf{p}_1, \dots, \mathbf{p}_J)$ se anula. Como implicación, sobre la base de la proposición precedente y de las propiedades de las convergencias en ley y en probabilidad, se obtiene el siguiente procedimiento de contraste:

Teorema 2.3.3 Para contrastar al nivel de significación nominal α la hipótesis nula:

$$H_0 : \tilde{E}(\mathcal{X}_1|P_1) = \dots = \tilde{E}(\mathcal{X}_J|P_J)$$

frente a la alternativa de que algunos de esos valores esperados difieren, H_0 debe rechazarse siempre que:

$$2(n_1 + \dots + n_J) \frac{\sum_{j=1}^J n_j [D_W^\varphi(\bar{\mathcal{X}}_j, \bar{\mathcal{X}})]^2}{\sum_{j=1}^J \sum_{i=1}^{n_j} [D_W^\varphi(\mathcal{X}_{ji}, \bar{\mathcal{X}}_j)]^2} > \tau_\alpha,$$

donde τ_α es el percentil de orden $100(1 - \alpha)$ de la distribución de $\hat{\lambda}_1 \chi_{1,1}^2 + \dots + \hat{\lambda}_q \chi_{1,q}^2$ en la que $\hat{\lambda}_1, \dots, \hat{\lambda}_q$ son los autovalores muestrales no nulos de la matriz $(L_1 + \dots + L_J - J) \times (L_1 + \dots + L_J - J)$ dada por:

$$\mathbf{C}(\mathbf{f}_{n_1}, \dots, \mathbf{f}_{n_J}) \mathbf{H} \left(\frac{A(\mathbf{f})}{B(\mathbf{f})} \right) \mathbf{C}(\mathbf{f}_{n_1}, \dots, \mathbf{f}_{n_J})^t,$$

con:

$$\mathbf{C}(\mathbf{f}_{n_1}, \dots, \mathbf{f}_{n_J}) \mathbf{C}(\mathbf{f}_{n_1}, \dots, \mathbf{f}_{n_J})^t = [I_{\mathcal{X}}^F(\mathbf{f}_{n_1}, \dots, \mathbf{f}_{n_J})]^{-1},$$

y la matriz hessiana muestral dada por:

$$\mathbf{H} \left(\frac{A(\mathbf{f})}{B(\mathbf{f})} \right) = [h_{i_j k_{j'}}(\mathbf{f})]$$

con $\mathbf{f} = (\mathbf{f}_{n_1}, \dots, \mathbf{f}_{n_J})$.

Además, la probabilidad de rechazar H_0 bajo H_a converge a 1 cuando $n_j \rightarrow \infty$ para todo $j \in \{1, \dots, J\}$.

Observación 2.3.1 Se debe mencionar que el test asintótico ANOVA (Teorema 2.3.3) para $J = 2$ y el asintótico de dos muestras (Teorema 2.2.4) presentados en esta memoria no son equivalentes, ya que mientras el primero tiene en cuenta la variabilidad, el segundo no va a tomarla en consideración.

Ejemplo 2.3.1 Los alumnos de cierta diplomatura de la Universidad de Oviedo han sido clasificados en tres poblaciones, Ω_1 , Ω_2 y Ω_3 , atendiendo a la frecuencia con la que practican deporte (*nunca, esporádicamente y a menudo* respectivamente).

En dichas poblaciones se ha analizado el tiempo que dedican a ver la televisión diariamente, $\tilde{x}_1 = \text{aproximadamente 1 hora o menos}$, $\tilde{x}_2 = \text{entre 1 y 2 horas aproximadamente}$ y $\tilde{x}_3 = \text{aproximadamente 3 horas o más}$. Dichos valores se han descrito mediante los siguientes números difusos (con escala de abscisas en minutos) (ver Figura 2.7):

$$\tilde{x}_1 = \begin{cases} 1 & \text{en } [0, 60] \\ 1 - S(60, 70) & \text{en } [60, 70] \\ 0 & \text{en otro caso} \end{cases}$$

$$\tilde{x}_2 = \begin{cases} S(50, 60) & \text{en } [50, 60] \\ 1 & \text{en } [60, 120] \\ 1 - S(120, 130) & \text{en } [120, 130] \\ 0 & \text{en otro caso} \end{cases}$$

$$\tilde{x}_3 = \begin{cases} S(110, 120) & \text{en } [110, 120] \\ 1 & \text{en } [120, 1440] \\ 0 & \text{en otro caso} \end{cases}$$

Para las muestras consideradas, los datos difusos observados se han recogido en la Tabla 2.7.

Para contrastar la hipótesis nula H_0 de igualdad de los valores esperados difusos del tiempo dedicado a ver la televisión en las poblaciones Ω_1 , Ω_2 y Ω_3 , cuando W y φ se identifican con la medida de Lebesgue en $[0, 1]$, y aplicando las ideas de Pardo and Zografos (2000) para la aproximación de

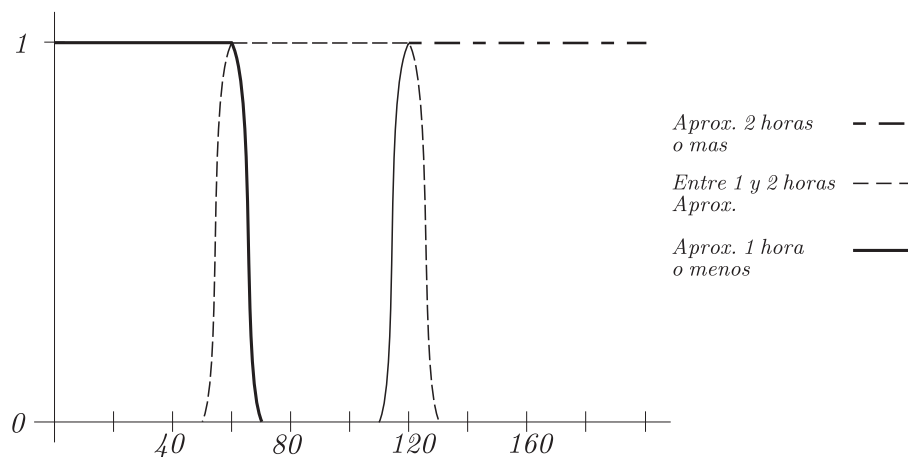


Fig. 2.4: Valores de la variable tiempo de TV

Tabla 2.7: Datos muestrales del tiempo dedicado a la TV

$\Omega_j \setminus \mathcal{X}$	\tilde{x}_1	\tilde{x}_2	\tilde{x}_3
Ω_1	20	12	4
Ω_2	16	24	20
Ω_3	14	12	2

los autovalores muestrales, se concluye que el p -valor vendrá dado (aproximadamente) por $p = 0.008$, de modo que H_0 se rechazaría para los niveles de significación usuales, por lo que podría afirmarse que, en la población genérica considerada, hay diferencias en el tiempo medio dedicado a ver la televisión según la frecuencia con la que hacen deporte.

2.3.3 Test ANOVA unifactorial bootstrap para variables aleatorias difusas simples

Si se denota por $\bar{\mathcal{X}}$ la media global $(n_1 + n_2)^{-1}(n_1 \mathcal{X}_1 \oplus n_2 \mathcal{X}_2)$, se puede probar que

$$\frac{n_1 n_2}{n_1 + n_2} \left[D_W^\varphi(\bar{\mathcal{X}}_1, \bar{\mathcal{X}}_2) \right]^2 = n_1 \left[D_W^\varphi(\bar{\mathcal{X}}_1, \bar{\mathcal{X}}) \right]^2 + n_2 \left[D_W^\varphi(\bar{\mathcal{X}}_2, \bar{\mathcal{X}}) \right]^2$$

y en consecuencia el bootstrap propuesto para dos muestras es válido como test bootstrap para el ANOVA de dos muestras.

De esta forma, el modelo bootstrap se puede generalizar al caso de J poblaciones de la siguiente manera. Sean $(\Omega_j, \mathcal{A}_j, P_j)$ espacios de probabilidad y \mathcal{X}_j variables aleatorias difusas asociadas a ellos y que toman los valores $\tilde{x}_{j1}, \dots, \tilde{x}_{jL_j}$ con probabilidades p_{j1}, \dots, p_{jL_j} ($j = 1, \dots, J$). Se considera una muestra aleatoria de cada una de esas poblaciones de tamaño n_j y se denotan los vectores aleatorios de frecuencias por $\mathbf{f}_{n_j} = (f_{n_j1}, \dots, f_{n_jL_j})$ ($j = 1, \dots, J$).

Para conseguir poblaciones bootstrap con media común a partir de las muestras obtenidas, se suma a cada muestra la media de las restantes, esto es, \mathcal{X}_j^* tomará los valores $\tilde{x}_{j1} \oplus \bar{\mathcal{X}}_{-j}, \dots, \tilde{x}_{jL_j} \oplus \bar{\mathcal{X}}_{-j}$ con distribución $\mathbf{f}_{n_j} = (f_{n_j1}, \dots, f_{n_jL_j})$ ($j = 1, \dots, J$) donde $\bar{\mathcal{X}}_{-j}$ representa la media global de las $J - 1$ muestras que se obtuvieron al excluir la j -ésima. Así, se remuestreará de esas poblaciones obteniendo de nuevo muestras de tamaño n_j y se denotarán los vectores aleatorios asociados las nuevas muestras por $\mathbf{f}_{n_j}^* = (f_{n_j1}^*, \dots, f_{n_jL_j}^*)$ ($j = 1, \dots, J$). El estadístico del contraste será

$$T_{(n_1, \dots, n_J)} = \frac{\sum_{j=1}^J n_j \left[D_W^\varphi(\tilde{E}(\mathcal{X}_j^* | \mathbf{f}_{n_j}^*), \tilde{E}(\mathcal{X}^* | \mathbf{f}^*)) \right]^2}{\sum_{j=1}^J \hat{S}^2(\mathcal{X}_j^* | \mathbf{f}_{n_j}^*) / n_j},$$

con $\tilde{E}(\mathcal{X}^* | \mathbf{f}^*)$ la media global de las J muestras bootstrap. La distribución de $T_{(n_1, \dots, n_J)}$ puede aproximarse por el método de Monte Carlo, y:

Teorema 2.3.4 *Para contrastar al nivel de significación nominal $\alpha \in [0, 1]$ la hipótesis nula:*

$$H_0 : \tilde{E}(\mathcal{X}_1|\mathbf{p}_1) = \dots = \tilde{E}(\mathcal{X}_J|\mathbf{p}_J)$$

frente a la alternativa de que no todas esas medias difusas sean coincidentes, H_0 debe rechazarse siempre que

$$\frac{\sum_{j=1}^J n_j \left[D_W^\varphi(\tilde{E}(\mathcal{X}_j|\mathbf{f}_{n_j}), \tilde{E}(\mathcal{X}|\mathbf{f})) \right]^2}{\sum_{j=1}^J \hat{S}^2(\mathcal{X}_j|\mathbf{f}_{n_j})/n_j} > z_\alpha,$$

donde z_α es el $100(1 - \alpha)$ percentil de la distribución de $T_{(n_1, \dots, n_J)}$ y $\tilde{E}(\mathcal{X}|\mathbf{f})$ es la media global de las J muestras.

Observación 2.3.2 Para una discusión detallada sobre la potencia del test precedente se seguirían ideas análogas a las de la Observación 2.1.2.

Ejemplo 2.3.2 Se consideran las poblaciones Ω_1 , Ω_2 y Ω_3 del Ejemplo 2.3.1 y los datos muestrales de la Tabla 2.7. Al aplicar la técnica bootstrap para el test del Teorema 2.3.4 se obtiene un p -valor igual a 0.026. En consecuencia, a nivel 0.05 se rechazaría la igualdad de valores esperados en esas tres poblaciones, pero podría aceptarse si se considerara un nivel de significación igual a 0.025 ó 0.01.

2.3.4 Estudios de simulación: comparación de técnicas de contraste ANOVA para variables aleatorias difusas

En este caso se han realizado únicamente simulaciones del test bootstrap, ya que a la vista de los resultados previos no cabe esperar mejora con el asintótico empírico y, de esta manera, se simplifican notablemente los cálculos a realizar.

Tabla 2.8: **Simulación de los tests asintóticos. ANOVA**

(n_1, n_2, n_3)	(30,30,30)	(30,30,100)	(30,100,100)	(100,100,100)
A_1	5.11	4.74	4.49	4.97
A_2	5.09	4.69	4.48	5.01

Se han restringido las simulaciones al caso de 3 poblaciones con tamaños muestrales 30 y 100 ya que únicamente se pretende realizar un análisis introductorio. Se han realizado 40000 repeticiones en las mismas condiciones que anteriormente. En la Tabla 2.8 se comparan los resultados obtenidos por el test bootstrap para el ANOVA en el contexto difuso (A_1) con el análogo en el contexto real (A_2) para el λ -promedio de las variables aleatorias difusas originales.

2.4 Valoración final y problemas abiertos

Como valoración final del Capítulo 2 cabe señalar las contribuciones realizadas en cuanto a:

- el desarrollo de una metodología de estadística inferencial para el análisis y tratamiento de parámetros difusos a partir de datos experimentales con valores de conjunto difuso, basada en la noción de variable aleatoria difusa y en la métrica generalizada D_W^φ ; esta metodología se ha centrado fundamentalmente en los contrastes sobre valores esperados (difusos) para una, dos y varias muestras, construyéndose diferentes técnicas exactas y asintóticas de aplicación sencilla en problemas reales;
- el empleo de técnicas bootstrap es especialmente interesante y novedoso en el caso estudiado en este capítulo, por cuanto la falta de

modelos que se ajusten bien a una parte importante de distribuciones de variables aleatorias difusas en los problemas reales limita la aplicabilidad de los métodos exactos y suele mejorar la potencia de los asintóticos;

- los métodos basados en el bootstrap se han comparado con los asintóticos mediante procedimientos de simulación recientes (Colubi *et al.*, 2002) específicos para variables aleatorias difusas.

En relación con la investigación del Capítulo 2, se plantean varios problemas de interés para un futuro próximo. Entre ellos cabe citar:

- Como en el capítulo anterior, los estudios asintóticos del presente capítulo podrían desarrollarse siguiendo las ideas de Körner (2000) para el contraste relativo al valor esperado de una variable aleatoria difusa que, como ya se apuntó, no es muy realista puesto que obliga a conocer ciertos parámetros poblacionales. Sin embargo, el empleo de técnicas bootstrap permitiría aproximar las ideas de Körner dando lugar a un procedimiento útil y extensible a dos y más muestras.
- El problema de construir métodos de contraste para hipótesis unilaterales relativas a valores esperados de variables aleatorias difusas conlleva un inconveniente añadido: la ausencia de un criterio de ordenación de números difusos que esté universalmente aceptado y conduzca a un orden/preorden completo. En Montenegro *et al.* (2002) se recoge un estudio de introducción a este problema, en el que se establecen algunos procedimientos para tales contrastes en el caso de una muestra. El principal punto débil de estos procedimientos reside en la consideración de un criterio específico de ordenación en $\mathcal{F}_c(\mathbb{R})$, de modo que las conclusiones inferenciales sólo serán válidas para dicho criterio, lo que supone una especial contrariedad cuando se acepta la

“igualdad” entre valores difusos (ya que la mayoría de los criterios de ordenación consideran ‘indiferentes’ números difusos no coincidentes). Un problema abierto de interés en relación con estos estudios es el de discutir la influencia del criterio de ordenación sobre las conclusiones y la potencia de los tests.

- Por último, la metodología iniciada en este capítulo puede continuarse involucrando otros parámetros difusos asociados a variables aleatorias difusas y desarrollando para los mismos procedimientos de contraste exactos, asintóticos y bootstrap. Aunque el empleo de la métrica D_W^φ sería teóricamente válida para esos contrastes, cabe esperar que surja una gran complejidad en su empleo cuando se trabaje con parámetros difusos distintos del valor esperado, por lo que muy posiblemente deba recurrirse a otras estrategias.

Epílogo

En la presente memoria se han desarrollado procedimientos para el análisis y tratamiento estadístico de datos experimentales con valores de conjunto (de intervalo o difuso). Los procedimientos se han basado en la consideración de los modelos probabilísticos formalizados por los conjuntos aleatorios compactos y convexos y por la extensión de éstos mediante las variables aleatorias difusas, así como en ciertas métricas sobre los espacios de valores de conjunto considerados.

Los problemas abiertos específicos de cada capítulo se han apuntado al final del mismo, pero existen un buen número de problemas que afectan a ambos y que se resumen en las tres líneas siguientes:

- La conveniencia de disponer de *modelos, operativos y ajustables a un buen número de situaciones reales, para las distribuciones de conjuntos aleatorios y variables aleatorias difusas*, en forma análoga a los modelos tradicionales de Bernoulli, binomial, Poisson, gamma, etc. para las variables aleatorias con valores reales.

Para ello, se determinarán situaciones habituales en la práctica en las que el uso de conjuntos con valores de intervalo o difusos sea conveniente para formalizar los datos experimentales. Se elegirán los experimentos empíricos representativos y se determinarán los parámetros

o características que hacen que esos experimentos se puedan incluir en clases generales.

Posteriormente, se establecerán modelos teóricos, examinando sus características principales (valor esperado, variabilidad etc.) y se intentará caracterizar su distribución de la manera más sencilla posible.

Se validará la utilidad práctica de los modelos encontrados. Algunos ejemplos metodológicos ayudarán a determinar su adaptación.

- La *simulación de tipo Montecarlo a partir de los modelos de distribuciones establecidos para conjuntos aleatorios y variables aleatorias difusas*.

Estos procedimientos de simulación serán útiles en relación con: la determinación de los tamaños muestrales que garanticen buenas aproximaciones al aplicar las técnicas asintóticas; la aproximación de la potencia de aquellas técnicas para las que el cálculo de su expresión exacta no es viable; la comparación de técnicas basadas en diferentes aproximaciones para un mismo tipo de inferencia.

Para ello, se simularán los modelos establecidos con el fin de estudiar la fiabilidad de las técnicas asintóticas y bootstrap de esta memoria. Así, se podrán determinar (entre otros) los tamaños muestrales idóneos para poder aplicar cada una de ellas fijados unos límites de error. Cuando sea posible, se intentarán justificar teóricamente las conclusiones obtenidas para casos particulares.

- El *análisis de la robustez de los métodos construidos* en función de la caracterización elegida para los datos difusos, o de los estadísticos que sirven de base para tales métodos.

Para ello, se llevará a cabo un estudio teórico/práctico de la robustez de las técnicas usuales sobre la base de los modelos paramétricos co-

munes por perturbación bien de los parámetros que determinan las clases, bien de las formas básicas de los valores difusos que componen esas clases. En este sentido, tiene especial interés comprobar si un suavizado de las funciones que identifican los valores difusos (que convierta, por ejemplo, un número difuso triangular en una π -curva) puede afectar o no a ciertas conclusiones estadísticas.

Por otro lado, se estudiará el efecto de la elección de las medidas de ponderación W y φ en la potencia de los procedimientos de contraste basados en las métricas dependientes de ellas.

Referencias y Bibliografía Consultada

- Akahira, M. (2002). Confidence intervals for the difference of means: application to the Behrens-Fisher type problem. *Statist. Papers* **43**, 273–284.
- Arnold, B.F. (1996). An approach to fuzzy hypothesis testing. *Metrika* **44**, 119–126.
- Aumann, R.J. (1965). Integrals of set-valued functions. *J. Math. Anal. Appl.* **12**, 1–12.
- Babu, G. J. and Singh, K. (1983). Inference on means using the bootstrap. *Ann. Statist.* **11**, 999-1003.
- Beran, R. and Srivastava, M.S. (1985). Bootstrap tests and confidence regions for functions of a covariance matrix. *Ann. Statist.* **13**, 95–115.
- Bertoluzza, C., Corral, N. and Salas A. (1995). On a new class of distances between fuzzy numbers. *Math. & Soft Comput.* **2**, 71–84.
- Billingsley, P. (1995). *Probability and Measure* (3rd ed.). John Wiley & Sons, New York.

- Casals, M.R., Gil, M.A. and Gil, P. (1986). The fuzzy decision problem: an approach to the problem of testing statistical hypotheses with fuzzy information. *Eur. J. Oper. Res.* **27**, 371–382.
- Casella, G. and Berger, R.L. (1990). *Statistical Inference*. Wadsworth & Brooks/Cole, Pacific Grove.
- Chatterjee, S., Price, B. (1991). *Regression Analysis by Example*. John Wiley & Sons, New York.
- Colubi, A., Fernández-García, C. and Gil, M.A. (2002). Simulation of random fuzzy variables: an empirical approach to statistical/probabilistic studies with fuzzy experimental data. *IEEE Trans. Fuzzy Syst.* **10**, 384–390.
- Colubi, A., López-Díaz, M., Domínguez-Menchero, J.S. and Gil, M.A. (1999). A generalized strong law of large numbers. *Probab. Theory Relat. Fields* **114**, 401–417.
- Colubi, A., Domínguez-Menchero, J. S., López-Díaz, M. and Ralescu, D. A. (2002). On the formalization of fuzzy random variables. *Inform. Sci.* **133**, 3–6.
- Colubi, A., Domínguez-Menchero, J. S., López-Díaz, M. and Ralescu, D. A. (2002). A $D_E[0, 1]$ -representation of random upper semicontinuous functions. *Proc. Am. Math. Soc.* **130**, 3237–3242.
- Cox, E. (1994). *The Fuzzy Systems Handbook*. Academic Press, Cambridge.
- Cressie, N. (1978). A strong limit theorem for random sets. *Adv. Appl. Probab., Special Supplement* **10**, 36–46.
- Cressie, N. and Pardo, L. (2002). Model checking in loglinear models using ϕ -divergences and MLEs. *J. Statist. Plan. Infer.* **103**, 437–453.

- Cressie, N. (1993). *Statistical for Spatial Data*. John Wiley & Sons, New York.
- David, F.N. (1938). *Tables of the Correlation Coefficient*. Cambridge Univ. Press. London and New York.
- Davis, A.W. (1977). A differential equation approach to linear combinations of independent chi-squares. *J. Amer. Statist. Assoc.* **72**, 212–214.
- Diamond, P. (1988). Fuzzy least squares. *Inform. Sci.* **46**, 141–157.
- Diamond, P. (1990). Least Squares Fitting of Compact Set-Valued Data. *J. Math. Anal. Appl.* **147**, 531–544.
- Diamond, P. and Kloeden P. (1990). Metric spaces of fuzzy sets. *Fuzzy Sets and Systems* **35**, 241–249.
- Diamond P. and Kloeden P. (1994). *Metric Spaces of Fuzzy Sets: Theory and Applications*. World Scientific, Singapore.
- Dubois, D. and Prade, H. (1978). Operations on fuzzy numbers. *Int. J. Syst. Sci.* **9**, 613–626.
- Dubois, D. and Prade, H. (1980). *Fuzzy Sets and Systems: Theory and Applications*. Academic Press, New York.
- Dubois, D. and Prade, H. (1987). The mean value of a fuzzy number. *Fuzzy Sets and Systems* **24**, 279–300.
- Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman and Hall, New York.
- Gil, M.A. and Casals, M.R. (1988). An operative extension of the likelihood ratio test from fuzzy data. *Statist. Pap.* **29**, 191–203.

- Gil, M. A. and López-Díaz, M. (1996). Fundamentals and Bayesian analyses of decision problems with fuzzy-valued utilities. *Int. J. Approx. Reas.* **15**, 203–224.
- Grad, A. and Solomon, H. (1955). Distribution of quadratic forms and some applications. *Ann. Math. Statist.* **26**, 464–477.
- Guenter, W.C. (1964). *Analysis of Variance*. Englewood Cliffs, New Jersey.
- Hall, P. and Martin, M.A. (1988). On bootstrap resampling and iteration. *Biometrika* **75**, 661–671.
- Hotelling, H. (1953). New light on the correlation coefficient and its transforms. *J. Roy. Statist. Soc., Serie B* **15**, 19-3-225, discussion 225–232.
- Imhof, J.P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika* **48**, 419–426.
- Jensen, D.R. and Solomon, H. (1972). A Gaussian approximation to the distribution of a definite quadratic form. *J. Amer. Statist. Assoc.* **67**, 898–902.
- Jobson, J.D. (1991). *Applied Multivariate Data Analysis*. Springer-Verlag, New York.
- Johnson, N.L. and Kotz, S. (1968). Tables of distributions of positive definite quadratic forms in central normal variables. *Sankhya, Serie B* **30**, 303–314.
- Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions. Vol 1*. John Wiley & Sons, New York.
- Johnson, N.L., Kotz, S. and Balakrishnan, N. (1995). *Continuous Univariate Distributions. Vol 2*. John Wiley & Sons, New York.

- Kacprzyk, J. and Fedrizzi, M. (1992). *Fuzzy Regression Analysis*. Physica-Verlag, Heidelberg.
- Kendall, D.G. (1974). Foundations of a theory of random sets. In *Stochastic Geometry* (E.F. Harding and D.G. Kendall, Eds.), 3–9. John Wiley & Sons, New York.
- Klement, E.P., Puri M.L. and Ralescu, D.A. (1986). Limit theorems for fuzzy random variables. *Proc. R. Soc, Lond. A* **407**, 171–182.
- Körner, R. (1997a). On the variance of fuzzy random variables. *Fuzzy Sets and Systems* **92**, 83–93.
- Körner, R. (1997b). *Linear Models with Random Fuzzy Variables*. PhD Thesis. Freiberg University of Mining and Technology.
- Körner, R. (2000). An asymptotic α -test for the expectation of random fuzzy variables. *J. Statist. Plan. Infer.* **83**, 331–346.
- Körner, R. and Näther, W. (1998). Linear regression with random fuzzy variables: extended classical estimates, best linear estimates, least squares estimates. *Inform. Sci.* **109**, 95–118.
- Körner, R. and Näther, W. (2002). On the variance of random fuzzy variables. In *Statistical Modeling, Analysis and Management of Fuzzy Data* (C. Bertoluzza, M.A. Gil and D.A. Ralescu, Eds.). Physica-Verlag, Heidelberg. pp. 22–39.
- Kruse, R. (1987). On the variance of random sets. *J. Math. Anal. Appl.* **122**, 469–473.
- Kruse, R. and Meyer, K.D.(1987). *Statistics with Vague Data*. Reidel Publ. Co., Dordrecht.

- Kruse, R., Gebhardt, J. and Gil, M.A. (1999). Fuzzy Statistics. In *Wiley Encyclopedia of Electrical and Electronics Engineering* (J.G. Webster, Ed.), 181-196. John Wiley & Sons, New York.
- Kwakernaak, H. (1978). Fuzzy random variables. Part I: Definitions and theorems. *Inform. Sci.* **15**, 1–29.
- Kwakernaak, H. (1979). Fuzzy random variables. Part II: Algorithms and examples for the discrete case. *Inform. Sci.* **17**, 253–278.
- Lehmann, E.H. (1983). *Theory of Point Estimation*. John Wiley & Sons, New York.
- Lindman, H.R. (1992). *Analysis of Variance in Experimental Design*. Springer-Verlag, New York.
- López-Díaz, M. and Gil, M.A. (1997). Constructive definitions of fuzzy random variables. *Statist. Probab. Lett.* **36**, 135–143.
- López-Díaz, M. and Gil, M.A. (1998a). Approximating integrably bounded fuzzy random variables in terms of the “generalized” Hausdorff metric. *Inform. Sci.* **104**, 279–291.
- López-Díaz, M. and Gil, M.A. (1998b). The λ -average value and the fuzzy expectation of a fuzzy random variable. *Fuzzy Sets and Systems* **99**, 347–352.
- López-Díaz, M. and Gil, M.A. (1998c). Reversing the order of integration in iterated expectations of fuzzy random variables, and statistical applications. *J. Statist. Plan. Infer.* **74**, 11–29.
- Lubiano Gómez, M.A. (1999). *Medidas de Variación para Elementos Aleatorios Imprecisos*. PhD Thesis, University of Oviedo.

- Lubiano, M.A., Alonso, M.C. and Gil, M.A. (1999). Statistical inferences on the S-mean squared dispersion of a fuzzy random variable. *Proc. EUROFUSE-SIC'99*, 532–537.
- Lubiano, M.A. and Gil, M.A. (1999). Estimating the expected value of fuzzy random variables in random samplings from finite populations. *Statist. Pap.* **40**, 277–295.
- Lyashenko, N.N. (1980). The statistics of random compacts in the Euclidean space. *Zap. Nauchn. Semin. Leningr. Otd. Mat. Inst. Steklova* **98**, 115–139.
- Mardia, K.V., Kent, J.T. and Bibby, J.M. (1979). *Multivariate analysis*. Academic Press.
- Martínez López, I. (1991). *Aproximación de Algunas Familias de Índices de Variación en Poblaciones Finitas*. Tesis Doctoral. Universidad de Oviedo.
- Matheron, G. (1975). *Random Sets and Integral Geometry*. John Wiley & Sons, New York.
- Molchanov, I.S. (1993). *Limit Theorems for Unions of Random Closed Sets*. Lect. Notes Math. **1561**. Springer-Verlag, Berlin.
- Molchanov, I. (1998). A limit theorem for solutions of inequalities. *Scand. J. Statist.* **25**, 235–242.
- Molchanov, I. (1999). On strong laws of large numbers for random upper semicontinuous functions. *J. Math. Anal. Appl.* **235**, 349–355.
- Montenegro, M., Casals, M.R. and Gil, M.A. (1998). Estimating fuzzy parameters of fuzzy random variables in a Bayesian context. *Proc. IP-MU'98*, **II**, 1002–1007.

- Montenegro, M., Colubi, A., Casals, M.R. and Gil, M.A. (2002). Test of one-sided hypotheses on the expected value of a fuzzy random variable. In *Soft Methods in Probability, Statistics and Data Analysis* (P. Grzegorzewski, O. Hryniewicz and M.A. Gil, Eds.), 228–235. Physica-Verlag, Heidelberg.
- Morales, D. and Pardo, L. (2000). New smooth test statistics of goodness-of-fit for categorized composite null hypotheses. *Test* **9**, 173–190.
- Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*. Wiley, New York.
- Näther, W. (1988). On expectation and variance of random fuzzy sets with application to linear regression with vague observations. *BUSEFAL* **36**, 82–91.
- Näther, W. and Albrecht, M. (1990). Linear regression with random fuzzy observations. *Statistics* **21**, 521–531.
- Näther, W. (1997). Linear statistical inference for random fuzzy data. *Statistics* **29**, 221–240.
- Odeh, R.E. (1982). Critical values of the sample product-moment correlation coefficient in the bivariate normal distribution. *Commun. Statist.-Simul. & Comput.* **11**, 1–26.
- Pardo, J.A., Pardo, L. and Zografos, K. (2002). Minimum ϕ -divergence estimators with constraints in multinomial populations. *J. Statist. Plan. Infer.* **104**, 221–237.
- Pardo, L. and Zografos, K. (2000). Goodness of fit tests with missclassified data based on ϕ -divergences *Biom. J.* **42**, 223–237.

- Pardo, L., Pardo, M.C. and Zografos, K. (2001). Minimum φ -divergence estimator for homogeneity in multinomial populations. *Sankhya, Ser. A* **63**, 72–92.
- Peña Sánchez de Rivera, D. (1989). *Estadística. Modelos y Métodos. 2. Modelos Lineales y Series Temporales*. Alianza Universidad Textos, Madrid.
- Peña Sánchez de Rivera, D. (2002). *Análisis de Datos Multivariantes*. McGraw Hill Interamericana, Madrid.
- Puri, M.L. and Ralescu, D.A. (1985). The concept of normality for fuzzy random variables. *Ann. Probab.* **13**, 1373–1379.
- Puri, M.L. and Ralescu, D. (1986). Fuzzy random variables. *J. Math. Anal. Appl.* **114**, 409–422.
- Ralescu, A. and Ralescu, D.A. (1986). Fuzzy sets in statistical inference. In *The Mathematics of Fuzzy Systems* (A. Di Nola and A.G.S. Ventre, Eds.), 273–283. Verlag TÜV Rheinland, Köln.
- Ralescu, D.A. (1995a). Inequalities for fuzzy random variables. *Proc. 26th Iranian Math. Confer.*, 333–335.
- Ralescu, D.A. (1995b). Fuzzy random variables revisited. *Proc. IFES'95 and Fuzzy IEEE Joint Conference* **2**, 993–1000.
- Rao, C.R. (1973). *Linear Statistical Inference and Its Applications* (2nd ed.). John Wiley & Sons, New York.
- Satterthwaite, F.E. (1946). An approximate distribution of estimates of variance components. *Biometrics* **2**, 110–114.
- Scheffé, H. (1959). *The Analysis of Variance*. John Wiley & Sons, New York.
- Searle, S.R. (1971). *Linear Models*. John Wiley & Sons, New York.

- Seber, G.A.F. (1984). *Multivariate Observations*. John Wiley & Sons, New York.
- Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. John Wiley & Sons, New York.
- Shao, J. and Tu, D. (1995). *The Jackknife and the Bootstrap*. Springer Series in Statistics. Springer-Verlag, New York.
- Stoyan, D.(1998). Random sets: Models and Statistics. *Int. Stat. Rev.* **66**, 1–27.
- Stoyan, D., Kendall, W.S. and Mecke, J. (1995). *Stochastic Geometry and its Applications*. J. Wiley & Sons, Chichester.
- Subrahmaniam, K. and Subrahmaniam, K. (1983). Some extensions to miss F. N. David's tables of sample correlation coefficient: distribution function and percentiles. *Sankhya, Serie B* **45**, 75–147.
- Sutton, C.D.(1993). Computer-intensive methods for test about the mean of an asymmetrical distribution. *J. Amer. Statist. Assoc.* **88** , 802–810.
- Taheri, S.M. and Behboodian, J. (1999). Neyman-Pearson Lemma for fuzzy hypothesis testing. *Metrika* **49**, 3–17.
- Yager, R. R. (1981). A procedure for ordering fuzzy subsets of the unit interval. *Inform. Sci.* **24**, 143–161.
- Zadeh, L.A. (1975). The concept of a linguistic variable and its application to approximate reasoning, Part 1. *Inform. Sci.* **8**, 199–249; Part 2. *Inform. Sci.* **8**, 301–353; Part 3. *Inform. Sci.* **9**, 43–80.

Índice Temático

aplicación α -nivel, 5

coeficiente de determinación entre CAs en relación lineal de Tipo 1, 46

anulación e independencia estocástica, 49

condiciones equivalentes para los valores extremos, 48

variación explicada por relación lineal de Tipo 1 entre CAs, 47

coeficiente de determinación entre CAs en relación lineal de Tipo 2, 50

anulación e independencia estocástica, 54

condiciones equivalentes para los valores extremos, 52

variación explicada por relación lineal de Tipo 1 entre CAs, 47

conjunto aleatorio

con valores en $\mathcal{K}_c(\mathbb{R})$ (CA), 4

integrabilidad acotada de un, 4

normal, 4

simple, 5

valor esperado de un, 4

contraste ANOVA para VADs

normales con varianzas coincidentes, 155

simples (test asintótico), 160

simples (test bootstrap), 165

contraste bilateral del valor esperado de una VAD

normal con varianza desconocida, 109

normal con varianza conocida, 110

simple (test asintótico), 122

simple (test bootstrap), 127, 129

contraste de igualdad de valores esperados de dos VADs

normales con varianzas coincidentes, 136

normales con varianzas conocidas, 136

simples (test asintótico), 142

simples con varianzas poblaciones coincidentes (test bootstrap), 146

simples (test bootstrap), 148, 149

distancia

d_H de Hausdorff en $\mathcal{K}(\mathbb{R}^p)$, 3

d_W entre elementos de $\mathcal{K}_c(\mathbb{R})$, 7

D_W^φ entre elementos de $\mathcal{F}_c(\mathbb{R})$, 10

espacio

$\mathcal{F}_c(\mathbb{R})$, 2

$\mathcal{F}(\mathbb{R}^p)$, 8

$\mathcal{K}_c(\mathbb{R})$, 2

$\mathcal{K}(\mathbb{R}^p)$, 3

expresión de la métrica de Hausdorff en $\mathcal{K}_c(\mathbb{R})$, 4

expresión simplificada del valor esperado de un CA, 4

función combinación lineal convexa de extremos del intervalo A , f_A , 7

integrabilidad acotada

de conjuntos aleatorios, 4

de variables aleatorias difusas, 6

métrica

d_H de Hausdorff en $\mathcal{K}(\mathbb{R}^p)$, 3

d_W o W -distancia entre elementos de $\mathcal{K}_c(\mathbb{R})$, 7

D_W^φ o (W, φ) -distancia entre elementos de $\mathcal{F}_c(\mathbb{R})$, 10

momentos de las variables f_X , 17

problema de regresión lineal entre CAs de Tipo 1

anulación de la pendiente en la relación óptima, 23

coeficiente de determinación, 46

función objetivo del, 19, 20, 22

independencia estocástica, 24

solución (relación óptima), 19

variación explicada, 47

variación residual, 46

problema de regresión lineal entre CAs de Tipo 2

anulación de la pendiente en la relación óptima, 23

duplicidad de las soluciones, 35

función objetivo del, 26, 27, 37

independencia estocástica, 38

solución (relación óptima), 28

unicidad de la solución, 35

variación explicada, 51

variación residual, 50

producto

de un conjunto por un escalar, 2

de un conjunto difuso por un escalar, 2

punto medio del intervalo A , $\text{mid } A$, 9

relación entre los momentos de f_X y los de $\text{mid } X$ y $\text{spr } X$, 17

semiamplitud del intervalo A , $\text{spr } A$, 9

suma

de conjuntos difusos, 2

de Minkowski de conjuntos, 2

valor esperado

de un conjunto aleatorio, 4

de una variable aleatoria difusa, 6

valor de intervalo en $\mathcal{K}_c(\mathbb{R})$

función de combinación lineal convexa de los extremos,

punto medio, 9

semiamplitud, 9

variable aleatoria difusa

con valores en $\mathcal{F}_c(\mathbb{R})$ (VAD), 5

integrabilidad acotada de una, 6

normal, 6

simple, 6

valor esperado de una, 6

variación explicada en el problema de regresión de Tipo 1 entre CAs, 47

variación explicada en el problema de regresión de Tipo 2 entre CAs, 51