

Universidad de Oviedo

# Variabilidad y centralidad para elementos aleatorios

Ignacio Cascos Fernández

*Departamento de Estadística e I.O.  
y Didáctica de la Matemática*

Tesis Doctoral realizada  
bajo la dirección de los profesores  
Miguel López Díaz y  
María Ángeles Gil Álvarez



D. 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. Ignacio Cascos Fernández, bajo la dirección de los Doctores D. Miguel López Díaz y Dña. María Ángeles Gil Álvarez, con el título “Variabilidad y centralidad para elementos aleatorios”.

Oviedo, 19 de enero de 2004

Fdo. Pedro Gil Álvarez



D. Miguel López Díaz, Profesor Titular de Universidad del Departamento de Estadística e I.O. y D.M. de la Universidad de Oviedo y Dña. María Ángeles Gil Álvarez, Catedrática de Universidad del Departamento de Estadística e I.O. y D.M. de la Universidad de Oviedo

CERTIFICAN:

Que la memoria titulada “Variabilidad y centralidad para elementos aleatorios” para optar al Grado de Doctor por la Universidad de Oviedo, ha sido realizada por D. Ignacio Cascos Fernández bajo su dirección, dando su conformidad para que sea presentada para su defensa en la Universidad de Oviedo.

Oviedo, 19 de enero de 2004

Fdo. Miguel López Díaz      Fdo. María Ángeles Gil Álvarez



*A mi madre*





# Agradecimientos

La Profesora Ángeles Gil me ha dado la oportunidad de escribir una Tesis Doctoral y me ha animado en todo momento para llevar la tarea a buen término. De no haber sido por su sugerencia inicial para realizar el doctorado y los ánimos que siempre me ha dado, es muy posible que nunca hubiese realizado estudios de postgrado. Por todo esto le doy las gracias.

Desearía expresar mi más profundo agradecimiento a dos personas cuya influencia en esta memoria es fundamental, los Profesores Miguel López e Ilya Molchanov. Sin los consejos, colaboración y ayuda incondicional de Miguel López, además de su pacientísima lectura del manuscrito original, esta memoria no se podría haber escrito. Tampoco hubiera sido posible su escritura sin los problemas propuestos y sugerencias para su resolución de Ilya Molchanov. De ambos, además de su contrastada valía científica, me gustaría destacar su calidad humana que los convierte en ejemplos que intentaré seguir.

La práctica totalidad de mi etapa como estudiante de doctorado he sido becario en el Departamento de Estadística e I.O. y D.M. de la Universidad de Oviedo a cuyos miembros, empezando por su Director, el Profesor Pedro Gil, quiero agradecer el afable trato que me han dispensado. Gran parte de esta memoria la he escrito como Ayudante en el Departamento de Estadística e I.O. de la Universidad Pública de Navarra, donde he tenido una acogida extraordinariamente cordial por parte de todos los integrantes del departamento. Quiero agradecer a la Universidad Pública de Navarra la oportunidad que me ha dado de trabajar y dar clases en ella, agradecimiento que

personifico en la Directora del Departamento de Estadística e I.O., la Profesora Carmen García Olaverri.

A lo largo de mis estudios de doctorado he disfrutado de estancias, siempre con el Profesor Molchanov, en el Departamento de Estadística de la Universidad de Glasgow y en el Departamento de Estadística Matemática y Ciencias Actuariales de la Universidad de Berna. Quiero expresar mi gratitud a todos los miembros de dichos departamentos, especialmente a Katsiaryna Kaval de la Universidad de Glasgow y al Profesor Lutz Dümbgen de la Universidad de Berna. También deseo dar las gracias a Tomáš Mrkvička con quien coincidí como visitante en la Universidad de Glasgow.

Hay un extenso número de personas de la comunidad estadística y matemática a las que debo agradecer su disponibilidad en el envío de separatas, siempre que se lo he solicitado, así como la paciencia que han tenido para explicarme sus trabajos y dedicarme su tiempo. No citaré ningún nombre, por temor a dejarme varios sin mencionar, pero a todos ellos les doy las gracias.

Mis compañeros becarios del Departamento de Estadística e I.O. y D.M. de la Universidad de Oviedo, Susana Díaz y Enrique Miranda han colaborado en esta memoria creando un ambiente de trabajo muy sano. De ellos quiero también destacar el apoyo que me han ofrecido en todo momento. Del mismo modo debo agradecer a otro antiguo becario del departamento, Pedro Terán, tanto la ayuda científica que me ha dispensado cuando se lo he solicitado como algunas discusiones muy enriquecedoras.

Deseo también agradecer a mis compañeros de doctorado del Departamento de Matemáticas de la Universidad de Oviedo sus consejos y amistad. Principalmente a Maribel que es una extraordinaria compañera y cuyas sugerencias tras la lectura del manuscrito original han contribuido a mejorar esta memoria.

He dejado para último lugar a mi familia, cuyo apoyo no sólo a lo largo de los cuatro años de período doctoral, sino de siempre es sin lugar a ninguna duda el más decisivo para llevar esta empresa o cualquier otra a buen término. De este modo deseo darles las gracias a mi hermano, a mis abuelos y por supuesto a mi madre, a quien está dedicada esta memoria.

# Prólogo

Cuando disponemos de dos distribuciones de probabilidad multivariante o dos conjuntos de datos con el mismo número de variables, la comparación de ambas probabilidades, o ambos conjuntos de datos según su variabilidad puede sernos muy útil. Así, si resumimos la riqueza de un conjunto de individuos en una serie de variables, podremos comparar diferentes poblaciones para estudiar en cuál de ellas existen mayores desigualdades sociales. Si los datos que almacenamos en dichas variables son relativos al comportamiento de una cierta máquina o los resultados en bolsa de un conjunto de empresas, podremos estudiar el comportamiento de qué máquina o de qué conjunto de empresas es más inestable.

Es también útil para un análisis descriptivo de datos la obtención de conjuntos de puntos representativos de la población y, con mayor generalidad, la obtención de conjuntos de puntos que son representativos de la población en cierto grado. Para el caso unidimensional la mediana y los intervalos intercuantílicos suelen ser los estadísticos elegidos para desempeñar tal función, sin embargo en el caso multivariante existen diferentes candidatos y la elección de un estadístico u otro depende de la situación particular que estemos considerando.

En esta memoria estudiamos la variabilidad de distribuciones de probabilidad multidimensionales utilizando ordenaciones estocásticas convexas e índices de desigualdad y relacionamos el problema de las comparaciones estocásticas con el de la obtención de conjuntos de puntos centrales. Construimos un nuevo marco para obtener regiones centrales con respecto a una

distribución de probabilidad y demostramos que dicho marco generaliza varios casos particulares de regiones centrales ya clásicas. Además, proponemos varias definiciones para estudiar el grado de centralidad de un punto con respecto a una distribución de probabilidad.

La memoria está dividida en cinco capítulos, además del Capítulo 0. En el Capítulo 0 se enuncian conceptos previos que resultan útiles a lo largo de toda la memoria. En concreto, se presentan resultados sobre la convergencia débil de probabilidades, los conjuntos aleatorios, zonoides y zonotopos, funciones de profundidad y regiones centrales y finalmente órdenes estocásticos.

Los Capítulos 1 y 2 tratan la variabilidad de los elementos aleatorios. A lo largo del Capítulo 1 se estudia un nuevo orden estocástico convexo para vectores aleatorios, el orden *simétrico* y se extiende, conjuntamente con otros órdenes convexos clásicos, al caso de los conjuntos aleatorios. Finalmente, se propone una nueva métrica de probabilidad. La variabilidad de los conjuntos aleatorios se estudia en el Capítulo 2 mediante una nueva familia de índices de desigualdad cuyas propiedades más importantes y su relación con los órdenes estocásticos convexos son analizadas. Para finalizar, examinamos el comportamiento de dichos índices de desigualdad mediante un ejemplo práctico.

La centralidad de un punto con respecto a una distribución de probabilidad se trata en los Capítulos 3, 4 y 5. En los Capítulos 3 y 4 se propone un nuevo marco que engloba un gran número de regiones centrales, y por tanto funciones de profundidad, clásicas y también incluye las regiones centrales inspiradas en el orden estocástico *simétrico* del Capítulo 1. La idea básica consiste en hacer que la profundidad de un punto con respecto a una distribución de probabilidad venga determinada por otra probabilidad que se obtiene a partir de la original redistribuyendo la masa a lo largo de su soporte. Así se definen las *regiones centrales recortadas* y las *regiones centrales integrales*. De estas últimas es importante destacar su relación con la familia de los órdenes estocásticos integrales. Por último, en el Capítulo 5 se definen tres nuevas funciones de profundidad, todas ellas con la particularidad de que vienen determinadas por un número de observaciones independientes de

un vector aleatorio. Además, debido a su relación con las nuevas funciones de profundidad, se resuelve el problema del recubrimiento de la circunferencia con semicircunferencias y se define un nuevo orden estocástico según la variabilidad para vectores aleatorios.



# Índice general

Agradecimientos	ix
Prólogo	xi
<b>0. Preliminares</b>	<b>1</b>
0.1. Convergencia débil de probabilidades . . . . .	1
0.2. Conjuntos aleatorios . . . . .	4
0.2.1. Esperanza de Aumann . . . . .	6
0.3. Zonoides y zonotopos . . . . .	8
0.4. Funciones de profundidad y regiones centrales . . . . .	10
0.4.1. Profundidad semiespacial o de Tukey . . . . .	12
0.4.2. Regiones centrales del zonoide . . . . .	13
0.5. Órdenes estocásticos . . . . .	14
0.5.1. Órdenes estocásticos integrales . . . . .	15
0.5.2. Órdenes estocásticos por inclusión de conjuntos . . . . .	18
<b>1. Orden estocástico según la variabilidad</b>	<b>21</b>
1.1. Orden simétrico . . . . .	22
1.1.1. Propiedades preliminares y caracterización . . . . .	23
1.1.2. El orden simétrico por inclusión de conjuntos . . . . .	25
1.1.3. Relación con otros órdenes estocásticos . . . . .	27
1.1.4. La relación ' $\leq_{\text{sym}}$ ' como orden estocástico integral . . . . .	29
1.1.5. Condiciones de igualdad estocástica . . . . .	30

1.2.	Órdenes para conjuntos aleatorios . . . . .	33
1.2.1.	Orden simétrico . . . . .	33
1.2.2.	Orden lineal convexo . . . . .	35
1.2.3.	Orden creciente positivo lineal convexo . . . . .	38
1.3.	Métrica de probabilidad . . . . .	43
1.3.1.	Métrica ideal de orden 2 . . . . .	46
1.3.2.	Relación con otras métricas de probabilidad . . . . .	50
1.4.	Comentarios y problemas abiertos . . . . .	51
<b>2.</b>	<b>Índices de desigualdad para conjuntos aleatorios</b>	<b>53</b>
2.1.	Preliminares . . . . .	54
2.2.	Índices de desigualdad . . . . .	57
2.3.	Propiedades . . . . .	59
2.4.	Criterios de convergencia . . . . .	63
2.4.1.	Continuidad . . . . .	64
2.4.2.	Aproximación . . . . .	66
2.5.	Ejemplo . . . . .	69
2.6.	Comentarios y problemas abiertos . . . . .	73
<b>3.</b>	<b>Regiones centrales recortadas</b>	<b>75</b>
3.1.	Regiones centrales y profundidad multivariante . . . . .	77
3.2.	El recorte de una probabilidad . . . . .	78
3.2.1.	Propiedades generales de la familia $P^\alpha$ . . . . .	78
3.2.2.	Comportamiento empírico de $P^\alpha$ . . . . .	82
3.3.	Regiones centrales recortadas . . . . .	87
3.4.	Propiedades . . . . .	88
3.4.1.	Propiedades generales . . . . .	89
3.4.2.	Resultados asintóticos . . . . .	92
3.5.	Ejemplos . . . . .	93
3.5.1.	Intervalos intercuantílicos . . . . .	93
3.5.2.	Regiones centrales del zonoide . . . . .	94
3.5.3.	Regiones centrales de la envolvente de zonoides . . . . .	96



3.5.4. Regiones centrales conexas, pero no convexas . . . . .	99
3.6. Comentarios y problemas abiertos . . . . .	100
<b>4. Regiones centrales integrales</b>	<b>103</b>
4.1. Definición . . . . .	104
4.1.1. Preliminares . . . . .	105
4.2. Propiedades . . . . .	106
4.2.1. Generalidades . . . . .	106
4.2.2. Regiones anidadas . . . . .	107
4.2.3. Regiones convexas . . . . .	107
4.2.4. Regiones equivariantes afines . . . . .	110
4.2.5. Regiones cerradas . . . . .	112
4.2.6. Regiones acotadas . . . . .	113
4.3. Resultados asintóticos . . . . .	114
4.4. Ejemplos . . . . .	123
4.4.1. Regiones centrales semiespaciales . . . . .	123
4.4.2. Regiones centrales del zonoide . . . . .	126
4.4.3. Regiones centrales de la envolvente de zonoides . . . . .	128
4.4.4. Regiones centrales integrales y recortadas . . . . .	130
4.5. Comentarios y problemas abiertos . . . . .	131
<b>5. Profundidad según un número de observaciones</b>	<b>135</b>
5.1. Introducción . . . . .	136
5.2. Profundidades relacionadas con el grado . . . . .	138
5.2.1. Propiedades . . . . .	139
5.2.2. Ejemplos y aplicaciones . . . . .	144
5.3. Esperanza de la envolvente convexa . . . . .	148
5.3.1. Propiedades . . . . .	149
5.3.2. Orden estocástico en variabilidad inducido . . . . .	151
5.4. Comentarios y problemas abiertos . . . . .	155
<b>Epílogo</b>	<b>157</b>

xviii

*Índice general*

**Bibliografía** **159**

**Índice de figuras** **167**

**Lista de símbolos** **169**

**Índice alfabético** **173**

# Capítulo 0

## Preliminares

En este capítulo se recogen las nociones y resultados fundamentales que son necesarios para el desarrollo de la memoria. Se estructura en cinco secciones; en la primera se introduce la convergencia débil de probabilidades, y a continuación se estudian sucesivamente los conjuntos aleatorios, los zonoides y zonotopos y las funciones de profundidad y regiones centrales. La última sección está dedicada a los órdenes estocásticos.

### 0.1. Convergencia débil de probabilidades

A lo largo de esta memoria trabajaremos con distribuciones de probabilidad definidas sobre  $\mathcal{B}_d$ , la  $\sigma$ -álgebra de Borel asociada a la topología inducida por la métrica euclídea en  $\mathbb{R}^d$ . Al conjunto de tales distribuciones de probabilidad lo denotaremos por  $\mathbb{P}$ .

A continuación damos varias definiciones y resultados sobre probabilidades, en concreto sobre la convergencia débil de probabilidades, particularizados a  $\mathbb{P}$  y extraídos en su mayoría de Billingsley (1999).

*Definición 0.1.1.* Decimos que una sucesión de probabilidades  $\{P_n\}_{n \in \mathbb{N}} \subset \mathbb{P}$  converge débilmente hacia  $P \in \mathbb{P}$  si para toda función real continua y acotada  $f$ , se cumple  $\int_{\mathbb{R}^d} f dP_n \rightarrow \int_{\mathbb{R}^d} f dP$ . A la convergencia débil la denotamos por  $P_n \Rightarrow P$ .

Del mismo modo, dos probabilidades son iguales si la integral de toda función real continua y acotada respecto de ellas coincide.

Existen varias condiciones equivalentes a la convergencia débil, las más importantes se enuncian en el *Teorema de Portmanteau* (Teorema 2.1 en Billingsley, 1999). A continuación se establece dicho resultado en el caso particular de  $\mathbb{P}$ .

**Teorema 0.1.1.** Sean  $P, P_n \in \mathbb{P}$  para todo  $n \in \mathbb{N}$ , las siguientes cinco condiciones son equivalentes:

- i.  $P_n \Rightarrow P$ ;
- ii.  $\lim_n \int f dP_n = \int f dP$  para cualquier función real  $f$  uniformemente continua y acotada;
- iii.  $\limsup_n P_n(F) \leq P(F)$  para cualquier conjunto cerrado  $F \subset \mathbb{R}^d$ ;
- iv.  $\liminf_n P_n(G) \geq P(G)$  para cualquier conjunto abierto  $G \subset \mathbb{R}^d$ ;
- v.  $\lim_n P_n(A) = P(A)$  para cualquier conjunto de  $P$ -continuidad  $A \subset \mathbb{R}^d$ , un conjunto  $A \in \mathcal{B}_d$  se dice que es un conjunto de  $P$ -continuidad si  $P(\partial A) = 0$ , donde  $\partial A$  denota a la frontera de  $A$ .

Una probabilidad  $P$  definida sobre la  $\sigma$ -álgebra de Borel asociada a una topología se dice *regular* si para cualquier conjunto  $A$  de dicha  $\sigma$ -álgebra y cualquier  $\varepsilon > 0$ , existen un conjunto cerrado  $F$  y otro abierto  $G$  tales que  $F \subset A \subset G$  y  $P(G \setminus F) < \varepsilon$ . La condición de regularidad nos resultará muy útil para probar la pertenencia de diversas probabilidades a ciertas clases.

**Teorema 0.1.2 (Teorema 1.1 en Billingsley, 1999).** Toda probabilidad en un espacio métrico es regular.

En consecuencia, toda probabilidad  $P \in \mathbb{P}$  es regular.

El concepto de familias *tight* aparecerá con cierta frecuencia a lo largo de esta memoria.

*Definición 0.1.2.* Una familia  $\Pi \subset \mathbb{P}$  se dice *tight* si para todo  $\varepsilon > 0$  existe un conjunto compacto  $K \subset \mathbb{R}^d$  tal que  $P(K) > 1 - \varepsilon$  para todo  $P \in \Pi$ . Decimos que una probabilidad  $P$  es *tight* si la familia formada exclusivamente por  $P$  es *tight*.

El Teorema 1.4 en Billingsley (1999) establece que toda probabilidad en un espacio métrico completo y separable es *tight*. Como  $\mathbb{R}^d$  con la métrica euclídea es completo y separable, se cumple el siguiente resultado.

**Teorema 0.1.3.** *Toda probabilidad de  $\mathbb{P}$  es tight.*

Una vez que sabemos que toda probabilidad de  $\mathbb{P}$  es *tight*, el primer ejemplo de un conjunto de probabilidades que satisface la propiedad de ser *tight* es el formado por los todos los elementos de una sucesión débilmente convergente. El Teorema 8 en el Apéndice III de Billingsley (1999) establece que dada una sucesión de probabilidades *tight* en un espacio métrico que converge débilmente a otra probabilidad *tight*, el conjunto formado por todas las probabilidades de la sucesión y la probabilidad límite es *tight*. Como según el Teorema 0.1.3 tenemos que toda probabilidad de  $\mathbb{P}$  es *tight*, podemos especializar dicho resultado para las probabilidades de  $\mathbb{P}$ .

**Teorema 0.1.4.** *Si  $P, P_n \in \mathbb{P}$  para todo  $n \in \mathbb{N}$  y  $P_n \Rightarrow P$ , entonces el conjunto  $\{P, P_1, P_2, \dots\}$  es tight.*

### La topología de la convergencia débil

Se puede construir una topología en  $\mathbb{P}$  de tal modo que la convergencia con respecto a ella sea la convergencia débil. Los conjuntos de la forma

$$\{Q \in \mathbb{P} : Q(F_i) < P(F_i) + \varepsilon, i = 1, \dots, k\}$$

(donde  $P \in \mathbb{P}$ , los conjuntos  $F_i \subset \mathbb{R}^d$  son cerrados,  $\varepsilon > 0$  y  $k \in \mathbb{N}$ ), constituyen una base de la topología llamada *topología de la convergencia débil en  $\mathbb{P}$*  que es aquella para la que la convergencia débil coincide con la convergencia con respecto a ella misma.

Un conjunto se dice *relativamente compacto* si toda sucesión suya contiene alguna subsucesión convergente. El *Teorema de Prohorov* (Teorema 6.1 en Billingsley, 1999) establece que todo conjunto de probabilidades definidas en un espacio métrico que satisface la condición de ser tight es relativamente compacto. Tal condición ha de interpretarse con respecto a la topología de la convergencia débil.

**Teorema 0.1.5.** *Si  $\Pi \subset \mathbb{P}$  es tight, entonces es relativamente compacto.*

El *Teorema de Prohorov* tiene numerosas consecuencias en el contexto de la topología de la convergencia débil. Así el Teorema 6 en el Apéndice III de Billingsley (1999) establece una serie de resultados de los que nos interesa la siguiente particularización.

**Teorema 0.1.6.** *Si  $\Pi \subset \mathbb{P}$  es tight, entonces su clausura con respecto a la topología de la convergencia débil es compacta también con respecto a la topología de la convergencia débil.*

## 0.2. Conjuntos aleatorios

Un *conjunto aleatorio cerrado* en  $\mathbb{R}^d$  es una aplicación de cierto espacio de probabilidad  $(\Omega, \mathcal{A}, \Pr)$  en la familia de conjuntos cerrados en  $\mathbb{R}^d$  medible respecto de la  $\sigma$ -álgebra de Borel asociada a la topología “hit-or-miss” definida en la familia de los conjuntos cerrados. Esta  $\sigma$ -álgebra, a la que denotamos por  $\sigma_h$ , está engendrada por las familias de conjuntos cerrados que dado un conjunto compacto fijo, no tienen ningún elemento común con él, es decir,

$$\sigma_h := \sigma(\{ \{ F \subset \mathbb{R}^d \text{ cerrado} : F \cap K = \emptyset \} : K \subset \mathbb{R}^d \text{ compacto} \}).$$

Como referencias básicas en la Teoría de los Conjuntos Aleatorios cabe destacar las monografías de Matheron (1975), Molchanov (1993), Stoyan *et al.* (1995) y las breves recopilaciones de Stoyan (1998) y Molchanov (1999).

La *suma de Minkowski* de dos conjuntos  $A, B \subset \mathbb{R}^d$  se denota por  $A + B$  y es el conjunto formado por la suma de cada par de elementos de  $A$  y  $B$ , es decir,  $A + B := \{a + b : a \in A, b \in B\}$ . Dados dos conjuntos aleatorios cerrados, su unión, intersección y suma de Minkowski conservan la medibilidad y son, por tanto, conjuntos aleatorios cerrados, véase la Sección 1.5 en Molchanov (1993). Otro ejemplo de un conjunto aleatorio cerrado es el conjunto  $\{\xi\}$ , siempre que  $\xi$  sea un vector aleatorio.

Denotamos por  $\mathcal{K}$  a la familia de subconjuntos compactos no vacíos de  $\mathbb{R}^d$ . Sobre dicha clase se define la *distancia de Hausdorff*, que se denota por  $d_H$ , y para dos elementos  $A, B \in \mathcal{K}$ , viene dada por cualquiera de la dos siguientes expresiones equivalentes,

$$\begin{aligned} d_H(A, B) &:= \max \left\{ \sup_{a \in A} \inf_{b \in B} \|a - b\|, \sup_{b \in B} \inf_{a \in A} \|a - b\| \right\} \\ &= \inf \{r > 0 : A \subset B^r \text{ y } B \subset A^r\}, \end{aligned}$$

donde  $\|\cdot\|$  es la norma euclídea y  $A^r$  es el  $r$ -entorno del conjunto  $A$ , es decir, la suma de Minkowski de  $A$  y la bola cerrada centrada en el origen y de radio  $r$ , esto es,  $A^r := A + B_r$ .

El conjunto  $\mathcal{K}$  dotado de la métrica de Hausdorff forma un espacio métrico completo y separable, véase Debreu (1967).

Un *conjunto aleatorio compacto* en  $\mathbb{R}^d$  es un conjunto aleatorio cerrado que toma casi seguro conjuntos compactos como imagen. Una aplicación  $X : \Omega \rightarrow \mathcal{K}$  es un conjunto aleatorio compacto si y sólo si es medible respecto de la  $\sigma$ -álgebra de Borel asociada a la topología generada por la distancia de Hausdorff, véase, por ejemplo, Debreu (1967) y Stoyan (1998).

Un conjunto  $A \subset \mathbb{R}^d$  se dice *convexo* si para cualesquiera  $a, b \in A$  y  $\lambda \in (0, 1)$ , se tiene que  $\lambda a + (1 - \lambda)b \in A$ . Si un conjunto aleatorio cerrado es casi seguro convexo, decimos que se trata de un *conjunto aleatorio cerrado y convexo*, si un conjunto aleatorio cerrado es casi seguro compacto y convexo, decimos que se trata de un *conjunto aleatorio compacto y convexo*. Denotamos por  $\mathcal{K}_c$  a la familia de subconjuntos no vacíos compactos y convexos de  $\mathbb{R}^d$ , es decir a la familia de *cuerpos convexos* de  $\mathbb{R}^d$ .

Un vector aleatorio  $\xi$  es una *selección* de un conjunto aleatorio  $X$  si  $\xi \in X$  *c.s.* donde *c.s.* denota casi seguro. Si estamos trabajando con varias probabilidades y puede surgir alguna ambigüedad, tras *c.s.* escribiremos entre corchetes la probabilidad con respecto a la que se da el casi seguro. Todo conjunto aleatorio cerrado  $X$  no vacío *c.s.* tiene, al menos, una selección.

### 0.2.1. Esperanza de Aumann

Si al menos una selección de  $X$  es integrable, podemos definir la *esperanza de Aumann* de  $X$ , a la que denotamos por  $\mathbb{E}X$ , como el conjunto de esperanzas de todas sus selecciones integrables, es decir

$$\mathbb{E}X := \{\mathbb{E}\xi : \xi \text{ selección integrable de } X\},$$

donde  $\mathbb{E}\xi$  denota a la esperanza del vector aleatorio  $\xi$ . Aumann (1965) realizó esta misma construcción para definir la integral de una función multivaluada y Molchanov (1998) recopila y compara diferentes definiciones de esperanzas para conjuntos aleatorios.

La esperanza de Aumann de un conjunto aleatorio convexo es convexa y dado cualquier conjunto aleatorio definido sobre un espacio de probabilidad no atómico, su esperanza de Aumann también será convexa, en tal caso  $\mathbb{E}X = \mathbb{E}\text{co}(X)$  donde  $\text{co}(X)$  denota a la *envolvente convexa* de  $X$ , es decir al menor conjunto convexo que contiene a  $X$ .

La esperanza de Aumann es equivariante afín, es decir, dado un conjunto aleatorio  $X$  en  $\mathbb{R}^d$ , una matriz  $A \in \mathbb{R}^{d \times d}$  y un vector  $b \in \mathbb{R}^d$ , se cumple  $\mathbb{E}[AX + b] = A\mathbb{E}X + b$ .

Dado  $X : \Omega \rightarrow \mathcal{K}$  un conjunto aleatorio compacto no vacío,  $X$  se dice *integrablemente acotado* si  $\mathbb{E}\|X\|$  es finito, donde  $\|X\| := \sup\{\|x\| : x \in X\}$  es la variable aleatoria *magnitud* del conjunto aleatorio  $X$ . La medibilidad de la variable aleatoria  $\|X\|$  y la de  $d_H(X, Y)$  donde  $X, Y : \Omega \rightarrow \mathcal{K}$  son conjuntos aleatorios compactos no vacíos ha sido demostrada por Hiai y



Umegaki (1977). Si  $X : \Omega \longrightarrow \mathcal{K}$  es un conjunto aleatorio compacto integrablemente acotado, entonces su esperanza de Aumann  $\mathbb{E}X$  es un conjunto compacto.

Otra forma alternativa para el cálculo de la esperanza de Aumann de un conjunto aleatorio está basada en la función soporte. La *función soporte* de un conjunto  $K \subset \mathbb{R}^d$  se define para cualquier  $u \in \mathbb{R}^d$  como  $h(K, u) := \sup\{\langle x, u \rangle : x \in K\}$ , donde  $\langle \cdot, \cdot \rangle$  denota al producto escalar. La función soporte de la esperanza de Aumann de un conjunto aleatorio en cualquier dirección es la esperanza de la función soporte del conjunto aleatorio en la misma dirección, véase pág. 15 en Molchanov (1993). Es decir, si al menos una selección de  $X$  es integrable, existe su esperanza de Aumann y para cualquier  $u \in \mathbb{R}^d$  se cumple

$$\mathbb{E}h(X, u) = h(\mathbb{E}X, u). \quad (1)$$

Obsérvese que esta relación es válida para cualquier conjunto aleatorio cerrado con tal de que sea no vacío y alguna de sus selecciones sea integrable. Además, como la función soporte caracteriza los conjuntos cerrados convexos, la relación (1) determina la esperanza de Aumann de un conjunto aleatorio siempre que ésta sea convexa, por consiguiente en casos como cuando el espacio de probabilidad es no atómico o el conjunto aleatorio es convexo.

El siguiente resultado es una particularización del Teorema 4.1 en Hiai y Umegaki (1977) y ofrece una cota de la distancia de Hausdorff entre las esperanzas de Aumann de dos conjuntos aleatorios integrablemente acotados definidos sobre un mismo espacio de probabilidad.

**Teorema 0.2.1.** *Si  $X, Y : \Omega \longrightarrow \mathcal{K}$  son dos conjuntos aleatorios integrablemente acotados, entonces se cumple*

$$d_H(\mathbb{E}X, \mathbb{E}Y) \leq \mathbb{E}d_H(X, Y).$$

### 0.3. Zonoides y zonotopos

Los zonoides son una clase de cuerpos convexos, en concreto forman la menor clase cerrada para la suma de Minkowski y la distancia de Hausdorff que contiene a todos los segmentos. La Matemática y más específicamente el Análisis Convexo les ha prestado gran atención. Entre otros, Bolker (1969) y (1971) y Goodey y Weil (1993) han estudiado profundamente los zonoides, mientras que Rockafellar (1970) y Schneider (1993) han escrito sendas monografías sobre Análisis Convexo, la segunda de las cuales dedica una sección al estudio de los zonoides.

Estos nuevos cuerpos convexos son centralmente simétricos. Un conjunto  $K \subset \mathbb{R}^d$  se dice *centralmente simétrico* si existe un punto  $\theta \in K$  tal que para cualquier  $x \in K$ , el punto simétrico a  $x$  respecto de  $\theta$  pertenece a  $K$ , es decir,  $2\theta - x \in K$ .

A la envolvente convexa de un conjunto finito de puntos se le conoce como *politopo*. Los *zonotopos* son una familia especial de politopos y los *zonoides* límites de sucesiones de zonotopos.

*Definición 0.3.1.* Un *zonotopo* es una suma de Minkowski de un número finito de segmentos.

*Definición 0.3.2.* Los *zonoides* son cuerpos convexos centralmente simétricos que pueden escribirse como el límite en el sentido de Hausdorff de una sucesión de zonotopos o equivalentemente como el rango de una medida vectorial.

Es claro, a partir de estas definiciones, que todo zonotopo es un zonoide.

En  $\mathbb{R}^2$  todo cuerpo convexo centralmente simétrico es un zonoide, siendo éstos, en cambio, escasísimos en  $\mathbb{R}^3$  y dimensiones superiores.

Koshevoy y Mosler (1998) definen el zonoide de un vector aleatorio, o más específicamente el *zonoide de momentos de un vector aleatorio* y su *zonoide elevado* y los caracterizan como la esperanza de Aumann de ciertos segmentos aleatorios.

*Definición 0.3.3.* Dado un vector aleatorio  $\xi$  en  $\mathbb{R}^d$  que induce una distribución de probabilidad  $P_\xi$ , definimos su *zonoide de momentos*, denotado  $Z(P_\xi)$ , como

$$Z(P_\xi) := \left\{ \int_{\mathbb{R}^d} xg(x)dP_\xi(x) : g : \mathbb{R}^d \longrightarrow [0, 1] \text{ medible} \right\} = \mathbb{E}[0, \xi],$$

donde  $[0, \xi]$  denota al segmento que une el origen con  $\xi$  y es, por tanto, un conjunto aleatorio compacto y convexo, mientras que  $\mathbb{E}[0, \xi]$  es su esperanza de Aumann. Fijado un vector aleatorio  $\xi$ , nos referiremos a  $Z(P_\xi)$  simplemente como su *zonoide*.

El *zonoide elevado* de  $\xi$  es el *zonoide de momentos* del vector aleatorio en  $\mathbb{R}^{d+1}$  que tiene la primera componente igual a 1 y luego el vector aleatorio  $\xi$ , es decir  $(1, \xi)$ , lo denotamos por  $\tilde{Z}(P_\xi)$  y satisface,

$$\begin{aligned} \tilde{Z}(P_\xi) &:= \left\{ \left( \int_{\mathbb{R}^d} g(x)dP_\xi(x), \int_{\mathbb{R}^d} xg(x)dP_\xi(x) \right) : g : \mathbb{R}^d \longrightarrow [0, 1] \text{ medible} \right\} \\ &= \mathbb{E}[0, (1, \xi)]. \end{aligned}$$

Tanto el *zonoide* como el *zonoide elevado* de un vector aleatorio son convexos y contienen al origen de coordenadas. Si además el vector aleatorio tiene primer momento finito, entonces son *zonoides* de acuerdo con la Definición 0.3.2, por tanto son conjuntos compactos y además centralmente simétricos. El *zonoide*  $Z(P_\xi)$  es centralmente simétrico respecto del punto  $\mathbb{E}\xi/2$  y el *zonoide elevado*  $\tilde{Z}(P_\xi)$  respecto del punto  $1/2(1, \mathbb{E}\xi)$ . Por último, el *zonoide elevado* de un vector aleatorio con primer momento finito caracteriza su distribución de probabilidad.

Los *zonoides* a los que da lugar una distribución de probabilidad discreta son simples *zonotopos* y pueden ayudarnos a entender la estructura de estos nuevos cuerpos convexos. Si  $P$  es una distribución de probabilidad discreta

en  $\mathbb{R}^d$  tal que para  $x_1, \dots, x_n \in \mathbb{R}^d$  cumple  $P(\{x_i\}) = p_i > 0$  con  $\sum_{i=1}^n p_i = 1$ , entonces da lugar al siguiente zonoide y zonoide elevado,

$$Z(P) = \left\{ \sum_{i=1}^n \alpha_i x_i p_i : 0 \leq \alpha_i \leq 1, \forall i \in \{1, \dots, n\} \right\},$$

$$\tilde{Z}(P) = \left\{ \left( \sum_{i=1}^n \alpha_i p_i, \sum_{i=1}^n \alpha_i x_i p_i \right) : 0 \leq \alpha_i \leq 1, \forall i \in \{1, \dots, n\} \right\}.$$

Los zonoides elevados de las distribuciones de las proyecciones unidimensionales de un vector aleatorio son las proyecciones unidimensionales del zonoide elevado de la distribución original. A partir del zonoide elevado de una distribución  $P_\xi$  es fácil obtener los zonoides elevados de las distribuciones  $P_{\langle \xi, u \rangle}$  para cualquier  $u \in \mathbb{R}^d$ .

**Teorema 0.3.1 (Proposición 2.24 en Mosler, 2002).** *Dado  $u \in \mathbb{R}^d$  y  $P \in \mathbb{P}$ , si  $P_u$  es una probabilidad en  $\mathbb{R}$  tal que  $P_u(A) = P(\{x : \langle x, u \rangle \in A\})$  para todo  $A \in \mathcal{B}_1$ , entonces*

$$\tilde{Z}(P_u) = \tilde{Z}(P) \begin{pmatrix} 1 & 0 \\ 0 & u \end{pmatrix}.$$

## 0.4. Funciones de profundidad y regiones centrales

Dado un punto  $x \in \mathbb{R}^d$ , las funciones de profundidad le asignan su grado de centralidad con respecto a una nube de puntos o una probabilidad  $P \in \mathbb{P}$ . En los últimos años, las funciones de profundidad y las regiones centrales han recibido gran atención por parte de la comunidad estadística. Zuo y Serfling (2000a) proponen condiciones que deberían satisfacer todas las funciones de profundidad y comparan algunas profundidades clásicas. Los mismos autores Zuo y Serfling (2000c) estudian algunas propiedades de los contornos de las regiones centrales. En este punto hay que reseñar que fue Liu (1990) la primera en introducir sistemáticamente las propiedades que entendía que una

función de profundidad debería satisfacer, con el fin de estudiarlas para el caso concreto de la *profundidad simplicial*.

Según la Definición 2.1 en Zuo y Serfling (2000a), una función de profundidad estadística se define como sigue.

*Definición 0.4.1.* Una *función de profundidad estadística* es una aplicación acotada no negativa  $D(\cdot; \cdot) : \mathbb{R}^d \times \mathbb{P} \longrightarrow \mathbb{R}$  que satisface:

- i. *Invarianza afín.*  $D(Ax + b; P_{A\xi+b}) = D(x; P_\xi)$  se cumple para cualquier vector aleatorio  $\xi$  en  $\mathbb{R}^d$ , cualquier matriz no singular  $A \in \mathbb{R}^{d \times d}$  y cualquier  $b \in \mathbb{R}^d$ ;
- ii. *Maximalidad en el centro.*  $D(\theta; P) = \sup_{x \in \mathbb{R}^d} D(x; P)$  se cumple para cualquier  $P \in \mathbb{P}$  con centro en  $\theta$ ;
- iii. *Monotonía con respecto al punto más profundo.* Para cualquier  $P \in \mathbb{P}$  con punto más profundo  $\theta$  y para todo  $\alpha \in [0, 1]$ , se cumple  $D(x; P) \leq D(\theta + \alpha(x - \theta); P)$ ;
- iv. *Se anula en el infinito.*  $D(x; P) \longrightarrow 0$  cuando  $\|x\| \rightarrow +\infty$ , para cualquier  $P \in \mathbb{P}$ .

El *centro* de una distribución de probabilidad será un punto que satisface alguna definición de estimador de localización o simetría multivariante.

Toda función de profundidad tiene una familia de regiones centrales asociada. La *región central de profundidad  $\alpha$  asociada con la función de profundidad  $D(\cdot; P)$*  se denota por  $D^\alpha(P)$  y se define como,

$$D^\alpha(P) := \{x \in \mathbb{R}^d : D(x; P) \geq \alpha\}.$$

Las propiedades que una familia de regiones centrales debe satisfacer según el Teorema 3.1 en Zuo y Serfling (2000c) son:

- i. *Equivarianza afín.*  $D^\alpha(P_{A,b}) = AD^\alpha(P) + b$ , donde  $P_{A,b}$  es una probabilidad en  $\mathbb{R}^d$  tal que  $P_{A,b}(C) = P(\{x : Ax + b \in C\})$  para todo  $C \in \mathcal{B}_d$ , se deduce de la invarianza afín de la profundidad;
- ii. *Anidamiento.* Si  $\alpha \geq \beta$ , se tiene  $D^\alpha(P) \subset D^\beta(P)$ ;
- iii. *Conexión.*  $D^\alpha(P)$  es conexa, se deduce de la monotonía con respecto al punto más profundo de la profundidad;
- iv. *Compacidad.*  $D^\alpha(P)$  es compacta.

Además de estas propiedades, a efectos prácticos, es interesante imponer algunas condiciones más fuertes, por ejemplo que las regiones centrales sean *convexas*, condición más fuerte que la *conexión*. Recordamos que un conjunto  $A \subset \mathbb{R}^d$  es *convexo* si no existe ningún par de conjuntos abiertos disjuntos  $V_1, V_2 \subset \mathbb{R}^d$  que satisfagan  $A \cap V_1 \neq \emptyset$ ,  $A \cap V_2 \neq \emptyset$  y  $A \subset V_1 \cup V_2$ .

### 0.4.1. Profundidad semiespacial o de Tukey

La profundidad semiespacial o de Tukey mide el grado de centralidad de un punto con respecto a una distribución de probabilidad considerando el conjunto de semiespacios cerrados que contienen al punto. Fue introducida por Tukey (1975) para medir el grado de centralidad de un punto con respecto a un conjunto finito de puntos.

*Definición 0.4.2.* Dado  $x \in \mathbb{R}^d$ , la *profundidad semiespacial* de  $x$  con respecto a  $P \in \mathbb{P}$ , a la que denotamos por  $\text{HD}(x; P)$ , se define como,

$$\text{HD}(x; P) := \inf\{P(H) : H \text{ semiespacio cerrado, } x \in H\}.$$

Las regiones centrales asociadas con la profundidad semiespacial se definen para cualquier  $\alpha \in (0, 1]$  como el conjunto de puntos cuya profundidad semiespacial es al menos  $\alpha$ , es decir

$$\text{HD}^\alpha(P) := \{x \in \mathbb{R}^d : \text{HD}(x; P) \geq \alpha\}, \quad \alpha \in (0, 1].$$

### 0.4.2. Regiones centrales del zonoide

Las regiones centrales del zonoide fueron introducidas por Koshevoy y Mosler (1997b) y están íntimamente relacionadas con el concepto del zonoide elevado.

*Definición 0.4.3.* Para cualquier  $\alpha \in (0, 1]$ , las *regiones centrales del zonoide* se definen como

$$\text{ZD}^\alpha(P) := \left\{ \int_{\mathbb{R}^d} xg(x)dP(x) : \right. \\ \left. g : \mathbb{R}^d \longrightarrow [0, \alpha^{-1}], \text{ medible y } \int_{\mathbb{R}^d} g(x)dP(x) = 1 \right\},$$

La función de profundidad inducida por una familia de regiones centrales es aquella que asigna a cada  $x \in \mathbb{R}^d$  el supremo de los índices  $\alpha$  para los que  $x$  pertenece a la región central de grado  $\alpha$ . Las regiones centrales del zonoide de la Definición 0.4.3 inducen la función de profundidad que aparece a continuación,

$$\text{ZD}(x; P) := \sup\{\alpha \in (0, 1] : x \in \text{ZD}^\alpha(P)\}.$$

Las regiones centrales del zonoide se pueden obtener de un modo muy sencillo a partir del zonoide elevado.

**Teorema 0.4.1 (Proposición 2.2 en Koshevoy y Mosler, 1997b).** *Da-da una probabilidad  $P \in \mathbb{P}$ , si para cualquier  $C \subset \mathbb{R}^{d+1}$ ,  $\text{proj}_\alpha(C)$  es la proyección sobre las  $d$  últimas coordenadas de la intersección de  $C$  con el hiperplano  $\{x \in \mathbb{R}^{d+1} : x_1 = \alpha\}$ , se cumple la siguiente relación entre el zonoide elevado y las regiones centrales del zonoide,*

$$\text{ZD}^\alpha(P) = \alpha^{-1}\text{proj}_\alpha(\tilde{Z}(P)).$$

## 0.5. Órdenes estocásticos

Antes de introducir la Teoría de los Órdenes Estocásticos, debemos establecer lo que es una *relación de orden (parcial)*. Una relación binaria ' $\leq$ ' definida en un conjunto arbitrario  $\mathcal{S}$  se dice que es un *orden* si satisface las tres condiciones siguientes:

- i. *Reflexividad*:  $x \leq x$  para todo  $x \in \mathcal{S}$ ;
- ii. *Transitividad*: si  $x \leq y$  e  $y \leq z$ , entonces  $x \leq z$ ;
- iii. *Antisimetría*: si  $x \leq y$  e  $y \leq x$ , entonces  $x = y$ .

Una relación reflexiva y transitiva se dice que es un *preorden*.

Los *órdenes estocásticos* son *relaciones de orden parcial* entre distribuciones de probabilidad de variables aleatorias, vectores aleatorios o, en general, cualquier tipo de elementos aleatorios. Se aplican en numerosos campos de investigación científica, como puedan ser Economía, Ciencias Actuariales, Teoría de Colas, Teoría de la Decisión, Teoría de la Fiabilidad, Epidemiología, etc.

A pesar de que los órdenes estocásticos ordenan distribuciones de probabilidad, lo normal en la literatura al respecto es hablar de desigualdad entre elementos aleatorios, y no entre distribuciones de probabilidad. Así, dados  $\xi, \eta$  dos vectores aleatorios con distribuciones de probabilidad  $P_\xi, P_\eta$ , si estas probabilidades están ordenadas respecto de algún orden ' $\leq$ ', diremos que  $\xi$  es *menor que*  $\eta$  respecto del orden ' $\leq$ ' y escribiremos  $\xi \leq \eta$ . Es importante reseñar que la extensión de cualquier relación de orden entre distribuciones de probabilidad a vectores aleatorios es, en general, sólo un *preorden* entre dichos vectores, es decir una relación reflexiva y transitiva, pero no antisimétrica.

Como referencias básicas en el campo de los órdenes estocásticos cabe destacar las monografías de Shaked y Shanthikumar (1994) y Müller y Stoyan (2002).



### 0.5.1. Órdenes estocásticos integrales

Muchos órdenes estocásticos se definen a partir de relaciones de desigualdad de esperanzas de transformaciones de variables o vectores aleatorios a través de las funciones de cierta familia. Concretamente, llamamos *orden estocástico integral* a aquel para el que existe una familia de funciones  $\mathcal{F}$  tal que dos elementos aleatorios  $\xi, \eta$  están ordenados respecto del orden ' $\leq_{\mathcal{F}}$ ', relación que denotamos por  $\xi \leq_{\mathcal{F}} \eta$ , si y sólo si  $\mathbb{E}f(\xi) \leq \mathbb{E}f(\eta)$  para toda función  $f \in \mathcal{F}$  para la que existan dichas esperanzas. A la familia  $\mathcal{F}$  se le denomina familia de *generadores* del orden ' $\leq_{\mathcal{F}}$ '.

Si la familia  $\mathcal{F}$  está compuesta por funciones convexas, decimos que el orden estocástico integral inducido por  $\mathcal{F}$  pertenece a la familia de los *órdenes estocásticos convexos*.

A continuación citamos cinco de los órdenes estocásticos de uso más común entre vectores aleatorios en  $\mathbb{R}^d$ , y los describiremos como órdenes integrales:

- El *orden estocástico habitual*:  $\xi \leq_{\text{st}} \eta$  si y sólo si  $\mathbb{E}f(\xi) \leq \mathbb{E}f(\eta)$  para toda función  $f$  creciente en cada una de sus componentes. Si  $\xi, \eta$  son variables aleatorias, esto equivale a la ordenación puntual de sus funciones de distribución  $F_{\xi}(\cdot) \geq F_{\eta}(\cdot)$ ;
- El *orden convexo*:  $\xi \leq_{\text{cx}} \eta$  si y sólo si  $\mathbb{E}f(\xi) \leq \mathbb{E}f(\eta)$  para toda función  $f$  convexa;
- El *orden creciente convexo*:  $\xi \leq_{\text{icx}} \eta$  si y sólo si  $\mathbb{E}f(\xi) \leq \mathbb{E}f(\eta)$  para toda función  $f$  creciente en cada una de sus componentes y convexa;
- El *orden lineal convexo*:  $\xi \leq_{\text{lcx}} \eta$  si y sólo si  $\langle \xi, u \rangle \leq_{\text{cx}} \langle \eta, u \rangle$  para todo  $u \in \mathbb{R}^d$ . La familia de generadores será  $\mathcal{F}_{\text{lcx}} = \{c \circ l : l : \mathbb{R}^d \rightarrow \mathbb{R} \text{ lineal, } c : \mathbb{R} \rightarrow \mathbb{R} \text{ convexa}\}$ ;
- El *orden creciente positivo lineal convexo*:  $\xi \leq_{\text{iplcx}} \eta$  si y sólo si  $\langle \xi, u \rangle \leq_{\text{icx}} \langle \eta, u \rangle$  para todo  $u \in \mathbb{R}_+^d$ . La familia de generadores será  $\mathcal{F}_{\text{iplcx}} = \{c \circ l : l : \mathbb{R}^d \rightarrow \mathbb{R} \text{ lineal creciente, } c : \mathbb{R} \rightarrow \mathbb{R} \text{ creciente convexa}\}$ .

Recordamos que una función  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  es *convexa* si su epigrafo es convexo o equivalentemente se tiene que para cualesquiera  $x, y \in \mathbb{R}^d$  y  $\lambda \in (0, 1)$  se cumple  $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ . Una función se dice *estrictamente convexa* si la desigualdad anterior es estricta.

Obsérvese que una función  $l : \mathbb{R}^d \rightarrow \mathbb{R}$  es lineal creciente si es de la forma  $l(x) = \langle u, x \rangle$  con  $u \in \mathbb{R}_+^d$  para todo  $x \in \mathbb{R}^d$ , donde  $\mathbb{R}_+^d$  es el primer cuadrante cerrado.

Müller (1997b) estudia en profundidad las propiedades de los órdenes integrales deduciéndolas de las de los conjuntos de funciones que los generan. En dicho trabajo se resuelve el problema de la caracterización de conjuntos generadores maximales y se razona que no tiene sentido la búsqueda de conjuntos generadores minimales. En este punto nosotros sólo señalaremos que cuanto mayor sea la familia de los generadores, más restricciones se imponen y, por lo tanto, el orden convexo es más fuerte que el lineal convexo y el orden estocástico habitual que el orden creciente convexo y que el orden creciente positivo lineal convexo.

*Nota 0.5.1.* Para variables aleatorias el orden lineal convexo y el orden convexo coinciden, y lo mismo sucede con el orden creciente convexo y el orden creciente positivo lineal convexo.

*Nota 0.5.2.* La comparación de dos variables aleatorias según un orden estocástico suele ser muy sencilla. Tal comparación es bastante más difícil en el caso de los vectores aleatorios. A ello y al hecho de que en muchos casos son las mezclas lineales las únicas que nos interesan, se debe el exhaustivo estudio de los *órdenes lineales*. Un orden estocástico entre vectores aleatorios es lineal cuando se verifica si y sólo si se verifica la correspondiente relación de orden entre cada par de variables aleatorias obtenidas tras aplicar la misma transformación lineal de  $\mathbb{R}^d$  en  $\mathbb{R}$  sobre ambos vectores.

*Nota 0.5.3.* Obsérvese que la definición de un orden generado por la familia de composiciones de funciones lineales con funciones crecientes convexas no tiene sentido ya que, como razonó Scarsini (1998), daría lugar al orden lineal convexo. Para generalizar el orden creciente convexo a través de transfor-

maciones lineales tenemos que considerar sólo las transformaciones lineales crecientes, también llamadas positivas.

*Nota 0.5.4.* La definición de órdenes convexos o según la variabilidad para variables aleatorias que no tienen primer momento finito conlleva serias dificultades y además no tiene sentido ya que si una variable aleatoria  $\xi$  no tiene primer momento finito y  $f$  es convexa, entonces  $\mathbb{E}f(\xi)$  es finito si y sólo si  $f$  es constante, como razonan Müller y Stoyan (2002), pág. 15. Por este motivo, al trabajar con órdenes convexos es necesario hacer la suposición de que trabajamos con vectores aleatorios con primer momento finito.

A continuación repasamos algunas propiedades de los órdenes estocásticos más utilizados.

### Orden estocástico habitual

**Teorema 0.5.1 (Teorema 2.1 en Bhattacharjee y Sethuraman, 1990).**

*Si dos variables aleatorias satisfacen  $\xi \leq_{st} \eta$  y para algún  $\alpha \neq 0$  se verifica  $\mathbb{E}\xi^\alpha = \mathbb{E}\eta^\alpha$ , entonces  $\xi$  y  $\eta$  tienen la misma distribución.*

### Órdenes estocásticos convexos

**Teorema 0.5.2 (Teorema 1.5.3 en Müller y Stoyan, 2002).** *Dos variables aleatorias  $\xi, \eta$  están ordenadas según el orden convexo como  $\xi \leq_{cx} \eta$  si y sólo si lo están respecto del orden creciente convexo,  $\xi \leq_{icx} \eta$  y tienen la misma esperanza,  $\mathbb{E}\xi = \mathbb{E}\eta$ .*

Puesto que el orden lineal convexo aplicado sobre vectores aleatorios implica la ordenación de cada par de componentes marginales respecto del orden convexo, el orden lineal convexo implica la igualdad de las esperanzas.

**Corolario 0.5.3.** *Si dos vectores aleatorios  $\xi, \eta$  están ordenados según el orden lineal convexo  $\xi \leq_{lcx} \eta$ , entonces tienen la misma esperanza,  $\mathbb{E}\xi = \mathbb{E}\eta$ .*

En el siguiente resultado aparece por primera vez una notación que se utilizará con frecuencia a lo largo de la memoria. Dado  $x \in \mathbb{R}$ , denotamos

por  $x_+$  a la parte positiva de  $x$ , es decir  $x_+ := \max\{0, x\}$ , mientras que  $x_-$  denota a la parte negativa de  $x$ , es decir  $x_- := \max\{0, -x\}$ , se cumple por tanto que  $x = x_+ - x_-$ .

**Teorema 0.5.4 (Teorema 1.5.7 en Müller y Stoyan, 2002).** *Dos variables aleatorias  $\xi, \eta$  están ordenadas según el orden creciente convexo como  $\xi \leq_{\text{icx}} \eta$  si y sólo si*

$$\mathbb{E}(\xi - t)_+ \leq \mathbb{E}(\eta - t)_+ \text{ para todo } t \in \mathbb{R}.$$

**Teorema 0.5.5 (Teorema 1 en Scarsini, 1998).** *Si las componentes marginales de dos vectores aleatorios satisfacen  $\mathbb{E}\xi_i^2 = \mathbb{E}\eta_i^2$  para todo  $i \in \{1, \dots, d\}$  y  $\xi \leq_{\text{icx}} \eta$ , entonces  $\xi$  y  $\eta$  tienen la misma distribución.*

**Teorema 0.5.6 (Teorema 1.5.14 en Müller y Stoyan, 2002).** *Para cada par de variables aleatorias  $\xi, \eta$  que satisfacen  $\xi \leq_{\text{icx}} \eta$ , existe una tercera variable aleatoria  $\zeta$  tal que*

$$\xi \leq_{\text{st}} \zeta \leq_{\text{cx}} \eta.$$

**Teorema 0.5.7 (Teorema 3.4.4 en Müller y Stoyan, 2002).** *Si  $\xi$  y  $\eta$  son dos vectores aleatorios en  $\mathbb{R}^d$  cuyas componentes marginales son independientes y satisfacen  $\xi_i \leq_{\text{icx}} \eta_i$  para todo  $i \in \{1, \dots, d\}$ , entonces  $\xi \leq_{\text{icx}} \eta$ .*

## 0.5.2. Órdenes estocásticos por inclusión de conjuntos

Koshevoy y Mosler (1998) definieron el *orden del zonoide* denotado ' $\leq_Z$ ' mediante una relación de contenido de zonoide, es decir,

$$\xi \leq_Z \eta \text{ si y sólo si } Z(P_\xi) \subset Z(P_\eta).$$

Es claro que el orden del zonoide es sólo un preorden entre las distribuciones de probabilidad de dos vectores aleatorios. No satisface la propiedad de antisimetría porque el zonoide de un vector aleatorio no caracteriza su distribución. El orden del zonoide es mucho menos restrictivo que los demás

órdenes convexos, y por tanto muchos vectores aleatorios no comparables respecto de otros órdenes lo son respecto del orden del zonoide. Dall’Aglío y Scarsini (2003) han utilizado el orden del zonoide para comparar dos vectores aleatorios según el grado de dependencia lineal entre sus componentes.

El siguiente resultado puede obtenerse fácilmente.

**Lema 0.5.8.** *El orden del zonoide es un orden integral y sus generadores son de la forma,*

$$\mathcal{F}_Z = \{l_+ : l : \mathbb{R}^d \longrightarrow \mathbb{R} \text{ lineal}\}, \text{ donde } l_+(x) = \langle x, u \rangle_+ \text{ para algún } u \in \mathbb{R}^d.$$

En efecto, dado un vector aleatorio  $\xi$  y  $u \in \mathbb{R}^d$  no nulo, la función soporte del zonoide en la dirección de  $u$  será,

$$h(Z(P_\xi), u) = h(\mathbb{E}[0, \xi], u) = \mathbb{E}h([0, \xi], u) = \mathbb{E} \max\{\langle \xi, u \rangle, 0\} = \mathbb{E}\langle \xi, u \rangle_+.$$

La relación de inclusión para conjuntos convexos equivale a la de menor o igual de las funciones soporte y a partir de ahí el resultado es claro.

*Nota 0.5.5.* Podemos considerar al orden del zonoide como un orden convexo, ya que las funciones de  $\mathcal{F}_Z$  son convexas.

Para concluir con la relación de resultados básicos que se utilizarán a lo largo de esta memoria, se citará la siguiente relación entre órdenes estocásticos y zonoides. El orden lineal convexo puede caracterizarse también mediante una relación de inclusión de zonoides, en este caso a través de una relación de contenido de zonoides elevados.

**Teorema 0.5.9 (Teorema 5.2 en Koshevoy y Mosler, 1998).** *Dados  $\xi, \eta$  dos vectores aleatorios,*

$$\xi \leq_{\text{lcx}} \eta \text{ si y sólo si } \tilde{Z}(P_\xi) \subset \tilde{Z}(P_\eta).$$



# Capítulo 1

## Orden estocástico según la variabilidad

El orden estocástico habitual es una generalización del orden existente en  $(\mathbb{R}, \leq)$  y por lo tanto compara vectores aleatorios según la relación *menor o igual* en cada componente. Los órdenes convexos, que son los órdenes de los que nos ocuparemos, comparan vectores aleatorios según su *variabilidad*. La variabilidad de un elemento aleatorio describe el riesgo ante un determinado resultado incierto. Así, si dos variables aleatorias con la misma esperanza describen el resultado de dos inversiones diferentes, cualquiera que desee evitar riesgos invertirá en aquella descrita por la variable aleatoria con menor variabilidad. Es por esto que los órdenes convexos se utilizan en la toma de decisiones bajo riesgo como se realiza en las Ciencias Actuariales. También se usan en las Ciencias Sociales para comparar la desigualdad asociada a diversas variables, tales como riqueza, salario, etc.

En este capítulo definimos un nuevo orden estocástico convexo, el *orden simétrico* (Casco y Molchanov, 2003) del que damos las propiedades esenciales y lo relacionamos con otros órdenes estocásticos. Además, presentamos un nuevo cuerpo convexo, la envolvente de zonoides, que determina completamente la distribución de vectores aleatorios con primer momento finito. A partir de ciertas caracterizaciones de varios órdenes estocásticos convexos, en

concreto del orden simétrico, del orden lineal convexo y del orden creciente positivo lineal convexo en términos de inclusiones de conjuntos, los generalizamos a relaciones entre conjuntos aleatorios y finalmente definimos una métrica que, bajo ciertas condiciones, es *ideal para el esquema de la suma de vectores aleatorios* (Casco, 2003).

## 1.1. Orden simétrico

En esta sección introduciremos un nuevo orden estocástico integral convexo para conjuntos aleatorios, el orden *simétrico*. Estudiaremos sus propiedades y relaciones con otros órdenes estocásticos así como condiciones de igualdad estocástica para vectores aleatorios ordenados respecto del orden simétrico. Además, introduciremos un nuevo cuerpo convexo que sirve para caracterizar el orden simétrico mediante la relación de contenido.

A lo largo del presente capítulo hacemos dos suposiciones. En primer lugar, asumimos que todos los vectores aleatorios con los que trabajamos tienen primer momento finito, ya que como se razonó en la Nota 0.5.4, los órdenes estocásticos convexos sólo tienen sentido en este caso. Por otra parte, suponemos que todos los elementos aleatorios, tanto vectores como conjuntos están definidos sobre espacios de probabilidad no atómicos. Así, las esperanzas de Aumann que obtenemos serán siempre conjuntos convexos. Para evitar esta segunda suposición bastaría con considerar las esperanzas de Aumann de las envolventes convexas de los conjuntos aleatorios.

Dado un vector aleatorio  $\xi$  y  $r > 0$ , consideramos el conjunto aleatorio compacto  $\{\xi\} \cup B_r$  donde recordamos que  $B_r$  es la bola cerrada centrada en el origen de radio  $r$  y comparamos la esperanza de Aumann de este tipo de conjuntos aleatorios según la relación de inclusión.

*Definición 1.1.1.* Sean  $\xi$  y  $\eta$  dos vectores aleatorios en  $\mathbb{R}^d$ . Decimos que  $\xi$  es más pequeño que  $\eta$  respecto del *orden simétrico*, relación que denotamos por  $\xi \leq_{\text{sym}} \eta$  si

$$\mathbb{E}(\{\xi\} \cup B_r) \subset \mathbb{E}(\{\eta\} \cup B_r)$$

para todo  $r > 0$ .



En el Teorema 1.1.5 se demostrará que la relación que acabamos de definir es, en efecto, un orden estocástico.

### 1.1.1. Propiedades preliminares y caracterización

En el caso unidimensional, cuando trabajamos con variables aleatorias, el orden simétrico es equivalente al orden creciente convexo de las partes positiva y negativa de las variables aleatorias a comparar. Es ésta la razón por la que denominamos a la relación ' $\leq_{\text{sym}}$ ' orden simétrico, el que las partes positiva y negativa de una variable aleatoria juegan un papel simétrico.

**Teorema 1.1.1.** *Dadas dos variables aleatorias  $\xi, \eta$ , se cumple  $\xi \leq_{\text{sym}} \eta$  si y sólo si*

$$\xi_+ \leq_{\text{icx}} \eta_+ \quad \text{y} \quad \xi_- \leq_{\text{icx}} \eta_- .$$

*Demostración.* Dada una variable aleatoria  $\xi$ , se cumple

$$\begin{aligned} \mathbb{E}(\{\xi\} \cup B_r) &= \mathbb{E}[\text{mín}\{\xi, -r\}, \text{máx}\{\xi, r\}] \\ &= [-r - \mathbb{E}(-\xi - r)_+, r + \mathbb{E}(\xi - r)_+] . \end{aligned}$$

Entonces  $\xi \leq_{\text{sym}} \eta$  si y sólo si  $\mathbb{E}(\xi - r)_+ \leq \mathbb{E}(\eta - r)_+$  y  $\mathbb{E}(-\xi - r)_+ \leq \mathbb{E}(-\eta - r)_+$  para todo  $r > 0$ . Esto se cumple si y sólo si  $\mathbb{E}(\xi_+ - r)_+ \leq \mathbb{E}(\eta_+ - r)_+$  y  $\mathbb{E}(\xi_- - r)_+ \leq \mathbb{E}(\eta_- - r)_+$  para todo  $r \in \mathbb{R}$ , lo que, por el Teorema 0.5.4, es equivalente a  $\xi_+ \leq_{\text{icx}} \eta_+$  y  $\xi_- \leq_{\text{icx}} \eta_-$ .  $\square$

La relación ' $\leq_{\text{sym}}$ ' es de hecho un orden estocástico *lineal*.

**Teorema 1.1.2.** *Dados dos vectores aleatorios cualesquiera,  $\xi \leq_{\text{sym}} \eta$  si y sólo si  $\langle \xi, u \rangle \leq_{\text{sym}} \langle \eta, u \rangle$  para todo  $u \in \mathbb{R}^d$ .*

*Demostración.* Obsérvese que para cualquier  $u \in \mathbb{R}^d$  y  $r > 0$  se tiene que

$$\begin{aligned} h(\mathbb{E}(\{\xi\} \cup B_r), u) &= \mathbb{E}h(\{\xi\} \cup B_r, u) \\ &= \mathbb{E} \text{máx}\{\langle \xi, u \rangle, r\|u\|\} = r\|u\| + \mathbb{E}(\langle \xi, u \rangle - r\|u\|)_+ . \end{aligned}$$

Por lo tanto, razonando como en el Teorema 1.1.1, la relación  $\xi \leq_{\text{sym}} \eta$  se satisface si y sólo si  $\langle \xi, u \rangle_+ \leq_{\text{icx}} \langle \eta, u \rangle_+$  es cierto para todo  $u \in \mathbb{R}^d$ . Aplicando de nuevo el mismo resultado, esto es equivalente a  $\langle \xi, u \rangle \leq_{\text{sym}} \langle \eta, u \rangle$ .  $\square$

**Corolario 1.1.3.** *Los tres enunciados siguientes son equivalentes:*

- i.  $\xi \leq_{\text{sym}} \eta$ ;
- ii.  $\langle \xi, u \rangle_+ \leq_{\text{icx}} \langle \eta, u \rangle_+$  para todo  $u \in \mathbb{R}^d$ ;
- iii.  $A\xi \leq_{\text{sym}} A\eta$  para cualquier matriz  $A \in \mathbb{R}^{d \times d}$ .

La ordenación de dos vectores aleatorios según el orden simétrico, implica ordenación de cada par de componentes.

**Corolario 1.1.4.** *Si se satisface  $\xi \leq_{\text{sym}} \eta$  para dos vectores aleatorios  $\xi = (\xi_1, \dots, \xi_d)$  y  $\eta = (\eta_1, \dots, \eta_d)$ , entonces  $\xi_i \leq_{\text{sym}} \eta_i$  para todo  $i = 1, \dots, d$ .*

El recíproco del Corolario 1.1.4 no se cumple, como se puede ver en el siguiente ejemplo.

*Ejemplo 1.1.1.* Sea  $\xi$  un vector aleatorio degenerado en  $(1, -1)$  y  $\eta$  otro vector aleatorio, también degenerado, en su caso en  $(1, -2)$ . Es claro que  $\xi_1 \leq_{\text{sym}} \eta_1$  y  $\xi_2 \leq_{\text{sym}} \eta_2$ , pero no se cumple que  $\xi$  sea menor que  $\eta$  respecto del orden simétrico. Si tomamos  $(2, 1) \in \mathbb{R}^2$ , entonces  $\langle \xi, (2, 1) \rangle$  es degenerado en 1 y  $\langle \eta, (2, 1) \rangle$  es degenerado en 0 con lo que  $\langle \xi, (2, 1) \rangle$  no es menor que  $\langle \eta, (2, 1) \rangle$  respecto del orden simétrico.

Para finalizar con el tratamiento preliminar de la relación ' $\leq_{\text{sym}}$ ' se demostrará que el orden simétrico es, en efecto, una relación de orden parcial entre distribuciones de probabilidad de vectores aleatorios.

**Teorema 1.1.5.** *La relación ' $\leq_{\text{sym}}$ ' es un orden parcial cuando la consideramos como una relación entre distribuciones de probabilidad de vectores aleatorios.*

*Demostración.* La reflexividad y la transitividad son obvias, sólo falta probar la antisimetría. Primero la demostraremos para variables aleatorias y luego para vectores en  $\mathbb{R}^d$  con  $d > 1$ .

Dadas dos variables aleatorias  $\xi, \eta$ , si  $\xi \leq_{\text{sym}} \eta$  y  $\eta \leq_{\text{sym}} \xi$ , por el Teorema 1.1.1 tenemos que  $\xi_+ \leq_{\text{icx}} \eta_+$ ,  $\xi_- \leq_{\text{icx}} \eta_-$ ,  $\eta_+ \leq_{\text{icx}} \xi_+$  y  $\eta_- \leq_{\text{icx}} \xi_-$  y por la antisimetría del orden creciente convexo se cumple que  $\xi_+$  y  $\eta_+$  siguen la misma distribución y lo mismo sucede para  $\xi_-$  y  $\eta_-$ . Por lo tanto  $\xi$  y  $\eta$  siguen la misma distribución.

Sean ahora  $\xi, \eta$  dos vectores aleatorios en  $\mathbb{R}^d$  tales que  $\xi \leq_{\text{sym}} \eta$  y  $\eta \leq_{\text{sym}} \xi$  y sea  $u \in \mathbb{R}^d$ . Por la linealidad de la relación ' $\leq_{\text{sym}}$ ' que se demuestra en el Teorema 1.1.2, tenemos que se satisface  $\langle \xi, u \rangle \leq_{\text{sym}} \langle \eta, u \rangle$  y la relación recíproca,  $\langle \eta, u \rangle \leq_{\text{sym}} \langle \xi, u \rangle$ . Ahora, por la primera parte de este resultado,  $\langle \xi, u \rangle$  y  $\langle \eta, u \rangle$  siguen la misma distribución y como esto sucede para cualquier  $u \in \mathbb{R}^d$ , se cumple que  $\xi$  y  $\eta$  siguen la misma distribución.  $\square$

### 1.1.2. El orden simétrico por inclusión de conjuntos

Como probaron Koshevoy y Mosler (1998), el orden lineal convexo puede caracterizarse mediante la inclusión de zonoides elevados. Veamos qué relación tiene el orden simétrico con el concepto del zonoide elevado.

**Proposición 1.1.6.** *Sean  $\xi$  y  $\eta$  dos vectores aleatorios, entonces  $\xi \leq_{\text{sym}} \eta$  si y sólo si*

$$h(\tilde{Z}(P_\xi), (-r, u)) \leq h(\tilde{Z}(P_\eta), (-r, u)) \quad \text{para todo } r \geq 0 \text{ y } u \in \mathbb{R}^d.$$

*Demostración.* El resultado se sigue directamente de

$$\begin{aligned} h(\mathbb{E}(\{\xi\} \cup B_r), u) &= r\|u\| + \mathbb{E} \max\{\langle \xi, u \rangle - r\|u\|, 0\} \\ &= r\|u\| + \mathbb{E} h([0, (1, \xi)], (-r\|u\|, u)) \\ &= r\|u\| + h(\tilde{Z}(P_\xi), (-r\|u\|, u)), \quad \text{para todo } u \in \mathbb{R}^d, r > 0. \end{aligned}$$

Como  $\xi$  tiene primer momento finito, el conjunto aleatorio  $\{\xi\} \cup B_r$  es integrablemente acotado y por lo tanto su esperanza de Aumann es un conjunto

compacto, al igual que el zonoide elevado de  $\xi$ . Finalmente, la función soporte  $h$  de un conjunto compacto es una función real continua y por tanto la relación deseada también se cumple para  $r = 0$ .  $\square$

El orden simétrico puede caracterizarse a partir de la relación de contenido de ciertos conjuntos.

*Definición 1.1.2.* Dado  $\xi$  un vector aleatorio definimos su *envolvente de zonoides*, denotada por  $W(P_\xi)$ , como el siguiente conjunto convexo,

$$W(P_\xi) := \text{co}(\tilde{Z}(P_\xi) \cup \{1\} \times Z(P_\xi)).$$

Si  $\xi$  tiene primer momento finito, de la compacidad del zonoide y de la del zonoide elevado se deriva la de  $W(P_\xi)$  que será, por tanto, un cuerpo convexo.

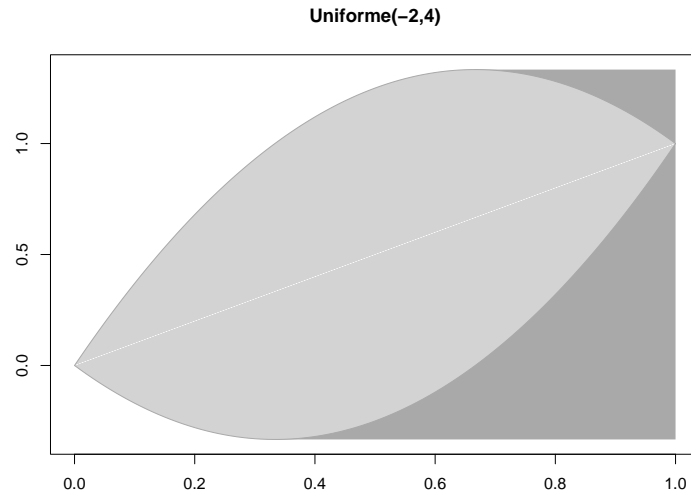


Figura 1.1: La región de color gris claro es el zonoide elevado de una variable uniforme en  $(-2, 4)$  y toda la región coloreada su envolvente de zonoides.

**Corolario 1.1.7.** *El orden simétrico queda caracterizado por la relación de contenido de las envolventes de zonoides.*

*Demostración.* Es inmediato que

$$h(W(P_\xi), (-r, u)) = h(\tilde{Z}(P_\xi), (-r, u)) \quad (1.1)$$

para todo  $r \geq 0$  y  $u \in \mathbb{R}^d$ . Finalmente, por la Proposición 1.1.6, la relación de contenido de envolventes de zonoides caracteriza el orden simétrico.  $\square$

Sobre la base del Corolario 1.1.7 y merced a la antisimetría del orden simétrico que se prueba en el Teorema 1.1.5, obtenemos el siguiente resultado.

**Corolario 1.1.8.** *Dado un vector aleatorio  $\xi$  con primer momento finito, su envolvente de zonoides,  $W(P_\xi)$ , caracteriza su distribución.*

Al igual que Koshevoy y Mosler (1997b) definen regiones centrales inspiradas en el zonoide elevado, nosotros también podemos definir una familia de regiones centrales inspiradas en la envolvente de zonoides.

*Definición 1.1.3.* Dado  $\alpha \in (0, 1]$  y una distribución de probabilidad  $P \in \mathbb{P}$ , las *regiones centrales de la envolvente de zonoides* se definen como,

$$WD^\alpha(P) := \alpha^{-1} \text{proj}_\alpha(W(P)),$$

donde, al igual que en el Teorema 0.4.1, para cualquier  $C \subset \mathbb{R}^{d+1}$ ,  $\text{proj}_\alpha(C)$  es la proyección sobre las  $d$  últimas coordenadas de la intersección de  $C$  con el hiperplano  $\{x \in \mathbb{R}^{d+1} : x_1 = \alpha\}$ .

### 1.1.3. Relación con otros órdenes estocásticos

El siguiente resultado sitúa el orden simétrico dentro de los órdenes estocásticos convexos.

**Teorema 1.1.9.** *Dados dos vectores aleatorios  $\xi$  y  $\eta$ , se cumple la siguiente cadena de implicaciones:*

$$\xi \leq_{\text{cx}} \eta \Rightarrow \xi \leq_{\text{lcx}} \eta \Rightarrow \xi \leq_{\text{sym}} \eta \Rightarrow \xi \leq_Z \eta.$$

*Demostración.* La primera implicación de la izquierda es inmediata, pero nos sirve para completar el esquema. Fijamos un  $u \in \mathbb{R}^d$  arbitrario. Si  $\xi \leq_{\text{lcx}} \eta$ , entonces  $\langle \xi, u \rangle \leq_{\text{cx}} \langle \eta, u \rangle$  lo que implica  $\langle \xi, u \rangle \leq_{\text{icx}} \langle \eta, u \rangle$  y por lo tanto  $\langle \xi, u \rangle_+ \leq_{\text{icx}} \langle \eta, u \rangle_+$ . Este último paso es cierto porque si  $f$  es una función creciente convexa, entonces  $f \circ \text{id}_+$  también será creciente convexa, donde  $\text{id}$  es la función identidad. Como la relación se cumple para cualquier  $u \in \mathbb{R}^d$ , de acuerdo con el apartado *ii.* del Corolario 1.1.3, esto es equivalente a  $\xi \leq_{\text{sym}} \eta$ .

Si  $\xi \leq_{\text{sym}} \eta$ , entonces  $\mathbb{E}(\{\xi\} \cup B_r) \subset \mathbb{E}(\{\eta\} \cup B_r)$  para todo  $r > 0$  y por tanto,

$$Z(P_\xi) = \mathbb{E}(\{\xi\} \cup \{0\}) \subset \mathbb{E}(\{\eta\} \cup \{0\}) = Z(P_\eta),$$

o equivalentemente  $\xi \leq_Z \eta$ . □

**Teorema 1.1.10.** *Dados dos vectores aleatorios  $\xi$  y  $\eta$ , se cumple  $\xi \leq_{\text{lcx}} \eta$  si y sólo si  $\xi \leq_{\text{sym}} \eta$  y  $\mathbb{E}\xi = \mathbb{E}\eta$ .*

*Demostración.* Por el Teorema 1.1.9, la relación  $\xi \leq_{\text{lcx}} \eta$  implica  $\xi \leq_{\text{sym}} \eta$ . Además, si dos vectores aleatorios están ordenados según el orden lineal convexo, entonces por el Corolario 0.5.3, sus esperanzas coinciden.

Para demostrar la implicación recíproca, consideremos dos vectores aleatorios  $\xi$  y  $\eta$  tales que  $\xi \leq_{\text{sym}} \eta$  y  $\mathbb{E}\xi = \mathbb{E}\eta$ .

Como el zonoide elevado  $\tilde{Z}(P_\xi)$  es centralmente simétrico respecto del punto  $1/2(1, \mathbb{E}\xi)$ , de la Proposición 1.1.6 se deduce

$$\begin{aligned} h(\tilde{Z}(P_\xi), (r, u)) &= h(\tilde{Z}(P_\xi), (-r, -u)) - \langle (1, \mathbb{E}\xi), (-r, -u) \rangle \\ &\leq h(\tilde{Z}(P_\eta), (-r, -u)) - \langle (1, \mathbb{E}\eta), (-r, -u) \rangle \\ &= h(\tilde{Z}(P_\eta), (r, u)) \end{aligned}$$

para cualesquiera  $r > 0$  y  $u \in \mathbb{R}^d$ . Por lo tanto  $h(\tilde{Z}(P_\xi), \cdot) \leq h(\tilde{Z}(P_\eta), \cdot)$ , con lo que el zonoide elevado de  $\xi$  es un subconjunto del zonoide elevado de  $\eta$  y finalmente, según el Teorema 0.5.9, se tiene  $\xi \leq_{\text{lcx}} \eta$ . □

Como ya se ha comentado en la Nota 0.5.1, los órdenes lineal convexo y convexo son equivalentes sobre variables aleatorias, por lo tanto los resultados sobre el orden lineal convexo se cumplen para el orden convexo entre variables aleatorias.

### 1.1.4. La relación ‘ $\leq_{\text{sym}}$ ’ como orden estocástico integral

El orden simétrico para variables aleatorias puede considerarse como un orden estocástico integral. Sus generadores son la familia  $\mathcal{F}_{\text{sym}}^1$  de funciones convexas  $f : \mathbb{R} \rightarrow \mathbb{R}$  que alcanzan su mínimo en cero. Los generadores del orden simétrico entre vectores aleatorios son la familia de funciones que se obtienen como composición de  $f \in \mathcal{F}_{\text{sym}}^1$  y una función lineal de  $\mathbb{R}^d$  en  $\mathbb{R}$ .

**Teorema 1.1.11.** *La familia*

$$\mathcal{F}_{\text{sym}}^1 = \{f : \mathbb{R} \rightarrow \mathbb{R} \text{ convexa y } f(0) \leq f(x) \text{ para todo } x \in \mathbb{R}\}$$

*genera el orden simétrico como orden estocástico integral entre variables aleatorias y dado  $d > 1$ , la familia*

$$\mathcal{F}_{\text{sym}}^d = \{f \circ l : l : \mathbb{R}^d \rightarrow \mathbb{R} \text{ lineal y } f \in \mathcal{F}_{\text{sym}}^1\}$$

*genera el orden simétrico entre vectores aleatorios en  $\mathbb{R}^d$ , es decir, para cualesquiera vectores aleatorios  $\xi$  y  $\eta$  en  $\mathbb{R}^d$ ,*

$$\xi \leq_{\text{sym}} \eta \text{ si y sólo si } \mathbb{E}f(\xi) \leq \mathbb{E}f(\eta) \text{ para todo } f \in \mathcal{F}_{\text{sym}}^d.$$

*Demostración.* Sea  $f \in \mathcal{F}_{\text{sym}}^1$ , definimos  $f_1 = (f - f(0))I_{(-\infty, 0]}$  y  $f_2 = (f - f(0))I_{[0, +\infty)}$ . Es claro que  $f_1$  es decreciente convexa,  $f_2$  es creciente convexa y  $f = f_1 + f_2 + f(0)$ . Dada una variable aleatoria  $\xi$ , por la linealidad del operador esperanza,

$$\mathbb{E}f(\xi) = \mathbb{E}f_1(\xi) + \mathbb{E}f_2(\xi) + f(0) = \mathbb{E}f_1(-\xi_-) + \mathbb{E}f_2(\xi_+) + f(0).$$

Como  $f_1$  es decreciente convexa,  $f_1 \circ (-\text{id})$  será creciente convexa, y como por el Teorema 1.1.1 el orden simétrico queda caracterizado por el orden creciente convexo en la parte positiva y negativa de una variable aleatoria, tenemos que si  $\xi \leq_{\text{sym}} \eta$ , entonces  $\mathbb{E}f(\xi) \leq \mathbb{E}f(\eta)$ . Para probar que la familia  $\mathcal{F}_{\text{sym}}^1$  es suficiente para caracterizar el orden simétrico, basta tener en cuenta

que todas las funciones convexas decrecientes hasta cero y luego constantes y las convexas constantes hasta cero y luego crecientes están contenidas en ella y aplicar de nuevo el Teorema 1.1.1.

Se demuestra que el conjunto  $\mathcal{F}_{\text{sym}}^d$  es familia de generadores por la linealidad del orden simétrico que se probó en el Teorema 1.1.2.  $\square$

### 1.1.5. Condiciones de igualdad estocástica

A continuación enunciamos dos condiciones de igualdad estocástica para el orden simétrico. La primera es más general e involucra la igualdad de la suma de los volúmenes de los zonoides y zonoides elevados de los vectores aleatorios. La segunda sólo se puede aplicar a variables aleatorias positivas, para las que el orden simétrico coincide con el orden creciente convexo.

Koshevoy y Mosler (1998) justifican que la medida de Lebesgue  $d$ -dimensional, el volumen, del zonoide de un vector aleatorio en  $\mathbb{R}^d$ , al que denotamos por  $V_d(Z(P_\xi))$ , es proporcional al valor absoluto del determinante de la matriz aleatoria  $M_\xi$  en  $\mathbb{R}^{d \times d}$  cuyas columnas son  $d$  copias independientes del vector aleatorio  $\xi$ ,

$$V_d(Z(P_\xi)) = \frac{1}{d!} \mathbb{E} |\det M_\xi|.$$

En el caso particular del zonoide elevado de una variable aleatoria, la expresión anterior resulta ser la *diferencia media de Gini* de la variable aleatoria, es decir

$$V_2(\tilde{Z}(P_\xi)) = \frac{1}{2} \int_{\mathbb{R}} \int_{\mathbb{R}} |x - y| dP_\xi(x) dP_\xi(y).$$

La diferencia media de Gini es un índice de desigualdad (Capítulo 2) entre variables aleatorias que mide la dispersión absoluta. Está mucho más popularizado el uso del *índice de Gini* que mide la dispersión relativa y es igual a la diferencia media de Gini dividida entre el valor absoluto de la esperanza de la variable aleatoria. Estas relaciones de los volúmenes de los zonoides con índices de desigualdad las utilizan Koshevoy y Mosler (1997a) para definir índices de desigualdad para vectores aleatorios.



Obsérvese que el volumen del zonoide de una variable aleatoria es igual a su primer momento absoluto,

$$V_1(Z(P_\xi)) = \mathbb{E}|\xi| = \int_{\mathbb{R}} |x| dP_\xi(x).$$

**Teorema 1.1.12.** *Dados  $\xi, \eta$  dos vectores aleatorios con primeros momentos finitos, si  $\xi \leq_{\text{sym}} \eta$  y la suma de los volúmenes de sus correspondientes zonoides y zonoides elevados coinciden, entonces  $\xi$  y  $\eta$  siguen la misma distribución.*

*Demostración.* Como  $Z(P_\xi)$  es la proyección ortogonal de  $\tilde{Z}(P_\xi)$  sobre el hiperplano  $\{x \in \mathbb{R}^{d+1} : x_1 = 1\}$  y gracias a la simetría de  $\tilde{Z}(P_\xi)$ , un simple razonamiento geométrico muestra que

$$V_{d+1}(W(P_\xi)) = \frac{1}{2} \left( V_d(Z(P_\xi)) + V_{d+1}(\tilde{Z}(P_\xi)) \right).$$

Según el Corolario 1.1.7, se cumple  $\xi \leq_{\text{sym}} \eta$  si y sólo si  $W(P_\xi) \subset W(P_\eta)$ . Por hipótesis estos dos cuerpos convexos tienen el mismo volumen, por lo tanto  $W(P_\xi) = W(P_\eta)$ , por lo que según el Corolario 1.1.8, los vectores  $\xi$  y  $\eta$  siguen la misma distribución.  $\square$

**Teorema 1.1.13.** *Sean  $\xi$  y  $\eta$  dos variables aleatorias positivas. Si satisfacen  $\xi \leq_{\text{sym}} \eta$  y  $\mathbb{E}\xi^2 = \mathbb{E}\eta^2$  entonces  $\xi$  y  $\eta$  tienen la misma distribución.*

*Demostración.* Como  $\xi$  y  $\eta$  son positivas, por el Teorema 1.1.1,  $\xi \leq_{\text{sym}} \eta$  implica  $\xi \leq_{\text{icx}} \eta$ . A partir del Teorema 0.5.6 se sigue la existencia de una variable aleatoria  $\zeta$  tal que  $\xi \leq_{\text{st}} \zeta \leq_{\text{cx}} \eta$ . Como la función  $f(x) = x^2$  es tanto convexa como creciente en  $\mathbb{R}_+$ , se cumple  $\mathbb{E}\xi^2 \leq \mathbb{E}\zeta^2 \leq \mathbb{E}\eta^2$ . Los extremos de la anterior cadena de desigualdades son iguales y por lo tanto se da la igualdad global. Tenemos que  $\xi \leq_{\text{st}} \zeta$  y tanto  $\xi$  como  $\zeta$  son variables aleatorias no negativas con  $\mathbb{E}\xi^2 = \mathbb{E}\zeta^2$ , podemos aplicar por tanto el Teorema 0.5.1 y deducir que  $\xi$  y  $\zeta$  siguen la misma distribución. Como  $\zeta \leq_{\text{cx}} \eta$  y  $\mathbb{E}\zeta^2 = \mathbb{E}\eta^2$ , del Teorema 0.5.5 deducimos que  $\eta$  y  $\zeta$  siguen también la misma distribución. Finalmente,  $\xi$  y  $\eta$  están idénticamente distribuidas.  $\square$

**Corolario 1.1.14.** Sean  $\xi$  y  $\eta$  dos variables aleatorias positivas. Si satisfacen  $\xi \leq_{\text{icx}} \eta$  y  $\mathbb{E}\xi^2 = \mathbb{E}\eta^2$ , entonces  $\xi$  y  $\eta$  siguen la misma distribución.

El siguiente ejemplo muestra que hay distribuciones de probabilidad que no son comparables respecto de los órdenes convexos habituales, pero sí lo son respecto del orden simétrico.

*Ejemplo 1.1.2.* Sean  $\xi$  y  $\eta$  dos variables aleatorias uniformemente distribuidas en  $(-1, 1)$  y  $(-3, 2)$  respectivamente. No es posible compararlas respecto del orden convexo porque tienen diferentes esperanzas. Tampoco son comparables respecto del orden creciente convexo. Sin embargo, se cumple  $\xi \leq_{\text{sym}} \eta$  como se deduce de la relación de contenido de las envolventes de zonoides que están dibujadas en la Figura 1.2.

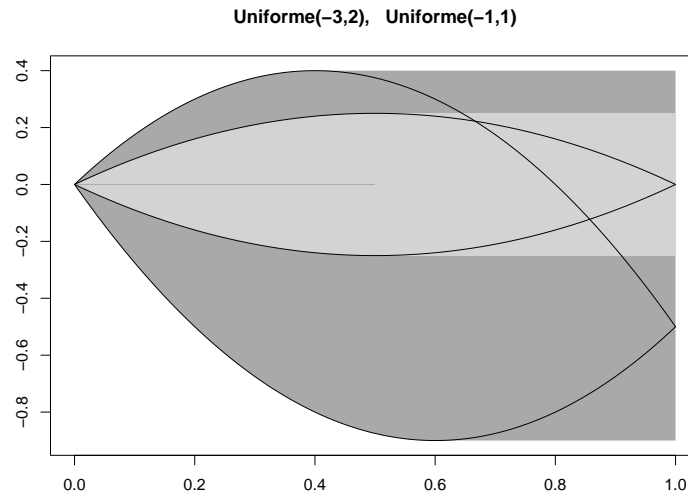


Figura 1.2: La región pintada en gris claro es la envolvente de zonoides de una variable uniforme en  $(-1, 1)$ , toda la región coloreada es la envolvente de zonoides de una variable uniforme en  $(-3, 2)$ .

## 1.2. Órdenes para conjuntos aleatorios

A lo largo de la presente sección se extenderán algunos órdenes según la variabilidad a conjuntos aleatorios. Varios autores han estudiado ordenaciones de distribuciones de probabilidad de conjuntos aleatorios, como por ejemplo Stoyan y Stoyan (1980). Nuestro objetivo es extender al marco de los conjuntos aleatorios órdenes estocásticos convexos bien estudiados para vectores aleatorios.

### 1.2.1. Orden simétrico

La Definición 1.1.1 puede extenderse para comparar conjuntos aleatorios compactos.

*Definición 1.2.1.* Un conjunto aleatorio compacto  $X$  se dice *simétricamente más pequeño en distribución* que otro conjunto aleatorio  $Y$  y se denota  $X \leq_{\text{sym}} Y$ , si

$$\mathbb{E}(X \cup B_r) \subset \mathbb{E}(Y \cup B_r)$$

para todo  $r > 0$ .

*Nota 1.2.1.* La relación para conjuntos aleatorios compactos introducida en la Definición 1.2.1 es reflexiva y transitiva, pero no es antisimétrica, por lo tanto no define un orden parcial en el conjunto de las distribuciones de probabilidad de los conjuntos aleatorios compactos. Por ejemplo, si  $X = \{1\}$  e  $Y = [0, 1]$  son dos conjuntos deterministas, entonces se satisface tanto  $X \leq_{\text{sym}} Y$  como  $Y \leq_{\text{sym}} X$ , pero  $X$  e  $Y$  siguen distribuciones diferentes.

El siguiente resultado relaciona el orden simétrico entre conjuntos aleatorios con el orden simétrico entre las partes positivas de sus funciones soporte en cualquier dirección, o lo que es equivalente, con el orden creciente convexo entre las partes positivas de sus funciones soporte en cualquier dirección.

**Teorema 1.2.1.** *Dados dos conjuntos aleatorios compactos  $X$  e  $Y$ , se cumple  $X \leq_{\text{sym}} Y$  si y sólo si  $h(X, u)_+ \leq_{\text{sym}} h(Y, u)_+$  para todo  $u \in \mathbb{R}^d$ .*

*Demostración.* Dado un conjunto aleatorio compacto  $X$  y  $r > 0$ , se tiene

$$h(\mathbb{E}(X \cup B_r), u) = \mathbb{E}h(X \cup B_r, u) = r\|u\| + \mathbb{E}(h(X, u) - r\|u\|)_+.$$

Dados dos conjuntos aleatorios compactos  $X$  e  $Y$ , se cumple  $X \leq_{\text{sym}} Y$  si y sólo si  $\mathbb{E}(h(X, u) - r)_+ \leq \mathbb{E}(h(Y, u) - r)_+$  para todo  $r > 0$  y  $u \in \mathbb{R}^d$  o, equivalentemente,  $\mathbb{E}(h(X, u)_+ - r)_+ \leq \mathbb{E}(h(Y, u)_+ - r)_+$  para todo  $r \in \mathbb{R}$  y  $u \in \mathbb{R}^d$ . De acuerdo con el Teorema 0.5.4, esto se cumple si y sólo si  $h(X, u)_+ \leq_{\text{icx}} h(Y, u)_+$  para todo  $u \in \mathbb{R}^d$ . Como se deduce a partir de la caracterización del orden simétrico del Teorema 1.1.1, para variables aleatorias no negativas, el orden creciente convexo y el orden simétrico son equivalentes y esta equivalencia concluye la demostración.  $\square$

Finalmente, relacionamos el orden simétrico entre conjuntos aleatorios con el orden simétrico entre algunos pares de sus selecciones.

**Teorema 1.2.2.** *Sean  $X$  e  $Y$  dos conjuntos aleatorios compactos. Si para cada selección  $\xi$  de  $X$  existe una selección  $\eta$  de  $Y$  tal que  $\xi \leq_{\text{sym}} \eta$ , entonces  $X \leq_{\text{sym}} Y$ .*

*Demostración.* Sea  $r > 0$ , demostraremos que  $\mathbb{E}(X \cup B_r) \subset \mathbb{E}(Y \cup B_r)$ . Para cualquier  $a \in \mathbb{E}(X \cup B_r)$ , existe una selección  $\xi_1$  de  $X \cup B_r$  tal que  $a = \mathbb{E}\xi_1$ . Sea  $\xi$  una selección de  $X$  tal que  $\xi = \xi_1$  si  $\|\xi_1\| > r$  y en otro caso  $\xi$  toma cualquier valor de  $X$  con tal de que se garantice la medibilidad. Como  $h(\{\xi_1\} \cup B_r, u) = r\|u\| + \text{máx}\{\langle \xi_1, u \rangle - r\|u\|, 0\}$  y

$$\text{máx}\{\langle \xi_1, u \rangle - r, 0\} \leq \text{máx}\{\langle \xi, u \rangle - r, 0\},$$

tenemos  $\mathbb{E}(\{\xi_1\} \cup B_r) \subset \mathbb{E}(\{\xi\} \cup B_r)$ . Por hipótesis existe una selección  $\eta$  de  $Y$  tal que  $\xi \leq_{\text{sym}} \eta$ . Finalmente,  $\mathbb{E}(\{\xi\} \cup B_r) \subset \mathbb{E}(\{\eta\} \cup B_r) \subset \mathbb{E}(Y \cup B_r)$  y como  $\xi_1$  es una selección de  $\{\xi_1\} \cup B_r$ , se cumple  $a \in \mathbb{E}(\{\xi_1\} \cup B_r)$ , con lo que  $a \in \mathbb{E}(Y \cup B_r)$  y por tanto  $\mathbb{E}(X \cup B_r) \subset \mathbb{E}(Y \cup B_r)$ .  $\square$

La implicación recíproca a la probada en el Teorema 1.2.2 no es cierta como se verá posteriormente en el Ejemplo 1.2.1.

### 1.2.2. Orden lineal convexo

En primer lugar vamos a dar dos caracterizaciones diferentes para el orden convexo entre variables aleatorias.

**Lema 1.2.3.** *Dadas dos variables aleatorias  $\xi$  y  $\eta$ , se cumple  $\xi \leq_{\text{cx}} \eta$  si y sólo si  $\xi \leq_{\text{icx}} \eta$  y  $-\xi \leq_{\text{icx}} -\eta$ .*

*Demostración.* Si se satisface  $\xi \leq_{\text{cx}} \eta$ , entonces se cumple  $\xi \leq_{\text{icx}} \eta$  y también  $-\xi \leq_{\text{cx}} -\eta$ , por lo tanto tenemos que  $-\xi \leq_{\text{icx}} -\eta$ .

Si se satisface  $\xi \leq_{\text{icx}} \eta$ , entonces se cumple  $\mathbb{E}\xi \leq \mathbb{E}\eta$ , y si se satisface  $-\xi \leq_{\text{icx}} -\eta$ , entonces se cumple  $-\mathbb{E}\xi \leq -\mathbb{E}\eta$ . Por lo tanto, tenemos que  $\mathbb{E}\xi = \mathbb{E}\eta$ , lo que conjuntamente con la relación  $\xi \leq_{\text{icx}} \eta$  implica  $\xi \leq_{\text{cx}} \eta$  por el Teorema 0.5.2.  $\square$

**Lema 1.2.4.** *Dadas dos variables aleatorias  $\xi$  y  $\eta$ , se cumple  $\xi \leq_{\text{cx}} \eta$  si y sólo si*

$$\mathbb{E}([\xi - s, \xi + s] \cup [-r, r]) \subset \mathbb{E}([\eta - s, \eta + s] \cup [-r, r])$$

para todo  $r, s > 0$ .

*Demostración.* Obviamente se tiene que

$$\mathbb{E}([\xi - s, \xi + s] \cup [-r, r]) = [-r - \mathbb{E}(-\xi + s - r)_+, r + \mathbb{E}(\xi + s - r)_+].$$

Para que  $\mathbb{E}([\xi - s, \xi + s] \cup [-r, r]) \subset \mathbb{E}([\eta - s, \eta + s] \cup [-r, r])$  se cumpla para cualquier  $r, s > 0$ , es condición necesaria y suficiente el que  $\mathbb{E}(\xi - t)_+ \leq \mathbb{E}(\eta - t)_+$  y  $\mathbb{E}(-\xi - t)_+ \leq \mathbb{E}(-\eta - t)_+$  para cualquier  $t \in \mathbb{R}$ . Por el Teorema 0.5.4 esto se cumple si y sólo si  $\xi \leq_{\text{icx}} \eta$  y  $-\xi \leq_{\text{icx}} -\eta$ , lo que es equivalente a  $\xi \leq_{\text{cx}} \eta$  de acuerdo con el Lema 1.2.3.  $\square$

Apoyándonos en este último resultado, damos una nueva caracterización del orden lineal convexo entre vectores aleatorios.

**Teorema 1.2.5.** *Dados dos vectores aleatorios  $\xi$  y  $\eta$ , se cumple  $\xi \leq_{\text{lcx}} \eta$  si y sólo si  $\mathbb{E}(\{\xi\}^s \cup B_r) \subset \mathbb{E}(\{\eta\}^s \cup B_r)$  para todo  $r, s > 0$ .*

*Demostración.* Dado cualquier  $u \in \mathbb{R}^d$ ,

$$\{\langle x, u \rangle : x \in \mathbb{E}(\{\xi\}^s \cup B_r)\} = \mathbb{E}([\langle \xi, u \rangle - s\|u\|, \langle \xi, u \rangle + s\|u\|] \cup [-r\|u\|, r\|u\|]).$$

Por lo tanto  $\mathbb{E}(\{\xi\}^s \cup B_r) \subset \mathbb{E}(\{\eta\}^s \cup B_r)$  se cumple para todo  $r, s > 0$  si y sólo si

$$\mathbb{E}([\langle \xi, u \rangle - s, \langle \xi, u \rangle + s] \cup [-r, r]) \subset \mathbb{E}([\langle \eta, u \rangle - s, \langle \eta, u \rangle + s] \cup [-r, r])$$

para todo  $u \in \mathbb{R}^d$  y  $r, s > 0$ . Por el Lema 1.2.4, esto es equivalente a  $\langle \xi, u \rangle \leq_{\text{cx}} \langle \eta, u \rangle$  para todo  $u \in \mathbb{R}^d$ , y por lo tanto  $\xi \leq_{\text{icx}} \eta$ .  $\square$

Una expresión similar a la usada en el Teorema 1.2.5 lleva al siguiente orden estocástico para conjuntos aleatorios compactos, más restrictivo que el de la Definición 1.2.1.

*Definición 1.2.2.* Sean  $X, Y$  dos conjuntos aleatorios compactos. Definimos el *orden lineal convexo entre conjuntos aleatorios* denotado por  $X \leq_{\text{slcx}} Y$ , como

$$\mathbb{E}(X^s \cup B_r) \subset \mathbb{E}(Y^s \cup B_r) \quad \text{para todo } s, r > 0. \quad (1.2)$$

El subíndice en el símbolo ' $\leq_{\text{slcx}}$ ' es un abreviatura de "set linear convex".

*Nota 1.2.2.* La relación ' $\leq_{\text{slcx}}$ ' no es un orden sobre el conjunto de distribuciones de probabilidad de los conjuntos aleatorios compactos, sino un preorden. Para ver que no cumple la propiedad antisimétrica, basta con considerar los intervalos aleatorios  $X$  e  $Y$ , de tal modo que  $X$  toma los valores  $[-1, 0]$  y  $[0, 1]$  con igual probabilidad, e  $Y$  toma los valores  $[-1, 1]$  y  $\{0\}$  también con igual probabilidad. Estos dos intervalos aleatorios cumplen  $X \leq_{\text{slcx}} Y$  e  $Y \leq_{\text{slcx}} X$ .

El siguiente resultado interpreta (1.2) utilizando las funciones soporte de  $X$  e  $Y$ .

**Teorema 1.2.6.** *Sean  $X$  e  $Y$  dos conjuntos aleatorios compactos. Se cumple  $X \leq_{\text{slcx}} Y$  si y sólo si  $h(X, u) \leq_{\text{icx}} h(Y, u)$  para todo  $u \in \mathbb{R}^d$ .*

*Demostración.* Obsérvese que  $X \leq_{\text{slcx}} Y$  si y sólo si  $\mathbb{E}h(X^s \cup B_r, u) \leq \mathbb{E}h(Y^s \cup B_r, u)$  para todo  $u \in \mathbb{R}^d$  y  $r, s > 0$ . Esto se satisfará si y sólo si  $\mathbb{E}(h(X, u) - t)_+ \leq \mathbb{E}(h(Y, u) - t)_+$  para todo  $u \in \mathbb{R}^d$  y  $t \in \mathbb{R}$  o, equivalentemente,  $h(X, u) \leq_{\text{icx}} h(Y, u)$  para todo  $u \in \mathbb{R}^d$  de acuerdo con el Teorema 0.5.4.  $\square$

Puesto que el orden creciente convexo entre variables aleatorias implica la ordenación de sus esperanzas, si dos conjuntos aleatorios están ordenados respecto del orden ' $\leq_{\text{slcx}}$ ', entonces la esperanza de Aumann del conjunto más pequeño estará contenida en la esperanza de Aumann del otro.

**Corolario 1.2.7.** *Si  $X \leq_{\text{slcx}} Y$ , entonces  $\mathbb{E}X \subset \mathbb{E}Y$  y  $X^r \leq_{\text{slcx}} Y^r$  para todo  $r > 0$ .*

Sobre la base del Corolario 1.2.7 y del Teorema 0.5.2 se obtiene el siguiente resultado.

**Corolario 1.2.8.** *Sean  $X$  e  $Y$  dos conjuntos aleatorios compactos. Se cumplen  $X \leq_{\text{slcx}} Y$  y  $\mathbb{E}X = \mathbb{E}Y$  si y sólo si  $h(X, u) \leq_{\text{cx}} h(Y, u)$  para todo  $u \in \mathbb{R}^d$ .*

Para finalizar esta sección enunciamos dos resultados que relacionan el orden ' $\leq_{\text{slcx}}$ ' entre dos conjuntos aleatorios y el orden lineal convexo entre algunos pares de sus selecciones. Obsérvese que de acuerdo con el Teorema 1.2.5, el orden lineal convexo es en realidad el orden ' $\leq_{\text{slcx}}$ ' cuando consideramos a los vectores aleatorios como conjuntos aleatorios con un único elemento.

**Teorema 1.2.9.** *Sean  $X$  e  $Y$  dos conjuntos aleatorios compactos. Si para cada selección  $\xi$  de  $X$  existe una selección  $\eta$  de  $Y$  tal que  $\xi \leq_{\text{lcx}} \eta$ , entonces  $X \leq_{\text{slcx}} Y$ .*

*Demostración.* Sean  $s, r > 0$  y  $a \in \mathbb{E}(X^s \cup B_r)$ . Existe entonces una selección  $\xi_1$  de  $X^s \cup B_r$ , tal que  $\mathbb{E}\xi_1 = a$ . Definimos una selección  $\xi_2$  de  $X^s$  igual a  $\xi_1$  cuando  $\|\xi_1\| > r$  y tomando cualquier valor en  $X^s$  que asegure su medibilidad

en otro caso. Como  $\xi_2$  es una selección de  $X^s$ , entonces existe una selección  $\xi \in X$  tal que  $\xi_2$  es una selección de  $\{\xi\}^s$ . Se cumple

$$a = \mathbb{E}\xi_1 \in \mathbb{E}(\{\xi_1\} \cup B_r) \subset \mathbb{E}(\{\xi_2\} \cup B_r) \subset \mathbb{E}(\{\xi\}^s \cup B_r).$$

Como  $\xi$  es una selección de  $X$ , por hipótesis existe  $\eta$ , selección de  $Y$ , tal que  $\xi \leq_{\text{lcx}} \eta$ . Entonces, según el Teorema 1.2.5, se cumple  $\mathbb{E}(\{\xi\}^s \cup B_r) \subset \mathbb{E}(\{\eta\}^s \cup B_r)$  y por ser  $\eta$  una selección de  $Y$ , tenemos  $\mathbb{E}(\{\eta\}^s \cup B_r) \subset \mathbb{E}(Y^s \cup B_r)$ , por tanto  $\mathbb{E}(X^s \cup B_r) \subset \mathbb{E}(Y^s \cup B_r)$ .  $\square$

**Teorema 1.2.10.** Sean  $X, Y$  dos conjuntos aleatorios tales que  $\mathbb{E}X \subset \mathbb{E}Y$ . Si para toda selección  $\eta$  de  $Y$  existe una selección  $\xi$  de  $X$  tal que  $\xi \leq_{\text{lcx}} \eta$ , entonces  $X \leq_{\text{slcx}} Y$ .

*Demostración.* Sea  $\eta$  una selección de  $Y$ , entonces existe  $\xi$ , selección de  $X$  tal que  $\xi \leq_{\text{lcx}} \eta$ , por lo tanto, según el Corolario 0.5.3, se cumple que  $\mathbb{E}\xi = \mathbb{E}\eta$ , y de ahí  $\mathbb{E}Y \subset \mathbb{E}X$ . Esto conjuntamente con  $\mathbb{E}X \subset \mathbb{E}Y$  implica que las esperanzas de  $X$  e  $Y$  coinciden.

Sea  $u \in \mathbb{R}^d$ , tomamos  $\eta$  una selección de  $Y$  tal que  $\langle \eta, u \rangle = h(Y, u)$ , entonces  $\langle \mathbb{E}\eta, u \rangle = h(\mathbb{E}Y, u) = h(\mathbb{E}X, u)$ . Sabemos que existe  $\xi$  una selección de  $X$  tal que  $\xi \leq_{\text{lcx}} \eta$ , por lo tanto  $\mathbb{E}\xi = \mathbb{E}\eta$  y  $\langle \mathbb{E}\xi, u \rangle = \langle \mathbb{E}\eta, u \rangle = h(\mathbb{E}X, u)$ . Esto quiere decir que  $\langle \xi, u \rangle = h(X, u)$  c.s. y finalmente  $h(X, u) \leq_{\text{cx}} h(Y, u)$  lo que garantiza  $X \leq_{\text{slcx}} Y$  de acuerdo con el Corolario 1.2.8.  $\square$

### 1.2.3. Orden creciente positivo lineal convexo

Un conjunto  $C \subset \mathbb{R}^d$  se dice *superior* si dado  $x \in C$  y  $x \leq y$  donde la relación menor o igual se entiende componente a componente, entonces  $y \in C$  (véase pág. 82 en Müller y Stoyan, 2002). Equivalentemente  $C$  es superior si  $C + \mathbb{R}_+^d = C$  donde recordamos que  $\mathbb{R}_+^d$  es el primer cuadrante cerrado. Un conjunto  $C \subset \mathbb{R}^d$  se dice *inferior* si  $-C$  es superior.

Dado  $H \subset \mathbb{R}^d$  un semiespacio cerrado,  $H$  será superior si existen  $v \in \mathbb{R}_+^d$  y  $a \in \mathbb{R}$  tales que  $H = \{x \in \mathbb{R}^d : \langle x, v \rangle \geq a\}$ . La función soporte de  $H$



evaluada en cualquier dirección  $u \in \mathbb{R}^d \setminus \{0\}$  distinta de  $-v$  será  $+\infty$  y se satisface que  $h(H, -v) = -a$ .

A continuación, caracterizaremos el orden creciente positivo lineal convexo entre vectores aleatorios a partir de una relación de contenido de ciertas esperanzas de Aumann.

**Teorema 1.2.11.** *Los siguientes enunciados son equivalentes:*

- i.*  $\xi \leq_{\text{iplcx}} \eta$ ;
- ii.*  $\mathbb{E}(\{\eta\} \cup H) \subset \mathbb{E}(\{\xi\} \cup H)$  para todo semiespacio cerrado superior  $H$ ;
- iii.*  $\mathbb{E}(\{\xi\} \cup H) \subset \mathbb{E}(\{\eta\} \cup H)$  para todo semiespacio cerrado inferior  $H$ .

*Demostración.* Sea  $H$  un semiespacio cerrado superior, entonces  $H = \{x \in \mathbb{R}^d : \langle x, v \rangle \geq a\}$  para ciertos  $v \in \mathbb{R}_+^d$  y  $a \in \mathbb{R}$ . Sea  $u \in \mathbb{R}^d \setminus \{0, -v\}$ , entonces  $h(\mathbb{E}(\{\xi\} \cup H), u) = +\infty$ . Si estudiamos la función soporte en la dirección de  $-v$  obtenemos que

$$h(\mathbb{E}(\{\xi\} \cup H), -v) = \mathbb{E} \max\{\langle \xi, -v \rangle, -a\} = \mathbb{E}(\langle \xi, -v \rangle + a)_+ - a.$$

Ahora la relación de contenido  $\mathbb{E}(\{\eta\} \cup H) \subset \mathbb{E}(\{\xi\} \cup H)$  se cumple para todo semiespacio cerrado superior  $H$  si y sólo si  $\mathbb{E}(\langle \eta, -v \rangle + a)_+ \leq \mathbb{E}(\langle \xi, -v \rangle + a)_+$  para todo  $a \in \mathbb{R}$  y todo  $v \in \mathbb{R}_+^d$ . Según el Teorema 0.5.4, esto es equivalente a  $\langle \eta, -v \rangle \leq_{\text{icx}} \langle \xi, -v \rangle$  para todo  $v \in \mathbb{R}_+^d$ , es decir  $\langle \xi, v \rangle \leq_{\text{icx}} \langle \eta, v \rangle$  para todo  $v \in \mathbb{R}_+^d$ , por lo tanto se cumple  $\xi \leq_{\text{iplcx}} \eta$  y la equivalencia entre *i.* y *ii.* queda probada.

Para demostrar que *iii.* es equivalente a los dos enunciados anteriores, sea  $\xi \leq_{\text{iplcx}} \eta$  o lo que es lo mismo,  $-\eta \leq_{\text{iplcx}} -\xi$ . Por la primera parte de este resultado, dicha desigualdad es cierta si y sólo si  $\mathbb{E}(\{-\xi\} \cup H) \subset \mathbb{E}(\{-\eta\} \cup H)$  para todo semiespacio cerrado superior  $H$ , en consecuencia  $\mathbb{E}(\{\xi\} \cup -H) \subset \mathbb{E}(\{\eta\} \cup -H)$  para todo semiespacio cerrado superior  $H$ , equivalentemente  $\mathbb{E}(\{\xi\} \cup H) \subset \mathbb{E}(\{\eta\} \cup H)$  para todo semiespacio cerrado inferior  $H$  con lo que obtenemos el enunciado *iii.*  $\square$

El orden creciente positivo lineal convexo está relacionado con el zonoide elevado según nos indica el siguiente resultado.

**Proposición 1.2.12.** *Dados  $\xi, \eta$  dos vectores aleatorios en  $\mathbb{R}^d$ ,*

$$\xi \leq_{\text{iplcx}} \eta \text{ si y sólo si } h(\tilde{Z}(P_\xi), (r, u)) \leq h(\tilde{Z}(P_\eta), (r, u)) \quad \forall r \in \mathbb{R}, u \in \mathbb{R}_+^d.$$

*Demostración.* Dados dos vectores aleatorios  $\xi$  y  $\eta$ , se cumple  $\xi \leq_{\text{iplcx}} \eta$  si y sólo si  $\mathbb{E}(\langle \xi, u \rangle + r)_+ \leq \mathbb{E}(\langle \eta, u \rangle + r)_+$  para todo  $u \in \mathbb{R}_+^d$  y  $r \in \mathbb{R}$  o lo que es lo mismo  $\mathbb{E}h([0, (1, \xi)], (r, u)) \leq \mathbb{E}h([0, (1, \eta)], (r, u))$  o equivalentemente  $h(\tilde{Z}(P_\xi), (r, u)) \leq h(\tilde{Z}(P_\eta), (r, u))$  para todo  $u \in \mathbb{R}_+^d$  y  $r \in \mathbb{R}$ .  $\square$

Inspirándonos en el Teorema 1.2.11 podemos dar dos definiciones de relaciones *crecientes lineal convexas* para conjuntos aleatorios.

*Definición 1.2.3.* Sean  $X, Y$  dos conjuntos aleatorios compactos en  $\mathbb{R}^d$ , definimos la relación *superior creciente lineal convexa* ' $\leq_{\text{uilcx}}$ ' y la relación *inferior creciente lineal convexa* ' $\leq_{\text{lilcx}}$ ' del siguiente modo:

- i.  $X \leq_{\text{uilcx}} Y$  si y sólo si  $\mathbb{E}(X \cup H) \subset \mathbb{E}(Y \cup H)$  para todo semiespacio cerrado inferior  $H$ ;
- ii.  $X \leq_{\text{lilcx}} Y$  si y sólo si  $\mathbb{E}(Y \cup H) \subset \mathbb{E}(X \cup H)$  para todo semiespacio cerrado superior  $H$ .

Obsérvese que el orden superior creciente lineal convexo se define a partir de semiespacios inferiores, mientras que el orden inferior creciente lineal convexo se define a partir de semiespacios superiores.

En la demostración del Teorema 1.2.11 utilizamos la caracterización  $\xi \leq_{\text{iplcx}} \eta$  si y sólo si  $-\eta \leq_{\text{iplcx}} -\xi$ . Las nuevas relaciones dadas en la Definición 1.2.3 satisfacen una propiedad similar.

**Proposición 1.2.13.** *Dados  $X, Y$  dos conjuntos aleatorios compactos, se cumple  $X \leq_{\text{uilcx}} Y$  si y sólo si  $-Y \leq_{\text{lilcx}} -X$ .*

*Demostración.* La demostración es inmediata. Basta considerar las definiciones de los órdenes superior creciente lineal convexo e inferior creciente lineal convexo.  $\square$

**Proposición 1.2.14.** *Dos conjuntos aleatorios compactos  $X$  e  $Y$  satisfacen  $X \leq_{\text{uilcx}} Y$  si y sólo si  $h(X, u) \leq_{\text{icx}} h(Y, u)$  para todo  $u \in \mathbb{R}_+^d$ .*

*Demostración.* Sean  $u \in \mathbb{R}_+^d$  y  $a \in \mathbb{R}$ , el conjunto  $H = \{x \in \mathbb{R}^d : \langle x, u \rangle \leq a\}$  es un semiespacio cerrado inferior y se cumple

$$h(\mathbb{E}(X \cup H), u) = \mathbb{E} \max\{h(X, u), a\} = \mathbb{E}(h(X, u) - a)_+ + a.$$

Finalmente, se cumple  $X \leq_{\text{uilcx}} Y$  si y sólo si  $\mathbb{E}(h(X, u) - a)_+ \leq \mathbb{E}(h(Y, u) - a)_+$  para todo  $a \in \mathbb{R}$  y  $u \in \mathbb{R}_+^d$  o equivalentemente, según el Teorema 0.5.4,  $h(X, u) \leq_{\text{icx}} h(Y, u)$  para todo  $u \in \mathbb{R}_+^d$ .  $\square$

**Proposición 1.2.15.** *Dos conjuntos aleatorios  $X$  e  $Y$  satisfacen  $X \leq_{\text{lilcx}} Y$  si y sólo si  $h(Y, u) \leq_{\text{icx}} h(X, u)$  para todo  $u \in \mathbb{R}_-^d$ .*

*Demostración.* Sea  $X \leq_{\text{lilcx}} Y$ , según la Proposición 1.2.13,  $-Y \leq_{\text{uilcx}} -X$  y entonces, por la Proposición 1.2.14, tenemos  $h(-Y, u) \leq_{\text{icx}} h(-X, u)$  para todo  $u \in \mathbb{R}_+^d$  o equivalentemente  $h(Y, u) \leq_{\text{icx}} h(X, u)$  para todo  $u \in \mathbb{R}_-^d$ .  $\square$

Como consecuencia del Teorema 1.2.6 y de las Proposiciones 1.2.14 y 1.2.15 se obtiene el siguiente resultado.

**Proposición 1.2.16.** *Si  $X \leq_{\text{slcx}} Y$ , entonces  $X \leq_{\text{uilcx}} Y$  e  $Y \leq_{\text{lilcx}} X$ .*

Finalmente, veremos cómo se relacionan los órdenes superior creciente lineal convexo e inferior creciente lineal convexo entre conjuntos aleatorios con el orden creciente positivo lineal convexo entre algunos pares de sus selecciones.

**Teorema 1.2.17.** *Sean  $X$  e  $Y$  dos conjuntos aleatorios compactos. Si para cada selección  $\xi$  de  $X$  existe una selección  $\eta$  de  $Y$  tal que  $\xi \leq_{\text{iplcx}} \eta$ , entonces  $X \leq_{\text{uilcx}} Y$ .*

*Demostración.* Sea  $H$  un semiespacio cerrado inferior, por lo tanto existen  $u \in \mathbb{R}_+^d$  y  $a \in \mathbb{R}$  tales que  $H = \{x \in \mathbb{R}^d : \langle x, u \rangle \leq a\}$ . Veremos que  $\mathbb{E}(X \cup H) \subset \mathbb{E}(Y \cup H)$ . Para cualquier  $a \in \mathbb{E}(X \cup H)$ , existe una selección  $\xi_1$

de  $X \cup H$  tal que  $a = \mathbb{E}\xi_1$ . Sea  $\xi$  una selección de  $X$  tal que  $\xi = \xi_1$  si  $\xi_1 \notin H$  y en otro caso  $\xi$  toma cualquier valor de  $X$  con tal de que se garantice la medibilidad. Como  $h(\{\xi_1\} \cup H, u) = \max\{\langle \xi_1, u \rangle - a, 0\} + a$  y

$$\max\{\langle \xi_1, u \rangle - a, 0\} \leq \max\{\langle \xi, u \rangle - a, 0\}$$

tenemos  $\mathbb{E}(\{\xi_1\} \cup H) \subset \mathbb{E}(\{\xi\} \cup H)$ . Por hipótesis existe una selección  $\eta$  de  $Y$  tal que  $\xi \leq_{\text{iplcx}} \eta$ . Finalmente, se cumple  $\mathbb{E}(\{\xi\} \cup H) \subset \mathbb{E}(\{\eta\} \cup H) \subset \mathbb{E}(Y \cup H)$ .  $\square$

**Teorema 1.2.18.** *Sean  $X$  e  $Y$  dos conjuntos aleatorios compactos. Si para cada selección  $\eta$  de  $Y$  existe una selección  $\xi$  de  $X$  tal que  $\xi \leq_{\text{iplcx}} \eta$ , entonces  $X \leq_{\text{lilcx}} Y$ .*

*Demostración.* Sea  $-\eta$  una selección de  $-Y$ , entonces  $\eta$  es una selección de  $Y$  y, por hipótesis, existe  $\xi$ , selección de  $X$ , tal que  $\xi \leq_{\text{iplcx}} \eta$ . El vector aleatorio  $-\xi$  es una selección de  $-X$  y se cumple  $-\eta \leq_{\text{iplcx}} -\xi$ . Como esto sucede para cualquier selección de  $-Y$ , entonces, según el Teorema 1.2.17, se tiene que  $-Y \leq_{\text{uilcx}} -X$  y, según la Proposición 1.2.13,  $X \leq_{\text{lilcx}} Y$ .  $\square$

En general los recíprocos de los Teoremas 1.2.2, 1.2.9, 1.2.10, 1.2.17 y 1.2.18 no se cumplen.

*Ejemplo 1.2.1.* Sean  $X, Y$  dos conjuntos aleatorios con la siguiente distribución,

$$\begin{aligned} \Pr(X = \{(2, 0)\}) &= \Pr(X = \{(0, 2)\}) = 1/2; \\ \Pr(Y = \text{co}\{(1, 0), (0, 1)\}) &= \Pr(Y = \text{co}\{(3, 0), (0, 3)\}) = 1/2. \end{aligned}$$

Es claro que  $X \leq_{\text{slcx}} Y$  y por lo tanto también se cumplen las relaciones  $X \leq_{\text{sym}} Y$ ,  $X \leq_{\text{uilcx}} Y$  y  $X \leq_{\text{lilcx}} Y$ . Sin embargo para una selección  $\eta$  de  $Y$  tal que  $\Pr(\eta = (1, 0)) = \Pr(\eta = (3, 0)) = 1/2$ , la única selección de  $X$  a la que denotamos por  $\xi$ ,  $X = \{\xi\}$ , no es menor que  $\eta$  ni respecto del orden lineal convexo puesto que sus esperanzas son distintas, ni del creciente positivo lineal convexo. Basta considerar  $u = (0, 1) \in \mathbb{R}^2$  para obtener

$\Pr(\langle \eta, u \rangle = 0) = 1$  y  $\Pr(\langle \xi, u \rangle = 0) = \Pr(\langle \xi, u \rangle = 2) = 1/2$  y por lo tanto  $1 = \mathbb{E}\langle \xi, u \rangle > \mathbb{E}\langle \eta, u \rangle = 0$ , por lo que  $\langle \xi, u \rangle$  no puede ser más pequeño que  $\langle \eta, u \rangle$  respecto del orden creciente convexo.

Además, para  $\xi$ , la selección de  $X$ , no existe ninguna selección de  $Y$  que la mayor ni respecto del orden creciente positivo lineal convexo, ni del simétrico y por lo tanto tampoco del lineal convexo. Si consideramos  $u = (0, 1)$ ,  $v = (1, 0) \in \mathbb{R}^2$ , se cumple que para cualquier selección  $\eta$  de  $Y$ , las variables aleatorias  $\langle \eta, u \rangle$  y  $\langle \eta, v \rangle$  toman valores positivos y lo mismo ocurre con las mismas transformaciones aplicadas sobre  $\xi$ . Es suficiente demostrar que no existe ninguna selección  $\eta$  de  $Y$  cuyas transformaciones sean mayores que las de  $\xi$  según el orden creciente convexo. Si  $\langle \xi, u \rangle \leq_{\text{icx}} \langle \eta, u \rangle$ , entonces  $\Pr(\langle \eta, u \rangle \geq 2) = \Pr(\eta \in \text{co}\{(0, 3), (1, 2)\}) \geq 1/2$ , si  $\langle \xi, v \rangle \leq_{\text{icx}} \langle \eta, v \rangle$ , entonces  $\Pr(\langle \eta, v \rangle \geq 2) = \Pr(\eta \in \text{co}\{(3, 0), (2, 1)\}) \geq 1/2$ , pero ambas circunstancias no pueden ocurrir a la vez, así que no existe ninguna selección de  $Y$  en tales condiciones.

### 1.3. Métrica de probabilidad

Existe una estrecha relación entre las métricas de probabilidad y los órdenes estocásticos que queda patente en varios trabajos recientes como los de Müller (1997a) y muy especialmente Boutsikas y Vaggelatou (2002).

Una *métrica de probabilidad (simple)* es un funcional  $d$  de pares de vectores aleatorios que toma valores en el intervalo  $[0, +\infty]$  y constituye una distancia entre las distribuciones de probabilidad de los vectores aleatorios, véase la Definición 2.3.1 en Rachev (1991).

*Definición 1.3.1.* Un funcional  $d : \Xi \times \Xi \rightarrow [0, +\infty]$ , donde  $\Xi$  es un espacio de vectores aleatorios en  $\mathbb{R}^d$ , se dice *métrica de probabilidad simple* si para cualesquiera vectores aleatorios  $\xi, \eta$  y  $\zeta$ , se cumple

- i.  $d(\xi, \eta) = 0$  si y sólo si  $\xi$  y  $\eta$  siguen la misma distribución;
- ii.  $d(\xi, \eta) = d(\eta, \xi)$ ;
- iii.  $d(\xi, \eta) \leq d(\xi, \zeta) + d(\zeta, \eta)$ .

Decimos que una métrica de probabilidad es *ideal* o más específicamente *ideal para el esquema de la suma de vectores aleatorios*, véase la Definición 14.2.1 en Rachev (1991) si satisface las propiedades de *homogeneidad* y *regularidad*. Estas propiedades la hacen especialmente propicia para el estudio de la convergencia de sucesiones de sumas parciales de vectores aleatorios independientes e idénticamente distribuidos.

*Definición 1.3.2.* Decimos que una métrica de probabilidad  $d$  es una *métrica de probabilidad ideal de orden  $s$* , si satisface las siguientes dos propiedades para cualesquiera vectores aleatorios  $\xi, \eta$  y  $\zeta$ :

- i. *Homogeneidad de orden  $s$ .*  $d(c\xi, c\eta) \leq |c|^s d(\xi, \eta)$  para cualquier  $c \neq 0$ ;
- ii. *Regularidad.*  $d(\xi + \zeta, \eta + \zeta) \leq d(\xi, \eta)$  para cualquier  $\zeta$  independiente de  $\xi$  y  $\eta$ .

A continuación definimos una distancia entre distribuciones de probabilidad inspirada en el orden simétrico.

*Definición 1.3.3.* Dados dos vectores aleatorios  $\xi$  y  $\eta$  en  $\mathbb{R}^d$ , definimos

$$m(\xi, \eta) := \int_0^\infty d_1(\mathbb{E}(\{\xi\} \cup B_r), \mathbb{E}(\{\eta\} \cup B_r)) dr,$$

donde

$$d_1(A, B) = \int_{S^{d-1}} |h(A, u) - h(B, u)| \mu(du)$$

es la distancia  $L^1$  entre las funciones soporte de dos conjuntos compactos no vacíos  $A, B \in \mathcal{K}$  y  $\mu$  es la medida de Hausdorff  $(d-1)$ -dimensional normalizada sobre la esfera unidad  $S^{d-1}$  centrada en el origen en  $\mathbb{R}^d$ .

**Teorema 1.3.1.** *El funcional de pares de vectores aleatorios  $m$  es una métrica de probabilidad.*

*Demostración.* Sean  $\xi, \eta$  y  $\zeta$  tres vectores aleatorios. Como  $d_1$  es una distancia entre cuerpos convexos, se satisface  $m(\xi, \eta) = 0$  si y sólo si  $\mathbb{E}(\{\xi\} \cup B_r) = \mathbb{E}(\{\eta\} \cup B_r)$  para todo  $r > 0$ , debido a la continuidad en  $r$  de las esperanzas

anteriores. Se tiene por tanto que  $\xi \leq_{\text{sym}} \eta$  y  $\eta \leq_{\text{sym}} \xi$  y por la antisimetría del orden estocástico simétrico, se cumple que  $\xi$  y  $\eta$  siguen la misma distribución. Además,  $m(\xi, \eta) = m(\eta, \xi)$  y  $m(\xi, \eta) \leq m(\xi, \zeta) + m(\zeta, \eta)$  por ser  $d_1$  una distancia entre cuerpos convexos.  $\square$

### Condición de finitud

Vamos a demostrar que dos vectores aleatorios cualesquiera con segundos momentos finitos son comparables respecto de la métrica  $m$ . En primer lugar daremos una cota para la distancia entre cualquier vector aleatorio y un vector aleatorio degenerado en el origen.

**Proposición 1.3.2.** *La distancia entre el vector aleatorio  $\xi$  y un vector aleatorio degenerado en el origen está acotada por el segundo momento absoluto de  $\xi$  dividido entre 2, esto es,*

$$m(\xi, 0) \leq \frac{1}{2} \mathbb{E} \|\xi\|^2.$$

*Demostración.*

$$\begin{aligned} m(\xi, 0) &= \int_0^\infty d_1(\mathbb{E}(\xi \cup B_r), B_r) dr \leq \int_0^\infty d_H(\mathbb{E}(\xi \cup B_r), B_r) dr \\ &= \int_0^\infty \sup_{u \in S^{d-1}} \mathbb{E}(\langle \xi, u \rangle_+ - r)_+ dr \leq \int_0^\infty \mathbb{E}(\sup_{u \in S^{d-1}} \langle \xi, u \rangle_+ - r)_+ dr \\ &= \int_0^\infty \mathbb{E}(\|\xi\| - r)_+ dr = \int_0^\infty \int_r^\infty (x - r) dP_{\|\xi\|}(x) dr \\ &= \int_0^\infty \int_0^x (x - r) dr dP_{\|\xi\|}(x) = \int_0^\infty \frac{x^2}{2} dP_{\|\xi\|}(x) = \frac{1}{2} \mathbb{E} \|\xi\|^2. \end{aligned}$$

$\square$

Ahora, sin más que aplicar la desigualdad triangular de  $m$ , deducimos que la distancia entre dos vectores cualesquiera con segundos momentos finitos es finita.

**Corolario 1.3.3.** *Si dos vectores aleatorios  $\xi$  y  $\eta$  tienen segundos momentos finitos, entonces  $m(\xi, \eta)$  es finita y*

$$m(\xi, \eta) \leq \frac{1}{2}(\mathbb{E}\|\xi\|^2 + \mathbb{E}\|\eta\|^2).$$

### 1.3.1. Métrica ideal de orden 2

El resultado principal que obtenemos sobre la métrica  $m$  establece que, bajo ciertas condiciones, ésta puede considerarse como una métrica de probabilidad ideal de orden 2. En primer lugar demostraremos que es homogénea de orden 2.

**Teorema 1.3.4.** *La métrica de probabilidad  $m$  es homogénea de orden 2, es decir, para cualesquiera vectores aleatorios  $\xi, \eta$  y cualquier  $c \neq 0$ ,*

$$m(c\xi, c\eta) = |c|^2 m(\xi, \eta).$$

*Demostración.* Sea  $c \neq 0$ ,

$$\begin{aligned} m(c\xi, c\eta) &= \int_0^\infty d_1(\mathbb{E}(c\xi \cup B_r), \mathbb{E}(c\eta \cup B_r)) dr \\ &= \int_0^\infty d_1(c\mathbb{E}(\xi \cup B_{r/|c|}), c\mathbb{E}(\eta \cup B_{r/|c|})) dr \\ &= \int_0^\infty |c| d_1(\mathbb{E}(\xi \cup B_{r/|c|}), \mathbb{E}(\eta \cup B_{r/|c|})) dr \\ &= |c| \int_0^\infty d_1(\mathbb{E}(\xi \cup B_{r/|c|}), \mathbb{E}(\eta \cup B_{r/|c|})) |c| d(r/|c|) \\ &= |c|^2 \int_0^\infty d_1(\mathbb{E}(\xi \cup B_s), \mathbb{E}(\eta \cup B_s)) ds \\ &= |c|^2 m(\xi, \eta). \end{aligned}$$

□

Antes de probar la regularidad de la métrica de probabilidad  $m$ , damos un resultado preliminar



**Lema 1.3.5.** Sean  $\xi, \eta$  dos vectores aleatorios en  $\mathbb{R}^d$ , entonces

$$m(\xi, \eta) = \int_{S^{d-1}} m(\langle \xi, u \rangle, \langle \eta, u \rangle) \mu(du).$$

*Demostración.*

$$\begin{aligned} m(\xi, \eta) &= \int_0^\infty \int_{S^{d-1}} |\mathbb{E}(\langle \xi, u \rangle - r)_+ - \mathbb{E}(\langle \eta, u \rangle - r)_+| \mu(du) dr \\ &= \int_{S^{d-1}} \int_0^\infty |\mathbb{E}(\langle \xi, u \rangle - r)_+ - \mathbb{E}(\langle \eta, u \rangle - r)_+| dr \mu(du) \\ &= \int_{S^{d-1}} \int_0^\infty \frac{1}{2} \left( |\mathbb{E}(\langle \xi, u \rangle - r)_+ - \mathbb{E}(\langle \eta, u \rangle - r)_+| \right. \\ &\quad \left. + |\mathbb{E}(\langle \xi, -u \rangle - r)_+ - \mathbb{E}(\langle \eta, -u \rangle - r)_+| \right) dr \\ &= \int_{S^{d-1}} m(\langle \xi, u \rangle, \langle \eta, u \rangle) \mu(du). \end{aligned}$$

□

Antes de probar la regularidad de la métrica  $m$ , veremos un resultado previo, según el cual la distancia entre la traslación de dos vectores aleatorios se puede acotar en función de la distancia entre estos vectores aleatorios, la longitud de la traslación y la distancia entre sus valores esperados. Su demostración se dividirá en tres partes, en primer lugar consideraremos el caso de variables aleatorias a las que sumamos un número positivo, a continuación el caso en que sumamos un número negativo a dos variables aleatorias y finalmente el caso general para vectores aleatorios en  $\mathbb{R}^d$  a los que se les suma un punto de  $\mathbb{R}^d$  cualquiera.

**Lema 1.3.6.** Sean  $\xi$  y  $\eta$  dos vectores aleatorios en  $\mathbb{R}^d$  con primer momento finito, y sea  $a \in \mathbb{R}^d$ , entonces

$$m(\xi + a, \eta + a) \leq m(\xi, \eta) + \frac{\|a\|}{2} \|\mathbb{E}\xi - \mathbb{E}\eta\|.$$

*Demostración.* Sea  $d = 1$  y  $a > 0$ , entonces

$$\begin{aligned}
m(\xi + a, \eta + a) &= \int_0^\infty \frac{1}{2} \left( |\mathbb{E}(\xi - (r - a))_+ - \mathbb{E}(\eta - (r - a))_+| \right. \\
&\quad \left. + |\mathbb{E}(\xi + (r + a))_- - \mathbb{E}(\eta + (r + a))_-| \right) dr \\
&= m(\xi, \eta) + \frac{1}{2} \int_{-a}^0 |\mathbb{E}(\xi - l)_+ - \mathbb{E}(\eta - l)_+| dl \\
&\quad - \frac{1}{2} \int_0^a |\mathbb{E}(\xi + l)_- - \mathbb{E}(\eta + l)_-| dl \\
&= m(\xi, \eta) + \frac{1}{2} \int_{-a}^0 |\mathbb{E}\xi - \mathbb{E}\eta + \mathbb{E}(\xi - l)_- - \mathbb{E}(\eta - l)_-| dl \\
&\quad + \frac{1}{2} \int_0^{-a} |\mathbb{E}(\xi - t)_- - \mathbb{E}(\eta - t)_-| dt \\
&= m(\xi, \eta) + \frac{1}{2} \int_{-a}^0 |\mathbb{E}\xi - \mathbb{E}\eta + \mathbb{E}(\xi - l)_- - \mathbb{E}(\eta - l)_-| dl \\
&\quad - \frac{1}{2} \int_{-a}^0 |\mathbb{E}(\xi - t)_- - \mathbb{E}(\eta - t)_-| dt \\
&\leq m(\xi, \eta) + \frac{a}{2} |\mathbb{E}\xi - \mathbb{E}\eta|.
\end{aligned}$$

Si  $a < 0$ , entonces

$$\begin{aligned}
m(\xi + a, \eta + a) &= m(-\xi - a, -\eta - a) \\
&\leq m(-\xi, -\eta) + \frac{|a|}{2} |-\mathbb{E}\xi + \mathbb{E}\eta| \\
&= m(\xi, \eta) + \frac{|a|}{2} |\mathbb{E}\xi - \mathbb{E}\eta|.
\end{aligned}$$

Sea  $d \geq 2$  y  $a \in \mathbb{R}^d$ , entonces

$$\begin{aligned}
m(\xi + a, \eta + a) &= \int_{S^{d-1}} m(\langle \xi, u \rangle + \langle a, u \rangle, \langle \eta, u \rangle + \langle a, u \rangle) \mu(du) \\
&\leq \int_{S^{d-1}} \left( m(\langle \xi, u \rangle, \langle \eta, u \rangle) + \frac{|\langle a, u \rangle|}{2} |\mathbb{E}\langle \xi, u \rangle - \mathbb{E}\langle \eta, u \rangle| \right) \mu(du) \\
&\leq \int_{S^{d-1}} \left( m(\langle \xi, u \rangle, \langle \eta, u \rangle) + \frac{\|a\|}{2} \|\mathbb{E}\xi - \mathbb{E}\eta\| \right) \mu(du) \\
&= m(\xi, \eta) + \frac{\|a\|}{2} \|\mathbb{E}\xi - \mathbb{E}\eta\|.
\end{aligned}$$

□

Si en vez de sumar un punto fijo de  $\mathbb{R}^d$  sumamos un vector aleatorio independiente de aquellos que estamos considerando, obtenemos un resultado similar.

**Lema 1.3.7.** Sean  $\xi, \eta, \zeta$  vectores aleatorios en  $\mathbb{R}^d$  tales que  $\zeta$  es independiente de  $\xi$  y  $\eta$  y éstos tienen primer momento finito, entonces

$$m(\xi + \zeta, \eta + \zeta) \leq m(\xi, \eta) + \frac{\mathbb{E}\|\zeta\|}{2} \|\mathbb{E}\xi - \mathbb{E}\eta\|.$$

*Demostración.*

$$\begin{aligned}
m(\xi + \zeta, \eta + \zeta) &= \int_{S^{d-1}} m(\langle \xi, u \rangle + \langle \zeta, u \rangle, \langle \eta, u \rangle + \langle \zeta, u \rangle) \mu(du) \\
&\leq \int_{\mathbb{R}^d} \left( m(\xi, \eta) + \frac{\|x\|}{2} \|\mathbb{E}\xi - \mathbb{E}\eta\| \right) dP_\zeta(x) \\
&= m(\xi, \eta) + \frac{\mathbb{E}\|\zeta\|}{2} \|\mathbb{E}\xi - \mathbb{E}\eta\|.
\end{aligned}$$

□

El Lema 1.3.7 nos indica hasta qué punto la métrica  $m$  es una métrica de probabilidad regular.

**Corolario 1.3.8.** La métrica de probabilidad  $m$  es regular cuando mide la distancia entre vectores aleatorios con la misma esperanza finita, es decir, si  $\mathbb{E}\xi = \mathbb{E}\eta$ , entonces  $m(\xi + \zeta, \eta + \zeta) \leq m(\xi, \eta)$  para cualquier vector aleatorio  $\zeta$  independiente de  $\xi$  y  $\eta$ .

Finalmente, el Teorema 1.3.4 y el Corolario 1.3.8 demuestran que, bajo ciertas condiciones, la métrica  $m$  es una métrica de probabilidad ideal.

**Corolario 1.3.9.** *La métrica de probabilidad  $m$  considerada sobre vectores aleatorios con la misma esperanza es una métrica de probabilidad ideal de orden 2.*

### 1.3.2. Relación con otras métricas de probabilidad

Para finalizar esta sección y el capítulo relacionaremos la métrica de probabilidad  $m$  con la *métrica ideal de Zolotarev de orden 2* definida por Zolotarev (1983).

*Definición 1.3.4.* Para cualquier par de variables aleatorias  $(\xi, \eta)$ , su *métrica ideal de Zolotarev de orden 2* se define como,

$$Z_2(\xi, \eta) = \int_{-\infty}^{\infty} |\mathbb{E}(\xi - r)_+ - \mathbb{E}(\eta - r)_+| dr.$$

La métrica de probabilidad  $m$  está muy relacionada con la métrica ideal de Zolotarev como se pone de manifiesto en el siguiente resultado.

**Proposición 1.3.10.** *Sean  $\xi$  y  $\eta$  dos variables aleatorias, entonces*

- i.  $m(\xi, \eta) \leq \frac{1}{2}(Z_2(\xi, \eta) + Z_2(-\xi, -\eta))$ ;*
- ii. si  $\mathbb{E}\xi = \mathbb{E}\eta$  finita, entonces  $m(\xi, \eta) = \frac{1}{2}Z_2(\xi, \eta)$ .*

*Demostración.*

$$\begin{aligned} m(\xi, \eta) &= \int_0^{\infty} \frac{1}{2} \left( |\mathbb{E}(\xi - r)_+ - \mathbb{E}(\eta - r)_+| + |\mathbb{E}(\xi + r)_- - \mathbb{E}(\eta + r)_-| \right) dr \\ &= \int_0^{\infty} \frac{1}{2} \left( |\mathbb{E}(\xi - r)_+ - \mathbb{E}(\eta - r)_+| + |\mathbb{E}(-\xi - r)_+ - \mathbb{E}(-\eta - r)_+| \right) dr \\ &\leq \int_{-\infty}^{\infty} \frac{1}{2} \left( |\mathbb{E}(\xi - r)_+ - \mathbb{E}(\eta - r)_+| + |\mathbb{E}(-\xi - r)_+ - \mathbb{E}(-\eta - r)_+| \right) dr \\ &= \frac{1}{2} (Z_2(\xi, \eta) + Z_2(-\xi, -\eta)). \end{aligned}$$

Esto demuestra la primera parte del enunciado. Sean ahora  $\mathbb{E}\xi = \mathbb{E}\eta$ , entonces

$$\begin{aligned}
m(\xi, \eta) &= \int_0^\infty \frac{1}{2} \left( |\mathbb{E}(\xi - r)_+ - \mathbb{E}(\eta - r)_+| + |\mathbb{E}(\xi + r)_- - \mathbb{E}(\eta + r)_-| \right) dr \\
&= \frac{1}{2} \int_0^\infty |\mathbb{E}(\xi - r)_+ - \mathbb{E}(\eta - r)_+| dr \\
&\quad + \frac{1}{2} \int_0^\infty |-\mathbb{E}\xi + \mathbb{E}\eta + \mathbb{E}(\xi + r)_+ - \mathbb{E}(\eta + r)_+| dr \\
&= \frac{1}{2} \int_0^\infty |\mathbb{E}(\xi - r)_+ - \mathbb{E}(\eta - r)_+| dr \\
&\quad - \frac{1}{2} \int_0^{-\infty} |\mathbb{E}(\xi - t)_+ - \mathbb{E}(\eta - t)_+| dt \\
&= \frac{1}{2} Z_2(\xi, \eta).
\end{aligned}$$

□

La regularidad de la métrica de probabilidad  $m$  para variables aleatorias podría deducirse fácilmente del enunciado *ii.* de la Proposición 1.3.10 y del hecho de que  $Z_2$  es una métrica de probabilidad ideal.

## 1.4. Comentarios y problemas abiertos

Hemos definido un orden estocástico convexo para vectores aleatorios que es más débil que el orden lineal convexo y por lo tanto permite comparar algunos pares de vectores aleatorios con distinta esperanza según su variabilidad. Gracias a las caracterizaciones de varios órdenes estocásticos convexos mediante la inclusión de las esperanzas de Aumann de ciertos conjuntos aleatorios, hemos extendido estos órdenes para comparar conjuntos aleatorios según su variabilidad. Además, hemos justificado estas relaciones a partir de las ordenaciones de ciertos pares de selecciones de los conjuntos aleatorios. Finalmente, se utiliza el desarrollo construido para trabajar con el orden simétrico para definir una nueva distancia entre distribuciones de probabilidad que es una métrica ideal para la suma de vectores aleatorios.

Todos los órdenes estocásticos para conjuntos aleatorios que se definen en la Sección 1.2 se construyen a partir relaciones de contenido de esperanzas de Aumann de conjuntos aleatorios que se definen como uniones de conjuntos. En el futuro tenemos pensado estudiar la posibilidad de definir un orden estocástico a partir de relaciones entre esperanzas de Aumann de conjuntos aleatorios contruidos como intersecciones de conjuntos.

Finalmente, deseamos seguir estudiando la métrica definida en la Sección 1.3 para obtener nuevas cotas para ella y relaciones con otras métricas. Nuestro objetivo es utilizarla en la construcción de teoremas límite para la suma de vectores aleatorios, donde se podría explotar el hecho de que se trata de una métrica ideal. Además, gracias a su relación con el orden simétrico, bajo ciertas condiciones, podrían obtenerse sucesiones convergentes de vectores aleatorios ordenados. En definitiva, deseamos seguir la línea de trabajo propuesta por Boutsikas y Vaggelatou (2002).

## Capítulo 2

# Índices de desigualdad para conjuntos aleatorios

En el Capítulo 1 se ha descrito un nuevo orden en variabilidad para vectores aleatorios y al mismo tiempo se ha analizado el comportamiento de la extensión de varios órdenes en variabilidad de vectores a conjuntos aleatorios. Los órdenes estocásticos son relaciones de orden *parcial*, no podemos por tanto utilizarlos para comparar dos distribuciones de probabilidad cualesquiera ya sean de vectores aleatorios o de conjuntos aleatorios. Para solventar este problema se utilizan los índices de desigualdad. Éstos cuantifican la variabilidad asociada a un elemento aleatorio y así nos permiten ordenar dichos elementos según su variabilidad sin más que recurrir al orden del valor de sus índices en  $\mathbb{R}$ .

En origen los índices de desigualdad se construyeron para medir la variabilidad o dispersión relativa que presenta una variable aleatoria positiva. Con el tiempo han aparecido versiones de índices de desigualdad que nos permiten medir la variabilidad de variables aleatorias que toman valores en toda la recta real, e incluso vectores aleatorios. Recientemente, a partir del estudio de variables que vienen dadas en términos de fluctuaciones o rangos y son modeladas como conjuntos aleatorios, han aparecido índices de desigualdad

para estudiar la variabilidad de dichos conjuntos. Estos índices de desigualdad para conjuntos aleatorios no toman necesariamente un valor numérico sino que a veces se construyen de modo que tomen como valor un intervalo.

Los índices de desigualdad se emplean con frecuencia en estudios de Ciencias Sociales, Economía o Industria. Con ellos se compara la desigualdad en las distribuciones de ingresos o riqueza en diferentes poblaciones y se cuantifica la concentración industrial. Para entender mejor cómo medir la desigualdad, podemos destacar, entre otros, los trabajos de Eichhorn y Gehrig (1982), Atkinson (1983) y Sen (1995).

A lo largo de este capítulo, en primer lugar extendemos una familia de índices de desigualdad, los  $f$ -índices de desigualdad, a vectores e intervalos aleatorios. Para extenderlos a vectores aleatorios se toma el índice de desigualdad de una transformación lineal creciente de los vectores aleatorios. En el caso de los intervalos aleatorios, se estudian índices de desigualdad que toman valor de intervalo (Cascos y López-Díaz, 2003). Finalmente, el índice de desigualdad de un conjunto aleatorio es el índice de desigualdad de un intervalo aleatorio que se obtiene como transformación lineal de dicho conjunto. Se estudia la continuidad de tales índices de desigualdad y se da una cota de la distancia entre los índices de desigualdad de dos conjuntos aleatorios. El capítulo concluye con un ejemplo de aplicación práctica de dichos resultados (Cascos *et al.*, 2004).

## 2.1. Preliminares

A lo largo de este capítulo se estudia un caso particular de conjuntos, los intervalos compactos reales. Denotamos por  $\mathcal{K}_c(\mathbb{R})$  a la familia de subconjuntos compactos y convexos de  $\mathbb{R}$ , es decir a la familia de intervalos cerrados y acotados. A un intervalo  $A \in \mathcal{K}_c(\mathbb{R})$  lo denotamos como  $A = [\text{mín } A, \text{máx } A]$ . Dados  $A, B \in \mathcal{K}_c(\mathbb{R})$  y  $\lambda \in \mathbb{R}_+$ , las operaciones entre intervalos con las que trabajaremos son la suma de Minkowski, el producto por un escalar positivo y el cociente de intervalos. Siempre que  $B$  no contenga al origen,  $0 \notin B$ ,



definimos el cociente de  $A$  entre  $B$  como el conjunto de cocientes de elementos de  $A$  entre elementos de  $B$ . Merced a la aritmética de intervalos estas operaciones toman expresiones muy sencillas,

$$A + B = \{a + b : a \in A, b \in B\} = [\text{mín } A + \text{mín } B, \text{máx } A + \text{máx } B];$$

$$\lambda A = \{\lambda a : a \in A\} = [\lambda \text{mín } A, \lambda \text{máx } A], \quad \lambda > 0;$$

$$A/B = \{a/b : a \in A, b \in B\} = [\text{mín } A/\text{máx } B, \text{máx } A/\text{mín } B], \quad \text{mín } B > 0.$$

Además, la distancia de Hausdorff también adopta una expresión muy sencilla para intervalos compactos,

$$d_H(A, B) = \text{máx}\{|\text{mín } A - \text{mín } B|, |\text{máx } A - \text{máx } B|\}.$$

Dada una función  $g : \mathbb{R}^d \longrightarrow \mathbb{R}$  y un conjunto  $A \subset \mathbb{R}^d$ , denotamos por  $g(A)$  al conjunto de todas las transformaciones por  $g$  de los elementos de  $A$ , así

$$g(A) := \{g(a) : a \in A\}.$$

Dados dos intervalos  $A$  y  $B$ , decimos que el intervalo  $A$  es menor o igual que el intervalo  $B$  y lo denotamos  $A \leq B$  si  $\text{mín } A \leq \text{mín } B$  y  $\text{máx } A \leq \text{máx } B$ .

Un *intervalo aleatorio compacto* será también un caso particular de conjunto aleatorio compacto. Así, si  $(\Omega, \mathcal{A}, \text{Pr})$  es un espacio de probabilidad, decimos que  $X : \Omega \longrightarrow \mathcal{K}_c(\mathbb{R})$  es un *intervalo aleatorio compacto* si es medible respecto de la  $\sigma$ -álgebra de Borel asociada a la topología generada por  $d_H$  en  $\mathcal{K}_c(\mathbb{R})$ . La medibilidad de esta aplicación es equivalente a la del vector aleatorio  $(\text{mín } X, \text{máx } X)$ , véase por ejemplo Cascos (2001).

Al igual que la distancia de Hausdorff adopta una expresión muy sencilla para el caso de intervalos compactos, también la esperanza de Aumann de un intervalo aleatorio toma una expresión muy sencilla. Siempre que  $\text{mín } X$  y  $\text{máx } X$  sean integrables, se cumple

$$\mathbb{E}X = [\mathbb{E} \text{mín } X, \mathbb{E} \text{máx } X].$$

A continuación definiremos espacios de vectores y conjuntos aleatorios.

Dado  $p \in [1, +\infty]$ , el espacio  $L^p(\text{Pr})$  – más brevemente  $L^p$  – de vectores aleatorios en  $\mathbb{R}^d$  viene dado por

$$L^p := \{\xi : \Omega \longrightarrow \mathbb{R}^d : \xi \text{ vector aleatorio, } \|\xi\|_p < \infty\}$$

y su distancia asociada es  $d^p(\xi, \eta) := \|\xi - \eta\|_p$ , donde  $\|\cdot\|_p$  denota a la norma  $L^p$  de un vector aleatorio, es decir,  $\|\xi\|_p := (\mathbb{E}\|\xi\|^p)^{1/p}$  si  $p \in [1, +\infty)$  y si  $p = \infty$ ,  $\|\xi\|_\infty := \text{ess sup } \|\xi\|$  es su supremo esencial. Al espacio  $L^p$  de vectores aleatorios en  $(0, +\infty)^d$  lo denotamos por  $L_+^p$ . Obsérvese que esta notación no especifica la dimensión en la que trabajamos, así si consideramos variables aleatorias,  $d = 1$ , utilizamos la misma notación.

El espacio  $\mathcal{L}^p$  de conjuntos aleatorios compactos y convexos se define como

$$\mathcal{L}^p := \{X : \Omega \longrightarrow \mathcal{K}_c : X \text{ conjunto aleatorio, } \|X\|_p < \infty\},$$

donde la norma  $\mathcal{L}^p$  de un conjunto aleatorio es  $\|X\|_p = \mathbb{E}(\|X\|^p)^{1/p}$ , siendo  $\|X\|$  la variable aleatoria magnitud del conjunto aleatorio  $X$  y su distancia asociada se define como

$$\mathcal{D}^p(X, Y) := \|d_H(X, Y)\|_p.$$

La distancia de Hausdorff entre dos conjuntos aleatorios compactos  $X, Y : \Omega \longrightarrow \mathcal{K}$  es medible como se explica en la Sección 0.2 y en consecuencia la distancia  $\mathcal{D}^p$  está bien definida. Al espacio  $\mathcal{L}^p$  de conjuntos aleatorios que toman como valores subconjuntos de  $(0, +\infty)^d$ , lo denotamos por  $\mathcal{L}_+^p$ . Si consideramos intervalos aleatorios,  $d = 1$ , como de nuevo la notación no especifica la dimensión en la que trabajamos, hablaremos de los espacios  $\mathcal{L}^p$  y  $\mathcal{L}_+^p$  de intervalos aleatorios. Dado  $X$  un intervalo aleatorio,  $X \in \mathcal{L}^p$  si y sólo si  $\text{mín } X, \text{máx } X \in L^p$  y si el intervalo aleatorio es positivo, es decir, toma como valores subconjuntos de  $(0, +\infty)$ , entonces  $X \in \mathcal{L}_+^p$  si y sólo si  $\text{máx } X \in L^p$ .

## 2.2. Índices de desigualdad

Los índices de desigualdad cuantifican la desigualdad o dispersión, habitualmente relativa, que presenta un elemento aleatorio. Se suelen calcular para variables aleatorias positivas.

El  $f$ -índice de desigualdad (*generalizado*) está inspirado en la divergencia generalizada de Csiszár entre dos distribuciones de probabilidad (Csiszár, 1967) y ha sido estudiado, entre otros, por Cascos *et al.* (2002). Se define para una función convexa  $f : (0, +\infty) \rightarrow \mathbb{R}$  tal que  $f(1) = 0$  y una variable aleatoria positiva  $\xi$  como el valor real

$$i_f(\xi) := \mathbb{E}f\left(\frac{\xi}{\mathbb{E}\xi}\right)$$

siempre que tal valor exista. En algunos trabajos se supone la convexidad estricta o monotonía de  $f$ .

Ejemplos particulares de  $f$ -índices de desigualdad son los *índices de desigualdad aditivamente descomponibles* de Bourguignon (1979) cuyo índice de orden  $\alpha \in \mathbb{R} \setminus \{0, 1\}$  está inducido por

$$f_\alpha(x) = \frac{1}{\alpha(\alpha - 1)}(x^\alpha - 1).$$

Nosotros consideraremos  $(f, l)$ -índices de desigualdad denotados por  $i_{f,l}$  definidos sobre vectores aleatorios en  $(0, +\infty)^d$  donde  $l$  es una aplicación lineal estrictamente creciente de  $\mathbb{R}^d$  en  $\mathbb{R}$  y  $f$  es, como anteriormente, una función convexa  $f : (0, +\infty) \rightarrow \mathbb{R}$  tal que  $f(1) = 0$ . Estos índices de desigualdad asignan a un vector aleatorio positivo el  $f$ -índice de desigualdad de la variable aleatoria positiva  $l(\xi)$ ,

$$i_{f,l}(\xi) := i_f(l(\xi)) = \mathbb{E}f\left(\frac{l(\xi)}{l(\mathbb{E}\xi)}\right).$$

Podemos extender las definiciones anteriores a conjuntos aleatorios compactos y convexos sin más que interpretar la esperanza como esperanza de

Aumann y el cociente como cociente de intervalos, de un modo similar al empleado por López-García *et al.* (2000) para extender otra familia de índices de desigualdad a los intervalos aleatorios.

*Definición 2.2.1.* Si  $f : (0, +\infty) \rightarrow \mathbb{R}$  es una función convexa tal que  $f(1) = 0$  y  $X \in \mathcal{L}_+^1$ , su  $f$ -índice de desigualdad con valor de intervalo se define como,

$$I_f(X) := \mathbb{E}f\left(\frac{X}{\mathbb{E}X}\right)$$

siempre y cuando esa esperanza exista.

Del mismo modo que se extienden los  $f$ -índices de desigualdad para variables aleatorias a índices de desigualdad para vectores aleatorios utilizando una transformación lineal estrictamente creciente de  $\mathbb{R}^d$  en  $\mathbb{R}$ , podemos extender los  $f$ -índices de desigualdad para intervalos aleatorios a índices de desigualdad para conjuntos aleatorios compactos y convexos.

*Definición 2.2.2.* Si  $l : \mathbb{R}^d \rightarrow \mathbb{R}$  es una función lineal estrictamente creciente,  $f : (0, +\infty) \rightarrow \mathbb{R}$  una función convexa tal que  $f(1) = 0$  y  $X$  es un conjunto aleatorio compacto y convexo en  $\mathbb{R}^d$  tal que  $X \in \mathcal{L}_+^1$ , su  $(f, l)$ -índice de desigualdad con valor de intervalo se define como,

$$I_{f,l}(X) := I_f(l(X)) = \mathbb{E}f\left(\frac{l(X)}{l(\mathbb{E}X)}\right).$$

siempre y cuando esa esperanza exista.

No supondría ningún inconveniente extender estos índices a conjuntos aleatorios compactos cualesquiera, no necesariamente convexos sin más que considerar la envolvente convexa de los conjuntos aleatorios.

Para finalizar esta sección es conveniente destacar que todos los conceptos sobre índices de desigualdad que se acaban de definir son generalización unos de otros. En primer lugar, dado un intervalo aleatorio  $X$ , y una función convexa  $f$  en las condiciones apropiadas, se tiene que  $I_f(X) = I_{f,\text{id}}(X)$ , siendo correcta la última expresión por la linealidad de la aplicación identidad. Además, el  $f$ -índice de desigualdad del intervalo aleatorio unipuntual

formado por una variable aleatoria es el conjunto unipuntual formado por el  $f$ -índice de la variable aleatoria y finalmente el  $(f, l)$ -índice de desigualdad del conjunto aleatorio unipuntual formado por un vector aleatorio es el conjunto unipuntual formado por el  $(f, l)$ -índice del vector aleatorio.

## 2.3. Propiedades

Para obtener propiedades interesantes de los índices de desigualdad, a lo largo de esta sección suponemos que la función  $f$  es monótona y estrictamente convexa. La transformación lineal  $l_u$  vendrá determinada por  $u \in (0, +\infty)^d$ , es decir  $l_u(x) = \langle x, u \rangle$ . Además, a lo largo de toda esta sección supondremos que los conjuntos aleatorios tienen índices de desigualdad acotados. En estas condiciones el  $(f, l_u)$ -índice de desigualdad de un conjunto aleatorio  $X \in \mathcal{L}_+^1$  adopta una forma mucho más simple. Si  $f$  es creciente,

$$\begin{aligned} I_{f, l_u}(X) &= \left[ \mathbb{E}f\left(\frac{\min l_u(X)}{\max l_u(\mathbb{E}X)}\right), \mathbb{E}f\left(\frac{\max l_u(X)}{\min l_u(\mathbb{E}X)}\right) \right] \\ &= \left[ \mathbb{E}f\left(\frac{-h(X, -u)}{\mathbb{E}h(X, u)}\right), \mathbb{E}f\left(\frac{h(X, u)}{-\mathbb{E}h(X, -u)}\right) \right], \end{aligned}$$

mientras que si  $f$  es decreciente,

$$\begin{aligned} I_{f, l_u}(X) &= \left[ \mathbb{E}f\left(\frac{\max l_u(X)}{\min l_u(\mathbb{E}X)}\right), \mathbb{E}f\left(\frac{\min l_u(X)}{\max l_u(\mathbb{E}X)}\right) \right] \\ &= \left[ \mathbb{E}f\left(\frac{h(X, u)}{-\mathbb{E}h(X, -u)}\right), \mathbb{E}f\left(\frac{-h(X, -u)}{\mathbb{E}h(X, u)}\right) \right]. \end{aligned}$$

Antes de estudiar las propiedades de la familia de  $(f, l)$ -índices como índices de desigualdad, damos dos resultados previos.

**Lema 2.3.1.** *Dado un intervalo aleatorio positivo  $X$  y una función  $f$  monótona y estrictamente convexa, entonces  $I_f(X)$  es un único punto si y sólo si  $X = \{\xi\}$  c.s. para cierta variable aleatoria  $\xi$ .*

*Demostración.* Es claro que si  $X = \{\xi\}$  c.s., entonces  $I_f(X) = \{i_f(\xi)\}$ , esto es, se trata de un conjunto unipuntual. Veamos que también se satisface la implicación contraria.

Supongamos que no existe ninguna variable aleatoria  $\xi$  tal que  $X = \{\xi\}$  c.s., entonces  $\mathbb{E} \min X < \mathbb{E} \max X$ . Si  $f$  es creciente, como además es estrictamente convexa, el crecimiento será estricto y así, por la desigualdad anterior, se satisface que  $f(x/\mathbb{E} \max X) < f(x/\mathbb{E} \min X)$ , como además  $\min X \leq \max X$ , de nuevo por ser  $f$  creciente, tenemos que  $\mathbb{E}f(\min X/\mathbb{E} \max X) < \mathbb{E}f(\max X/\mathbb{E} \min X)$  y por lo tanto  $I_f(X)$  no será un conjunto unipuntual. Si  $f$  fuese decreciente, el razonamiento sería análogo.  $\square$

Este resultado se puede generalizar al caso de los conjuntos aleatorios. Diremos que *un conjunto aleatorio  $X$  es normal a  $u \in \mathbb{R}^d$*  si cada una de sus realizaciones está contenida en un hiperplano normal a  $u$ , es decir

$$\Pr(\{\omega : X(\omega) \subset H_\omega \text{ con } H_\omega \text{ hiperplano normal a } u\}) = 1.$$

**Proposición 2.3.2.** *Dado un conjunto aleatorio  $X \in \mathcal{L}_+^1$ , una función  $f$  monótona y estrictamente convexa y otra función  $l_u$  lineal estrictamente creciente tal que  $X$  no es normal a  $u$ , entonces  $I_{f,l_u}(X)$  es un único punto si y sólo si  $X = \{\xi\}$  c.s. para cierto vector aleatorio  $\xi$ .*

A continuación se enuncian algunas de las propiedades de los  $(f, l)$ -índices de desigualdad para conjuntos aleatorios, cuyas demostraciones se obtienen de manera inmediata.

### No negatividad promedio

**Proposición 2.3.3.** *Los índices  $I_{f,l}$  satisfacen la no negatividad promedio, es decir, para cualquier conjunto aleatorio  $X$ , si  $f$  es creciente,*

$$\min I_{f,l}(X) \mathbb{E} \max l(X) + \max I_{f,l}(X) \mathbb{E} \min l(X) \geq 0,$$

*mientras que si  $f$  es decreciente,*

$$\max I_{f,l}(X) \mathbb{E} \max l(X) + \min I_{f,l}(X) \mathbb{E} \min l(X) \geq 0.$$

### Minimalidad

**Proposición 2.3.4.** *Dado un conjunto aleatorio  $X$ , se cumple  $I_{f,l_u}(X) = 0$  si y sólo si  $X$  es degenerado en algún punto de  $\mathbb{R}^d$  o está completamente distribuido en un hiperplano normal a  $u$ .*

### Independencia de la media

**Proposición 2.3.5.** *Para cualquier  $\lambda > 0$ , se cumple  $I_{f,l}(\lambda X) = I_{f,l}(X)$ .*

### Traslaciones

**Proposición 2.3.6.** *Dado  $x \in \mathbb{R}_+^d$ , se cumple  $I_{f,l}(X + x) \leq I_{f,l}(X)$ .*

### Relación con los órdenes estocásticos en variabilidad

Los  $(f, l)$ -índices de desigualdad para conjuntos aleatorios están relacionados con los órdenes estocásticos según la variabilidad para conjuntos aleatorios que se han definido en la Sección 1.2, es decir, con el orden simétrico, con el lineal convexo y con los órdenes superior lineal convexo e inferior lineal convexo.

**Proposición 2.3.7.** *Dados dos conjuntos aleatorios  $X, Y \in \mathcal{L}_+^1$  tales que  $\mathbb{E}X = \mathbb{E}Y$ , los tres siguientes enunciados son equivalentes:*

- i.  $X \leq_{\text{uilcx}} Y$ ;*
- ii.  $\max I_{f,l}(X) \leq \max I_{f,l}(Y)$  para cualquier  $f$  creciente;*
- iii.  $\min I_{f,l}(X) \leq \min I_{f,l}(Y)$  para cualquier  $f$  decreciente.*

*Demostración.* Se tiene que  $\mathbb{E}X = \mathbb{E}Y$ , así según el Teorema 0.5.2 y la Proposición 1.2.14, se cumple  $X \leq_{\text{uilcx}} Y$  si y sólo si  $h(X, u) \leq_{\text{cx}} h(Y, u)$  para cualquier  $u \in \mathbb{R}_+^d$  o, equivalentemente,  $\max l_u(X) \leq_{\text{cx}} \max l_u(Y)$  para cualquier  $l_u$  lineal creciente.

Si  $f$  es creciente, se cumple

$$\max I_{f,l_u}(X) = \mathbb{E}f\left(\frac{\max l_u(X)}{\mathbb{E} \min l_u(X)}\right) = \mathbb{E}f\left(\frac{h(X, u)}{-\mathbb{E}h(X, -u)}\right).$$

Como la función  $f$  es convexa, también lo es  $g(x) = f(x/\mathbb{E} \min l_u(X))$ . La observación de que las esperanzas son iguales  $\mathbb{E}X = \mathbb{E}Y$  finaliza la demostración de la equivalencia entre *i.* y *ii.* La equivalencia con *iii.* se demuestra de un modo similar.  $\square$

**Proposición 2.3.8.** *Dados dos conjuntos aleatorios  $X, Y \in \mathcal{L}_+^1$  tales que  $\mathbb{E}X = \mathbb{E}Y$ , los tres siguientes enunciados son equivalentes:*

- i.*  $X \leq_{\text{lilcx}} Y$ ;
- ii.*  $\min I_{f,l}(X) \leq \min I_{f,l}(Y)$  para cualquier  $f$  creciente;
- iii.*  $\max I_{f,l}(X) \leq \max I_{f,l}(Y)$  para cualquier  $f$  decreciente.

*Demostración.* La demostración sigue la misma estructura que la de la Proposición 2.3.7.  $\square$

Como consecuencia de las Proposiciones 2.3.7 y 2.3.8 obtenemos el siguiente resultado.

**Corolario 2.3.9.** *Dados dos conjuntos aleatorios  $X, Y \in \mathcal{L}_+^1$  tales que  $\mathbb{E}X = \mathbb{E}Y$ , entonces  $X \leq_{\text{uilcx}} Y$  e  $Y \leq_{\text{lilcx}} X$  si y sólo si  $I_{f,l}(X) \leq I_{f,l}(Y)$  para cualesquiera  $f$  y  $l$ .*

A partir de la Proposición 1.2.16 y del Corolario 2.3.9 el siguiente resultado es inmediato.

**Corolario 2.3.10.** *Dados dos conjuntos aleatorios  $X, Y \in \mathcal{L}_+^1$  tales que  $\mathbb{E}X = \mathbb{E}Y$  y  $X \leq_{\text{slcx}} Y$ , entonces  $I_{f,l}(X) \leq I_{f,l}(Y)$  para cualquier  $f$ .*



Antes de obtener el último resultado que relaciona índices de desigualdad con órdenes estocásticos para conjuntos aleatorios, presentamos una relación entre el orden simétrico y el superior creciente lineal convexo para conjuntos aleatorios que utilizaremos en dicho resultado.

**Lema 2.3.11.** *Dados dos conjuntos aleatorios compactos  $X$  e  $Y$  que toman como valores subconjuntos de  $(0, +\infty)^d$  y tales que  $X \leq_{\text{sym}} Y$ , entonces  $X \leq_{\text{uilcx}} Y$ .*

*Demostración.* Sea  $u \in \mathbb{R}_+^d$ , como  $X \leq_{\text{sym}} Y$ , según el Teorema 1.2.1 se satisface la relación  $h(X, u)_+ \leq_{\text{sym}} h(Y, u)_+$  o equivalentemente  $h(X, u)_+ \leq_{\text{icx}} h(Y, u)_+$ . Como  $X$  e  $Y$  toman valores en  $(0, +\infty)^d$ , la función soporte en la dirección de  $u$  toma siempre valores positivos, con lo que tenemos  $h(X, u) \leq_{\text{icx}} h(Y, u)$  o equivalentemente según la Proposición 1.2.14, se tiene  $X \leq_{\text{uilcx}} Y$ .  $\square$

Sobre la base de la Proposición 2.3.7 y del Lema 2.3.11 se obtiene el siguiente resultado.

**Corolario 2.3.12.** *Dados dos conjuntos aleatorios  $X, Y \in \mathcal{L}_+^1$  tales que  $\mathbb{E}X = \mathbb{E}Y$  y se satisface la relación  $X \leq_{\text{sym}} Y$ , entonces  $\max I_{f,l}(X) \leq \max I_{f,l}(Y)$  para cualquier  $f$  creciente y  $\min I_{f,l}(X) \leq \min I_{f,l}(Y)$  para cualquier  $f$  decreciente.*

## 2.4. Criterios de convergencia

En esta sección se estudiará en primer lugar la continuidad de los  $(f, l)$ -índices de desigualdad como aplicaciones del espacio de los conjuntos aleatorios compactos y convexos positivos  $\mathcal{L}_+^p$  y luego se buscarán cotas a la distancia entre los índices de desigualdad de dos conjuntos aleatorios expresadas en función de la distancia entre dichos conjuntos. Este desarrollo tiene por objetivo construir un marco adecuado para el ejemplo que aparece en la Sección 2.5 y con el que se justificarán los resultados obtenidos.

### 2.4.1. Continuidad

El resultado principal sobre continuidad de los índices de desigualdad para conjuntos aleatorios es el Corolario 2.4.4. Se obtiene como generalización inmediata de la continuidad de los índices de desigualdad para intervalos aleatorios que se demuestra en el Teorema 2.4.3.

Antes de demostrar la continuidad de los índices de desigualdad, necesitamos dos resultados auxiliares. Denotaremos por  $O(\cdot)$  a la “ $O$  mayúscula” de Landau, es decir dadas dos funciones  $f, g$ , escribimos ‘ $f(x) = O(g(x))$  cuando  $x \rightarrow +\infty$ ’ si existen constantes  $C, k, M$  tales que  $|f(x)| \leq C + k|g(x)|$  para todo  $x > M$ .

**Lema 2.4.1.** *Sea  $p \in [1, +\infty)$  y  $g : \mathbb{R} \rightarrow \mathbb{R}$  una función continua tal que  $g(x) = O(|x|^p)$  cuando  $|x| \rightarrow +\infty$ . Si  $\|\xi_n - \xi\|_p \rightarrow 0$  donde  $\xi, \xi_n \in L^p$  para todo  $n \in \mathbb{N}$  son variables aleatorias, entonces  $\|g(\xi_n) - g(\xi)\|_1 \rightarrow 0$ .*

*Demostración.* Si  $\|\xi_n - \xi\|_p \rightarrow 0$  con  $\xi, \xi_n \in L^p$  para todo  $n \in \mathbb{N}$ , entonces la convergencia  $L^p$  de  $\{\xi_n\}_n$  implica su convergencia en probabilidad. Por la continuidad de  $g$ , la convergencia en probabilidad también se cumple para la sucesión  $\{g(\xi_n)\}_n$ .

Como  $\|\xi_n - \xi\|_p \rightarrow 0$ , entonces la sucesión de variables aleatorias  $\{|\xi_n|^p\}_n$  es uniformemente integrable. Por otra parte  $g(x) = O(|x|^p)$  cuando  $|x| \rightarrow +\infty$  y  $g$  es continua en  $\mathbb{R}$ , y por lo tanto acotada en cualquier compacto  $[-M, M]$ , se cumple entonces que hay ciertas constantes  $k, C > 0$  tales que  $|g(\xi_n)| \leq C + k|\xi_n|^p$  y por lo tanto  $\{|g(\xi_n)|\}_n$  es uniformemente integrable. Finalmente, la integrabilidad uniforme y la convergencia en probabilidad implican  $\|g(\xi_n) - g(\xi)\|_1 \rightarrow 0$ .  $\square$

Podemos extender este resultado a intervalos aleatorios.

**Lema 2.4.2.** *Sea  $p \in [1, +\infty)$  y  $g : \mathbb{R} \rightarrow \mathbb{R}$  una función continua tal que  $g(x) = O(|x|^p)$  cuando  $|x| \rightarrow +\infty$ . Si  $\mathcal{D}^p(X_n, X) \rightarrow 0$  donde  $X, X_n \in \mathcal{L}^p$  para todo  $n \in \mathbb{N}$  son intervalos aleatorios, entonces  $\mathcal{D}^1(g(X_n), g(X)) \rightarrow 0$ .*

*Demostración.* Como  $g$  es continua y  $X$  toma como valores intervalos compactos, entonces  $g(X) = [\text{mín } g(X), \text{máx } g(X)]$ . Sabemos que existen dos constantes positivas  $k$  y  $C$  tales que  $|g(x)| \leq C + k|x|^p$  y por lo tanto las normas  $\mathcal{L}^1$  de los valores extremos de  $g(X)$  están acotadas, es decir,  $\text{máx}\{\|\text{mín } g(X)\|_1, \|\text{máx } g(X)\|_1\} \leq C + k\|\text{máx}\{|\text{mín } X|, |\text{máx } X|\}^p\|_1$  que es finito porque  $X \in \mathcal{L}^p$ . Por lo tanto  $g(X) \in \mathcal{L}^1$  y de un modo análogo se puede demostrar que  $g(X_n) \in \mathcal{L}^1$  para todo  $n \in \mathbb{N}$ .

Como  $\|d_H(X_n, X)\|_p \rightarrow 0$ , entonces  $\{d_H(X_n, X)\}_n$  y  $\{d_H(g(X_n), g(X))\}_n$  convergen en probabilidad a 0. Para demostrar que  $\{d_H(g(X_n), g(X))\}_n = \{\text{máx}\{|\text{máx } g(X_n) - \text{máx } g(X)|, |\text{mín } g(X_n) - \text{mín } g(X)|\}\}_n$  es uniformemente integrable, sólo hay que probar la integrabilidad uniforme de  $\{\text{mín } g(X_n) - \text{mín } g(X)\}_n$  y de  $\{\text{máx } g(X_n) - \text{máx } g(X)\}_n$ . La de la segunda sucesión de variables es inmediata a partir de las siguientes desigualdades,

$$\begin{aligned} |\text{máx } g(X_n) - \text{máx } g(X)| &\leq |\text{máx } g(X_n)| + |\text{máx } g(X)| \\ &\leq C + k \text{máx}\{|\text{mín } X_n|, |\text{máx } X_n|\}^p + |\text{máx } g(X)|, \end{aligned}$$

y entonces de  $\|\text{mín } X_n - \text{mín } X\|_p \rightarrow 0$  y  $\|\text{máx } X_n - \text{máx } X\|_p \rightarrow 0$  se sigue la integrabilidad uniforme de  $\{\text{máx}\{|\text{mín } X_n|, |\text{máx } X_n|\}^p\}_n$  y como  $\text{máx } g(X) \in L^1$ , entonces  $\{|\text{máx } g(X_n) - \text{máx } g(X)|\}_n$  es uniformemente integrable. De manera análoga se prueba que  $\{\text{mín } g(X_n) - \text{mín } g(X)\}_n$  es uniformemente integrable y esto conjuntamente con la convergencia en probabilidad de  $\{d_H(g(X_n), g(X))\}_n$  a 0 implican  $\mathcal{D}^1(g(X_n), g(X)) \rightarrow 0$ .  $\square$

En primer lugar probamos la continuidad de los índices  $I_f$  de intervalos aleatorios.

**Teorema 2.4.3.** *Sea  $f : [0, +\infty) \rightarrow \mathbb{R}$  una función convexa tal que  $f(1) = 0$  y  $f(x) = O(|x|^p)$  cuando  $|x| \rightarrow +\infty$  para algún  $p \in [1, +\infty)$ . Entonces,  $I_f : \mathcal{L}_+^p \rightarrow \mathcal{K}_c(\mathbb{R})$  es  $\mathcal{D}^p$ -continua.*

*Demostración.* Sea  $X, X_n \in \mathcal{L}_+^p$  para todo  $n \in \mathbb{N}$  tal que  $\mathcal{D}^p(X_n, X) \rightarrow 0$ . Como las esperanzas de las variables aleatorias  $\min X_n, \min X, \max X_n$  y  $\max X$  son positivas, entonces

$$\left\| \frac{\min X_n}{\mathbb{E} \max X_n} - \frac{\min X}{\mathbb{E} \max X} \right\|_p \rightarrow 0 \quad \text{y} \quad \left\| \frac{\max X_n}{\mathbb{E} \min X_n} - \frac{\max X}{\mathbb{E} \min X} \right\|_p \rightarrow 0$$

y por lo tanto  $\mathcal{D}^p(X_n/\mathbb{E}X_n, X/\mathbb{E}X) \rightarrow 0$ .

Por otra parte, la convexidad de  $f$  garantiza su continuidad y por el Lema 2.4.2 tenemos que  $\mathcal{D}^1(f(X_n/\mathbb{E}X_n), f(X/\mathbb{E}X)) \rightarrow 0$ . Finalmente, como  $d_H(I_f(X_n), I_f(X)) \leq \mathcal{D}^1(f(X_n/\mathbb{E}X_n), f(X/\mathbb{E}X))$  se verifica la continuidad.  $\square$

El resultado anterior se generaliza inmediatamente a los conjuntos aleatorios.

**Corolario 2.4.4.** *Sea  $f : [0, +\infty) \rightarrow \mathbb{R}$  una función convexa tal que  $f(1) = 0$  y  $f(x) = O(|x|^p)$  cuando  $|x| \rightarrow +\infty$  para algún  $p \in [1, +\infty)$  y sea  $l : \mathbb{R}^d \rightarrow \mathbb{R}$  lineal estrictamente creciente. Entonces  $I_{f,l} : \mathcal{L}_+^p \rightarrow \mathcal{K}_c(\mathbb{R})$  es  $\mathcal{D}^p$ -continua.*

## 2.4.2. Aproximación

En la búsqueda de cotas para la distancia entre los índices de desigualdad de dos conjuntos aleatorios, el resultado principal es el Corolario 2.4.8, que se obtiene como consecuencia inmediata del Teorema 2.4.7, en el que se acota la distancia entre los índices de desigualdad de dos intervalos aleatorios.

Antes de pasar a estos resultados obtenemos dos lemas previos que acotan la distancia entre las transformaciones de dos intervalos por ciertas funciones. Denotamos por  $\mathcal{C}^1$  al conjunto de funciones de  $\mathbb{R}$  en  $\mathbb{R}$  con derivada continua.

**Lema 2.4.5.** *Sea  $g : \mathbb{R} \rightarrow \mathbb{R}$  una función de clase  $\mathcal{C}^1$  y sean  $A, B$  dos intervalos compactos, entonces*

$$d_H(g(A), g(B)) \leq d_H(A, B) \max_{z \in \text{co}(A \cup B)} \{|g'(z)|\}.$$

*Demostración.* Sean  $x \in A$ ,  $y \in B$ , es claro que

$$|g(x) - g(y)| \leq |x - y| \max_{z \in \text{co}\{x,y\}} \{|g'(z)|\} \leq |x - y| \max_{z \in \text{co}(A \cup B)} \{|g'(z)|\},$$

ya que la derivada de la función  $g$  es continua, y la envolvente convexa de dos intervalos compactos sigue siendo compacta, con lo que la función alcanzará un máximo en  $\text{co}(A \cup B)$ . Ahora

$$\sup_{x \in A} \inf_{y \in B} |g(x) - g(y)| \leq \sup_{x \in A} \inf_{y \in B} |x - y| \max_{z \in \text{co}(A \cup B)} \{|g'(z)|\},$$

siendo la desigualdad cierta si intercambiamos  $A$  por  $B$ , con lo que obtenemos,

$$d_H(g(A), g(B)) \leq d_H(A, B) \max_{z \in \text{co}(A \cup B)} \{|g'(z)|\}.$$

□

Si la función  $g$  fuese convexa, su derivada tomaría valores extremos en la frontera de un conjunto, por lo que se obtiene el siguiente resultado.

**Corolario 2.4.6.** *Sea  $g : \mathbb{R} \rightarrow \mathbb{R}$  una función convexa de clase  $\mathcal{C}^1$  y sean  $A, B$  dos intervalos compactos, entonces*

$$d_H(g(A), g(B)) \leq d_H(A, B) \max_{z \in A \cup B} \{|g'(z)|\}.$$

Recordamos que nuestro objetivo es encontrar una cota superior para la distancia de Hausdorff entre los índices de desigualdad de dos conjuntos aleatorios y para ello empezamos estudiando el caso de los intervalos aleatorios.

Denotamos por  $q$  al *conjugado* de  $p$ , es decir,  $q = p/(p-1)$  si  $p \in (1, +\infty)$ ,  $q = +\infty$  si  $p = 1$  y  $q = 1$  si  $p = +\infty$ .

**Teorema 2.4.7.** *Sean  $X, Y$  dos intervalos aleatorios en  $\mathcal{L}_+^p$  para algún  $p \in [1, +\infty]$ . Sea  $f : (0, +\infty) \rightarrow \mathbb{R}$  una función convexa en  $\mathcal{C}^1$  tal que  $f(1) = 0$ . Si tanto  $I_f(X)$  como  $I_f(Y)$  existen y  $f'(X/\mathbb{E}X), f'(Y/\mathbb{E}Y) \in \mathcal{L}^q$ , donde  $q$  es el conjugado de  $p$ , entonces*

$$d_H(I_f(X), I_f(Y)) \leq \mathcal{D}^p\left(\frac{X}{\mathbb{E}X}, \frac{Y}{\mathbb{E}Y}\right) \left[ \left\| f'\left(\frac{X}{\mathbb{E}X}\right) \right\|_q + \left\| f'\left(\frac{Y}{\mathbb{E}Y}\right) \right\|_q \right].$$

*Demostración.* El  $f$ -índice de desigualdad de un intervalo aleatorio  $X$  es la esperanza de Aumann de una transformación de  $X$  y así, según el Teorema 0.2.1, se cumple

$$d_H(I_f(X), I_f(Y)) \leq \mathbb{E} \left[ d_H \left( f \left( \frac{X}{\mathbb{E}X} \right), f \left( \frac{Y}{\mathbb{E}Y} \right) \right) \right].$$

Por el Corolario 2.4.6,

$$\begin{aligned} d_H(I_f(X), I_f(Y)) &\leq \mathbb{E} \left[ d_H \left( \frac{X}{\mathbb{E}X}, \frac{Y}{\mathbb{E}Y} \right) \max_{z \in X/\mathbb{E}X \cup Y/\mathbb{E}Y} |f'(z)| \right] \\ &\leq \mathbb{E} \left[ d_H \left( \frac{X}{\mathbb{E}X}, \frac{Y}{\mathbb{E}Y} \right) \max_{x \in X/\mathbb{E}X} |f'(z)| \right] \\ &\quad + \mathbb{E} \left[ d_H \left( \frac{X}{\mathbb{E}X}, \frac{Y}{\mathbb{E}Y} \right) \max_{y \in Y/\mathbb{E}Y} |f'(z)| \right]. \end{aligned}$$

Finalmente aplicando la desigualdad de Hölder obtenemos,

$$d_H(I_f(X), I_f(Y)) \leq \left\| d_H \left( \frac{X}{\mathbb{E}X}, \frac{Y}{\mathbb{E}Y} \right) \right\|_p \left[ \left\| f' \left( \frac{X}{\mathbb{E}X} \right) \right\|_q + \left\| f' \left( \frac{Y}{\mathbb{E}Y} \right) \right\|_q \right]$$

□

**Corolario 2.4.8.** Sean  $X, Y$  dos conjuntos aleatorios en  $\mathcal{L}_+^p$  para algún  $p \in [1, +\infty]$ . Sea  $f : (0, +\infty) \rightarrow \mathbb{R}$  una función convexa en  $\mathcal{C}^1$  tal que  $f(1) = 0$  y  $l : \mathbb{R}^d \rightarrow \mathbb{R}$  una aplicación lineal estrictamente creciente. Si tanto  $I_{f,l}(X)$  como  $I_{f,l}(Y)$  existen y  $f'(l(X)/l(\mathbb{E}X)), f'(l(Y)/l(\mathbb{E}Y)) \in \mathcal{L}^q$ , donde  $q$  es el conjugado de  $p$ , entonces

$$\begin{aligned} d_H(I_{f,l}(X), I_{f,l}(Y)) \\ \leq \mathcal{D}^p \left( \frac{l(X)}{l(\mathbb{E}X)}, \frac{l(Y)}{l(\mathbb{E}Y)} \right) \left[ \left\| f' \left( \frac{l(X)}{l(\mathbb{E}X)} \right) \right\|_q + \left\| f' \left( \frac{l(Y)}{l(\mathbb{E}Y)} \right) \right\|_q \right]. \end{aligned}$$

Este resultado puede especializarse para vectores aleatorios,

**Corolario 2.4.9.** Sean  $\xi, \eta$  dos vectores aleatorios en  $L_+^p$  para algún  $p \in [1, +\infty]$ . Sea  $f : (0, +\infty) \rightarrow \mathbb{R}$  una función convexa en  $\mathcal{C}^1$  tal que  $f(1) = 0$  y  $l : \mathbb{R}^d \rightarrow \mathbb{R}$  una aplicación lineal estrictamente creciente. Si tanto  $i_{f,l}(\xi)$  como  $i_{f,l}(\eta)$  existen y  $f'(l(\xi)/l(\mathbb{E}\xi)), f'(l(\eta)/l(\mathbb{E}\eta)) \in L^q$  donde  $q$  es el conjugado de  $p$ , entonces

$$\begin{aligned} |i_{f,l}(\xi) - i_{f,l}(\eta)| &\leq \left\| \frac{l(\xi)}{l(\mathbb{E}\xi)} - \frac{l(\eta)}{l(\mathbb{E}\eta)} \right\|_p \left[ \left\| f' \left( \frac{l(\xi)}{l(\mathbb{E}\xi)} \right) \right\|_q + \left\| f' \left( \frac{l(\eta)}{l(\mathbb{E}\eta)} \right) \right\|_q \right] \\ &\leq 2 \left[ \frac{\|f'(\xi/\mathbb{E}\xi)\|_q + \|f'(\eta/\mathbb{E}\eta)\|_q}{l(\mathbb{E}\xi)l(\mathbb{E}\eta)} \right] \|l\|^2 \|\xi - \eta\|_p \min\{\|\xi\|_p, \|\eta\|_p\} \end{aligned}$$

## 2.5. Ejemplo

Como ejemplo de aplicación de los resultados obtenidos analizaremos el comportamiento de los índices de desigualdad de ciertos intervalos aleatorios.

Nuestro objetivo es estudiar la variabilidad relativa del ‘rango del nivel de glucosa en sangre en los días de la semana para un paciente diabético’. A lo largo de una semana se producen alteraciones en el ritmo de vida de una persona y en consecuencia los registros vitales varían. Hemos tomado el período de una semana como el más corto que puede mostrar una desigualdad significativa.

Modelaremos el rango del nivel de glucosa en sangre como un intervalo aleatorio discreto,  $X = [\text{mín } X, \text{máx } X]$  sobre el espacio de probabilidad  $(E, \mathcal{P}(E), P)$  de los siete días de la semana,  $E = \{e_1, \dots, e_7\}$  todos con la misma probabilidad,  $P(\{e\}) = 1/7$ . De este modo las variables aleatorias  $\text{mín } X(e)$  y  $\text{máx } X(e)$  representarán los niveles mínimo y máximo de glucosa en el día  $e$ .

Los datos disponibles se pueden representar como aparece en la Figura 2.1.

Podemos aproximar este intervalo aleatorio mediante una sucesión de intervalos aleatorios  $\{X_n\}_{n \in \mathbb{N}}$  sobre  $(E, \mathcal{P}(E), P)$  estando cada una de sus realizaciones definida como la envolvente convexa de  $n$  mediciones del nivel de glucosa en diferentes momentos del día. Más específicamente, para

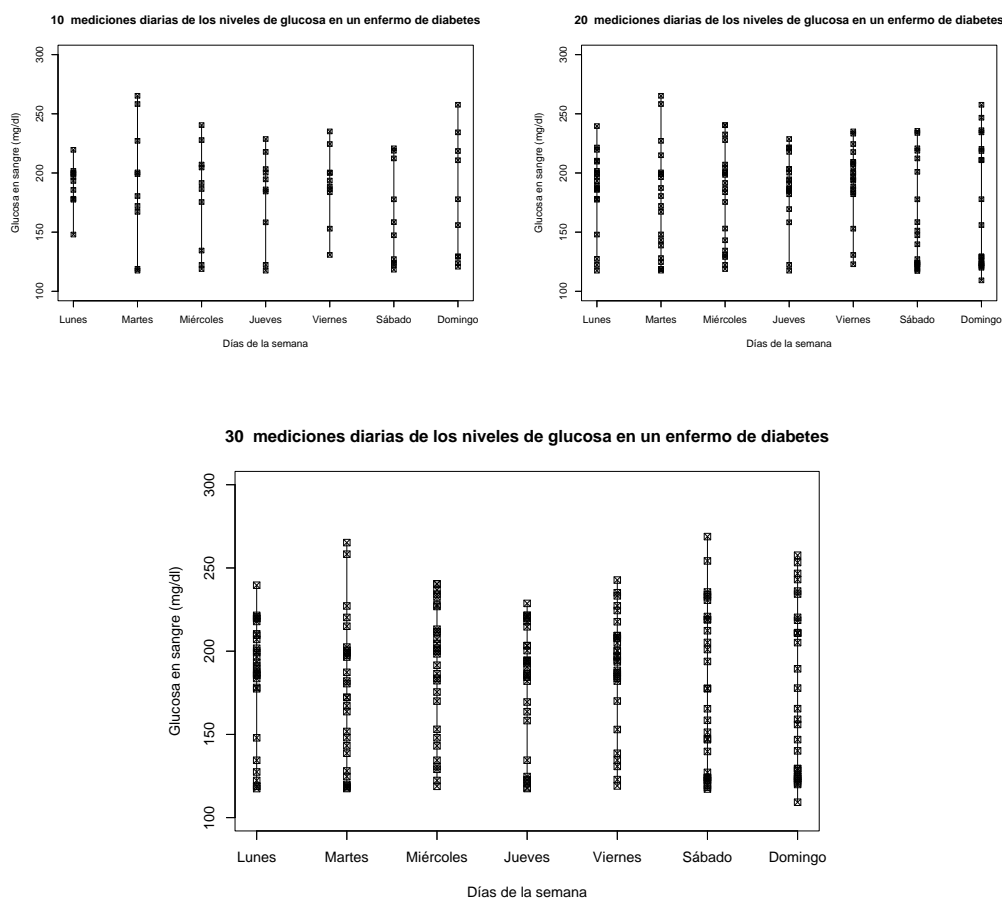


Figura 2.1: Evolución de la glucosa en sangre a lo largo de una semana. Cada símbolo  $\boxtimes$  indica el resultado de una medición del nivel de glucosa en sangre en el día correspondiente.

cada día  $e$  del estudio tomamos diferentes mediciones del nivel de glucosa en sangre,  $\xi_1^e, \xi_2^e, \dots, \xi_n^e$  y aproximamos  $X(e)$  por  $X_n(e)$  donde  $\min X_n(e) = \min\{\xi_1^e, \dots, \xi_n^e\}$  y  $\max X_n(e) = \max\{\xi_1^e, \dots, \xi_n^e\}$ .

En este punto hacemos la suposición de que fijado  $e \in E$ , las variables  $\{\xi_i^e\}_i$  son independientes y están idénticamente distribuidas en el intervalo



$X(e)$ . Además, suponemos que existen dos constantes  $\gamma, k > 0$  tales que para todo  $x \in [0, \gamma]$ , la probabilidad de que  $\xi_1^e$  esté a una distancia de  $\max X(e)$  menor o igual que  $x$  es al menos  $kx$  y lo mismo para  $\min X(e)$ . Es decir, supuestas  $\xi_1^e, \xi_2^e, \dots$  definidas en un espacio  $(\Omega, \mathcal{A}, \Pr)$ , se cumple  $\Pr(|\xi_1^e - \max X(e)| \leq x) \geq kx$  y  $\Pr(|\xi_1^e - \min X(e)| \leq x) \geq kx$ . Obsérvese que la familia de distribuciones que verifica esta condición es suficientemente amplia. Para estas variables aleatorias el orden de convergencia de la distancia máxima de los valores mínimo y máximo de  $n$  copias independientes de  $\xi_1^e$  a los extremos del intervalo  $X(e)$  es casi seguro  $O(n^{-1} \log n)$ , como obtienen Dümbgen y Walther (1996). Esto nos ayudará a mantener cierta velocidad de convergencia en nuestras aproximaciones del índice de desigualdad al valor objetivo.

La distancia de Hausdorff entre los intervalos  $X_n(e)$  y  $X(e)$  es igual a  $\max\{\min X_n(e) - \min X(e), \max X(e) - \max X_n(e)\}$  y por lo tanto se tiene que  $d_H(X_n(e), X(e)) = O(n^{-1} \log n)$  *c.s.*

Para cualquier función  $f$  que satisfaga los requerimientos para definir un índice de desigualdad, según el Teorema 2.4.7 y por el hecho de que  $X$  está acotado ya que  $X(e_j) \in \mathcal{K}_c(\mathbb{R})$  para  $j \in \{1, \dots, 7\}$ , el orden de convergencia de los índices de desigualdad  $\{I_f(X_n)\}_n$  hacia  $I_f(X)$  coincidirá con el de  $\{X_n/\mathbb{E}X_n\}_n$  hacia  $X/\mathbb{E}X$  en el espacio  $\mathcal{L}^1$  y tenemos que

$$d_H(X_n(e)/\mathbb{E}X_n, X(e)/\mathbb{E}X) \leq (2/\mathbb{E} \min X) d_H(X_n(e), X(e)).$$

Ahora obtenemos fácilmente el orden de convergencia de  $\{X_n/\mathbb{E}X_n\}_n$  hacia  $X/\mathbb{E}X$ ,

$$\mathcal{D}^1(X_n/\mathbb{E}X_n, X/\mathbb{E}X) = O(n^{-1} \log n) \quad \textit{c.s.} \quad (2.1)$$

y del mismo modo  $d_H(I_f(X_n), I_f(X)) = O(n^{-1} \log n)$  *c.s.* Este hecho nos ayudará a controlar el tamaño muestral a considerar para lograr una reducción de algún orden fijo en la cota de error.

Para ilustrar este ejemplo con datos numéricos, hemos utilizado el programa AIDA-online2 [<http://www.2aida.net>]. Se trata de un simulador de

los efectos de la insulina en el nivel de glucosa en sangre con el que se pueden analizar los efectos que producen las alteraciones en la dieta y en el suministro de insulina sobre el nivel de glucosa en sangre en pacientes diabéticos virtuales. Se ha escogido un paciente diabético virtual específico y sobre él se han observado los 96 valores disponibles del nivel de glucosa en sangre al día durante los 7 días de la semana.

Sobre la base de la envolvente convexa de los 96 valores disponibles para un día  $e$ , definimos el índice de desigualdad poblacional. Tomamos una función de la familia que generan los índices aditivamente descomponibles (Bourguignon, 1979), en concreto la que genera el llamado índice hiperbólico,  $f_{-1}(x) = (1/x - 1)/2$ . El índice  $I_f(X)$  vendrá dado por,

$$I_f(X) = \left[ \sum_{i=1}^7 \frac{\min X(e_i)}{14\mathbb{E} \max X} - \frac{1}{2}, \sum_{i=1}^7 \frac{\max X(e_i)}{14\mathbb{E} \min X} - \frac{1}{2} \right] = [-0,2719, 0,6017]$$

A partir de la población anterior y para cada día de la semana se ha simulado una muestra de tamaño  $n$ , variando de 1 a 10,  $n = 20$  y  $n = 30$ , donde para cada día los tiempos de la observación han sido elegidos aleatoriamente de entre los 96 disponibles. En la Figura 2.1 se puede ver el resultado obtenido para tamaños muestrales  $n = 10, 20$  y  $30$ .

Dado un  $n$  cualquiera y las 7 muestras correspondientes, podemos obtener el valor mínimo y máximo para cada día de la semana y posteriormente calcular el índice de desigualdad hiperbólico. La Figura 2.2 contiene los valores del índice para las muestras simuladas. Estos valores reflejan la variabilidad del nivel de glucosa a lo largo de la semana. El extremo superior viene determinado por los niveles máximos del nivel de glucosa en cada día divididos entre el extremo inferior del nivel medio para normalizarlos, el extremo inferior por los niveles mínimos normalizados al dividirlos entre el extremo superior del nivel medio. Podrían utilizarse para comparar la variabilidad en dos semanas diferentes. La velocidad de convergencia hacia el índice poblacional y, por tanto la mejora en la aproximación que se logra al aumentar el tamaño de la muestra, ya se razonó en (2.1).

$n$	mín $I_{f_{-1}}(X_n)$	máx $I_{f_{-1}}(X_n)$
1	0'0221	0'0221
2	-0'0861	0'1416
3	-0'1391	0'2306
4	-0'1866	0'3270
5	-0'2064	0'3736
6	-0'2296	0'4363
7	-0'2353	0'4554
8	-0'2367	0'4611
9	-0'2399	0'4713
10	-0'2468	0'4978
20	-0'2627	0'5585
30	-0'2699	0'5917

Figura 2.2: Índices de desigualdad para  $f_{-1}(x) = (1/x - 1)/2$ .

## 2.6. Comentarios y problemas abiertos

Hemos propuesto índices de desigualdad para conjuntos aleatorios, derivado sus propiedades esenciales y los hemos relacionado con las ordenaciones estocásticas según la variabilidad del Capítulo 1

Los valores que toman los índices de desigualdad para conjuntos aleatorios están completamente condicionados por la función  $f$  que los induce y por la dirección  $u$  en que estudiamos la variabilidad. Así, estos parámetros deberían ser escogidos dependiendo del problema que queremos estudiar. Sería interesante hacer estudios de la robustez de los índices según la función que los induce.

Los índices de desigualdad que se han propuesto para conjuntos aleatorios son una generalización de índices de desigualdad para variables aleatorias. Las variables aleatorias difusas constituyen una generalización de los conjuntos aleatorios y se podrían adaptar nuestros índices de desigualdad

para conjuntos aleatorios para que reflejen la variabilidad de una variable aleatoria difusa.

Se pueden estudiar índices de desigualdad definidos de manera análoga a los que estudiamos nosotros, pero utilizando otras definiciones de esperanza para conjuntos aleatorios distintas de la esperanza de Aumann. También se podría hacer un estudio semejante al de la Sección 2.4 para métricas distintas de la de Hausdorff.

## Capítulo 3

# Regiones centrales recortadas

La recta real está dotada de un orden total por el que podemos comparar dos números reales cualesquiera respecto de la relación menor o igual. Dada una distribución de probabilidad inducida por una variable aleatoria, el orden menor o igual también nos permite hablar del punto más central, o conjunto de puntos más centrales, con respecto a dicha distribución de probabilidad. Tales puntos son las *medianas* y satisfacen que la probabilidad de que la variable aleatoria tome un valor menor o igual que ellas es, al menos, un medio y la probabilidad de que tome un valor mayor o igual es también, al menos un medio. Las medianas, además de esta propiedad satisfacen otras muy interesantes, como por ejemplo, ser el punto para el que la esperanza de la distancia euclídea a la variable aleatoria es mínima. Del mismo modo que consideramos el conjunto de medianas o puntos más centrales, podemos considerar regiones de puntos con cierto “grado de centralidad” o *profundidad*, así el intervalo comprendido entre el cuantil  $\alpha$  y el cuantil  $(1 - \alpha)$  con  $\alpha \in (0, 1/2]$  será más pequeño en torno a la mediana cuanto mayor sea  $\alpha$  y define regiones de profundidad  $\alpha$ .

Si trabajamos en  $\mathbb{R}^d$ , la relación menor o igual entendida ahora componente a componente deja de ser un orden total, por lo que el concepto de punto más central con respecto a una distribución de probabilidad inducida por un vector aleatorio en  $\mathbb{R}^d$ , empieza a ser ambiguo. Surgen entonces nume-

rosas definiciones de medianas multidimensionales, muchas de ellas consideradas por Small (1990) y cada una satisfaciendo algunas de las propiedades de la mediana unidimensional. Hay también numerosas regiones centrales para distribuciones multivariantes, algunas generalizan el concepto de intervalo intercuantílico y otras están inspiradas en estimadores de localización distintos de los cuantiles.

Varios autores como Barnett (1976), Eddy (1983 y 1985) y Liu y Singh (1992) han considerado el problema de la ordenación de datos multidimensionales como paso previo a su estudio estadístico. Lo habitual es realizar tal ordenación conforme a la centralidad con respecto a una distribución de probabilidad y existe un gran interés en la comunidad estadística por el estudio de regiones centrales y funciones de profundidad. Así, entre otros, Massé y Theodorescu (1994), Koshevoy y Mosler (1997b) y Zuo y Serfling (2000c) estudian diferentes regiones centrales multivariantes, mientras que Liu (1990), Liu *et al.* (1999) y Zuo y Serfling (2000a) se centran en el estudio de las profundidades de los puntos. Del mismo modo que se puede estudiar la profundidad de un punto también se pueden estudiar cuantiles multidimensionales, véase por ejemplo, Kolchinskii (1997). La peculiaridad de los cuantiles multidimensionales es que tienen en cuenta no sólo el grado de centralidad, sino también la localización del punto respecto de la distribución y además suelen conservar propiedades de los cuantiles unidimensionales.

En este capítulo estudiamos *regiones centrales recortadas* inducidas por un estimador de localización. En primer lugar definimos el *recorte* de una distribución de probabilidad como el conjunto de probabilidades que tienen cierto grado de relación con dicha distribución de probabilidad. Utilizando estos recortes de probabilidad, la región central recortada se define como el conjunto de todos los valores de ese estimador de localización sobre el conjunto de probabilidades del recorte. Derivamos las propiedades de estas familias de regiones centrales de las de los estimadores de localización que las inducen y probamos que generalizan otras regiones clásicas (Casco y López-Díaz, 2004a).

### 3.1. Regiones centrales y profundidad multivariante

Tal como se vio en la Sección 0.4, las regiones centrales son conjuntos anidados formados por puntos de cierta profundidad con respecto a una determinada distribución de probabilidad. Los conceptos de *función de profundidad* y *región central* son duales y así a partir de una función de profundidad podemos definir regiones centrales y las familias de regiones centrales inducen funciones de profundidad. Nosotros nos ocuparemos del estudio de las regiones centrales.

En esta memoria se han definido las regiones centrales del zonoide, las de la envolvente de zonoides y las de la profundidad semiespacial. Las primeras tienen como región más central el conjunto unipuntual formado por la esperanza de una variable aleatoria y van creciendo a medida que concentramos la probabilidad en ciertas regiones de la distribución original y hallamos todas las esperanzas de las nuevas distribuciones. Las regiones centrales de la envolvente de zonoides están íntimamente relacionadas con las del zonoide. Finalmente, las regiones centrales semiespaciales tienen como conjunto más central el formado por los puntos para los que el ínfimo del conjunto de probabilidades de cualquier semiespacio cerrado que los contenga es máximo. Ésta es una de las propiedades de la mediana de las variables aleatorias y, en su caso, el máximo de estos ínfimos es siempre al menos  $1/2$ . La profundidad de los restantes puntos viene dada también por el mismo ínfimo, siendo en consecuencia estas regiones generalizaciones de los intervalos comprendidos entre los cuantiles unidimensionales  $\alpha$  y  $(1 - \alpha)$  cuando  $\alpha \in (0, 1/2]$ .

Para  $d = 1$  los *intervalos intercuantílicos* constituyen la familia más conocida de regiones centrales. Si  $\alpha \in (0, 1/2]$  y  $P$  es una probabilidad en  $\mathbb{R}$ , los intervalos intercuantílicos son de la forma

$$[\inf\{x : P((-\infty, x]) \geq \alpha\}, \sup\{x : P([x, +\infty)) \geq \alpha\}]. \quad (3.1)$$

## 3.2. El recorte de una probabilidad

Recordamos que  $\mathbb{P}$  denota al conjunto de probabilidades sobre  $(\mathbb{R}^d, \mathcal{B}_d)$  donde  $\mathcal{B}_d$  es la  $\sigma$ -álgebra de Borel en  $\mathbb{R}^d$  y el conjunto  $\mathbb{P}$  está dotado de la topología de la convergencia débil, a la que denotamos por ' $\Rightarrow$ '.

Con el objetivo de definir regiones centrales de cierta profundidad, comenzamos introduciendo una familia de probabilidades construida a partir de una distribución de probabilidad original. Los elementos de la familia tienen una relación de cierto grado con respecto a la probabilidad original y es este grado el que más adelante determinará la profundidad de un punto.

*Definición 3.2.1.* Dado  $\alpha \in (0, 1]$ , definimos el  $\alpha$ -recorte de la probabilidad  $P$  como el conjunto de probabilidades dominadas por  $\alpha^{-1}P$  y lo denotamos por  $P^\alpha$ , esto es

$$P^\alpha := \{Q \in \mathbb{P} : Q \leq \alpha^{-1}P\}.$$

A continuación estudiamos las propiedades de esta nueva familia de probabilidades. En primer lugar analizamos sus propiedades más generales, principalmente las topológicas y luego pasamos a estudiar el comportamiento del  $\alpha$ -recorte de una distribución de probabilidad empírica.

### 3.2.1. Propiedades generales de la familia $P^\alpha$

Los conjuntos  $P^\alpha$  son convexos y crecen desde  $P^1 = \{P\}$  a medida que disminuye  $\alpha$ .

**Lema 3.2.1.** *Dado  $\alpha \in (0, 1]$  y una probabilidad  $P \in \mathbb{P}$ , el conjunto  $P^\alpha$  satisface las siguientes propiedades:*

- i.  $P^\alpha$  es convexo;*
- ii. si  $\alpha \leq \beta$ , entonces  $P^\alpha \supset P^\beta$ ;*
- iii.  $P^1 = \{P\}$ .*



**Corolario 3.2.2.** *El conjunto  $P^\alpha$  es conexo por caminos.*

*Demostración.* Dados cualesquiera  $Q_1, Q_2 \in P^\alpha$ , basta considerar la aplicación  $f : [0, 1] \rightarrow P^\alpha$  definida como  $f(\lambda) = (1 - \lambda)Q_1 + \lambda Q_2$ . Es inmediato demostrar que esta aplicación es continua para la topología de la convergencia débil.  $\square$

Los conjuntos  $P^\alpha$  son tight y compactos.

**Lema 3.2.3.** *Para todo  $\alpha \in (0, 1]$  y  $P \in \mathbb{P}$ , la clase  $P^\alpha$  es tight.*

*Demostración.* Como ya se razonó en la Sección 0.1, toda probabilidad  $P \in \mathbb{P}$  es tight y por la Definición 0.1.2 para  $\varepsilon > 0$  existe un compacto  $K \subset \mathbb{R}^d$  tal que  $P(K) > 1 - \varepsilon\alpha$  o equivalentemente  $P(\mathbb{R}^d \setminus K) < \varepsilon\alpha$ .

Así, dados  $\varepsilon > 0$  y  $Q \in P^\alpha$  se cumple  $Q(\mathbb{R}^d \setminus K) \leq \alpha^{-1}P(\mathbb{R}^d \setminus K) \leq \varepsilon$  y por lo tanto  $Q(K) > 1 - \varepsilon$ , con lo que la familia  $P^\alpha$  es tight.  $\square$

**Lema 3.2.4.** *Para todo  $\alpha \in (0, 1]$  y  $P \in \mathbb{P}$ ,  $P^\alpha$  es compacto.*

*Demostración.* Como  $P^\alpha$  es tight, según el Teorema 0.1.6, la clausura de  $P^\alpha$ ,  $\overline{P^\alpha}$  es compacta. Así que es suficiente con demostrar que  $\overline{P^\alpha} = P^\alpha$ , es decir, que  $P^\alpha$  es cerrado. Obsérvese que la clausura se toma para la topología de la convergencia débil.

Sea  $Q \in \overline{P^\alpha}$ , entonces existe una sucesión  $\{Q_n\}_n \subset P^\alpha$  tal que  $Q_n \Rightarrow Q$ .

Sea  $G \subset \mathbb{R}^d$  un conjunto abierto, por lo tanto  $Q(G) \leq \liminf_n Q_n(G) \leq \alpha^{-1}P(G)$ . Si  $A \in \mathcal{B}_d$ , como según el Teorema 0.1.2,  $P$  es regular, podemos tomar una sucesión decreciente de conjuntos abiertos en  $\mathbb{R}^d$ ,  $\{G_n\}_n$  de tal modo que para todo  $n \in \mathbb{N}$  se cumple  $A \subset G_n$  y  $P(G_n \setminus A) < 1/n$  y así  $Q(A) \leq Q(G_n) \leq \alpha^{-1}P(G_n)$ . Por la continuidad de la probabilidad  $Q(A) \leq \alpha^{-1}P(A)$  de donde  $Q \in P^\alpha$ , y por lo tanto,  $P^\alpha$  es compacto.  $\square$

Continuamos el estudio de las propiedades preliminares de  $P^\alpha$  con dos lemas que nos resultarán muy útiles más adelante y sirven de base para el Teorema 3.2.7. Probaremos por un lado que si tenemos una sucesión de probabilidades que converge débilmente y construimos otra sucesión de probabilidades que también converja débilmente y de tal modo que cada elemento de

la nueva sucesión pertenece al  $\alpha$ -recorte de la correspondiente probabilidad de la sucesión original, entonces el límite de la nueva sucesión pertenecerá al  $\alpha$ -recorte del límite de la sucesión original. En el otro resultado auxiliar se prueba que fijado  $\alpha \in (0, 1]$ , el conjunto de  $\alpha$ -recortes de una sucesión de probabilidades débilmente convergente es tight. El Teorema 3.2.7 demuestra que toda sucesión de probabilidades pertenecientes a los  $\alpha$ -recortes de una sucesión de probabilidades convergente, tiene una subsucesión que converge a alguna probabilidad del  $\alpha$ -recorte de la probabilidad límite.

**Lema 3.2.5.** Sean  $P, P_n \in \mathbb{P}$  para todo  $n \in \mathbb{N}$  tales que  $P_n \Rightarrow P$  y sea  $\alpha \in (0, 1]$ . Si  $\{Q_n\}_n \subset \mathbb{P}$  es una sucesión convergente,  $Q_n \Rightarrow Q$ , con  $Q_n \in P_n^\alpha$  para todo  $n \in \mathbb{N}$ , entonces  $Q \in P^\alpha$ .

*Demostración.* Sea  $G \subset \mathbb{R}^d$  un conjunto abierto, entonces

$$\begin{aligned} Q(G) &\leq \liminf_n Q_n(G) \leq \liminf_n \alpha^{-1} P_n(G) \\ &\leq \liminf_n \alpha^{-1} P_n(\overline{G}) \leq \limsup_n \alpha^{-1} P_n(\overline{G}) \leq \alpha^{-1} P(\overline{G}). \end{aligned}$$

Tomamos ahora un conjunto  $F \subset \mathbb{R}^d$  cerrado y consideramos la sucesión de abiertos de  $\mathbb{R}^d$ ,  $\{G_n\}_n$  en la que cada elemento está definido como  $G_n = \{x : d(x, F) < 1/n\}$ , donde  $d(x, F) = \inf\{\|x - y\| : y \in F\}$ , entonces  $Q(F) \leq Q(G_n) \leq \alpha^{-1} P(\overline{G_n})$  para todo  $n \in \mathbb{N}$ . Tomando límites obtenemos que  $Q(F) \leq \alpha^{-1} P(F)$ .

Finalmente, para un conjunto  $A \in \mathcal{B}_d$  cualquiera, como tanto  $P$  como  $Q$  son regulares, basta tomar una sucesión creciente de conjuntos cerrados  $\{F_n\}_n \subset \mathbb{R}^d$  tales que  $F_n \subset A$  y  $Q(A \setminus F_n) + P(A \setminus F_n) < 1/n$  y por lo tanto,  $Q(A) = \lim_n Q(F_n) \leq \lim_n \alpha^{-1} P(F_n) = \alpha^{-1} P(A)$  lo que finaliza la demostración.  $\square$

**Lema 3.2.6.** Sean  $P, P_n \in \mathbb{P}$  para todo  $n \in \mathbb{N}$  tales que  $P_n \Rightarrow P$  y sea  $\alpha \in (0, 1]$ , entonces  $\cup_{n=1}^\infty P_n^\alpha$  es tight.

*Demostración.* Como  $P_n \Rightarrow P$ , según el Teorema 0.1.4, el conjunto  $\{P, P_1, P_2, \dots\}$  es tight, condición que mantendrá si quitamos un elemento, con lo que

$\cup_{n=1}^{\infty} \{P_n\}$  es tight. Así, por la Definición 0.1.2, para  $\varepsilon > 0$  existe un compacto  $K \subset \mathbb{R}^d$  tal que  $P_n(K) > 1 - \varepsilon\alpha$  para todo  $n \in \mathbb{N}$  o equivalentemente  $P(\mathbb{R}^d \setminus K) < \varepsilon\alpha$  para todo  $n \in \mathbb{N}$ .

Dados  $\varepsilon > 0$  y  $Q \in \cup_{n=1}^{\infty} P_n^\alpha$  se cumple que  $Q \in P_i^\alpha$  para algún  $i \in \mathbb{N}$  y por lo tanto  $Q(\mathbb{R}^d \setminus K) \leq \alpha^{-1} P_i(\mathbb{R}^d \setminus K) \leq \varepsilon$  con lo que finalmente obtenemos que  $Q(K) > 1 - \varepsilon$  obteniéndose el resultado deseado.  $\square$

**Teorema 3.2.7.** Sean  $P, P_n \in \mathbb{P}$  para todo  $n \in \mathbb{N}$  tales que  $P_n \Rightarrow P$  y sea  $\alpha \in (0, 1]$ , entonces para cualquier sucesión  $\{Q_n\}_n$  con  $Q_n \in P_n^\alpha$  para cada  $n \in \mathbb{N}$ , existe una subsucesión  $\{Q_{n_k}\}_k \subset \{Q_n\}_n$  y  $Q \in P^\alpha$  tales que la subsucesión converge hacia  $Q$ , esto es  $Q_{n_k} \Rightarrow Q$ .

*Demostración.* Como  $P_n \Rightarrow P$  y  $\alpha \in (0, 1]$ , según el Lema 3.2.6, el conjunto  $\cup_{n=1}^{\infty} P_n^\alpha$  es también tight. Dada una sucesión  $\{Q_n\}_n$  con  $Q_n \in P_n^\alpha$  para cada  $n \in \mathbb{N}$ , como  $\cup_{n=1}^{\infty} P_n^\alpha$  es tight, existe una subsucesión convergente. Así, tenemos  $\{Q_{n_k}\}_k$  tal que  $Q_{n_k} \Rightarrow Q$  para algún  $Q \in \mathbb{P}$  y por el Lema 3.2.5, se cumple  $Q \in P^\alpha$ .  $\square$

Dada una sucesión de probabilidades que converge débilmente hacia otra probabilidad y fijado un  $\alpha \in (0, 1]$ , el límite superior de los  $\alpha$ -recortes de las probabilidades de la sucesión está contenido en el  $\alpha$ -recorte de la probabilidad límite.

Obsérvese que ésta es la primera vez en que aparece el *límite superior de una sucesión de conjuntos* a lo largo de la presente memoria. Tal límite superior es simplemente el conjunto de todos los límites de las subsucesiones convergentes originadas a partir de sucesiones con un elemento en cada conjunto de la sucesión.

**Corolario 3.2.8.** Sean  $P, P_n \in \mathbb{P}$  para todo  $n \in \mathbb{N}$  tales que  $P_n \Rightarrow P$  y sea  $\alpha \in (0, 1]$ , entonces

$$\limsup_n P_n^\alpha \subset P^\alpha.$$

### 3.2.2. Comportamiento empírico de $P^\alpha$

El resultado más importante que obtenemos sobre el comportamiento del recorte de una probabilidad empírica es el Teorema 3.2.15. Según él, cualquier probabilidad de un  $\alpha$ -recorte de la probabilidad poblacional pertenece, casi seguro, al límite superior de la sucesión de los  $\alpha$ -recortes de las probabilidades empíricas y además, el límite superior de la sucesión de los  $\alpha$ -recortes de las probabilidades empíricas está contenido en el  $\alpha$ -recorte de la probabilidad poblacional casi seguro. Cabe también destacar la Proposición 3.2.14 en la que dada una probabilidad cualquiera perteneciente al  $\alpha$ -recorte de la probabilidad poblacional, se prueba que existe casi seguro, y se construye explícitamente, una sucesión de probabilidades cuyos elementos pertenecen a los  $\alpha$ -recortes de la sucesión de las probabilidades empíricas y de tal modo que una subsucesión suya converge hacia la probabilidad del  $\alpha$ -recorte de la probabilidad poblacional que habíamos considerado.

Dada una distribución de probabilidad  $P \in \mathbb{P}$ , obtenemos una muestra aleatoria simple  $\xi_1(\cdot), \xi_2(\cdot), \dots$  de vectores aleatorios en un espacio de probabilidad  $(\Omega, \mathcal{A}, \Pr)$  de tal modo que todos ellos siguen la distribución  $P$ . Con esta muestra construimos  $P_{n,\cdot}$ , la probabilidad empírica asociada a  $P$ , es decir, para cualquier  $A \in \mathcal{B}_d$ , se define

$$P_{n,\cdot}(A) := \frac{1}{n} \sum_{i=1}^n I_A(\xi_i(\cdot)),$$

donde  $I_A$  es la función indicador del conjunto  $A$ .

Obsérvese que si  $Q \in P^\alpha$ , entonces  $Q$  es absolutamente continua con respecto a  $P$ , y por lo tanto existe una *derivada de Radon-Nikodym de  $Q$  con respecto a  $P$* . Esta derivada es una aplicación  $g : \mathbb{R}^d \rightarrow [0, +\infty)$  de clase  $L^1(P)$  tal que para todo  $A \in \mathcal{B}_d$ , se cumple  $Q(A) = \int_A g dP$ . De ahora en adelante como  $Q$  es una probabilidad y  $Q \leq \alpha^{-1}P$ , no supone ninguna restricción considerar una derivada de Radon-Nikodym  $g$  satisfaciendo  $0 \leq g \leq \alpha^{-1}$ .

Por otra parte, como  $Q \in \mathbb{P}$ , tenemos que

$$\|g\|_{L^1(P)} = \int g dP = 1 = \|g(\xi_1(\cdot))\|_{L^1(\text{Pr})}$$

y es claro que  $\|g\|_{L^1(P_{n,\cdot})} \longrightarrow \|g\|_{L^1(P)}$  c.s. [Pr]. Obsérvese que en la notación del Capítulo 2,  $\|g(\xi_1(\cdot))\|_{L^1(\text{Pr})} = \|g(\xi_1)\|_1$ .

Dada una variable aleatoria  $\eta$  definida en un espacio de probabilidad  $(\Omega, \mathcal{A}, \text{Pr})$  y  $r \in \mathbb{R} \setminus \{0\}$ , decimos que  $\eta$  está distribuida en el retículo  $R_r = \{nr : n \in \mathbb{Z}\}$  si  $R_r$  es el menor retículo de este tipo que satisface  $\text{Pr}(\eta \in R_r) = 1$ . Si el soporte de  $\eta$  no está contenido en ningún retículo de este tipo, decimos que está distribuida en  $R_0 = \mathbb{R}$ . La Definición 3.37 en Breiman (1968) establece que si  $\eta_1, \eta_2, \eta_3 \dots$  es una sucesión de variables aleatorias independientes e idénticamente distribuidas, todas ellas definidas en un espacio de probabilidad  $(\Omega, \mathcal{A}, \text{Pr})$ , su sucesión de sumas parciales  $\eta_1, \eta_1 + \eta_2, \eta_1 + \eta_2 + \eta_3, \dots$  es *recurrente* si se verifica  $\text{Pr}(\eta_1 + \dots + \eta_n \in I \text{ i.o.}) = 1$  para todo intervalo acotado  $I$  tal que  $R_r \cap I \neq \emptyset$ , donde  $R_r$  es el retículo en el que está distribuida  $\eta_1$  e *i.o.* indica que ocurre para un número infinito de  $n$ 's.

A continuación enunciaremos un resultado sobre recurrencia que se utilizará para estudiar el comportamiento de la sucesión de  $\alpha$ -recortes empíricos.

**Teorema 3.2.9 (Teorema 3.38 en Breiman, 1968).** *En las condiciones anteriores, si  $\mathbb{E}\eta_1 = 0$ , entonces la sucesión de sumas parciales  $\eta_1, \eta_1 + \eta_2, \eta_1 + \eta_2 + \eta_3 \dots$  es recurrente.*

La recurrencia nos sirve para obtener el siguiente resultado ya en el marco de nuestro trabajo.

**Lema 3.2.10.** *Sea  $S_n(\cdot) = \sum_{i=1}^n (g(\xi_i(\cdot)) - 1)$ , entonces  $\text{Pr}(S_n(\cdot) \geq 0 \text{ i.o.}) = 1$ , donde  $S_n(\cdot) \geq 0 \text{ i.o.}$  es el conjunto*

$$\{\omega \in \Omega : S_n(\omega) \geq 0 \text{ ocurre para un número infinito de } n \text{'s}\}.$$

*Demostración.* Como  $\mathbb{E}g(\xi_1(\cdot)) - 1 = 0$ , según el Teorema 3.2.9 la sucesión  $S_1(\cdot), S_2(\cdot), \dots$  es recurrente y por lo tanto  $\text{Pr}(S_n(\cdot) \geq 0 \text{ i.o.}) = 1$ . Obsérvese

que el intervalo  $[0, +\infty)$  debe cortar forzosamente al retículo en que esté distribuido  $g(\xi_1(\cdot)) - 1$ .  $\square$

Como consecuencia inmediata obtenemos.

**Corolario 3.2.11.**  $\Pr(\|g\|_{L^1(P_{n,\cdot})} \geq \|g\|_{L^1(P)} \text{ i.o.}) = 1$ .

*Demostración.*

$$\begin{aligned} \Pr(S_n(\cdot) \geq 0 \text{ i.o.}) &= \Pr\left(\sum_{i=1}^n g(\xi_i(\cdot)) \geq n \text{ i.o.}\right) \\ &= \Pr\left(\frac{1}{n} \sum_{i=1}^n g(\xi_i(\cdot)) \geq 1 \text{ i.o.}\right) = \Pr(\|g\|_{L^1(P_{n,\cdot})} \geq \|g\|_{L^1(P)} \text{ i.o.}). \end{aligned}$$

$\square$

Definimos un tipo de funciones que nos resultarán útiles en la obtención de resultados previos a la Proposición 3.2.14 y al Teorema 3.2.15 donde se describe el comportamiento del recorte de una probabilidad empírica.

Una función  $f$  se dice *semicontinua superiormente* si sus conjuntos nivel del tipo  $\{x : f(x) < \gamma\}$  son abiertos para todo  $\gamma \in \mathbb{R}$ . En la Definición 4.1.4 se define el concepto de función semicontinua inferiormente que está muy relacionado con éste.

En Billingsley (1999), pág. 17 se establece que si una función  $f$  es acotada y semicontinua superiormente y tenemos  $P, P_n \in \mathbb{P}$  para todo  $n \in \mathbb{N}$  tal que  $P_n \Rightarrow P$ , entonces

$$\limsup_n \int f dP_n \leq \int f dP.$$

El siguiente resultado es una especialización del *Teorema de Caratheodory-Vitali* que puede encontrarse, por ejemplo, en el Teorema 2.6.2 en Rudin (1987).

**Teorema 3.2.12 (Teorema de Caratheodory-Vitali).** *Dada  $f : \mathbb{R}^d \rightarrow [0, +\infty)$  con  $f \in L^1(P)$ , existe una sucesión de funciones acotadas y semicontinuas superiormente  $\{u_m\}_m$ ,  $u_m : \mathbb{R}^d \rightarrow [0, +\infty)$  con  $0 \leq u_m \leq f$ , tales que  $\|f - u_m\|_{L^1(P)} \rightarrow 0$ .*

A partir del resultado anterior obtenemos la siguiente relación entre integrales respecto de la probabilidad empírica y poblacional.

**Lema 3.2.13.** *Dada  $f : \mathbb{R}^d \rightarrow [0, +\infty)$  con  $f \in L^1(P)$ , entonces se cumple c.s. [Pr] que para todo  $F \subset \mathbb{R}^d$  cerrado,*

$$\limsup_n \int_F f dP_{n,\cdot} \leq \int_F f dP.$$

*Demostración.* Sea  $F \subset \mathbb{R}^d$  un conjunto cerrado y  $\{u_m\}_m$  una sucesión de funciones en las condiciones del Teorema 3.2.12. Como  $P_{n,\cdot} \Rightarrow P$  c.s. [Pr] y  $u_m I_F$  está acotada y es superiormente semicontinua para todo  $m \in \mathbb{N}$ , tenemos que

$$\limsup_n \int_F u_m dP_{n,\cdot} = \limsup_n \int u_m I_F dP_{n,\cdot} \leq \int u_m I_F dP \quad \text{c.s. [Pr].}$$

Como  $u_m \leq f$ , tenemos c.s. [Pr] que

$$\limsup_n \int_F u_m dP_{n,\cdot} \leq \int_F f dP.$$

Por otra parte

$$\left| \int_F f dP_{n,\cdot} - \int_F u_m dP_{n,\cdot} \right| = \int_F (f - u_m) dP_{n,\cdot} \leq \int (f - u_m) dP_{n,\cdot}.$$

Como  $\{u_m\}_m$  es una familia numerable de funciones, tenemos c.s. [Pr] que para cualquier  $m \in \mathbb{N}$ ,

$$\int (f - u_m) dP_{n,\cdot} = \|(f - u_m)\|_{L^1(P_{n,\cdot})} \xrightarrow{n \rightarrow \infty} \|(f - u_m)\|_{L^1(P)}.$$

Finalmente tenemos que casi seguro [Pr]

$$\begin{aligned} \limsup_n \int_F f dP_{n,\cdot} &= \limsup_n \left( \int_F (f - u_m) dP_{n,\cdot} + \int_F u_m dP_{n,\cdot} \right) \\ &\leq \limsup_n \int_F (f - u_m) dP_{n,\cdot} + \limsup_n \int_F u_m dP_{n,\cdot} \\ &\leq \|f - u_m\|_{L^1(P)} + \int_F f dP, \end{aligned}$$

para todo  $m \in \mathbb{N}$  y  $F \subset \mathbb{R}^d$  cerrado, lo que gracias al Teorema 3.2.12 nos conduce al resultado deseado.  $\square$

**Proposición 3.2.14.** *Si  $\alpha \in (0, 1]$  y  $Q \in P^\alpha$ , entonces existe casi seguro [Pr] una sucesión  $\{Q_{n_k, \cdot}\}_k \subset \mathbb{P}$  tal que  $Q_{n_k, \cdot} \in P_{n_k, \cdot}^\alpha$  para todo  $k \in \mathbb{N}$  y  $Q_{n_k, \cdot} \Rightarrow Q$ . Además si  $g$  es la derivada de Radon-Nikodym de  $Q$  con respecto a  $P$ , los elementos de esta subsucesión son de la forma,*

$$Q_{n_k, \cdot}(A) = \frac{1}{\|g\|_{L^1(P_{n_k, \cdot})}} \int_A g dP_{n_k, \cdot} \quad \text{para todo } A \in \mathcal{B}_d.$$

*Demostración.* Sea  $\alpha \in (0, 1]$ , se cumple  $Q(A) = \int_A g dP$  para todo  $A \in \mathcal{B}_d$ .

Definimos  $Q_{n, \cdot} : \mathcal{B}_d \rightarrow [0, +\infty)$  como

$$Q_{n, \cdot}(A) = \frac{1}{\|g\|_{L^1(P_{n, \cdot})}} \int_A g dP_{n, \cdot} \quad \text{para todo } A \in \mathcal{B}_d.$$

Esta aplicación está bien definida c.s [Pr], al menos a partir de un  $n$  suficientemente grande.

Además se trata de una probabilidad,  $Q_{n, \cdot} \in \mathbb{P}$  y por el Lema 3.2.13 tenemos que casi seguro [Pr] para todo  $F \subset \mathbb{R}^d$  cerrado,

$$\limsup_n Q_{n, \cdot}(F) = \limsup_n \frac{1}{\|g\|_{L^1(P_{n, \cdot})}} \int_F g dP_{n, \cdot} \leq \int_F g dP = Q(F),$$

y según el Teorema 0.1.1,  $Q_{n, \cdot} \Rightarrow Q$  c.s. [Pr].

Por otra parte para cualquier  $A \in \mathcal{B}_d$ ,

$$Q_{n, \cdot}(A) \leq \frac{1}{\alpha \|g\|_{L^1(P_{n, \cdot})}} P_{n, \cdot}(A),$$

pero según el Corolario 3.2.11,  $\Pr(\{\omega : \|g\|_{L^1(P_{n, \omega})} \geq 1 \text{ i.o.}\}) = 1$  y por lo tanto existe casi seguro [Pr] una subsucesión  $\{Q_{n_k, \cdot}\}_k \subset \{Q_{n, \cdot}\}_n$  (la subsucesión puede depender de  $\omega$ ) tal que para todo  $A \in \mathcal{B}_d$ ,

$$Q_{n_k, \cdot}(A) \leq \frac{1}{\alpha \|g\|_{L^1(P_{n_k, \cdot})}} P_{n_k, \cdot}(A) \leq \alpha^{-1} P_{n_k, \cdot}(A),$$

y finalmente  $Q_{n_k, \cdot} \in P_{n_k, \cdot}^\alpha$  para todo  $k \in \mathbb{N}$  lo que completa la demostración.  $\square$



**Teorema 3.2.15.** Dada  $P_{n,\cdot}$  la distribución empírica asociada con  $P$  y  $\alpha \in (0, 1]$ ,

i. si  $Q \in P^\alpha$ , entonces  $Q \in \limsup_n P_{n,\cdot}^\alpha$  c.s. [Pr];

ii.  $\limsup_n P_{n,\cdot}^\alpha \subset P^\alpha$  c.s. [Pr].

*Demostración.* La primera parte de este resultado es inmediata a partir de la Proposición 3.2.14. En cuanto a la segunda, dado  $Q \in \limsup_n P_{n,\cdot}^\alpha$ , existe una sucesión  $\{Q_{n_k}\}_k \subset \mathbb{P}$  con  $Q_{n_k} \in P_{n_k,\cdot}^\alpha$  y tal que  $Q_{n_k} \Rightarrow Q$ . Tenemos que  $P_{n,\cdot} \Rightarrow P$  c.s. [Pr], con lo que por el Lema 3.2.5 obtenemos que  $Q \in P^\alpha$  c.s. [Pr].  $\square$

### 3.3. Regiones centrales recortadas

Dada una probabilidad  $P \in \mathbb{P}$  denotamos por  $L(P)$  a un estimador de localización de  $P$  que toma cualquier tipo de valor, ya sea un único punto o un conjunto.

*Definición 3.3.1.* Para  $\alpha \in (0, 1]$  y un estimador de localización  $L$ , definimos como *región central  $\alpha$ -recortada de  $P$  inducida por  $L$*  al conjunto de estimadores de localización de todas las probabilidades de  $P^\alpha$ , al que denotamos por  $D_L^\alpha(P)$ , es decir,

$$D_L^\alpha(P) := \bigcup_{Q \in P^\alpha} L(Q) = L(P^\alpha).$$

La *función de profundidad* inducida por las regiones centrales recortadas se define de la manera habitual.

*Definición 3.3.2.* La *profundidad* de un punto  $z \in \mathbb{R}^d$  con respecto a una probabilidad  $P$  y un estimador de localización  $L$  se define como,

$$D_L(z; P) := \sup\{\alpha \in (0, 1] : z \in D_L^\alpha(P)\}.$$

A continuación se enuncian dos condiciones sobre estimadores de localización con el objetivo de derivar algunas conclusiones sobre las regiones centrales recortadas inducidas por ellos.

Dada una probabilidad  $P \in \mathbb{P}$  y una sucesión de probabilidades  $\{P_n\}_n \subset \mathbb{P}$  que convergen débilmente hacia  $P$ ,  $P_n \Rightarrow P$ , decimos que un estimador de localización  $L$  satisface la propiedad,

C1 si  $\limsup_n L(P_n) = L(P)$ ;

C2 si  $L(P)$  es compacto y  $\lim_n d_H(L(P_n), L(P)) = 0$ .

Claramente la condición C2 implica la condición C1.

### 3.4. Propiedades

En esta sección proponemos condiciones suficientes sobre los estimadores de localización que dan lugar a las regiones centrales recortadas para que éstas satisfagan algunas propiedades de especial interés extraídas de Zuo y Serfling (2000c).

Se comienza con propiedades generales de las familias de regiones centrales, así en el Teorema 3.4.1 se prueba que todas las regiones centrales recortadas son *anidadas*, en el Teorema 3.4.2 se imponen condiciones para obtener regiones centrales *convexas* y el Teorema 3.4.3 demuestra que cuando un estimador de localización es equivariante afín, entonces las regiones centrales que induce son también *equivariantes afines*. Utilizando las condiciones C1 y C2, los Teoremas 3.4.4 y 3.4.5 hacen referencia a la *conexión* de las regiones centrales recortadas, el Teorema 3.4.6 a su condición de conjuntos *cerrados* y el Teorema 3.4.7 y el Corolario 3.4.8 a su *compacidad*.

Para finalizar en el Teorema 3.4.9 y el Corolario 3.4.10 consideramos el *comportamiento asintótico* de las regiones centrales recortadas.

### 3.4.1. Propiedades generales

#### Regiones anidadas

**Teorema 3.4.1.** Sean  $\alpha, \beta \in (0, 1]$  tales que  $\alpha \leq \beta$ ,  $P \in \mathbb{P}$  y  $L$  cualquier estimador de localización, entonces  $D_L^\alpha(P) \supset D_L^\beta(P)$ .

*Demostración.* Según el Lema 3.2.1 se cumple  $P^\alpha \supset P^\beta$  y el resultado es inmediato.  $\square$

#### Regiones Convexas

**Teorema 3.4.2.** Si para dos distribuciones de probabilidad cualesquiera,  $Q_1, Q_2 \in \mathbb{P}$  y cualquier  $\lambda \in (0, 1)$ , se cumple  $\lambda L(Q_1) + (1 - \lambda)L(Q_2) \subset L(\lambda Q_1 + (1 - \lambda)Q_2)$ , entonces las regiones centrales recortadas inducidas por  $L$  son convexas.

*Demostración.* Sean  $x, y \in D_L^\alpha(P)$ , entonces  $x \in L(Q_1)$  e  $y \in L(Q_2)$  para algunas probabilidades  $Q_1, Q_2 \in P^\alpha$ . Dado  $\lambda \in (0, 1)$ , se cumple  $\lambda x + (1 - \lambda)y \in \lambda L(Q_1) + (1 - \lambda)L(Q_2) \subset L(\lambda Q_1 + (1 - \lambda)Q_2)$  y como por el Lema 3.2.1,  $P^\alpha$  es convexo, se cumple  $\lambda Q_1 + (1 - \lambda)Q_2 \in P^\alpha$ .  $\square$

#### Equivarianza Afín

**Teorema 3.4.3.** Dado cualquier estimador de localización  $L$  equivariante afín, las regiones centralres  $D_L^\alpha(P)$  son también equivariantes afines.

*Demostración.* Dada una matriz  $A \in \mathbb{R}^{d \times d}$  no singular,  $b \in \mathbb{R}^d$  y  $z \in D_L^\alpha(P_{A,b})$ , existe una probabilidad  $Q \in (P_{A,b})^\alpha$  tal que  $z \in L(Q)$ . Construimos la aplicación  $Q' : \mathcal{B}_d \rightarrow \mathbb{R}$  definida como  $Q'(B) = Q(AB + b)$ . Como  $A$  es no singular  $Q'$  será una probabilidad,  $Q' \in \mathbb{P}$ . Además  $Q'(B) = Q(AB + b) \leq \alpha^{-1} P_{A,b}(AB + b) = \alpha^{-1} P(B)$  y por lo tanto  $Q' \in P^\alpha$ . Finalmente, se cumple  $(Q')_{A,b}(B) = Q'(A^{-1}(B - b)) = Q(B)$  y así tenemos que  $z \in L((Q')_{A,b}) = AL(Q') + b$  lo que implica que  $z \in AD_L^\alpha(P) + b$ .

Recíprocamente, si  $z \in AD_L^\alpha(P) + b$ , entonces existe  $x \in L(Q)$  para alguna probabilidad  $Q \in P^\alpha$  tal que  $z = Ax + b$ . Como  $L(Q_{A,b}) = AL(Q) + b$ , se ha de cumplir  $z \in L(Q_{A,b})$ . Finalmente, se tiene que

$$Q_{A,b}(B) = Q(A^{-1}(B - b)) \leq \alpha^{-1}P(A^{-1}(B - b)) = \alpha^{-1}P_{A,b}(B)$$

y así  $Q_{A,b} \in (P_{A,b})^\alpha$  con lo que se cumple  $z \in D_L^\alpha(P_{A,b})$ .  $\square$

### Regiones Conexas

**Teorema 3.4.4.** *Si un estimador de localización  $L$  satisface la condición C1 y  $L(P)$  es conexo para cualquier  $P \in \mathbb{P}$ , entonces  $D_L^\alpha(P)$  es conexo.*

*Demostración.* En el caso trivial, cuando  $D_L^\alpha(P) = \emptyset$ , el resultado es obvio. En otro caso, supongamos que  $D_L^\alpha(P)$  no es conexo, así habrá dos conjuntos disjuntos abiertos  $V_1, V_2 \subset \mathbb{R}^d$  tales que  $D_L^\alpha(P) \cap V_1 \neq \emptyset$ ,  $D_L^\alpha(P) \cap V_2 \neq \emptyset$  y  $D_L^\alpha(P) \subset V_1 \cup V_2$ .

Dados  $a \in V_1 \cap D_L^\alpha(P)$  y  $b \in V_2 \cap D_L^\alpha(P)$ , habrá  $Q_a, Q_b \in P^\alpha$  tales que  $a \in L(Q_a)$  y  $b \in L(Q_b)$ .

Definimos los conjuntos de probabilidades  $A_P = \{Q \in P^\alpha : L(Q) \subset V_1\}$  y  $B_P = \{Q \in P^\alpha : L(Q) \subset V_2\}$ . Como  $L$  toma valores conexos, entonces cualquier  $Q \in P^\alpha$  pertenece bien a  $A_P$  o bien a  $B_P$ .

Sea  $Q_\lambda = (1 - \lambda)Q_a + \lambda Q_b$  que pertenece a  $P^\alpha$  por ser este conjunto convexo según el Lema 3.2.1. Si  $\{\lambda_n\}_n \subset [0, 1]$  satisface  $\lim_n \lambda_n = \lambda$  entonces obviamente  $Q_{\lambda_n} \Rightarrow Q_\lambda$ .

Definimos  $\mu_A = \sup\{\mu : Q_\lambda \in A_P, \forall \lambda \in [0, \mu]\}$  y  $\mu_B = \inf\{\mu : Q_\lambda \in B_P, \forall \lambda \in [\mu, 1]\}$ . Como  $L$  verifica C1, tenemos  $L(Q_{\mu_A}) \subset V_1$  y  $L(Q_{\mu_B}) \subset V_2$ . Por otra parte no es posible que simultáneamente  $\mu_A = 1$  y  $\mu_B = 0$ .

Supongamos que  $\mu_A < 1$  (si  $\mu_B > 0$  el razonamiento sería análogo). Existe, por tanto, una sucesión  $\{\lambda_n\}_n \subset [0, 1]$  con  $\lambda_n > \mu_A$  tal que  $\lambda_n \rightarrow \mu_A$  y  $L(Q_{\lambda_n}) \subset V_2$ . Como  $L$  satisface C1,  $L(Q_{\mu_A}) = \limsup L(Q_{\lambda_n})$  lo que contradice que  $V_1$  y  $V_2$  sean conjuntos abiertos disjuntos.  $\square$

Si consideramos estimadores de localización con valor unipuntual, obtenemos el siguiente resultado.

**Teorema 3.4.5.** *Si  $L : \mathbb{P} \rightarrow \mathbb{R}^d$  es continua, entonces  $D_L^\alpha(P)$  es conexo por caminos.*

*Demostración.* Cualquier transformación continua de un conjunto conexo por caminos es conexa por caminos y por el Corolario 3.2.2,  $P^\alpha$  es conexo por caminos.  $\square$

### Regiones Cerradas

**Teorema 3.4.6.** *Si  $L$  satisface C1, entonces  $D_L^\alpha(P)$  es cerrado.*

*Demostración.* Sea  $\{x_n\}_n \subset D_L^\alpha(P)$  una sucesión convergente tal que  $\lim x_n = x$ . Como para todo  $n$ , se cumple  $x_n \in D_L^\alpha(P)$ , entonces existe una sucesión de probabilidades  $\{Q_n\}_n \subset P^\alpha$  tal que  $x_n \in L(Q_n)$ . Por el Lema 3.2.3,  $P^\alpha$  es tight y en consecuencia  $\{Q_n\}_n$  también es tight. Según el Teorema 0.1.5 existe una subsucesión  $\{Q_{n_k}\}_k \subset \{Q_n\}_n$  que converge débilmente a cierto  $Q \in \mathbb{P}$ , esto es  $Q_{n_k} \Rightarrow Q$  y por el Lema 3.2.4, se concluye que  $Q \in P^\alpha$ . Ahora  $x \in \limsup_k L(Q_{n_k})$  y como  $L$  satisface C1, concluimos  $x \in L(Q) \subset D_L^\alpha(P)$ .  $\square$

### Regiones Compactas

**Teorema 3.4.7.** *Si  $L$  satisface la condición C2, entonces  $D_L^\alpha(P)$  es compacto.*

*Demostración.* Por la condición C2 tenemos que  $L(\overline{\mathbb{C}}) \subset \overline{L(\mathbb{C})}$  para todo  $\mathbb{C} \subset \mathbb{P}$  y por lo tanto  $L$  es continua como aplicación de  $\mathbb{P}$  en el espacio de conjuntos compactos no vacíos de  $\mathbb{R}^d$  con respecto a la métrica de Hausdorff. Finalmente,  $L(P^\alpha)$  es compacto ya que por el Lema 3.2.4,  $P^\alpha$  es compacto y la igualdad  $L(P^\alpha) = D_L^\alpha(P)$  concluye la demostración.  $\square$

Claramente la condición C1 no es suficiente para garantizar la acotación de las regiones centrales recortadas. Por otra parte el Teorema 3.4.7 se puede particularizar para estimadores unipuntuales y continuos.

**Corolario 3.4.8.** *Si  $L : \mathbb{P} \rightarrow \mathbb{R}^d$  es continua, entonces  $D_L^\alpha(P)$  es compacto.*

### 3.4.2. Resultados asintóticos

El resultado principal que obtenemos sobre el comportamiento asintótico de las regiones centrales recortadas establece que si un estimador de localización satisface la condición C1, las regiones centrales recortadas de una distribución poblacional coinciden casi seguro con el límite superior de las regiones centrales recortadas de su distribución empírica asociada.

**Teorema 3.4.9.** *Si  $L$  satisface la condición C1, entonces*

$$D_L^\alpha(P) = \limsup_n D_L^\alpha(P_{n,\cdot}) \quad c.s. \text{ [Pr]}$$

*Demostración.* Sea  $D$  un conjunto numerable y denso en  $D_L^\alpha(P)$ . Obsérvese que la existencia de dicho conjunto está garantizada al ser  $\mathbb{R}^d$  segundo numerable. Dado  $x \in D$ , para algún  $Q \in P^\alpha$ ,  $x \in L(Q)$  y por el Teorema 3.2.15, existe una sucesión  $\{Q_{n_k,\cdot}\}_k$  tal que  $Q_{n_k,\cdot} \in P_{n_k,\cdot}^\alpha$  y  $Q_{n_k,\cdot} \Rightarrow Q$  c.s. [Pr]. Como  $L$  satisface C1,  $\limsup_k L(Q_{n_k,\cdot}) = L(Q)$  c.s. [Pr] y entonces existe una subsucesión  $\{x_{n_{k_l}}\}_l$  tal que  $x_{n_{k_l}} \in L(Q_{n_{k_l},\cdot})$  y  $\lim_l x_{n_{k_l}} = x$ , por lo tanto  $x \in \limsup_n D_L^\alpha(P_{n,\cdot})$  c.s. [Pr]. Por la numerabilidad de  $D$  tenemos que  $D \subset \limsup_n D_L^\alpha(P_{n,\cdot})$  c.s. [Pr].

Como  $\limsup D_L^\alpha(P_{n,\cdot})$  es cerrado,  $D_L^\alpha(P) \subset \overline{D} \subset \limsup D_L^\alpha(P_{n,\cdot})$  c.s. [Pr].

Para probar la otra inclusión, sea  $x \in \limsup D_L^\alpha(P_{n,\cdot})$ . Existe una sucesión  $\{x_{n_k}\}_k$  tal que  $x_{n_k} \in D_L^\alpha(P_{n_k,\cdot})$  y  $x = \lim_k x_{n_k}$ . Así para todo  $k \in \mathbb{N}$ ,  $x_{n_k} \in L(Q_{n_k,\cdot})$  para algún  $Q_{n_k,\cdot} \in P_{n_k,\cdot}^\alpha$ . Como  $P_{n,\cdot} \Rightarrow P$  c.s. [Pr], entonces según el Lema 3.2.6, el conjunto  $\cup_{n=1}^\infty P_n^\alpha$  es también tight c.s. [Pr]. En consecuencia existe c.s. [Pr] una subsucesión  $\{Q_{n_{k_l},\cdot}\}_l$  que converge débilmente a cierta probabilidad  $Q \in \mathbb{P}$  y según el Lema 3.2.5 concluimos  $Q \in P^\alpha$ . Como  $L$  verifica la condición C1,  $x \in \limsup L(Q_{n_{k_l},\cdot}) = L(Q)$  y por lo tanto  $\limsup D_L^\alpha(P_{n,\cdot}) \subset D_L^\alpha(P)$  c.s. [Pr].  $\square$

**Corolario 3.4.10.** *Si  $L : \mathbb{P} \rightarrow \mathbb{R}^d$  es continua, entonces*

$$D_L^\alpha(P) = \limsup_n D_L^\alpha(P_{n,\cdot}) \quad c.s. \text{ [Pr]}$$

## 3.5. Ejemplos

Veremos que la familia de regiones centrales recortadas generaliza los *intervalos intercuantílicos* (para variables aleatorias), las *regiones centrales del zonoide* y las *regiones centrales de la envolvente de zonoides*. Para finalizar el capítulo se dará un ejemplo de su estimador de localización que induce regiones centrales conexas, pero no convexas.

### 3.5.1. Intervalos intercuantílicos

Para  $d = 1$ , los intervalos intercuantílicos de (3.1) están íntimamente relacionados con las regiones centrales recortadas inducidas por la mediana. Sea  $\text{Me}$  el intervalo mediano, es decir, el conjunto de todos los puntos que satisfacen la definición de mediana.

**Teorema 3.5.1.** *Dada una probabilidad  $P$  sobre  $\mathbb{R}$ ,  $\text{Me}$  el intervalo mediano y cualquier  $\alpha \in (0, 1/2]$ ,*

$$\left[ \inf\{x : P((-\infty, x]) \geq \alpha\}, \sup\{x : P([x, +\infty)) \geq \alpha\} \right] = D_{\text{Me}}^{2\alpha}(P).$$

*Demostración.* Es sencillo probar que el ínfimo y el supremo de la fórmula anterior son en realidad un mínimo y un máximo. Así dado  $y \in \left[ \inf\{x : P((-\infty, x]) \geq \alpha\}, \sup\{x : P([x, +\infty)) \geq \alpha\} \right]$ , entonces  $P((-\infty, y]) \geq \alpha$  y  $P([y, +\infty)) \geq \alpha$ . Supongamos que  $P(\{y\}) = 0$ , definimos la aplicación  $Q : \mathcal{B}_d \rightarrow \mathbb{R}$  que asigna a cada  $A \in \mathcal{B}_d$ ,

$$Q(A) = \frac{1}{2P((-\infty, y])} P(A \cap (-\infty, y]) + \frac{1}{2P([y, +\infty))} P(A \cap [y, +\infty))$$

que es claramente una probabilidad,  $Q \in \mathbb{P}$  y además  $Q \leq (2\alpha)^{-1}P$ , con lo que  $Q \in P^{2\alpha}$  e  $y \in D_{\text{Me}}^{2\alpha}(P)$ . Si  $P$  tiene un átomo en  $y$ , la demostración de esta inclusión es análoga, sólo hay que distribuir la masa de tal modo que la mitad quede a la izquierda y la mitad a la derecha de  $y$ .

Por otra parte sea  $y \in D_{\text{Me}}^{2\alpha}(P)$ , entonces existe  $Q \in P^{2\alpha}$  tal que  $Q((-\infty, y]) \geq 1/2$  y  $Q([y, +\infty)) \geq 1/2$ , como  $Q \leq (2\alpha)^{-1}P$  se cumple

que  $P((-\infty, y]) \geq \alpha$  y  $P([y, +\infty)) \geq \alpha$ , por lo tanto  $\inf\{x : P((-\infty, x]) \geq \alpha\} \leq y \leq \sup\{x : P([x, +\infty)) \leq 1 - \alpha\}$ .  $\square$

### 3.5.2. Regiones centrales del zonoide

Las regiones centrales del zonoide establecidas en la Definición 0.4.3 pertenecen a la familia de regiones centrales recortadas.

**Teorema 3.5.2.** *Si tomamos la esperanza como estimador de localización, obtenemos las regiones centrales del zonoide como regiones centrales recortadas, es decir,  $ZD^\alpha(P) = D_{\mathbb{E}}^\alpha(P)$  para cualquier  $\alpha \in (0, 1]$ ,  $P \in \mathbb{P}$ .*

*Demostración.* Es claro que  $ZD^\alpha(P) = \{\int x dQ(x) : Q \in P^\alpha\}$ , la función  $g$  de la Definición 0.4.3 jugará el papel de la derivada de Radon-Nikodym de  $Q$  con respecto a  $P$ .  $\square$

Aunque éstas son propiedades conocidas de Koshevoy y Mosler (1997b) (Teoremas 5.3, 5.4 y 5.7), de los resultados que se han probado a lo largo del presente capítulo se deduce fácilmente que las regiones centrales del zonoide son *anidadas* (Teorema 3.4.1), *convexas* (Teorema 3.4.2 y linealidad de la integral) y *equivariantes afines* (Teorema 3.4.3 y equivarianza afín de la esperanza).

Para dibujar las regiones centrales del zonoide de una distribución empírica, o en general de una distribución que asigna a cada uno de los puntos de un conjunto finito la misma probabilidad, podemos hacer uso del siguiente resultado.

**Teorema 3.5.3 (Teorema 4.1 en Koshevoy y Mosler, 1997b).** *Sea  $P \in \mathbb{P}$  y  $\{x_1, \dots, x_n\} \subset \mathbb{R}^d$  tal que  $P(\{x_i\}) = 1/n$  para  $i \in \{1, \dots, n\}$ . Si  $\alpha \in [k/n, (k+1)/n]$  para cualquier  $k \in \{1, \dots, n-1\}$ , entonces*



$$\text{ZD}^\alpha(P) = \text{co} \left\{ \frac{1}{\alpha n} \sum_{j=1}^k x_{i_j} + \left(1 - \frac{k}{\alpha n}\right) x_{i_{k+1}} : \{i_1, \dots, i_{k+1}\} \subset N \right\},$$

donde  $N = \{1, \dots, n\}$ . Si  $\alpha \in (0, 1/n]$ ,

$$\text{ZD}^\alpha(P) = \text{co}\{x_1, \dots, x_n\}.$$

El Teorema 3.5.3 se demuestra de un modo sencillo utilizando el Teorema 3.5.2 y la caracterización de los cuerpos convexos a través su función soporte.

*Ejemplo 3.5.1.* Si consideramos la distribución de probabilidad  $P$  que asigna probabilidad  $1/4$  a cada uno de los puntos  $(1, 1)$ ,  $(1, -1)$ ,  $(-1, 1)$  y  $(-1, -1)$ , obtenemos las regiones centrales del zonoide que aparecen en la Figura 3.1.

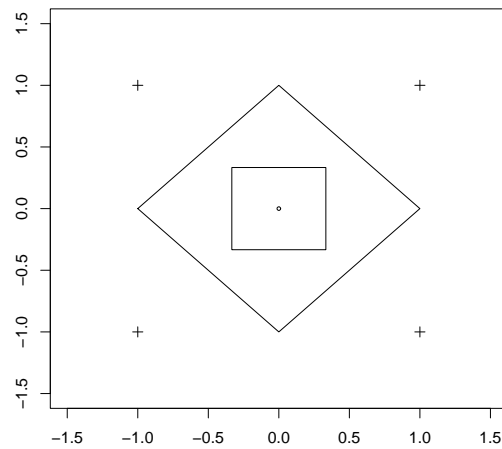


Figura 3.1: Contornos de  $\text{ZD}^1(P)$ ,  $\text{ZD}^{3/4}(P)$  y  $\text{ZD}^{1/2}(P)$ .

### 3.5.3. Regiones centrales de la envolvente de zonoides

Las regiones centrales de la envolvente de zonoides establecidas en la Definición 1.1.3 pertenecen a la familia de regiones centrales recortadas.

**Lema 3.5.4.** *Dada cualquier probabilidad  $P \in \mathbb{P}$  y  $\alpha \in (0, 1]$ ,*

$$\alpha^{-1}\text{proj}_\alpha(W(P)) = \bigcup_{Q \in P^\alpha} Z(Q).$$

*Demostración.* Dado  $z \in \mathbb{R}^d$ , las siguientes condiciones son claramente equivalentes:  $z \in \alpha^{-1}\text{proj}_\alpha(W(P))$ ;  $(\alpha, \alpha z) \in W(P) = \text{co}\{\tilde{Z}(P) \cup \{1\} \times Z(P)\}$ ; puesto que  $Z(P)$  es la proyección sobre las  $d$  últimas coordenadas de  $\tilde{Z}(P)$ , se tiene  $(\beta, \alpha z) \in \tilde{Z}(P)$  para algún  $\beta \leq \alpha$ ;  $\alpha\beta^{-1}z \in ZD^\beta(P)$  para algún  $\beta \leq \alpha$ ;  $z \in \beta\alpha^{-1}ZD^\beta(P)$  para algún  $\beta \leq \alpha$ ;  $z \in \beta\alpha^{-1}\{\int x dQ(x) : Q \in P^\beta\}$  para algún  $\beta \leq \alpha$ ;  $z \in \cup_{0 < \beta \leq \alpha} \cup_{Q \in P^\beta} \beta\alpha^{-1}Z(Q) = \cup_{Q \in P^\alpha} Z(Q)$ . Para probar esta última igualdad sólo tenemos que comprobar que  $\cup_{Q \in P^\beta} \beta Z(Q) \subset \cup_{Q \in P^\alpha} \alpha Z(Q)$  para cualquier  $0 < \beta \leq \alpha$ . Sea  $Q_\beta \in P^\beta$ , entonces  $Q_\alpha = \beta\alpha^{-1}Q_\beta + (\alpha - \beta)\alpha^{-1}(1 - \beta)^{-1}(P - \beta Q_\beta) \in P^\alpha$ , y claramente para cualquier  $u \in \mathbb{R}^d$ ,  $\int \beta\alpha^{-1}\langle x, u \rangle_+ dQ_\beta(x) \leq \int \langle x, u \rangle_+ dQ_\alpha(x)$ , por lo tanto  $\beta Z(Q_\beta) \subset \alpha Z(Q_\alpha)$  y esto finaliza la demostración.  $\square$

**Teorema 3.5.5.** *Si tomamos el zonoide como estimador de localización, entonces obtenemos las regiones centrales  $WD^\alpha(P)$  como regiones centrales recortadas, es decir,  $WD^\alpha(P) = D_Z^\alpha(P)$  para cualquier  $\alpha \in (0, 1]$ ,  $P \in \mathbb{P}$ .*

*Demostración.* Por definición de regiones centrales recortadas,  $D_Z^\alpha(P) = \cup_{Q \in P^\alpha} Z(Q)$  y según el Lema 3.5.4,  $\cup_{Q \in P^\alpha} Z(Q) = \alpha^{-1}\text{proj}_\alpha(W(P))$ . Como por definición  $\alpha^{-1}\text{proj}_\alpha(W(P)) = WD^\alpha(P)$ , concluimos  $D_Z^\alpha(P) = WD^\alpha(P)$ .  $\square$

A partir de los resultados de este capítulo es fácil deducir que estas regiones centrales son *anidadas* (Teorema 3.4.1) y *convexas* (Teorema 3.4.2 y linealidad de la integral). Como el zonoide de una distribución no es equivariente afín, no sorprende que las regiones centrales de la envolvente de zonoides tampoco lo sean.

Podemos obtener un resultado análogo al Teorema 3.5.3 que nos simplifique la tarea de dibujar las regiones de la envolvente de zonoides de una distribución de probabilidad que asigna la misma masa a cada uno de los puntos de un conjunto finito.

**Teorema 3.5.6.** *Sea  $P \in \mathbb{P}$  y  $\{x_1, \dots, x_n\} \subset \mathbb{R}^d$  tal que  $P(\{x_i\}) = 1/n$  para  $i \in \{1, \dots, n\}$ . Si  $\alpha \in (k/n, (k+1)/n]$  para cualquier  $k \in \{0, 1, \dots, n-1\}$ , entonces*

$$\text{WD}^\alpha(P) = \text{co} \left\{ \frac{1}{\alpha n} \sum_{j=1}^k x_{i_j} + \left(1 - \frac{k}{\alpha n}\right) x_{i_{k+1}} : \{i_1, \dots, i_{k+1}\} \subset N^* \right\}, \quad (3.2)$$

donde  $N^* = \{-n, \dots, -1, 1, \dots, n\}$  y  $x_{-i} = 0$  para todo  $i \in \{1, \dots, n\}$ .

Observar que la expresión anterior contiene el caso particular  $\alpha \in (0, 1/n]$ , en cuyo caso la envolvente de zonoides es la envolvente convexa del conjunto  $\{0, x_1, \dots, x_n\}$ ,

$$\text{WD}^\alpha(P) = \text{co}\{0, x_1, \dots, x_n\}.$$

*Demostración.* Como ambos conjuntos son convexos, es suficiente probar que sus funciones soporte coinciden. Si  $Q \in P^\alpha$ , entonces  $Q(\{x_i\}) \leq (\alpha n)^{-1}$  para  $i \in \{1, \dots, n\}$  y  $Q(\{x\}) = 0$  si  $x \notin \{x_1, \dots, x_n\}$ . Sea  $u \in \mathbb{R}^d \setminus \{0\}$ ,

$$\begin{aligned} h(\text{WD}^\alpha(P), u) &= \sup_{Q \in P^\alpha} h(Z(Q), u) \\ &= \sup \left\{ \sum_{i=1}^n \langle x_i, u \rangle_+ q_i : 0 \leq q_i \leq \frac{1}{\alpha n}, \sum_{i=1}^n q_i = 1 \right\}. \end{aligned}$$

Si ordenamos los puntos  $\{x_1, \dots, x_n\}$  de tal modo que  $\langle x_{i_1}, u \rangle \geq \langle x_{i_2}, u \rangle \geq \dots \geq \langle x_{i_n}, u \rangle$ , el supremo anterior se alcanzará dotando de masa  $(\alpha n)^{-1}$  a los  $k$  primeros puntos  $\{x_{i_1}, \dots, x_{i_k}\}$  siendo  $k$  el mayor entero tal que  $k/(\alpha n) < 1$ , es decir, satisface las condiciones  $k/(\alpha n) < 1$  y  $(k+1)/(\alpha n) \geq 1$ ,

por tanto  $\alpha \in (k/n, (k+1)/n]$ . La masa restante  $1 - k/(\alpha n)$  se le asignará al punto siguiente,  $x_{i_{k+1}}$ . La función soporte de la envolvente de zonoides en la dirección  $u$  adoptará la forma

$$h(\text{WD}^\alpha(P), u) = \frac{1}{\alpha n} \sum_{j=1}^k \langle x_{i_j}, u \rangle_+ + \left(1 - \frac{k}{\alpha n}\right) \langle x_{i_{k+1}}, u \rangle_+,$$

que coincide con la función soporte en dirección  $u$  del conjunto a la derecha en la expresión (3.2).  $\square$

*Ejemplo 3.5.2.* Con la distribución de probabilidad del Ejemplo 3.5.1, obtenemos las regiones centrales de la envolvente de zonoides que aparecen en la Figura 3.2.

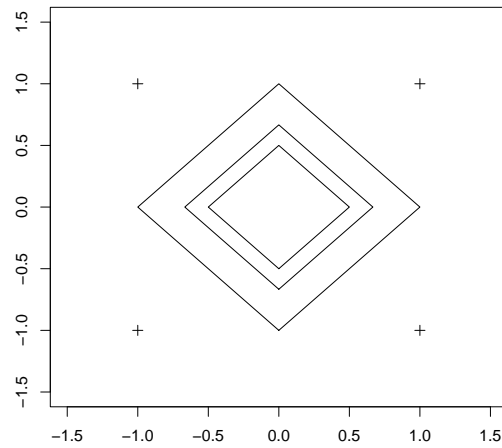


Figura 3.2: Contornos de  $\text{WD}^1(P)$ ,  $\text{WD}^{3/4}(P)$  y  $\text{WD}^{1/2}(P)$ .

### 3.5.4. Regiones centrales conexas, pero no convexas

Tomamos cualquier distribución de probabilidad discreta sobre  $\mathbb{R}^d$ . Sea  $I \subset \mathbb{N}$  y  $x_i \in \mathbb{R}^d$  tales que  $P(\{x_i\}) = p_i > 0$  para todo  $i \in I$  y con  $\sum_{i \in I} p_i = 1$ . Como estimador de localización, asignamos a tal distribución el valor medio de una versión reescalada suya,

$$M(P) = \sum_{i \in I} x_i \frac{\sqrt{p_i}}{\sum_{j \in I} \sqrt{p_j}}.$$

A partir de los resultados de este capítulo es fácil probar que estas regiones centrales son *anidadas*, *equivariantes afines* y *conexas por caminos*. En la gráfica del siguiente ejemplo observamos que estas regiones centrales no son necesariamente convexas.

*Ejemplo 3.5.3.* Con la distribución de probabilidad del Ejemplo 3.5.1 y el estimador de localización  $M$ , obtenemos las regiones centrales que aparecen en la Figura 3.3.

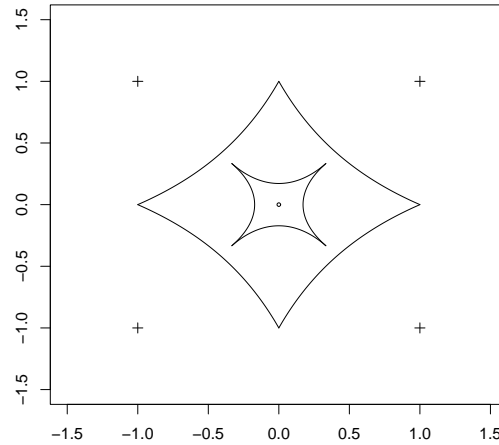


Figura 3.3: Contornos de  $D_M^1(P)$ ,  $D_M^{3/4}(P)$  y  $D_M^{1/2}(P)$ .

### 3.6. Comentarios y problemas abiertos

Además de algunos resultados particulares sobre las regiones centrales de la envolvente de zonoides, la aportación más importante que se ha hecho en este capítulo es la definición del recorte de una probabilidad, la obtención de diversos resultados relativos a tal recorte y su utilización para definir regiones centrales. Creemos que los recortes de una probabilidad constituyen un marco apropiado para estudiar la profundidad multivariante y prueba de ello es que varias funciones de profundidad clásicas pueden englobarse dentro de las construidas a partir de los recortes de una probabilidad.

El problema más clásico que surge cuando se define una nueva familia de funciones de profundidad es la definición de  $L$ -estimadores multivariantes y la obtención de su convergencia asintótica. Tales estimadores, en su forma más general y siguiendo la notación que hemos propuesto para las regiones centrales recortadas, adoptarían la expresión

$$\sum_{i=1}^n f(D_L(\xi_i(\cdot); P_{n,\cdot}))\xi_i(\cdot),$$

donde  $f$  es una función real de variable real y, como anteriormente,  $\xi_1(\cdot), \xi_2(\cdot), \dots, \xi_n(\cdot)$  es una muestra aleatoria simple de variables aleatorias independientes idénticamente distribuidas con distribución empírica  $P_n$ .

Una de las propiedades esenciales de un estimador de localización es su robustez. Las regiones centrales construidas para una probabilidad son estimadores de localización que toman valor de conjunto. Puede ser interesante estudiar la *estabilidad* de una profundidad cuando toda la masa se concentra en alguna parte de su soporte, o en general cuando la información la obtenemos a partir de algún elemento de  $P^\alpha$ . Así, si  $D(\cdot; \cdot)$  es una profundidad cualquiera, para  $x \in \mathbb{R}^d$  y  $P \in \mathbb{P}$ , podemos estudiar

$$\sup_{\alpha \in (0,1]} \{\alpha |D(x; P) - D(x; Q)| : Q \in P^\alpha\}.$$

Cuando menor fuese el valor que tomara la expresión anterior en  $x$ , mayor sería la estabilidad de la profundidad en ese punto. Esta idea puede ser uti-

lizada para definir funciones de influencia como las utilizadas en los trabajos de Romanazzi (2001) y Chen y Tyler (2002).

El estudio de las funciones de profundidad es un método descriptivo enmarcado en el Análisis Multivariante no Paramétrico. Como método descriptivo tiene especial interés cuando trabajamos con distribuciones de probabilidad que tienen algún tipo de simetría multivariante o son, al menos, unimodales. Existen, no obstante, intentos de aplicar las funciones de profundidad a métodos como el Análisis Cluster. Creemos que el marco de los recortes de una probabilidad puede ser apropiado para el Análisis Cluster ya que si una distribución de probabilidad  $P \in \mathbb{P}$  describe dos agrupaciones, éstas podrían analizarse como un elemento de  $P^\alpha$  y otro de  $P^{1-\alpha}$ .

Para finalizar, puede ser interesante relacionar los recortes de una probabilidad con los órdenes estocásticos. De este modo podría decirse que una distribución es menor que otra según un determinado orden estocástico y en cierto grado. Sean  $P, Q \in \mathbb{P}$  y ' $\leq_*$ ' un orden estocástico cualquiera. Si

$$\beta = \sup\{\alpha \in (0, 1] : R \leq_* Q \text{ con } R \in P^\alpha\} \quad (3.3)$$

podemos decir que  $P$  es menor que  $Q$  según el orden  $\leq_*$  en grado  $\beta$ . Obsérvese que en la expresión (3.3) utilizamos una notación diferente a la empleada hasta ahora para los órdenes estocásticos, en la que utilizamos probabilidades en lugar de elementos aleatorios, pero esto no supone ningún inconveniente.





## Capítulo 4

# Regiones centrales integrales

A lo largo de la presente memoria se han expuesto numerosas relaciones entre órdenes estocásticos y regiones centrales. Así, el orden lineal convexo está relacionado con las regiones centrales del zonoide; se ha definido una familia de regiones centrales, la de la envolvente de zonoides, basada en el orden simétrico y existe una estrecha relación entre el orden estocástico habitual y las regiones centrales semiespaciales. Más específicamente, el orden estocástico habitual generaliza la relación de orden *menor o igual* y las regiones centrales semiespaciales son una generalización de los intervalos intercuantílicos que se definen a partir del orden en  $\mathbb{R}$ . Inspirándonos en estas relaciones y en el hecho de que la familia de órdenes estocásticos integrales es una de las más prolíficas, definimos las *regiones centrales integrales* inducidas por una familia de funciones.

Al igual que en el Capítulo 3, la profundidad vendrá determinada por la pertenencia de una probabilidad a cierto recorte de la probabilidad original, pero ahora el estimador de localización está inducido por una familia de conjuntos de funciones y será el conjunto de puntos tales que su transformación por cada una de las funciones de la familia es menor o igual que la integral de la función respecto de alguna probabilidad del recorte.

En primer lugar obtenemos las propiedades de las regiones centrales integrales a partir de propiedades de las familias de funciones que las generan.

A continuación, estudiamos el comportamiento asintótico de las regiones integrales y para finalizar, damos varios ejemplos de regiones centrales integrales (Casco y López-Díaz, 2004b).

## 4.1. Definición

Las letras caligráficas  $\mathcal{F}$  y  $\mathcal{G}$  denotarán conjuntos de funciones de  $\mathbb{R}^d$  en  $\mathbb{R}$  e  $I, J$  denotarán conjuntos de índices.

Dada una probabilidad  $P$  en  $\mathbb{R}^d$  y  $\alpha \in (0, 1]$ , consideramos las regiones centrales de nivel  $\alpha$  definidas como conjuntos de puntos para los que la imagen a través de cualquier función de cierta familia no es mayor que la integral de la misma función respecto de una probabilidad de  $P^\alpha$ . A estas regiones centrales las llamaremos *regiones centrales integrales*.

*Definición 4.1.1.* Dado  $\alpha \in (0, 1]$  y un conjunto de funciones  $\mathcal{F}$ , definimos la *región central integral de nivel  $\alpha$  para  $P$  respecto del conjunto  $\mathcal{F}$* , denotada por  $D_{\mathcal{F}}^\alpha(P)$ , como

$$\begin{aligned} D_{\mathcal{F}}^\alpha(P) &:= \left\{ x \in \mathbb{R}^d : \text{existe } Q_x \in P^\alpha, f(x) \leq \int f dQ_x, \forall f \in \mathcal{F} \right\} \\ &= \bigcup_{Q \in P^\alpha} \bigcap_{f \in \mathcal{F}} f^{-1} \left( \left( -\infty, \int f dQ \right] \right). \end{aligned}$$

Dada una familia de conjuntos de funciones  $\{\mathcal{F}_i\}_{i \in I}$ , definimos

$$D_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P) := \bigcap_{i \in I} D_{\mathcal{F}_i}^\alpha(P).$$

La *función de profundidad* inducida por las regiones centrales integrales se define del modo habitual.

*Definición 4.1.2.* La *profundidad* de un punto  $x \in \mathbb{R}^d$  con respecto a una probabilidad  $P$  y una familia de conjuntos de funciones  $\{\mathcal{F}_i\}_{i \in I}$  viene dada por,

$$D_{\{\mathcal{F}_i\}_{i \in I}}(x; P) := \sup\{\alpha \in (0, 1] : x \in D_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P)\}.$$

### 4.1.1. Preliminares

Definimos en primer lugar algunas clases de funciones y a continuación clases de conjuntos de funciones.

*Definición 4.1.3.* Una función  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  se dice *cuasi-convexa* si para cualesquiera  $x, y \in \mathbb{R}^d$  y  $\lambda \in (0, 1)$ ,  $f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\}$ . Se dice que la función es *estrictamente cuasi-convexa* si la desigualdad es estricta,  $f(\lambda x + (1 - \lambda)y) < \max\{f(x), f(y)\}$ .

*Definición 4.1.4.* Una función  $f$  se dice *inferiormente semicontinua* si sus conjuntos nivel del tipo  $\{x : f(x) \leq \gamma\}$  son cerrados para todo  $\gamma \in \mathbb{R}$ .

En el Capítulo 3 se definió cuándo una función es *superiormente semicontinua*. Este concepto está estrechamente relacionado con la semicontinuidad inferior de tal modo que una función  $f$  es superiormente semicontinua si su opuesta  $-f$  es inferiormente semicontinua.

A continuación, se introducen dos definiciones aplicables a conjuntos de funciones, en primer lugar la *comonotonía*, véase Denneberg (1994).

*Definición 4.1.5.* Dado un conjunto de funciones  $\mathcal{F}$ , decimos que es *comonótono* si para cualesquiera  $f, g \in \mathcal{F}$  no hay ningún par  $x, y \in \mathbb{R}^d$  tal que  $f(x) < f(y)$  y  $g(x) > g(y)$ .

El concepto de *semi-convexidad* que introducimos a continuación es nuevo, lo utilizamos para generar regiones integrales convexas.

*Definición 4.1.6.* Dado un conjunto de funciones  $\mathcal{F}$ , decimos que es *semi-convexo* si para cualesquiera  $x, y \in \mathbb{R}^d$  y  $\lambda \in (0, 1)$ , bien  $f(\lambda x + (1 - \lambda)y) \leq f(x)$  para todo  $f \in \mathcal{F}$  o bien  $f(\lambda x + (1 - \lambda)y) \leq f(y)$  para todo  $f \in \mathcal{F}$ .

## 4.2. Propiedades

Se propondrán condiciones suficientes sobre las familias de conjuntos de funciones que inducen regiones centrales integrales para que estas regiones centrales satisfagan ciertas propiedades ya consideradas por Zuo y Serfling (2000c). En el Teorema 4.2.1 se enuncian algunas propiedades generales de las regiones centrales integrales. En el Teorema 4.2.2 se prueba que todas las regiones centrales integrales son *anidadas*. Los Teoremas 4.2.3 y 4.2.7 establecen condiciones suficientes para obtener regiones integrales *convexas*. Finalmente, en los Teoremas 4.2.8 y 4.2.9 se hace un estudio similar para la obtención de regiones centrales *equivariantes afines*, en los Teoremas 4.2.10 y 4.2.11 para obtener regiones centrales *cerradas* y en el Teorema 4.2.13 para obtener regiones centrales *acotadas*.

### 4.2.1. Generalidades

**Teorema 4.2.1.** *Sea  $P \in \mathbb{P}$  una probabilidad,  $\alpha \in (0, 1]$ ,  $\mathcal{F}, \mathcal{G}$  conjuntos de funciones,  $I, J$  conjuntos de índices y  $\{\mathcal{F}_i\}_{i \in I}, \{\mathcal{G}_i\}_{i \in I}$  familias de conjuntos de funciones.*

- i. Si  $\mathcal{F} \subset \mathcal{G}$ , entonces  $D_{\mathcal{F}}^{\alpha}(P) \supset D_{\mathcal{G}}^{\alpha}(P)$ ;*
- ii. si  $\mathcal{F}_i \subset \mathcal{G}_i$  para todo  $i \in I$ , entonces  $D_{\{\mathcal{F}_i\}_{i \in I}}^{\alpha}(P) \supset D_{\{\mathcal{G}_i\}_{i \in I}}^{\alpha}(P)$ ;*
- iii. si  $\cup_{i \in I} \mathcal{F}_i \subset \mathcal{G}$ , entonces  $D_{\{\mathcal{F}_i\}_{i \in I}}^{\alpha}(P) \supset D_{\mathcal{G}}^{\alpha}(P)$ ;*
- iv. si  $I \subset J$ , entonces  $D_{\{\mathcal{F}_i\}_{i \in I}}^{\alpha}(P) \supset D_{\{\mathcal{F}_j\}_{j \in J}}^{\alpha}(P)$ .*

Junto con estas propiedades, hay que reseñar que la inclusión de funciones constantes en el conjunto de generadores no añade ninguna restricción adicional y no altera las regiones centrales.

### 4.2.2. Regiones anidadas

Las regiones centrales integrales forman una familia anidada.

**Teorema 4.2.2.** Sean  $\alpha, \beta \in (0, 1]$  tales que  $\alpha \leq \beta$ ,  $P \in \mathbb{P}$  y  $\{\mathcal{F}_i\}_{i \in I}$  una familia de conjuntos de funciones, entonces  $D_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P) \supset D_{\{\mathcal{F}_i\}_{i \in I}}^\beta(P)$ .

*Demostración.* Como  $\alpha \leq \beta$ , por el Lema 3.2.1 se cumple  $P^\alpha \supset P^\beta$  y entonces el resultado es inmediato.  $\square$

### 4.2.3. Regiones convexas

Las funciones convexas inducen regiones centrales integrales convexas.

**Teorema 4.2.3.** Si todas las funciones de  $\{\mathcal{F}_i\}_{i \in I}$  son convexas, entonces  $D_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P)$  es convexo para todo  $\alpha \in (0, 1]$  y cualquier  $P \in \mathbb{P}$ .

*Demostración.* Dado un conjunto  $\mathcal{F}_i$  de funciones convexas, una probabilidad  $P \in \mathbb{P}$ ,  $x, y \in D_{\mathcal{F}_i}^\alpha(P)$  y  $\lambda \in (0, 1)$ , veremos que  $\lambda x + (1 - \lambda)y \in D_{\mathcal{F}_i}^\alpha(P)$ .

Como  $x, y \in D_{\mathcal{F}_i}^\alpha(P)$ , existen dos probabilidades  $Q_x, Q_y \in P^\alpha$  tales que  $f(x) \leq \int f dQ_x$  y  $f(y) \leq \int f dQ_y$  para todo  $f \in \mathcal{F}_i$ .

Dada cualquier  $f \in \mathcal{F}_i$ , por su convexidad y por el hecho de que  $x, y \in D_{\mathcal{F}_i}^\alpha(P)$ , tenemos

$$\begin{aligned} f(\lambda x + (1 - \lambda)y) &\leq \lambda f(x) + (1 - \lambda)f(y) \\ &\leq \lambda \int f dQ_x + (1 - \lambda) \int f dQ_y = \int f d(\lambda Q_x + (1 - \lambda)Q_y). \end{aligned}$$

Ahora como por el Lema 3.2.1  $P^\alpha$  es convexo,  $\lambda Q_x + (1 - \lambda)Q_y \in P^\alpha$  y como la desigualdad se verifica para cualquier  $f \in \mathcal{F}_i$ , entonces  $D_{\mathcal{F}_i}^\alpha(P)$  es convexo, basta con considerar  $Q_{\lambda x + (1 - \lambda)y} = \lambda Q_x + (1 - \lambda)Q_y$ . Finalmente  $D_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P)$  es una intersección de conjuntos convexas y por lo tanto convexo.  $\square$

Antes de obtener el siguiente resultado sobre la convexidad de las regiones centrales integrales, damos unos resultados previos en los que relacionamos la nueva noción de *semi-convexidad* establecida en la Definición 4.2.4 con la *cuasi-convexidad* y la *comonotonía*.

**Lema 4.2.4.** *Si  $\mathcal{F}$  es comonótona y todas sus funciones son cuasi-conexas, entonces es semi-conveja.*

*Demostración.* Supongamos que  $\mathcal{F}$  es comonótona y todas sus funciones son cuasi-conexas. Sean  $x, y \in \mathbb{R}^d$  y  $\lambda \in (0, 1)$ , si para todo  $f \in \mathcal{F}$  se cumpliera  $f(x) = f(y)$ , entonces el resultado estaría probado ya que por la cuasi-conexidad  $f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\} = f(x)$ . Si para una función  $g \in \mathcal{F}$ , se tiene  $g(x) < g(y)$  (equivalentemente  $g(y) < g(x)$ ), entonces por la comotonía de  $\mathcal{F}$ , se satisface  $f(x) \leq f(y)$  (equivalentemente  $f(y) \leq f(x)$ ) para todo  $f \in \mathcal{F}$  y finalmente por la cuasi-conexidad de  $f$ , obtenemos  $f(\lambda x + (1 - \lambda)y) \leq \max\{f(x), f(y)\} \leq f(y)$  (equivalentemente  $\leq f(x)$ ).  $\square$

Desafortunadamente el resultado recíproco no se verifica. Aunque la semi-conexidad de un conjunto de funciones claramente implica la cuasi-conexidad de todos sus elementos, no es una condición suficiente para la comotonía. El siguiente ejemplo ilustra este comentario.

*Ejemplo 4.2.1.* Sean  $f$  y  $g$  dos funciones reales de variable real definidas como sigue,

$$f(x) = \begin{cases} 0 & \text{si } x < 0, \\ 1 & \text{si } x = 0, \\ 2 + x & \text{si } x > 0, \end{cases} \quad \text{y} \quad g(x) = \begin{cases} 0 & \text{si } x < -1, \\ -1 - x & \text{si } -1 \leq x \leq 0, \\ 2 + x & \text{si } x > 0. \end{cases}$$

Claramente el conjunto  $\{f, g\}$  es semi-convejo, ya que si  $x, y \leq 1$ , entonces  $f(\lambda x + (1 - \lambda)y) \leq f(\min\{x, y\})$  y del mismo modo  $g(\lambda x + (1 - \lambda)y) \leq g(\min\{x, y\})$ , mientras que si  $x > 1$  ó  $y > 1$ , entonces  $f(\lambda x + (1 - \lambda)y) \leq f(\max\{x, y\})$  y  $g(\lambda x + (1 - \lambda)y) \leq g(\max\{x, y\})$ . Cada una de estas funciones será cuasi-conveja, pero no son comonótonas, obsérvese que por ejemplo  $f(-1/2) = 0 < 1 = f(0)$ , pero  $g(-1/2) = -1/2 > -1 = g(0)$ .

Si la familia de funciones que consideramos contiene exclusivamente funciones continuas, entonces se tiene el recíproco del Lema 4.2.4.

**Lema 4.2.5.** *Dada una familia de funciones continuas  $\mathcal{F}$ , la familia es semi-conveja si y sólo si todas sus funciones son cuasi-conexas y la familia es comonótona.*

*Demostración.* Por el Lema 4.2.4 sólo tenemos que demostrar que dada una familia semi-convexa de funciones continuas, todas las funciones son cuasi-convexas y además la familia es comonótona. La cuasi-convexidad de las funciones es clara, vamos a probar la comonotonía. Sean  $x, y \in \mathbb{R}^d$  y  $f$  una función cualquiera de la familia, supongamos que  $f(x) < f(y)$ . Tenemos que probar que para cualquier función  $g$  de la familia,  $g(x) \leq g(y)$ . Como  $f$  es continua, entonces no se puede cumplir que para cualquier  $\lambda \in (0, 1)$ ,  $f(\lambda x + (1 - \lambda)y) \leq f(x)$ , sin embargo la familia de funciones es semi-convexa, así que tal condición se tiene que dar para un elemento de  $\{x, y\}$  fijo para todas las funciones que tendrá que ser  $y$ , por lo tanto, para cualquier  $\lambda \in (0, 1)$ , se cumple  $g(\lambda x + (1 - \lambda)y) \leq g(y)$  y ahora por la continuidad de  $g$ , finalmente obtenemos  $g(x) \leq g(y)$ .  $\square$

Presentamos otra condición suficiente para la semi-convexidad.

**Lema 4.2.6.** *Si  $\mathcal{F}$  es comonótona y tiene una función estrictamente cuasi-convexa, entonces es semi-convexa.*

*Demostración.* Sean  $x, y \in \mathbb{R}^d$ ,  $\lambda \in (0, 1)$  y  $g \in \mathcal{F}$  una función estrictamente cuasi-convexa. Como  $g(\lambda x + (1 - \lambda)y) < \max\{g(x), g(y)\}$ , suponemos que  $g(x) \leq g(y)$  (si  $g(y) < g(x)$  se realiza un razonamiento análogo), entonces  $g(\lambda x + (1 - \lambda)y) < g(y)$  y por la comonotonía de  $\mathcal{F}$ , se cumple  $f(\lambda x + (1 - \lambda)y) \leq f(y)$  para todo  $f \in \mathcal{F}$ .  $\square$

Finalmente, establecemos un nuevo resultado por el que se obtiene la convexidad de las regiones centrales integrales generadas por familias de conjuntos de funciones semi-convexas.

**Teorema 4.2.7.** *Si para todo  $i \in I$ , el conjunto  $\mathcal{F}_i$  es semi-convexo, entonces  $D_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P)$  es convexo para todo  $\alpha \in (0, 1]$  y cualquier  $P \in \mathbb{P}$ .*

*Demostración.* Demostraremos que todos los  $D_{\mathcal{F}_i}^\alpha(P)$  son convexos y en consecuencia su intersección  $D_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P)$  será convexa.

Dado  $i \in I$ ,  $x, y \in D_{\mathcal{F}_i}^\alpha(P)$  y  $\lambda \in (0, 1)$ , existen  $Q_x, Q_y \in P^\alpha$  tales que  $f(x) \leq \int f dQ_x$  y  $f(y) \leq \int f dQ_y$  para todo  $f \in \mathcal{F}_i$ . Por la semi-convexidad de  $\mathcal{F}$ , bien  $f(\lambda x + (1 - \lambda)y) \leq f(x)$  para todo  $f \in \mathcal{F}_i$ , o bien  $f(\lambda x + (1 - \lambda)y) \leq f(y)$  para todo  $f \in \mathcal{F}_i$ . Si  $f(\lambda x + (1 - \lambda)y) \leq f(x)$  para todo  $f \in \mathcal{F}_i$ , obtenemos el resultado deseado tomando  $Q_{\lambda x + (1 - \lambda)y} = Q_x$  ya que  $f(\lambda x + (1 - \lambda)y) \leq f(x) \leq \int f dQ_x$  para todo  $f \in \mathcal{F}_i$  y por lo tanto  $\lambda x + (1 - \lambda)y \in D_{\mathcal{F}_i}^\alpha(P)$ . Si  $f(\lambda x + (1 - \lambda)y) \leq f(y)$  para todo  $f \in \mathcal{F}_i$ , entonces es suficiente considerar  $Q_{\lambda x + (1 - \lambda)y} = Q_y$ .  $\square$

#### 4.2.4. Regiones equivariantes afines

Los dos siguientes resultados estudian la *equivarianza afín* de las regiones centrales integrales. En primer lugar estudiamos las transformaciones por una matriz  $A$  cuadrada cualquiera y luego por una matriz no singular.

Dada una matriz  $A \in \mathbb{R}^{d \times d}$ , denotamos por  $T_{A,b}$  a la transformación afín  $T_{A,b} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  definida como  $T_{A,b}(x) := Ax + b$ .

**Teorema 4.2.8.** *Para cualesquiera  $A \in \mathbb{R}^{d \times d}$ ,  $b \in \mathbb{R}^d$ ,  $\alpha \in (0, 1]$ ,  $P \in \mathbb{P}$  y  $\mathcal{F}, \mathcal{G}$  conjuntos de funciones, se satisfacen las siguientes afirmaciones:*

- i. *si  $\mathcal{F} \supset \{g \circ T_{A,b} : g \in \mathcal{G}\}$ , entonces  $AD_{\mathcal{F}}^\alpha(P) + b \subset D_{\mathcal{G}}^\alpha(P_{A,b})$ ;*
- ii. *si para todo  $j \in I$  existe  $i^j \in I$  tal que  $\mathcal{F}_{i^j} \supset \{f \circ T_{A,b} : f \in \mathcal{F}_j\}$ , entonces  $AD_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P) + b \subset D_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P_{A,b})$ .*

*Demostración.* Sean  $\mathcal{F}, \mathcal{G}$  dos conjuntos de funciones tales que  $\mathcal{F} \supset \{g \circ T_{A,b} : g \in \mathcal{G}\}$ . Para cualquier  $x \in D_{\mathcal{F}}^\alpha(P)$  habrá alguna probabilidad  $Q \in P^\alpha$  tal que  $f(x) \leq \int f dQ$  para todo  $f \in \mathcal{F}$ . Dado  $g \in \mathcal{G}$ , como  $g \circ T_{A,b} \in \mathcal{F}$ , tenemos que

$$g(Ax + b) \leq \int g \circ T_{A,b} dQ = \int g dQ_{A,b}, \quad (4.1)$$

donde  $Q_{A,b}$  se define a partir de  $Q$  del mismo modo que  $P_{A,b}$  a partir de  $P$ . Como  $Q_{A,b} \in (P_{A,b})^\alpha$  y la relación (4.1) se satisface para cualquier  $g \in \mathcal{G}$ , se cumple  $Ax + b \in D_{\mathcal{G}}^\alpha(P_{A,b})$ .



Para probar la afirmación *ii.*, suponemos que para todo  $j \in I$ , existe  $i^j \in I$  tal que  $\mathcal{F}_{i^j} \supset \{f \circ T_{A,b} : f \in \mathcal{F}_j\}$ . Dado  $j \in I$ , por la primera parte de este resultado,  $AD_{\mathcal{F}_{i^j}}^\alpha(P) + b \subset D_{\mathcal{F}_j}^\alpha(P_{A,b})$  y por lo tanto,  $\bigcap_{j \in I} (AD_{\mathcal{F}_{i^j}}^\alpha(P) + b) \subset \bigcap_{j \in I} D_{\mathcal{F}_j}^\alpha(P_{A,b})$ . Por otra parte,  $T_{A,b}(\bigcap_{j \in I} D_{\mathcal{F}_{i^j}}^\alpha(P)) \subset \bigcap_{j \in I} T_{A,b}(D_{\mathcal{F}_{i^j}}^\alpha(P))$  y finalmente por el apartado *iv.* del Teorema 4.2.1,

$$AD_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P) + b \subset AD_{\{\mathcal{F}_{i^j}\}_{j \in I}}^\alpha(P) + b \subset \bigcap_{j \in I} D_{\mathcal{F}_j}^\alpha(P_{A,b}).$$

□

**Teorema 4.2.9.** *Para cualquier matriz no singular  $A \in \mathbb{R}^{d \times d}$  y  $b \in \mathbb{R}^d$ , se satisfacen las siguientes afirmaciones:*

*i.* si  $\mathcal{F} = \{g \circ T_{A,b} : g \in \mathcal{G}\}$ , entonces  $AD_{\mathcal{F}}^\alpha(P) + b = D_{\mathcal{G}}^\alpha(P_{A,b})$ ;

*ii.* si para todo  $j \in I$ , existe  $i^j \in I$  tal que  $\mathcal{F}_{i^j} = \{f \circ T_{A,b} : f \in \mathcal{F}_j\}$  y para todo  $i \in I$ , existe  $j_i \in I$  tal que  $\mathcal{F}_i = \{f \circ T_{A,b} : f \in \mathcal{F}_{j_i}\}$ , entonces  $AD_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P) + b = D_{\{\mathcal{F}_{i^j}\}_{j \in I}}^\alpha(P_{A,b})$ .

*Demostración.* Si  $\mathcal{F} = \{g \circ T_{A,b} : g \in \mathcal{G}\}$ , entonces  $\mathcal{G} = \{f \circ T_{A,b}^{-1} : f \in \mathcal{F}\}$ , donde  $T_{A,b}^{-1} = T_{A^{-1}, -A^{-1}b}$ . Como  $\mathcal{F} = \{g \circ T_{A,b} : g \in \mathcal{G}\}$ , según la Proposición 4.2.8,  $AD_{\mathcal{F}}^\alpha(P) + b \subset D_{\mathcal{G}}^\alpha(P_{A,b})$  y por la misma razón  $\mathcal{G} = \{f \circ T_{A,b}^{-1} : f \in \mathcal{F}\}$  implica  $A^{-1}D_{\mathcal{G}}^\alpha(P_{A,b}) - A^{-1}b \subset D_{\mathcal{F}}^\alpha(P)$ . Finalmente, tenemos que

$$D_{\mathcal{G}}^\alpha(P_{A,b}) = A(A^{-1}D_{\mathcal{G}}^\alpha(P_{A,b}) - A^{-1}b) + b \subset AD_{\mathcal{F}}^\alpha(P) + b \subset D_{\mathcal{G}}^\alpha(P_{A,b}),$$

lo que implica  $AD_{\mathcal{F}}^\alpha(P) + b = D_{\mathcal{G}}^\alpha(P_{A,b})$ .

Supongamos que se satisfacen las hipótesis del apartado *ii.*, según la primera parte del resultado,  $AD_{\mathcal{F}_i}^\alpha(P) + b = D_{\mathcal{F}_{j_i}}^\alpha(P_{A,b})$  y  $AD_{\mathcal{F}_{j_i}}^\alpha(P) + b = D_{\mathcal{F}_i}^\alpha(P_{A,b})$ . Ahora

$$\begin{aligned} AD_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P) + b &= A\left(\bigcap_{i \in I} D_{\mathcal{F}_i}^\alpha(P)\right) + b = \bigcap_{i \in I} (AD_{\mathcal{F}_i}^\alpha(P) + b) \\ &= \bigcap_{i \in I} D_{\mathcal{F}_{j_i}}^\alpha(P_{A,b}) \supset \bigcap_{j \in I} D_{\mathcal{F}_j}^\alpha(P_{A,b}) \\ &= D_{\{\mathcal{F}_{i^j}\}_{j \in I}}^\alpha(P_{A,b}) = \bigcap_{j \in I} (AD_{\mathcal{F}_{i^j}}^\alpha(P) + b) \supset AD_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P) + b, \end{aligned}$$

con lo que tenemos la igualdad global  $AD_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P) + b = D_{\{\mathcal{F}_{i^j}\}_{j \in I}}^\alpha(P_{A,b})$ . □

### 4.2.5. Regiones cerradas

Las regiones integrales inducidas por familias de conjuntos comonótonos de funciones inferiormente semicontinuas son cerradas.

**Teorema 4.2.10.** *Si todas las funciones de  $\{\mathcal{F}_i\}_{i \in I}$  son inferiormente semicontinuas y para todo  $i \in I$  el conjunto  $\mathcal{F}_i$  es comonótono, entonces  $D_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P)$  es cerrado para todo  $\alpha \in (0, 1]$  y cualquier  $P \in \mathbb{P}$ .*

*Demostración.* Dado  $P \in \mathbb{P}$  y  $\alpha \in (0, 1]$ , si para cualquier  $i \in I$ , el conjunto  $D_{\mathcal{F}_i}^\alpha(P)$  fuese cerrado, entonces su intersección,  $D_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P)$  seguiría siendo cerrada.

Como  $\mathcal{F}_i$  es comonótono, los conjuntos  $\{f(x)\}_{x \in \mathbb{R}^d}$  con  $f \in \mathcal{F}_i$  están idénticamente ordenados, es decir, dadas  $f, g \in \mathcal{F}_i$  y  $x, y \in \mathbb{R}^d$ , si  $f(x) < f(y)$ , entonces por la comonotonía de  $\mathcal{F}_i$ , se cumple  $g(x) \leq g(y)$ . Por lo tanto  $\int f dQ$  es máximo para toda  $f \in \mathcal{F}_i$  en cierta  $Q_i \in P^\alpha$ , aquella que sitúa la mayor probabilidad posible en los puntos de  $\mathbb{R}^d$  en los que las funciones de  $\mathcal{F}_i$  toman mayores valores. La región central de profundidad  $\alpha$  se puede describir ahora como

$$\begin{aligned} D_{\mathcal{F}_i}^\alpha(P) &= \bigcup_{Q \in P^\alpha} \left\{ x \in \mathbb{R}^d : f(x) \leq \int f dQ, \forall f \in \mathcal{F}_i \right\} \\ &= \left\{ x \in \mathbb{R}^d : f(x) \leq \int f dQ_i, \forall f \in \mathcal{F}_i \right\} \\ &= \bigcap_{f \in \mathcal{F}_i} \left\{ x \in \mathbb{R}^d : f(x) \leq \int f dQ_i \right\}, \end{aligned}$$

donde los conjuntos  $\{x \in \mathbb{R}^d : f(x) \leq \int f dQ_i\}$  son cerrados por la semicontinuidad inferior de las funciones de  $\mathcal{F}_i$ , con lo que su intersección  $D_{\mathcal{F}_i}^\alpha(P)$  también sera cerrada.  $\square$

El siguiente resultado impone condiciones más habituales para la obtención de regiones integrales cerradas. Así, las regiones centrales inducidas por familias de conjuntos de funciones continuas y acotadas son también cerradas.

**Teorema 4.2.11.** *Si todas las funciones de  $\{\mathcal{F}_i\}_{i \in I}$  son continuas y acotadas, entonces  $D_{\{\mathcal{F}_i\}_{i \in I}}^\alpha(P)$  es cerrado para todo  $\alpha \in (0, 1]$  y cualquier  $P \in \mathbb{P}$ .*

*Demostración.* Dado  $P \in \mathbb{P}$  y  $\alpha \in (0, 1]$ , es suficiente probar que para cualquier  $i \in I$ , el conjunto  $D_{\mathcal{F}_i}^\alpha(P)$  es cerrado. Sea  $\{x_n\}_n$  una sucesión de elementos de  $D_{\mathcal{F}_i}^\alpha(P)$  que converge a cierto  $x \in \mathbb{R}^d$ . Para todo  $n \in \mathbb{N}$  se tiene que  $x_n \in D_{\mathcal{F}_i}^\alpha(P)$  y por lo tanto existe  $Q_n \in P^\alpha$  que satisface

$$f(x_n) \leq \int f dQ_n \text{ para todo } f \in \mathcal{F}_i. \quad (4.2)$$

Según el Lema 3.2.3, la clase  $P^\alpha$  es tight, y como  $\{Q_n\}_n \subset P^\alpha$ , según el Teorema 0.1.5, existe una subsucesión  $\{Q_{n_k}\}_k$  débilmente convergente, esto es  $Q_{n_k} \Rightarrow_k Q$  para cierto  $Q \in \mathbb{P}$  y como según el Lema 3.2.4, la clase  $P^\alpha$  es compacta, se satisface que  $Q \in P^\alpha$ .

Finalmente, por la continuidad de las funciones  $f \in \mathcal{F}_i$  y como  $\lim_n x_n = x$ , se satisface que  $\lim_n f(x_n) = f(x)$ . Además por la convergencia débil  $Q_{n_k} \Rightarrow_k Q$  y por la continuidad y acotación de las funciones  $f \in \mathcal{F}_i$ , tenemos que  $\lim_k \int f dQ_{n_k} = \int f dQ$ , con lo que de la ecuación (4.2) se deduce que  $f(x) \leq \int f dQ$  para todo  $f \in \mathcal{F}_i$  y como  $Q \in P^\alpha$  se tiene que  $x \in D_{\mathcal{F}_i}^\alpha(P)$  y por lo tanto este conjunto es cerrado.  $\square$

#### 4.2.6. Regiones acotadas

La última propiedad general que consideraremos de las regiones integrales antes de pasar a estudiar su comportamiento asintótico es su acotación. En primer lugar se obtendrá una condición suficiente para la acotación de las regiones integrales de variables aleatorias y luego extenderemos este resultado a dimensiones mayores.

**Lema 4.2.12.** *Sea  $d = 1$  y  $\mathcal{F}$  un conjunto de funciones crecientes tales que  $\int f dP < +\infty$ . Si para todo  $a \in \mathbb{R}$  existe  $f \in \mathcal{F}$  tal que  $f$  es estrictamente creciente en algún  $b > a$ , entonces  $\sup D_{\mathcal{F}}^\alpha(P) < +\infty$  para todo  $\alpha \in (0, 1]$ .*

*Demostración.* Sea  $\alpha \in (0, 1]$ , como toda función  $f \in \mathcal{F}$  es creciente,  $\int f dQ$  toma su valor máximo en algún  $Q \in P^\alpha$  independiente de la elección de  $f \in \mathcal{F}$ . Además  $\int f dP < +\infty$  y por lo tanto,  $\int f dQ < +\infty$ . Si hay alguna función  $f \in \mathcal{F}$  no acotada, entonces  $f(y) > \int f dQ$  para algún  $y \in \mathbb{R}$  y el resultado estaría demostrado. Si toda  $f \in \mathcal{F}$  es acotada, tomamos  $a \in \mathbb{R}$  tal que  $Q((-\infty, a)) > 0$  y  $g \in \mathcal{F}$  estrictamente creciente en algún  $b > a$ . Como  $g$  es creciente y acotada,  $\lim_{x \rightarrow \infty} g(x) = k < +\infty$ , y  $\int g dQ \leq g(a)Q((-\infty, a)) + kQ([a, +\infty)) < k$ . Finalmente, hay algún  $y \in \mathbb{R}$  tal que  $g(y)$  está lo suficientemente próximo a  $k$  y  $g(y) > \int f dQ$  de donde se deduce que  $\sup D_{\mathcal{F}}^\alpha(P) \leq y$ .  $\square$

**Teorema 4.2.13.** Sean  $\{u_i\}_{i \in I} \subset \mathbb{R}^d$  ortogonales a las caras de un poliedro convexo con  $\langle u_i, x - y_i \rangle < 0$  para cualquier  $x$  en el interior del poliedro e  $y_i$  en la cara ortogonal a  $u_i$ . Sea  $\mathcal{F}$  un conjunto de funciones de variable real crecientes y tales que para cualquier  $a \in \mathbb{R}$  tenemos  $f \in \mathcal{F}$  estrictamente creciente en algún punto  $b > a$ .

Sea  $J$  un nuevo conjunto de índices tal que  $I \subset J$ , consideramos una familia de conjuntos de funciones  $\{\mathcal{F}_j\}_{j \in J}$  donde para cada  $i \in I$ , se cumple  $\mathcal{F}_i \supset \{f \circ l_i : l_i(x) = \langle x, u_i \rangle, f \in \mathcal{F}\}$ .

En estas condiciones, si  $\int f \circ l_i dP < +\infty$  para todo  $f \in \mathcal{F}$ , entonces  $D_{\{\mathcal{F}_j\}_{j \in J}}^\alpha(P)$  es acotado para todo  $\alpha \in (0, 1]$ .

*Demostración.* Este resultado es inmediato a partir de la Proposición 4.2.1 y el Lema 4.2.12.  $\square$

### 4.3. Resultados asintóticos

A lo largo de la presente sección se estudia el comportamiento asintótico de las regiones centrales integrales. En primer lugar se recurre a un ejemplo para mostrar que, en general, no se tiene la convergencia directa de las regiones centrales integrales de una distribución de probabilidad empírica a las regiones integrales de su correspondiente distribución poblacional. Para su-

perar este inconveniente, utilizaremos dos herramientas nuevas, las familias de funciones de *Glivenko-Cantelli* y una nueva región central integral.

Los resultado más importantes que se obtienen son los Teoremas 4.3.7 y 4.3.8 en los que se describen las regiones centrales integrales de una distribución poblacional a partir de regiones integrales empíricas, en primer lugar para regiones integrales inducidas por un conjunto de funciones y luego para regiones inducidas por una familia de conjuntos de funciones.

Como el conjunto  $D_{\mathcal{F}}^{\alpha}(P)$  se define a partir de ciertas integrales de funciones de  $\mathcal{F}$  y de la evaluación de esas funciones en un punto, parece natural que haya que exigir la continuidad de los elementos de  $\mathcal{F}$  para obtener un buen comportamiento asintótico.

En el Ejemplo 4.3.1 se verá que ni la continuidad uniforme ni la acotación uniforme de los elementos de  $\mathcal{F}$  unidas a la numerabilidad de  $\mathcal{F}$  son condiciones suficientes para garantizar que el límite de una subsucesión de  $\{D_{\mathcal{F}}^{\alpha}(P_{n,\cdot})\}_n$  sea igual a  $D_{\mathcal{F}}^{\alpha}(P)$ . Es por tanto necesario desarrollar nuevas herramientas para caracterizar  $D_{\mathcal{F}}^{\alpha}(P)$  a partir de regiones aleatorias basadas en distribuciones empíricas.

*Ejemplo 4.3.1.* Sea  $d = 1$ ,  $\alpha = 1$  y  $P$  la probabilidad inducida por una distribución uniforme en  $(-1, 1)$ . Construimos la familia  $\mathcal{F}$  con funciones de la forma  $f_{\lambda}$ , con  $\lambda \in \Lambda = ([-1, 1] \cap \mathbb{Q}) \setminus \{0\}$  definidas como

$$f_{\lambda}(x) = \begin{cases} -1 & \text{si } x < -\lambda, \\ \lambda^{-1}x & \text{si } -\lambda \leq x \leq \lambda, \\ 1 & \text{si } x > \lambda, \end{cases} \quad \text{si } \lambda > 0, \text{ y } f_{\lambda} = -f_{-\lambda} \text{ si } \lambda < 0.$$

Tenemos que  $\int f_{\lambda} dP = 0$  y  $f_{\lambda}(0) = 0$  para todo  $\lambda \in \Lambda$ . Como para cualquier  $x \neq 0$  existe  $\lambda \in \Lambda$  tal que  $f_{\lambda}(x) > 0$ , entonces  $D_{\mathcal{F}}^1(P) = \{0\}$ . Por otra parte, es sencillo probar que  $D_{\{f_1, f_{-1}\}}^1(P_{n,\cdot}) = \{(1/n) \sum_{i=1}^n \xi_i(\cdot)\}$  c.s. [Pr], es decir, es un conjunto con un único elemento. Por lo tanto sólo puede ocurrir una de las dos siguientes circunstancias, bien  $D_{\mathcal{F}}^1(P_{n,\cdot}) = \{(1/n) \sum_{i=1}^n \xi_i(\cdot)\}$  c.s. [Pr], o bien  $D_{\mathcal{F}}^1(P_{n,\cdot}) = \emptyset$  c.s. [Pr].

Si  $(1/n) \sum_{i=1}^n \xi_i(\cdot)$  pertenece a  $D_{\mathcal{F}}^1(P_{n,\cdot})$  para algún  $n \in \mathbb{N}$ , entonces para todo  $\lambda \in \Lambda$

$$f_{\lambda} \left( \frac{1}{n} \sum_{i=1}^n \xi_i(\cdot) \right) \leq \frac{1}{n} \sum_{i=1}^n f_{\lambda}(\xi_i(\cdot)) \quad \text{y} \quad f_{-\lambda} \left( \frac{1}{n} \sum_{i=1}^n \xi_i(\cdot) \right) \leq \frac{1}{n} \sum_{i=1}^n f_{-\lambda}(\xi_i(\cdot)),$$

lo que implica que  $f_{\lambda}((1/n) \sum_{i=1}^n \xi_i(\cdot)) = (1/n) \sum_{i=1}^n f_{\lambda}(\xi_i(\cdot))$  para todo  $\lambda \in \Lambda$ . Como  $(1/n) \sum_{i=1}^n \xi_i(\cdot) \neq 0$  c.s. [Pr], para algún  $\lambda_0$  suficientemente próximo a 0 se cumple que  $f_{\lambda_0}((1/n) \sum_{i=1}^n \xi_i(\cdot)) = 1$ . Esto implica que  $f_{\lambda_0}(\xi_i(\cdot)) = 1$  para todo  $i \in \{1, \dots, n\}$  y por lo tanto, bien todo  $\xi_i(\cdot)$  es positivo o bien todo  $\xi_i(\cdot)$  es negativo, entonces

$$\Pr \left( \frac{1}{n} \sum_{i=1}^n \xi_i(\cdot) \in D_{\mathcal{F}}^1(P_{n,\cdot}) \right) \leq \left( \frac{1}{2} \right)^n.$$

Ahora, según el *Lema de Borel-Cantelli*, tenemos que

$$\Pr \left( \frac{1}{n} \sum_{i=1}^n \xi_i(\cdot) \in D_{\mathcal{F}}^1(P_{n,\cdot}) \text{ i.o.} \right) = 0$$

donde recordamos que *i.o.* significa que ocurre para un número infinito de  $n$ 's. Finalmente, se tiene que  $D_{\mathcal{F}}^1(P_{n,\cdot}) = \emptyset$  c.s. [Pr] excepto para a lo sumo un número finito de ellas.

Obsérvese que en este caso la familia  $\mathcal{F}$  satisface propiedades “excelentes” como son el ser numerable y con funciones uniformemente continuas y uniformemente acotadas.

A continuación resumimos algunas nociones básicas sobre *clases de Glivenko-Cantelli* extraídas de Talagrand (1987) y (1996).

Dado un espacio de probabilidad completo  $(\mathbb{R}^d, \Sigma, P)$ , consideramos la probabilidad producto  $P^n$  sobre  $(\mathbb{R}^d)^n$ , la probabilidad producto  $P^\infty$  sobre  $(\mathbb{R}^d)^\infty$  y la probabilidad exterior de  $P^n$  a la que denotaremos como  $P^{n*}$ .

Si  $s = \{s_i\}_{i=1}^\infty \in (\mathbb{R}^d)^\infty$ , dada  $\mathcal{Z}$  una familia de funciones de  $\mathbb{R}^d$  en  $\mathbb{R}$  tal que  $\mathcal{Z} \subset L^1(P)$ , definimos

$$d_n(s) := \sup_{f \in \mathcal{Z}} \left| \frac{1}{n} \sum_{i=1}^n f(s_i) - \int f dP \right|. \quad (4.3)$$

*Definición 4.3.1.* Una familia de funciones  $\mathcal{Z}$  se dice que es una *clase de Glivenko-Cantelli (para  $P$ )* si  $d_n(s) \rightarrow 0$  c.s. [Pr].

Asociados al concepto de familia de Glivenko-Cantelli están los conceptos de familia *estable* y de *orden acotado*.

*Definición 4.3.2.* Una familia de funciones  $\mathcal{Z}$  es *estable* si para cada  $\alpha < \beta$ , y cada conjunto  $A \in \Sigma$  con  $P(A) > 0$ , existe  $n > 0$  tal que

$$P^{2n*}(\{(s_1, \dots, s_n, t_1, \dots, t_n) \in A^{2n} : \text{existe } f \in \mathcal{Z} \text{ con } f(s_i) < \alpha, f(t_i) > \beta, \forall i \leq n\}) < P(A)^{2n}.$$

*Definición 4.3.3.* Una familia de funciones  $\mathcal{Z}$  es *de orden acotado* si existe una aplicación  $h \in L^1(P)$  tal que  $|f| \leq h$  para todo  $f \in \mathcal{Z}$ .

El resultado principal que utilizaremos para estudiar el comportamiento asintótico de las regiones centrales integrales es el Teorema 2 en Talagrand (1987) que está enunciado a continuación.

**Teorema 4.3.1 (Teorema 2 en Talagrand, 1987).** *Para un subconjunto  $\mathcal{Z}$  de  $L^1(P)$  las siguientes condiciones son equivalentes:*

- i.  $\mathcal{Z}$  es una clase de Glivenko-Cantelli y  $\{\int f dP : f \in \mathcal{Z}\}$  es acotado;*
- ii.  $\mathcal{Z}$  es estable y de orden acotado.*

De ahora en adelante las funciones de  $\mathcal{F}$  se supondrán medibles con respecto a  $\mathcal{B}_d$  y denotaremos por  $(\mathbb{R}^d, \overline{\mathcal{B}_d}, \overline{P})$  a la completación del espacio de probabilidad.

**Proposición 4.3.2.** *Dado un espacio de probabilidad  $(\mathbb{R}^d, \mathcal{B}_d, P)$ , si  $\mathcal{F}$  es una clase de Glivenko-Cantelli para  $\overline{P}$ , entonces*

$$D_n(\omega) := \sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^n f(\xi_i(\omega)) - \int f dP \right| \rightarrow 0 \quad \text{c.s. [Pr]}$$

*Demostración.* En primer lugar suponemos que  $P$  es completa. Sea  $A = \{\omega \in \Omega : D_n(\omega) \not\rightarrow 0\}$  y sea  $B = \{(\xi_i(\omega))_{i=1}^\infty : \omega \in A\}$ , claramente para todo  $s \in B$  tenemos que  $d_n(s) \not\rightarrow 0$ , donde  $d_n$  se definió en (4.3). Como  $\mathcal{F}$  es una clase de Glivenko-Cantelli, entonces existe un conjunto  $C$  medible tal que  $B \subset C$  y  $P^\infty(C) = 0$ , por lo tanto  $A \subset ((\xi_i)_{i=1}^\infty)^{-1}(C)$  con  $\Pr(((\xi_i)_{i=1}^\infty)^{-1}(C)) = 0$ , lo que finaliza la demostración cuando  $P$  es completa.

Si  $P$  no es completa, tomamos su compleción  $\bar{P}$ , obsérvese que  $\int f dP = \int f d\bar{P}$  para todo  $f \in \mathcal{F}$  porque  $f$  es medible respecto de  $\mathcal{B}_d$ .

Sea  $\eta : \Omega \rightarrow \mathbb{R}^d$  medible respecto de  $\mathcal{A}$  y  $\bar{\mathcal{B}}_d$  tal que la probabilidad inducida por  $\eta$  sobre  $(\mathbb{R}^d, \bar{\mathcal{B}}_d)$  es  $\bar{P}$ . Sean  $\eta_1, \eta_2, \dots, \eta_n$   $n$  copias independientes de  $\eta$ . Definimos

$$\bar{D}_n(\omega) := \sup_{f \in \mathcal{F}} \left| \frac{1}{n} \sum_{i=1}^n f(\eta_i(\omega)) - \int f dP \right|,$$

y razonando de un modo parecido a como lo hemos hecho anteriormente, podemos demostrar que  $\bar{D}_n \rightarrow 0$  c.s. [Pr]. Como  $\xi_1$  y  $\eta_1$  siguen la misma distribución como aplicaciones medibles respecto de  $\mathcal{B}_d$ , obtenemos el resultado deseado.  $\square$

**Proposición 4.3.3.** *Dado un espacio de probabilidad  $(\mathbb{R}^d, \mathcal{B}_d, P)$ , si  $\mathcal{F}$  es una clase de Glivenko-Cantelli para  $\bar{P}$  formada por funciones uniformemente acotadas y  $g$  es la derivada de Radon-Nikodym de  $Q \in P^\alpha$  con respecto a  $P$ , entonces  $\{gf : f \in \mathcal{F}\}$  es una clase de Glivenko-Cantelli para  $\bar{P}$ .*

*Demostración.* Por la acotación uniforme de  $\mathcal{F}$  y como  $g \leq \alpha^{-1}$ , se tiene que  $\{gf : f \in \mathcal{F}\}$  es de orden acotado.

Por otra parte, como  $g \in L^1(\bar{P})$ , entonces claramente  $\{g\}$  es una clase de Glivenko-Cantelli para  $\bar{P}$  y según el Teorema 4.3.1,  $\{g\}$  es estable. Según la Proposición 24 en Talagrand (1987),  $\{gf : f \in \mathcal{F}\}$  es estable, lo que finaliza la demostración sin más que considerar de nuevo el Teorema 4.3.1.  $\square$



**Corolario 4.3.4.** Dado un espacio de probabilidad  $(\mathbb{R}^d, \mathcal{B}_d, P)$ , si  $\mathcal{F}$  es una clase de Glivenko-Cantelli para  $\bar{P}$  formada por funciones uniformemente acotadas y  $g$  es la derivada de Radon-Nikodym de  $Q \in P^\alpha$  con respecto a  $P$ , entonces

$$\sup_{f \in \mathcal{F}} \left| \int f g dP_{n,\cdot} - \int f g dP \right| \longrightarrow 0 \quad \text{c.s. [Pr]}$$

Para obtener las regiones integrales poblacionales como límite de regiones empíricas, introducimos el concepto de *regiones centrales  $\varepsilon$ -integrales*

*Definición 4.3.4.* Dado  $\alpha \in (0, 1]$ ,  $\varepsilon > 0$  y  $\mathcal{F}$  un conjunto de funciones, definimos la *región central  $\varepsilon$ -integral de nivel  $\alpha$  para  $P$  respecto del conjunto  $\mathcal{F}$*  y la denotamos por  $D_{\mathcal{F}}^{\alpha, \varepsilon}(P)$ , como

$$D_{\mathcal{F}}^{\alpha, \varepsilon}(P) := \left\{ x \in \mathbb{R}^d : \text{existe } Q_x \in P^\alpha, f(x) - \varepsilon \leq \int f dQ_x, \forall f \in \mathcal{F} \right\}.$$

Obsérvese que  $D_{\mathcal{F}}^\alpha(P) \subset D_{\mathcal{F}}^{\alpha, \varepsilon}(P)$  para todo  $\varepsilon > 0$ .

Ahora estamos en condiciones de demostrar el siguiente resultado.

**Proposición 4.3.5.** Dado un espacio de probabilidad  $(\mathbb{R}^d, \mathcal{B}_d, P)$ , si  $\mathcal{F}$  es una clase de Glivenko-Cantelli para  $\bar{P}$  formada por funciones uniformemente acotadas, entonces para todo  $\alpha \in (0, 1]$  tenemos

$$D_{\mathcal{F}}^\alpha(P) \subset \bigcap_{\varepsilon > 0} \limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n,\cdot}) \quad \text{c.s. [Pr]}.$$

*Demostración.* Sea  $D$  un conjunto numerable y denso en  $D_{\mathcal{F}}^\alpha(P)$  cuya existencia está garantizada debido a la segundo numerabilidad de  $\mathbb{R}^d$ . Dado  $x \in D$ , entonces existe  $Q \in P^\alpha$  tal que  $f(x) \leq \int f dQ$  para todo  $f \in \mathcal{F}$ . Según el Teorema 3.2.7 existe una subsucesión formada por elementos  $Q_{n_k, \cdot} \in P_{n_k, \cdot}^\alpha$  para todo  $k \in \mathbb{N}$ , tal que  $Q_{n_k, \cdot} \Rightarrow Q$  c.s. [Pr].

Por otra parte

$$\sup_{f \in \mathcal{F}} \left| \int f dQ_{n_k, \cdot} - \int f dQ \right| = \sup_{f \in \mathcal{F}} \left| \int \frac{fg}{\|g\|_{L^1(P_{n_k, \cdot})}} dP_{n_k, \cdot} - \int fg dP \right|,$$

siendo  $g$  la derivada de Radon-Nikodym de  $Q$  con respecto a  $P$ .

Según el Corolario 4.3.4, dado  $\varepsilon > 0$  existe  $k_0$  tal que para todo  $k \geq k_0$  (obsérvese que  $\|g\|_{L^1(P_{n_k, \cdot})} \rightarrow 1$  c.s. [Pr])

$$\sup_{f \in \mathcal{F}} \left| \int f dQ_{n_k, \cdot} - \int f dQ \right| < \frac{\varepsilon}{2} \quad \text{c.s. [Pr]},$$

y entonces para todo  $f \in \mathcal{F}$

$$f(x) - \varepsilon \leq \int f dQ_{n_k, \cdot} \quad \text{c.s. [Pr]}.$$

Tenemos, por tanto, que  $x \in D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n_k, \cdot})$  para todo  $k \geq k_0$  y por lo tanto c.s. [Pr], se cumple  $x \in \limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \cdot})$  para todo  $\varepsilon > 0$ , equivalentemente  $x \in \bigcap_{\varepsilon > 0} \limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \cdot})$  c.s. [Pr] y como  $D$  es numerable,

$$D \subset \bigcap_{\varepsilon > 0} \limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \cdot}) \quad \text{c.s. [Pr]}.$$

Finalmente, como  $\limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \cdot})$  es cerrado, entonces la intersección  $\bigcap_{\varepsilon > 0} \limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \cdot})$  también será cerrada y tenemos que

$$D_{\mathcal{F}}^{\alpha}(P) \subset \overline{D} \subset \bigcap_{\varepsilon > 0} \limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \cdot}) \quad \text{c.s. [Pr]}.$$

□

La siguiente proposición establece condiciones bajo las cuales se obtiene el contenido inverso al de la Proposición 4.3.5.

**Proposición 4.3.6.** *Si  $\mathcal{F}$  es un conjunto de funciones continuas y acotadas, entonces para todo  $\alpha \in (0, 1]$*

$$\bigcap_{\varepsilon > 0} \limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \cdot}) \subset D_{\mathcal{F}}^{\alpha}(P) \quad \text{c.s. [Pr]}.$$

*Demostración.* Se cumple que  $P_{n, \cdot} \Rightarrow P$  c.s. [Pr], sea  $\Omega_0 \subset \Omega$  un conjunto de probabilidad 1 en el que se tiene la convergencia, es decir,  $\Pr(\Omega_0) = 1$  y para todo  $\omega \in \Omega_0$  se verifica  $P_{n, \omega} \Rightarrow P$ .

Tomamos  $\omega \in \Omega_0$  fijo. Dado  $x_{\omega} \in \bigcap_{\varepsilon > 0} \limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \omega})$ , para  $\varepsilon \in \mathbb{Q} \cap (0, +\infty)$  existe una sucesión, que puede depender de  $\varepsilon$  tanto en el punto

que se toma dentro de cada  $D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \omega})$  como en los índices que se tomen,  $(x_{\omega}^{\varepsilon})_{n_{k(\varepsilon)}} \in D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n_{k(\varepsilon)}, \omega})$  tal que  $x_{\omega} = \lim_{k(\varepsilon)} (x_{\omega}^{\varepsilon})_{n_{k(\varepsilon)}}$ , por lo tanto

$$f((x_{\omega}^{\varepsilon})_{n_{k(\varepsilon)}}) - \varepsilon \leq \int f dQ_{n_{k(\varepsilon)}, \omega}^{(\varepsilon)}$$

para todo  $f \in \mathcal{F}$  y algún  $Q_{n_{k(\varepsilon)}, \omega}^{(\varepsilon)} \in P_{n_{k(\varepsilon)}, \omega}^{\alpha}$

Tomamos la sucesión  $\{\varepsilon_m\}_m$  con  $\varepsilon_m = 1/m$ , puesto que  $\{Q_{n_{k(\frac{1}{m})}, \omega}^{(\frac{1}{m})}\}_k \subset P_{n_{k(\frac{1}{m})}, \omega}^{\alpha}$ , según el Teorema 3.2.7 existe una subsucesión (a la que denotaremos del mismo modo) tal que

$$Q_{n_{k(\frac{1}{m})}, \omega}^{(\frac{1}{m})} \xrightarrow{k(\frac{1}{m})} Q_{m, \omega}$$

para algún  $Q_{m, \omega} \in P^{\alpha}$ .

Por la continuidad y acotación de  $f$

$$f(x) - \frac{1}{m} \leq \int f dQ_{m, \omega}, \text{ para todo } m \in \mathbb{N} \text{ y } f \in \mathcal{F}.$$

Como  $P^{\alpha}$  es tight, habrá una subsucesión de  $\{Q_{m, \omega}\}_m \subset P^{\alpha}$  con cierto límite  $Q_{\omega}$  y por el Lema 3.2.4,  $Q_{\omega} \in P^{\alpha}$ . De nuevo por la continuidad y la acotación de las funciones de  $\mathcal{F}$ ,

$$f(x) \leq \int f dQ_{\omega}, \text{ para todo } f \in \mathcal{F},$$

y por lo tanto  $x_{\omega} \in D_{\mathcal{F}}^{\alpha}(P)$ . Así  $\bigcap_{\varepsilon > 0} \limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \omega}) \subset D_{\mathcal{F}}^{\alpha}(P)$  y como esto se tiene para todo  $\omega \in \Omega_0$  que es un conjunto de probabilidad 1, se cumple que  $\bigcap_{\varepsilon > 0} \limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \cdot}) \subset D_{\mathcal{F}}^{\alpha}(P)$  c.s. [Pr].  $\square$

Finalmente podemos caracterizar el conjunto  $D_{\mathcal{F}}^{\alpha}(P)$  a partir de ciertos conjuntos empíricos.

**Teorema 4.3.7.** *Si  $\mathcal{F}$  es una clase de Glivenko-Cantelli para  $\bar{P}$  formada por funciones continuas y uniformemente acotadas, entonces para todo  $\alpha \in (0, 1]$  tenemos que casi seguro [Pr]*

$$\bigcap_{\varepsilon > 0} \limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \cdot}) = D_{\mathcal{F}}^{\alpha}(P).$$

La convergencia asintótica puede extenderse al caso de regiones centrales integrales generadas por una familia de conjuntos de funciones.

**Teorema 4.3.8.** *Si para todo  $i \in I \subset \mathbb{N}$ ,  $\mathcal{F}_i$  es una clase de Glivenko-Cantelli para  $\bar{P}$  formada por funciones continuas y uniformemente acotadas, entonces para todo  $\alpha \in (0, 1]$  tenemos que casi seguro [Pr]*

$$\bigcap_{i \in I} \bigcap_{\varepsilon > 0} \limsup_n D_{\mathcal{F}_i}^{\alpha, \varepsilon}(P_n, \cdot) = D_{\{\mathcal{F}_i\}_{i \in I}}^{\alpha}(P).$$

Para finalizar esta sección, volvemos al Ejemplo 4.3.1 y calculamos  $D_{\mathcal{F}}^1(P)$  utilizando el Teorema 4.3.7 y las regiones centrales  $\varepsilon$ -integrales empíricas.

*Ejemplo 4.3.2.* Sea  $P$  la distribución de probabilidad uniforme en  $(-1, 1)$  y la familia de funciones  $\mathcal{F} = \{f_{\lambda} : \lambda \in \Lambda\}$  como la del Ejemplo 4.3.1. Obsérvese que la compleción de  $P$  es  $\bar{P} = 2^{-1}\mu$  donde  $\mu$  es la medida de Lebesgue en  $\mathbb{R}$ .

Según la Proposición 2.1 en Talagrand (1996), un conjunto numerable de funciones  $\mathcal{F}$  es una clase de Glivenko-Cantelli para  $\bar{P}$  si para todo  $\varepsilon > 0$ ,  $\mathcal{F}$  puede recubrirse con un número finito de corchetes de tamaño menor que  $\varepsilon$ . Un corchete queda definido por los aplicaciones  $f \leq g$  según  $[f, g] := \{h : f \leq h \leq g\}$  y su tamaño es  $\int (g - f) d\bar{P}$ .

Dado  $\varepsilon > 0$ , sea  $0 < a \leq \varepsilon/4$ ,  $m_1 = a^{-1}$  y  $m_{i+1} = m_i/(1 + 2m_i a)$  para  $i \in \{1, \dots, k\}$  donde  $k$  es tal que  $m_k \leq (1 + a)^{-1}$ . Podemos considerar las funciones

$$\underline{g}_i(x) = \begin{cases} -1 & \text{si } x < -m_i^{-1} + a, \\ m_i(x - a) & \text{si } -m_i^{-1} + a \leq x \leq m_i^{-1} + a, \\ 1 & \text{si } x > m_i^{-1} + a, \end{cases} \quad y$$

$$\bar{g}_i(x) = \begin{cases} -1 & \text{si } x < -m_i^{-1} - a, \\ m_i(x - a) & \text{si } -m_i^{-1} - a \leq x \leq m_i^{-1} - a, \\ 1 & \text{si } x > m_i^{-1} - a, \end{cases}$$

conjuntamente con  $\bar{g}_{-i} = -\underline{g}_i$  y  $\underline{g}_{-i} = -\bar{g}_i$ . Como los corchetes  $[\underline{g}_i, \bar{g}_i]$  con  $i \in \{-k, \dots, k\} \setminus \{0\}$  recubren la familia  $\mathcal{F}$  y satisfacen  $\int (\bar{g}_i - \underline{g}_i) d\bar{P} \leq \varepsilon$ , se cumple que  $\mathcal{F}$  es una clase de Glivenko-Cantelli para  $\bar{P}$ .

Según el Teorema 4.3.7  $D_{\mathcal{F}}^{\alpha}(P) = \bigcap_{\varepsilon > 0} \limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \cdot})$  c.s. [Pr].

Sea ahora  $\varepsilon > 0$ , los elementos de  $D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \cdot})$  son los números reales que satisfacen  $f_{\lambda}(x) - \varepsilon \leq (1/n) \sum_{i=1}^n f_{\lambda}(\xi_i(\cdot))$  para todo  $\lambda \in \Lambda$ . Como  $f_{\lambda} = -f_{-\lambda}$ , la condición anterior es equivalente a  $f_{\lambda}(x) - \varepsilon \leq (1/n) \sum_{i=1}^n f_{\lambda}(\xi_i(\cdot)) \leq f_{\lambda}(x) + \varepsilon$ .

Como  $\mathcal{F}$  es una clase de Glivenko-Cantelli para  $\bar{P}$ , entonces  $\sup_{\lambda \in \Lambda} |(1/n) \sum_{i=1}^n f_{\lambda}(\xi_i(\cdot))|$  converge a 0 y por lo tanto  $\limsup_n D_{\mathcal{F}}^{\alpha, \varepsilon}(P_{n, \cdot}) = \{0\}$  c.s. [Pr] y finalmente  $D_{\mathcal{F}}^1(P) = \{0\}$ .

## 4.4. Ejemplos

Según la familia de funciones que consideremos, sus regiones centrales integrales nos darán diferente información sobre una distribución de probabilidad.

Los dos primeros ejemplos de regiones centrales integrales que proponemos están generados por conjuntos de funciones muy simples. A continuación estudiamos ciertos ejemplos de regiones centrales clásicas dentro del marco de las regiones centrales integrales.

*Ejemplo 4.4.1.* Sea  $\mathcal{F}$  el conjuntos de todas las funciones constantes, para cualquier  $P \in \mathbb{P}$  y  $\alpha \in (0, 1]$ ,  $D_{\mathcal{F}}^{\alpha}(P) = \mathbb{R}^d$ .

*Ejemplo 4.4.2.* Sea  $\xi$  un vector aleatorio en  $\mathbb{R}^d$ , si  $\mathcal{F}$  tiene la función indicador del complementario del soporte de  $\xi$  como único elemento,  $\mathcal{F} = \{I_{\mathbb{R}^d \setminus \text{supp } \xi}\}$ , entonces para cualquier  $\alpha \in (0, 1]$ , se cumple  $D_{\mathcal{F}}^{\alpha}(P_{\xi}) = \text{supp } \xi$ .

### 4.4.1. Regiones centrales semiespaciales

Las regiones centrales semiespaciales dadas en la Sección 0.4.1 pertenecen a familia de regiones centrales integrales.

**Teorema 4.4.1.** *Las familias de conjuntos de funciones  $\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}$  y  $\{\mathcal{G}_u\}_{u \in \mathbb{R}^d}$  con  $\mathcal{F}_u = \{f \circ l_u : f : \mathbb{R} \rightarrow \mathbb{R} \text{ no decreciente}, l_u(x) = \langle x, u \rangle\}$  y  $\mathcal{G}_u = \{I_{(a, +\infty)} \circ l_u : l_u(x) = \langle x, u \rangle, a \in \mathbb{R}\}$  generan las regiones centrales*

semiespaciales como regiones centrales integrales, es decir,  $\text{HD}^\alpha(P) = D_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P) = D_{\{\mathcal{G}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$  para cualquier  $\alpha \in (0, 1]$ ,  $P \in \mathbb{P}$ .

*Demostración.* Como  $\mathcal{F}_u \supset \mathcal{G}_u$  para todo  $u \in \mathbb{R}^d$ , según el Teorema 4.2.1,  $D_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P) \subset D_{\{\mathcal{G}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$  para todo  $\alpha \in (0, 1]$ . Tomamos ahora  $\alpha \in (0, 1]$  fijo.

En primer lugar demostraremos que  $\text{HD}^\alpha(P) \subset D_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha$ . Sea  $z \in \text{HD}^\alpha(P)$  y  $v \in \mathbb{R}^d \setminus \{0\}$ , definimos el semiespacio  $H$  como  $H = \{x \in \mathbb{R}^d : \langle x, v \rangle \geq \langle z, v \rangle\}$ , como  $z$  está en la frontera de  $H$ ,  $P(H) \geq \alpha$ . Sea ahora  $Q = P|_H$ , claramente  $Q \in P^\alpha$  y  $f(\langle z, v \rangle) \leq \int f(\langle x, v \rangle) dQ(x)$  para cualquier  $f$  no decreciente, por lo tanto  $z \in D_{\mathcal{F}_v}^\alpha(P)$ . Como el mismo argumento puede aplicarse a cualquier  $u \in \mathbb{R}^d \setminus \{0\}$ ,  $z \in \bigcap_{u \in \mathbb{R}^d} D_{\mathcal{F}_u}^\alpha(P) = D_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$ .

Ahora demostraremos que  $D_{\{\mathcal{G}_u\}_{u \in \mathbb{R}^d}}^\alpha(P) \subset \text{HD}^\alpha(P)$ . Sea  $z \in D_{\{\mathcal{G}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$  y  $H$  un semiespacio cerrado tal que  $z$  pertenece a su frontera, entonces  $H = \{x \in \mathbb{R}^d : \langle x, v \rangle \geq \langle z, v \rangle\}$  para algún  $v \in \mathbb{R}^d$ . Según la definición de regiones centrales integrales, existe  $Q \in P^\alpha$  tal que  $f(\langle z, v \rangle) \leq \int f(\langle x, v \rangle) dQ(x)$  para todo  $f \in \{I_{(a, +\infty)} : a \in \mathbb{R}\}$ . Si tomamos la sucesión apropiada de funciones indicadores es fácil ver que,

$$\begin{aligned} 1 &\leq Q(\{x \in \mathbb{R}^d : \langle x, v \rangle \geq \langle z, v \rangle\}) \\ &\leq \alpha^{-1} P(\{x \in \mathbb{R}^d : \langle x, v \rangle \geq \langle z, v \rangle\}) = \alpha^{-1} P(H), \end{aligned}$$

es decir,  $P(H) \geq \alpha$ . Como eso se cumple para cualquier semiespacio cerrado con  $z$  en su frontera, entonces  $z \in \text{HD}^\alpha(P)$ .

Finalmente, se cumple

$$\text{HD}^\alpha(P) \subset D_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P) \subset D_{\{\mathcal{G}_u\}_{u \in \mathbb{R}^d}}^\alpha(P) \subset \text{HD}^\alpha(P)$$

y se obtiene la igualdad global.  $\square$

Aunque las regiones centrales ya han sido sobradamente estudiadas y todos los resultados que comentamos a continuación son conocidos, su tratamiento dentro del marco de las regiones centrales integrales resulta muy sencillo.

El anidamiento de la familia de regiones centrales semiespaciales se deduce del Teorema 4.2.2, la convexidad puede deducirse del Teorema 4.2.7. De los Teoremas 4.2.9 y 4.2.8 podemos deducir respectivamente que las regiones centrales semiespaciales son equivariantes afines, es decir  $AHD^\alpha(P) + b = HD^\alpha(P_{A,b})$  si  $A \in \mathbb{R}^{d \times d}$  no es singular, y  $AHD^\alpha(P) + b \subset HD^\alpha(P_{A,b})$  para cualquier  $A \in \mathbb{R}^{d \times d}$ . Además las regiones centrales semiespaciales son compactas porque la familia de conjuntos de funciones  $\{\mathcal{G}_u\}_{u \in \mathbb{R}^d}$  satisface las hipótesis de los Teoremas 4.2.10 y 4.2.13.

*Ejemplo 4.4.3.* Hemos simulado una muestra de 10 realizaciones de una normal bidimensional  $N_2(0, I)$  y dibujado algunas de las regiones centrales semiespaciales de su distribución empírica en la Figura 4.1.

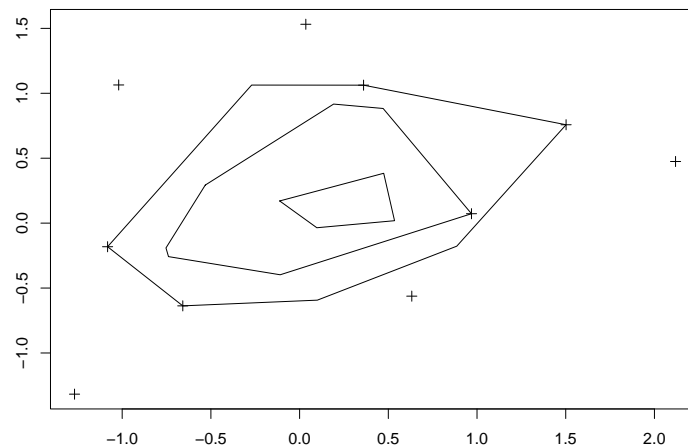


Figura 4.1: Contornos de  $HD^{4/10}(P)$ ,  $HD^{3/10}(P)$  y  $HD^{2/10}(P)$ .

### 4.4.2. Regiones centrales del zonoide

Según hemos visto en el Teorema 3.5.2, las regiones centrales del zonoide se pueden escribir como  $ZD^\alpha(P) = \{\int x dQ(x) : Q \in P^\alpha\}$ .

Antes de demostrar que las regiones centrales del zonoide pertenecen a la familia de regiones centrales integrales, analizamos el comportamiento de sus proyecciones unidimensionales. Recordamos que  $P_u$  es una probabilidad en  $\mathbb{R}$  definida por  $P_u(A) = P(\{x : \langle x, u \rangle \in A\})$  para cualquier  $A \in \mathcal{B}_1$ .

**Lema 4.4.2.** *Para todo  $u \in \mathbb{R}^d$ ,  $ZD^\alpha(P_u) = \{\langle z, u \rangle : z \in ZD^\alpha(P)\}$ .*

*Demostración.* Este resultado es inmediato a partir de la representación de las regiones centrales del zonoide a partir del zonoide elevado y el Teorema 0.3.1.  $\square$

**Teorema 4.4.3.** *El conjunto de funciones  $\mathcal{F} = \{l : \mathbb{R}^d \rightarrow \mathbb{R} \text{ lineal}\}$  y la familia de conjuntos  $\mathcal{F}_u = \{l_u : l_u(x) = \langle x, u \rangle\}$ ,  $u \in \mathbb{R}^d$  inducen cada uno de ellos las regiones centrales del zonoide como regiones centrales integrales, es decir,  $ZD^\alpha(P) = D_{\mathcal{F}}^\alpha(P) = D_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$  para cualquier  $\alpha \in (0, 1]$ ,  $P \in \mathbb{P}$ .*

*Demostración.* Primero probaremos que  $ZD^\alpha(P) = D_{\mathcal{F}}^\alpha(P)$  y a continuación, como para todo  $u \in \mathbb{R}^d$ , tenemos que  $\mathcal{F}_u \subset \mathcal{F}$ , según el Teorema 4.2.1 se cumple  $ZD^\alpha(P) \subset D_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$ . Finalizaremos la demostración probando que  $ZD^\alpha(P) \supset D_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$ .

Sea  $z \in ZD^\alpha(P)$ , por la definición de regiones centrales integrales tiene que existir  $Q \in P^\alpha$  tal que  $z = \int x dQ(x)$ , equivalentemente  $\langle z, u \rangle \leq \int \langle x, u \rangle dQ(x)$  para todo  $u \in \mathbb{R}^d$ , y finalmente se tiene que  $z \in D_{\mathcal{F}}^\alpha(P)$ .

Sea  $z \in D_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$ , entonces claramente  $\langle z, u \rangle \in D_{\{\lambda x\}_{\lambda \in \mathbb{R}}}^\alpha(P_u)$  para todo  $u \in \mathbb{R}^d$  lo que por la primera parte de este resultado implica  $\langle z, u \rangle \in ZD^\alpha(P_u)$  para todo  $u \in \mathbb{R}^d$ . Por la caracterización previa de las regiones centrales del zonoide como regiones centrales integrales y según el Teorema 4.2.3 estas regiones centrales son convexas y entonces por el Lemma 4.4.2  $z \in ZD^\alpha(P)$ . Obsérvese que a partir de los resultados del Capítulo 3 ya sabíamos que las regiones centrales del zonoide eran convexas.  $\square$



El anidamiento de las regiones centrales del zonoide puede deducirse del Teorema 4.2.2, la convexidad del Teorema 4.2.3 y el hecho de que todos los elementos de  $\mathcal{F}$  son convexos. La equivarianza afín se sigue del Teorema 4.2.9. Estas tres primeras conclusiones ya habían sido obtenidas en el Capítulo 3 a partir de la caracterización de las regiones centrales del zonoide como regiones centrales recortadas. A partir de la caracterización como regiones integrales además obtenemos que como la familia de conjuntos  $\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}$  satisface las hipótesis de los Teoremas 4.2.10 y 4.2.13, las regiones centrales del zonoide son compactas, propiedad que ya ha sido demostrada en el Teorema 5.4 de Koshevoy y Mosler (1997b).

*Ejemplo 4.4.4.* Para la distribución de probabilidad empírica del Ejemplo 4.4.3 obtenemos las regiones centrales del zonoide que aparecen en la Figura 4.2.

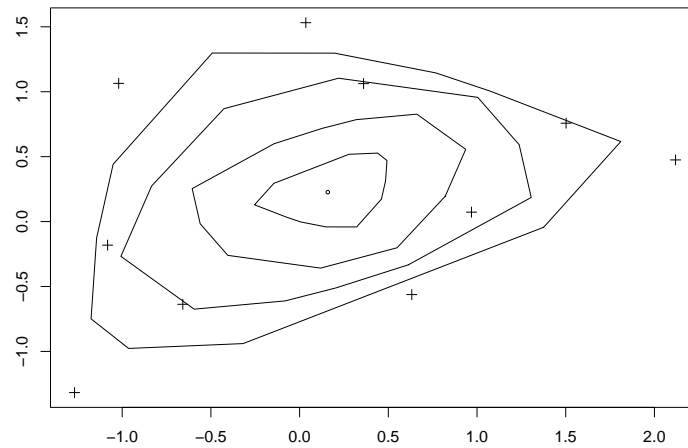


Figura 4.2: Contornos de  $ZD^1(P)$ ,  $ZD^{4/5}(P)$ ,  $ZD^{3/5}(P)$ ,  $ZD^{2/5}(P)$  y  $ZD^{1/5}(P)$ .

### 4.4.3. Regiones centrales de la envolvente de zonoides

Recordamos que según se vio en el Lema 3.5.4, las regiones centrales de la envolvente de zonoides vienen dadas por  $WD^\alpha(P) = \cup_{Q \in P^\alpha} Z(Q)$ .

El zonoide de una distribución de probabilidad se puede caracterizar a partir del siguiente resultado.

**Lema 4.4.4.** *Dado  $P \in \mathbb{P}$ ,  $z \in Z(P)$  si y sólo si  $\langle z, u \rangle_+ \leq \int \langle x, u \rangle_+ dP(x)$  para todo  $u \in \mathbb{R}^d$ .*

*Demostración.* Sabemos que  $Z(P)$  es convexo y además contiene al origen,  $0 \in Z(P)$ , por lo tanto  $z \in Z(P)$  si y sólo si  $Z(\delta_z) = [0, z] \subset Z(P)$  donde  $\delta_z$  es una distribución de probabilidad degenerada en  $z$ . Finalmente, por la caracterización del orden del zonoide como orden estocástico integral dada en el Lema 0.5.8, esto es equivalente a  $\langle z, u \rangle_+ \leq \int \langle x, u \rangle_+ dP(x)$  para todo  $u \in \mathbb{R}^d$ .  $\square$

**Lema 4.4.5.** *Para cualquier  $u \in \mathbb{R}^d$ ,  $WD^\alpha(P_u) = \{\langle z, u \rangle : z \in WD^\alpha(P)\}$ .*

*Demostración.* Como consecuencia de la representación del cuerpo convexo  $W$  según el zonoide elevado y el Teorema 0.3.1,

$$W(P_u) = W(P) \begin{pmatrix} 1 & 0 \\ 0 & u \end{pmatrix}$$

para todo  $u \in \mathbb{R}^d$ . Finalmente por la relación  $WD^\alpha(P) = \alpha^{-1} \text{proj}_\alpha(W(P))$  obtenemos el resultado deseado.  $\square$

Las regiones centrales de la envolvente de zonoides pertenecen a la familia de regiones centrales integrales.

**Teorema 4.4.6.** *El conjunto de funciones  $\mathcal{F} = \{l_+ : l : \mathbb{R}^d \rightarrow \mathbb{R} \text{ lineal}\}$  y la familia de conjuntos  $\mathcal{F}_u = \{(l_u)_+ : l_u(x) = \langle x, u \rangle\}$ ,  $u \in \mathbb{R}^d$  inducen cada uno de ellos las regiones centrales  $WD^\alpha(P)$  como regiones centrales integrales, es decir,  $WD^\alpha(P) = D_{\mathcal{F}}^\alpha(P) = D_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$  para cualquier  $\alpha \in (0, 1]$ ,  $P \in \mathbb{P}$ .*

*Demostración.* En primer lugar demostraremos que  $\text{WD}^\alpha(P) = \text{D}_{\mathcal{F}}^\alpha(P)$  y a continuación, como para todo  $u \in \mathbb{R}^d$ , se cumple  $\mathcal{F}_u \subset \mathcal{F}$ , entonces por el apartado *iii.* del Teorema 4.2.1 se tendrá  $\text{WD}^\alpha(P) \subset \text{D}_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$  y sólo faltará probar que  $\text{WD}^\alpha(P) \supset \text{D}_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$ .

Sea  $z \in \text{WD}^\alpha(P)$ , según el Lema 3.5.4 esto es equivalentemente a la existencia de  $Q \in P^\alpha$  tal que  $z \in \text{Z}(Q)$  y de acuerdo con el Lema 4.4.4 se cumple  $\langle z, u \rangle_+ \leq \int \langle x, u \rangle_+ dQ(x)$  para todo  $u \in \mathbb{R}^d$  con lo que finalmente se tiene que  $z \in \text{D}_{\mathcal{F}}^\alpha(P)$ .

Sea  $z \in \text{D}_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$ , por definición  $\langle z, u \rangle_+ \in \text{D}_{\{(\lambda x)_+\}_{\lambda \in \mathbb{R}}}(P_u)$  para todo  $u \in \mathbb{R}^d$  y por la primera parte de este resultado esto implica  $\langle z, u \rangle_+ \in \text{WD}^\alpha(P_u)$  para todo  $u \in \mathbb{R}^d$ . Por la caracterización previa de las regiones centrales de la envolvente de zonoides como regiones integrales y el Teorema 4.2.3, estas regiones centrales son convexas y entonces por el Lema 4.4.5,  $z \in \text{ZD}^\alpha(P)$ . Obsérvese que a partir de los resultados del Capítulo 3 ya sabíamos que las regiones centrales de la envolvente de zonoides eran convexas.  $\square$

Las regiones centrales de la envolvente de zonoides forman una familia de regiones anidadas y convexas como se deduce de los Teoremas 4.2.2 y 4.2.3 y del hecho de que los elementos de  $\mathcal{F}$  sean convexas. Obsérvese que estas dos primeras consecuencias ya habían sido obtenidas en el Capítulo 3. A partir de los Teoremas 4.2.10 y 4.2.13 y por la caracterización de las regiones centrales de la envolvente de zonoides como  $\text{D}_{\{\mathcal{F}_u\}_{u \in \mathbb{R}^d}}^\alpha(P)$  éstas son compactas. Finalmente, por el Teorema 4.2.9 y el hecho de que para cualquier matriz no singular  $A \in \mathbb{R}^{d \times d}$ , se cumple que  $\mathcal{F} = \{f \circ T_{A,0} : f \in \mathcal{F}\}$ , tenemos  $\text{WD}^\alpha(P_{A,0}) = A\text{WD}^\alpha(P)$ . Claramente estas regiones centrales no son equivariantes en localización.

*Ejemplo 4.4.5.* Para la distribución de probabilidad empírica del Ejemplo 4.4.3 obtenemos las regiones centrales de la envolvente de zonoides que aparecen en la Figura 4.3.

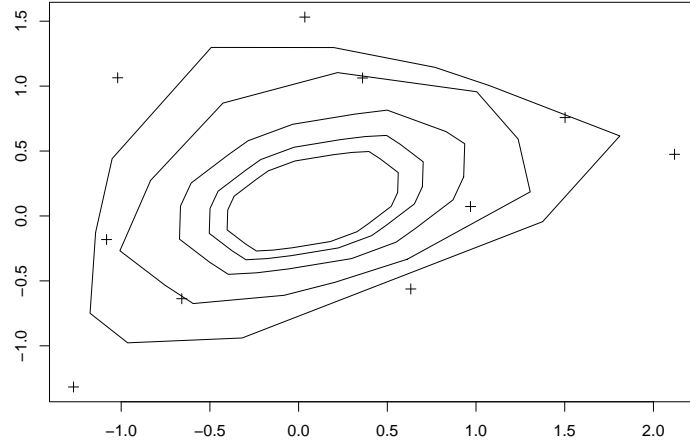


Figura 4.3: Contornos de  $WD^1(P)$ ,  $WD^{4/5}(P)$ ,  $WD^{3/5}(P)$ ,  $WD^{2/5}(P)$  y  $WD^{1/5}(P)$ .

#### 4.4.4. Regiones centrales integrales y recortadas

Las regiones centrales integrales inducidas por un único conjunto de funciones pertenecen a la familia de regiones centrales recortadas. El estimador de localización se definirá a partir de las integrales de las funciones del conjunto de generadores. Regiones centrales de este tipo son, por ejemplo, las del zonoide y las de la envolvente de zonoides.

**Teorema 4.4.7.** *Cualquier región central integral generada por un único conjunto de funciones es una región central recortada para el estimador de localización apropiado.*

*Demostración.* Sea  $\mathcal{F}$  un conjunto de funciones, para cualquier  $P \in \mathbb{P}$  definimos  $L^{\mathcal{F}}(P) = \cap_{f \in \mathcal{F}} f^{-1}((-\infty, \int f(x)dP(x)])$ , entonces  $D_{\mathcal{F}}^{\alpha}(P) = D_{L^{\mathcal{F}}}^{\alpha}(P)$ .  $\square$

## 4.5. Comentarios y problemas abiertos

Existen varios trabajos que relacionan las ordenaciones estocásticas con las regiones centrales. Por afinidad con la línea seguida en esta memoria, destacamos los estudios de Koshevoy y Mosler (1997b) y (1998). Como los órdenes estocásticos integrales son la familia más estudiada de ordenaciones estocásticas, hemos utilizado el marco que proporcionan los recortes de una probabilidad para definir la nueva familia de regiones centrales integrales. De este modo ahondamos en las relaciones de los órdenes estocásticos con las regiones centrales.

En el futuro pretendemos seguir trabajando con problemas relacionados con el estudio del grado de centralidad de un punto con respecto a una probabilidad. En concreto, una de nuestras ideas es construir una generalización de este problema y analizar el grado de afinidad de una probabilidad con respecto a un funcional de capacidad.

Dado un conjunto aleatorio cerrado  $X$ , su funcional de capacidad se denota por  $T_X$  y se define para cualquier compacto no vacío en  $\mathbb{R}^d$ ,  $K \in \mathcal{K}$  como

$$T_X(K) := \Pr(X \cap K \neq \emptyset).$$

La distribución de un conjunto aleatorio cerrado queda caracterizada por su funcional de capacidad.

El grado de afinidad de una probabilidad con respecto a un funcional de capacidad podría medirse mediante funcionales que satisfacen condiciones semejantes a las que se proponen en la Definición 0.4.1 para una función de profundidad estadística. Serían funcionales acotados, preferiblemente normalizados al intervalo  $[0, 1]$  que cumplen las siguientes condiciones:

- i. *invarianza por transformaciones ortogonales*, se consideran transformaciones afines definidas por matrices ortogonales porque son las que preservan las distancias y por lo tanto respetan las especiales condiciones geométricas de los conjuntos aleatorios;

- ii. *maximalidad en el centro*, donde los centros de un funcional de capacidad serían probabilidades que cumplen ciertas condiciones que se consideren convenientes en cada caso. Las distribuciones de probabilidad de las selecciones de un conjunto aleatorio con dicho funcional de capacidad podrían ser, por ejemplo, centros;
- iii. *conservación del grado de afinidad mínimo bajo combinaciones lineales convexas*, es decir, fijado un funcional de capacidad, el grado de afinidad de una probabilidad que se obtiene como combinación lineal convexa de otras dos probabilidades es, al menos, el menor de los grados de afinidad de estas probabilidades;
- iv. *convergencia hacia cero cuando la probabilidad “se aleja arbitrariamente del origen”*, este alejamiento del origen puede entenderse, por ejemplo, como un crecimiento arbitrariamente grande del primer momento de la probabilidad.

El grado de afinidad de una probabilidad  $P \in \mathbb{P}$  con respecto a un funcional de capacidad  $T_X$  podría adoptar una expresión del tipo de las dos que aparecen a continuación,

$$(1 + \inf\{\lambda > 0 : P(K) \leq T(K^\lambda) \text{ para todo } K \in \mathcal{K}\})^{-1}; \quad (4.4)$$

$$2(1 + \inf\{\lambda > 0 : P(K) \leq T(K^\lambda) + \lambda \text{ para todo } K \in \mathcal{K}\})^{-1} - \frac{1}{2}. \quad (4.5)$$

Las aplicaciones de estos grados de afinidad pueden ser muy diversas. Principalmente permiten ordenar distribuciones de probabilidad con arreglo a un funcional de capacidad. A continuación se explica un ejemplo de aplicación práctica:

Supongamos que la sección de recursos humanos de una empresa ha de seleccionar al individuo más idóneo, dentro de los que se presentan para ocupar cierto puesto. Si el comportamiento de cada uno de estos individuos en las pruebas de selección es modelado mediante una distribución de probabilidad y el conjunto de individuos idóneos está modelado según un funcional

de capacidad, entonces ordenando a los individuos según su grado de afinidad con respecto a dicho funcional de capacidad, habremos obtenido una ordenación según la idoneidad para ocupar el puesto que se ofertaba.





## Capítulo 5

# Profundidad según un número de observaciones

En los Capítulos 3 y 4 se han estudiado regiones centrales, con respecto a una probabilidad, inducidas por ciertos funcionales de probabilidades pertenecientes a recortes de la probabilidad original. En el primer caso estos funcionales eran estimadores de localización cualesquiera y en el segundo venían dados por familias de conjuntos de funciones.

A lo largo del presente capítulo se considerarán funciones de profundidad con respecto a una probabilidad cuyo valor en un punto viene determinado por el número de observaciones independientes de un vector aleatorio que sigue la distribución de probabilidad anterior necesarias para que el punto esté contenido en su envolvente convexa. Tales funciones de profundidad son, por su propia filosofía, diferentes a las consideradas con anterioridad.

El procedimiento del “convex hull peeling”, véase por ejemplo Eddy (1982), consiste en trazar la envolvente convexa de una nube de puntos, eliminar los puntos de la nube en la frontera de la envolvente, volver a trazar la envolvente convexa y así sucesivamente para obtener la capa a la que cierto punto pertenece. Este procedimiento es muy habitual en la práctica para estudiar el grado de “exterioridad” de un punto con respecto a un conjunto de observaciones. Las nuevas profundidades están inspiradas en el “convex

hull peeling” y en los trabajos de Liu (1990) y Chiu y Molchanov (2003). El primer inconveniente que nos encontramos al estudiarlas es la dificultad de su cálculo en la práctica. Para algunas de ellas este cálculo se basa en la distribución del número de semiesferas, equivalentemente hemisferios, no necesariamente uniformemente distribuidas, que se necesitan para recubrir una esfera. Nosotros resolvemos este problema en  $\mathbb{R}^2$ .

En primer lugar, proponemos el estudio de funciones de profundidad dadas por el número de observaciones independientes de un vector aleatorio necesarias para contener a un punto en su envolvente convexa con cierta probabilidad y por el número medio de observaciones independientes de un vector aleatorio necesarias para contener a un punto en su envolvente convexa y damos además un método para calcular estas profundidades en  $\mathbb{R}^2$ . A continuación, estudiamos regiones centrales definidas como la esperanza de Aumann de la envolvente convexa de un conjunto de copias independientes de un vector aleatorio y su relación con el orden lineal convexo.

## 5.1. Introducción

Las construcciones que se realizan en este capítulo están inspiradas en las de Liu (1990) y Chiu y Molchanov (2003). En este último trabajo los autores consideran el problema de la pertenencia de un punto a la envolvente convexa de ciertos puntos de un proceso, mientras que Liu (1990) considera el problema de la pertenencia del punto a un simplex cuyos vértices son copias independientes de un vector aleatorio.

A lo largo del presente capítulo  $\xi$  denotará un vector aleatorio genérico en  $\mathbb{R}^d$  definido en un espacio de probabilidad  $(\Omega, \mathcal{A}, \Pr)$ , mientras que  $\xi_1, \xi_2, \dots$  serán vectores aleatorios independientes distribuidos igual que  $\xi$  y definidos en el mismo espacio de probabilidad. Más brevemente diremos que  $\xi_1, \xi_2, \dots$  son *copias independientes de  $\xi$* .

Si tenemos  $n$  copias independientes de una variable aleatoria  $\xi$ , denotamos por  $\xi_{i:n}$  al  $i$ -ésimo estadístico ordenado con  $1 \leq i \leq n$ .

Chiu y Molchanov (2003) definen el *grado de un punto concreto de un proceso puntos* como la variable aleatoria  $N$  que representa el menor número de puntos del proceso comenzando por el más cercano al punto concreto de tal modo que dicho punto está contenido en el interior de su envolvente convexa.

Liu (1990) define la *profundidad simplicial* de un punto  $x \in \mathbb{R}^d$  con respecto a  $P_\xi$ , la probabilidad inducida por el vector aleatorio  $\xi$ , a la que denotamos por  $\text{SD}(x; P_\xi)$ , como la probabilidad de que  $x$  esté contenido en la envolvente convexa de un simplex cuyos vértices son  $d + 1$  copias independientes de  $\xi$ , es decir,

$$\text{SD}(x; P_\xi) := \Pr(x \in \text{co}\{\xi_1, \dots, \xi_{d+1}\}).$$

*Definición 5.1.1.* Definimos la variable aleatoria *grado de un punto  $x$  con respecto a una distribución de probabilidad  $P_\xi$*  como

$$\text{grado}(x; P_\xi) := \min\{n \in \mathbb{N} : x \in \text{co}\{\xi_1, \dots, \xi_n\}\}.$$

Obsérvese que esta variable aleatoria toma valores en  $\mathbb{N} \cup \{+\infty\}$ , pero como razona Billingsley (1986), pág. 185 esto no supone ningún inconveniente para su medibilidad.

La profundidad simplicial y la variable aleatoria grado están íntimamente relacionadas. Así, una se puede expresar en función de la otra como

$$\text{SD}(x; P_\xi) = \Pr(\text{grado}(x; P_\xi) \leq d + 1).$$

La noción de simetría multivariante respecto de una distribución de probabilidad está muy relacionada con la de estimador de localización y por lo tanto con la de punto más profundo, como se razona en los trabajos de Small (1990) y Zuo y Serfling (2000b). A lo largo del presente capítulo haremos uso de la *simetría esférica* y la *simetría angular*.

*Definición 5.1.2.* Dado  $\xi$  un vector aleatorio  $d$ -dimensional y  $x \in \mathbb{R}^d$ , decimos que:

- i. la distribución de  $\xi$  es *esféricamente simétrica* respecto de  $x$  si toda transformación lineal ortogonal de  $\xi - x$  sigue la misma distribución;
- ii. la distribución  $\xi$  es *angularmente simétrica* en torno a  $x$  si  $(\xi - x)/\|\xi - x\|$  y  $-(\xi - x)/\|\xi - x\|$  siguen la misma distribución.

## 5.2. Profundidades relacionadas con el grado

En esta sección se definen dos profundidades relacionadas con la variable aleatoria grado, la *profundidad de conteo* y la *profundidad del grado medio*. El resultado más importante que se obtiene sobre ellas es que ambas son funciones de profundidad estadística en el sentido de la Definición 0.4.1. Además, presentamos varios ejemplos en los que se calculan dichas profundidades apoyándonos en la relación existente entre la variable aleatoria grado y el problema del recubrimiento de una esfera mediante semiesferas, problema que resolvemos en el caso de  $\mathbb{R}^2$ .

*Definición 5.2.1.* La *profundidad de conteo* de un punto con respecto a una distribución de probabilidad en  $\mathbb{R}^d$  es el inverso del mínimo número de copias independientes de un vector aleatorio distribuido según la probabilidad necesarias para que el punto pertenezca a su envolvente convexa con cierta probabilidad  $p \in (0, 1)$ . Sea  $x \in \mathbb{R}^d$ , entonces

$$D_p(x; P_\xi) := \left( \min\{n \in \mathbb{N} : \Pr(x \in \text{co}\{\xi_1, \dots, \xi_n\}) \geq p\} \right)^{-1}.$$

Al igual que sucede con la profundidad simplicial, podemos relacionar la profundidad de conteo con la variable aleatoria grado,

$$D_p(x; P_\xi) = \left( \min\{n \in \mathbb{N} : \Pr(\text{grado}(x; P_\xi) \leq n) \geq p\} \right)^{-1}.$$

*Definición 5.2.2.* La *profundidad del grado medio* se define como el inverso del grado medio de un punto con respecto a una distribución de probabilidad. Sea  $x \in \mathbb{R}^d$ , entonces

$$\text{ED}(x; P_\xi) := \left( \mathbb{E}\text{grado}(x; P_\xi) \right)^{-1}.$$

En la definición de la *profundidad de conteo* y en la de la *profundidad del grado medio* tomamos  $(+\infty)^{-1} = 0$ .

Como el grado mide el alejamiento del punto al centro de una distribución de probabilidad, cualquier transformación decreciente de su media o cuantil  $p$  sería un candidato sensato para función de profundidad. Hemos elegido el inverso por simplicidad.

### 5.2.1. Propiedades

Veamos que las nociones de profundidad de conteo y profundidad del grado medio son en efecto funciones de profundidad estadística en el sentido de la Definición 0.4.1.

**Proposición 5.2.1.** *Una transformación afín aplicada simultáneamente a un vector aleatorio y a un punto nunca disminuye ni la profundidad de conteo ni la del grado medio.*

*Demostración.* Sea  $A \in \mathbb{R}^{d \times d}$  y  $b, x \in \mathbb{R}^d$ , si  $x \in \text{co}\{\xi_1, \dots, \xi_n\}$ , entonces  $Ax + b \in \text{co}\{A\xi_1 + b, \dots, A\xi_n + b\}$ , por lo tanto  $\Pr(x \in \text{co}\{\xi_1, \dots, \xi_n\}) \leq \Pr(Ax + b \in \text{co}\{A\xi_1 + b, \dots, A\xi_n + b\})$ . Finalmente  $\text{ED}(Ax + b; P_{A\xi+b}) \geq \text{ED}(x; P_\xi)$  y  $D_p(Ax + b; P_{A\xi+b}) \geq D_p(x; P_\xi)$  para cualquier  $p \in (0, 1)$ .  $\square$

**Proposición 5.2.2.** *La profundidad de conteo y la del grado medio son invariantes afines.*

*Demostración.* Si  $A$  no es singular,  $x \in \text{co}\{\xi_1, \dots, \xi_n\}$  si y sólo si  $Ax + b \in \text{co}\{A\xi_1 + b, \dots, A\xi_n + b\}$  con lo que tenemos  $\text{ED}(Ax + b; P_{A\xi+b}) = \text{ED}(x; P_\xi)$  y  $D_p(Ax + b; P_{A\xi+b}) = D_p(x; P_\xi)$  para cualquier  $p \in (0, 1)$ .  $\square$

Liu (1990) demuestra que la profundidad simplicial es decreciente en rayos que parten del punto de simetría angular de una distribución siempre que ésta sea absolutamente continua. Tal propiedad la cumplen también la profundidad de conteo y la del grado medio. Como la demostración es similar a la del Teorema 3 en Liu (1990), utilizaremos una notación muy semejante.

**Proposición 5.2.3.** *Si la distribución de  $\xi$  es absolutamente continua y angularmente simétrica en torno a  $\theta$ , entonces para cualesquiera  $\alpha \in [0, 1]$ ,  $x \in \mathbb{R}^d$  y  $n \in \mathbb{N}$ , se cumple*

$$\Pr(\text{grado}(\theta + \alpha(x - \theta); P_\xi) \leq n) \geq \Pr(\text{grado}(x; P_\xi) \leq n).$$

*Demostración.* Para hacer la demostración más sencilla, tomamos  $\theta$  el punto de simetría angular igual al origen,  $\theta = 0$ , lo que no supone ninguna restricción ya que del mismo modo que según la Proposición 5.2.2 las profundidades establecidas son invariantes afines, la variable aleatoria grado también lo es.

Queremos demostrar que para cualquier  $\alpha \in [0, 1]$ , se cumple  $\Pr(x \in \text{co}\{\xi_1, \dots, \xi_n\}) \leq \Pr(\alpha x \in \text{co}\{\xi_1, \dots, \xi_n\})$  para cualquier  $n \in \mathbb{N}$ . Como la distribución de  $\xi$  es absolutamente continua, la probabilidad  $P_\xi$  de cualquier hiperplano es 0. Así, si  $n \leq d$ ,  $\Pr(x \in \text{co}\{\xi_1, \dots, \xi_n\}) = \Pr(\alpha x \in \text{co}\{\xi_1, \dots, \xi_n\}) = 0$ . Tomemos, por tanto,  $n \geq d + 1$ .

Consideramos el segmento orientado que va de  $\alpha x$  a  $x$  y los sucesos que nos indiquen bien que entra o que sale del politopo aleatorio  $\text{co}\{\xi_1, \dots, \xi_n\}$ ,

$$A_{\text{in}} := \{\text{el segmento orientado de } \alpha x \text{ a } x \text{ entra en } \text{co}\{\xi_1, \dots, \xi_n\}\}$$

$$A_{\text{out}} := \{\text{el segmento orientado de } \alpha x \text{ a } x \text{ sale de } \text{co}\{\xi_1, \dots, \xi_n\}\}.$$

Obsérvese que como la probabilidad de todo hiperplano es 0, con probabilidad 1 la dimensión afín de  $\text{co}\{\xi_1, \dots, \xi_n\}$  es  $d$ .

Dados  $d$  puntos  $\{x_1, \dots, x_d\} \in \mathbb{R}^d$  con dimensión afín  $d - 1$ , denotamos por  $H(\{x_1, \dots, x_d\})$  al semiespacio cerrado que tiene por frontera la clausura afín de los  $d$  puntos,  $\text{aff}\{x_1, \dots, x_d\}$ , y contiene al origen de coordenadas. Si  $0 \in \text{aff}\{x_1, \dots, x_d\}$ , el semiespacio cerrado  $H(\{x_1, \dots, x_d\})$  será cualquiera de los dos que delimita  $\text{aff}\{x_1, \dots, x_d\}$ . Denotamos por  $C$  al conjunto de  $d$ -tuplas de elementos de  $\mathbb{R}^d$  cuya envolvente convexa interseca con el segmento  $\text{co}\{x, \alpha x\}$ ,  $C = \{(x_1, \dots, x_d) : x_i \in \mathbb{R}^d, \text{co}\{x_1, \dots, x_d\} \cap \text{co}\{x, \alpha x\} \neq \emptyset\}$ .

Para cada subconjunto  $S$  de  $d$  elementos de  $\{1, 2, \dots, n\}$ , el suceso  $A_{\text{in}}^S$  consiste en que la envolvente convexa de las  $d$  copias independientes del vector aleatorio  $\{\xi_i : i \in S\}$ , interseca con el segmento orientado que une  $\alpha x$

con  $x$  y de entre las restantes copias independientes de  $\xi$ , no hay ninguna en  $H(\{\xi_i : i \in S\})$ , es decir  $\{\xi_i : i \in \{1, \dots, n\} \setminus S\} \subset \mathbb{R}^d \setminus H(\{\xi_i : i \in S\})$ .

En las mismas condiciones el suceso  $A_{\text{out}}^S$  consiste en que la envolvente convexa de las  $d$  copias independientes del vector aleatorio,  $\{\xi_i : i \in S\}$ , interseca con el segmento orientado que une  $\alpha x$  con  $x$  y las restantes copias independientes de  $\xi$  están todas en  $H(\{\xi_i : i \in S\})$ , es decir  $\{\xi_i : i \in \{1, \dots, n\} \setminus S\} \subset H(\{\xi_i : i \in S\})$ .

Para  $\beta \in \{\alpha, 1\}$ , denotamos por  $B_\beta$  al suceso " $\beta x$  pertenece a la envolvente convexa de las  $n$  copias independientes de  $\xi$ ", esto es  $B_\beta = \{\beta x \in \text{co}\{\xi_1, \xi_2, \dots, \xi_n\}\}$ , entonces tenemos que

$$\Pr(\alpha x \in \text{co}\{\xi_1, \dots, \xi_n\}) - \Pr(x \in \text{co}\{\xi_1, \dots, \xi_n\}) = \Pr(B_\alpha \setminus B_1) - \Pr(B_1 \setminus B_\alpha).$$

Estos nuevos sucesos son fácilmente expresables a partir de los anteriores,

$$B_\alpha \setminus B_1 = A_{\text{out}} \setminus A_{\text{in}}, \quad B_1 \setminus B_\alpha = A_{\text{in}} \setminus A_{\text{out}}.$$

Podemos simplificar la expresión que tenemos, obteniendo,

$$\begin{aligned} & \Pr(\alpha x \in \text{co}\{\xi_1, \dots, \xi_n\}) - \Pr(x \in \text{co}\{\xi_1, \dots, \xi_n\}) \\ &= \Pr(A_{\text{out}}) - \Pr(A_{\text{out}} \cap A_{\text{in}}) - (\Pr(A_{\text{in}}) - \Pr(A_{\text{in}} \cap A_{\text{out}})) \\ &= \Pr(A_{\text{out}}) - \Pr(A_{\text{in}}). \end{aligned}$$

Ahora como  $\binom{n}{d}$  es el número de conjuntos diferentes de  $d$  elementos que podemos tomar con  $n$  elementos diferentes, dado  $S \subset \{1, \dots, n\}$  con  $d$  elementos,  $\Pr(A_{\text{out}}) = \binom{n}{d} \Pr(A_{\text{out}}^S)$ , mientras que  $\Pr(A_{\text{in}}) = \binom{n}{d} \Pr(A_{\text{in}}^S)$ . Por otra parte se cumple

$$\Pr(A_{\text{out}}^S) = \int_C \Pr(\xi \in H(\{x_1, \dots, x_d\}))^{n-d} dP_\xi(x_1) \dots dP_\xi(x_d),$$

mientras que

$$\Pr(A_{\text{in}}^S) = \int_C \Pr(\xi \notin H(\{x_1, \dots, x_d\}))^{n-d} dP_\xi(x_1) \dots dP_\xi(x_d).$$

Por la simetría angular de la distribución de  $\xi$  respecto del origen y el hecho de que el origen pertenece a los semiespacios  $H(\{x_1, \dots, x_n\})$ , se cumple  $\Pr(\xi \in H(\{x_1, \dots, x_d\})) \geq 1/2$ , y por tanto  $\Pr(\xi \notin H(\{x_1, \dots, x_d\}))^{n-d} \leq \Pr(\xi \in H(\{x_1, \dots, x_d\}))^{n-d}$ , de donde  $\Pr(A_{\text{out}}^S) \geq \Pr(A_{\text{in}}^S)$ , en consecuencia  $\Pr(A_{\text{out}}) \geq \Pr(A_{\text{in}})$  y finalmente

$$\Pr(\alpha x \in \text{co}\{\xi_1, \dots, \xi_n\}) \geq \Pr(x \in \text{co}\{\xi_1, \dots, \xi_n\}).$$

□

Las dos diferencias esenciales con el Teorema 3 en Liu (1990) es que nosotros hemos tenido que considerar politopos en lugar de símlices y hemos tenido que negar la condición de que haya alguna realización de la variable aleatoria en ciertos semiespacios, en lugar de trabajar con la condición positiva.

De la Proposición 5.2.3 se obtiene de modo inmediato el siguiente resultado que aparece en términos de la profundidad de conteo y de la del grado medio.

**Corolario 5.2.4.** *Si la distribución de  $\xi$  es absolutamente continua y angularmente simétrica en torno a  $\theta$ , entonces para cualesquiera  $p \in (0, 1)$ ,  $\alpha \in [0, 1]$  y  $x \in \mathbb{R}^d$ , se cumple  $D_p(x; P_\xi) \leq D_p(\theta + \alpha(x - \theta); P_\xi)$  y  $\text{ED}(x; P_\xi) \leq \text{ED}(\theta + \alpha(x - \theta); P_\xi)$ .*

El Corolario 5.2.4 es fundamental para la justificación del estudio de profundidades relacionadas con el grado. En el caso unidimensional relaciona los puntos de profundidad máxima con la mediana. Dada cualquier mediana de una variable aleatoria  $\xi$  que no constituya un átomo de la probabilidad inducida  $P_\xi$ , la distribución de  $\xi$  será angularmente simétrica en torno a ella. Así, si tenemos una variable aleatoria cuya distribución es absolutamente continua, su mediana será el punto de profundidad máxima con respecto a la profundidad de conteo y a la del grado medio. Si la distribución de  $\xi$  no es absolutamente continua, esto no se cumple necesariamente como se ve en el siguiente ejemplo.



*Ejemplo 5.2.1.* Sea  $\xi$  una variable aleatoria discreta que toma los valores  $-1$  y  $1$  con probabilidad  $0'45$  y el valor  $0$  con probabilidad  $0'1$ . Claramente la mediana de  $\xi$  es  $0$ , pero los puntos de máxima profundidad de conteo para cualquier  $p$  mayor que  $0'1$  son  $-1$  y  $1$ . Con la profundidad del grado medio sucede lo mismo, mediante un sencillo cómputo podemos calcular que las profundidades de los puntos  $-1$  y  $1$  son  $\text{ED}(-1; P_\xi) = \text{ED}(1; P_\xi) = 0'45$ , mayores que la profundidad de la mediana,  $\text{ED}(0; P_\xi) = 11/29$ .

Puesto que según el Corolario 5.2.4, las profundidades de conteo y del grado medio alcanzan su valor máximo en el punto de simetría angular de las probabilidades absolutamente continuas y son decrecientes en rayos que parten de dicho punto, para demostrar que son funciones de profundidad, sólo nos falta comprobar que convergen hacia cero para sucesiones de puntos cuya norma diverge a  $+\infty$ .

**Proposición 5.2.5.** *Las profundidades de conteo y del grado medio convergen a cero para puntos cuya norma diverge a  $+\infty$ .*

*Demostración.* Para cualquier  $n \in \mathbb{N}$ , se cumple

$$\Pr(x \in \text{co}\{\xi_1, \dots, \xi_n\}) \leq n\Pr(\|\xi\| \geq \|x\|) = nP_\xi(\mathbb{R}^d \setminus \text{int}(B_{\|x\|})),$$

donde  $\text{int}(B_{\|x\|})$  denota la bola abierta centrada en el origen y de radio  $\|x\|$ . Según el Teorema 0.1.3 todas las probabilidades de  $\mathbb{P}$  son tight, por lo tanto también lo es  $P_\xi$  y así  $nP_\xi(\mathbb{R}^d \setminus \text{int}(B_{\|x\|}))$  converge a  $0$  cuando  $\|x\| \rightarrow +\infty$ . Es entonces claro que para cualquier  $p \in (0, 1)$ ,

$$\lim_{\|x\| \rightarrow \infty} \text{ED}(x; P_\xi) = 0 \quad \text{y} \quad \lim_{\|x\| \rightarrow \infty} D_p(x; P_\xi) = 0$$

□

**Teorema 5.2.6.** *La profundidad de conteo y la del grado medio son funciones de profundidad estadísticas en el sentido de la Definición 0.4.1.*

### 5.2.2. Ejemplos y aplicaciones

A continuación explicamos cómo calcular las profundidades relacionadas con el grado a través del estudio del recubrimiento de una esfera con semiesferas, problema que resolvemos completamente en  $\mathbb{R}^2$ . Además, presentamos algunos ejemplos de los valores que toman estas profundidades para puntos que satisfacen ciertas condiciones de simetría multivariante.

#### Relación del grado con el recubrimiento de esferas

Dados  $x, y_1, \dots, y_n \in \mathbb{R}^d$  con  $y_i \neq x$  para todo  $i \in \{1, \dots, n\}$ , el punto  $x$  pertenece a la envolvente convexa del conjunto  $\{y_1, \dots, y_n\}$  si y sólo si todo semiespacio cerrado que contenga a  $x$  contiene algún punto del conjunto. Equivalentemente, dada  $S^{d-1}(x)$  la esfera de radio unidad centrada en  $x$ , las semiesferas de  $S^{d-1}(x)$  con centro en las proyecciones de  $y_i$  sobre  $S^{d-1}(x)$ , es decir, en las intersecciones de los rayos con origen en  $x$  y que pasan por cada  $y_i$  con la esfera unidad centrada en  $x$ , recubren la esfera  $S^{d-1}(x)$ .

*Ejemplo 5.2.2.* Sea  $\xi$  un vector aleatorio en  $\mathbb{R}^d$  tal que  $(\xi - x)/\|\xi - x\|$  es esféricamente simétrico respecto del origen y  $P_\xi(\{x\}) = 0$ . Las proyecciones de las observaciones del vector aleatorio  $\xi$  en la esfera  $S^{d-1}(x)$  están uniformemente distribuidas sobre tal esfera. En el Teorema 1.5 en Hall (1988) se calcula la probabilidad de que  $n$  semiesferas independientes y uniformemente distribuidas recubran completamente la esfera, y dicha probabilidad en nuestro marco nos proporciona la distribución de la variable aleatoria grado,

$$\Pr(\text{grado}(x; P_\xi) \leq n) = 1 - 2^{-n+1} \sum_{i=0}^{d-1} \binom{n-1}{i}, \quad \text{para } n \geq d. \quad (5.1)$$

A partir de la fórmula (5.1) se puede calcular la profundidad del grado medio que será el inverso del número esperado de semiesferas uniformemente distribuidas necesarias para recubrir una esfera y que aparece, por ejemplo, en Hall (1988), pág. 69

$$\text{ED}(x; P_\xi) = (\mathbb{E}\text{grado}(x; P_\xi))^{-1} = \frac{1}{2d+1}.$$

Además, si un vector aleatorio  $\xi$  en  $\mathbb{R}^d$  y  $x \in \mathbb{R}^d$  satisfacen las condiciones anteriores y otro vector aleatorio  $\eta$  en  $\mathbb{R}^{d-1}$  e  $y \in \mathbb{R}^{d-1}$  también satisfacen las mismas condiciones, Wendel (1962) obtiene una relación que expresada en términos de la variable aleatoria grado adopta la siguiente forma

$$\begin{aligned} \Pr(\text{grado}(x; P_\xi) \leq n) \\ = \frac{1}{2} \left( \Pr(\text{grado}(x; P_\xi) \leq n-1) + \Pr(\text{grado}(y; P_\eta) \leq n-1) \right). \end{aligned}$$

De esta fórmula y de (5.1) se deduce fácilmente la siguiente relación que es de utilidad para calcular la profundidad de conteo para cualquier  $p \in (0, 1)$ . Si  $n \geq d+1$ , se cumple

$$\Pr(\text{grado}(x; P_\xi) \leq n) = \Pr(\text{grado}(x; P_\xi) \leq n-1) + 2^{-n+1} \binom{n-2}{d-1}.$$

### Grado en $\mathbb{R}^2$

Sea  $\xi$  un vector aleatorio en  $\mathbb{R}^2$ , veremos cómo calcular la profundidad de conteo y del grado medio de cualquier punto  $x \in \mathbb{R}^2$  con respecto a  $\xi$  cuando  $P_\xi(\{x\}) = 0$ . Si  $P_\xi(\{x\}) \neq 0$ , el cómputo de las profundidades anteriores puede realizarse con la misma técnica, pero es algo más engorroso. Como estas dos funciones de profundidad son equivariantes respecto de las traslaciones, de hecho son equivariantes afines, es suficiente explicar cómo se calculan para  $x = 0$ .

Sean  $S_-^1$  y  $S_+^1$  dos semicircunferencias complementarias de  $S^1$ , la esfera unidad centrada en el origen en  $\mathbb{R}^2$ , es decir, la circunferencia unidad centrada en el origen, tales que si  $H \in S_-^1 \cap S_+^1$  es la recta que definen, se cumple  $P_\xi(H) = 0$ . El grado de 0 con respecto a  $\xi$  es a lo sumo  $n$  si y sólo si las primeras  $n$  semicircunferencias con centro en la proyección de una observación del vector aleatorio  $\xi$  recubren la circunferencia  $S^1$ . Definimos dos sucesos,

$$\begin{aligned} A_n &:= \{\text{las primeras } n \text{ semicircunferencias recubren la circunferencia}\}, \\ B_i^n &:= \{\text{de las primeras } n \text{ proyecciones, exactamente } i \text{ pertenecen a } S_-^1\}. \end{aligned}$$

Por la ley de la probabilidad total,  $\Pr(A_n) = \sum_{i=1}^{n-1} \Pr(A_n|B_i^n)\Pr(B_i^n)$ .

Es claro que  $\Pr(B_i^n)$  es la probabilidad de  $i$  éxitos en  $n$  realizaciones independientes, con  $q = P_\xi(S_-^1)$  la probabilidad de un éxito,

$$\Pr(B_i^n) = \Pr(\text{Binom}(n, q) = i) = \binom{n}{i} q^i (1-q)^{n-i},$$

donde  $\text{Binom}(n, q)$  es una variable aleatoria con distribución binomial de parámetros  $n$  y  $q$  definida en el espacio de probabilidad  $(\Omega, \mathcal{A}, \Pr)$ . Para todo  $i \in \mathbb{N}$ , definimos la variable aleatoria  $\eta_i$  como el ángulo de la  $i$ -ésima proyección sobre  $S_-^1$  medido empezando en el punto  $(0, -1)$  y la variable aleatoria  $\zeta_i$  como el ángulo de la  $i$ -ésima proyección sobre  $S_+^1$  medido empezando en el punto  $(0, 1)$ , ambas en la dirección de las agujas del reloj. Se tiene que

$$\begin{aligned} \Pr(A_n|B_i^n) &= 1 - \Pr(\{\eta_{i:i} < \zeta_{1:n-i}\} \cup \{\eta_{1:i} > \zeta_{n-i:n-i}\}) \\ &= 1 - \Pr(\eta_{i:i} < \zeta_{1:n-i}) - \Pr(\eta_{i:i} < \zeta_{1:n-i}). \end{aligned}$$

*Ejemplo 5.2.3.* Sea  $\xi$  un vector aleatorio en  $\mathbb{R}^2$  angularmente simétrico en torno a  $x$  y tal que  $P_\xi(\{x\}) = 0$ . Para cualesquiera dos semicircunferencias complementarias  $S_-^1(x)$  y  $S_+^1(x)$  que consideremos, por la simetría,  $\eta$  y  $\zeta$  seguirán la misma distribución y la probabilidad de que  $\xi$  esté en cualquiera de las semicircunferencias  $S_-^1(x)$  y  $S_+^1(x)$  es la misma,  $q = 1/2$ .

*Ejemplo 5.2.4.* Sea  $\xi$  un vector aleatorio en  $\mathbb{R}^2$  tal que  $P_\xi(\{x\}) = 0$  y las proyecciones de  $\xi$  sobre la circunferencia  $S^1(x)$  están uniformemente distribuidas sobre cierta semicircunferencia con probabilidad  $q$  y uniformemente distribuidas sobre la semicircunferencia opuesta con probabilidad  $1 - q$ .

Un ejemplo práctico de tal situación se obtiene con un vector aleatorio  $\rho$  que sigue distribución normal  $N_2(0, I)$  donde  $I$  es la matriz identidad y otra variable aleatoria  $\tau$  independiente de  $\rho$  y que satisface,  $\Pr(\tau = -1) = q$ ,  $\Pr(\tau = 1) = 1 - q$ . Si  $u \in \mathbb{R}^d \setminus \{0\}$ , el vector aleatorio  $\rho \text{signo}(\langle \rho, u \rangle) \cdot \tau$ , donde  $\text{signo}(x) = x/|x|$  si  $x \neq 0$  y  $\text{signo}(0) = 0$ , sigue una distribución como la del vector aleatorio  $\xi$  que acabamos de describir.

Las variables aleatorias  $\eta$  y  $\zeta$  seguirán una distribución uniforme en  $(0, \pi)$  y  $\Pr(\eta_{i:i} < \zeta_{1:n-i}) = \binom{n}{i}^{-1}$ , con lo que obtenemos

$$\Pr(\text{grado}(x; P_\xi) \leq n) = 1 - q^n - (1 - q)^n - 2 \sum_{i=1}^{n-1} q^i (1 - q)^{n-i}, \quad (5.2)$$

$$\text{ED}(x; P_\xi) = (\mathbb{E}\text{grado}(x; P_\xi))^{-1} = \frac{q(1 - q)}{1 + q - q^2}. \quad (5.3)$$

Además, podemos obtener una fórmula muy simple que evita muchos cálculos al calcular la profundidad de conteo con respecto a cierto  $p \in (0, 1)$ .

**Lema 5.2.7.** *Dado un vector aleatorio  $\xi$  en las condiciones anteriores, se cumple  $\Pr(\text{grado}(x; P_\xi) \leq 2) = 0$  y para  $n \geq 3$ ,*

$$\Pr(\text{grado}(x; P_\xi) \leq n) = \frac{1}{2} \left( \Pr(\text{grado}(x; P_\xi) \leq n-1) + 1 - q^{n-1} - (1 - q)^{n-1} \right).$$

*Demostración.* Denotamos por  $p_n$  a la probabilidad de que el grado de un punto  $x$  respecto de una distribución de probabilidad  $P_\xi$  sea menor o igual que  $n$ , esto es  $p_n = \Pr(\text{grado}(x; P_\xi) \leq n)$ . Según la expresión (5.2) se cumple,

$$\begin{aligned} p_n &= 1 - q^n - (1 - q)^n - 2 \sum_{i=1}^{n-1} q^i (1 - q)^{n-i} \\ &= 1 - q^n + (1 - q)^n - 2 \sum_{i=0}^{n-1} q^i (1 - q)^{n-i}. \end{aligned}$$

Podemos desarrollar el sumatorio de la fórmula anterior para que aparezca  $p_{n-1}$  en su expresión, así

$$\begin{aligned} 2 \sum_{i=0}^{n-1} q^i (1 - q)^{n-i} &= (1 - q) 2 \sum_{i=0}^{n-1} q^i (1 - q)^{(n-1)-i} \\ &= (1 - q) \left( 2(1 - q)^{n-1} + 2q^{n-1} + 2 \sum_{i=1}^{(n-1)-1} q^i (1 - q)^{(n-1)-i} \right) \\ &= (1 - q) (1 - p_{n-1} + (1 - q)^{n-1} + q^{n-1}). \end{aligned}$$

Como  $q$  y  $1 - q$  juegan el mismo papel, podemos intercambiarlas, así obtenemos una fórmula que se deriva de las expresiones anteriores para  $p_n$  y luego la misma fórmula intercambiando  $q$  y  $1 - q$ ,

$$\begin{aligned} p_n &= 1 - q^n + (1 - q)^n - (1 - q)(1 - p_{n-1} + (1 - q)^{n-1} + q^{n-1}), \\ p_n &= 1 - (1 - q)^n + q^n - q(1 - p_{n-1} + q^{n-1} + (1 - q)^{n-1}). \end{aligned}$$

Sumando estas dos expresiones obtenemos

$$2p_n = 1 + p_{n-1} - (1 - q)^{n-1} - q^{n-1},$$

lo que finaliza la demostración.  $\square$

### 5.3. Esperanza de la envolvente convexa

Presentamos una nueva familia de regiones centrales, junto con su función de profundidad estadística asociada. La relación de contenido de estas regiones centrales induce un orden estocástico según la variabilidad más débil que el orden lineal convexo.

Dado un vector aleatorio  $\xi$  con primer momento finito, consideramos la sucesión de conjuntos aleatorios  $\{\text{co}\{\xi_1, \dots, \xi_n\}\}_{n \in \mathbb{N}}$ , donde los vectores aleatorios  $\xi_i$  son copias independientes de  $\xi$ .

Es importante observar que a lo largo de la presente sección se supondrá que todos los vectores aleatorios con los que trabajemos tienen primer momento finito.

*Definición 5.3.1.* La *profundidad de la envolvente convexa* de un punto con respecto a un vector aleatorio con primer momento finito es el mínimo número de copias independientes de sí mismo necesarias para que el punto pertenezca a la esperanza de Aumann de su envolvente convexa. Dado  $x \in \mathbb{R}^d$ , tenemos

$$D_{\text{ch}}(x; P_\xi) := \left( \min\{n \in \mathbb{N} : x \in \mathbb{E}\text{co}\{\xi_1 \dots \xi_n\}\} \right)^{-1}.$$

De nuevo tomamos  $(+\infty)^{-1} = 0$ .

Las *regiones centrales de la envolvente convexa* se definen del modo habitual, dado  $\alpha \in (0, 1]$ ,

$$D_{\text{ch}}^{\alpha}(P_{\xi}) := \{x \in \mathbb{R}^d : D_{\text{ch}}(x; P_{\xi}) \geq \alpha\}. \quad (5.4)$$

Si denotamos por  $\lfloor a \rfloor$  a la parte entera de  $a$ , entonces a partir de la Definición 5.3.1 y de la fórmula (5.4) es claro que para cualquier  $\alpha \in (0, 1]$ , se cumple

$$D_{\text{ch}}^{\alpha}(P_{\xi}) = \mathbb{E}\text{co}\{\xi_1, \dots, \xi_{\lfloor \alpha^{-1} \rfloor}\}.$$

### 5.3.1. Propiedades

Veamos que la profundidad de la envolvente convexa es una función de profundidad estadística.

**Teorema 5.3.1.** *La familia de regiones centrales inducida por la profundidad de la envolvente convexa satisface las siguientes propiedades:*

- i.  $D_{\text{ch}}^1(P_{\xi}) = \{\mathbb{E}\xi\}$ ;
- ii.  $D_{\text{ch}}^{\alpha}(P_{\xi}) \subset D_{\text{ch}}^{\beta}(P_{\xi})$  para todo  $\beta \leq \alpha$ ;
- iii.  $D_{\text{ch}}^{\alpha}(P_{A\xi+b}) = AD_{\text{ch}}^{\alpha}(P_{\xi}) + b$ ;
- iv.  $D_{\text{ch}}^{\alpha}(P_{\xi})$  es compacto y convexo.

De las propiedades de las regiones centrales deducimos las propiedades de la función de profundidad.

**Corolario 5.3.2.** *Sea  $\xi$  un vector aleatorio en  $\mathbb{R}^d$  con primer momento finito, la profundidad de la envolvente convexa satisface las siguientes propiedades:*

- i.  $D_{\text{ch}}(Ax + b; P_{A\xi+b}) \geq D_{\text{ch}}(x; P_{\xi})$  y si  $A$  es una matriz no singular, se satisface la igualdad  $D_{\text{ch}}(Ax + b; P_{A\xi+b}) = D_{\text{ch}}(x; P_{\xi})$ ;
- ii.  $D_{\text{ch}}(\mathbb{E}\xi; P_{\xi}) = \sup_{x \in \mathbb{R}^d} D_{\text{ch}}(x; P_{\xi})$ ;

iii.  $D_{\text{ch}}(x; P_{\xi}) \leq D_{\text{ch}}(\mathbb{E}\xi + \alpha(x - \mathbb{E}\xi); P_{\xi})$ , para todo  $\alpha \in [0, 1]$ ;

iv.  $D_{\text{ch}}(x; P_{\xi}) \longrightarrow 0$  cuando  $\|x\| \rightarrow +\infty$ .

**Corolario 5.3.3.** *La profundidad de la envolvente convexa es una función de profundidad estadística en el sentido de la Definición 0.4.1.*

Veamos el tipo de regiones centrales que obtenemos para esta nueva profundidad.

*Ejemplo 5.3.1.* Si tomamos la distribución de probabilidad  $P$  que asigna probabilidad  $1/4$  a cada uno de los puntos  $(1, 1)$ ,  $(1, -1)$ ,  $(-1, 1)$  y  $(-1, -1)$ , obtenemos las regiones centrales que aparecen en la Figura 5.1 para la profundidad de la envolvente convexa.

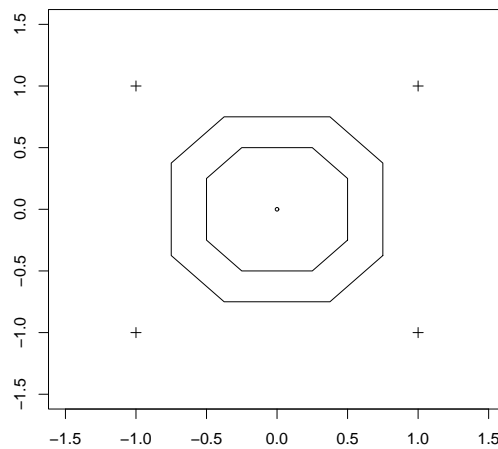


Figura 5.1: Contornos de  $D_{\text{ch}}^1(P)$ ,  $D_{\text{ch}}^{1/2}(P)$  y  $D_{\text{ch}}^{1/3}(P)$ .



### 5.3.2. Orden estocástico en variabilidad inducido

El punto más profundo con respecto a la profundidad de la envolvente convexa es la esperanza del vector aleatorio. Debería, por tanto, existir alguna relación con otra función de profundidad que satisface la misma propiedad, la *profundidad del zonoide*. Tal relación se encuentra en el orden estocástico definido por la relación de contenido de las regiones centrales. La profundidad del zonoide induce el orden lineal convexo. A continuación definimos el orden de la envolvente convexa a partir de la relación de contenido de las regiones centrales de la profundidad de la envolvente convexa.

*Definición 5.3.2.* Definimos el *orden de la envolvente convexa* como

$$\xi \leq_{\text{ch}} \eta \quad \text{si} \quad D_{\text{ch}}^{\alpha}(P_{\xi}) \subset D_{\text{ch}}^{\alpha}(P_{\eta}) \quad \text{para todo } \alpha \in [0, 1].$$

*Nota 5.3.1.* El orden de la envolvente convexa puede caracterizarse de un modo más simple como  $\xi \leq_{\text{ch}} \eta$  si y sólo si  $\mathbb{E}\text{co}\{\xi_1, \dots, \xi_n\} \subset \mathbb{E}\text{co}\{\eta_1, \dots, \eta_n\}$  para todo  $n \in \mathbb{N}$ .

El orden de la envolvente convexa es un orden estocástico en el sentido de que satisface las propiedades de reflexividad, transitividad y antisimetría.

**Proposición 5.3.4.** *Dados  $\xi, \eta$  dos vectores aleatorios en  $\mathbb{R}^d$ , se tienen las siguientes propiedades:*

- i. Reflexividad,  $\xi \leq_{\text{ch}} \xi$ ;*
- ii. Transitividad, si  $\xi \leq_{\text{ch}} \eta$  y  $\eta \leq_{\text{ch}} \zeta$ , entonces  $\xi \leq_{\text{ch}} \zeta$ ;*
- iii. Antisimetría, si  $\xi \leq_{\text{ch}} \eta$  y  $\eta \leq_{\text{ch}} \xi$ , entonces  $P_{\xi} = P_{\eta}$ .*

*Demostración.* Como el orden de la envolvente convexa se define a través de inclusiones de conjuntos y la inclusión de conjuntos es reflexiva y transitiva, se verifican los dos primeros enunciados. La antisimetría se satisface por el hecho de que la sucesión de esperanzas de Aumann de la envolvente convexa de conjuntos formados por una sucesión de copias independientes de un vector aleatorio con esperanza finita caracteriza su distribución, como demostró Vitale (1987).  $\square$

### Propiedades del orden estocástico

En primer lugar daremos unas propiedades generales del orden estocástico de la envolvente convexa que enunciamos sin demostración.

**Proposición 5.3.5.** *El orden de la envolvente convexa verifica las siguientes propiedades:*

- i. si  $\xi \leq_{\text{ch}} \eta$ , entonces  $\mathbb{E}\xi = \mathbb{E}\eta$ ;
- ii. si  $\mathbb{E}\xi = 0$ , entonces  $\xi \leq_{\text{ch}} a\xi$  para todo  $a \geq 1$ ;
- iii. si  $\xi \leq_{\text{ch}} \eta$ , entonces  $A\xi + b \leq_{\text{ch}} A\eta + b$  para todo  $A \in \mathbb{R}^{d \times d}$  y  $b \in \mathbb{R}^d$ ;
- iv. si  $\xi \leq_{\text{ch}} \eta$ , entonces  $\text{co}\{\text{supp } \xi\} \subset \text{co}\{\text{supp } \eta\}$ .

Donde  $\text{supp } \xi$  denota al soporte de la variable aleatoria  $\xi$ .

El orden de la envolvente convexa para variables aleatorias puede caracterizarse a partir de desigualdades en las esperanzas de los estadísticos ordenados extremos.

**Lema 5.3.6.** *Dadas  $\xi, \eta$  dos variables aleatorias,  $\xi \leq_{\text{ch}} \eta$  si y sólo si  $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$  y  $\mathbb{E}\xi_{1:n} \geq \mathbb{E}\eta_{1:n}$  para todo  $n \in \mathbb{N}$ .*

Podemos demostrar que se trata de un orden estocástico lineal.

**Lema 5.3.7.** *Dados  $\xi, \eta$  dos vectores aleatorios en  $\mathbb{R}^d$*

$$\xi \leq_{\text{ch}} \eta \text{ si y sólo si } \langle \xi, u \rangle \leq_{\text{ch}} \langle \eta, u \rangle, \text{ para todo } u \in \mathbb{R}^d.$$

*Demostración.* Por definición se tiene que  $\xi \leq_{\text{ch}} \eta$  si y sólo si  $\mathbb{E}\text{co}\{\xi_1, \dots, \xi_n\} \subset \mathbb{E}\text{co}\{\eta_1, \dots, \eta_n\}$  para todo  $n \in \mathbb{N}$  lo que, por la relación entre la esperanza de Aumann y la función soporte, es equivalente a

$$\mathbb{E}\text{máx}\{\langle \xi_1, u \rangle, \dots, \langle \xi_n, u \rangle\} \leq \mathbb{E}\text{máx}\{\langle \eta_1, u \rangle, \dots, \langle \eta_n, u \rangle\}$$

para todo  $n \in \mathbb{N}$  y  $u \in \mathbb{R}^d$ . Si conjuntamente con cualquier  $u \in \mathbb{R}^d$  tomamos su opuesto,  $-u$ , la relación entre valores máximos se convierte en una relación entre valores mínimos y obtenemos la equivalencia con  $\langle \xi, u \rangle \leq_{\text{ch}} \langle \eta, u \rangle$  para todo  $u \in \mathbb{R}^d$ .  $\square$

Hoeffding (1953) prueba que la sucesión de esperanzas de los máximos de  $n$  variables aleatorias independientes e idénticamente distribuidas,  $\{\mathbb{E}\xi_{n:n}\}_{n \in \mathbb{N}}$  caracteriza su distribución. Dadas dos variables aleatorias  $\xi$  y  $\eta$ , las condiciones  $\mathbb{E}\xi = \mathbb{E}\eta$  y  $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$  para todo  $n \in \mathbb{N}$  no son suficientes para tener la ordenación  $\xi \leq_{\text{ch}} \eta$ .

*Ejemplo 5.3.2.* Dada  $\zeta$  una variable aleatoria que sigue una distribución exponencial con media 1, sean  $\xi = 1 - \zeta$  y  $\eta = \zeta - 1$ . Es fácil comprobar que

$$\begin{aligned} \mathbb{E}\xi_{n:n} = 1 - \mathbb{E}\zeta_{1:n} = 1 - \frac{1}{n} \quad , \quad \mathbb{E}\xi_{1:n} = 1 - \mathbb{E}\zeta_{n:n} = - \sum_{i=2}^n \frac{1}{i}; \\ \mathbb{E}\eta_{n:n} = \mathbb{E}\zeta_{n:n} - 1 = \sum_{i=2}^n \frac{1}{i} \quad , \quad \mathbb{E}\eta_{1:n} = \mathbb{E}\zeta_{n:n} - 1 = \frac{1}{n} - 1. \end{aligned}$$

Entonces claramente  $\mathbb{E}\xi = \mathbb{E}\eta = 0$  y  $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$  para todo  $n \in \mathbb{N}$ , pero  $\mathbb{E}\xi_{1:3} = -5/6 < -2/3 = \mathbb{E}\eta_{1:3}$  y es falso que  $\xi \leq_{\text{ch}} \eta$ . Se puede probar que tampoco se cumple  $\eta \leq_{\text{ch}} \xi$ .

El orden de la envolvente convexa es un orden según la variabilidad estrictamente más débil que el orden lineal convexo. Este resultado se probará en el Teorema 5.3.11 y para ello necesitaremos algunos lemas previos.

**Lema 5.3.8.** Sean  $\xi_1, \dots, \xi_n$  y  $\eta_1, \dots, \eta_n$  dos conjuntos de variables aleatorias independientes y  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  creciente convexa, entonces

$$\text{si } \xi_i \leq_{\text{icx}} \eta_i \text{ para todo } i, \text{ se cumple } f(\xi_1, \dots, \xi_n) \leq_{\text{icx}} f(\eta_1, \dots, \eta_n).$$

*Demostración.* Sean  $\xi_i, \eta_i$  y  $f$  cumpliendo las condiciones impuestas en el enunciado del resultado. Sea  $g : \mathbb{R} \rightarrow \mathbb{R}$  una función creciente y convexa, entonces  $g \circ f : \mathbb{R}^n \rightarrow \mathbb{R}$  es creciente convexa.

Como para todo  $i$ , se cumple  $\xi_i \leq_{\text{icx}} \eta_i$  y tanto las variables  $\xi_1, \xi_2, \dots, \xi_n$  como las variables  $\eta_1, \eta_2, \dots, \eta_n$  son independientes, por Teorema 0.5.7, se tiene que  $(\xi_1, \dots, \xi_n) \leq_{\text{icx}} (\eta_1, \dots, \eta_n)$ . Por lo tanto  $\mathbb{E}(g \circ f)(\xi_1, \dots, \xi_n) \leq \mathbb{E}(g \circ f)(\eta_1, \dots, \eta_n)$  y finalmente  $f(\xi_1, \dots, \xi_n) \leq_{\text{icx}} f(\eta_1, \dots, \eta_n)$ .  $\square$

**Lema 5.3.9.** Si dos variables aleatorias  $\xi, \eta$  satisfacen  $\xi \leq_{\text{icx}} \eta$ , entonces  $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$  para todo  $n \in \mathbb{N}$ .

*Demostración.* Sean  $\xi_1, \dots, \xi_n$  y  $\eta_1, \dots, \eta_n$  copias independientes de  $\xi$  y  $\eta$  respectivamente. Como  $\text{máx} : \mathbb{R}^n \rightarrow \mathbb{R}$  es creciente y convexa, según el Lema 5.3.8  $\text{máx}\{\xi_1, \dots, \xi_n\} \leq_{\text{icx}} \text{máx}\{\eta_1, \dots, \eta_n\}$  y finalmente  $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$  para todo  $n \in \mathbb{N}$ .  $\square$

**Lema 5.3.10.** Si dos variables aleatorias  $\xi, \eta$  satisfacen  $\xi \leq_{\text{cx}} \eta$ , entonces  $\xi \leq_{\text{ch}} \eta$ .

*Demostración.* Sea  $\xi \leq_{\text{cx}} \eta$ , entonces  $\xi \leq_{\text{icx}} \eta$  y  $-\xi \leq_{\text{icx}} -\eta$ . Según el Teorema 5.3.9  $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$  y  $\mathbb{E}(-\xi)_{n:n} \leq \mathbb{E}(-\eta)_{n:n}$  o equivalentemente  $\mathbb{E}\xi_{1:n} \geq \mathbb{E}\eta_{1:n}$  para todo  $n \in \mathbb{N}$  lo que implica  $\xi \leq_{\text{ch}} \eta$ .  $\square$

**Teorema 5.3.11.** Si dos vectores aleatorios  $\xi, \eta$  satisfacen  $\xi \leq_{\text{lcx}} \eta$ , entonces  $\xi \leq_{\text{ch}} \eta$ .

*Demostración.* Dado  $\xi \leq_{\text{lcx}} \eta$ , se cumple  $\langle \xi, u \rangle \leq_{\text{cx}} \langle \eta, u \rangle$  para todo  $u \in \mathbb{R}^d$ , según el Teorema 5.3.10  $\langle \xi, u \rangle \leq_{\text{ch}} \langle \eta, u \rangle$  para todo  $u \in \mathbb{R}^d$ , lo que por el Lema 5.3.7 implica  $\xi \leq_{\text{ch}} \eta$ .  $\square$

Para finalizar, mostramos un ejemplo de dos variables aleatorias ordenadas respecto del orden de la envolvente convexa, pero que no están ordenados según el orden convexo.

*Ejemplo 5.3.3.* Sea  $\xi$  una variable aleatoria discreta que toma los valores  $-2/3$  y  $2/3$ , ambos con probabilidad  $1/2$  y sea  $\eta$  otra variable aleatoria uniformemente distribuida en el intervalo  $(-1, 1)$ .

Es fácil comprobar que  $\mathbb{E}\xi_{n:n} = (2^{n-1} - 1)/(3 \cdot 2^{n-2})$  y  $\mathbb{E}\eta_{n:n} = (n - 1)/(n + 1)$  y por lo tanto se cumple  $\mathbb{E}\xi_{n:n} \leq \mathbb{E}\eta_{n:n}$  para todo  $n$  y como  $\xi$  y  $\eta$  están simétricamente distribuidas en torno al 0 esto es suficiente para probar que  $\xi \leq_{\text{ch}} \eta$ .

Si tomamos la función convexa  $f(x) = x_+$ , se tiene que  $\mathbb{E}f(\xi) = 1/3 > 1/4 = \mathbb{E}f(\eta)$  y por lo tanto no se satisface  $\xi \leq_{\text{cx}} \eta$ .

## 5.4. Comentarios y problemas abiertos

A lo largo de este capítulo se han introducido, en primer lugar funciones de profundidad inspiradas en el número de observaciones independientes de un vector aleatorio necesarias para contener a un punto en su envolvente convexa y en segundo lugar una función de profundidad determinada por el número de realizaciones independientes de un vector aleatorio necesarias para que un punto esté contenido en la esperanza de Aumann del conjunto aleatorio formado por la envolvente convexa de esos vectores aleatorios. El estudio del primer tipo de funciones de profundidad nos ha servido para ahondar en la investigación sobre problemas de recubrimiento de esferas mediante semiesferas, que resolvemos en el caso general sobre el plano. El segundo tipo de funciones de profundidad nos ha llevado a un nuevo orden estocástico según la variabilidad, el de la envolvente convexa. Este orden es más débil que el orden lineal convexo, al igual que el orden simétrico que estudiábamos en el Capítulo 1. Una de las diferencias con el orden simétrico es que, para que dos vectores aleatorios puedan estar ordenados según el orden de la envolvente convexa, es necesario que tengan la misma esperanza

Los problemas que dejamos abiertos en este tema y estudiaremos más adelante son los siguientes:

Sería interesante poder describir cómo se obtiene la profundidad del grado medio y de conteo de un punto con respecto a una distribución en  $\mathbb{R}^d$  cuando  $d \geq 3$ , problema que involucraría resolver completamente el del recubrimiento de la esfera con semiesferas.

Pretendemos obtener el comportamiento asintótico de las funciones de profundidad introducidas, tanto la de la esperanza de la envolvente convexa como la del grado medio y la de conteo y estudiar bajo qué condiciones las profundidades con respecto a una distribución de probabilidad empírica convergen a las profundidades con respecto a la distribución poblacional correspondiente.

Finalmente, continuaremos el estudio del orden estocástico de la envolvente convexa. Existe otro modo alternativo, distinto al seguido en esta

memoria, para demostrar que este orden, aplicado sobre variables aleatorias, es más débil que el orden convexo. Basta con utilizar los funcionales propuestos por Ramos y Sordo (2003) y resultados clásicos sobre estadísticos ordenados que se encuentran en Galambos y Kotz (1978), David (1981) o Arnold *et al.* (1992). Los funcionales propuestos por Ramos y Sordo (2003) para una variable aleatoria  $\xi$  son de la forma

$$I_\omega(\xi) = \int_0^1 w(p)(F_\xi^{-1}(p) - \mathbb{E}\xi) dp, \quad (5.5)$$

donde  $F_\xi^{-1}$  es la función cuantil de  $\xi$ , es decir  $F_\xi^{-1} := \inf\{x : F_\xi(x) \geq t\}$  para  $0 \leq t \leq 1$  y las funciones  $\omega : [0, 1] \rightarrow \mathbb{R}$  pertenecen a cierta familia. Si para dos variables aleatorias  $\xi$  y  $\eta$ , se tiene que  $I_\omega(\xi) \leq I_\omega(\eta)$  para todas las funciones  $\omega$  de cierta familia, se define una ordenación estocástica. La relación de estos índices con el orden de la envolvente convexa se basa en que si  $\omega$  es la función  $\omega(x) = nx^{n-1}$ , entonces el índice que induce es la esperanza del máximo de  $n$  copias independientes de la variable aleatoria  $\xi$ , mientras que  $\omega(x) = n(1-x)^{n-1}$  induce como índice la esperanza del mínimo de  $n$  copias independientes de la variable aleatoria  $\xi$ .

Continuando con el estudio del orden estocástico de la envolvente convexa, sería muy interesante saber si se trata de un orden estocástico integral y podría generalizarse el problema para estudiar qué órdenes estocásticos integrales pueden describirse a partir de desigualdades de los funcionales de Ramos y Sordo (2003) y qué órdenes del tipo de los de estos autores son órdenes estocásticos integrales.

# Epílogo

Comenzaré reflejando mis sensaciones sobre la investigación que he llevado a cabo. A continuación resumiré las aportaciones que considero más interesantes dentro de las que se han hecho en esta memoria y comentaré el trabajo que tengo previsto acometer con inmediatez.

Los dos grandes temas sobre los que gira esta memoria son los Órdenes Estocásticos y las Funciones de Profundidad. Estas dos áreas de estudio reciben en la actualidad gran atención por parte de la comunidad estadística y probabilista. Los órdenes estocásticos han adquirido gran importancia gracias a sus numerosas aplicaciones tanto prácticas en campos como la Economía, las Ciencias Actuariales o la Epidemiología, como teóricas en la Teoría de la Probabilidad y en la Teoría de la Aproximación para obtener teoremas límite. En cuanto a las funciones de profundidad, su estudio está enmarcado en el Análisis Multivariante no Paramétrico, área en la que se están haciendo grandes progresos. A título personal, creo que de esta memoria deben destacarse las relaciones que se establecen entre estos dos temas y la utilización de los conjuntos aleatorios, o más específicamente, la esperanza de Aumann como una herramienta que nos permite avanzar en ambos. Los resultados sobre ordenaciones estocásticas no tienen visos de aplicación inmediata, pero constituyen un estudio de base que puede ser muy útil en el futuro.

Resumidas en tres puntos, las aportaciones más novedosas que se presentan en esta memoria son:

- Las comparaciones según la variabilidad para algunos tipos de elementos aleatorios. En concreto, vectores aleatorios con distinta esperanza a

través del orden simétrico y conjuntos aleatorios compactos a través del orden simétrico, del lineal convexo y de los órdenes lineales crecientes convexos y de los índices de desigualdad para conjuntos aleatorios.

- Un nuevo marco de trabajo para la obtención de funciones de profundidad y regiones centrales a partir de los recortes de una probabilidad. En su contexto se construyen las regiones centrales recortadas y las integrales y se obtienen varias de sus propiedades. Estas familias de regiones centrales engloban otras regiones centrales clásicas facilitando el estudio de sus propiedades.
- Gracias al estudio de las funciones de profundidad basadas en un número de observaciones de un vector aleatorio, se explica cómo calcular en el caso general la distribución del número de semicircunferencias necesarias para recubrir una circunferencia y además se obtiene un nuevo orden estocástico según la variabilidad que permite comparar algunos pares de vectores aleatorios no comparables según el orden lineal convexo. Además, el orden de la envolvente convexa nos sirve para insistir en la relación de las regiones centrales con las ordenaciones estocásticas.

Los proyectos en los que estoy trabajando y están más avanzados son, por un lado, el estudio de las funciones de profundidad propuestas en el Capítulo 5 que tengo pendiente finalizar con el Profesor Molchanov y por otro lado el estudio de diferentes grados de afinidad de distribuciones de probabilidad con respecto a un funcional de capacidad con el Profesor López Díaz, codirector de mi Tesis Doctoral. Este tipo de funcionales se han descrito en la Sección 4.5.



# Bibliografía

- Arnold, B.C., Balakrishnan, N. y Nagaraja, H.N. (1992). *A first course in order statistics*. Series in Probability and Mathematical Statistics. John Wiley and sons, New York.
- Atkinson, A. (1983). *The Economics of Inequality*. Clarendon Press, Oxford.
- Aumann, R. (1965). Integrals of set-valued functions. *Journal of Mathematical Analysis and its Applications*, **12**:1–12.
- Barnett, V. (1976). The ordering of multivariate data (with discussion). *Journal of the Royal Statistical Society. Series A. General*, **139**(3):318–354.
- Bhattacharjee, M. y Sethuraman, J. (1990). Families of life distributions characterized by two moments. *Journal of Applied Probability*, **27**(3):720–725.
- Billingsley, P. (1986). *Probability and Measure*. Segunda edición, Series in Probability and Mathematical Statistics. John Wiley and sons, New York.
- Billingsley, P. (1999). *Convergence of probability measures*. Segunda edición, Series in Probability and Mathematical Statistics. John Wiley and sons, New York.
- Bolker, E. (1969). A class of convex bodies. *Transactions of the American Mathematical Society*, **145**:323–345.

- Bolker, E. (1971). The zonoid problem. *American Mathematical Monthly*, **78**:529–531.
- Bourguignon, F. (1979). Decomposable income inequality measures. *Econometrica*, **47**(4):901–920.
- Boutsikas, M. y Vaggelatou, E. (2002). On the distance between convex-ordered random variables, with applications. *Advances in Applied Probability*, **34**(2):349–374.
- Breiman, L. (1968). *Probability*. Addison-Wesley Publishing Company, Reading, Mass.-London-Don Mills, Ont.
- Cascos, I. (2001). Resultados asintóticos sobre índices de desigualdad de conjuntos aleatorios. Trabajo de Investigación de Tercer Ciclo, Universidad de Oviedo.
- Cascos, I. (2003). An ideal probability metric based on the symmetric order. *Actas del XXIII Seminar on Stability Problems for Stochastic Models*.
- Cascos, I. y López-Díaz, M. (2003). Diferentes modelos de índices de desigualdad para intervalos aleatorios. *Actas del 27 Congreso Nacional de Estadística e Investigación Operativa*, 1362–1369.
- Cascos, I. y López-Díaz, M. (2004a). A new family of central regions, location trimmed regions, *enviado para su publicación*.
- Cascos, I. y López-Díaz, M. (2004b). Integral trimmed regions, *enviado para su publicación*.
- Cascos, I., López-Díaz, M. y Gil, M.A. (2002). On the variation of the  $f$ -inequality of a random variable. En *Soft methods in probability, statistics and data analysis* (Grzegorzewski, P., Hryniewicz, O., Gil, M.A. eds.) 98–104. *Advances in soft computing*. Physica-Verlag, Heidelberg.

- Cascos, I., López-Díaz, M., y Gil, M.A. (2004). Convergence criteria for interval-valued inequality indices. *Statistics*, **38**(1):59–66.
- Cascos, I. y Molchanov, I. (2003). A stochastic order for random vectors and random sets based on the Aumann expectation. *Statistics and Probability Letters*, **63**(3):295–305.
- Chen, Z. y Tyler, D. (2002). The influence function and maximum bias of Tukey's median. *Annals of Statistics*, **30**(6):1737–1759.
- Chiu, S. y Molchanov, I. (2003). A new graph related to directional characteristics of point processes. *Advances in Applied Probability*, **35**(1):47–55.
- Csiszár, I. (1967). Information-type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica*, **2**:299–318.
- Dall'Aglio, M. y Scarsini, M. (2003). Zonoids, linear dependence, and size-biased distributions on the simplex. *Advances in Applied Probability*, **35**:871–884.
- David, H.A. (1981). *Order Statistics*. Segunda edición, Series in Probability and Mathematical Statistics. John Wiley and sons, New York.
- Debreu, G. (1967). Integration of correspondences. *Proceedings of the Fifth Berkeley Symposium in Mathematical Statistics and Probability*, **2**:351–372.
- Denneberg, D. (1994). *Non-Additive Measure and Integral*. Kluwer Academic Publishers.
- Dümbgen, L. y Walther, G. (1996). Rates of convergence for random approximations of convex sets. *Advances in Applied Probability*, **28**(2):384–393.
- Eddy, W. (1982). Convex hull peeling. En *COMPSTAT (IASC)* 42–47. Physica-Verlag, Vienna.

- Eddy, W. (1983). Set-valued orderings for bivariate data. En *Stochastic Geometry, Geometric Statistics and Stereology* (Ambartsumian, R., Weil, W. eds.) 79–90. Teubner, Leipzig.
- Eddy, W. (1985). Ordering of multivariate data. En *Computer Science and Statistics* (Billard, L. ed.) 25–30. The Interface.
- Eichhorn, W. y Gehrig, W. (1982). Measurement of inequality in economics. En *Applied Mathematics-Optimization and Operations Research* (Korte, B. ed.) 657–693. North-Holland, Amsterdam.
- Galambos, J. y Kotz, S. (1978). *Characterizations of Probability Distributions*. Lecture Notes in Mathematics **675**. Springer-Verlag, Berlin.
- Goodey, P. y Weil, W. (1993). Zonoids and generalisations. En *Handbook of Convex Geometry*, 1297–1326. Amsterdam: North-Holland.
- Hall, P. (1988). *Introduction to the Theory of Coverage Processes*. John Wiley, New York.
- Hiai, F. y Umegaki, H. (1977). Integrals, conditional expectations, and martingales of multivalued functions. *Journal of Multivariate Analysis*, **7**(1):149–182.
- Hoeffding, W. (1953). On the distribution of the expected values of the order statistics. *Annals of Mathematical Statistics*, **24**:93–100.
- Koltchinskii, V. (1997). M-estimation, convexity and quantiles. *Annals of Statistics*, **25**(2):435–477.
- Koshevoy, G. y Mosler, K. (1996). The Lorenz zonoid of a multivariate distribution. *Journal of the American Statistical Association*, **91**(434):873–882.
- Koshevoy, G. y Mosler, K. (1997a). Multivariate Gini indices. *Journal of Multivariate Analysis*, **60**(2):252–276.

- Koshevoy, G. y Mosler, K. (1997b). Zonoid trimming for multivariate distributions. *Annals of Statistics*, **25**(5):1998–2017.
- Koshevoy, G. y Mosler, K. (1998). Lift zonoids, random convex hulls and the variability of random vectors. *Bernoulli*, **4**(3):377–399.
- Liu, R. (1990). On a notion of data depth based on random simplices. *Annals of Statistics*, **18**(1):405–414.
- Liu, R., Parelius, J., y Singh, K. (1999). Multivariate analysis by data depth: descriptive statistics, graphics and inference. *Annals of Statistics*, **27**(3):783–840.
- Liu, R. y Singh, K. (1992). Ordering directional data: concepts of data depth on circles and spheres. *Annals of Statistics*, **20**(3):1468–1484.
- López-García, H., López-Díaz, M. y Gil, M. (2000). Interval-valued quantification of the inequality associated with a random set. *Statistics and Probability Letters*, **46**(2):149–159.
- Massé, J. y Theodorescu, R. (1994). Halfplane trimming for bivariate distributions. *Journal of Multivariate Analysis*, **48**(2):188–202.
- Matheron, G. (1975). *Random Sets and Integral Geometry*. John Wiley and sons, New York.
- Molchanov, I. (1993). *Limit Theorems for Unions of Random Closed Sets*. Lecture Notes in Mathematics **1561**. Springer-Verlag, Berlin.
- Molchanov, I. (1998). Averaging of random sets and binary images. *CWI Quarterly*, **11**:371–384.
- Molchanov, I. (1999). Random sets in view of image filtering applications. En *Nonlinear Filters for Image Processing* (Dougherty E.R., Astola J. eds.), capítulo 10, 419–447. SPIE.

- Mosler, K. (2002). *Multivariate dispersion, central regions and depth. The lift zonoid approach*. Lecture Notes in Statistics **165**. Springer-Verlag, Berlin.
- Müller, A. (1997a). Integral probability metrics and their generating classes of functions. *Advances in Applied Probability*, **29**(2):429–443.
- Müller, A. (1997b). Stochastic orders generated by integrals: a unified study. *Advances in Applied Probability*, **29**(2):414–428.
- Müller, A. y Stoyan, D. (2002). *Comparison Methods for Stochastic Models and Risks*. Wiley Series in Probability and Statistics. John Wiley and sons, Ltd.
- Rachev, S. (1991). *Probability Metrics and the Stability of Stochastic Models*. John Wiley and sons, Ltd., Chichester.
- Ramos, H.M. y Sordo, M.A. (2003). Dispersion measures and dispersive orderings. *Statistics and Probability Letters*, **61**(2):123–131.
- Rockafellar, R. (1970). *Convex Analysis*. Princeton University Press.
- Romanazzi, M. (2001). Influence function of the halfspace depth. *Journal of Multivariate Analysis*, **77**(1):138–161.
- Rudin, W. (1987). *Real and complex analysis*. Tercera edición, McGraw-Hill Book Co., New York.
- Scarsini, M. (1998). Multivariate convex orderings, dependence, and stochastic equality. *Journal of Applied Probability*, **35**(1):93–103.
- Schneider, R. (1993). *Convex bodies: the Brunn-Minkowski theory*. *Encyclopedia of Mathematics and its Applications* **44**. Cambridge University Press, Cambridge.
- Sen, A. (1995). *Inequality Reexamined*. Harvard University Press, Cambridge.

- Shaked, M. y Shanthikumar, J. (1994). *Stochastic Orders and Their Applications*. Academic Press, Boston.
- Small, C. (1990). A survey of multidimensional medians. *International Statistical Review*, **58**(3):263–277.
- Stoyan, D. (1998). Random sets: Models and statistics. *International Statistical Review*, **66**:1–27.
- Stoyan, D., Kendall, W. y Mecke, J. (1995). *Stochastic Geometry and its Applications*. John Wiley and sons, Chichester.
- Stoyan, H. y Stoyan, D. (1980). On some partial orderings of random closed sets. *Mathematische Operationsforschung und Statistik Series Optimization*, **11**(1):145–154.
- Talagrand, M. (1987). The Glivenko-Cantelli problem. *Annals of Probability*, **15**(3):837–870.
- Talagrand, M. (1996). The Glivenko-Cantelli problem, ten years later. *Journal of Theoretical Probability*, **9**(2):371–384.
- Tukey, J. (1975). Mathematics and the picturing of data. En *Proceedings of the International Congress of Mathematicians*, Vancouver, **2**:523–531.
- Vitale, R. (1987). Expected convex hulls, order statistics, and Banach space probabilities. *Acta Applicandae Mathematicae*, **9**:97–102.
- Wendel, J.G. (1962). A problem in geometric probability. *Mathematica Scandinavica*, **11**:109–111.
- Zolotarev, V. (1983). Probability metrics. *Theory of Probability and its Applications*, **28**:278–302.
- Zuo, Y. y Serfling, R. (2000a). General notions of statistical depth functions. *Annals of Statistics*, **28**(2):461–482.

- Zuo, Y. y Serfling, R. (2000b). On the performance of some robust nonparametric location measures relative to a general notion of multivariate symmetry. *Journal of Statistical Planning and Inference*, **84**:55–79.
- Zuo, Y. y Serfling, R. (2000c). Structural properties and convergence results for contours of sample statistical depth functions. *Annals of Statistics*, **28**(2):483–499.



# Índice de figuras

1.1.	La región de color gris claro es el zonoide elevado de una variable uniforme en $(-2, 4)$ y toda la región coloreada su envolvente de zonoides. . . . .	26
1.2.	La región pintada en gris claro es la envolvente de zonoides de una variable uniforme en $(-1, 1)$ , toda la región coloreada es la envolvente de zonoides de una variable uniforme en $(-3, 2)$ . . . . .	32
2.1.	Evolución de la glucosa en sangre a lo largo de una semana. Cada símbolo $\boxtimes$ indica el resultado de una medición del nivel de glucosa en sangre en el día correspondiente. . . . .	70
2.2.	Índices de desigualdad para $f_{-1}(x) = (1/x - 1)/2$ . . . . .	73
3.1.	Contornos de $ZD^1(P)$ , $ZD^{3/4}(P)$ y $ZD^{1/2}(P)$ . . . . .	95
3.2.	Contornos de $WD^1(P)$ , $WD^{3/4}(P)$ y $WD^{1/2}(P)$ . . . . .	98
3.3.	Contornos de $D_M^1(P)$ , $D_M^{3/4}(P)$ y $D_M^{1/2}(P)$ . . . . .	99
4.1.	Contornos de $HD^{4/10}(P)$ , $HD^{3/10}(P)$ y $HD^{2/10}(P)$ . . . . .	125
4.2.	Contornos de $ZD^1(P)$ , $ZD^{4/5}(P)$ , $ZD^{3/5}(P)$ , $ZD^{2/5}(P)$ y $ZD^{1/5}(P)$ . . . . .	127
4.3.	Contornos de $WD^1(P)$ , $WD^{4/5}(P)$ , $WD^{3/5}(P)$ , $WD^{2/5}(P)$ y $WD^{1/5}(P)$ . . . . .	130
5.1.	Contornos de $D_{ch}^1(P)$ , $D_{ch}^{1/2}(P)$ y $D_{ch}^{1/3}(P)$ . . . . .	150



# Lista de símbolos

$(\Omega, \mathcal{A}, \text{Pr})$	espacio de probabilidad
$:=$	igual por definición
$\mathbb{P}$	conjunto de probabilidades en $\mathbb{R}^d$
$\mathcal{B}_d$	$\sigma$ -álgebra de Borel asociada a $\mathbb{R}^d$
<i>c.s.</i> [Pr], <i>c.s.</i>	casi seguro con respecto a la probabilidad Pr
$\ \cdot\ $	norma euclídea
$\langle \cdot, \cdot \rangle$	producto escalar
$P, Q$	probabilidades de $\mathbb{P}$
$P_\xi$	probabilidad inducida por el vector aleatorio $\xi$
$P_u$	probabilidad definida en la página 10
$P_{A,b}$	probabilidad definida en la página 12
$\Rightarrow$	convergencia débil de probabilidades
id	función identidad
$\mathcal{K}$	familia de subconjuntos compactos no vacíos de $\mathbb{R}^d$
$\mathcal{K}_c$	familia de subconjuntos compactos y convexos de $\mathbb{R}^d$
$\mathcal{K}_c(\mathbb{R})$	familia de intervalos reales compactos
$\mathbb{R}_+^d$	primer cuadrante cerrado
$\mathbb{E}$	esperanza de un vector aleatorio o bien esperanza de Aumann de un conjunto aleatorio
$\mathcal{F}, \mathcal{G}$	familias de funciones de $\mathbb{R}^d$ en $\mathbb{R}$
$I, J$	conjuntos de índices

$B_r$	bola cerrada centrada en el origen y de radio $r$
$S^{d-1}$	esfera unidad en $\mathbb{R}^d$ centrada en el origen
$S^{d-1}(x)$	esfera unidad en $\mathbb{R}^d$ centrada en $x$
$A^r$	$r$ -entorno de $A$
$I_A$	función indicador del conjunto $A$
$i_f(\cdot)$	$f$ -índice de desigualdad para variables aleatorias
$i_{f,l}(\cdot)$	$(f, l)$ -índice de desigualdad para vectores aleatorios
$I_f(\cdot)$	$f$ -índice de desigualdad para intervalos aleatorios
$I_{f,l}(\cdot)$	$(f, l)$ -índice de desigualdad para conjuntos aleatorios,
$L^p(\text{Pr}), L^p$	espacio de vectores aleatorios en $\mathbb{R}^d$
$d^p(\cdot, \cdot)$	distancia asociada con $L^p(\text{Pr})$
$L_+^p$	espacio de vectores aleatorios en $(0, +\infty)^d$
$\mathcal{L}^p$	espacio de conjuntos aleatorios compactos y convexos en $\mathbb{R}^d$
$\mathcal{D}^p(\cdot, \cdot)$	distancia asociada con $\mathcal{L}^p$
$\mathcal{L}_+^p$	espacio de conjuntos aleatorios compactos y convexos en $\mathbb{R}^d$ , que toman como valores subconjuntos de $(0, +\infty)^d$
$\ \cdot\ _p$	norma asociada con $L^p(\text{Pr})$ o bien $\mathcal{L}^p$
$\ \cdot\ _{L^1(P)}$	norma asociada con $L^1(P)$
$O(\cdot)$	“ $O$ mayúscula” de Landau
$P^\alpha$	$\alpha$ -recorte de una probabilidad
$T_{A,b}$	transformación afín, $T_{A,b}(x) := Ax + b$
$Z(\cdot)$	zonoide de una probabilidad
$\tilde{Z}(\cdot)$	zonoide elevado de una probabilidad
$W(\cdot)$	envolvente de zonoides de una probabilidad
$\text{HD}(\cdot; \cdot)$	función de profundidad semiespacial o de Tukey
$\text{HD}^\alpha(\cdot)$	región central semiespacial de nivel $\alpha$
$\text{ZD}(\cdot; \cdot)$	función de profundidad del zonoide
$\text{ZD}^\alpha(\cdot)$	región central del zonoide de nivel $\alpha$
$\text{WD}^\alpha(\cdot)$	región central de la envolvente de zonoides de nivel $\alpha$

$D_{\mathcal{F}}(\cdot; \cdot)$	función de profundidad integral
$D_{\mathcal{F}}^{\alpha}(\cdot)$	región central integral de nivel $\alpha$
$D_{\{\mathcal{F}_i\}_{i \in I}}(\cdot; \cdot)$	función de profundidad integral
$D_{\{\mathcal{F}_i\}_{i \in I}}^{\alpha}(\cdot)$	región central integral de nivel $\alpha$
$D_{\mathcal{F}}^{\alpha, \varepsilon}(\cdot)$	región central $\varepsilon$ -integral de nivel $\alpha$
$D_L(\cdot; \cdot)$	función de profundidad respecto al estimador de localización L
$D_L^{\alpha}(\cdot)$	región central $\alpha$ -recortada
$SD(\cdot; \cdot)$	función de profundidad simplicial
$ED(\cdot; \cdot)$	función de profundidad del grado medio
$D_p(\cdot; \cdot)$	función de profundidad de conteo
$D_{\text{ch}}(\cdot; \cdot)$	función de profundidad de la envolvente convexa
$D_{\text{ch}}^{\alpha}(\cdot)$	región central de la envolvente convexa de nivel $\alpha$
$\leq_{\mathcal{F}}$	orden estocástico integral cuyos generadores son la familia $\mathcal{F}$
$\leq_{\text{st}}$	orden estocástico habitual para vectores aleatorios
$\leq_{\text{cx}}$	orden estocástico convexo para vectores aleatorios
$\leq_{\text{icx}}$	orden estocástico creciente convexo para vectores aleatorios
$\leq_{\text{lcx}}$	orden estocástico lineal convexo para vectores aleatorios
$\leq_{\text{iplcx}}$	orden estocástico creciente positivo lineal convexo para vectores aleatorios
$\leq_Z$	orden estocástico del zonoide para vectores aleatorios
$\leq_{\text{ch}}$	orden estocástico de la envolvente convexa para vectores aleatorios
$\leq_{\text{sym}}$	orden estocástico simétrico para vectores y conjuntos aleatorios
$\leq_{\text{slcx}}$	orden estocástico lineal convexo para conjuntos aleatorios
$\leq_{\text{uilcx}}$	orden estocástico superior creciente lineal convexo para conjuntos aleatorios
$\leq_{\text{lilcx}}$	orden estocástico inferior creciente lineal convexo para conjuntos aleatorios
$d_H$	distancia de Hausdorff
$h(\cdot, \cdot)$	función soporte

$x_+$	parte positiva de $x$
$x_-$	parte negativa de $x$
$Z_2(\cdot, \cdot)$	métrica ideal de Zolotarev de orden 2
$\mathcal{C}^1$	conjunto de funciones reales de variable real derivables y con derivada continua
$\text{co}(A)$	envolvente convexa de conjunto $A$
$\text{proj}_\alpha$	aplicación definida en la página 13
$\text{supp } \xi$	soporte del vector aleatorio $\xi$
$\xi_{i:n}$	$i$ -ésimo estadístico ordenado en una muestra de $n$ variables aleatorias independientes distribuidos como $\xi$
$\text{Binom}(n, q)$	variable aleatoria con distribución binomial de parámetros $n$ y $q$
$\partial A$	frontera del conjunto $A$
$\text{aff } A$	clausura afín del conjunto $A$
$\det$	determinante
$V_d$	medida de Lebesgue $d$ -dimensional
$\delta_z$	distribución de probabilidad degenerada en $z$
$[a]$	parte entera de $a$

# Índice alfabético

$\alpha$ -recorte de una probabilidad, 78

clase de Glivenko-Cantelli, 117

comonotonía, 105

conjunto

centralmente simétrico, 8

conexo, 12

convexo, 5

inferior, 38

superior, 38

conjunto aleatorio

cerrado, 4

cerrado y convexo, 5

compacto, 5

compacto y convexo, 5

convergencia débil, 1

cuasi-convexidad, 105

cuerpo convexo, 5

derivada de Radon-Nikodym, 82

diferencia media de Gini, 30

distancia de Hausdorff, 5

entorno,  $r$ -entorno, 5

envolvente convexa, 6

envolvente de zonoides, 26

esperanza de Aumann, 6

familia

de Glivenko-Cantelli, ver clase de Glivenko-Cantelli

de orden acotado, 117

- estable, 117
- función
  - convexa, 16
  - estrictamente convexa, 16
- función de profundidad
  - (definición), 11
  - de conteo, 138
  - de la envolvente convexa, 148
  - del grado medio, 138
  - integral, 104
  - respecto a un estimador de localización, 87
  - semiespacial, 12
  - simplicial, 137
- función soporte, 7
- grado
  - de un punto en un proceso, 137
  - variable aleatoria grado, 137
- índice de Gini, 30
- índices de desigualdad
  - $(f, l)$ -índices de desigualdad para conjuntos aleatorios, 58
  - $(f, l)$ -índices de desigualdad para vectores aleatorios, 57
  - $f$ -índices de desigualdad para intervalos aleatorios, 58
  - $f$ -índices de desigualdad para variables aleatorias, 57
  - aditivamente descomponibles, 57
- intervalo aleatorio compacto, 55
- intervalo intercuantílico, 77
- Lema de Borel-Cantelli, 116
- límite superior de una sucesión de conjuntos, 81
- métrica de probabilidad



- de Zolotarev, 50
  - homogénea de orden  $s$ , 44
  - ideal de orden  $s$ , 44
  - inspirada en el orden simétrico, 44
  - regular, 44
  - simple, 43
- orden estocástico
- convexo, 15
  - creciente convexo, 15
  - creciente positivo lineal convexo, 15
  - de la envolvente convexa, 151
  - de la familia de los órdenes convexos, 15
  - de la familia de los órdenes lineales, 16
  - del zonoide, 18
  - habitual, 15
  - inferior creciente lineal convexo para conjuntos aleatorios, 40
  - integral, 15
  - lineal convexo para conjuntos aleatorios, 36
  - lineal creciente convexo, 15
  - simétrico para conjuntos aleatorios, 33
  - simétrico para vectores aleatorios, 22
  - superior creciente lineal convexo para conjuntos aleatorios, 40
- politopo, 8
- preorden, 14
- probabilidad
- regular, 2
  - tight, 3
- recurrencia, 83
- regiones centrales
- $\varepsilon$ -integrales, 119

- de la envolvente convexa, 149
- de la envolvente de zonoides, 27
- del zonoide, 13
- integrales, 104
- recortadas, 87
- semiespaciales, 12
- relación de orden (parcial), 14
- relativamente compacto, 4
  
- selección de un conjunto aleatorio, 6
- semi-convexidad, 105
- semicontinuidad
  - inferior, 105
  - superior, 84
- simetría multivariante
  - angular, 138
  - esférica, 138
- suma de Minkowski, 5
  
- Teorema
  - de Caratheodory-Vitali, 84
  - de Portmanteau, 2
  - de Prohorov, 4
- topología de la convergencia débil, 3
  
- zonoide, 8
- zonoide de un vector aleatorio, 9
- zonoide elevado de un vector aleatorio, 9
- zonotopo, 8