Universidad de Oviedo

Facultad de Ciencias

Trabajo de Fin de Grado

Extensiones de la preferencia estadística al caso multivariante

Autor: *Julián Ros Martínez*

Tutores : *Raúl Pérez Fernández e Ignacio Montes Gutiérrez*

18 de junio de 2024

Índice general

Capítulo 1

Introducción

Los órdenes estocásticos surgen en el contexto de la teoría de decisión como una herramienta para escoger de forma racional entre distintas alternativas modelizadas mediante variables aleatorias. El orden estocástico más frecuente en la literatura es la dominancia estocástica [2, 11, 17], que se basa en la comparación de funciones de distribución. Sin embargo, la dominancia estocástica posee algunos inconvenientes, siendo los más notorios la incomparabilidad en algunos contextos y el uso de información parcial. La preferencia estadística es otro orden estocástico introducido por De Schuymer et al. en 2003 [6, 7]. Este orden estocástico no posee algunos de los inconvenientes de la dominancia estocástica, lo que hace que en ciertos contextos sea una herramienta suplementaria o incluso superior a la dominancia estocástica. Estas ventajas, en gran medida, vienen dadas por ser definida a través de una relación probabilística que establece un grado de preferencia entre cero y uno. La preferencia estadística posee además una interesante relación con la mediana de la distribución conjunta, lo cual la contrapone a la dominancia estocástica que posee una estrecha relación con la media.

El desarrollo de los órdenes estocásticos ha sido realizado principalmente para variables aleatorias unidimensionales, aunque debido al extensivo uso de la dominancia estocástica se pueden encontrar distintas extensiones al campo multivariante [17, 23]. Sin embargo, no se han identificado extensiones multivariantes ampliamente aceptadas de la preferencia estadística.

Por ello, el objetivo principal de este trabajo es extender esta noción al caso bivariante. Para llevar a cabo este objetivo, el trabajo se organiza como sigue. En primer lugar, en el Capítulo 2 se introducirán algunas nociones necesarias para el resto del trabajo como son las estructuras de preferencia o la teoría de cópulas. A continuación, en el Capítulo 3 se abordarán los órdenes estocásticos unidimensionales. Se comenzará definiendo el concepto genérico de orden estocástico para posteriormente estudiar la dominancia estocástica y sus propiedades. Esto motivará la introducción de la preferencia estadística, donde de nuevo se realizará un repaso de sus propiedades y limitaciones. Por último, se estudiarán ambos órdenes estocásticos en algunas de las distribuciones estadísticas clásicas, analizando en particular el papel que juega la cópula que determina la estructura de dependencia. En el Capítulo 4 se abordará el objetivo principal de este trabajo: extender la preferencia estadística para la comparación de vectores aleatorios. Para ello se propondrán tres extensiones bivariantes distintas de la dominancia estocástica: una extensión mediante la agregación de la preferencia estadística univariante de las distribuciones marginales, una extensión mediante la agregación de las distribuciones marginales de cada vector aleatorio y una extensión puramente bivariante. Posteriormente, se estudiará cómo las distintas extensiones se relacionan entre sí. Finalmente, motivados por la estrecha relación de la preferencia estadística unidimensional con la mediana, se estudiará cómo se relaciona en el caso bivariante. Por último, en el Capítulo 5 se realizará un resumen de lo tratado a lo largo del trabajo y se concluirá con una enumeración de líneas futuras de investigación.

Capítulo 2

Preliminares

En este capítulo se introducirán nociones previas que se utilizarán durante el desarrollo del trabajo. Si bien algunas de ellas resultan muy básicas, como podrían ser las relaciones binarias o las nociones de estadística, el objetivo en este capítulo será aclarar la notación utilizada. Otras nociones, en cambio, no serán tan elementales, por lo que se definirán con detalle y se acompañarán de ejemplos que faciliten su comprensión.

2.1. Estructuras de preferencia

A lo largo de todo el trabajo se van a comparar variables aleatorias. Por ello, en primer lugar vamos a introducir la estructura matemática que se utilizará. Comenzamos introduciendo un concepto que aparece frecuentemente en prácticamente toda rama de las matemáticas, la relación binaria.

Definición 2.1. *[19, Definición 3.1][Relación binaria] Dado A un conjunto finito de elementos, una relación binaria $\mathcal{R}$ sobre A es un subconjunto del producto cartesiano $A \times A$, es decir, un conjunto de pares ordenados (a, b) tales que a y b pertenecen a A. Cada par ordenado del conjunto $\mathcal{R}$ se denota por $a\mathcal{R}b$ (otras notaciones comunes son $\mathcal{R}(a, b)$ ó $(a, b) \in \mathcal{R}$).*

En algunos casos, dada una relación binaria $\mathcal{R}$, resulta conveniente definir la relación inversa o recíproca $\breve{\mathcal{R}}$ [21, pág. 97]:

$$\breve{\mathcal{R}} = \{(a, b) \in A \times A : b\mathcal{R}a\}.$$

En función de cómo se defina, una relación binaria puede cumplir distintas propiedades. A continuación se enumeran las más relevantes.

Definición 2.2. *(Propiedades de relaciones binarias) Dado un conjunto A y una relación binaria $\mathcal{R}$, diremos que $\mathcal{R}$ es:*

- ***Reflexiva**, si $a\mathcal{R}a$ para todo elemento a del conjunto A.*
- ***Irreflexiva**, si $a\not\mathcal{R}a$ para todo elemento a del conjunto A.*
- ***Simétrica**, si $a\mathcal{R}b$ implica que $b\mathcal{R}a$ para todo elemento a, b del conjunto A.*
- ***Asimétrica**, si $a\mathcal{R}b$ implica que $b\not\mathcal{R}a$ para todo elemento a, b del conjunto A.*
- ***Antisimétrica**, si $a\mathcal{R}b$ y $b\mathcal{R}a$ implica que $a = b$ para todo elemento a, b del conjunto A.*
- ***Completa**, si $a\mathcal{R}b$ ó $b\mathcal{R}a$ para todo elemento a, b del conjunto A con a distinto de b.*
- ***Fuertemente completa**, si $a\mathcal{R}b$ ó $b\mathcal{R}a$ para todo elemento a, b del conjunto A.*
- ***Transitiva**, si $a\mathcal{R}b$ y $b\mathcal{R}c$ entonces $a\mathcal{R}c$ para todo elemento a, b, c del conjunto A.*

La relación de preferencia dará lugar a diferentes estructuras en función de las propiedades que satisfaga.

Definición 2.3. *Diremos que una relación binaria $\mathcal{R}$ es una relación de equivalencia, denotada como E, si es reflexiva, simétrica y transitiva.*

Dado un elemento a del conjunto A y una relación de equivalencia E, definimos su clase de equivalencia, y la denotamos por $[a]$, al conjunto:

$$[a] = \{b \in A : a\mathcal{R}b\}.$$

Es directo comprobar que las distintas clases de equivalencia de a son disjuntas, de modo que una relación de equivalencia divide el conjunto en sus distintas clases de equivalencia. Una vez introducido el concepto de relación binaria, procedemos a definir la estructura de preferencia. Las relaciones binarias serán las herramientas utilizadas para definirla. La idea es que cada una de ellas desempeñe un criterio de comparación (indiferencia, igualdad, preferencia, etc.). La siguiente definición formaliza esta idea:

Definición 2.4. *[19, Definición 5.1][Estructura de preferencia] Una estructura de preferencia es una colección de relaciones binarias $\{\mathcal{R}_1, \mathcal{R}_2, ..., \mathcal{R}_n\}$ definidas sobre un conjunto A que satisfacen:*

- *Para cada par (a, b) en $A \times A$, al menos una de las relaciones binarias se satisfacen. Es decir, existe un $k \in \{1, \ldots, n\}$ tal que $a\mathcal{R}_k b$.*

- *Para cada par (a,b) en $A\times A$, si una de las relaciones se satisface, entones ninguna otra se satisface. Es decir, si $a\mathcal{R}_k b$ entonces $a\not\mathcal{R}_l b$ para todo l distinto de k.*

La primera propiedad fuerza que todo par (a,b) de $A \times A$ pertenezca a uno de estos subconjuntos. La segunda propiedad fuerza a que los subconjuntos $\mathcal{R}_k$ de $A \times A$ sean disjuntos. En consecuencia, una estructura de preferencia determina una partición de $A \times A$.
Todo esto puede resultar muy abstracto, de modo que vamos a dar un ejemplo.

Ejemplo 2.1. *Consideremos A como el conjunto de números naturales $\mathbb{N}$ y las relaciones binarias $\{>,<,=\}$. Entonces para cada dos números naturales (n,m) en $\mathbb{N}\times\mathbb{N}$ tendremos que $n > m$, $n < m$ ó $n = m$. Por tanto siempre se cumple una de las tres relaciones binarias. Además, las relaciones son excluyentes entre sí, luego el segundo punto también se cumple.*
Las relaciones binaria "$>$", "$<$" son irreflexivas, asimétricas y transitivas, mientras que "$=$" es reflexiva, simétrica y transitiva luego ambas relaciones están bien caracterizadas por sus propiedades. Por último, podríamos definir la relación binaria $\{\geq\} \equiv \{>\}\cup\{=\}$ pudiendo caracterizar las otras tres de la siguiente manera:

- *$a = b$ si y solo si $a \geq b$ y $b \geq a$.*
- *$a > b$ si y solo si $a \geq b$ y $b \not\geq a$.*
- *$a < b$ si y solo si $a \not\geq b$ y $b \geq a$.*

De este modo, podemos caracterizar toda la estructura de preferencia a través únicamente de "$\geq$".

En el ejemplo que acabamos de considerar, hemos definido la estructura de preferencia más sencilla posible, una en la que existen dos únicas posibilidades en la comparación de dos elementos:

- Uno es preferido al otro ($n > m$ ó $m > n$).
- Existe una indiferencia entre ambos ($n = m$).

Además, en este ejemplo encontramos una propiedad que si bien no sería necesaria de acuerdo con la Definición 2.4, resulta muy útil. Esta propiedad sería la posibilidad de caracterizar la estructura de preferencia a través de una única relación binaria. A esta última se le denominará **relación binaria característica** y se denotará por $\hat{\mathcal{R}}$.
A continuación, vamos a formalizar esta sencilla estructura.

Definición 2.5. *[19, Definición 5.2][Estructura $\langle P, I\rangle$] Una estructura $\langle P, I\rangle$ del conjunto A es una estructura de preferencia de A dotada de la terna de relaciones binarias $\{P, \breve{P}, I\}$ tales que:*

- *P es asimétrica.*
- *I es reflexiva y simétrica.*
- *$\breve{P} = \{(a, b) \in A \times A : bPa\}$.*

Al subconjunto de todos los pares ordenados (a, b) en $A \times A$ tales que "a es preferido a b" lo denominaremos como *relación de preferencia* y se denotará por P, mientras que el subconjunto de todos los pares ordenados (a, b) en $A \times A$ tales que "a es indiferente a b", se denominará *relación de indiferencia* y se denotará por I. Notemos que, en este caso, I es el complementario de $P \cup \breve{P}$.
Si definimos $\hat{\mathcal{R}} = P \cup I$, entonces podríamos construir el resto de relaciones binarias como:

$$aPb \iff a\hat{\mathcal{R}}b \text{ y } b\not\hat{\mathcal{R}}a, \quad a\breve{P}b \iff a\not\hat{\mathcal{R}}b \text{ y } b\hat{\mathcal{R}}a, \quad aIb \iff a\hat{\mathcal{R}}b \text{ y } b\hat{\mathcal{R}}a.$$

En función de las propiedades de la relación binaria característica, podremos definir órdenes de preferencia más fuertes o más débiles. A continuación se definen dos de ellos.

Definición 2.6. *Dado un conjunto A y una estructura de preferencia $\langle P, I\rangle$ dotada de una relación binaria característica $\hat{\mathcal{R}}$, diremos que esta define un orden total si es reflexiva, antisimétrica, completa y transitiva.*

En esta estructura, la indiferencia únicamente se da cuando los objetos son idénticos debido a la asimetría, lo cual resulta muy exigente. La siguiente estructura debilita esta condición.

Definición 2.7. *Dado un conjunto A y una estructura de preferencia $\langle P, I\rangle$ dotada de una relación binaria característica $\hat{\mathcal{R}}$, diremos que esta define un orden débil si es reflexiva, completa y transitiva.*

Al eliminar la condición de antisimetría, la indiferencia pasa a formar una clase de equivalencia no trivial (en el orden total también forma clases de equivalencia, pero estas son de un único elemento).
Como indicamos, la estructura de preferencia $\langle P, I\rangle$ define la estructura de preferencia más sencilla que es posible considerar. A cambio, esta únicamente establece dos posibilidades en la comparación de dos objetos: indiferencia o preferencia. Sin embargo, podrían existir otras posibilidades como la indecisión o la incomparabilidad.

Para ello, existen estructuras más complejas que sí tienen en cuenta esto. En particular, vamos a introducir una estructura con incomparabilidad. Para ello, se añade una tercera relación de equivalencia (la incomparabilidad) a la terna $\{P, \breve{P}, I\}$ que se tenía cuando la incomparabilidad no era posible.

Definición 2.8. *[19, Sección 5.2.2][Estructura $\langle P, I, J\rangle$] Una estructura $\langle P, I, J\rangle$ del conjunto A es una estructura de preferencia de A dotada de la cuaterna de relaciones binarias $\{P, \breve{P}, I, J\}$ tales que:*

- *P es asimétrica.*
- *I es reflexiva y simétrica.*
- *$\breve{P} = \{(a, b) \in A \times A : bPa\}$.*
- *J es simétrica e irreflexiva.*
- *$aJb \iff a\not{P}b,\ b\not{P}a\,, a\not{I}b$.*

A la relación binaria J se le conoce como incomparabilidad. Por otro lado, la relación binaria característica sigue siendo $P \cup I$ de modo que la relación binaria J se definiría como:

$$aJb \iff a\not{\hat{\mathcal{R}}}b \quad y \quad b\not{\hat{\mathcal{R}}}a.$$

Al haber introducido la incomparabilidad, la relación binaria característica deja de ser completa, por tanto, los órdenes definidos dejan de ser válidos. Por ello, definimos dos nuevos órdenes.

Definición 2.9. *Dado un conjunto A y una estructura de preferencia $\langle P, I, J\rangle$ dotada de una relación binaria característica $\hat{\mathcal{R}}$, diremos que esta define un orden parcial si es reflexiva, antisimétrica y transitiva.*

Definición 2.10. *Dado un conjunto A y una estructura de preferencia $\langle P, I, J\rangle$ dotada de una relación binaria característica $\hat{\mathcal{R}}$, diremos que esta define un cuasi-orden débil ó preorden si es reflexiva y transitiva.*

2.2. Análisis Matemático

En esta sección vamos a recordar conceptos básicos del análisis matemático que serán de utilidad a lo largo del trabajo. En principio estos conceptos deberían ser conocidos, el objetivo de enumerarlos es simplemente evitar la ambigüedad que pueda existir entre algunos de ellos.

Definición 2.11 (Monotonía de las funciones)**.** *Dada una función* $f : U \subseteq \mathbb{R} \to \mathbb{R}$*, diremos que:*

1. *f es creciente si para cada $x, y \in U$ tales que $x < y$, entonces $f(x) \leq f(y)$.*
2. *f es decreciente si para cada $x, y \in U$ tales que $x < y$, entonces $f(x) \geq f(y)$.*
3. *f es estrictamente creciente si para cada $x, y \in U$ tales que $x < y$, entonces $f(x) < f(y)$.*
4. *f es estrictamente decreciente si para cada $x, y \in U$ tales que $x < y$, entonces $f(x) > f(y)$.*

Definición 2.12 (Concavidad y convexidad)**.** *Dada una función $f : U \subseteq \mathbb{R} \to \mathbb{R}$ con U un conjunto convexo, diremos que:*

1. *f es cóncava si para todo $x, y \in U$, $\alpha \in [0, 1]$ se tiene que*
$$f\big((1-\alpha)x + \alpha y\big) \geq (1-\alpha)f(x) + \alpha f(y).$$
2. *f es convexa si para todo $x, y \in U$, $\alpha \in [0, 1]$ se tiene que*
$$f\big((1-\alpha)x + \alpha y\big) \leq (1-\alpha)f(x) + \alpha f(y).$$
3. *f es estrictamente cóncava si para todo $x, y \in U$, $\alpha \in [0, 1]$ se tiene que*
$$f\big((1-\alpha)x + \alpha y\big) > (1-\alpha)f(x) + \alpha f(y).$$
4. *f es estrictamente convexa si para todo $x, y \in U$, $\alpha \in [0, 1]$ se tiene que*
$$f\big((1-\alpha)x + \alpha y\big) < (1-\alpha)f(x) + \alpha f(y).$$

Para funciones en la recta real que sean derivables, la monotonía y convexidad se puede caracterizar a partir de las derivadas. La siguiente es una generalización de esta idea.

Definición 2.13. *[4, Sección 14.1][13, Teorema 2.10][n-monotonía] Dada una función $f : U \subseteq \mathbb{R} \to \mathbb{R}$ con U convexo, diremos que f es n-monótona si es n veces diferenciable y para todo $m \leq n$ se cumple que $(-1)^{m+1}u^{(m)} \geq 0$ donde $u^{(m)}$ es la derivada $m-$ésima.*

2.3. Estadística

En esta sección vamos a recordar algunas nociones básicas de estadística unidimensional con el objetivo principal de dejar clara la notación utilizada en el trabajo.

Consideremos un espacio probabilístico $(\Omega, \mathcal{A}, P)$ donde Ω es el espacio muestral, $\mathcal{A}$ es una $\sigma-$álgebra sobre Ω y P una aplicación de $\mathcal{A}$ a $[0,1]$ denominada medida de probabilidad que debe cumplir los conocidos axiomas de Kolmogorov. Denotaremos las variables aleatorias por $X, Y, ...$ y sus funciones de distribución serán denotadas por $F_X, F_Y, ...$ Recordamos que la función de distribución de una variable aleatoria X se define como:

$$F_X(x) = P(X \leq x) = P\big((-\infty, x]\big) = P_X\big(\{\omega \in \Omega : X(\omega) \leq x\}\big), \quad \forall x \in \mathbb{R}.$$

Además, esta cumple las siguientes propiedades:

1. $\lim_{x\to\infty} F_X(x) = 1$.

2. $\lim_{x\to-\infty} F_X(x) = 0$.

3. F_X es una función creciente.

4. F_X es continua por la derecha.

Diremos que X es una variable aleatoria discreta si existe un conjunto S a lo sumo numerable, llamado soporte, tal que $P(S) = 1$ y para todo conjunto A medible se tiene que $P(A) = \sum_{x\in A\cap S} P(\{x\})$. Para variables aleatorias discretas, se define la función de masa ó probabilidad $P(\{k\})$ como $P(\{k\}) = P(X = k)$ para los puntos k en el conjunto S.
Por otra parte, diremos que X es una **variable aleatoria absolutamente continua** si existe una función f_X no negativa tal que:

$$F_X(x) = \int_0^x dx\, f_X(x), \quad \forall x \in \mathbb{R}.$$

Esta función f_X se llamará **función de densidad**, y se cumple que:

$$P(X \in A) = \int_A dx\, f_X(x),$$

para todo conjunto medible A en $\mathbb{R}$.

Por último, en algunos casos resulta útil trabajar con la **función de supervivencia**, denotada por $\bar{F}_X$ y definida como:

$$\bar{F}_X(x) = P(X > x) = 1 - F_X(x), \quad \forall x \in \mathbb{R}.$$

En algunos casos, cuando no exista ambigüedad, se prescindirá del subíndice X para las funciones de densidad, distribución y de supervivencia.
Introducimos ahora dos medidas de centralización que se utilizarán a lo largo del trabajo,

la esperanza y la mediana. La definición de esperanza desde el punto de vista de la teoría de la medida viene dada por:

$$E(X) = \int_{\Omega} dP\, X = \int_{\Omega} P(d\omega)\, X(\omega),$$

donde la integral es en el sentido de Lebesgue [3]. Para el caso particular de variables aleatorias X discretas ó absolutamente continuas con función de densidad f_X, esta definición se simplifica a:

$$E(X) = \sum_{k \in Sop(X)} k \cdot P(X = k), \quad E(X) = \int_{\mathbb{R}} dx\, x f_X(x),$$

donde $Sop(X)$ denota el soporte de la variable aleatoria discreta X.
Además, dada $u : \mathbb{R} \to \mathbb{R}$ una función medible, se define la esperanza de la variable aleatoria $u(X)$,

$$E\big(u(X)\big) = \int_{\Omega} dP\, u(X),$$

que en los casos particulares de variables aleatorias discretas y absolutamente continuas se simplifica a:

$$E(X) = \sum_{k \in Sop(X)} u(k) \cdot P(X = k), \quad E(X) = \int_{\mathbb{R}} dx\, u(x) f_X(x).$$

Un concepto que aparece con frecuencia junto con la esperanza es la varianza. Esta mide la dispersión de la variable aleatoria respecto de la esperanza y se define como,

$$Var(X) = E\big[(X - E(X))^2\big].$$

Otra medida de centralización algo menos común que la media es la mediana. Para una variable aleatoria X la denotaremos por $Me(X)$ y se define como el conjunto de valores t satisfaciendo:

$$P\big(X \leq t\big) \geq \frac{1}{2} \quad y \quad P\big(X \geq t\big) \geq \frac{1}{2}.$$

Observar que la esperanza puede no existir (también se suele decir que no es finita) mientras que la mediana siempre existe. Sin embargo, la esperanza en caso de existir es única, mientras que la mediana no tiene por qué serlo, esta puede ser un conjunto de valores.
Por otro lado, cuando se consideren vectores aleatorios de dimensión 2, se considerará un espacio de probabilidad $(\Omega, \mathcal{A}, P)$ donde los elementos son análogos al caso unidimensional. En general, los vectores aleatorios se denotarán por (X, Y) cuando solo exista uno o por (X_1, X_2), (Y_1, Y_2) cuando se quieran comparar dos vectores aleatorios. De manera

análoga se define la función de distribución conjunta de un vector aleatorio (X, Y):

$$\begin{aligned} F_{XY}(x,y) &= P(X \le x, Y \le y) = P_{X,Y}\big((-\infty, x] \times (-\infty, y]\big) \\ &= P\big(\{\omega \in \Omega : X(\omega) \le x, Y(\omega) \le y\}\big), \quad \forall (x,y) \in \mathbb{R}^2. \end{aligned}$$

Esta debe cumplir las siguientes propiedades:

- $\lim_{x,y\to\infty} F_{XY}(x,y) = 1$.
- $\lim_{x\to-\infty} F_{XY}(x,y) = \lim_{y\to-\infty} F_{XY}(x,y) = 0$.
- F_{XY} es continua por la derecha en todo punto de $\mathbb{R}^2$.
- Para todo $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ con $x_1 < x_2$ e $y_1 < y_2$ se cumple la desigualdad de los rectángulos:

$$F_{XY}(x_2,y_2) - F_{XY}(x_1,y_2) - F_{XY}(x_2,y_1) + F_{XY}(x_1,y_1) \ge 0.$$

Esta función de distribución contiene toda la información de la dependencia entre las variables aleatorias, en particular se cumple que “X e Y son independientes si y solo si $F_{XY}(x,y) = F_X(x)F_Y(y)$, para todo (x,y) en $\mathbb{R}^2$”.

Podemos también definir la función de supervivencia:

$$\overline{F}_{XY} = P(X > x, Y > y) = 1 - F_X(x) - F_Y(y) + F_{XY}(x,y), \quad \text{para todo } (x,y) \text{ en } \mathbb{R}^2.$$

De manera totalmente análoga al caso de variables aleatorias, diremos que un vector aleatorio (X,Y) es absolutamente continuo si existe una función de densidad f_{XY} no negativa tal que:

$$P\big((X,Y) \in A\big) = \int_A d(x,y)\, f_{XY}(x,y),$$

para todo conjunto medible A de $\mathbb{R}^2$. Si además X e Y son independientes $f_{XY}(x,y) = f_X(x)f_Y(y)$.

Ahora, vamos a recordar un resultado clásico para variables aleatorias unidimensionales.

Proposición 2.1. *Sea U una variable aleatoria con distribución uniforme en $[0,1]$ y F una función de distribución. Si $\hat{F}^{-1}$ es la inversa generalizada definida como $\hat{F}^{-1}(y) := \inf\{x : F(x) \ge y\}$, entonces:*

$$P(\hat{F}^{-1}(U) \le x) = F(x), \quad \forall x \in \mathbb{R}.$$

Recíprocamente, si X es una variable aleatoria continua con función de distribución F_X, entonces:

$$F_X(X) \sim U(0,1).$$

Este resultado cobra importancia cuando se desean simular variables aleatorias, pues haciendo uso de la inversa generalizada $\hat{F}_X^{-1}$ de la función de distribución, se tiene $X \sim \hat{F}_X^{-1}(U)$, de modo que a partir de valores simulados en una uniforme se pueden obtener valores simulados de X. Sin embargo, en nuestro caso utilizaremos este resultado más adelante en la teoría de cópulas.
Uno de los objetivos del uso de variables aleatorias es la modelización de experimentos aleatorios de la "vida real". Con ello, existen ciertas distribuciones modelo que estas variables aleatorias pueden seguir. Esta distribución modelo no tiene por qué corresponderse exactamente con el problema real, a veces es una idealización del mismo. Algunas de estas distribuciones se utilizarán a lo largo de este trabajo, luego se van a introducir estas distribuciones junto con algunas de sus propiedades más relevantes.

Distribución binomial

Diremos que una variable aleatoria X sigue una distribución binomial de parámetros p, n con $n \in \mathbb{N}$ y $p \in (0, 1)$, y la denotamos por $\mathcal{B}(n, p)$, si la probabilidad de que la variable aleatoria tome el valor $k \in \{0, 1, \ldots, n\}$ es:

$$P(X = k) = \binom{n}{k} p^k (1 - p)^k.$$

Además, para $X \sim \mathcal{B}(n, p)$ se cumple:

$$E(X) = n \cdot p, \quad Var(X) = np(1 - p).$$

Distribución de Poisson

Diremos que una variable aleatoria sigue una distribución de Poisson de parámetro λ, y la denotamos por $\mathcal{P}(\lambda)$ con $\lambda > 0$, si la probabilidad de que tome el valor $k \in \{0, 1, 2, ...\}$ es:

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}.$$

Su función de distribución viene dada por:

$$F(k) = \frac{\Gamma\big(\lfloor k + 1 \rfloor, \lambda\big)}{k!}, \quad k > 0,$$

donde $\Gamma(x, y)$ es la función gamma incompleta [1, Sección 26.4.21]. Además, para $X \sim \mathcal{P}(\lambda)$ se cumple:

$$E(X) = \lambda, \quad Var(X) = \lambda.$$

Distribución uniforme continua en $[a, b]$

Diremos que una variable aleatoria X tiene una distribución uniforme en el intervalo $[a, b]$, y lo denotaremos por $\mathcal{U}(a, b)$, si su función de densidad es de la forma:

$$f(x) = \frac{1}{b-a} \mathbb{1}_{[a,b]}(x), \quad \forall x \in \mathbb{R}.$$

Además, para $X \sim \mathcal{U}(a, b)$ se cumple:

$$E(X) = \frac{a+b}{2}, \quad Var(X) = \frac{(b-a)^2}{12}.$$

Un caso bastante importante ocurre cuando $a = 0$ y $b = 1$, que se corresponde con el intervalo $[0, 1]$, y que se conoce como distribución uniforme estándar y que se suele denotar por $\mathcal{U}(0, 1)$.

Distribución normal

Diremos que una variable aleatoria X tiene una distribución normal de media $\mu \in \mathbb{R}$ y desviación típica $\sigma > 0$, y lo denotaremos por $\mathcal{N}(\mu, \sigma)$, si su función de densidad es de la forma:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \forall x \in \mathbb{R}.$$

Un caso muy importante ocurre cuando $\mu = 0$, $\sigma = 1$, que se conoce como distribución **normal estándar** y su función de distribución se denota por $\Phi(x)$. Esta distribución tiene numerosas propiedades muy útiles, como por ejemplo:

Reproductividad Si X e Y son distribuciones normales independientes con parámetros (μ_X, σ_X) y (μ_Y, σ_Y) respectivamente, entonces se cumple que:

$$X + Y \sim \mathcal{N}\left(\mu_X + \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2}\right).$$

Transformaciones afines Si X sigue una distribución normal de parámetros (μ, σ), entonces se cumple que:

$$aX + b \sim \mathcal{N}(a\mu + b, |a|\sigma), \quad \forall a, b \in \mathbb{R}.$$

Tipificación Es un caso concreto cuando $a = 1/\sigma$, $b = -\mu/\sigma$,

$$\frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

Simetría La distribución es simétrica respecto a su media, de este modo se cumple que:

$$F_X(\mu_X + x) = 1 - F_X(\mu_X - x), \quad \forall x \in \mathbb{R}.$$

Distribución exponencial

Diremos que una variable aleatoria X tiene una distribución exponencial de parámetro $\lambda > 0$, y lo denotaremos por $Exp(\lambda)$, si su función de densidad es de la forma:

$$f(x) = \lambda e^{-\lambda x} \cdot \mathbb{1}_{[0,\infty)}(x), \quad \forall x \in \mathbb{R}.$$

Para $X \sim Exp(\lambda)$ se cumple:

$$E(X) = \frac{1}{\lambda}, \quad Var(X) = \frac{1}{\lambda^2}.$$

Una propiedad relevante de esta distribución es que bajo transformaciones lineales positivas la distribución sigue siendo exponencial

$$X \sim Exp(\lambda) \implies aX \sim Exp\Big(\frac{\lambda}{a}\Big), \quad a > 0.$$

Familias de localización

Diremos que una familia de distribuciones es una familia de localización si y solo si para todo par de variables aleatorias X, Y pertenecientes a la familia, existe un b tal que X e $Y + b$ tienen la misma distribución. Por tanto, todas las distribuciones de la familia serían traslaciones de una única distribución. Un ejemplo de esta clase de familias sería la normal con varianza fija, donde el parámetro μ define una traslación de la normal estándar.

Familias de escala

Diremos que una familia de distribuciones es una familia de escala si y solo si para todo par de variables aleatorias X, Y pertenecientes a la familia, existe un $a > 0$ tal que X y aY tienen la misma distribución. Por tanto, todas las distribuciones de la familia serían escalamientos de una única distribución. Ejemplos de esta clase de familias serían la exponencial y la normal con media fija donde los parámetros λ y σ son parámetros relacionados con la escala.

Familias de localización-escala

Diremos que una familia de distribuciones es una familia de localización-escala si y solo si para todo par de variables aleatorias X, Y pertenecientes a la familia, existe un $a > 0$ y un b tales que X y $aY + b$ tienen la misma distribución. Un ejemplo de esta clase de familias sería la normal donde el parámetro μ define un parámetro de traslación y el parámetro σ uno de escala.

Distribución normal bivariante

Diremos que un vector aleatorio (X, Y) tiene una distribución normal bivariante con vector de medias $\vec{\mu}^T = (\mu_X, \mu_Y)$ y matriz de varianzas-covarianzas $\Sigma = \begin{pmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{pmatrix}$; donde ρ es el coeficiente de correlación de Pearson entre X e Y y σ_X^2, σ_Y^2 son las varianzas de X e Y respectivamente, si su función de densidad es de la forma:

$$f(x,y) = \frac{1}{2\pi|\Sigma|^{1/2}} exp\Bigg(-\frac{1}{2}\Big((x,y) - (\mu_X,\mu_Y)\Big)\Sigma^{-1}\Big((x,y) - (\mu_X,\mu_Y)\Big)^T\Bigg).$$

En el caso que $\vec{\mu} = 0_{\mathbb{R}_2} = (0,0)^T$ entonces denotamos la función de densidad por $\phi_\Sigma(x,y)$ y la función de distribución por $\Phi_\Sigma(x,y)$. Este caso es importante ya que si (X,Y) tiene una distribución normal bivariante con vector de medias (μ_X, μ_Y) entonces $(X,Y) - (\vec{\mu}_X, \vec{\mu}_Y)$ sigue una distribución normal bivariante con vector de medias nulo.
Por otro lado, para las distribuciones marginales de una distribución normal bivariante se cumple que:

$$X \sim \mathcal{N}(\mu_X, \sigma_X), \quad Y \sim \mathcal{N}(\mu_Y, \sigma_Y).$$

Notar, que el recíproco no es cierto, es decir existen vectores aleatorios con marginales normales pero que la conjunta no es una distribución normal bivariante.
A modo de resumen, en la Tabla 2.1 podemos encontrar las funciones de densidad/masa de probabilidad, esperanzas y varianzas de las variables aleatorias unidimensionales descritas.

2.4. Cópulas

Cuando se estudian vectores aleatorios aparece el concepto de dependencia. Este es en principio bastante simple: dos variables aleatorias son dependientes si y solo si su función de distribución conjunta no se puede expresar mediante el producto de sus funciones de distribución marginales. Sin embargo, cuando se busca trabajar con variables aleatorias

Distribución	**Función de densidad/masa**	**Esperanza**	**Varianza**
$\mathcal{B}(n,p)$	$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},\ k=0,1,\ldots,n$	np	$np(1-p)$
$\mathcal{P}(\lambda)$	$P(X=k)=\dfrac{\lambda^k e^{-\lambda}}{k!},\ k=0,1,2,\ldots$	λ	λ
$U(a,b)$	$f(x)=\dfrac{1}{b-a}\mathbb{1}_{[0,1]}(x),\ x\in\mathbb{R}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
$\mathcal{N}(\mu,\sigma)$	$f(x)=\dfrac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\ x\in\mathbb{R}$	μ	σ^2
$Exp(\lambda)$	$f(x)=\lambda e^{-\lambda x}\mathbb{1}_{[0,\infty)}(x),\ x\in\mathbb{R}$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$

Tabla 2.1: Distribuciones con sus funciones de densidad/masa de probabilidad, esperanzas y varianzas

dependientes esta definición resulta demasiado general al no dar ninguna información sobre la dependencia más allá de que la igualdad no se satisfaga.
Por tanto, ¿cómo podemos estudiar la dependencia entre las variables aleatorias cuando no son independientes? Recordemos el caso de variables aleatorias independientes donde se tiene que $F_{XY}(x,y)=F_X(x)F_Y(y)$. Observamos que si definimos una función $C(u,v)=u\cdot v$ entonces:

$$F_{XY}(x,y)=C\Big(F_X(x),F_Y(y)\Big)=F_X(x)\cdot F_Y(y),\quad \forall(x,y)\in\mathbb{R}^2.$$

De este modo, podríamos intentar generalizar a otras funciones $C(u,v)$ que contengan toda la información de la dependencia entre las variables aleatorias independientemente de cuales sean sus distribuciones. A esta función $C(u,v)$ la llamaremos cópula. Sin embargo, no toda función puede ser una cópula, pues al generar esta la función de distribución conjunta, lo ha de hacer satisfaciendo las propiedades de la misma.

Definición 2.14 (Cópula bidimensional)**.** *Una cópula bidimensional es una función* $C:[0,1]^2\to[0,1]$ *satisfaciendo:*

- *Para todo* u, v *en* $[0,1]$*:*
  - $C(u,1)=u$ *y* $C(1,v)=v$.
  - $C(0,v)=C(u,0)=0$.
- *Para todo* u_1, u_2, v_1, v_2 *en* $[0,1]$ *tal que* $u_1\le u_2$ *y* $v_1\le v_2$*:*

$$C(u_2,v_2)-C(u_2,v_1)-C(u_1,v_2)+C(u_1,v_1)\ge 0.$$

Esta definición viene a decirnos que una cópula es una función de distribución bidimensional con marginales uniformes restringidas al dominio $[0,1]\times[0,1]$, es decir, podemos ver una cópula como la función de distribución en $[0,1]\times[0,1]$ de dos variables aleatorias uniformes U, V. Ahora, recordando la Proposición 2.1, podemos expresar la cópula a partir de las funciones de distribución del vector aleatorio y sus componentes:

$$C(u,v) = F_{XY}\big(\hat{F}_X^{-1}(u), \hat{F}_Y^{-1}(v)\big).$$

Hemos visto que las cópulas nos permiten expresar vectores aleatorios en función de la relación entre las variables. Por un lado, parece claro que toda cópula pueda definir un vector aleatorio, sin embargo, ¿es el recíproco cierto? El siguiente teorema responde a esta pregunta y da sentido al uso de cópulas.

Teorema 2.2. *[18, Teorema 2.3.3][Teorema de Sklar] Dado un vector aleatorio (X,Y) con función de distribución asociada F_{XY} y con marginales F_X, F_Y, entonces existe una cópula C tal que:*

$$F_{XY}(x,y) = C(F_X(x), F_Y(y)), \quad \forall (x,y) \in \mathbb{R}^2.$$

Si F_X y F_Y son continuas, entonces C es única; de otro modo, C está unívocamente determinada por $RanF_X \times RanF_Y$ donde $RanF \subseteq [0,1]$ es el conjunto de valores que toma la función de distribución F.
Recíprocamente, si C es una cópula y F_X y F_Y son funciones de distribución de las variables aleatorias X, Y respectivamente, entonces para todo x, y reales, $F_{XY}(x,y) = C(F_X(x), F_Y(y))$ es una función de distribución con marginales $F_X(x)$ y $F_Y(y)$.

Este es el teorema de central de la teoría de cópulas y tiene una implicación muy importante, pues asegura que la dependencia entre las variables aleatorias está completamente descrita por una función (la cópula) independientemente de las distribuciones marginales que sigan las variables aleatorias.
Hasta ahora, la teoría de cópulas se ha formulado mediante la función de distribución, pero ¿se podrá hacer lo mismo con la función de supervivencia? Consideremos un vector aleatorio (X,Y) ligado por una cópula $C(u,v)$.

$$\begin{aligned}
\overline{F}_{XY}(x,y) &= 1 - F_X(x) - F_Y(y) + F_{XY}(x,y) \\
&= \overline{F}_X(x) + \overline{F}_Y(y) - 1 + C(F_X(x), F_Y(y)) \\
&= \overline{F}_X(x) + \overline{F}_Y(y) - 1 + C(1 - \overline{F}_X(x), 1 - \overline{F}_Y(y)).
\end{aligned}$$

De este modo, si se considera la función $\overline{C}(u,v) = u + v - 1 + C(1-u, 1-v)$, entonces

$$\overline{F}_{XY}(x,y) = \overline{C}\big(\overline{F}_X(x), \overline{F}_Y(y)\big), \quad \forall (x,y) \in \mathbb{R}^2.$$

Al caracterizar la cópula la dependencia entre las variables, este resultado se podía esperar. Sin embargo, esta expresión nos será útil más adelante en un contexto distinto al de la función de supervivencia.

Definición 2.15. *[18, Sección 2.6] Dada una cópula $C(u,v)$, la función*

$$\overline{C}(u,v) = u + v - 1 + C(1-u, 1-v),$$

describe también una cópula, conocida como Cópula de Supervivencia y denotada como $\overline{C}(u,v)$.

Hemos visto cómo las cópulas nos permiten modelizar dos variables aleatorias en función únicamente de cómo depende una de la otra. Esta dependencia puede ser mayor o menor, siendo esta información provista por la cópula. Pero ¿cómo de grande puede ser esta dependencia? Consideremos por ejemplo dos variables aleatorias uniformes U_1, U_2 con $U_1 = U_2$, observamos que esta es la máxima dependencia positiva que puede haber entre dos uniformes. En particular, su cópula sería:

$$C(u,v) = P(U \leq u, V \leq v) = P(U \leq u, U \leq v) = P(U \leq \min\{u,v\}) = \min\{u,v\}.$$

Esta cópula se conoce como cópula del mínimo y generalmente se denota por $M(u,v)$. Las variables aleatorias vinculadas por esta cópula se dicen comonótonas. Para variables aleatorias arbitrarias se puede demostrar que si X e Y son comonótonas entonces existe una función T creciente tal que $T(X) = Y$ casi seguro.
En el otro lado, tendríamos el caso donde $V = 1 - U$, donde la dependencia es contraria, si U crece, V decrece; sería la máxima dependencia negativa entre dos uniformes. En este caso la cópula sería:

$$\begin{aligned} C(u,v) &= P(U \leq u, V \leq v) = P(U \leq u, U \geq 1 - v) \\ &= \begin{cases} 0, & si \quad u + v < 1 \\ P(1 - v \leq U \leq u) & si \quad u + v \geq 1, \end{cases} \\ &= \begin{cases} 0, & si \quad u + v < 1, \\ u + v - 1, & si \quad u + v \geq 1, \end{cases} = \max\{u + v - 1, 0\}. \end{aligned}$$

Esta cópula se conoce como cópula de Lukasiewicz y generalmente se denota por $W(u,v)$. Las variables aleatorias vinculadas por esta cópula se dicen contramonótonas. Para variables aleatorias arbitrarias se puede demostrar que si X e Y son contramonótonas entonces existe una función T decreciente tal que $T(X) = Y$ casi seguro.
Estos dos casos de dependencia extrema dan lugar a unas cotas entre las que debe estar cualquier cópula.

Teorema 2.3. *[18, Teorema 2.2.3] [Cota de Fréchet-Hoeffding] Para toda cópula $C(u,v)$*

se cumple que:

$$M(u,v) \leq C(u,v) \leq W(u,v), \quad \forall u,v \in [0,1] \times [0,1].$$

En la Figura 2.1 podemos observar la interpretación geométrica de la cota de Frechet-Hoeffding. Toda cópula estará contenida en la pirámide donde las caras posterior e inferior las delimita la cópula de Lukasiewicz ($C(u,v) = \text{máx}\{u+v-1,0\}$), mientras que las caras frontales están delimitadas por la cópula del mínimo ($C(u,v) = \text{mín}\{u,v\}$).

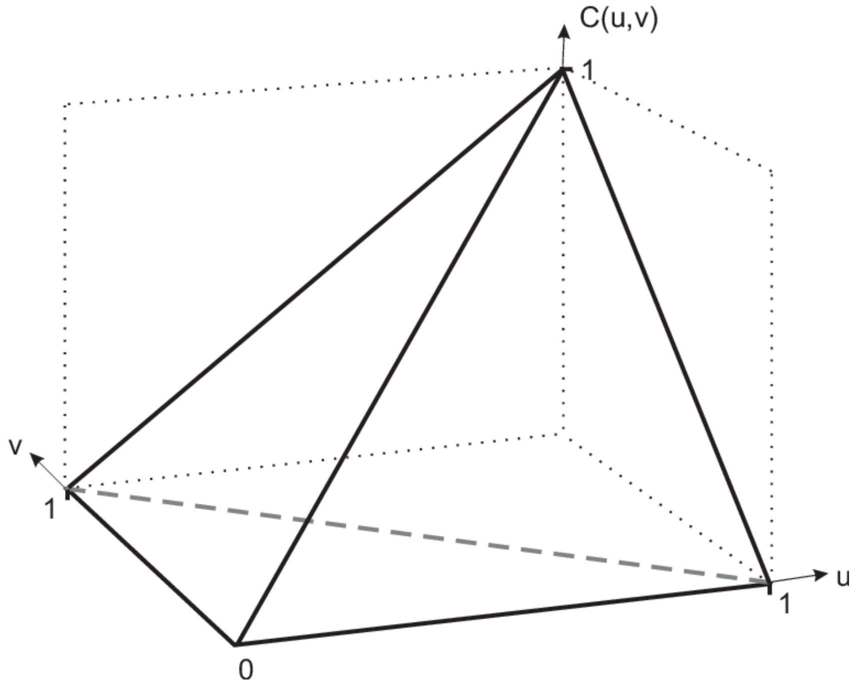

Figura 2.1: Interpretación geométrica de la cota de Frechet-Hoeffding. Créditos: T. Schmidt (2007) [22].

Pasamos ahora a introducir algunas de las cópulas más relevantes que se pueden encontrar en la literatura. Aunque ya la adelantamos previamente, en primer lugar introducimos la cópula del producto:

$$\Pi(u,v) = u \cdot v, \quad (u,v) \in [0,1] \times [0,1].$$

Es consecuencia directa del teorema de Sklar que dos variables aleatorias son independientes si y solo si su cópula es la cópula del producto, de ahí su importancia.
Tal y como hemos descrito a partir de la Proposición 2.1, podemos construir distintas cópulas a partir de distribuciones conocidas. Por ejemplo si quisiéramos determinar la cópula que vincula dos normales en una normal bivariante, podríamos hacerlo mediante un cópula, denominada *cópula Gaussiana*, de la siguiente manera

$$C_{\rho}^{Gauss}(u,v) = \Phi_{\Sigma}(\Phi^{-1}(u), \Phi^{-1}(v)),$$

donde Φ_Σ es la función de distribución de la normal bivariante con vector de medias $\vec{\mu} = 0_{\mathbb{R}^2}$ y matriz de varianzas covarianzas con unos en la diagonal y ρ en el resto, mientras que Φ la función de distribución de la normal estándar.
Tres casos particulares de esta cópula serían la cópula del producto cuando $\rho = 0$, la cópula de Lukasiewicz cuando $\rho = -1$ y la cópula del mínimo cuando $\rho = 1$.
Por último, introducimos una familia de cópulas que aparece con frecuencia en la literatura, las *cópulas Arquimedianas*. Estas cópulas son utilizadas debido a distintas razones: se construyen de manera sencilla, una gran variedad de cópulas pertenecen a esta familia y tienen algunas propiedades que facilitan su manejo.
En primer lugar, introducimos la noción de pseudo-inversa de una función.

Definición 2.16 (Función Pseudo-inversa). *Sea $\varphi : [0,1] \to [0,\infty]$ una función continua y estrictamente decreciente tal que $\varphi(1) = 0$. La función pseudo-inversa de φ, denotada por $\varphi^{[-1]}$ viene dada por:*

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & 0 \leq t \leq \varphi(0), \\ 0, & \varphi(0) \leq t \leq \infty. \end{cases}$$

Observación 1. *La pseudo-inversa $\varphi^{[-1]}$ es una función continua y decreciente en $[0,\infty]$. Además $\varphi^{[-1]}(\varphi(u)) = u$ en $[0,1]$ y :*

$$\varphi\Big(\varphi^{[-1]}(t)\Big) = \begin{cases} t, & 0 \leq t \leq \varphi(0), \\ \varphi(0), & \varphi(0) \leq t \leq \infty. \end{cases} = \min\{t, \varphi(0)\}.$$

Además, si $\varphi(0) = \infty$, entonces $\varphi^{[-1]} = \varphi^{-1}$.

Consideremos la cópula del producto $\Pi(u,v)$ y la función $\varphi(t) = -\log(t)$, entonces:

$$\varphi(\Pi(u,v)) = \varphi(u) + \varphi(v), \quad \forall u,v \in (0,1] \implies \Pi(u,v) = \varphi^{-1}\Big(\varphi(u) + \varphi(v)\Big).$$

Así, la cópula del producto ha quedado totalmente determinada por φ. Entonces, si encontramos otras funciones φ adecuadas podríamos construir nuevas cópulas de manera sencilla; es así es como se construyen las cópulas arquimedianas. El siguiente teorema nos indica qué clase de funciones hemos de buscar.

Teorema 2.4. *[18, Teorema 4.1.4] Consideremos una función continua y estrictamente decreciente $\varphi : [0,1] \to [0,\infty]$ con $\varphi(1) = 0$. Entonces la función*

$$C(u,v) = \begin{cases} \varphi^{[-1]}\Big(\varphi(u) + \varphi(v)\Big), & si \quad \varphi(u) + \varphi(v) \leq \varphi(0), \\ 0, & en\ otro\ caso. \end{cases}$$

es una cópula si y solo si φ es convexa.

A la función φ se le conoce como generador (generador aditivo). Si además $\varphi(0) = \infty$, entonces el generador se dice que es estricto y $C(u,v) = \varphi^{-1}(\varphi(u) + \varphi(v))$.
Consideremos por ejemplo el generador $\varphi(u) = \frac{1}{\theta}(u^{-\theta} - 1)$ con θ perteneciente a $[-1, \infty) \setminus \{0\}$. Si $\theta > 0$, el generador es estricto y la inversa viene dada por $\varphi^{-1} = (\theta t + 1)^{-1/\theta}$. Si θ pertenece a $[-1, 0)$, su pseudo-inversa es

$$\varphi^{[-1]}(t) = \begin{cases} (\theta t + 1)^{-1/\theta}, & si \quad 0 \leq t \leq -\frac{1}{\theta}, \\ 0, & si \quad -\frac{1}{\theta} \leq t \leq \infty. \end{cases}$$

De este modo se tiene que $C(u,v) = \Big(\text{máx}\{u^{-\theta} + v^{-\theta} - 1, 0\}\Big)^{-1/\theta}$, conocida como *Cópula de Clayton*. Si estudiamos ahora los comportamientos asintóticos:

- $\theta = -1$ obtenemos la cópula de Lukasiewicz.
- $\theta \to \infty$ obtenemos la cópula del mínimo.
- $\theta \to 0$ la cópula del producto.

Esta es la idea subyacente que encontramos en todas las familias de cópulas arquimedianas, interpolar entre otras cópulas.

2.5. Simetrías

En estadística, muchas de las distribuciones conocidas son simétricas. En el caso multidimensional esta propiedad no es tan clara, pues pueden existir distintas simetrías. En esta sección discutimos esta propiedad, especialmente en el caso multidimensional y su uso en cópulas.

Definición 2.17 (Variable aleatoria simétrica)**.** *Diremos que una variable aleatoria X es simétrica respecto de S_X si las funciones de distribución de $X - S_X$ y $S_X - X$ coinciden, es decir, para cualquier x perteneciente a $\mathbb{R}$, $P(X - S_X \leq x) = P(S_X - X \leq x)$.*
Si además X es una variable aleatoria continua con función de distribución F_X, esto se traduce en:

$$F_X(S_X + x) = 1 - F_X(S_X - x) = \overline{F}_X(S_X - x), \quad \forall x \in \mathbb{R}.$$

Si además la distribución tiene función de densidad $f_X(x)$, entonces se cumple que:

$$f_X(S_X + x) = f_X(S_X - x), \quad \forall x \in \mathbb{R}.$$

Cuando nos adentramos en el caso multidimensional, las simetrías pueden ser diversas, pues esta puede serlo respecto a distintos planos o incluso podríamos considerar la simetría de las marginales únicamente.

Definición 2.18. *Dado un vector aleatorio (X,Y) y un punto (S_X,S_Y) de $\mathbb{R}^2$, diremos que:*

1. *(X,Y) es marginalmente simétrico respecto de (S_X,S_Y) si X e Y son simétricas respecto de S_X y S_Y, respectivamente.*
2. *(X,Y) es radialmente simétrico respecto de (S_X,S_Y) si la función de distribución conjunta de $(X-S_X,Y-S_Y)$ coincide con la de (S_X-X,S_Y-Y).*

Observar que simetría marginal implica simetría radial pero el recíproco no es cierto en general. De hecho, el caso más relevante de estudio será el segundo; por tanto, cuando digamos que un vector aleatorio es simétrico, esto significará que es radialmente simétrico.

Ejemplo 2.2. *Consideremos un vector aleatorio (X,Y) que sigue una distribución normal bivariante. Vamos a ver que este vector aleatorio es radialmente simétrico respecto del vector de medias (μ_X,μ_Y). Tenemos que comprobar que las siguientes funciones de distribución son idénticas:*

$$F_{X-\mu_X,Y-\mu_Y}(x,y)=P(X-\mu_X\leq x,Y-\mu_Y\leq y)=\Phi_\Sigma(x,y),$$

$$F_{\mu_X-X,\mu_Y-Y}(x,y)=P(-X+\mu_X\leq x,-Y+\mu_Y\leq y)=\overline{\Phi}_\Sigma(-x,-y).$$

La función de densidad de la normal bivariante con medias nulas viene dada por:

$$\phi_\Sigma(x,y)=\frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}}exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{x^2}{\sigma_X^2}+\frac{y^2}{\sigma_Y^2}-\frac{2\rho xy}{\sigma_X\sigma_Y}\right)\right].$$

de donde se deduce que $\phi(x,y)=\phi(-x,-y)$. De este modo:

$$\Phi_\Sigma(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}\phi(s,t)dsdt=\int_{-y}^{\infty}\int_{-x}^{\infty}\phi(u,v)dudv=\overline{\Phi}_\Sigma(-x,-y),$$

para todo $(x,y)\in\mathbb{R}^2$, donde se ha realizado el cambio de variables $u=-t$, $v=-s$.

Partiendo de dos variables aleatorias simétricas, en la Definición 2.18 se han descrito dos posibles casos de vectores aleatorios con marginales simétricas: una donde la distribución conjunta no tenía por qué ser simétrica, y otra donde sí era simétrica respecto del plano $\{(x,y)\in\mathbb{R}^2 : x-S_X=y-S_Y\}$. En el primer caso, la cópula parece "romper" la simetría, mientras que en el segundo, la mantiene. Surge entonces la pregunta, ¿qué condiciones ha de tener la cópula para mantener la simetría?

Teorema 2.5. *[18, Teorema 2.7.3] Sean X e Y dos variables aleatorias continuas con funciones de distribución marginales F_X, F_Y, respectivamente, y conjunta F_{XY}, siendo C la cópula que las relaciona. Supongamos que X e Y son simétricas. Entonces, son equivalentes*

1. *(X, Y) es radialmente simétrica.*
2. *$C(u, v) = u + v - 1 + C(1 - u, 1 - v)$ para todo (u, v) en $[0, 1]^2$.*
3. *La cópula coincide con su cópula de supervivencia.*

Este teorema caracteriza las cópulas con simetría radial bajo una condición sencilla, que la cópula coincida con la de supervivencia. De este modo las distribuciones conocidas que son simétricas deberían cumplir esto.

Ejemplo 2.3. *Vamos a ver que la cópula Gaussiana coincide con su cópula de supervivencia, esto es consecuencia directa del Ejemplo 2.2 y del Teorema 2.5 pues:*

$$\overline{C}_{\rho}^{Gauss} = C_{\rho}^{Gauss}.$$

Otras cópulas relevantes que cumplen esta propiedad son la del mínimo, la de Lukasiewicz y la del producto.

Ejemplo 2.4. *Consideremos las cópulas del mínimo $M(u, v)$, Lukasiewicz $W(u, v)$ y del producto $\Pi(u, v)$.*

$$\begin{aligned}
\overline{M}(u, v) &= u + v - 1 + \text{mín}\{1 - u, 1 - v\} = u + v - \text{máx}\{u, v\} = \text{mín}\{u, v\} = M(u, v). \\
\overline{W}(u, v) &= u + v - 1 + \text{máx}\{1 - u + 1 - v - 1, 0\} = \text{máx}\{u + v - 1, 0\} = W(u, v). \\
\overline{\Pi}(u, v) &= u + v - 1 + (1 - u)(1 - v) = uv = \Pi(u, v).
\end{aligned}$$

Entre las cópulas arquimedianas, la única que es radialmente simétrica, es decir que coincide con su cópula de supervivencia, es la cópula de Frank [1, Proposición 4.2]. Sin embargo, estas poseen otra clase de simetría más débil.

Proposición 2.6 (Teorema 2.7.4)**.** *[18] Consideremos dos variables aleatorias X, Y con funciones de distribución F_X, F_Y respectivamente y función de distribución conjunta F_{XY} a partir de una cópula C. Entonces son equivalentes:*

1. *X e Y son intercambiables, es decir $(X, Y) \stackrel{\mathcal{D}}{=} (Y, X)$.*
2. *$F_X = F_Y$ y $C(u, v) = C(v, u)$ para todo u, v en $[0, 1]$.*

Observar que se ha introducido la notación "$\overset{\mathcal{D}}{=}$", esto denota que ambos vectores aleatorios (para variables aleatorias también se utilizará) siguen la misma distribución, sin embargo, no tienen por qué ser el mismo vector aleatorio.
Al definir las cópulas arquimedianas como $\varphi^{[-1]}\big(\varphi(u)+\varphi(v)\big)$, toda cópula arquimediana posee esta clase de simetría.

Capítulo 3

Órdenes estocásticos univariantes

En este capítulo abordaremos la comparación de variables aleatorias en el caso unidimensional. Empezaremos introduciendo el orden estocástico más frecuente en la literatura, la dominancia estocástica, para luego trabajar con el orden estocástico motivo de este trabajo, la preferencia estadística, y finalmente dar caracterizaciones para algunas distribuciones.
Dadas dos variables aleatorias X e Y definidas en el mismo espacio probabilístico $(\Omega, \mathcal{A}, P)$, definiremos una relación binaria $\succeq$ para comparar ambas variables en el ámbito deseado. Para ello podemos utilizar características de las variables aleatorias como podrían ser su función de distribución, esperanza o varianza.

Definición 3.1 (Orden estocástico)**.** *Dadas dos variables aleatorias X,Y y una relación binaria $\succeq$ para la comparación de ambas denominada **orden estocástico**, diremos que:*

- *X es **estrictamente preferida** a Y con respecto a "$\succeq$", denotado por $X \succ Y$, si $X \succeq Y$ e $Y \not\succeq X$.*
- *X e Y son **indiferentes** con respecto a "$\succeq$", denotado por $X \equiv Y$, si $X \succeq Y$ e $Y \succeq X$.*
- *X e Y son **incomparables** con respecto a "$\succeq$", denotado por $X \nsim Y$, si $X \not\succeq Y$ ni $Y \not\succeq X$.*

Dado un conjunto de variables aleatorias, denotado por $\mathcal{D}$, entonces $(\mathcal{D}, \succ, \prec, \equiv, \nsim)$ determina una estructura de preferencia $\langle P, I, J \rangle$. Si además la incomparabilidad no es posible, es decir no existe ningún par de elementos incomparables, entonces $(\mathcal{D}, \succ, \prec, \equiv)$ forma una estructura de preferencia sin elementos incomparables (estructura de preferencia $\langle P, I \rangle$).

Observación 2. *Se definen las relaciones binarias* $\prec$*,* $\preceq$ *como:*

$$X \prec Y \iff Y \succ X, \quad X \preceq Y \iff Y \succeq X.$$

Observación 3. *En la definición anterior* $\succeq$ *es la relación binaria característica de la estructura de preferencia.*

Comencemos introduciendo este concepto con un ejemplo simple. Supongamos que tenemos dos tragaperras y queremos escoger cuál usar. Un primer enfoque sería escoger la máquina con mayor media, que de acuerdo con la ley de los grandes números sería el enfoque más beneficioso a largo plazo. Podríamos ir un paso más allá, y en caso de igualdad de medias escoger la máquina con menor varianza, pues esto nos dará menos altibajos.

Ejemplo 3.1. *Esta relación binaria la definiríamos como:*

$$X \succeq Y \ si \begin{cases} E(X) > E(Y), \\ \qquad\qquad ó \\ E(X) = E(Y) \quad y \quad Var(X) \leq Var(Y). \end{cases}$$

Esta relación es reflexiva y transitiva, mientras que no es ni simétrica ni antisimétrica (podríamos considerar dos variables aleatorias distintas con misma media y varianza). Además, sería una relación no completa, pues si las esperanzas no existen (y, por tanto, las varianzas tampoco): $X \not\succeq Y$ *y* $Y \not\succeq X$*.*

3.1. Dominancia Estocástica

El primer orden estocástico del que hablaremos se centra en la función de distribución y su ámbito más común de aplicación es el económico. Para introducirlo, supongamos que tenemos dos posibles inversiones modelizadas mediante dos variables aleatorias y queremos decidir cuál de las dos nos dará mayor rentabilidad; entonces estadísticamente escogeríamos aquella inversión que mayor probabilidad concentre en los valores altos. Esta es la idea subyacente de este orden estocástico.

Definición 3.2 (Dominancia estocástica)**.** *Sean* X *e* Y *dos variables aleatorias con funciones de distribución* F_X *y* F_Y *respectivamente. Se dice que* X ***domina a*** Y ***estocásticamente de primer grado****, y se denota como* $X \succeq_{FSD} Y$*, si:*

$$F_X(t) = P(X \leq t) \leq P(Y \leq t) = F_Y(t), \quad \forall\, t \in \mathbb{R}.$$

Si además existe un $t \in \mathbb{R}$ tal que $F_X(t) < F_Y(t)$, entonces diremos que X ***domina estrictamente a*** *Y* ***estocásticamente*** *de primer grado y lo denotaremos por $X \succ_{FSD} Y$.*

Observación 4. *Si bien se ha definido la dominancia estocástica de primer grado por medio de la función de distribución, en la literatura se puede encontrar también definida a partir de la función de supervivencia $\bar{F}(t) = 1 - F(t)$ en cuyo caso la condición para que X domine estocásticamente de primer grado a Y es:*

$$\bar{F}_X(t) \geq \bar{F}_Y(t), \quad \forall\, t \in \mathbb{R}.$$

Sin embargo, es directo comprobar que ambas formulaciones son equivalentes.

Observación 5. *La relación binaria $\succeq_{FSD}$ es transitiva y reflexiva.*

Algunos autores consideran el conjunto de alternativas como el conjunto de distribuciones que las variables aleatorias pueden tomar. En este caso, la dominancia estocástica es además antisimétrica. Para verlo, observar que si $X \succeq_{FSD} Y$ e $Y \succeq_{FSD} X$ entonces $F_X(t) = F_Y(t)$ para todo t real. Como la función de distribución caracteriza la distribución de una variable aleatoria, entonces X e Y son iguales en distribución. Además, al añadir la condición de antisimetría la dominancia estocástica pasaría a definir un orden parcial. Por un lado, la dominancia estocástica resulta bastante simple y fácilmente computable. Sin embargo, esta no determina una relación completa pues es posible encontrar variables aleatorias incomparables.

Ejemplo 3.2. *Consideremos $X \sim \mathcal{N}(0,1)$ e $Y \sim \mathcal{N}(-1,1)$. Dado que $F_Y(y) = F_X(y+1)$, entonces $F_Y(t) > F_X(t)$ para todo $t \in \mathbb{R}$, de modo que $X \succ_{FSD} Y$. Gráficamente, en la Figura 3.1 observamos que en ningún momento las funciones de distribución se cruzan.*

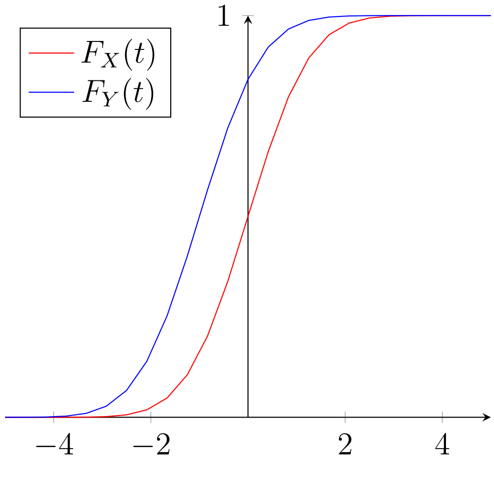

Figura 3.1: Funciones de distribución del Ejemplo 3.2 utilizadas para estudiar la dominancia estocástica.

Ejemplo 3.3. *Consideremos X una variable aleatoria degenerada en 0.4 e $Y \sim \mathcal{B}(0.2)$. Entonces observamos que $0 = F_X(0.2) < F_Y(0.2) = 0.8$, mientras que $1 = F_X(0.4) > F_Y(0.4) = 0.8$, de modo que $X \not\succ_{FSD} Y$ e $Y \not\succ_{FSD} X$. De este modo se concluye que estas dos variables son incomparables.*

Con estos dos ejemplos hemos visto, por un lado, un caso donde la dominancia era estricta y, por otro lado, uno donde la comparabilidad no era posible. La incomparabilidad supone un inconveniente cuando se busca comparar variables aleatorias, y por ese motivo existen otros grados de dominancia estocástica más débiles. Estos siguen la misma idea que la dominancia estocástica anteriormente descrita, dar mayor importancia a las regiones donde la variable aleatoria toma valores altos.

Definición 3.3 (Dominancia estocástica de segundo grado)**.** *Sean X e Y variables aleatorias con funciones de distribución F_X y F_Y respectivamente. Se dice que X domina a Y estocásticamente de segundo grado, y se denota como $X \succeq_{SSD} Y$, si:*

$$G_X^2(t) = \int_{-\infty}^{t} F_X(x)dx \leq \int_{-\infty}^{t} F_Y(y)dy = G_Y^2(t), \quad \forall\, t \in \mathbb{R}.$$

Del mismo modo que en la dominancia estocástica de primer grado, si además existe un t en $\mathbb{R}$ tal que $G_X^2(t) < G_Y^2(t)$, entonces diremos que X domina a Y estocásticamente de segundo grado y lo denotaremos por $X \succ_{SSD} Y$.

Esta definición tiene de nuevo el problema de la incomparabilidad, sin embargo, al ser menos restrictiva que la primera, existirán variables aleatorias comparables por segundo grado, pero no por primero.

Observación 6. *Dadas dos variables aleatorias X, Y entonces se cumple que:*

$$X \succ_{FSD} Y \implies X \succ_{SSD} Y.$$
$$X \equiv_{FSD} Y \implies X \equiv_{SSD} Y.$$

Veamos ahora un ejemplo donde utilizar este orden estocástico más débil resulta de utilidad.

Ejemplo 3.4. *Consideremos una variable aleatoria $X \sim U(0,1)$ e Y una variable aleatoria con función de densidad $f_Y(y) = \frac{3}{8}(y+1)^2 \mathbb{1}_{[-1,1]}(y)$. De este modo:*

$$F_X(x) = \begin{cases} 0, & \text{si } x < 0, \\ x, & \text{si } x \in [0,1), \\ 1, & \text{si } x \geq 1. \end{cases} \qquad F_Y(y) = \begin{cases} 0, & \text{si } y < -1, \\ \frac{1}{8}(x+1)^3, & \text{si } y \in [-1,1), \\ 1, & \text{si } y \geq 1. \end{cases}$$

Como $F_X(-0.5) = 0 < \frac{1}{64} = F_Y(-0.5)$*, mientras que* $F_X(0.5) = 0.5 > \frac{27}{64} = F_Y(0.5)$*, no se tiene dominancia estocástica. Ahora, integrando:*

$$G_X^2(x) = \begin{cases} 0, & \text{si } x < 0, \\ \frac{1}{2}x^2, & \text{si } x \in [0,1), \\ x - \frac{1}{2}, & \text{si } x \geq 1. \end{cases} \qquad G_Y^2(y) = \begin{cases} 0, & \text{si } y < -1, \\ \frac{1}{32}(y+1)^4, & \text{si } y \in [-1,1), \\ y - \frac{1}{2}, & \text{si } y \geq 1. \end{cases}$$

Es claro que en $(-1,1)$*,* $G_Y^2(t) > G_X^2(t)$*. Por otro lado, si calculamos los puntos de corte en* $[0,1]$*, obtenemos que únicamente se cortan en* $x = 1$*, de modo que en* $[0,1)$ *se cumple* $G_Y^2(y) > G_X^2(x)$*. Por último, en* $[1,\infty)$ *es claro que* $G_X^2(x) = G_Y^2(y)$*. Por tanto comprobamos que* $X \succ_{SSD} Y$*.*
Gráficamente, en la Figura 3.2 observamos que cuando trabajamos a primer orden existen dos intervalos, uno donde $F_Y(t) > F_X(t)$ *y otro donde* $F_Y(t) < F_X(t)$*. Mientras que cuando trabajamos a orden dos* $G_Y(t) \geq G_X(t)$ *para todo* t *en la recta real, siendo la desigualdad estricta en el intervalo (-1,1).*

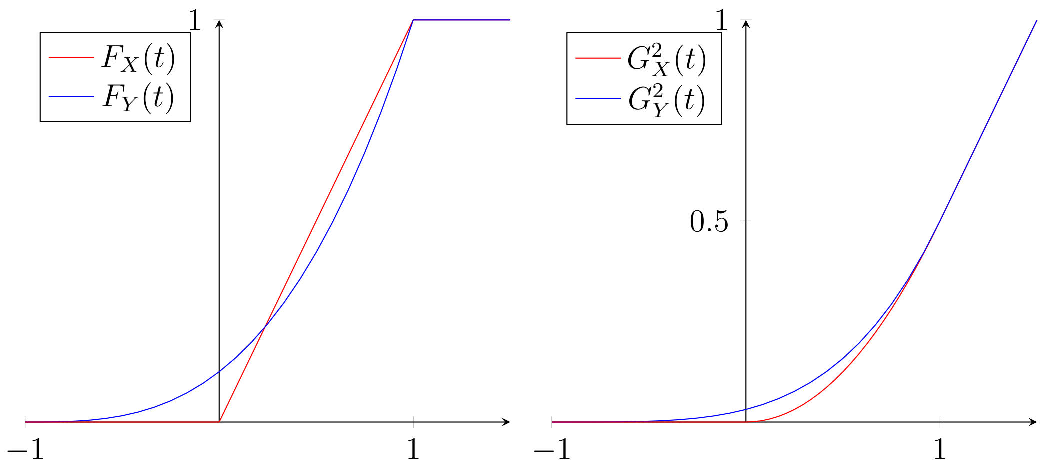

Figura 3.2: Funciones de distribución y sus primitivas del Ejemplo 3.4 utilizadas para estudiar la dominancia estocástica de primer y segundo grado.

Del mismo modo que hemos definido el segundo orden estocástico, podríamos ir definiendo órdenes más altos y, por tanto, más débiles. Para ello, debemos en primer lugar definir la función n-ésima acumulada. Esta se define recursivamente como sigue.

Definición 3.4 (Función n-ésima acumulada)**.** *Dada una variable aleatoria* X *y su función de distribución acumulada* F_X*, su función* n*-ésima acumulada se define como:*

$$G_X^n(t) = \int_{(-\infty,t]} G_X^{n-1}(x)dx = \int_{(-\infty,t]\times\mathbb{R}^{n-2}} F_X(x)(dx)^{n-1}, \quad n \geq 1, \ t \in \mathbb{R}.$$

En particular, para $n = 1$*,* $G_X^1 = F_X$*.*

Definición 3.5 (Dominancia estocástica de orden n). *Consideremos X e Y dos variables aleatorias con funciones de distribución F_X y F_Y respectivamente. Entonces, diremos que X domina a Y estocásticamente por grado n, y se denota como $X \succeq_{nSD} Y$, si se cumple:*

$$G_X^n(t) \leq G_Y^n(t), \ \ \forall\, t \in \mathbb{R}.$$

Además, diremos que X domina estrictamente a Y estocásticamente por grado n, y se denota como $X \succ_{nSD} Y$, si además se cumple:

$$G_X^n(t) < G_Y^n(t), \quad \textit{para algún } t \in \mathbb{R}.$$

De nuevo, en estas definiciones tenemos el problema de incomparabilidad, al igual que habíamos tenido con las anteriores definiciones. Observamos también que cuando $n = 2$, recuperamos la dominancia por segundo grado y cuando $n = 1$ la dominancia de primer grado o estocástica.

Observación 7. *Dadas dos variables aleatorias X e Y y dos números naturales n, m con $n<m$, entonces:*

$$X \succ_{nSD} Y \implies X \succ_{mSD} Y,$$
$$X \equiv_{nSD} Y \implies X \equiv_{mSD} Y.$$

Una de las propiedades más importantes de la dominancia estocástica es la caracterización a través de las esperanzas. Es más, estas nos van a permitir caracterizar el grado n-ésimo de dominancia. El siguiente resultado revela esta relación.

Teorema 3.1. *[11, 17][Caracterización de la dominancia estocástica por las medias] Dadas X e Y variables aleatorias con esperanzas finitas, entonces se cumple:*

1. $X \succeq_{FSD} Y \iff E(u(X)) \geq E(u(Y)),\ \forall u : \mathbb{R} \to \mathbb{R}$ *creciente.*

2. $X \succeq_{SSD} Y \iff E(u(X)) \geq E(u(Y)),\ \forall u : \mathbb{R} \to \mathbb{R}$ *creciente y cóncava (2-creciente).*

3. $X \succeq_{nSD} Y \iff E(u(X)) \geq E(u(Y)),\ \forall u : \mathbb{R} \to \mathbb{R}$ *n-creciente (Definición 2.13).*

Esta caracterización a través de las esperanzas podría recordar al Ejemplo 3.1 donde argumentábamos que de acuerdo con la ley de los grandes números escoger una estrategia que tenga mayor esperanza era razonable. En este caso, llevaríamos esta idea un paso más allá exigiéndolo para cualquier transformación creciente la cual nos conservaría el orden,

es decir, cualquier transformación tal que si $x > y$ entonces $g(x) \geq g(y)$.
Por último, existen algunas propiedades relevantes de la dominancia estocástica que son mencionadas a continuación.

Proposición 3.2. *[17, Teorema 1.2.13] Dadas X e Y variables aleatorias tales que $X \succeq_{FSD} Y$ y dada una función creciente $\varphi : \mathbb{R} \to \mathbb{R}$, entonces $\varphi(X) \succeq_{FSD} \varphi(Y)$.*

Proposición 3.3. *[17, Teorema 1.2.17] Dadas las variables aleatorias $\{X_k\}_{k=1}^n$, $\{Y_k\}_{k=1}^n$ independientes, si $X_k \succ_{FSD} Y_k$ para todo $k = 1, ..., n$, entonces $X_1 + ... + X_n \succ_{FSD} Y_1 + ... + Y_n$.*

Los resultados anteriores indican que la dominancia estocástica se preserva por transformaciones crecientes y por agregación (bajo independencia). El tercer y último resultado indica que también se preserva bajo la convergencia en ley.

Proposición 3.4. *[17, Teorema 1.2.14] Dadas las variables aleatorias $\{X_n\}_{n\in\mathbb{N}}$, $\{Y_n\}_{n\in\mathbb{N}}$ tales que $X_n \xrightarrow{\mathcal{L}} X$, $Y_n \xrightarrow{\mathcal{L}} Y$, si $X_n \succ_{FSD} Y_n$ para todo $n \in \mathbb{N}$, entonces $X \succ_{FSD} Y$.*

3.2. Preferencia Estadística

Hemos visto un primer acercamiento al orden estocástico con la dominancia estocástica y sus versiones más débiles de grado n. Sin embargo, todos estos órdenes presentan un gran problema: existen variables aleatorias incomparables. Además, aunque se puedan definir órdenes sucesivos cada vez más débiles mediante la dominancia estocástica de grado-n, a medida que aumenta el grado, el coste computacional crece con él significativamente.
Por ello introducimos en esta sección un nuevo orden estocástico, la preferencia estadística, que no presenta estos inconvenientes. Este orden estocástico fue introducido por De Schuymer et al. en 2003 [6, 7] y se define a partir de lo que se conoce como relación probabilística.

Definición 3.6 (Relación probabilística)**.** *Dada una serie de alternativas $\mathcal{D}$, una relación probabilística es una aplicación $\mathcal{Q} : \mathcal{D} \times \mathcal{D} \to [0,1]$, tal que $\mathcal{Q}(a,b) + \mathcal{Q}(b,a) = 1$ para cualquier par de alternativas $a, b \in \mathcal{D}$.*

En este caso, como estamos trabajando con variables aleatorias, el conjunto de alternativas $\mathcal{D}$ estará formado por variables aleatorias definidas en el mismo espacio de probabilidad $(\Omega, \mathcal{A}, P)$ y tomando valores en un espacio ordenado $(\Omega', \mathcal{A}')$.

Definición 3.7. *Dadas dos variables aleatorias X e Y definimos la aplicación:*

$$\begin{aligned}\mathcal{Q}(X,Y) :&\mathcal{D} \times \mathcal{D} \to [0,1] \\ &(X,Y) \hookrightarrow P(X > Y) + \frac{1}{2}P(X = Y).\end{aligned}$$

Proposición 3.5. *La aplicación $\mathcal{Q}$ definida previamente es una relación probabilísitica.*

Demostración. Esta aplicación toma valores en $[0,1]$, por tanto queda ver que $\mathcal{Q}(X,Y)+\mathcal{Q}(Y,X)=1$. Efectivamente, se tiene que:

$$\mathcal{Q}(X,Y)+\mathcal{Q}(Y,X)=P(X>Y)+P(Y>X)+P(X=Y)=1. \qquad \square$$

A través de esta aplicación vamos a definir un orden estocástico. Siguiendo la línea de la dominancia estocástica buscamos que X sea preferida a Y cuando la primera tienda a tomar valores más altos.

Definición 3.8 (Preferencia estadística)**.** *Dadas dos variables X e Y, se dice que:*

- *X es **preferida estadísticamente** a Y, y se denota por $X \succeq_{SP} Y$, si $\mathcal{Q}(X,Y) \geq 0.5$.*
- *X es **indiferente estadísticamente** a Y, y se denota por $X \equiv_{SP} Y$, si $\mathcal{Q}(X,Y) = 0.5$.*
- *X es **estrictamente preferida estadísticamente** a Y, y se denota por $X \succ_{SP} Y$, si $\mathcal{Q}(X,Y) > 0.5$.*

Ejemplo 3.5. *Volvamos ahora al Ejemplo 3.3. Habíamos visto que X e Y eran incomparables si considerábamos como orden estocástico la dominancia estocástica. Veamos ahora qué ocurre si consideramos la preferencia estadística:*

X/Y	0	1
0.4	0.8	0.2

$$\mathcal{Q}(X,Y)=P(X>Y)+\frac{1}{2}P(X=Y)=0.8 \Rightarrow X \succ_{SP} Y.$$

Observamos que la preferencia estadística nos permite comparar dos variables aleatorias que no podíamos comparar mediante la dominancia estocástica.

Ejemplo 3.6. *En la Figura 3.3 podemos observar la interpretación geométrica de la preferencia estadística, el plano $X=Y$ separa el espacio en dos regiones: una donde $X>Y$ y otra donde $X<Y$. Si el volumen de la función de densidad en la región donde $X>Y$ es mayor que el de la región donde $X<Y$ decimos que X es preferida estadísticamente a Y. Si, en cambio, ocurre que el volumen de la segunda región es mayor que el de la primera, entonces Y es preferida a X. Por último, si ambas regiones tienen el mismo volumen entonces son indiferentes. En este caso X es preferida a Y.*

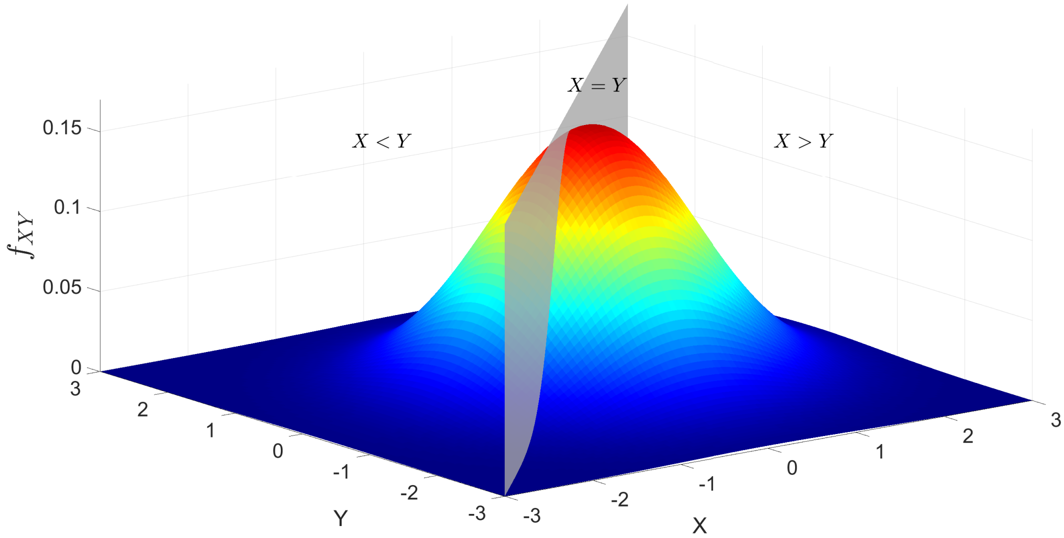

Figura 3.3: Interpretación geométrica de la preferencia estadística.

Este nuevo orden estocástico no deja lugar a la incomparabilidad, de modo que $(\mathcal{D}, \succ_{SP}, \equiv_{SP})$ determina una estructura de preferencia sin elementos incomparables. Esta incomparabilidad es consecuencia de la relación probabilística pues al tomar valores entre 0 y 1 siempre podemos establecer una desigualdad.

Observación 8. *La preferencia estadística constituye una relación binaria reflexiva y completa. Sin embargo, no es transitiva ni antisimétrica.*

Esta falta de transitivad hace que la preferencia estadística no defina ningún orden de los que se suelen considerar. El siguiente ejemplo pone de manifiesto la no transitividad de esta relación binaria.

Ejemplo 3.7. *Consideremos el espacio de probabilidad* $(\Omega, \mathcal{A}, P)$ *con* $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$ *y tres variables aleatorias* X, Y, Z*:*

Ω	ω_1	ω_2	ω_3	ω_4
P	0.1	0.2	0.3	0.4
X	-1	-1	1	1
Y	0	1	1	0.5
Z	2	0	2	0

De este modo:

$$\begin{aligned}
\mathcal{Q}(X,Y) &= P(X=1, Y=0.5) & +\frac{1}{2}P(X=1,Y=1) &= 0.55 & \Longrightarrow\ & X \succ_{SP} Y,\\
\mathcal{Q}(Y,Z) &= P(Y=1, Z=0) & +P(Y=0.5,Z=0) &= 0.6 & \Longrightarrow\ & Y \succ_{SP} Z,\\
\mathcal{Q}(X,Z) &= P(X=1, Z=0) & &= 0.4 & \Longrightarrow\ & X \prec_{SP} Z.
\end{aligned}$$

Por tanto, se comprueba que no hay transitividad.

El uso de una relación probabilística permite cuantificar la preferencia entre $[0,1]$ (Figura 3.4), cuanto más cercana sea $\mathcal{Q}(X,Y)$ a 1 más preferida será X a Y, mientras que cuanto más cerca esté de 0 mayor será la preferencia por Y, estableciendo en 0.5 la indiferencia entre ambas variables. Esto supone un nuevo valor añadido respecto a la dominancia estocástica, donde únicamente se establecía una preferencia.

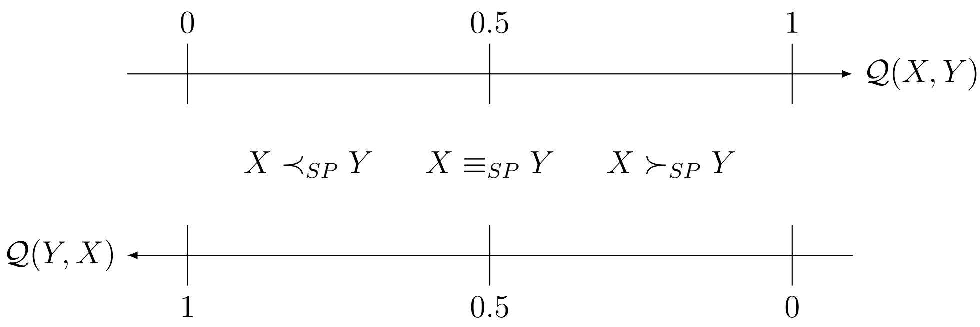

Figura 3.4: La preferencia estadística nos permite ponderar cúan preferida es una variable frente a otra en un rango de 0 a 1.

Utilizando propiedades sencillas de probabilidad podemos obtener una primera caracterización de la preferencia estadística:

Proposición 3.6. *[13, Lema 2.2] Dadas dos variables aleatorias X e Y, son equivalentes:*

(I) $X \succeq_{SP} Y$.

(II) $\mathcal{Q}(X,Y) \geq \mathcal{Q}(Y,X)$.

(III) $P(X \geq Y) \geq P(Y \geq X)$.

(IV) $P(X > Y) \geq P(Y < X)$.

Demostración. $\boxed{i) \to ii)}$ Utilizando que $Q(X,Y) = 1 - Q(Y,X)$:

$$X \succeq_{SP} Y \iff Q(X,Y) \geq 0.5 \iff Q(Y,X) \leq 0.5.$$

$\boxed{ii) \to iii)}$

$$\mathcal{Q}(X,Y) = P(X > Y) + \frac{1}{2}P(X = Y) = P(X \geq Y) - \frac{1}{2}P(X = Y).$$
$$\mathcal{Q}(Y,X) = P(Y > X) + \frac{1}{2}P(Y = X) = P(Y \geq X) - \frac{1}{2}P(Y = X).$$

Por hipótesis $\mathcal{Q}(X,Y) \geq \mathcal{Q}(Y,X)$, por otro lado $P(X = Y) = P(Y = X)$, de donde se deduce el resultado.

$\boxed{iii) \to iv)}$

$$P(X > Y) = P(X \geq Y) - P(X = Y).$$
$$P(Y > X) = P(Y \geq X) - P(Y = X).$$

Como por hipótesis $P(X \geq Y) \geq P(Y \geq X)$ y $P(X = Y) = P(Y = X)$, se deduce el resultado.

$\boxed{iv) \to i)}$

$$\mathcal{Q}(X,Y) = P(X > Y) + \frac{1}{2}P(X = Y).$$
$$\mathcal{Q}(Y,X) = P(Y > X) + \frac{1}{2}P(Y = X).$$

Como $P(X = Y) = P(Y = X)$, se deduce que $\mathcal{Q}(X,Y) \geq \mathcal{Q}(Y,X)$. Por otro lado $\mathcal{Q}(X,Y) + \mathcal{Q}(Y,X) = 1$. De modo donde se obtiene que $\mathcal{Q}(X,Y) \geq 0.5$. □

Vimos cómo la dominancia estocástica se conservaba bajo ciertas transformaciones, en particular bajo funciones crecientes. En la preferencia estadística podemos encontrar un resultado similar.

Proposición 3.7. *[14, Proposición 9] Dadas dos variables aleatorias X e Y y una función $\varphi : \mathbb{R} \to \mathbb{R}$, entonces:*

1. *Si φ es estrictamente creciente:*

$$X \succeq_{SP} Y \iff \varphi(X) \succeq_{SP} \varphi(Y).$$

2. *Si φ es estrictamente decreciente:*

$$X \succeq_{SP} Y \iff \varphi(X) \preceq_{SP} \varphi(Y).$$

Demostración. Se probará para el caso de funciones estrictamente decrecientes, el otro caso es análogo.

$\boxed{\Rightarrow}$ Por ser φ estrictamente decreciente:

$$\{\omega \,:\, X(\omega) > Y(\omega)\} = \{\omega \,:\, \varphi(X(\omega)) < \varphi(Y(\omega))\}.$$
$$\{\omega \,:\, X(\omega) < Y(\omega)\} = \{\omega \,:\, \varphi(X(\omega)) > \varphi(Y(\omega))\}.$$

De modo que $P(X > Y) = P(\varphi(Y) > \varphi(X))$ y $P(Y < X) = P(\varphi(X) > \varphi(Y))$.
Ahora, utilizando la caracterización anterior, $X \succeq_{SP} Y \iff P(X > Y) \geq P(Y > X)$ de donde se deduce que:

$$P(\varphi(Y) > \varphi(X)) > P(\varphi(X) > \varphi(Y)).$$

Y de nuevo utilizando la caracterización se deduce la implicación.
$\boxed{\Leftarrow}$ Por hipótesis $\varphi(X) \preceq_{SP} \varphi(Y)$ para cualquier función φ estrictamente decreciente, en particular, para $\varphi(t) = -t$. Utilizando la caracterización anterior:

$$-Y \succeq_{SP} -X \iff P(-Y \geq -X) \geq P(-X \geq -Y) \iff P(Y \leq X) \geq P(X \leq Y).$$

Donde una vez más, utilizando la caracterización de la Proposición 3.6, se llega al resultado. $\square$

Por otro lado, la preferencia estadística tiene una importante relación con la mediana de $X - Y$.

Teorema 3.8. *[14, Teorema 15] Dadas X e Y variables aleatorias definidas en un mismo espacio de probabilidad, entonces:*

1. $\sup Me(X-Y) > 0 \implies X \succeq_{SP} Y \implies \sup Me(X-Y) \geq 0$.

2. $X \succ_{SP} Y \implies Me(X-Y) \subseteq [0, \infty)$.

3. *El recíproco de 2. no es cierto, sin embargo:*

$$\text{ínf}\, Me(X-Y) > 0 \implies X \succ_{SP} Y.$$

4. *Si $P(X = Y) = 0$, entonces:*

$$X \succ_{SP} Y \iff \text{ínf}\, Me(X-Y) > 0.$$

Sin embargo, aunque $P(X = Y) = 0$, que $Me(X-Y)$ contenga al 0 no es equivalente a $\mathcal{Q}(X,Y) = 0.5$.

Esta relación con la mediana es consecuencia de cómo se define la preferencia estadística donde X es preferida a Y si $P(X - Y > 0) > P(X - Y < 0)$, por tanto, resulta lógico

que en caso de que sea preferida la mediana sea al menos mayor o igual que cero.
Por último, resultaría útil establecer una conexión entre la preferencia estadística y la dominancia estocástica. Además, resultaría razonable que esta conexión se diera, pues si la idea de la dominancia estocástica era dar mayor probabilidad a los valores más altos, la condición de la preferencia estadística de $P(X > Y) \geq P(Y < X)$ da una idea similar de este orden estocástico.
El siguiente teorema permite establecer bajo qué condiciones se da esta relación entre estos dos órdenes estocásticos.

Teorema 3.9. *[16, Teorema 45] Dadas X e Y variables aleatorias definidas en el mismo espacio de probabilidad, entonces $X \succeq_{FSD} Y$ implica $X \succeq_{SP} Y$ bajo cualquiera de las siguientes condiciones:*

1. *X e Y son independientes.*
2. *X e Y son variables aleatorias absolutamente continuas contramonótonas.*
3. *X e Y son variables aleatorias absolutamente continuas comonótonas.*
4. *X e Y son variables aleatorias discretas con soportes finitos contramonótonas.*
5. *X e Y son variables aleatorias discretas con soportes finitos comonótonas.*
6. *X e Y son variables aleatorias absolutamente continuas relacionadas mediante una cópula arquimediana estricta con generador dos veces diferenciable.*
7. *X e Y son variables aleatorias absolutamente continuas relacionadas mediante una cópula arquimediana nilpotente con generador dos veces diferenciable.*

Además, $X \succ_{FSD} Y$ implica $X \succ_{SP} Y$ en los casos 1, 3, 5 *y* 6

Sin embargo, este resultado no es cierto en general.

Ejemplo 3.8. *Consideremos dos variables aleatorias X, Y distribuidas de la siguiente manera:*

X/Y	-1	0	1	
-1	0	0.2	0	0.2
0	0	0.1	0.1	0.2
1	0.2	0	0.4	0.6
	0.2	0.3	0.5	

En primer lugar calculamos las funciones de distribución:

$$F_X(t) = \begin{cases} 0, & si \quad t < -1 \\ 0.2, & si \quad -1 \leq t < 0 \\ 0.4, & si \quad 0 \leq t < 1 \\ 1, & si \quad t \geq 1. \end{cases} \qquad F_Y(t) = \begin{cases} 0, & si \quad t < -1 \\ 0.2, & si \quad -1 \leq t < 0 \\ 0.5, & si \quad 0 \leq t < 1 \\ 1, & si \quad t \geq 1. \end{cases}$$

De modo que $F_X(t) \leq F_Y(t)$ para todo t real, además en $t = 0$ se tiene que $F_X(t) < F_Y(t)$, de modo que $X \succ_{FSD} Y$.
Por otro lado:

$$P(X > Y) = P(X = 0, Y = -1) + P(X = 1, Y = -1) + P(X = 1, Y = 0) = 0.2.$$
$$P(X < Y) = P(X = -1, Y = 0) + P(X = -1, Y = 1) + P(X = 0, Y = 1) = 0.3.$$

Por tanto, tenemos que $Y \succ_{SP} X$.

Cuando introdujimos las cópulas en el Capítulo 2, enunciamos la cota de Fréchet-Hoeffding (Teorema 2.3) que nos daba los dos casos más extremos de dependencia. Resulta interesante estudiar la preferencia estadística en estos contextos tan particulares.

Proposición 3.10. *[5, Proposición 7] Dadas X e Y dos variables aleatorias absolutamente continuas contramonótonas con funciones de distribución F_X y F_Y y funciones de densidad f_X, f_Y, entonces:*

$$\mathcal{Q}(X, Y) = \int_{x: F_X(x) + F_Y(x) \geq 1} f_X(x)\, dx,$$

ó, equivalentemente

$$\mathcal{Q}(X, Y) = F_Y(u) \text{ con } u \text{ tal que } F_X(u) + F_Y(u) = 1.$$

Esta caracterización es consecuencia de la dependencia totalmente negativa en la contramonotonía, es decir $Y = T(X)$ con T una función decreciente. Vamos a intentar dar una explicación intuitiva, para ello vamos a suponer que nos encontramos en el caso más sencillo en el que T es estrictamente decreciente y por tanto posee inversa y F_X,F_Y son estrictamente crecientes.
En primer lugar, supongamos que T tiene un punto fijo u, entonces la aplicación T cumple:

$$\begin{aligned} T(x) &> x, \quad \text{si} \quad x < u, \\ T(x) &= x, \quad \text{si} \quad x = u, \\ T(x) &< x, \quad \text{si} \quad x > u. \end{aligned}$$

Debido a que Y es una transformación estrictamente decreciente de X, entonces u constituirá la frontera entre $X > Y$ e $Y > X$ por lo que acabamos de ver. Así:

$$\{\omega \in \Omega : X(\omega) > Y(\omega)\} = \{\omega \in \Omega : X(\omega) > T(X(\omega))\} = \{\omega \in \Omega : X(\omega) > u\},$$
$$\{\omega \in \Omega : X(\omega) < Y(\omega)\} = \{\omega \in \Omega : X(\omega) < T(X(\omega))\} = \{\omega \in \Omega : X(\omega) < u\}.$$

Además este punto fijo cumple que $F_X(u) + F_Y(u) = 1$:

$$\left.\begin{aligned} &F_X(u) = P(X \leq u) \\ &F_Y(u) = P(T(X) \leq u) = P(X \geq T^{-1}(u)) = P(X \geq u) \end{aligned}\right\} F_X(u) + F_Y(u) = 1.$$

Con esto podemos ver de dónde viene la primera expresión, la equivalencia con la segunda se sigue porque

$$\mathcal{Q}(X,Y) = P(X > Y) = P(X > u) = 1 - F_X(u) = F_Y(u),$$

donde se ha utilizado que $F_X(u) + F_Y(u) = 1$ y por ser variables aleatorias absolutamente continuas, $P(X = Y) = 0$.
Si T no tuviera un punto fijo, entonces podrían darse dos casos $X(\omega) > Y(\omega)$ ó $X(\omega) < Y(\omega)$ para todo $\omega \in \Omega$. En ambos casos además las variables aleatorias tendrían soportes disjuntos, esto es consecuencia de que T no tenga ningún punto fijo.
Supongamos que $X(\omega) > Y(\omega)$ para todo ω de Ω. Si denotamos por $Sop(X)$ y $Sop(Y)$ el conjunto de puntos donde f_X y f_Y son estrictamente positivas, se cumple sup $Sop(Y) <$ ínf $Sop(X)$, y por tanto

$$F_X(t) + F_Y(t) = \begin{cases} 0, & \text{si } t < \text{ínf } Sop(Y), \\ F_Y(t), & \text{si } t \in Sop(Y), \\ 1, & \text{si } t \in \Big[\sup Sop(Y), \text{ínf } Sop(X)\Big], \\ 1 + F_X(t), & \text{si } t \in Sop(X), \\ 2, & \text{si } t \geq \sup Sop(X). \end{cases}$$

De ahí se sigue que en la Proposición 3.10 se cumpla que $\mathcal{Q}(X,Y) = 1$. Además, tomando u en el intervalo $\Big[\sup Sop(Y), \text{ínf } Sop(X)\Big]$ se tendría que $F_X(u) + F_Y(u) = 1$, $F_X(u) = 0$ y $F_Y(u) = 1$.
El caso en que $X(\omega) < Y(\omega)$ para todo ω de Ω sería totalmente análogo.
Gráficamente, podemos interpretar esta fórmula en la Figura 3.5 donde se observa que en u, $t_1 + t_2 = 1$.

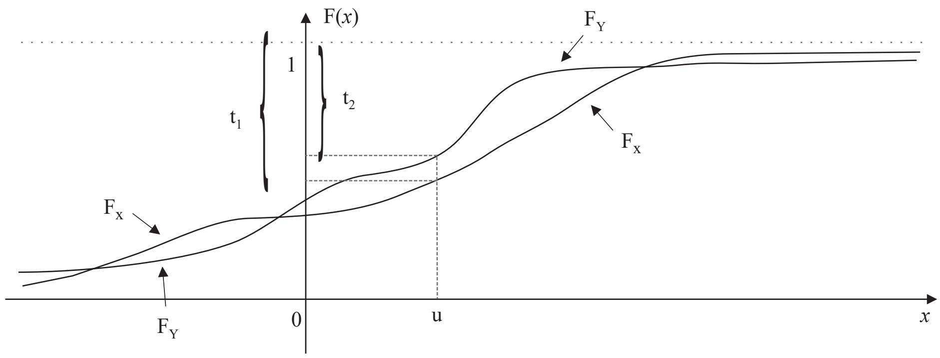

Figura 3.5: Interpretación gráfica de $\mathcal{Q}(X,Y)$ para dos variables contramonótonas. Créditos: DeBaets et al. (2007) [5].

Proposición 3.11. *[5, Proposición 6] Dadas X e Y dos variables aleatorias absolutamente continuas comonótonas con funciones de distribución F_X y F_Y y funciones de densidad f_X y f_Y, entonces:*

$$\mathcal{Q}(X,Y) = \int_{x:F_X(x)<F_Y(y)} f_X(x)dx + \frac{1}{2}\int_{x:F_X(x)=F_Y(y)} f_X(x)dx.$$

En particular,

$$P(X > Y) = \int_{x:F_X(x)<F_Y(y)} f_X(x)dx,$$
$$P(X = Y) = \int_{x:F_X(x)=F_Y(y)} f_X(x)dx.$$

Esta fórmula es consecuencia de la dependencia totalmente positiva en la comonotonía donde Y es una transformación T creciente de X. Vamos a dar una explicación intuitiva de este resultado, para ello, supondremos nos encontramos en el caso más sencillo donde T es estrictamente creciente y por tanto tiene inversa y F_X y F_Y también son estrictamente crecientes. Ahora:

$$\{\omega : X(\omega) > Y(\omega)\} = \{\omega : X(\omega) > T\big(X(\omega)\big)\}.$$

Para ver qué relación tiene esta expresión con la función de distribución, consideremos un intervalo (a,b) donde $F_X > F_Y$ y $F_X(a) = F_Y(a) \equiv F_a$, $F_X(b) = F_Y(b)$. Entonces como $F_X(a) = F_Y(b)$ se tiene que:

$$\left.\begin{aligned} &F_X(a) = P(X \leq a) \\ &F_Y(a) = P(Y \leq a) = P(T(X) \leq a) = P\big(X \leq T^{-1}(a)\big) \end{aligned}\right\} T^{-1}(a) = a.$$

Es decir, a sería un punto fijo de T. Por tanto en la región (a,b) donde $F_X > F_Y$, cualquier subintervalo (a, t) cumple estar en $\{X(\omega) : X(\omega) > Y(\omega)\}$ ya que como $F_X(t) > F_Y(t)$ y

$$F_X(t) = P(X \in (a,t)) + F_X(a) = P(a < X < t) + F_a,$$
$$F_Y(t) = P(Y \in (a,t)) + F_Y(a) = P(a < T(X) < t) + F_a = P(a \leq X \leq T^{-1}(t)) + F_a.$$

Entonces se tiene necesariamente que $t > T^{-1}(t)$. Es por esto que la fórmula lo que nos dice es que únicamente es necesario fijarse en lo que ocurre entre cada par de puntos de corte de F_X y F_Y consecutivas (Figura 3.6) para estudiar las condiciones de $X > Y$ y $X = Y$.

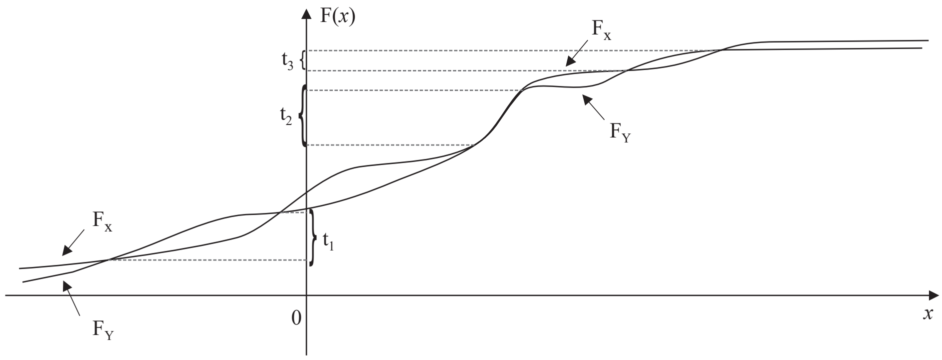

Figura 3.6: Interpretación gráfica de $\mathcal{Q}(X, Y)$ para variables comonótonas. Créditos: De-Baets et al. (2007) [5].

3.3. Caracterizaciones de la dominancia estocástica

En esta sección se darán caracterizaciones de la dominancia estocástica para distribuciones de la misma familia. Estas, nos permitirían además estudiar la preferencia estadística en los casos descritos en el Teorema 3.9.
Comenzamos dando un resultado para distribuciones normales.

Proposición 3.12. *Dadas dos variables aleatorias* $X \sim \mathcal{N}(\mu_X, \sigma)$ *e* $Y \sim \mathcal{N}(\mu_Y, \sigma)$, *entonces:*

$$X \succ_{FSD} Y \iff \mu_X > \mu_Y,$$
$$X \equiv_{FSD} Y \iff \mu_X = \mu_Y,$$
$$X \succeq_{FSD} Y \iff \mu_X \geq \mu_Y.$$

Demostración. Estandarizando (Φ denota la función de distribución de una distribución normal estándar), $F_X(x) = \Phi\Big(\dfrac{x - \mu_X}{\sigma}\Big)$, $F_Y(y) = \Phi\Big(\dfrac{y - \mu_Y}{\sigma}\Big)$. De modo que:

$$F_X(t) > F_Y(t) \iff \Phi\Big(\frac{t - \mu_X}{\sigma}\Big) > \Phi\Big(\frac{t - \mu_Y}{\sigma}\Big) \iff \mu_X > \mu_Y.$$

La segunda parte, se sigue directamente dado que ambas están idénticamente distribuidas. □

Ejemplo 3.9. *En el caso de que las varianzas no sean iguales, el resultado no es cierto, basta considerar* $X \sim \mathcal{N}(0,1)$ *e* $Y \sim \mathcal{N}(0,2)$*. Entonces:*

$$F_X(x) = \Phi(x) \quad y \quad F_Y(y) = \Phi\left(\frac{y}{2}\right).$$

En particular, se cumple:

$$\begin{aligned} &\text{En } t\text{=2}, && F_X(2) = \Phi(2) > \Phi(1) = F_Y(2). \\ &\text{En } t\text{=-2}, && F_X(-2) = \Phi(-2) < \Phi(-1) = F_Y(-2). \end{aligned}$$

En la Figura 3.7 podemos observar las funciones de distribución, encontrando un punto de corte en $t = 0$*.*

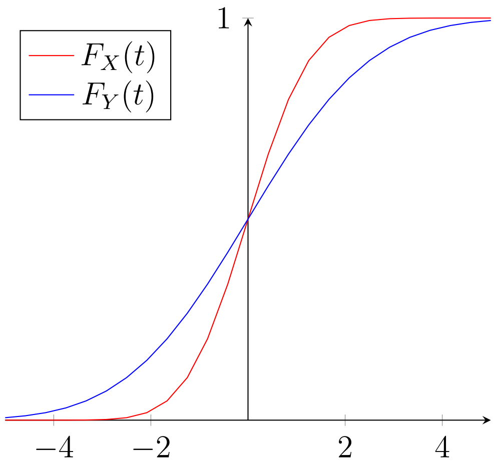

Figura 3.7: Funciones de distribución del Ejemplo 3.9.

Procedemos ahora con la distribución de Bernoulli.

Proposición 3.13. *Dadas* $X \sim \mathcal{B}(p_X)$ *e* $Y \sim \mathcal{B}(p_Y)$*, entonces:*

$$\begin{aligned} X \succ_{FSD} Y &\iff p_X > p_Y, \\ X \equiv_{FSD} Y &\iff p_X = p_Y, \\ X \succeq_{FSD} Y &\iff p_X \geq p_Y. \end{aligned}$$

Demostración. Calculamos en primer lugar las funciones de distribución:

$$F_X(x) = \begin{cases} 0, & \text{si } x < 0, \\ 1 - p_X, & \text{si } x \in [0,1), \\ 1, & \text{si } x \geq 1. \end{cases} \qquad F_Y(y) = \begin{cases} 0, & \text{si } y < 0, \\ 1 - p_Y, & \text{si } y \in [0,1), \\ 1, & \text{si } y \geq 1. \end{cases}$$

Entonces las funciones de distribución coinciden salvo en el intervalo $[0,1)$, donde $F_X(x)$ será menor que $F_Y(y)$ únicamente cuando $p_X > p_Y$. Si $p_X = p_Y$, entonces X e Y tienen la misma distribución y $X \equiv_{FSD} Y$. $\square$

Una vez considerado el caso más sencillo de las Bernoulli vamos a estudiar qué ocurre cuando consideramos la distribución binomial.

Proposición 3.14. *Dadas $X \sim \mathcal{B}(n, p_X)$ e $Y \sim \mathcal{B}(n, p_Y)$, entonces:*

$$\begin{aligned} X \succ_{FSD} Y &\iff p_X > p_Y, \\ X \equiv_{FSD} Y &\iff p_X = p_Y, \\ X \succeq_{FSD} Y &\iff p_X \geq p_Y. \end{aligned}$$

Demostración. El caso en que $p_X = p_Y$ es inmediato pues entonces X e Y están idénticamente distribuidas, de modo que $X \equiv_{FSD} Y$.
Vamos ahora a probar el caso con desigualdad estricta dado que la desigualdad con igual sería consecuencia de las otras dos.
$\boxed{\Leftarrow}$ Como una distribución binomial de parámetros n p se puede expresar como suma de n distribuciones de Bernoulli independientes de parámetro p:

$$\begin{aligned} X &= \sum_{k=1}^{n} X_k, \quad \text{con} \quad X_k \sim \mathcal{B}(p_X) \ \ \forall\, k = 1, ..., n. \\ Y &= \sum_{k=1}^{n} Y_k, \quad \text{con} \quad Y_k \sim \mathcal{B}(p_Y) \ \ \forall\, k = 1, ..., n. \end{aligned}$$

Por la Proposición 3.3 y la Proposición 3.13:

$$p_X > p_Y \iff \mathcal{B}(p_X) \succ_{FSD} \mathcal{B}(p_Y) \implies X \succ_{FSD} Y.$$

$\boxed{\Rightarrow}$ Por hipótesis, $F_X(t) \leq F_Y(t)$ para todo t real. Evaluemos las funciones de distribución en 0:

$$F_X(0) = P(X = 0) = (1 - p_X)^n, \quad F_Y(0) = P(Y = 0) = (1 - p_Y)^n. \tag{3.1}$$

Luego $F_X(t) \leq F_Y(t)$ si y solo si $p_X \geq p_Y$. Nos queda ver que la desigualdad es estricta. Observamos que si $p_X = p_Y$ entonces X e Y tienen la misma distribución luego $X \equiv_{FSD} Y$ lo que contradeciría la hipótesis de $X \succ_{FSD} Y$. Por tanto $p_X > p_Y$. $\square$

Ahora vamos a variar el parámetro correspondiente al número de repeticiones del experimento en la distribución binomial.

Proposición 3.15. *Dadas* $X \sim \mathcal{B}(n_X, p)$ *e* $Y \sim \mathcal{B}(n_Y, p)$*, entonces:*

$$\begin{aligned} X \succ_{FSD} Y &\iff n_X > n_Y. \\ X \equiv_{FSD} Y &\iff n_X = n_Y, \\ X \succeq_{FSD} Y &\iff n_X \geq n_Y. \end{aligned}$$

Demostración. El caso en que $n_X = n_Y$ es inmediato pues entonces X e Y están idénticamente distribuidas, de modo que $X \equiv_{FSD} Y$.
Vamos ahora a probar el caso con desigualdad estricta dado que la desigualdad con igual sería consecuencia de las otras dos. Expresando de nuevo la distribución binomial como suma de distribuciones de Bernoulli independientes:

$$\begin{aligned} X &= \sum_{k=1}^{n_X} X_k \quad \text{con} \quad X_k \sim \mathcal{B}(p) \ \ \forall\, k = 1, ..., n_X. \\ Y &= \sum_{k=1}^{n_Y} Y_k \quad \text{con} \quad Y_k \sim \mathcal{B}(p) \ \ \forall\, k = 1, ..., n_Y. \end{aligned}$$

$\boxed{\leftarrow}$ Supongamos en adelante que $n_X > n_Y$. Definimos la variable aleatoria degenerada en 0: $D \sim 0_{\mathcal{D}}$ independiente de $X_{k_X} \sim \mathcal{B}(p)$ y de $Y_{k_Y} \sim \mathcal{B}(p)$ para todo $k_X \in \{1, ..., n_X\}$, $k_Y \in \{1, ..., n_Y\}$. De modo que $X_{k_X} \succ_{FSD} 0_{\mathcal{D}}$ para todo $k_X \in \{1, ..., n_X\}$. Tomando una m.a.s. de $0_{\mathcal{D}}$ de tamaño $(n_X - n_Y)$, $\{D_{n_Y+1}, ..., D_{n_X}\}$, entonces:

$$\begin{aligned} X &= \sum_{k=1}^{n_X} X_k. \\ Y &= \sum_{k=1}^{n_Y} Y_k + \sum_{k=n_Y+1}^{n_X} D_k. \end{aligned}$$

De este modo, por la Proposición 3.3 $X \succeq_{FSD} Y$. Además como $F_Y(n_Y) = 1 > F_X(n_Y)$ entonces se tiene además que $X \succ_{FSD} Y$.

$\boxed{\rightarrow}$ Por hipótesis, $F_X(t) \leq F_Y(t)$ para todo t real, evaluemos las funciones de distribución en 0:

$$F_X(0) = P(X = 0) = (1-p)^{n_X}, \quad F_Y(0) = P(Y = 0) = (1 - p_Y)^{n_Y}. \tag{3.2}$$

Como $1 - p \leq 1$, $F_X(0) \leq F_Y(0)$ si y solo sí $n_X \geq n_Y$. Quedaría descartar el caso en que $n_X = n_Y$. Si esto ocurriera, entonces $X \equiv_{FSD} Y$ luego contradeciría la hipótesis de $X \succ_{FSD} Y$. Por tanto $n_X > n_Y$. □

Si tenemos en cuenta que para una binomial de parámetros n, p la esperanza es $n \cdot p$, entonces en ambos resultados la caracterización se podría hacer a través de las medias. Sin embargo, cuando uno de los parámetros no es común, el resultado deja de ser cierto.

Ejemplo 3.10. *Consideremos dos variables aleatorias X, Y con $X \sim \mathcal{B}(2, 0.49)$ e $Y \sim \mathcal{B}(4, 0.25)$. Si calculamos sus medias:*

$$E(X) = 0.98, \quad E(Y) = 1.$$

Si calculamos las funciones de distribución:

$$\begin{aligned} k = 0: \quad F_X(0) = 0.2601 \quad &< \quad 0.3164 = F_Y(0), \\ k = 2: \qquad\quad F_X(2) = 1 \quad &> \quad 0.9492 = F_Y(2). \end{aligned}$$

Luego no se puede establecer la comparación por medio de la dominancia estocástica.

Obtenemos también una caracterización de la dominancia estocástica en el caso de la distribución de Poisson.

Proposición 3.16. *Dadas $X \sim \mathcal{P}(\lambda_X)$ e $Y \sim \mathcal{P}(\lambda_Y)$, entonces:*

$$\begin{aligned} X \succ_{FSD} Y &\iff \lambda_X > \lambda_Y, \\ X \equiv_{FSD} Y &\iff \lambda_X = \lambda_Y, \\ X \succeq_{FSD} Y &\iff \lambda_X \geq \lambda_Y. \end{aligned}$$

Demostración. Si $\lambda_X = \lambda_Y$, entonces las variables aleatorias tienen la misma distribución y $X \equiv_{FSD} Y$.
Para una variable aleatoria $X \sim \mathcal{P}(\lambda)$:

$$F_X(k) = \frac{\Gamma(\lfloor k \rfloor, \lambda)}{\lfloor k \rfloor!} \quad \text{con} \quad \Gamma(a, x) = \int_x^\infty t^{a-1} e^{-t} dt.$$

Puesto que x y a son mayores que cero:

$$\frac{\partial \Gamma}{\partial x}(a, x) = -x^{a-1} e^{-x} < 0 \ (si\ x, a > 0).$$

La función de distribución es escalonada en k, luego estudiamos únicamente en k enteros. Por la monotonía de la función de distribución en λ se deduce:

$$F_X(k) < F_Y(k) \quad \forall k \in \mathbb{N} \iff \lambda_X > \lambda_Y. \qquad \square$$

La mayoría de los casos que hemos visto hasta ahora son familias de distribuciones donde el parámetro es un parámetro de escala, como es el caso de la desviación típica en la normal; o bien era un parámetro de localización, como es el caso de la media en la distribución normal. Ahora vamos a generalizar estas familias en la dominancia estocástica.

Proposición 3.17. *Consideremos dos variables aleatorias X, Y pertenecientes a una misma familia de localización, es decir, $X \stackrel{\mathcal{D}}{=} Y + b$ con b el parámetro de localización. Entonces:*

$$X \succ_{FSD} Y \iff b < 0.$$

Demostración. Calculamos la función de distribución:

$$F_X(x) = P(X \leq x) = P(Y + b \leq x) = P(Y \leq x - b) = F_Y(x - b).$$

De modo que por ser la función de distribución creciente:

$$F_X(t) \leq F_Y(t), \quad \forall t \in \mathbb{R} \iff F_Y(t - b) \leq F_Y(t), \quad \forall t \in \mathbb{R} \iff b \leq 0.$$

Por otro lado, la función de distribución no puede tomar el valor 1 en todo $\mathbb{R}$ de modo que han de existir al menos dos puntos t_1, t_2 con $t_1 > t_2$ tales que $F_Y(t_1) > F_Y(t_2)$. De modo que si $b < 0$ entonces existe un punto tal que $F_X(t) < F_Y(t)$. □

Observar que con este resultado hemos generalizado la Proposición 3.12 de las distribuciones normales cuando considerábamos la desviación típica fija. Otra familia que aparece recurrentemente en la literatura es la familia de escala, sin embargo, en este caso no se puede ser tan general.

Proposición 3.18. *Consideremos dos variables aleatorias X, Y pertenecientes a una misma familia de escala, es decir, X e $\frac{Y}{\theta}$ son iguales en distribución, donde θ es el parámetro de escala (mayor que cero). Entonces, si las distribuciones tienen soporte positivo:*

$$X \succ_{FSD} Y \iff \theta > 1.$$

Si, en cambio, las distribuciones tienen soporte negativo:

$$X \succ_{FSD} Y \iff \theta < 1.$$

Demostración. Vamos a verlo para el caso en que las distribuciones tienen soporte positivo, el otro caso es análogo. Como $F_X(x) = P(X \leq x) = P(Y \leq \theta x) = F_Y(\theta x)$, entonces por ser la función de distribución creciente y ser $x \geq 0$, se sigue que $F_X(t)$ será mayor o igual que $F_Y(t)$ si y solo si $\theta \geq 1$. Ahora, si $F_X \neq F_Y$, entonces existe un valor t_0 donde las funciones de distribución no coinciden, de donde se sigue que $F_X(t_0) = F_Y(\theta t_0) < F(t_0)$ si y solo si θ es mayor que uno. □

Sin embargo, el resultado no es cierto en general cuando utilizamos familias de distribución con soportes en $\mathbb{R}$. Un ejemplo de esto sería el Ejemplo 3.9.

Este resultado permite caracterizar a algunas familias, una bastante importante sería la exponencial.

Corolario 3.19. *Dadas dos variables aleatorias* $X \sim Exp(\lambda_X)$*,* $Y \sim Exp(\lambda_Y)$*, entonces:*

$$\begin{aligned} X \succ_{FSD} Y &\iff \lambda_X < \lambda_Y, \\ X \equiv_{FSD} Y &\iff \lambda_X = \lambda_Y, \\ X \succeq_{FSD} Y &\iff \lambda_X \leq \lambda_Y. \end{aligned}$$

Demostración. Si $\lambda_X = \lambda_Y$, entonces las variables aleatorias son iguales en distribución y por tanto $X \equiv_{FSD} Y$.
Consideremos ahora el caso de desigualdad estricta. Al ser la exponencial una familia de escala con soporte positivo:

$$X \stackrel{\mathcal{D}}{=} \frac{\lambda_X}{\lambda_Y} Y.$$

Utilizando la proposición 3.18:

$$X \succeq_{FSD} Y \iff \frac{\lambda_Y}{\lambda_X} > 1.$$

El último caso sería consecuencia de los dos anteriores. $\square$

3.4. Caracterizaciones de la preferencia estadística

En la sección anterior vimos cómo en el caso de normales con la misma varianza podíamos establecer un orden con la dominancia estocástica, sin embargo, cuando las varianzas no eran iguales, las variables podían ser incomparables. Como adelantamos, en la preferencia estadística la incomparabilidad no es posible, por ello, comenzamos esta sección estudiando el caso general de variables aleatorias normales independientes.

Teorema 3.20. *Dadas dos variables aleatorias* $X \sim \mathcal{N}(\mu_X, \sigma_X)$ *e* $Y \sim \mathcal{N}(\mu_Y, \sigma_Y)$ *independientes, entonces:*

$$\begin{aligned} X \succ_{SP} Y &\iff \mu_X > \mu_Y, \\ X \equiv_{SP} Y &\iff \mu_X = \mu_Y, \\ X \succeq_{SP} Y &\iff \mu_X \geq \mu_Y. \end{aligned}$$

Además, $\mathcal{Q}(X,Y) = \Phi\left(\frac{\mu_X - \mu_Y}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right)$.

Demostración. Utilizando la reproductividad de la normal:

$$X - Y \sim \mathcal{N}\Big(\mu_X - \mu_Y, \sqrt{\sigma_X^2 + \sigma_Y^2}\Big).$$

De este modo:

$$\mathcal{Q}(X,Y) = P(X > Y) = P(X - Y > 0) = 1 - \Phi\left(\frac{\mu_Y - \mu_X}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right) = \Phi\left(\frac{\mu_X - \mu_Y}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right).$$

De modo que:

$$\mathcal{Q}(X,Y) = 0.5 \iff \mu_X = \mu_Y,$$
$$\mathcal{Q}(X,Y) > 0.5 \iff \mu_X > \mu_Y,$$
$$\mathcal{Q}(X,Y) < 0.5 \iff \mu_X < \mu_Y. \qquad \square$$

Ya vimos en el Ejemplo 3.5 que la preferencia estadística nos permitía "romper" la incomparabilidad que aparecía en la dominancia estocástica cuando la desviación típica no era la misma. Ahora, podemos caracterizar una familia recurrente en la estadística, como es la de las distribuciones normales únicamente mediante su media, siempre y cuando sean independientes.

Observación 9. *Aunque en principio, "ser preferida" dependa únicamente de las medias, la varianza juega un peso importante en cuán preferida es X a Y. Supongamos dos variables aleatorias normales e independientes X,Y con $X \succ_{SP} Y$, utilizando la caracterización de $\mathcal{Q}$ obtenida en el anterior teorema, observamos que $\mathcal{Q}$ es estrictamente decreciente en σ_Y^2 , es decir, cuanto mayor sea la varianza de Y, menos preferida será X a Y. En particular, dada la continuidad de Φ:*

$$\lim_{\sigma_Y^2 \to \infty} \mathcal{Q}(X,Y) = 1 - \lim_{\sigma_Y^2 \to \infty} \Phi\left(\frac{\mu_Y - \mu_X}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right) = 1 - \Phi\left(\lim_{\sigma_Y^2 \to \infty} \frac{\mu_Y - \mu_X}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right) = 1 - \Phi(0) = 0.5.$$

Del mismo modo, se puede comprobar que $\mathcal{Q}$ es estrictamente decreciente en σ_X^2 y además $\lim_{\sigma_X^2 \to \infty} \mathcal{Q}(X,Y) = 0.5$.
De modo que la varianza juega un importante y ambigüo papel en la preferencia, pues si X es preferida a Y, mayor será la preferencia por X cuanto menor sea su varianza; sin embargo, a Y le afecta de manera contraria, menor será la preferencia por X cuanto mayor sea la varianza de Y.

Continuamos estudiando distribuciones normales, en particular, vamos a estudiar dos casos extremos de dependencia: comonotonía y contramonotonía.

Proposición 3.21. *Dadas dos variables aleatorias $X \sim \mathcal{N}(\mu_X, \sigma_X)$ y $Y \sim \mathcal{N}(\mu_Y, \sigma_Y)$*

comonótonas, entonces:

$$
\begin{aligned}
X \succ_{SP} Y &\iff \mu_X > \mu_Y,\\
X \equiv_{SP} Y &\iff \mu_X = \mu_Y,\\
X \succeq_{SP} Y &\iff \mu_X \geq \mu_Y.
\end{aligned}
$$

Además, $\mathcal{Q}(X,Y) = \Phi\left(\dfrac{\mu_X - \mu_Y}{|\sigma_X - \sigma_Y|}\right)$.

Demostración. Como las variables aleatorias son comonótonas, entonces

$$F_{XY}(x,y) = \min\{F_X(x), F_Y(y)\} = \min\left\{\Phi\Big(\frac{x-\mu_X}{\sigma_X}\Big), \Phi\Big(\frac{y-\mu_Y}{\sigma_Y}\Big)\right\}.$$

Recordando la Proposición 3.11:

$$\mathcal{Q}(X,Y) = \int_{t:F_X(t)<F_Y(t)} f_X(t)dt + \frac{1}{2}\int_{t:F_X(t)=F_Y(t)} f_X(t)dt.$$

Ahora si $\mu_X \neq \mu_Y$ y $\sigma_X \neq \sigma_Y$, F_X y F_Y únicamente tienen un punto de corte, de modo que la segunda integral vale 0 y el punto de corte es

$$t_0 \equiv \frac{\mu_X\sigma_Y - \mu_Y\sigma_X}{\sigma_Y - \sigma_X}.$$

Estudiando ahora las funciones de distribución:

$$F_X(t) < F_Y(t) \iff \Phi\Big(\frac{t-\mu_X}{\sigma_X}\Big) < \Phi\Big(\frac{t-\mu_Y}{\sigma_Y}\Big) \iff \begin{cases} t < t_0 & si \quad \sigma_Y > \sigma_X,\\ t > t_0 & si \quad \sigma_Y < \sigma_X.\end{cases}$$

Si $\sigma_Y > \sigma_X$:

$$\mathcal{Q}(X,Y) = \int_{-\infty}^{t_0} f_X(t)dt = F_X(t_0) = \Phi\Big(\frac{t_0-\mu_X}{\sigma_X}\Big) = \Phi\Big(\frac{\mu_X-\mu_Y}{\sigma_Y-\sigma_X}\Big).$$

Así:

$$
\begin{aligned}
X \succ_{SP} Y \iff \Phi\Big(\frac{\mu_X-\mu_Y}{\sigma_Y-\sigma_X}\Big) > 0.5 \iff \mu_X > \mu_Y.\\
X \equiv_{SP} Y \iff \mu_X = \mu_Y.
\end{aligned}
$$

Si $\sigma_X > \sigma_Y$:

$$\mathcal{Q}(X,Y) = \int_{t_0}^{\infty} f_X(t)dt = 1 - F_X(t_0) = 1 - \Phi\Big(\frac{t_0-\mu_X}{\sigma_X}\Big) = 1 - \Phi\Big(\frac{\mu_Y-\mu_X}{\sigma_X-\sigma_Y}\Big).$$

Así:

$$X \succ_{SP} Y \iff \Phi\Big(\frac{\mu_Y - \mu_X}{\sigma_X - \sigma_Y}\Big) < 0.5 \iff \mu_X > \mu_Y.$$

$$X \equiv_{SP} Y \iff \mu_X = \mu_Y.$$

En el caso en el que $\sigma_X = \sigma_Y \equiv \sigma$ y $\mu_X \neq \mu_Y$, entonces no se cortan en ningún punto y por tanto:

$$F_X < F_Y \iff \Phi\Big(\frac{t-\mu_X}{\sigma}\Big) < \Phi\Big(\frac{t-\mu_Y}{\sigma}\Big) \iff \mu_X > \mu_Y.$$

De modo que en este caso:

$$\mathcal{Q}(X,Y) = \begin{cases} 0, & \text{si } \mu_X < \mu_Y, \\ 1, & \text{si } \mu_X > \mu_Y. \end{cases}$$

Si $\mu_X = \mu_Y$ y $\sigma_X = \sigma_Y$, entonces

$$\mathcal{Q}(X,Y) = \frac{1}{2}\int_{t:F_X(t)=F_Y(t)} f_X(t)dt = 0.5.$$

Por tanto, $X \equiv_{SP} Y$.

Por último, observar que podríamos expresar la $\mathcal{Q}$ en general para todos los casos como:

$$\mathcal{Q}(X,Y) = \Phi\left(\frac{\mu_X - \mu_Y}{|\sigma_X - \sigma_Y|}\right). \qquad \square$$

Proposición 3.22. *Dadas dos variables aleatorias $X \sim \mathcal{N}(\mu_X, \sigma_X)$ y $Y \sim \mathcal{N}(\mu_Y, \sigma_Y)$ contramonótonas, entonces:*

$$X \succ_{SP} Y \iff \mu_X > \mu_Y,$$
$$X \equiv_{SP} Y \iff \mu_X = \mu_Y,$$
$$X \succeq_{SP} Y \iff \mu_X \geq \mu_Y.$$

Además, $\mathcal{Q}(X,Y) = \Phi\left(\frac{\mu_X - \mu_Y}{\sigma_X + \sigma_Y}\right)$.

Demostración. Como X e Y son contramonótonas, entonces

$$F_{XY}(x,y) = \text{máx}\Big\{F_X(x) + F_Y(y) - 1, 0\Big\} = \text{máx}\Big\{\Phi\Big(\frac{x-\mu_X}{\sigma_X}\Big) + \Phi\Big(\frac{y-\mu_Y}{\sigma_Y}\Big) - 1, 0\Big\}.$$

Recordando la Proposición 3.10:

$$\mathcal{Q}(X,Y) = F_Y(u) \quad \text{con } u \text{ un punto tal que} \quad F_X(u) + F_Y(u) = 1.$$

Como $F_X(x) = \Phi\Big(\dfrac{x-\mu_X}{\sigma_X}\Big)$ y $F_Y(y) = \Phi\Big(\dfrac{y-\mu_Y}{\sigma_Y}\Big)$, entonces:

$$F_X(u) = 1 - F_Y(u) \iff \frac{u-\mu_X}{\sigma_X} = -\frac{u-\mu_Y}{\sigma_Y} \iff u = \frac{\mu_Y\sigma_X + \mu_X\sigma_Y}{\sigma_X+\sigma_Y}.$$

De este modo:

$$\mathcal{Q}(X,Y) = \Phi\Big(\frac{\mu_X-\mu_Y}{\sigma_X+\sigma_Y}\Big),$$

de donde se deduce el resultado. □

Hemos estudiado la preferencia estadística en tres casos, variables independientes, variables contramonótonas y variables comonótonas. Todos estos casos tienen en común que la caracterización se hace únicamente a través de las medias, es decir es independiente de las varianzas. Podemos observar además que:

$$\mathcal{Q}^{contra}(X,Y) \leq \mathcal{Q}^{ind}(X,Y) \leq \mathcal{Q}^{com}(X,Y).$$

Recordando que estos tres casos de dependencia serían los casos de $\rho \in \{-1, 0, 1\}$ de una distribución normal bivariante, parece indicar que existe una relación entre el coeficiente de correlación y el grado de preferencia. El siguiente resultado generaliza esto.

Teorema 3.23. *[13, Teorema 3.82] Consideremos un vector aleatorio (X,Y) que sigue una distribución normal bivariante. Entonces*

$$\mathcal{Q}(X,Y) = \Phi\left(\frac{\mu_X-\mu_Y}{\sqrt{\sigma_X^2+\sigma_Y^2-2\rho\sigma_X\sigma_Y}}\right). \tag{3.3}$$

Luego:

$$\begin{aligned} \mathcal{Q}(X,Y) = 0.5 &\iff \mu_X = \mu_Y, \\ \mathcal{Q}(X,Y) > 0.5 &\iff \mu_X > \mu_Y, \\ \mathcal{Q}(X,Y) < 0.5 &\iff \mu_X < \mu_Y. \end{aligned}$$

Este resultado confirma la hipótesis de la dependencia con el coeficiente de correlación y además caracteriza no solo los tres casos extremos de la normal bivariante, sino además cualquiera caso de una normal bivariante.

En gran medida, que se pueda caracterizar una familia de distribuciones biparamétrica mediante únicamente uno solo de sus parámetros se debe a la simetría de la función de densidad conjunta de X e Y. El siguiente resultado generaliza esto a otras familias no necesariamente normales. Este resultado se basa en estudiar qué clases de cópulas conservan la simetría radial que como se vio en el Capítulo 2, que eran aquellas cópulas que coincidían con su cópula de supervivencia. Ejemplos de esta clase de cópulas serían la gaussiana (y por tanto cópula independiente, cópula del mínimo y cópula de Lukasiewicz) o entre las cópulas arquimedianas la cópula de Frank.

Proposición 3.24. *Consideremos dos variables aleatorias absolutamente continuas X e Y con distribuciones simétricas respecto a S_X y S_Y respectivamente. Si su cópula coincide con la cópula de supervivencia, entonces:*

$$
\begin{aligned}
X \succ_{SP} Y &\iff S_X > S_Y, \\
X \equiv_{SP} Y &\iff S_X = S_Y, \\
X \succeq_{SP} Y &\iff S_X \geq S_Y.
\end{aligned}
$$

Demostración. Como X e Y son distribuciones simétricas con respecto a S_X y S_Y respectivamente y $C(u,v) = \overline{C}(u,v)$, se sigue por la Proposición 2.5 que $f_{XY}(x,y)$ es radialmente simétrica respecto a (S_X, S_Y).
Definimos ahora $A = \{(x,y) \in \mathbb{R}^2 : x - S_X > y - S_Y\}$. Por la simetría radial:

$$
\int_A f_{XY}(x,y)d(x,y) = \frac{1}{2}\int_{\mathbb{R}^2} f_{XY}(x,y)d(x,y) = \frac{1}{2}.
$$

El vector director de la frontera de A es el mismo que el de la frontera de $B = \{(x,y) \in \mathbb{R}^2 : x > y\}$, esto se puede observar en la Figura 3.8. De modo que si $S_X \neq S_Y$, $A \subset B$ ó $A \supset B$. En particular:

$$
A \supsetneq B \iff S_X > S_Y.
$$

De este modo:

$$
P(X > Y) = \int_B f_{XY}(x,y)d(x,y) > 0.5 \iff A \supsetneq Q \iff S_X > S_Y.
$$

Si $S_X = S_Y$, entonces $B = A$, de modo que

$$
P(X > Y) = \int_B f_{XY}(x,y)d(x,y) = 0.5.
$$

En la Figura 3.8 se ilustra el caso en que $B \subset A$ siendo las circunferencias las curvas de nivel de la función de densidad. Observamos como A contiene la mitad de la probabilidad pues corta en dos secciones simétricas todas las curvas de nivel. □

Observamos que cuando la familia tenga esperanzas finitas entonces el eje de simetría serán las medias como veíamos que ocurría para normales en los casos gaussianos (Teorema 3.23). De este modo, este resultado es más general pues nos permite caracterizar incluso familias como la Cauchy la cual si bien es simétrica no tiene esperanza finita.
En particular, la cópula del producto coincide con su cópula de supervivencia. De este modo, este resultado nos permite deducir que en cualquier familia de localización-escala

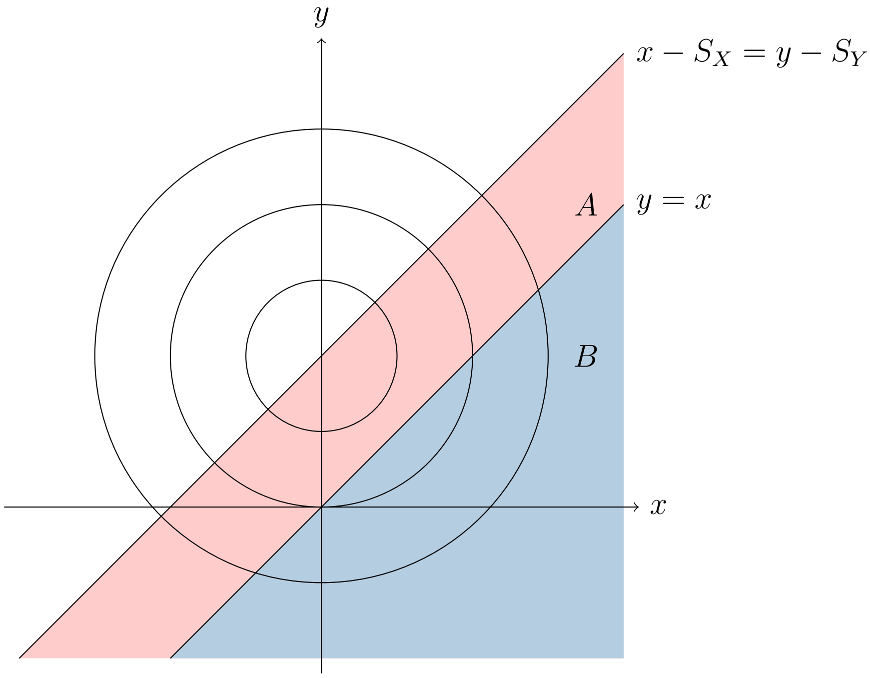

Figura 3.8: Curvas de nivel de una función de densidad bivariante con simetría radial junto con las dos regiones construidas en la demostración del Teorema 3.24.

con esperanzas finitas y simétrica, dadas dos variables aleatorias independientes, podemos caracterizar la preferencia estadística únicamente mediante sus medias. Sin embargo, esto deja de ser cierto cuando la familia no es simétrica, tal y como se ve en el ejemplo a continuación.

Ejemplo 3.11. *Consideremos la familia de localización-escala*

$$\left\{ \frac{E-b}{a} \middle| E \sim Exp(1); a \in \mathbb{R}^+, b \in \mathbb{R} \right\}.$$

Tomamos ahora Y perteneciente a la familia descrita y X una copia de E independiente. Entonces sus funciones de densidad:

$$f_X(x) = e^{-x}\mathbb{1}_{(0,\infty]}(x), \quad f_Y(y) = ae^{-ay-b}\mathbb{1}_{(-b/a,\infty]}(y).$$

Ahora estudiamos la preferencia:
Si $b < 0$,

$$P(X > Y) = \int_{X>Y} f_X(x)f_Y(y)d(x,y) = \int_{-b/a}^{\infty}\int_{-b/a}^{x} ae^{-x}e^{-ya-b}dydx = e^{b/a}\frac{a}{a+1}.$$

Si $b > 0$,

$$\begin{aligned} P(X > Y) &= \int_{X>Y} f_X(x) f_Y(y) d(x,y) \\ &= \int_0^\infty \int_0^x a e^{-x} e^{-ya-b} dy dx + \int_0^\infty \int_{-b/a}^x a e^{-x} e^{-ya-b} dy dx \\ &= 1 - \frac{e^{-b}}{a+1}. \end{aligned}$$

Por otro lado, $E(Y) = \dfrac{1-b}{a}$.
Si ahora tomamos $a = \dfrac{5}{2}$ y $b = -1$, entonces:

$$\mathcal{Q}(X,Y) = 0.4788 < 0.5 \implies X \prec_{SP} Y.$$

Sin embargo, $E(X) = 1$ y $E(Y) = \dfrac{4}{5}$.

Por otro lado, cuando las marginales son simétricas pero la cópula no tiene simetría radial, el resultado tampoco es cierto. En el siguiente ejemplo se ha comprobado por simulación para el caso de normales.

Ejemplo 3.12. *Consideremos dos variables aleatorias X e Y que siguen ambas una normal estándar, es decir $X, Y \sim \mathcal{N}(0,1)$, relacionadas mediante la cópula de Khoudraji (toma parámetros a_1, a_2) formada por las cópulas de Gumbel (la cual toma parámetro θ) y la cópula del producto):*

$$C(u,v) = \left[\max\{u^{-\theta(1-a_1)} + v^{-\theta(1-a_2)} - 1;, 0\}\right]^{-1/\theta} u^{a_1} v^{a_2}, \quad \forall (u,v) \in [0,1] \times [0,1].$$

En particular cuando $a_1 = 0.4$, $a_2 = 1$ y $\theta = 6$ se obtiene (por simulación) $\mathcal{Q}(X,Y) \approx 0.4366$. Esta diferencia se explica muy bien gráficamente en la Figura 3.9, pues se ha perdido la simetría respecto al eje $Y = X$ y la mayoría de la probabilidad se encuentra en el semiplano inferior.

Una familia de cópulas bastante recurrente es la de las cópulas arquimedianas, que poseen ciertas propiedades que facilitan su manejo. Una de estas propiedades es que son simétricas, es decir $C(u,v) = C(v,u)$. El siguiente resultado, aunque es bastante simple, es aplicable a esta familia de cópulas.

Proposición 3.25. *Consideremos dos variables aleatorias X e Y idénticamente distribuidas. Si su cópula tiene simetría de la forma $C(u,v) = C(v,u)$, entonces se cumple que:*

$$X \equiv_{SP} Y.$$

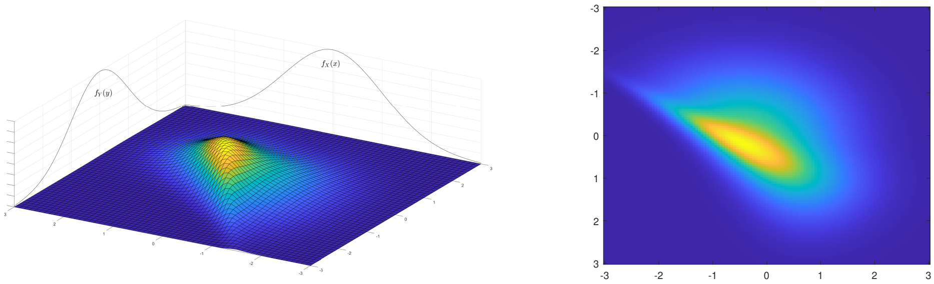

Figura 3.9: Distribución de densidad conjunta de dos normales estándar relacionadas mediante la cópula de Khoudraji con valores $a_1 = 0.4$ y $a_2 = 1$ con subcópulas de Gumbel $\theta = 6$ y del producto.

Demostración. Por un lado, la cópula es simétrica mientras que por otro X e Y están idénticamente distribuidas, entonces de acuerdo con la Proposición 2.6 el vector aleatorio (X, Y) tiene la misma distribución que (Y, X). Por tanto $P(X > Y) = P(Y > X)$ y así $\mathcal{Q}(X, Y) = \mathcal{Q}(Y, X)$. $\square$

El siguiente ejemplo pretende demostrar que este resultado no es cierto en general.

Ejemplo 3.13. *Consideremos dos variables aleatorias U, V ambas con distribución uniforme en $(0, 1)$. Construimos su distribución conjunta mediante la cópula:*

$$C(u, v) = uv - u^3 v(1-u)(1-v), \quad \forall (u, v) \in [0, 1] \times [0, 1].$$

Esta cópula es claramente asimétrica pues $C(1, 0.5) = 0.25$ mientras que $C(0.5, 1) = 0.4375$. Al ser las distribuciones uniformes estándar, la cópula coincide con la función de distribución en el cuadrado $[0, 1] \times [0, 1]$, de modo que su función de distribución conjunta viene dada por:

$$F(u, v) = \begin{cases} 0, & \text{si } u \leq 0, \text{ ó } v \leq 0, \\ uv - u^3 v(1-u)(1-v), & \text{si } u, v \in [0, 1], \\ v, & \text{si } u \geq 1, v \in [0, 1], \\ u, & \text{si } u \in [0, 1], v \geq 1. \end{cases}$$

La función de densidad conjunta es:

$$f(u, v) = \frac{\partial}{\partial u}\frac{\partial}{\partial v} F(u, v) = \Big(1 + u^3(4 - 8v) + u^2(-3 + 6v)\Big) \mathbb{1}_{[0,1]}(u, v).$$

Ahora calculamos la preferencia estadística:

$$\begin{aligned}\mathcal{Q}(X,Y) = P(X > Y) &= \int_{X>Y} d(u,v)\, f(u,v) \\ &= \int_0^1 \int_0^u dvdu\,(1 + u^3(4-8v) + u^2(6v-3)) = \frac{29}{60} < 0.5\,.\end{aligned}$$

De modo que comprobamos que a pesar de ser distribuciones idénticas $X \prec_{SP} Y$.

Capítulo 4

Preferencia estadística bivariante

Una vez introducidos dos órdenes estocásticos unidimensionales distintos, vamos a extender la preferencia estadística al caso multidimensional. De este modo, trabajaremos en el espacio probabilístico $(\Omega, \mathcal{A}, P)$, denotando el conjunto de vectores aleatorios en este espacio como $\mathcal{D}^{(2)}$. Salvo que se indique lo contrario, las vectores aleatorios a comparar estarán en el mismo espacio probabilístico.

Observación 10. *Cuando estudiábamos órdenes estocásticos unidimensionales, siempre definíamos una desigualdad entre dos cantidades. En este caso, tratábamos con escalares lo cual permitía hacerlo de la manera usual. Sin embargo, cuando comparamos cantidades vectoriales, resulta necesario establecer un criterio de qué significa "ser mayor que". El criterio escogido será que $\vec{X} > \vec{Y}$ cuando la desigualdad se cumpla componente a componente, es decir $X_1 > Y_1$ y $X_2 > Y_2$. De este modo se deduce que $\vec{X} \not> \vec{Y}$ cuando alguna de las componentes no cumpla la desigualdad, es decir $X_1 \leq Y_1$ ó $X_2 \leq Y_2$. Del mismo modo, definiremos $\vec{X} \geq \vec{Y}$ si $X_1 \geq Y_1$ y $X_2 \geq Y_2$ y por tanto $\vec{X} \not\geq \vec{Y}$ cuando $X_1 < Y_1$ ó $X_2 < Y_2$.*

Cuando buscamos extender un concepto, consideremos una definición o caracterización e intentamos ver cómo se comportaría en el caso a extender. Sin embargo, si existen distintas caracterizaciones o definiciones, consecuentemente existirán distintas extensiones en función de la caracterización que se utilice.
Pongamos por ejemplo el caso de la dominancia estocástica. Si tomáramos la definición de que $X \succeq_{FSD} Y$ si $\bar{F}_X(t) \geq \bar{F}_Y(t)$ para todo t en $\mathbb{R}$, entonces la extensión natural para vectores aleatorios de dimensión 2 sería:

$$\vec{X} \succeq^{w}_{FSD} \vec{Y} \quad \text{si} \quad \bar{F}_{\vec{X}}(t_1, t_2) \geq \bar{F}_{\vec{Y}}(t_1, t_2), \quad \forall t_1, t_2 \in \mathbb{R}.$$

Sin embargo, vimos en el Teorema 3.1 que X fuera preferido a Y era equivalente a que

$E(u(X)) \geq E(u(Y))$ para toda función u creciente. De este modo, podríamos definir la extensión:

$$\vec{X} \succeq^s_{FSD} \vec{Y} \quad \text{si} \quad E\Big(u(\vec{X})\Big) \geq E\Big(u(\vec{Y})\Big) \quad \text{para toda } u : \mathbb{R}^2 \to \mathbb{R} \text{ creciente.}$$

Sin embargo, estas dos extensiones que en el caso univariante son equivalentes, en este caso no lo serían.

Ejemplo 4.1. *[20, Ejemplo 1] Consideremos los vectores aleatorios $\vec{X}$ e $\vec{Y}$ cuyas distribuciones de probabilidad vienen dadas por:*

(X_1, X_2)	$P_{\vec{X}}$
$(0,0)$	0.5
$(1,1)$	0.5

(Y_1, Y_2)	$P_{\vec{Y}}$
$(0,1)$	0.5
$(1,0)$	0.5

Entonces:

$$\bar{F}_{\vec{X}}(t_1,t_2) = \begin{cases} 1 & si \quad t_1, t_2 < 0, \\ 0.5 & si \quad t_1 \in [0,1),\, t_2 < 1 \quad ó \quad t_1 < 1,\, t_2 \in [0,1), \\ 0 & si \quad t_1 \geq 1 \text{ ó } t_2 \geq 1. \end{cases}$$

$$\bar{F}_{\vec{Y}}(t_1,t_2) = \begin{cases} 1 & si \quad t_1, t_2 < 0, \\ 0.5 & si \quad t_1 < 0,\, t_2 \in [0,1) \quad ó \quad t_1 \in [0,1),\, t_2 < 0, \\ 0 & si \quad t_1, t_2 \geq 0. \end{cases}$$

Se comprueba que $\bar{F}_{\vec{X}}(t_1,t_2) > \bar{F}_{\vec{Y}}(t_1,t_2)$ de modo que $\vec{X} \succeq^w_{FSD} \vec{Y}$. Si consideramos ahora la función creciente $u(t_1,t_2) = \max\{t_1,t_2\}$, entonces $E\big(u(\vec{X})\big) = 0.5$ mientras que $E\big(u(\vec{Y})\big) = 1$, por tanto $\vec{X} \not\succeq^s_{FSD} \vec{Y}$.

Como acabamos de ver, distintas condiciones que en el caso univariante son equivalentes dan lugar a distintas extensiones bivariantes. En la literatura podemos encontrar principalmente tres extensiones de dominancia estocástica al campo multivariante. La primera, quizá la menos común, está basada en la función de supervivencia o de distribución, que si bien en el caso unidimensional daban lugar a la misma definición, en el caso bidimensional no es así, como por ejemplo [17] donde se denomina “dominancia ortante” (orthant dominance); la introducida previamente a través de la función de supervivencia se conoce como dominancia estocástica débil [20].

En segundo lugar, podemos encontrar una extensión basada en las esperanzas. Si bien esta fue introducida por Sriboonchittaet al. en [23], esta es equivalente a la introducida en por Levhari [10] y se conoce como dominancia estocástica fuerte.

Podemos encontrar una última clase de orden estocástico en la literatura, la dominancia

estocástica lineal introducida por Dentcheva y Ruszczýnski [8], donde se dice que $\vec{X}$ domina linealmente a $\vec{Y}$ por primer grado, y se denota por $\vec{X} \succeq_{FSD}^{lin} \vec{Y}$, si $\vec{c}^T\vec{X} \succeq_{FSD} \vec{c}^T\vec{Y}$ para todo $\vec{c}$ en $\mathbb{R}^2_+$ [1] donde $\succeq_{FSD}$ denota la dominancia estocástica unidimensional.
En general, se podría demostrar (por ejemplo en [20, Teorema 5]) que $\vec{X} \succeq_{FSD}^{s} \vec{Y}$ implica $\vec{X} \succeq_{FSD}^{w} \vec{Y}$ y $\vec{X} \succeq_{FSD}^{lin} \vec{Y}$, es decir, es una condición más fuerte que las otras dos.

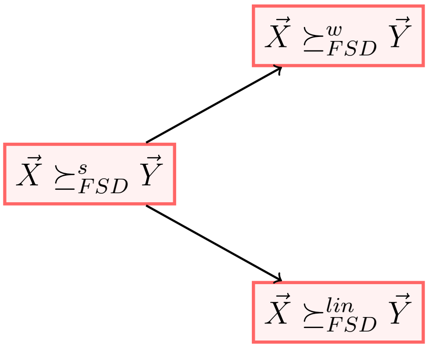

Figura 4.1: Relaciones entre las distintas extensiones de la dominancia estocástica.

En el resto de este capítulo, el objetivo será extender la preferencia estadística a la comparación de vectores aleatorios. Una vez visto cómo se puede extender la dominancia estocástica al caso bivariante, esto nos da una idea del trabajo que vamos a realizar para la preferencia estadística, dando distintas extensiones para posteriormente definir distintos órdenes y analizar las relaciones entre ellos.
En particular, se van a considerar tres extensiones distintas:

- En primer lugar, daremos una noción de la preferencia estadística multivariante en la cual estudiamos la preferencia univariante componente a componente para posteriormente agregar los valores obtenidos.
- En segundo lugar, realizaremos un proceso inverso al anterior, realizando primero una agregación de las componentes de cada vector aleatorio obteniendo así dos variables aleatorias unidimensionales las cuales comparamos mediante la preferencia estadística univariante.
- Por último, daremos una noción puramente bivariante.

En todos los casos, esta comparación se realizará a través de una aplicación $\mathcal{Q}^* : \mathcal{D} \times \mathcal{D} \to [0,1]$ siguiendo la idea del caso univariante.

[1] $\mathbb{R}^2_+ = \{(x,y) \in \mathbb{R}^2 : x, y > 0\}$.

4.1. Agregación de la preferencia estadística univariante

Consideremos dos vectores aleatorios $\vec{X}$, $\vec{Y}$ de dimensión 2. Seguramente el enfoque más sencillo que se podría plantear es considerar las preferencias estadísticas unidimensionales componente a componente y realizar su promedio:

$$\frac{1}{2}\big(\mathcal{Q}(X_1,Y_1)+\mathcal{Q}(X_2,Y_2)\big).$$

Este enfoque plantea dos ventajas fundamentalmente, es muy fácilmente computable y, como veremos, continúa siendo una relación probabilística. Esto último resulta útil pues de este modo la estructura de preferencia no tiene elementos incomparables. Sin embargo, si bien estudiar las distribuciones marginales aporta simplicidad, se pierde toda la información sobre la posible dependencia existente entre las distintas componentes, pudiendo resultar en una definición excesivamente sencilla.
Si bien este enfoque podría ser generalizado con cualquier medida de centralización, no solamente la media; la moda o la mediana serían indicadores bastante pobres cuando se trata de vectores aleatorios bidimensionales. Sin embargo, sí podríamos considerar aplicaciones distintas de la media que posean propiedades similares a esta. En particular, exigiremos que las funciones sean idempotentes ($f(x,x)=x$) y monótonas crecientes. Un ejemplo de esta clase de funciones sería una media ponderada, de utilidad cuando, por el motivo que sea, existe una componente que posee mayor interés.

Definición 4.1. *Sean $\vec{X}$, $\vec{Y}$ dos vectores aleatorios de dimensión 2 y $M:[0,1]^2\to[0,1]$ una función estrictamente creciente e idempotente. Entonces definimos la aplicación:*

$$\mathcal{Q}^M(\vec{X},\vec{Y})=M\Big(\mathcal{Q}(X_1,Y_1),\mathcal{Q}(X_2,Y_2)\Big). \tag{4.1}$$

Una vez introducida la aplicación a través de la cual se realiza la comparación, podemos definir la primera extensión de preferencia estadística.

Definición 4.2 (Preferencia estadística mediante agregación de la preferencia estadística univariante)**.** *Sean $\vec{X}$, $\vec{Y}$ dos vectores aleatorios de dimensión 2 y Q^M la aplicación definida en la Ecuación (4.1). Entonces diremos que:*

- *$\vec{X}$ es M preferida estadísticamente a $\vec{Y}$, y se denota por $\vec{X}\succeq_{SP}^M\vec{Y}$, si $\mathcal{Q}^M(\vec{X},\vec{Y})\geq\mathcal{Q}^M(\vec{Y},\vec{X})$.*
- *$\vec{X}$ es M indiferente estadísticamente a $\vec{Y}$, y se denota por $\vec{X}\equiv_{SP}^M\vec{Y}$, si $\mathcal{Q}^M(\vec{X},\vec{Y})=\mathcal{Q}^M(\vec{Y},\vec{X})$.*

- $\vec{X}$ *es M estrictamente preferida estadísticamente , y se denota por* $\vec{X} \succ_{SP}^{M} \vec{Y}$*, si* $\mathcal{Q}^M(\vec{X},\vec{Y}) > \mathcal{Q}^M(\vec{Y},\vec{X})$.

Observación 11. *En este caso la relación binaria característica sería* $\succeq_{SP}^{M}$*. Esta relación es reflexiva y fuertemente completa. Sin embargo, al definirse a partir de la preferencia unidimensional, no es transitiva.*

En primer lugar, esta noción define una estructura de preferencia sin elementos incomparables la relación binaria característica es fuertemente completa. De este modo, $(\mathcal{D}^{(2)}, \succ_{SP}^{M}, \prec_{SP}^{M}, \equiv_{SP}^{M})$ determina una estructura de preferencia sin elementos incomparables $\langle P, I\rangle$. En segundo lugar, si bien esta noción de preferencia estadística no será una relación probabilística en general (como se verá en el Ejemplo 4.2), sí existirán algunas funciones M para las cuales lo sea (tal y como se demuestra a continuación).

Proposición 4.1. *La aplicación* $\mathcal{Q}^M$ *definida en la definición 4.1 es una relación probabilística si y solo si* $M(a,b) + M(1-a, 1-b) = 1$ *para todo* $a, b \in [0,1]$.
En particular, si M es la media aritmética ($M(a,b) = \mathrm{am}(a,b) = \frac{a+b}{2}$ *para todo* $a, b \in [0,1]$*) entonces es una relación probabilística.*

Demostración.

$$\begin{aligned}
\mathcal{Q}^M(\vec{X},\vec{Y}) &+ \mathcal{Q}^M(\vec{Y},\vec{X}) = \\
&= M\Big(Q(X_1,Y_1,), Q(X_2,Y_2)\Big) + M\Big(Q(Y_1,X_1), Q(Y_2,X_2)\Big) \\
&= M\Big(Q(X_1,Y_1), Q(X_2,Y_2)\Big) + M\Big(1-Q(X_1,Y_1), 1-Q(X_2,Y_2)\Big).
\end{aligned}$$

Entonces si $\mathcal{Q}^M$ es una relación probabilísica entonces se sigue que:

$$1 = M\Big(Q(X_1,Y_1), Q(X_2,Y_2)\Big) + M\Big(1-Q(X_1,Y_1), 1-Q(X_2,Y_2)\Big).$$

Por otro lado si $M\Big(Q(X_1,Y_1), Q(X_2,Y_2)\Big) + M\Big(1-Q(X_1,Y_1), 1-Q(X_2,Y_2)\Big) = 1$ entonces

$$\mathcal{Q}^M(\vec{X},\vec{Y}) = 1.$$

Si consideramos la media aritmética $M = \mathrm{am}$, entonces por linealidad se obtiene:

$$\mathrm{am}(a,b) + \mathrm{am}(1-a, 1-b) = \mathrm{am}(a,b) + 1 - \mathrm{am}(a,b) = 1. \qquad \square$$

Ejemplo 4.2. *Consideremos el espacio probabilístico* $(\Omega, \mathcal{P}(\Omega), P)$ *y dos vectores aleato-*

rios (X_1, X_2) *y* (Y_1, Y_2) *dados por:*

Ω	ω_1	ω_2	ω_3
P	$1/4$	$1/2$	$1/4$
(X_1, X_2)	$(1,2)$	$(2,1)$	$(2,1)$
(Y_1, Y_2)	$(1,1)$	$(2,2)$	$(3,3)$

Calculamos las preferencias estadísticas unidimensionales:

$$\mathcal{Q}(X_1, Y_1) = \frac{1}{2}P(\{\omega_1, \omega_2\}) = \frac{3}{8},$$
$$\mathcal{Q}(X_2, Y_2) = P(\{\omega_1\}) = \frac{1}{2}.$$

De este modo, si consideramos la media aritmética $M = \mathrm{am}$ *se obtiene:*

$$\mathcal{Q}^{\mathrm{am}}(\vec{X}, \vec{Y}) = \frac{7}{16}, \quad y \quad \mathcal{Q}^{\mathrm{am}}(\vec{Y}, \vec{X}) = \frac{9}{16},$$

cumpliéndose además que $\mathcal{Q}^{\mathrm{am}}(\vec{X}, \vec{Y}) + \mathcal{Q}^{\mathrm{am}}(\vec{Y}, \vec{X}) = 1$. *Por tanto,* $\vec{Y} \succ_{SP}^{am} \vec{X}$.
Sin embargo, si consideramos la media geométrica (es decir $\mathrm{gm}(a,b) = \sqrt{ab}$ *para todo* $a, b \in [0,1]$*) entonces* $\mathcal{Q}^{gm}(\vec{X}, \vec{Y}) = \dfrac{\sqrt{3}}{4}$ *y* $\mathcal{Q}^{\mathrm{gm}}(\vec{X}, \vec{Y}) = \dfrac{\sqrt{5}}{4}$ *y por tanto también se cumple* $\vec{Y} \succ_{SP}^{gm} \vec{X}$. *Ahora bien, a pesar de que la media geométrica es creciente e idempotente, no se cumple la propiedad dada en la Proposición 4.1 y por tanto*

$$\mathcal{Q}^{gm}(\vec{Y}, \vec{X}) + \mathcal{Q}^{gm}(\vec{Y}, \vec{X}) = \frac{\sqrt{5} + \sqrt{3}}{4} \neq 1,$$

de modo que Q^{gm} *no es una relación probabilística.*

Por último vamos a dar una caracterización similar a la Proposición 3.6 para el caso particular en que M es la media aritmética.

Proposición 4.2. *Dados dos vectores aleatorios* $\vec{X}$, $\vec{Y}$ *de dimensión* 2, *entonces son equivalentes:*

1. $\mathcal{Q}^{am}(\vec{X}, \vec{Y}) > \dfrac{1}{2}$.
2. $\mathcal{Q}^{am}(\vec{X}, \vec{Y}) > \mathcal{Q}^{am}(\vec{Y}, \vec{X})$.
3. $P(X_1 > Y_1) + P(X_2 > Y_2) > P(Y_1 > Y_1) + P(Y_2 > X_2)$.
4. $P(X_1 \geq Y_1) + P(X_2 \geq Y_2) > P(Y_1 \geq Y_1) + P(Y_2 \geq X_2)$.

Demostración. $\boxed{1. \leftrightarrow 2.}$ Utilizando la Proposición 4.1 ($\mathcal{Q}^{am}$ es una relación probabilística):

$$\mathcal{Q}^{am}(\vec{X},\vec{Y}) > \mathcal{Q}^{am}(\vec{Y},\vec{X}) = 1 - \mathcal{Q}^{am}(\vec{X},\vec{Y}) \iff \mathcal{Q}^{am}(\vec{X},\vec{Y}) > \frac{1}{2}.$$

$\boxed{2. \leftrightarrow 3.}$ Por hipótesis $\mathcal{Q}^{am}(\vec{X},\vec{Y}) - \mathcal{Q}^{am}(\vec{Y},\vec{X}) > 0$, de modo que

$$0 < \mathcal{Q}^{am}(\vec{X},\vec{Y}) - \mathcal{Q}^{am}(\vec{Y},\vec{X}) = \frac{1}{2}\sum_{k=1}^{2}\Big(P(X_k > Y_k) - P(Y_k > X_k)\Big),$$

de donde se deduce el resultado.
$\boxed{3. \leftrightarrow 4.}$ A partir de la igualdad

$$\begin{aligned}
&\sum_{k=1}^{2}\Big(P(X_k > Y_k) - P(Y_k > X_k)\Big)\\
&\quad=\sum_{k=1}^{2}\Big(P(X_k > Y_k) + P(X_k = Y_k) - P(X_k = Y_k) - P(Y_k > X_k)\Big)\\
&\quad=\sum_{k=1}^{2}\Big(P(X_k \geq Y_k) - P(Y_k \geq X_k)\Big)
\end{aligned}$$

se deduce el resultado. □

4.2. Preferencia estadística mediante agregación de las componentes

La propuesta realizada en la sección anterior tenía el inconveniente de considerar cada una de las componentes del vector aleatorio por separado sin tener en cuenta la posible relación existente entre ellas. Con el fin de paliar este problema, intentaremos en esta sección dar una extensión que no pierda tanta información. Para ello extenderemos la idea de que la preferencia estadística divide el espacio en tres zonas, una donde la primera variable aleatoria toma valores más altos que la segunda ($\{\omega \in \Omega \mid X(\omega) > Y(\omega)\}$) en el caso unidimensional) y otra zona donde ocurre lo contrario ($\{\omega \in \Omega \mid X(\omega) < Y(\omega)\}$) y una última donde tenemos un "empate" ($\{\omega \in \Omega \mid X(\omega) = Y(\omega)\}$). Definiremos la preferencia estadística como la medida de la primera zona más la mitad de la zona del "empate" .
En particular, nos vamos a fijar en el plano formado por la suma de las marginales de cada vector aleatorio, es decir, comparamos $X_1 + X_2$ frente a $Y_1 + Y_2$. La agregación de estas marginales la denotamos por X_+, Y_+ (Figura 4.2).

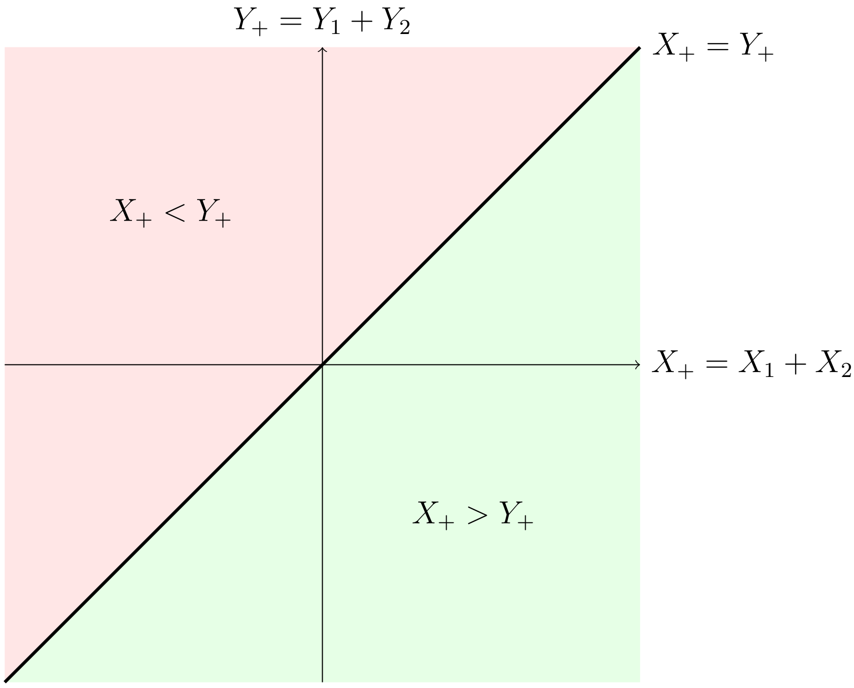

Figura 4.2: Representación gráfica de las zonas consideradas para $\mathcal{Q}^+$.

$$X_+ = X_1 + X_2 \quad \text{y} \quad Y_+ = Y_1 + Y_2.$$

Este enfoque conlleva una agregación de las marginales para posteriormente comparar mediante a un orden estocástico unidimensional, lo cual recuerda a la extensión lineal de la dominancia estocástica. Recordamos que en el caso unidimensional la frontera ente las dos regiones era la recta $\{\omega \in \Omega \mid X(\omega) = Y(\omega)\}$, que separa el espacio en dos semiplanos, uno donde $X > Y$, $\{\omega \in \Omega \mid X(\omega) > Y(\omega)\}$, y otro donde $Y < X$, $\{\omega \in \Omega \mid X(\omega) < Y(\omega)\}$. En este caso tendríamos análogamente las regiones de empate $X_+ = Y_+$, "victoria de X_+" $X_+ > Y_+$ y "derrota de X_+" $X_+ < Y_+$.

Definición 4.3. *Sean $\vec{X}, \vec{Y}$ dos vectores aleatorios de dimensión 2, entonces definimos la relación probabilística estadística bidimensional suma como:*

$$\mathcal{Q}^+(\vec{X}, \vec{Y}) = \mathcal{Q}(X_+, Y_+) = P(X_1 + X_2 > Y_1 + Y_2) + \frac{1}{2} P(X_1 + X_2 = Y_1 + Y_2).$$

Esta extensión esta íntimamente relacionada con la preferencia estadística unidimensional y es una relación probabilística como veremos más adelante. De este modo podríamos considerar que $\vec{X}$ es preferida a $\vec{Y}$ si $Q^+(\vec{X}, \vec{Y}) > 0.5$.
Cuando utilizamos la extensión Q^+ lo que al final se está computando es la relación probabilística unidimensional de una agregación de las componentes, donde en este caso la agregación se realiza mediante la aplicación $+ : \mathbb{R}^2 \to \mathbb{R}$ con $+(\vec{X}) = X_1 + X_2$. De este modo podríamos considerar otras aplicaciones como podría ser por ejemplo una suma ponderada si queremos dar mayor importancia a una componente que a otra. De forma

más general podemos considerar cualquier aplicación $g : \mathbb{R}^2 \to \mathbb{R}$ que sea estrictamente creciente. Imponer el crecimiento tiene dos explicaciones, por un lado garantiza que sea medible, mientras que por otro lado, similarmente a lo que vimos en la Proposición 3.7, preservará el orden existente entre las marginales.

Definición 4.4. *Sean $\vec{X}$, $\vec{Y}$ dos vectores aleatorios de dimensión 2 y $g : \mathbb{R}^2 \to \mathbb{R}$ una aplicación estrictamente creciente. Definimos la relación probabilística bidimensional por agregación como:*

$$\mathcal{Q}^g(\vec{X},\vec{Y}) = \mathcal{Q}\big(g(\vec{X}), g(\vec{Y})\big) = P\big(g(\vec{X}) > g(\vec{Y})\big) + \frac{1}{2}P\big(g(\vec{X}) = g(\vec{Y})\big). \tag{4.2}$$

Como ocurría con $\mathcal{Q}^+$, al estar agregando las componentes para realizar una comparación unidimensional, conservamos la propiedad de relación probabilística.

Proposición 4.3. *La aplicación $\mathcal{Q}^g(\vec{X},\vec{Y})$ definida en la Ecuación* (4.2) *es una relación probabilística.*

Demostración. Vimos que la preferencia estadística unidimensional $(\mathcal{Q}(X,Y))$ era una relación probabilística. De este modo:

$$\mathcal{Q}^g(\vec{X},\vec{Y}) + \mathcal{Q}^g(\vec{Y},\vec{X}) = \mathcal{Q}\big(g(\vec{X}), g(\vec{Y})\big) + \mathcal{Q}\big(g(\vec{Y}), g(\vec{X})\big) = 1. \qquad \square$$

Esta propiedad nos permite dar otra extensión de la preferencia estadística.

Definición 4.5 (Preferencia estadística por agregación de componentes)**.** *Dados dos vectores aleatorios bidimensionales $\vec{X}$, $\vec{Y}$, se dice que:*

- *$\vec{X}$ es preferida estadísticamente por agregación de g a $\vec{Y}$, y se denota por $\vec{X} \succeq^g_{SP} \vec{Y}$, si $\mathcal{Q}^g(\vec{X},\vec{Y}) \geq 0.5$.*
- *$\vec{X}$ es indiferente estadísticamente por agregación de g a $\vec{Y}$, y se denota por $\vec{X} \equiv^g_{SP} \vec{Y}$, si $\mathcal{Q}^g(\vec{X},\vec{Y}) = 0.5$.*
- *$\vec{X}$ es estrictamente preferida estadísticamente por agregación de g a $\vec{Y}$, y se denota por $\vec{X} \succ^g_{SP} \vec{Y}$, si $\mathcal{Q}^g(\vec{X},\vec{Y}) > 0.5$.*

Este nuevo orden estocástico no deja lugar a la incomparabilidad, de modo que $(\mathcal{D}^{(2)}, \succ^g_{SP}, \prec^g_{SP}, \equiv^g_{SP})$ determina una estructura de preferencia sin elementos incomparables. Esta definición nos permite dar una caracterización análoga a la mostrada en la Proposición 3.6.

Proposición 4.4. *Dados dos vectores aleatorios bidimensionales $\vec{X}$, $\vec{Y}$, son equivalentes:*

1. *$\vec{X} \succeq^g_{SP} \vec{Y}$.*

2. $\mathcal{Q}^g(\vec{X},\vec{Y}) \geq \mathcal{Q}^g(\vec{Y},\vec{X})$.

3. $P\big(g(\vec{X}) \geq g(\vec{Y})\big) \geq P\big(g(\vec{Y}) \geq g(\vec{X})\big)$.

4. $P\big(g(\vec{X}) > g(\vec{Y})\big) \geq P\big(g(\vec{Y}) > g(\vec{X})\big)$.

5. $\Big(g(\vec{X}) - g(\vec{Y})\Big) \succeq_{SP} 0_{\mathcal{D}}$ *donde* $0_{\mathcal{D}}$ *es un vector aleatorio degenerado en* $0_{\mathbb{R}^2}$.

Demostración. Por definición $\mathcal{Q}^g(\vec{X},\vec{Y}) = \mathcal{Q}(g(X_1,X_2), g(Y_1,Y_2))$, de modo que aplicando la Proposición 3.6 a $g(X_1,X_2)$ y $g(Y_1,Y_2)$ se deducen las cuatro primeras equivalencias. Vamos a probar que 1. $\iff$ 5. Basta notar que:

$$P\big(g(\vec{X}) > g(\vec{Y})\big) = P\big(g(\vec{X}) - g(\vec{Y}) > 0\big).$$
$$P\big(g(\vec{X}) = g(\vec{Y})\big) = P\big(g(\vec{X}) - g(\vec{Y}) = 0\big).$$

Se cumple que:

$$\begin{aligned} Q^g(\vec{X},\vec{Y}) = Q\big(g(\vec{X}), g(\vec{Y})\big) &= P\big(g(\vec{X}) > g(\vec{Y})\big) + \frac{1}{2}P\big(g(\vec{X}) = g(\vec{Y})\big) \\ &= P\big(g(\vec{X}) - g(\vec{Y}) > 0\big) + \frac{1}{2}P\big(g(\vec{X}) - g(\vec{Y}) = 0\big) = Q(g(\vec{X}) - g(\vec{Y}), 0_{\mathcal{D}}), \end{aligned}$$

y por tanto $\vec{X} \succ^g_{SP} \vec{Y}$ si y solo si $\Big(g(\vec{X}) - g(\vec{Y})\Big) \succeq_{SP} 0_{\mathcal{D}}$. □

4.3. Preferencia estadística inherentemente bivariante

Hasta ahora hemos visto dos posibles extensiones, en primer lugar ($\mathcal{Q}^M$) medíamos las preferencias estadísticas marginales para posteriormente agregarlas; y en segundo lugar, agregábamos las marginales para medir la preferencia ($\mathcal{Q}^g$). Si bien estos dos enfoques resultaban bastante sencillos, el primer enfoque utilizaba únicamente la distribución conjunta de X_k y Y_k mientras que el segundo utilizaba la distribución conjunta de $g(\vec{X})$ y $g(\vec{Y})$. Ninguno de estos utilizaba la distribución conjunta de $\vec{X}$ e $\vec{Y}$ lo que supone una pérdida de información. Es por ello que ahora vamos a plantear un enfoque sin este problema.

Definición 4.6. *Sean $\vec{X}$, $\vec{Y}$ dos vectores aleatorios bidimensionales, definimos la preferencia estadística inherentemente bivariante como:*

$$\begin{aligned} \mathcal{Q}^2(\vec{X},\vec{Y}) : &\mathcal{D} \times \mathcal{D} \to [0,1] \\ &(\vec{X},\vec{Y}) \hookrightarrow P(\vec{X} - \vec{Y} > 0). \end{aligned}$$

Esta extensión de preferencia estadística se basa en el primer sumando de la fórmula unidimensional $\mathcal{Q}(X,Y) = P(X > Y) + \frac{1}{2}P(X = Y)$ la cual viene a ser la parte determinante al establecer la preferencia pues el segundo sumando es común a $\mathcal{Q}(X,Y)$ y $\mathcal{Q}(Y,X)$.

Definición 4.7. *Consideremos dos vectores aleatorios $\vec{X}$, $\vec{Y}$ pertenecientes al mismo espacio de probabilidad, diremos que:*

- *$\vec{X}$ es fuertemente preferida a $\vec{Y}$, y lo denotaremos por $\vec{X} \succ^s_{SP} \vec{Y}$, si $\mathcal{Q}^2(\vec{X},\vec{Y}) > 0.5$.*
- *$\vec{X}$ es moderadamente preferida a $\vec{Y}$, y lo denotaremos por $\vec{X} \succ^m_{SP} \vec{Y}$, si $\mathcal{Q}^2(\vec{X},\vec{Y}) > \mathcal{Q}^2(\vec{Y},\vec{X})$.*
- *$\vec{X}$ es débilmente preferida a $\vec{Y}$, y lo denotaremos por $\vec{X} \succ^w_{SP} \vec{Y}$, si $\mathcal{Q}^2(\vec{Y},\vec{X}) < 0.5$.*
- *$\vec{X}$ es indiferente a $\vec{Y}$, y lo denotaremos por $\vec{X} \equiv_{SP} \vec{Y}$, si $\mathcal{Q}^2(\vec{X},\vec{Y}) = \mathcal{Q}^2(\vec{Y},\vec{X})$.*

Si estudiamos las distintas estructuras de preferencia:
$(\mathcal{D}^{(2)}, \succ^s_{SP}, \prec^s_{SP}, \equiv_{SP})$ determina una estructura de preferencia $\langle P, I, J\rangle$.
$(\mathcal{D}^{(2)}, \succ^m_{SP}, \prec^m_{SP}, \equiv_{SP})$ determina una estructura de preferencia sin elementos incomparables $\langle P, I\rangle$.
$(\mathcal{D}^{(2)}, \succ^w_{SP}, \prec^w_{SP}, \equiv_{SP})$ no determina una estructura de preferencia en general.

Observación 12. *Para los casos de preferencia moderada y fuerte se podría definir la relación binaria característica:*

$$\{\succeq^s_{SP}\} = \{\succ^s_{SP}\} \cup \{\equiv_{SP}\}, \quad \{\succeq^m_{SP}\} = \{\succ^m_{SP}\} \cup \{\equiv_{SP}\}.$$

Sin embargo, para la preferencia débil no es posible definirla pues no definen una estrutura de preferencia al no ser $\succ^w_{SP}$ y $\prec^w_{SP}$ disjuntas.
Por último, $\succeq^m_{SP}$ son relaciones binarias completas mientras que $\succeq^s_{SP}$ no lo es. En el caso de la preferencia débil, $\{\succ^w_{SP}\} \cup \{\prec^w_{SP}\} \cup \{\equiv_{SP}\}$ sí es una relación completa.

De nuevo podemos dar una interpretación geométrica, como se puede ver en la Figura 4.3. Ahora nos fijamos en el plano con abscisas $X_1 - Y_1$ y ordenadas $X_2 - Y_2$; en este caso existirían dos únicas zonas, una donde $\vec{X}$ "gana" a $\vec{Y}$ (en verde) y otra donde $\vec{Y}$ "gana" a $\vec{X}$ (en rojo), los ejes se considerarían como zona de "empate" mientras que los cuadrantes con desigualdades cruzadas de incomparabilidad. Sin embargo, al contrario que en la preferencia estadística unidimensional, las zonas de empate e incomparabilidad no son asignadas a partes iguales entre las dos $\mathcal{Q}^2$, en vez de eso definiremos distintos órdenes de preferencia. En particular, si la probabilidad del cuadrante donde $\vec{X} > \vec{Y}$ (cuadrante verde en la figura) es mayor que 0.5 entonces tendremos la preferencia fuerte; mientras que si únicamente tenemos que la medida de la zona donde $\vec{X} > \vec{Y}$ (verde en la figura) es mayor que la zona donde $\vec{Y} > \vec{X}$ (cuadrante rojo), entonces tendremos preferencia moderada; por último, si la zona donde $\vec{Y} > \vec{X}$ tiene medida menor de 0.5 entonces diremos que tenemos preferencia débil.

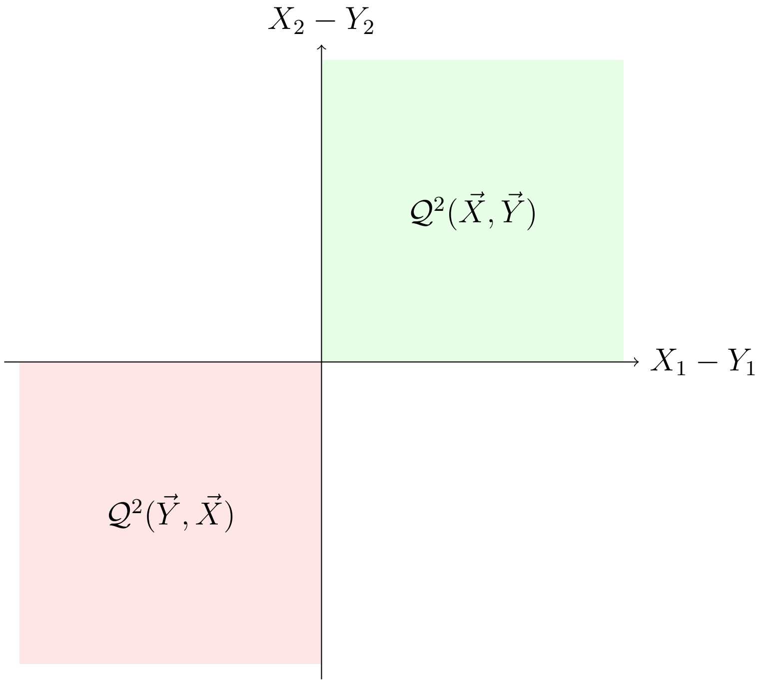

Figura 4.3: Interpretación geométrica de la preferencia inherentemente bivariante $\mathcal{Q}^2(\vec{X}, \vec{Y})$. Existen únicamente dos zonas de interés: $P(\vec{X} > \vec{Y})$ (verde) y $P(\vec{X} < \vec{Y})$ (rojo).

Se demuestra a continuación que, tal y como sus nombres sugieren, la preferencia fuerte implica la preferencia moderada, que a su vez implica la preferencia débil.

Teorema 4.5. *Dados dos vectores aleatorios $\vec{X}$, $\vec{Y}$ pertenecientes al mismo espacio de probabilidad, se cumple:*

$$\vec{X} \succ^s_{SP} \vec{Y} \implies \vec{X} \succ^m_{SP} \vec{Y} \implies \vec{X} \succ^w_{SP} \vec{Y}.$$

Demostración. $\boxed{\vec{X} \succ^s_{SP} \vec{Y} \implies \vec{X} \succ^m_{SP}}$ Como:

$$\{(x, y) \in \mathbb{R}^2 : x_1 < y_1,\, x_2 < y_2\} \subset \{(x, y) \in \mathbb{R}^2 : x_1 > y_1,\, x_2 > y_2\}^c,$$

entonces por la monotonía de la probabilidad se cumple:

$$\mathcal{Q}^2(\vec{Y}, \vec{X}) = P(\vec{X} < \vec{Y}) \leq 1 - P(\vec{X} > \vec{Y}) < 0.5 < \mathcal{Q}^2(\vec{X}, \vec{Y}).$$

$\boxed{\vec{X} \succ_{SP}^{m} \vec{Y} \implies \vec{X} \succ_{SP}^{w}}$ Análogamente al caso anterior:

$$\{(x,y) \in \mathbb{R}^2 : x_1 > y_1,\, x_2 > y_2\} \subset \{(x,y) \in \mathbb{R}^2 : x_1 < y_1,\, x_2 < y_2\}^c.$$

Ahora, utilizando la hipótesis y que $\mathcal{Q}^2(\vec{X},\vec{Y}) \leq 1 - \mathcal{Q}^2(\vec{Y},\vec{X})$:

$$\mathcal{Q}^2(\vec{Y},\vec{X}) < \mathcal{Q}^2(\vec{X},\vec{Y}) \leq 1 - \mathcal{Q}^2(\vec{Y},\vec{X}),$$

de donde se deduce que $\mathcal{Q}^2(\vec{Y},\vec{X}) < 0.5$. $\square$

Ejemplo 4.3. *Consideremos dos vectores aleatorios independientes* $\vec{X} \sim \mathcal{N}_2\Big(\vec{\mu}_X, \dfrac{\sigma^2}{2} I_2\Big)$ *e* $\vec{Y} \sim \mathcal{N}_2\Big(\vec{\mu}_Y, \dfrac{\sigma^2}{2} I_2\Big)$. *Vimos en el Teorema 3.20 que dadas dos normales independientes* $X \sim \mathcal{N}(\mu_X, \sigma_X)$ *e* $Y \sim \mathcal{N}(\mu_Y, \sigma_Y)$*:*

$$P(X > Y) = 1 - \Phi\left(\frac{\mu_Y - \mu_X}{\sqrt{\sigma_X^2 + \sigma_Y^2}}\right).$$

De este modo, si denotamos $u = \dfrac{\mu_Y^{(1)} - \mu_X^{(1)}}{\sigma}$, $v = \dfrac{\mu_Y^{(2)} - \mu_X^{(2)}}{\sigma}$, *entonces:*

$$\mathcal{Q}^2(\vec{X},\vec{Y}) = P(\vec{X} > \vec{Y}) = P(X_1 > Y_1)P(X_2 > Y_2) = (1 - \Phi(u))(1 - \Phi(v)),$$

y análogamente $\mathcal{Q}^2(\vec{Y},\vec{X}) = \Phi(u)\Phi(v)$. *Si representamos estas* $\mathcal{Q}^2$ *como funciones de u y v (Figura 4.4) podemos hacernos una idea de bajo qué condiciones una variable es preferida a otra. Cuando u y v son negativos, es decir* $\vec{\mu}_X > \vec{\mu}_Y$, *entonces en general* $\mathcal{Q}^2(\vec{X},\vec{Y}) > 0.5$ *mientras que cuando son positivos, es decir* $\vec{\mu}_X < \vec{\mu}_Y$, *en general* $\mathcal{Q}^2(\vec{Y},\vec{X}) > 0.5$. *Cuando u y v tienen signos cruzados pero tienen valores absolutos muy distantes entonces si el término negativo es el más alto (en valor absoluto) entonces* $\vec{X} \succ_{SP}^{m} \vec{Y}$ *mientras que si el término positivo es el que domina* $\vec{X} \prec_{SP}^{m} \vec{Y}$. *Lo interesante ocurre cuando los signos de u y v se cruzan y sus valores absolutos son bastante cercanos donde observamos una región de solape entre ambas la cual define la transición de* $\vec{X} \succ_{SP}^{m} \vec{Y}$ *a* $\vec{X} \prec_{SP}^{m} \vec{Y}$. *A continuación se dan algunos valores para* $\sigma^2 = 1$.

$\vec{\mu}_X$	$\vec{\mu}_y$	$\mathcal{Q}^2(\vec{X},\vec{Y})$	$\mathcal{Q}^2(\vec{Y},\vec{X})$
$(1,1)$	$(0,0)$	0.7079	0.0252
$(2,-1)$	$(0,0)$	0.1550	0.0191
$(\frac{1}{2},-1)$	$(0,0)$	0.1097	0.2596
$(1,-1)$	$(0,0)$	0.1335	0.1335

Observamos que en el primer caso donde $\vec{\mu}_X > \vec{\mu}_Y$ *tenemos que* $\vec{X}$ *es fuertemente pre-*

ferida a Y mientras que en el segundo donde $\vec{\mu}_X \not\succ \vec{\mu}_Y$ únicamente tenemos que $\vec{X}$ es moderadamente preferida a $\vec{Y}$. En el tercer caso tenemos que $\vec{X}$ es débilmente preferida a $\vec{Y}$, sin embargo $\vec{Y}$ es moderadamente preferida a $\vec{X}$, este caso da cuenta de lo poco exigente que es la condición de preferencia débil. Por último, tenemos un caso donde $\vec{X}$ indiferente a $\vec{Y}$ bajo preferencia moderada mientras que $\vec{X}$ si es débilmente preferida a $\vec{Y}$.

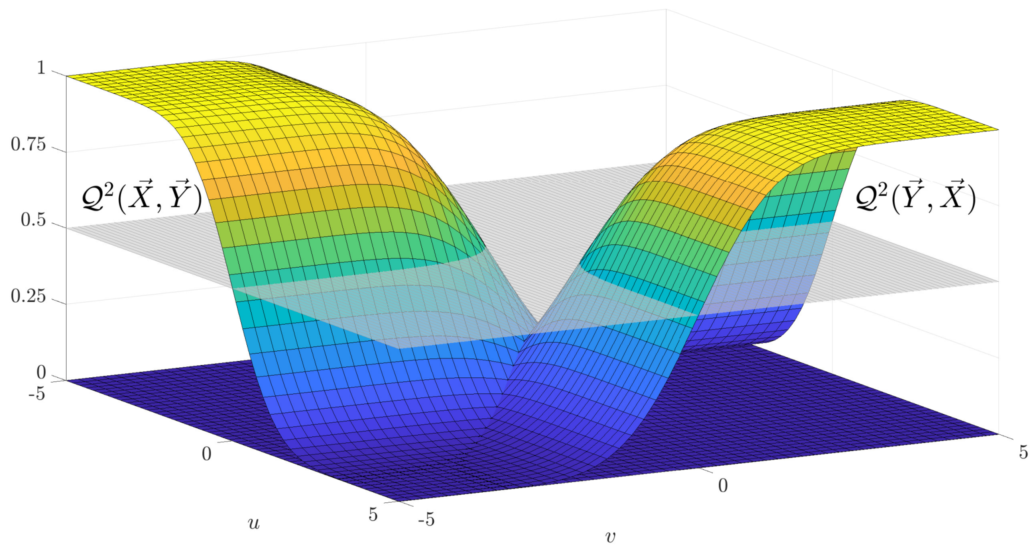

Figura 4.4: Representación gráfica de $\mathcal{Q}^2$ para normales independientes.

4.4. Relación entre las distintas extensiones

Una vez definidas tres extensiones distintas, dando lugar a cinco órdenes de preferencia, procedemos a estudiar cómo se relacionan entre sí. En la preferencia inherentemente bivariante, ya vimos que unas extensiones eran más fuertes que otras. Ahora comparamos las preferencias inherentemente bivariantes con la M-preferencia estadística.

Teorema 4.6. *Dados dos vectores aleatorios $\vec{X}$, $\vec{Y}$ pertenecientes al mismo espacio de probabilidad y $M : [0,1]^2 \to [0,1]$ una aplicación creciente e idempotente, entonces:*

$$\vec{X} \succ_{SP}^{s} \vec{Y} \implies \vec{X} \succ_{SP}^{M} \vec{Y} \implies \vec{X} \succ_{SP}^{w} \vec{Y}.$$

Demostración. Veamos en primer lugar la primera implicación. Por hipótesis $P(\vec{X} > \vec{Y}) > 0.5$, mientras que por otro lado:

$$0.5 < \mathcal{Q}^2(\vec{X},\vec{Y}) = P(\vec{X} > \vec{Y}) = P(\{X_1 > Y_1\} \cap \{X_2 > Y_2\}) \leq \begin{cases} P(X_1 > Y_1), \\ P(X_2 > Y_2). \end{cases}$$

Recordando que $\mathcal{Q}(X,Y) = P(X > Y) + \frac{1}{2}P(X = Y)$ se deduce que $\mathcal{Q}(X_1, Y_1) > 0.5$ y $\mathcal{Q}(X_2, Y_2) > 0.5$. Por último, al ser M una función idempotente y estrictamente creciente se tiene que

$$\mathcal{Q}^M(\vec{X}, \vec{Y}) = M\big(\mathcal{Q}(X_1, Y_1), \mathcal{Q}(X_2, Y_2)\big) > M(0.5, 0.5) = 0.5.$$

En segundo lugar, en la otra implicación se tiene por hipótesis que $\mathcal{Q}^M(\vec{X}, \vec{Y}) > 0.5$. Supongamos ahora por reducción al absurdo que $\vec{X} \not\succ^w_{SP} \vec{Y}$, entonces $P(\vec{X} < \vec{Y}) \geq 0.5$ y razonando exactamente igual que en la primera implicación, llegaríamos a que $\mathcal{Q}(Y_1, X_1), \mathcal{Q}(Y_2, X_2) \geq 0.5$, de modo que utilizando que $\mathcal{Q}$ es una relación probabilística se tendría que $\mathcal{Q}(X_1, Y_1) < 0.5$ y $\mathcal{Q}(X_2, Y_2) < 0.5$ y por tanto (M es idempotente y creciente) $\mathcal{Q}^M(\vec{X}, \vec{Y}) < 0.5$ contradiciendo la hipótesis. $\square$

Por tanto, hemos demostrado que $\mathcal{Q}^M$ ocupa un lugar intermedio entre la preferencia fuerte y la débil, un puesto similar al que tiene la preferencia moderada la cual, como veremos adelante no está en general relacionada con la M-preferencia estadísitca. Ahora nos queda estudiar qué lugar ocupa $\mathcal{Q}^g$.

Teorema 4.7. *Dados $\vec{X}$, $\vec{Y}$ dos vectores aleatorios y una aplicación $g : \mathbb{R}^2 \to \mathbb{R}$ estrictamente creciente, entonces:*

$$\vec{X} \succ^s_{SP} \vec{Y} \implies \vec{X} \succ^g_{SP} \vec{Y} \implies \vec{X} \succ^w_{SP} \vec{Y}.$$

Demostración. Comenzamos con la primera implicación, por hipótesis tenemos que

$$0.5 < P(\{\omega : X_1(\omega) > Y_1(\omega)\} \cap \{\omega : X_2(\omega) > Y_2(\omega)\}).$$

Por ser g una función estrictamente creciente, dado un ω arbitrario con $X_1(\omega) > Y_1(\omega)$ y $X_2(\omega) > Y_2(\omega)$, entonces $g(X_1(\omega), X_2(\omega)) > g(Y_1(\omega), Y_2(\omega))$. Deducimos que:

$$\{\omega : X_1(\omega) > Y_1(\omega)\} \cap \{\omega : X_2(\omega) > Y_2(\omega)\} \subseteq \{\omega : g(X_1(\omega), X_2(\omega)) > g(Y_1(\omega), Y_2(\omega))\}.$$

De este modo se concluye que

$$Q\big(g(\vec{X}), g(\vec{Y})\big) \geq P\big(g(\vec{X}) > g(\vec{Y})\big) \geq P(\vec{X} > \vec{Y}) > 0.5,$$

y por tanto $\vec{X} \succ^g_{SP} \vec{Y}$.
Pasamos ahora a la segunda implicación, de manera análoga al caso anterior tenemos que:

$$\{\omega : X_1(\omega) < Y_1(\omega)\} \cap \{\omega : X_2(\omega) < Y_2(\omega)\} \subseteq \{\omega : g(X_1(\omega), X_2(\omega)) < g(Y_1(\omega), Y_2(\omega))\}.$$

Por hipótesis:

$$P\big(\{\omega : g(X_1(\omega), X_2(\omega)) < g(Y_1(\omega), Y_2(\omega))\}\big) \\ + \frac{1}{2}P\big(\{\omega : g(X_1(\omega), X_2(\omega)) = g(Y_1(\omega), Y_2(\omega))\}\big) < 0.5.$$

En particular:

$$0.5 > P\big(\{\omega : g(X_1(\omega), X_2(\omega)) < g(Y_1(\omega), Y_2(\omega))\}\big) \geq P(\vec{Y} > \vec{X}),$$

concluyendo que $P(\vec{Y} > \vec{X}) < 0.5$. $\square$

Hemos visto cómo se relacionan estas extensiones entre sí, pero ¿cómo se relacionan las marginales con la preferencia unidimensional? Algunas de estas relaciones han sido ya vistas en alguna demostración, de todas formas las reunimos en el siguiente resultado.

Teorema 4.8. *Dados $\vec{X}$, $\vec{Y}$ dos vectores aleatorios en el mismo espacio de probabilidad, las siguientes relaciones se cumplen:*

1. *Si $\vec{X} \succ^s_{SP} \vec{Y}$, entonces $X_1 \succ_{SP} Y_1$ y $X_2 \succ_{SP} Y_2$.*
2. *Si $X_1 \succ_{SP} Y_1$, $X_2 \succ_{SP} Y_2$ y $X_1 - Y_1$ y $X_2 - Y_2$ son independientes, entonces $\vec{X} \succ^m_{SP} \vec{Y}$.*
3. *Si $X_1 \succ_{SP} Y_1$ y $X_2 \succ_{SP} Y_2$, entonces $\vec{X} \succ^w_{SP} \vec{Y}$.*
4. *Si $X_1 \succ_{SP} Y_1$ y $X_2 \succ_{SP} Y_2$, entonces $\vec{X} \succ^M_{SP} \vec{Y}$.*

Demostración. Veamos los distintos puntos:

1. Por hipótesis $P(\vec{X} > \vec{Y}) > 0.5$ de modo que:

   $$0.5 < P(\vec{X} > \vec{Y}) = P(\{X_1 > Y_1\} \cap \{X_2 > Y_2\}) \leq \begin{cases} P(X_1 > Y_1), \\ P(X_2 > Y_2). \end{cases}$$

   Recordando que para dos variables aleatorias unidimensionales X, Y, $\mathcal{Q}(X,Y) = P(X > Y) + \frac{1}{2}P(X = Y)$ deducimos que $X_1 \succ_{SP} Y_1$ y $X_2 \succ_{SP} Y_2$.

2. Utilizando la Proposición 3.6 tenemos que $P(X_1 > Y_1) > P(Y_1 > X_1)$ y $P(X_2 > Y_2) > P(Y_2 > X_2)$. De modo que al ser independientes:

   $$\mathcal{Q}^2(\vec{X}, \vec{Y}) = P(X_1 > Y_1) \cdot P(X_2 > Y_2) > P(X_1 < Y_1) \cdot P(X_2 < Y_2) = \mathcal{Q}^2(\vec{Y}, \vec{X}).$$

3. Por hipótesis se tiene que:

$$0.5 > \mathcal{Q}(Y_1, X_1) = P(X_1 < Y_1) + \frac{1}{2}P(X_1 = Y_1),$$
$$0.5 > \mathcal{Q}(Y_2, X_2) = P(X_2 < Y_2) + \frac{1}{2}P(X_2 = Y_2),$$

de modo que $P(Y_1 > X_1) < 0.5$, $P(Y_2 > X_2) < 0.5$ y por tanto

$$P(\vec{Y} > \vec{X}) = P\big(\{Y_1 > X_1\} \cap \{Y_2 > X_2\}\big) \leq \text{mín}\{P(Y_1 > X_1), P(Y_2 > X_2)\} < 0.5.$$

4. Sea $q = \text{mín}\{\mathcal{Q}(X_1, Y_1), \mathcal{Q}(X_2, Y_2)\} > 0.5$, entonces por ser M idempotente y estrictamente creciente se cumple:

$$\mathcal{Q}(\vec{X}, \vec{Y}) = M\big(\mathcal{Q}(X_1, Y_1), \mathcal{Q}(X_2, Y_2)\big) \geq M(q, q) = q > 0.5. \qquad \square$$

En estas dos últimas secciones hemos visto las relaciones entre las distintas extensiones de la preferencia estadística. Las relaciones demostradas se resumen en la Figura 4.5. Observamos que la preferencia fuerte es la condición más estricta de todas implicando todas las demás, mientras que con la débil ocurre todo lo contrario, todas las demás implican esta. Por otro lado, las preferencias centralizada, moderada y por agregación están en medio de todas las demás. En la Sección A veremos mediante contraejemplos que no existe ninguna relación más entre estas condiciones.

$$\begin{tikzcd}
 & \boxed{\vec{X} \succ_{SP}^{M} \vec{Y}} \arrow[dr] & \\
\boxed{\vec{X} \succ_{SP}^{s} \vec{Y}} \arrow[ur] \arrow[r] \arrow[dr] & \boxed{\vec{X} \succ_{SP}^{m} \vec{Y}} \arrow[r] & \boxed{\vec{X} \succ_{SP}^{w} \vec{Y}} \\
 & \boxed{\vec{X} \succ_{SP}^{g} \vec{Y}} \arrow[ur] &
\end{tikzcd}$$

Figura 4.5: Resumen de las relaciones entre todas las extensiones de la preferencia estadística.

4.5. Relación de las extensiones con la mediana

En el Teorema 3.8 vimos que la preferencia estadística unidimensional guarda una estrecha relación con la mediana. A continuación, vamos a estudiar qué relaciones guardan

las extensiones bivariantes de la preferencia estadística con la mediana. Para ello, nos encontramos con un problema ¿cómo se extiende la mediana a vectores aleatorios bidimensionales?
Existen distintos enfoques a este problema, nosotros definiremos la mediana de un vector aleatorio ($\vec{X}$) como el producto cartesiano de las medianas unidimensionales, es decir $Me(\vec{X}) = Me(X_1) \times Me(X_2)$. En función de si las medianas marginales son un intervalo ó un punto, este conjunto podrá ser un cuadrado, un segmento o un punto. Dadas las distintas formas de este conjunto, se hablará del ínfimo del mismo, refiriéndonos al producto cartesiano de los ínfimos: $(\text{ínf}\, Me(X_1), \text{ínf}\, Me(X_2))$, pudiendo reducir el problema a un único punto.
En particular, se consideran tres notaciones relacionadas con el ínfimo del conjunto de medianas comparado con el valor $(0,0) = 0_{\mathbb{R}^2}$:

1. $\text{ínf}\, Me(\vec{X}-\vec{Y}) > 0_{\mathbb{R}^2}$ denotará que las medianas en ambas componentes son mayores que 0.

2. $\text{ínf}\, Me(\vec{X}-\vec{Y}) \nleq 0_{\mathbb{R}^2}$ denotará que al menos una mediana es mayor que 0.

3. $\text{ínf}\, Me(\vec{X}-\vec{Y}) \not< 0_{\mathbb{R}^2}$ denotará que al menos una mediana es mayor o igual que 0.

Es claro que (1) implica (2), que a su vez implica (3).
Comenzaremos estudiando cómo se relaciona la preferencia estadística inherentemente bivariante con la mediana.

Teorema 4.9. *Dados dos vectores aleatorios bidimensionales $\vec{X}$, $\vec{Y}$ pertenecientes al mismo espacio de probabilidad, se cumple:*

1. $\vec{X} \succ^s_{SP} \vec{Y} \implies \text{ínf}\, Me(\vec{X}-\vec{Y}) > 0_{\mathbb{R}^2} \implies \vec{X} \succ^m_{SP} \vec{Y}$.

2. $\vec{X} \succ^m_{SP} \vec{Y} \implies \text{ínf}\, Me(\vec{X}-\vec{Y}) \not< 0_{\mathbb{R}^2}$.

3. $\text{ínf}\, Me(\vec{X}-\vec{Y}) \nleq 0_{\mathbb{R}^2} \implies \vec{X} \succ^w_{SP} \vec{Y}$.

Demostración. $\boxed{1.}$ Desarrollando la expresión de $\mathcal{Q}^2(\vec{X},\vec{Y})$ como se hizo en la sección anterior, obtenemos:

$$\mathcal{Q}^2(\vec{X},\vec{Y}) = P(\vec{X}-\vec{Y} > 0) = P(\{X_1 - Y_1 > 0\} \cap \{X_2 - Y_2 > 0\}) \leq \begin{cases} P(X_1 - Y_1 > 0), \\ P(X_2 - Y_2 > 0). \end{cases}$$

Por hipótesis $\mathcal{Q}^2(\vec{X},\vec{Y}) > 0.5$, entonces:

$$P(X_1 - Y_1 > 0) > 0.5 \iff \text{ínf}\, Me(X_1 - Y_1) > 0.$$

Análogamente, ínf $Me(X_2 - Y_2) > 0$.
Para la segunda implicación introducimos la notación $Z_1 = X_1 - Y_1$, $Z_2 = X_2 - Y_2$. Desarrollando $\mathcal{Q}^2(\vec{X}, \vec{Y})$:

$$\begin{aligned}
\mathcal{Q}^2(\vec{X}, \vec{Y}) &= P(Z_1 > 0) + P(Z_2 > 0) - P(\{Z_1 > 0\} \cup \{Z_2 > 0\}) \\
&= P(Z_1 > 0) + P(Z_2 > 0) - 1 + P(Z_1, Z_2 \leq 0) \\
&= P(Z_1 > 0) + P(Z_2 > 0) - 1 + \mathcal{Q}^2(\vec{Y}, \vec{X}) \\
&\quad + P(\{Z_1 = 0, Z_2 \leq 0\} \cup \{Z_1 \leq 0, Z_2 = 0\}).
\end{aligned}$$

Hemos llegado a la expresión:

$$\begin{aligned}
&\mathcal{Q}^2(\vec{X}, \vec{Y}) - \mathcal{Q}^2(\vec{Y}, \vec{X}) \\
&\qquad = P(Z_1 > 0) + P(Z_2 > 0) - 1 + P(\{Z_1 = 0, Z_2 \leq 0\} \cup \{Z_1 \leq 0, Z_2 = 0\}).
\end{aligned}$$

Por hipótesis ínf $Me(\vec{X} - \vec{Y}) > 0$, por tanto:

$$\left.\begin{array}{l} P(Z_1 > 0) > 0.5 \\ P(Z_2 > 0) > 0.5 \end{array}\right\} P(Z_1 > 0) + P(Z_2 > 0) - 1 > 0.$$

Por otro lado $P(\{Z_1 = 0, Z_2 \leq 0\} \cup \{Z_1 \leq 0, Z_2 = 0\}) \geq 0$, por lo que llegamos a que $\mathcal{Q}^2(\vec{X}, \vec{Y}) - \mathcal{Q}^2(\vec{Y}, \vec{X}) > 0$.
$\boxed{2.}$ Podemos desarrollar $P(\{Z_1 = 0, Z_2 \leq 0\} \cup \{Z_1 \leq 0, Z_2 = 0\})$ como:

$$\begin{aligned}
&P(Z_1 = 0, Z_2 \leq 0) + P(Z_1 \leq 0, Z_2 = 0) - P(Z_1 = Z_2 = 0) \\
&= P(Z_1 = 0) - P(Z_1 = 0, Z_2 > 0) + P(Z_2 = 0) - P(Z_1 > 0, Z_2 = 0) - P(Z_1 = Z_2 = 0).
\end{aligned}$$

Ahora como en el punto 1 dedujimos que

$$\begin{aligned}
&\mathcal{Q}^2(\vec{X}, \vec{Y}) - \mathcal{Q}^2(\vec{Y}, \vec{X}) \\
&\qquad = P(Z_1 > 0) + P(Z_2 > 0) - 1 + P(\{Z_1 = 0, Z_2 \leq 0\} \cup \{Z_1 \leq 0, Z_2 = 0\}).
\end{aligned}$$

Entonces:

$$\begin{aligned}
\mathcal{Q}^2(\vec{X}, \vec{Y}) - \mathcal{Q}^2(\vec{Y}, \vec{X}) &= P(Z_1 \geq 0) + P(Z_2 \geq 0) - 1 \\
&\quad - \Big[P(Z_1 = 0, Z_2 > 0) + P(Z_1 > 0, Z_2 = 0) + P(Z_1 = Z_2 = 0)\Big].
\end{aligned}$$

Como $\Big[P(Z_1 = 0, Z_2 > 0) + P(Z_1 > 0, Z_2 = 0) + P(Z_1 = Z_2 = 0)\Big] \geq 0$, entonces

$$0 < \mathcal{Q}^2(\vec{X}, \vec{Y}) - \mathcal{Q}^2(\vec{Y}, \vec{X}) \leq P(Z_1 \geq 0) + P(Z_2 \geq 0) - 1.$$

Supongamos por reducción al absurdo que ínf $Me(X_1 - Y_1) < 0$ e ínf $Me(X_2 - Y_2) < 0$. Entonces:

$$P(Z_1 < 0) \geq P(Z_1 \leq \text{ínf } Me(X_1 - Y_1)) \geq 0.5,$$

por lo que $P(Z_1 \geq 0) \leq 0.5$, y análogamente se tiene $P(Z_2 \geq 0) \leq 0.5$. Sin embargo, estas dos condiciones implican que $P(Z_1 \geq 0) + P(Z_2 \geq 0) \leq 1$, contradiciendo la hipótesis inicial. Por tanto, concluimos que al menos uno de los ínfimos de las medianas es no negativo.

$\boxed{3.}$ No es restrictivo suponer que ínf $Me(X_1 - Y_1) > 0$. Ahora, expandiendo $\mathcal{Q}^2(\vec{Y}, \vec{X})$:

$$\mathcal{Q}^2(\vec{Y}, \vec{X}) = P(X_1 - Y_1 < 0, X_2 - Y_2 < 0) = P(Z_1 < 0, Z_2 < 0) \leq P(Z_1 < 0).$$

De este modo:

$$P(Z_1 < 0) \leq P(Z_1 < \text{ínf } Me(Z_1)) \leq 0.5.$$

Supongamos que $0.5 = P(Z_1 < 0) = P(X_1 - Y_1 < 0)$, entonces $0 \geq$ ínf $Me(X_1 - Y_1)$ lo que contradice la hipótesis y por tanto $P(Z_1 < 0) < 0.5$.

Concluimos entonces que $\mathcal{Q}^2(\vec{Y}, \vec{X}) < 0.5$. □

Hemos visto la relación existente entre la mediana y la preferencia inherentemente bivariante, veamos ahora qué relación tiene con la preferencia por centralización.

Teorema 4.10. *Dados dos vectores aleatorios bidimensionales $\vec{X}$, $\vec{Y}$ pertenecientes al mismo espacio de probabilidad y $M : [0,1]^2 \to [0,1]$ una aplicación estrictamente creciente e idempotente, entonces:*

$$\text{ínf } Me(\vec{X} - \vec{Y}) > 0_{\mathbb{R}^2} \implies \vec{X} \succ^M_{SP} \vec{Y} \implies \text{ínf } Me(\vec{X} - \vec{Y}) \not< 0_{\mathbb{R}^2}. \tag{4.3}$$

Demostración. Queremos ver que $\mathcal{Q}^M(\vec{X}, \vec{Y}) > 0.5$. Por hipótesis ínf $Me(\vec{X} - \vec{Y}) > 0$ de modo que utilizando el Teorema 3.8, entonces $q = \text{mín}\{\mathcal{Q}(X_1, Y_1), \mathcal{Q}(X_2, Y_2)\} > 0.5$. Utilizando que la función M es estrictamente creciente e idempotente:

$$\mathcal{Q}^M(\vec{X}, \vec{Y}) = M\big(\mathcal{Q}(X_1, Y_1), \mathcal{Q}(X_2, Y_2)\big) \geq M(q, q) = q > 0.5.$$

Veamos ahora la segunda implicación. Supongamos por reducción al absurdo que ambas medianas son negativas. Usando de nuevo el Teorema 3.8, se tiene que $Q(X_1, Y_1) < 0.5$ y

$\mathcal{Q}(X_2, Y_2) < 0.5$. Entonces utilizando de nuevo las propiedades de la función M:

$$\mathcal{Q}^M(\vec{X}, \vec{Y}) = M\big(\mathcal{Q}(X_1, Y_1), \mathcal{Q}(X_2, Y_2)\big) \leq M(0.5, 0.5) = 0.5 \implies \mathcal{Q}^M(\vec{X}, \vec{Y}) \leq 0.5,$$

contradiciendo la hipótesis. □

En el caso de $\mathcal{Q}^g$, debido a su estrecha relación con el caso unidimensional, obtenemos un resultado totalmente análogo al Teorema 3.8.

Teorema 4.11. *Dados $\vec{X}$, $\vec{Y}$ dos vectores aleatorios bidimensionales y una aplicación $g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ estrictamente creciente, entonces:*

1. $\sup Me\big(g(\vec{X}) - g(\vec{Y})\big) > 0 \implies \vec{X} \succeq^g_{SP} Y \implies \sup Me\big(g(\vec{X}) - g(\vec{Y})\big) \geq 0.$
2. $\vec{X} \succ^g_{SP} \vec{Y} \implies Me\big(g(\vec{X}) - g(\vec{Y})\big) \subseteq [0, \infty).$
3. *El recíproco de 2. no es cierto, sin embargo:*

$$\text{ínf}\, Me\big(g(\vec{X}) - g(\vec{Y})\big) > 0 \implies \vec{X} \succeq^g_{SP} \vec{Y}.$$

4. *Si $P(g(\vec{X}) = g(\vec{Y})) = 0$, entonces:*

$$\vec{X} \succ^g_{SP} \vec{Y} \iff \text{ínf}\, Me\big(g(\vec{X}) - g(\vec{Y})\big) > 0.$$

Sin embargo, aunque $P(g(\vec{X}) = g(\vec{Y})) = 0$, $0 \in Me(g(\vec{X}) - g(\vec{Y}))$ no es equivalente a $\mathcal{Q}^g(\vec{X}, \vec{Y}) = 0.5$.

Demostración. El resultado es directo sin más que tener en cuenta que por definición $\mathcal{Q}^g(\vec{X}, \vec{Y}) = \mathcal{Q}\big(g(\vec{X}), g(\vec{Y})\big)$. □

Sin embargo, no hay implicaciones directas con las medianas, esto se demostrará mediante contraejemplos en el Capítulo A.
Previamente habíamos reunido las relaciones entre las distintas extensiones en la Figura 4.6, ahora añadimos a esta cómo se relacionan con la mediana en la Figura 4.6). Observamos que la condición ínf $Me(\vec{X} - \vec{Y}) \nless 0$ define una condición muy débil, al nivel de la preferencia débil. Por otro lado ínf $Me(\vec{X} - \vec{Y}) \nleq 0$ define una condición al nivel de la moderada y los dos definidas mediante agregación ($\mathcal{Q}^M$, $\mathcal{Q}^g$). Por último, ínf $Me(\vec{X} - \vec{Y}) > 0$ es una condición algo más fuerte que la preferencia moderada y por agregación de la preferencia unidimensional ($\mathcal{Q}^M$). Por último, indicar que todas las relaciones que no están indicadas en la Figura 4.6 no se dan. En el Apéndice A se exponen distintos contraejemplos de estas relaciones.

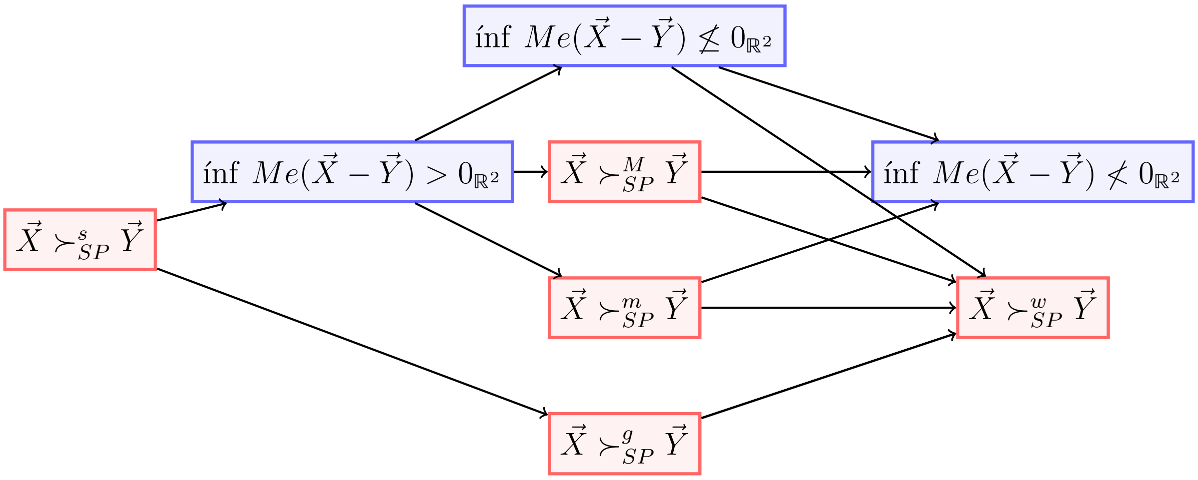

Figura 4.6: Resumen de las relaciones entre todas las extensiones de la preferencia estadística y la mediana.

Capítulo 5

Conclusiones

A lo largo de este trabajo se han estudiado dos órdenes estocásticos diferentes, la dominancia estocástica y la preferencia estadística, proponiendo finalmente tres extensiones bivariantes distintas del segundo de estos órdenes estocásticos.
En el caso unidimensional, la preferencia estadística constituye una buena alternativa a la dominancia estocástica al poder establecer una preferencia en casos de incomparabilidad. Este orden presenta otras ventajas como son el uso toda la información de la distribución conjunta y la graduación de la preferencia en una escala de cero a uno. El principal problema que posee la preferencia estadística es la falta de transitividad, lo cual impide establecer un orden. Esto no ocurre en la dominancia estocástica, la cual define un orden parcial (cuando se considera el conjunto de alternativas como el conjunto de distribuciones). Por tanto, la preferencia estadística no deber verse como una mejora de la dominancia estocástica, sino como otra herramienta complementaria la cual puede resultar útil en algunos contextos.
Para la preferencia estadística, se ha estudiado también su relación con la simetría radial de la función de densidad conjunta, lo que ha permitido caracterizar en este contexto la preferencia estadística a través únicamente del eje de simetría. Esto resulta útil, pues son numerosas las distribuciones con simetría radial en estadística.
El objetivo principal del trabajo era la extensión al caso bivariante, donde se han propuesto tres posibles extensiones. En primer lugar, se ha realizado una extensión mediante agregación de las preferencias estadísticas marginales. Esta extensión destaca por su sencillez. Al apoyarse en gran medida de la preferencia estadística univariante, permite el uso de todos los resultados univariantes previos. Esta además es en algunos contextos relevantes (como cuando se utiliza la media aritmética como función de agregación) una relación probabilística, lo cual es un valor añadido aunque no necesario, pues define una estructura de preferencia sin elementos incomparables. Su principal problema es el uso de información marginal. Esta sencillez bajo la premisa de un uso marginal de la información recuerda a la dominancia estocástica unidimensional, donde se tenía una situación

similar.
En segundo lugar, se ha realizado una extensión mediante la agregación de las componentes. El uso de esta sea seguramente más limitado que el de las otras extensiones. Esto se debe a que para su aplicación, debe tener sentido esta agregación de las componentes marginales del vector aleatorio, lo cual en algunos contextos puede no resultar razonable. Esta tiene la ventaja de nuevo de que se apoya en la preferencia estadística univariante, lo cual permitiría hacer uso de sus propiedades.
Por último, se ha propuesto una extensión puramente bivariante con tres posibles grados de preferencia: fuerte, moderada y débil. La preferencia fuerte es una condición muy exigente existiendo variables aleatorias incomparables cuando se hace uso de esta. Este es un problema idéntico al que poseía la dominancia estocástica unidimensional. Sin embargo, cuando este grado de preferencia se diera, este arrojaría unos argumentos muy sólidos para escoger una de las alternativas. La preferencia moderada seguramente constituya la extensión más interesante de todas las realizadas, pues no posee ninguna de las desventajas descritas hasta ahora. Esta utiliza la información completa de los vectores aleatorios, define una estructura de preferencia sin elementos incomparables y es bivariante pura. Uno de los problemas que se podrían considerar es que no define una relación probabilística. Sin embargo, al definir una estructura de preferencia sin elementos incomparables, no le es necesario definirla. Por último, la preferencia débil es la extensión más limitada de todas las realizadas. Esta ni siquiera define una estructura de preferencia, existiendo variables aleatorias las cuales son simultáneamente preferidas entre sí. Esto supone un gran problema, pues con los órdenes estocásticos se pretende establecer una decisión lógica entre dos alternativas y con este orden esto no es posible.
En el caso unidimensional, la preferencia estadística poseía una estrecha relación con la mediana [14]. En este trabajo se ha estudiado cómo era la relación en el caso bivariante, considerando la mediana como el producto cartesiano de las medianas marginales. Sin embargo, se podrían considerar otra clase de extensiones de la mediana (por ejemplo una extensión geométrica) y comprobar si sigue existiendo esta relación.
Otra rama que sería interesante de abordar sería la posible relación entre las extensiones bivariantes de la dominancia estocástica [17, 23] y la preferencia estadística. Resultaría razonable que bajo ciertas condiciones (por ejemplo independencia) unas extensiones implicaran otras, hecho que ya ocurría en el caso univariante (Teorema 7).
Aunque las extensiones multivariantes hayan sido realizadas únicamente para vectores aleatorios de dimensión 2, sus extensiones a casos de dimensión n serían directas a partir de las definiciones. Si se realizaran estas extensiones, seguramente la extensión n-dimensional de la preferencia fuerte sería la menos adecuada, pues la condición $P(\vec{X} - \vec{Y}) > 0.5$ sería muy exigente y en rara ocasión se daría. A la preferencia débil, le ocurriría todo lo contrario, se daría con bastante frecuencia lo que de nuevo la haría poco adecuada. El resto de las extensiones seguramente seguirían siendo adecuadas, pues

estas se basan en la comparación de $\mathcal{Q}^*(\vec{X},\vec{Y})$ frente $\mathcal{Q}^*(\vec{Y},\vec{X})$ y no en la comparación de $\mathcal{Q}^*(\vec{X},\vec{Y})$ frente a un valor entre cero y uno como ocurría con las otras dos extensiones. Finalmente, indicar que el contenido del último capítulo, el cual se ha desarrollado de manera original, será presentado en el congreso *Soft Methods in Probability and Statistics* (SMPS'2024). Este congreso tendrá lugar en Salzburgo (Austria) del 3 al 6 de septiembre de 2024, y sus actas se incluyen como capítulos de un libro de la editorial Springer.

Apéndice A

Contraejemplos

En Capítulo 4 se han visto 5 órdenes de preferencia y 3 órdenes para la mediaana y cómo se relacionan entre ellos. Ahora, nos queda demostrar que las implicaciones que no se han expuesto no son ciertas en general, para ello consideraremos distintos ejemplos en esta sección. Para facilitar al lector el entendimiento de esta sección se presenta a continuación un resumen con las distintas relaciones junto con los resultados o contraejemplos correspondientes.

$$\boxed{\vec{X} \succ^{w}_{SP} \vec{Y}}$$

No implica

- $\text{ínf}(Me(\vec{X}-\vec{Y})) \not< 0$ Ejemplo A.3.
- $\text{ínf}(Me(\vec{X}-\vec{Y})) \not\leq 0$ Ejemplo A.3.
- $\text{ínf}(Me(\vec{X}-\vec{Y})) > 0$ Ejemplo A.3.
- $\vec{X} \succ^{M}_{SP} \vec{Y}$ Ejemplo A.1.
- $\vec{X} \succ^{m}_{SP} \vec{Y}$ Ejemplo A.1 ($\vec{Y}$, $\vec{X}$ invertidas).
- $\vec{X} \succ^{g}_{SP} \vec{Y}$ Ejemplo A.1.
- $\vec{X} \succ^{s}_{SP} \vec{Y}$ Ejemplo A.1.

$$\boxed{\text{ínf}(Me(\vec{X}-\vec{Y})) \not< 0}$$

No implica

- $\text{ínf}(Me(\vec{X}-\vec{Y})) \not\leq 0$ Por definición.

- $\text{ínf}(Me(\vec{X}-\vec{Y})) > 0$ — Por definición.
- $\vec{X} \succ^{w}_{SP} \vec{Y}$ — Ejemplo A.2.
- $\vec{X} \succ^{M}_{SP} \vec{Y}$ — Ejemplo A.1.
- $\vec{X} \succ^{m}_{SP} \vec{Y}$ — Ejemplo A.1 ($\vec{Y}$, $\vec{X}$ invertidas).
- $\vec{X} \succ^{g}_{SP} \vec{Y}$ — Ejemplo A.1.
- $\vec{X} \succ^{s}_{SP} \vec{Y}$ — Ejemplo A.1.

$$\boxed{\text{ínf}(Me(X-Y)) \not\leq 0}$$

Implica

- $\text{ínf}(Me(\vec{X}-\vec{Y})) \not< 0$ — Por definición.
- $\vec{X} \succ^{w}_{SP} \vec{Y}$ — Teorema 4.9.

No implica

- $\text{ínf}(Me(\vec{X}-\vec{Y})) > 0$ — Por definición.
- $\vec{X} \succ^{M}_{SP} \vec{Y}$ — Ejemplo A.5.
- $\vec{X} \succ^{m}_{SP} \vec{Y}$ — Ejemplo A.5.
- $\vec{X} \succ^{g}_{SP} \vec{Y}$ — Ejemplo A.3 ($\vec{Y}$, $\vec{X}$ invertidas).
- $\vec{X} \succ^{s}_{SP} \vec{Y}$ — Ejemplo A.3 ($\vec{Y}$, $\vec{X}$ invertidas).

$$\boxed{\vec{X} \succ^{M}_{SP} \vec{Y}}$$

Implica

- $\text{ínf}(Me(\vec{X}-\vec{Y})) \not< 0$ — Teorema 4.10.
- $\vec{X} \succ^{w}_{SP} \vec{Y}$ — Teorema 4.6.

No implica

- $\text{ínf}(Me(\vec{X}-\vec{Y}))>0$ — Ejemplo A.1 ($\vec{Y}$, $\vec{X}$ invertidas).
- $\text{ínf}(Me(X-Y))\not\leq 0$ — Ejemplo A.1 ($\vec{Y}$, $\vec{X}$ invertidas).
- $\vec{X}\succ^{m}_{SP}\vec{Y}$ — Ejemplo A.1 ($\vec{Y}$, $\vec{X}$ invertidas).
- $\vec{X}\succ^{g}_{SP}\vec{Y}$ — Ejemplo A.4.
- $\vec{X}\succ^{s}_{SP}\vec{Y}$ — Ejemplo A.1 ($\vec{Y}$, $\vec{X}$ invertidas).

$$\boxed{\vec{X}\succ^{m}_{SP}\vec{Y}}$$

Implica

- $\text{ínf}(Me(\vec{X}-\vec{Y}))\not< 0$ — Teorema 4.9.
- $\vec{X}\succ^{w}_{SP}\vec{Y}$ — Teorema 4.5.

No implica

- $\text{ínf}(Me(\vec{X}-\vec{Y}))>0$ — Ejemplo A.1.
- $\text{ínf}(Me(X-Y))\not\leq 0$ — Ejemplo A.1.
- $\vec{X}\succ^{M}_{SP}\vec{Y}$ — Ejemplo A.1.
- $\vec{X}\succ^{g}_{SP}\vec{Y}$ — Ejemplo A.1.
- $\vec{X}\succ^{s}_{SP}\vec{Y}$ — Ejemplo A.1.

$$\boxed{\vec{X}\succ^{g}_{SP}\vec{Y}}$$

Implica

- $\vec{X}\succ^{w}_{SP}\vec{Y}$ — Teorema 4.7.

No implica

- $\text{ínf}(Me(\vec{X} - \vec{Y})) \nless 0$ Ejemplo A.3.
- $\text{ínf}(Me(\vec{X} - \vec{Y})) > 0$ Ejemplo A.3 ($\vec{Y}$, $\vec{X}$ invertidas).
- $\text{ínf}(Me(X - Y)) \nleq 0$ Ejemplo A.3 ($\vec{Y}$, $\vec{X}$ invertidas).
- $\vec{X} \succ_{SP}^{M} \vec{Y}$ Ejemplo A.3.
- $\vec{X} \succ_{SP}^{m} \vec{Y}$ Ejemplo A.3 ($\vec{Y}$, $\vec{X}$ invertidas).
- $\vec{X} \succ_{SP}^{s} \vec{Y}$ Ejemplo A.3 ($\vec{Y}$, $\vec{X}$ invertidas).

$$\boxed{\text{ínf}(Me(X - Y)) > 0}$$

Implica

- $\text{ínf}(Me(X - Y)) \nless 0$ Por definición.
- $\text{ínf}(Me(X - Y)) \nleq 0$ Por definición.
- $\vec{X} \succ_{SP}^{M} \vec{Y}$ Teorema 4.10.
- $\vec{X} \succ_{SP}^{m} \vec{Y}$ Teorema 4.9.
- $\vec{X} \succ_{SP}^{w} \vec{Y}$ Consecuencia de los Teoremas 4.9 y 4.5.

No implica

- $\vec{X} \succ_{SP}^{g} \vec{Y}$ Ejemplo A.3 ($\vec{Y}$, $\vec{X}$ invertidas).
- $\vec{X} \succ_{SP}^{s} \vec{Y}$ Ejemplo A.3 ($\vec{Y}$, $\vec{X}$ invertidas).

$$\boxed{\vec{X} \succ_{SP}^{s} \vec{Y}}$$

Implica

- $\text{ínf}(Me(X - Y)) \nless 0$ Consecuencia Teorema 4.9.
- $\text{ínf}(Me(X - Y)) \nleq 0$ Consecuencia Teorema 4.9.
- $\text{ínf}(Me(X - Y)) > 0$ Teorema 4.9.

- $\vec{X} \succ^{M}_{SP} \vec{Y}$ Teorema 4.6.
- $\vec{X} \succ^{m}_{SP} \vec{Y}$ Teorema 4.5.
- $\vec{X} \succ^{g}_{SP} \vec{Y}$ Teorema 4.7.
- $\vec{X} \succ^{w}_{SP} \vec{Y}$ Consecuencia de los Teoremas 4.9 y 4.5.

Ejemplo A.1. *Consideramos $g = +$, $M = \text{am}$ y los vectores aleatorios independientes $\vec{X}$ e $\vec{Y}$ con $\vec{Y} = 0_{\mathcal{D}^{(2)}}$ y $\vec{X}$ dado por:*

		X_2			
		-1	0	1	
	-2	0.1	0	0.3	0.4
X_1	0	0.3	0.1	0	0.4
	1	0	0	0.2	0.2
		0.4	0.1	0.5	

En primer lugar, $Y_1 + Y_2 = 0_{\mathcal{D}^{(1)}}$ y calculamos la distribución de $X_1 + X_2$:

$X_1 + X_2$	-3	-2	-1	0	1	2
	0.1	0	0.6	0.1	0	0.2

Calculamos ahora las distintas $\mathcal{Q}$:

$$\left.\begin{array}{r}\mathcal{Q}(X_1,Y_1) = P(X_1=1) + \frac{1}{2}P(X_1=0) = 0.4\\ \mathcal{Q}(X_2,Y_2) = P(X_2=1) + \frac{1}{2}P(X_2=0) = 0.55\end{array}\right\} \mathcal{Q}^{am}(\vec{X},\vec{Y}) = 0.475 \implies \vec{Y} \succ^{am}_{SP} \vec{X}.$$

$$\begin{array}{r}\mathcal{Q}^2(\vec{X},\vec{Y}) = P(X_1=1, X_2=1) = 0.2\\ \mathcal{Q}^2(\vec{Y},\vec{X}) = P(X_1=-2, X_2=-1) = 0.1\end{array} \implies \left\{\begin{array}{l}\vec{X} \succ^{m}_{SP} \vec{Y},\\ \vec{X} \succ^{w}_{SP} \vec{Y},\\ \vec{Y} \succ^{w}_{SP} \vec{X}.\end{array}\right.$$

$$\mathcal{Q}^+(\vec{X},\vec{Y}) = P(X_+ = 1) + P(X_+ = 2) + \frac{1}{2}P(X_+ = 0) = 0.25 \implies \vec{Y} \succ^{+}_{SP} \vec{X}.$$

Estudiemos ahora las medianas:

$$\left.\begin{array}{rcl}Me(X_1 - Y_1) & = & 0\\ Me(X_2 - Y_2) & = & [0,1]\end{array}\right\} \text{ínf } Me(\vec{X} - \vec{Y}) = (0,0),$$

$$\left.\begin{array}{rcl} Me(Y_1 - X_1) & = & 0 \\ Me(Y_2 - X_2) & = & [-1, 0] \end{array}\right\} \text{ínf } Me(\vec{Y} - \vec{X}) = (0, -1).$$

Por tanto tenemos que:

✓ $\vec{X} \succ^{w}_{SP} \vec{Y}$	✓ $\vec{Y} \succ^{w}_{SP} \vec{X}$
× $\vec{X} \succ^{M}_{SP} \vec{Y}$	✓ $\vec{Y} \succ^{M}_{SP} \vec{X}$
× $\vec{X} \succ^{g}_{SP} \vec{Y}$	✓ $\vec{Y} \succ^{g}_{SP} \vec{X}$
✓ $\vec{X} \succ^{m}_{SP} \vec{Y}$	× $\vec{Y} \succ^{m}_{SP} \vec{X}$
× $\vec{X} \succ^{s}_{SP} \vec{Y}$	× $\vec{Y} \succ^{s}_{SP} \vec{X}$
✓ ínf $Me(\vec{X} - \vec{Y}) \not< 0$	✓ ínf $Me(\vec{Y} - \vec{X}) \not< 0$
× ínf $Me(\vec{Y} - \vec{X}) \not\leq 0$	× ínf $Me(\vec{Y} - \vec{X}) \not\leq 0$
× ínf $Me(\vec{X} - \vec{Y}) > 0$	× ínf $Me(\vec{Y} - \vec{X}) > 0$

Como enunciamos anteriormente, este ejemplo nos permite probar que muchas relaciones no se cumplen, a destacar la relación que se muestra entre $\mathcal{Q}^M$ y $\mathcal{Q}^m$, demostrando que en general ninguna es más fuerte que la otra. Otro aspecto relevante es lo débil que es la condición de preferencia moderada, pues en este caso esta preferencia se cumple en ambos sentidos. Por último comprobamos también que $\vec{Y} \succ^{M}_{SP} \vec{X}$ no implica que $Y_1 \succ_{SP} Y_1$ e $Y_1 \succ_{SP} Y_2$.

Ejemplo A.2. *Consideremos $g = +$, $M = \mathrm{am}$ y el vector aleatorio $\vec{Y} = 0_{\mathcal{D}^{(2)}}$. Consideraremos también el vector aleatorio $\vec{X}$ que se define de la siguiente manera. Sea $\vec{U} \sim \mathcal{N}_2(0_{\mathbb{R}^2}, \mathbb{1})$, y definimos $\vec{X}$ como $X_1 = U_1$ y*

$$X_2 = \begin{cases} U_2 & si\ U_1 \cdot U_2 > 0. \\ -U_2 & si\ U_1 \cdot U_2 \leq 0. \end{cases}$$

Esto lo que quiere decir es que al cuadrante donde $X, Y > 0$ se le asigna la probabilidad correspondiente a este cuadrante, así como la correspondiente al cuadrante donde $X < 0, Y > 0$. De manera similar, al cuadrante donde $X, Y < 0$ se le asigna la probabilidad de este mismo cuadrante y del cuadrante donde $X < 0, Y > 0$.
De este modo, por definición, $X_1 \sim \mathcal{N}(0, 1)$. Por otro lado, para X_2 se cumple que, si

$x > 0$:

$$\begin{aligned} P(X_2 \leq x) &= P(U_1 \geq 0, U_2 \leq x) + P(U_1 \leq 0, U_2 \geq -x) \\ &= P(U_1 \geq 0)P(U_2 \leq x) + P(U_1 \leq 0)P(U_2 \geq -x) \\ &= \Phi(0)\Big(P(U_2 \leq x) + P(U_2 \leq x)\Big) = P(U_2 \leq x), \end{aligned}$$

mientras que si $x < 0$:

$$\begin{aligned} P(X_2 \leq x) &= P(U_1 \leq 0, U_2 \leq x) + P(U_1 \geq 0, U_2 \geq -x) \\ &= P(U_1 \leq 0)P(U_2 \leq x) + P(U_1 \geq 0)P(U_2 \geq -x) \\ &= \Phi(0)\Big(P(U_2 \leq x)) + P(U_2 \leq x)\Big) = P(U_2 \leq x), \end{aligned}$$

donde hemos vuelto a utilizar Φ para denotar la función de distribución de una normal estándar. Se observa que X_2 y U_2 tienen la misma distribución, por lo que $X_2 \sim \mathcal{N}(0,1)$. Por tanto $\vec{X}$ tiene marginales con distribución normal estándar, sin embargo la conjunta no es una normal bivariante pues se tiene que:

$$P(X_1 > 0, X_2 < 0) = P(X_1 > 0, X_2 < 0) = 0.$$

Si ahora calculamos las distintas $\mathcal{Q}^*$, se obtiene:

$$\begin{aligned} \mathcal{Q}^2(\vec{X},\vec{Y}) = P(\vec{X} > 0) = 0.5 \\ \mathcal{Q}^2(\vec{Y},\vec{X}) = P(\vec{X} < 0) = 0.5 \end{aligned} \implies \begin{cases} \vec{X} \equiv^s_{SP} \vec{Y}, \\ \vec{X} \equiv^m_{SP} \vec{Y}, \\ \vec{X} \equiv^w_{SP} \vec{Y}. \end{cases}$$

Debido a la simetría, se tiene que $Me(\vec{X}) = \vec{\mu} = 0_{\mathbb{R}^2} \not< 0_{\mathbb{R}^2}$. Por tanto, utilizando el Teorema 3.24:

$$\left.\begin{aligned} X_1 \equiv_{SP} Y_1 \\ X_2 \equiv_{SP} Y_2 \end{aligned}\right\} \mathcal{Q}^M(\vec{X},\vec{Y}) = 0.5 \implies \vec{X} \equiv^{am}_{SP} \vec{Y}.$$

$$\begin{aligned} \mathcal{Q}^+(\vec{X},\vec{Y}) &= P(X_1 + X_2 > 0) \\ &= P(U_1 + U_2 > 0, U_1 > 0, U_2 > 0) + P(U_1 - U_2 > 0, U_1 > 0, U_2 < 0) \\ &= P(U_1, U_2 > 0) + P(U_1 > 0, U_2 < 0) = 0.5, \end{aligned}$$

por lo que $\vec{X} \equiv^+_{SP} \vec{Y}$. En este caso tendríamos:

- $\times$ $\vec{X} \succ^w_{SP} \vec{Y}$
- $\times$ $\vec{X} \succ^M_{SP} \vec{Y}$

× $\vec{X} \succ^{g}_{SP} \vec{Y}$

× $\vec{X} \succ^{m}_{SP} \vec{Y}$

× $\vec{X} \succ^{s}_{SP} \vec{Y}$

✓ $\inf Me(\vec{X} - \vec{Y}) \nless 0$

× $\inf Me(\vec{Y} - \vec{X}) \nleq 0$

× $\inf Me(\vec{X} - \vec{Y}) > 0$

× $\vec{Y} \succ^{w}_{SP} \vec{X}$

× $\vec{Y} \succ^{M}_{SP} \vec{X}$

× $\vec{Y} \succ^{g}_{SP} \vec{X}$

× $\vec{Y} \succ^{m}_{SP} \vec{X}$

× $\vec{Y} \succ^{s}_{SP} \vec{X}$

✓ $\inf Me(\vec{Y} - \vec{X}) \nless 0$

× $\inf Me(\vec{Y} - \vec{X}) \nleq 0$

× $\inf Me(\vec{Y} - \vec{X}) > 0$

La importancia de este ejemplo radica en que hemos encontrado un ejemplo en el que $Me(\vec{X}) \nless 0$ no implica $\vec{X} \succ^{w}_{SP} \vec{Y}$, ni ninguna de las otras relaciones estudiadas.

Ejemplo A.3. *Consideremos $\vec{Y} = 0_{\mathcal{D}^{(1)}}$, $g = +$ y $M = \mathrm{am}$. Definimos $\vec{X}$ como:*

		X_2		
		-1	2	
X_1	-1	0.4	0.3	0.7
	2	0.3	0	0.3
		0.7	0.3	

Se cumple que $Y_1 + Y_2 = 0_{\mathcal{D}^{(2)}}$ mientras que la distribución de $X_1 + X_2$ viene dada por:

$X_1 + X_2$	-2	1	4
	0.4	0.6	0

Calculamos ahora las distintas $\mathcal{Q}^*$:

$$\left.\begin{aligned} \mathcal{Q}(X_1, Y_1) = P(X_1 = 2) = 0.3 \\ \mathcal{Q}(X_2, Y_2) = P(X_2 = 2) = 0.3 \end{aligned}\right\} \mathcal{Q}^{am}(\vec{X}, \vec{Y}) = 0.3 \implies \vec{Y} \succ^{am}_{SP} \vec{X}.$$

Los valores de Q^2 son:

$$\begin{aligned} \mathcal{Q}^2(\vec{X}, \vec{Y}) = P(X_1 = 2, X_2 = 2) = 0 \\ \mathcal{Q}^2(\vec{Y}, \vec{X}) = P(X_1 = -1, X_2 = -1) = 0.4 \end{aligned} \implies \begin{cases} \vec{Y} \succ^{m}_{SP} \vec{X}, \\ \vec{Y} \succ^{w}_{SP} \vec{X}, \\ \vec{X} \succ^{w}_{SP} \vec{Y}. \end{cases}$$

Y por último, los valores de Q^+ son:

$$\mathcal{Q}^+(\vec{X},\vec{Y}) = P(X_+ = 1) = 0.6 \implies \vec{X} \succ^+_{SP} \vec{Y}.$$

Estudiemos ahora las medianas:

$$\left.\begin{array}{rcl} Me(X_1 - Y_1) & = & -1 \\ Me(X_2 - Y_2) & = & -1 \end{array}\right\} \text{ínf } Me(\vec{X} - \vec{Y}) = (-1,-1),$$

$$\left.\begin{array}{rcl} Me(Y_1 - X_1) & = & 1 \\ Me(Y_2 - X_2) & = & 1 \end{array}\right\} \text{ínf } Me(\vec{Y} - \vec{X}) = (1,1).$$

Por tanto, tendríamos:

✓ $\vec{X} \succ^w_{SP} \vec{Y}$	✓ $\vec{Y} \succ^w_{SP} \vec{X}$
× $\vec{X} \succ^M_{SP} \vec{Y}$	✓ $\vec{Y} \succ^M_{SP} \vec{X}$
✓ $\vec{X} \succ^g_{SP} \vec{Y}$	× $\vec{Y} \succ^g_{SP} \vec{X}$
× $\vec{X} \succ^m_{SP} \vec{Y}$	✓ $\vec{Y} \succ^m_{SP} \vec{X}$
× $\vec{X} \succ^s_{SP} \vec{Y}$	× $\vec{Y} \succ^s_{SP} \vec{X}$
× $\text{ínf } Me(\vec{X} - \vec{Y}) \not< 0$	✓ $\text{ínf } Me(\vec{Y} - \vec{X}) \not< 0$
× $\text{ínf } Me(\vec{Y} - \vec{X}) \not\leq 0$	✓ $\text{ínf } Me(\vec{Y} - \vec{X}) \not\leq 0$
× $\text{ínf } Me(\vec{X} - \vec{Y}) > 0$	✓ $\text{ínf } Me(\vec{Y} - \vec{X}) > 0$

Observamos que tenemos que $\text{ínf } Me(\vec{Y}-\vec{X}) > 0$ *pero* $\vec{X} \succ^+_{SP} \vec{Y}$ *lo que permite comprobar que esa implicación no se da. Del mismo modo tenemos* $\vec{X} \succ^+_{SP} \vec{Y}$ *pero* $\text{ínf } Me(\vec{X}-\vec{Y}) < 0$*. También podemos ver que no existe relación entre* $\mathcal{Q}^M$ *y* $\mathcal{Q}^g$ *pues* $\vec{X} \succ^+_{SP} \vec{Y}$ *pero* $\vec{Y} \succ^g_{SP} \vec{X}$*. También destacar de nuevo lo poco exigente que es la preferencia débil ya* $\mathcal{Q}^2(\vec{X},\vec{Y}) = 0$ *y sin embargo,* $\vec{X} \succ^w_{SP} \vec{Y}$*. Por último observamos que tenemos preferencia en ambas marginales de* Y*, pero sin embargo* $\vec{X} \succ^+_{SP} Y$*.*

Ejemplo A.4. *Consideremos* $g = +$*,* $M = \text{am}$ *y los vectores aleatorios* $\vec{Y} = 0_{\mathcal{D}^{(1)}}$ *y* $\vec{X}$

dado por:

		X_2			
		-1	*0*	*1*	
X_1	*0*	*0.5*	*0.1*	*0.2*	*0.8*
	1	*0*	*0*	*0.2*	*0.2*
		0.5	*0.1*	*0.4*	

Se tiene que $Y_+ = Y_1 + Y_2 = 0_{\mathcal{D}^{(1)}}$ *mientras que la distribución de* $X_+ = X_1 + X_2$ *viene dada por:*

$X_1 + X_2$	-1	0	1	2
	0.5	0.1	0.2	0.2

Calculamos las distintas $\mathcal{Q}^*$. *En primer lugar:*

$$\left.\begin{array}{r}\mathcal{Q}(X_1,Y_1) = P(X_1=1) + \frac{1}{2}P(X_1=0) = 0.6\\ \mathcal{Q}(X_2,Y_2) = P(X_2=1) + \frac{1}{2}P(X_2=0) = 0.45\end{array}\right\} \mathcal{Q}^{am}(\vec{X},\vec{Y}) = 0.525 \implies \vec{X} \succ_{SP}^{am} \vec{Y}.$$

Las relaciones $\mathcal{Q}^2$ *toman los valores:*

$$\begin{array}{r}\mathcal{Q}^2(\vec{X},\vec{Y}) = P(X_1=1, X_2=1) = 0.2\\ \mathcal{Q}^2(\vec{Y},\vec{X}) = 0\end{array} \implies \left\{\begin{array}{l}\vec{X} \succ_{SP}^{m} \vec{Y},\\ \vec{X} \succ_{SP}^{w} \vec{Y},\\ \vec{Y} \succ_{SP}^{w} \vec{X}.\end{array}\right.$$

Y por último $\mathcal{Q}^+$ *es:*

$$\mathcal{Q}^+(\vec{X},\vec{Y}) = P(X_+=1) + P(X_+=2) + \frac{1}{2}P(X_+=0) = 0.45 \implies \vec{Y} \succ_{SP}^{+} \vec{X}.$$

Estudiemos ahora las medianas:

$$\left.\begin{array}{rcl}Me(X_1-Y_1) &=& 0\\ Me(X_2-Y_2) &=& [-1,0]\end{array}\right\} \text{ínf } Me(\vec{X}-\vec{Y}) = (0,-1),$$

$$\left.\begin{array}{rcl}Me(Y_1-X_1) &=& 0\\ Me(Y_2-X_2) &=& [0,1]\end{array}\right\} \text{ínf } Me(\vec{X}-\vec{Y}) = (0,0).$$

Por tanto, tendríamos:

✓ $\vec{X} \succ_{SP}^{w} \vec{Y}$ × $\vec{X} \succ_{SP}^{g} \vec{Y}$

✓ $\vec{X} \succ_{SP}^{M} \vec{Y}$ ✓ $\vec{X} \succ_{SP}^{m} \vec{Y}$

$\times$ $\vec{X} \succ^s_{SP} \vec{Y}$

$\checkmark$ $\inf Me(\vec{X} - \vec{Y}) \nless 0$

$\times$ $\inf Me(\vec{Y} - \vec{X}) \nleq 0$

$\times$ $\inf Me(\vec{X} - \vec{Y}) > 0$

$\checkmark$ $\vec{Y} \succ^w_{SP} \vec{X}$

$\times$ $\vec{Y} \succ^M_{SP} \vec{X}$

$\checkmark$ $\vec{Y} \succ^g_{SP} \vec{X}$

$\times$ $\vec{Y} \succ^m_{SP} \vec{X}$

$\times$ $\vec{Y} \succ^s_{SP} \vec{X}$

$\checkmark$ $\inf Me(\vec{Y} - \vec{X}) \nless 0$

$\times$ $\inf Me(\vec{Y} - \vec{X}) \nleq 0$

$\times$ $\inf Me(\vec{Y} - \vec{X}) > 0$

Ejemplo A.5. *Consideremos $g = +$, $M =$ am y los vectores aleatorios $\vec{Y} = 0_{\mathcal{D}^{(1)}}$ y $\vec{X}$ dado por:*

		X_2			
		-1	1	2	
	-1	0.3	0.1	0.1	0.5
X_1	0	0.1	0.1	0.1	0.3
	1	0	0.1	0.1	0.2
		0.4	0.3	0.3	

La distribución de $Y_+ = Y_1 + Y_2 = 0_{\mathcal{D}^{(2)}}$ mientras que la de $X_+ = X_1 + X_2$ es:

$X_1 + X_2$	-2	-1	0	1	2	3
	0.3	0.1	0.1	0.2	0.2	0.1

Calculamos las distintas $\mathcal{Q}^*$:

$$\left.\begin{aligned} \mathcal{Q}(X_1, Y_1) = P(X_1 = 1) + \frac{1}{2}P(X_1 = 0) &= 0.35 \\ \mathcal{Q}(X_2, Y_2) = P(X_2 = 1) + P(X_2 = 2) &= 0.6 \end{aligned}\right\} \mathcal{Q}^{am}(\vec{X}, \vec{Y}) = 0.475 \implies \vec{Y} \succ^{am}_{SP} \vec{X}.$$

Los valores de $\mathcal{Q}^2$ son:

$$\begin{aligned} \mathcal{Q}^2(\vec{X}, \vec{Y}) = \quad & P(X_1 = 1, X_2 = 1) + P(X_1 = 1, X_2 = 2) &= 0.2 \\ \mathcal{Q}^2(\vec{Y}, \vec{X}) = \quad & P(X_1 = -1, X_2 = -1) &= 0.3 \end{aligned} \implies \begin{cases} \vec{Y} \succ^m_{SP} \vec{X}, \\ \vec{X} \succ^w_{SP} \vec{Y}, \\ \vec{Y} \succ^w_{SP} \vec{X}. \end{cases}$$

Y por último, los valores de $\mathcal{Q}^+$ son:

$$\mathcal{Q}^+(\vec{X}, \vec{Y}) = P(X_+ = 1) + P(X_+ = 2) + P(X_+ = 3) + \frac{1}{2}P(X_+ = 0) = 0.55 \implies \vec{X} \succ^+_{SP} \vec{Y}.$$

Estudiemos ahora las medianas:

$$\left.\begin{array}{rcl} Me(X_1 - Y_1) & = & [-1, 0] \\ Me(X_2 - Y_2) & = & 1 \end{array}\right\} \text{ínf } Me(\vec{X} - \vec{Y}) = (-1, 1),$$

$$\left.\begin{array}{rcl} Me(Y_1 - X_1) & = & [0, 1] \\ Me(Y_2 - X_2) & = & -1 \end{array}\right\} \text{ínf } Me(\vec{X} - \vec{Y}) = (0, -1).$$

Por tanto, tendríamos:

✓ $\vec{X} \succ^w_{SP} \vec{Y}$	✓ $\vec{Y} \succ^w_{SP} \vec{X}$
× $\vec{X} \succ^M_{SP} \vec{Y}$	✓ $\vec{Y} \succ^M_{SP} \vec{X}$
✓ $\vec{X} \succ^g_{SP} \vec{Y}$	× $\vec{Y} \succ^g_{SP} \vec{X}$
× $\vec{X} \succ^m_{SP} \vec{Y}$	✓ $\vec{Y} \succ^m_{SP} \vec{X}$
× $\vec{X} \succ^s_{SP} \vec{Y}$	× $\vec{Y} \succ^s_{SP} \vec{X}$
✓ ínf $Me(\vec{X} - \vec{Y}) \nless 0$	✓ ínf $Me(\vec{Y} - \vec{X}) \nless 0$
✓ ínf $Me(\vec{Y} - \vec{X}) \nleq 0$	× ínf $Me(\vec{Y} - \vec{X}) \nleq 0$
× ínf $Me(\vec{X} - \vec{Y}) > 0$	× ínf $Me(\vec{Y} - \vec{X}) > 0$

Vamos a considerar un último contraejemplo para comparar las extensiones con la preferencia unidimensional.

Ejemplo A.6. *Consideremos las variables aleatorias $\vec{Y} = 0_{\mathcal{D}(1)}$ y $\vec{X}$:*

		X_2			
		-1	0	1	
	-1	0.3	0	0	0.3
X_1	0	0	0	0.3	0.3
	1	0	0.3	0.1	0.4
		0.3	0.3	0.4	

Calculamos las distintas $\mathcal{Q}^*$:

$$\mathcal{Q}(X_1, Y_1) = P(X_1 = 1) + \frac{1}{2}P(X_1 = 0) \quad = 0.55 \quad \Longrightarrow \quad X_1 \succ_{SP} Y_1.$$

$$\mathcal{Q}(X_2, Y_2) = P(X_2 = 1) + P(X_2 = 1) \quad = 0.55 \quad \Longrightarrow \quad X_2 \succ_{SP} Y_2.$$

Las relaciones Q^2 toman los siguientes valores:

$$\begin{array}{l} \mathcal{Q}^2(\vec{X},\vec{Y}) = P(X_1 = 1, X_2 = 1) = 0.1 \\ \mathcal{Q}^2(\vec{Y},\vec{X}) = P(X_1 = -1, X_2 = -1) = 0.3 \end{array} \implies \begin{cases} \vec{Y} \succ_{SP}^{m} \vec{X}, \\ \vec{X} \succ_{SP}^{w} \vec{Y}, \\ \vec{Y} \succ_{SP}^{w} \vec{X}. \end{cases}$$

De este modo tendríamos preferencia en ambas marginales de $\vec{X}$ sin embargo no hay ni preferencia moderada ni fuerte.

Bibliografía

[1] M. Abramowitz e I.A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York, USA: Dover Publications, 1972.

[2] F. Belzunce, C. Martinez Riquelme y J. Mulero. *An Introduction to Stochastic Orders*. 1st. San Diego, CA, USA: Academic Press, 2016.

[3] P. Billingsley. *Probability and Measure*. Hoboken, NJ, USA: John Wiley y Sons, 1995.

[4] G. Choquet. "Theory of capacities". En: *Annales de l'Institut Fourier* 5 (1954), págs. 131-295.

[5] H. De Meyer, B. De Baets y B. De Shuymer. "On the transitivity of the comonotonic and countermonotonic comparison of random variables". En: *Journal of Multivariate Analysis* 98.1 (2007), págs. 177-193.

[6] B. De Schuymer, H. De Meyer y B. De Baets. "A fuzzy approach to stochastic dominance of random variables". En: Lecture Notes in Artificial Intelligence, 2003, págs. 253-260.

[7] B. De Schuymer, H. De Meyer y B. De Baets. "Cycle-transitivity comparison of independent random variables". En: *Journal of Multivariate Analysis* 96.2 (2005), págs. 352-373.

[8] D. Dentcheva y A. Ruszczýnski. "Optimization with multivariate stochastic dominance constraints". En: *SIAM Journal on Optimization* 25.1 (2015), págs. 564-588.

[9] E.P. Klement, R. Mesiar y E. Pap. "Invariant copulas". En: *Kybernetika* 38.3 (2002), págs. 275-286.

[10] D. Levhari, J. Paroush y B. Peleg. "Efficiency analysis for multivariate distributions". En: *The Review of Economic Studies* 42.1 (1975), págs. 87-91.

[11] H. Levy. *Stochastic Dominance*. Dordrecht, Netherlands, 1998.

[12] H.B. Mann y D.R. Whitney. "On a test of whether one of two random variables is stochastically larger than the other". En: *Annals of Mathematical Statistics* 18 (1947), págs. 50-60.

[13] I. Montes. “Comparison of Alternatives under Uncertainty and Imprecision”. Tesis doct. Oviedo, Spain: Universidad de Oviedo, 2014.

[14] I. Montes, D. Martinetti, S. Díaz y S. Montes. “Interpretation of statistical preference in terms of location parameters”. En: *INFOR: Information Systems and Operational Research* 53.1 (2015), págs. 1-12.

[15] I. Montes, D. Martinetti, S. Díaz y S. Montes. “Statistical preference as a tool in consensus processes”. En: *Consensual Processes*. Springer Berlin Heidelberg, 2011, págs. 65-92.

[16] I. Montes y S. Montes. “Stochastic dominance and statistical preference for random variables coupled by an Archimedean copula or by the Fréchet-Hoeffding upper bound”. En: *Journal of Multivariate Analysis* 143 (2016), págs. 275-298.

[17] A. Müller y D. Stoyan. *Comparison Methods for Stochastic Models and Risks*. Chichester, UK: Wiley, 2002.

[18] R.B. Nelsen. *An Introduction to Copulas*. New York, NY, USA: Springer, 2010.

[19] M. Öztürké, A. Tsoukiàs, P. Vincke, J. Figueira, S. Greco y M. Ehrogott. “Preference modelling”. En: *Multiple Criteria Decision Analysis: State of the Art Surveys*. Springer New York, ene. de 2005, págs. 27-59.

[20] B. Petrová. “Multivariate Stochastic Dominance and its Application in Portfolio Optimization Problems”. Tesis doct. Prague, Czech Republic: Charles University, 2018.

[21] B. Russell. *The Principles of Mathematics*. Cambridge, UK: Cambridge University Press, 1903.

[22] T. Schmidt. “Coping with copulas”. En: *Copulas - From Theory to Application in Finance* (2007).

[23] S. Sriboonchitta, W. K. Wong, S. Dhompongsa y H. T. Nguyen. *Stochastic Dominance and Applications to Finance, Risk and Economics*. Boca Raton, FL, USA: Chapman y Hall/CRC, 2009.