



Universidad de Oviedo

ANÁLISIS COMPARATIVO DE TESTS DE  
DOMINANCIA ESTOCÁSTICA

Manuel Emilio Baliño Vilar

Supervisado por: Juan Baz González y Raúl Pérez Fernández

UNIVERSIDAD DE OVIEDO

Facultad de Ciencias

Grado en Matemáticas

Junio de 2024

# Índice general

<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	1
1.2. Objetivos . . . . .	2
1.3. Estructura del trabajo . . . . .	2
<b>2. Preliminares</b>	<b>4</b>
2.1. Primeras definiciones y resultados . . . . .	4
2.2. Distribuciones . . . . .	7
2.3. Resultados auxiliares . . . . .	10
2.4. Procesos estocásticos . . . . .	11
2.5. Teoría de cópulas . . . . .	12
<b>3. Dominancia estocástica</b>	<b>15</b>
3.1. Definiciones básicas . . . . .	15
3.2. Dominancia estocástica en familias de localización-escala . . . . .	18
<b>4. Tests de dominancia estocástica</b>	<b>28</b>
4.1. Test de Anderson . . . . .	29
4.2. Test de Davidson-Duclos . . . . .	31
4.3. Test de Barrett-Donald . . . . .	32
4.4. Test de Linton-Maasoumi-Whang . . . . .	34
4.5. Test de Kaur-Rao-Singh . . . . .	37
4.6. Test de Kolmogorov-Smirnov . . . . .	38

<b>5. Experimentación</b>	<b>39</b>
5.1. Metodología . . . . .	39
5.2. Independencia entre las variables . . . . .	40
5.2.1. Variables aleatorias con distribución normal . . . . .	40
5.2.2. Variables aleatorias con distribución uniforme . . . . .	42
5.2.3. Variables aleatorias con distribución exponencial . . . . .	44
5.3. Variables aleatorias comonótonas . . . . .	46
5.3.1. Variables aleatorias con distribución normal . . . . .	46
5.3.2. Variables aleatorias con distribución uniforme . . . . .	48
5.3.3. Variables aleatorias con distribución exponencial . . . . .	50
5.4. Variables aleatorias contramonótonas . . . . .	52
5.4.1. Variables aleatorias con distribución normal . . . . .	52
5.4.2. Variables aleatorias con distribución uniforme . . . . .	54
5.4.3. Variables aleatorias con distribución exponencial . . . . .	56
5.5. Síntesis . . . . .	58
<b>6. Conclusiones</b>	<b>60</b>
<b>Bibliografía</b>	<b>62</b>

# Capítulo 1

## Introducción

### 1.1. Motivación

El concepto de dominancia estocástica surge alrededor del año 1970 con el objetivo de establecer reglas de decisión para elegir entre dos activos de riesgo, ya que los criterios existentes hasta el momento, como el modelo de Markowitz, incurrían en paradojas incluso en ejemplos sencillos [10]. Dado que los activos de riesgo pueden ser modelizados como variables aleatorias y estas, a su vez, son caracterizadas por su función de distribución; es natural que dichas reglas de decisión se basen en el comportamiento de las funciones de distribución y que, intuitivamente, se elija la *mayor* entre dichas distribuciones, siendo esta la función de distribución que tiende a tomar valores más altos con una mayor probabilidad. De hecho, la formalización de esta idea es la definición de dominancia estocástica como se verá posteriormente.

Como en la práctica las funciones de distribución de las variables aleatorias que modelizan los activos financieros suelen ser desconocidas, para hacer las comparaciones se usan los llamados **tests de dominancia estocástica**, que a partir de dos muestras nos permiten contrastar la existencia de dominancia estocástica entre las funciones de distribución subyacentes. A lo largo de los años se han desarrollado múltiples tests, analizando en este trabajo algunos de los más populares: Anderson [1], Barrett-Donald [2], Davidson-Duclos [4], Kaur-Rao-Singh [7], Kolmogorov-Smirnov, Linton-Maasoumi-Whang [11]. Dichos tests se pueden clasificar en función de la forma en la que construyen el estadístico del contraste:

- Valores de las funciones de distribución sobre una malla de puntos (Anderson, Davidson-Duclos).
- Valores mínimos y/o máximos de la diferencia entre las funciones de distribución (Barrett-Donald, Kaur-Rao-Singh, Linton-Maasoumi-Whang, Kolmogorov-Smirnov).

Como se ha comentado, la dominancia estocástica surge en contextos relacionados con la economía; sin embargo, hoy en día, la dominancia estocástica se aplica en múltiples contextos [10]: economía, agricultura, medicina... Por tanto, la eficiencia de dichos tests es algo que merece la pena ser estudiado más allá del marco teórico.

## 1.2. Objetivos

Retomando lo dicho en el epígrafe anterior, la aplicación de los tests de dominancia estocástica es algo interesante para analizar. Por ello, en este trabajo se pretende comparar la potencia de distintos tests de dominancia estocástica (como, por ejemplo, **Anderson**, **Davidson-Duclos**, **Barrett-Donald**, **Kaur-Rao-Singh**, **Linton-Maasoumi-Whang** y **Kolmogorov-Smirnov**) en las tres situaciones más sencillas de relación (**independencia**, **comonotonía** y **contramonotonía**) entre dos variables aleatorias dentro de la misma familia de localización.

## 1.3. Estructura del trabajo

En los Capítulos 2 y 3 se exponen las definiciones y resultados más relevantes para el desarrollo de los tests. En particular, el Capítulo 2 se centra más en conceptos generales, mientras que el Capítulo 3 se enfoca únicamente en la dominancia estocástica. Más aún, en este último se dan diversas proposiciones para resaltar la importancia y extensión de este concepto y se desarrolla la teoría para la dominancia estocástica en familias de localización-escala. A lo largo del Capítulo 4 se desarrollan los diferentes tests, cada uno en su respectiva sección. En la primera sección del Capítulo 5 se indica la metodología seguida en las sucesivas, donde se realizan las simulaciones mediante el método de Monte Carlo para cada situación expuesta en los objetivos. Se cierra el capítulo con una síntesis final analizando los resultados

de la experimentación. Finalmente, en el Capítulo 6 se recogen las conclusiones del trabajo, los resultados obtenidos en el capítulo anterior y posibles líneas para una futura investigación.

## Capítulo 2

# Preliminares

En este capítulo se introducen: conceptos generales, distribuciones que se utilizarán en el trabajo, resultados necesarios para el desarrollo de los tests, algunas definiciones sobre procesos estocásticos y, por último, situaciones de dependencia que se tratarán en la experimentación.

### 2.1. Primeras definiciones y resultados

En este epígrafe, se definen conceptos y enuncian resultados ampliamente conocidos con el objetivo de que el trabajo sea autocontenido. De esta manera, las primeras definiciones, naturalmente, son las de espacio de probabilidad, variable aleatoria y función de distribución:

**Definición 2.1.** *Sea  $\Omega$  un conjunto,  $\sigma \in \mathcal{P}(\Omega)$  una  $\sigma$ -álgebra sobre  $\Omega$ . Entonces  $(\Omega, \sigma)$  es un **espacio medible**.*

*Dado  $(\Omega, \sigma)$  un espacio medible, se llama **probabilidad sobre  $(\Omega, \sigma)$**  a una aplicación  $P : (\Omega, \sigma) \rightarrow [0, 1]$  verificando:*

- $P(\Omega) = 1$ .
- *Dado  $\{A_n\}_{n \in \mathbb{N}} \in \sigma$  con  $A_n \cap A_m = \emptyset$  para todo  $n, m \in \mathbb{N}$  con  $n \neq m$ , entonces  $P(\bigcup_{n \in \mathbb{N}} A_n) = \sum_{n \in \mathbb{N}} P(A_n)$ .*

*A la terna  $(\Omega, \sigma, P)$  se le llama **espacio de probabilidad**.*

**Definición 2.2.** Sean  $(\Omega_1, \sigma_1, P)$  un espacio de probabilidad y  $(\Omega_2, \sigma_2)$  un espacio medible. Si  $X : (\Omega_1, \sigma_1, P) \rightarrow (\Omega_2, \sigma_2)$  es una aplicación verificando:

$$X^{-1}(A) \in \sigma_1, \quad \forall A \in \sigma_2.$$

se dice que  $X$  es una **variable aleatoria**.

Al conjunto de valores que puede tomar una variable aleatoria  $X$  se le dice **soporte de  $X$** .

Si  $\Omega_2 = C_1 \times \dots \times C_n$ , con  $\{C_i\}_{i=1}^n$  conjuntos, entonces se tiene que si  $(X_1, \dots, X_n) : (\Omega_1, \sigma_1, P) \rightarrow (\Omega_2, \sigma_2)$  es una aplicación verificando:

$$(X_1, \dots, X_n)^{-1}(A) \in \sigma_1, \quad \forall A \in \sigma_2,$$

se dice que  $(X_1, \dots, X_n)$  es un **vector aleatorio**.

Se usará la notación  $\mathbb{B}$  para identificar la  $\sigma$ -álgebra de Borel en  $\mathbb{R}$ , que es la  $\sigma$ -álgebra generada por los abiertos de  $\mathbb{R}$ , es decir, la menor que los contiene. A continuación se introduce el concepto de función de distribución de una variable aleatoria, que caracteriza la distribución de probabilidad de dicha variable aleatoria.

**Definición 2.3.** Sea  $X$  una variable aleatoria con imagen en  $(\mathbb{R}, \mathbb{B})$ . Entonces, la **función de distribución** de  $X$  es la función  $F_X : \mathbb{R} \rightarrow [0, 1]$ , dada por:

$$F_X(x) = P(X \leq x).$$

### Propiedades

- $F_X$  es monótona no decreciente, es decir, si  $x \leq y$ , entonces  $F_X(x) \leq F_X(y)$ .
- $\lim_{x \rightarrow +\infty} F_X(x) = 1$ .
- $\lim_{x \rightarrow -\infty} F_X(x) = 0$ .
- Es continua por la derecha:  $\lim_{x \rightarrow a^+} F_X(x) = F_X(a^+)$ .

Si  $X$  es una variable aleatoria continua, es decir su función de distribución es absolutamente continua, se puede definir la noción de función de densidad de  $X$ .

**Definición 2.4.** Sea  $X$  una variable aleatoria continua con llegada en  $(\mathbb{R}, \mathbb{B})$ . Entonces, la **función de densidad** de  $X$  es la función  $f_X : \mathbb{R} \rightarrow \mathbb{R}$ , dada por:

$$f_X(x) = \frac{dF_X(x)}{dx}.$$

Por definición, es claro que dada la función de densidad de  $X$  se puede obtener su función de distribución mediante:

$$P(X \leq x) = \int_{-\infty}^x f(t)dt.$$

De manera similar a  $\mathbb{R}$ , se usará la notación  $\mathbb{B}_{\mathbb{R}^2}$  para identificar la  $\sigma$ -álgebra de Borel en  $\mathbb{R}^2$ , que es la  $\sigma$ -álgebra generada por los abiertos de  $\mathbb{R}^2$ . A continuación se introduce el concepto de función de distribución de un vector aleatorio, que caracteriza la distribución de probabilidad de dicho vector aleatorio.

**Definición 2.5.** Sea  $(X, Y)$  un vector aleatorio bidimensional con llegada en  $(\mathbb{R}^2, \mathbb{B}_{\mathbb{R}^2})$ . Entonces, la **función de distribución** de  $(X, Y)$  es la función  $F_{(X,Y)} : \mathbb{R} \rightarrow [0, 1]$ , dada por:

$$F_{(X,Y)}(x, y) = P(X \leq x, Y \leq y).$$

### Propiedades

- Es monótona no decreciente en cada componente.
- $\lim_{x,y \rightarrow +\infty} F_{(X,Y)}(x, y) = 1$ .
- $\lim_{x \rightarrow -\infty} F_{(X,Y)}(x, y) = \lim_{y \rightarrow +\infty} F_{(X,Y)}(x, y) = 0$ .
- Es continua por la derecha en cada componente.
- Es 2-creciente: Para todo  $x_1, x_2, y_1, y_2 \in \mathbb{R}$  con  $x_1 \leq x_2$  e  $y_1 \leq y_2$ :

$$F_{(X,Y)}(x_2, y_2) + F_{(X,Y)}(x_1, y_1) - F_{(X,Y)}(x_1, y_2) - F_{(X,Y)}(x_2, y_1) \geq 0.$$

Dada una muestra cualquiera, se puede definir la llamada función de distribución empírica, que intuitivamente, se puede entender como un estimador funcional de la función de distribución de la variable aleatoria de la que proviene la muestra. A continuación, se presenta la definición formal de esta idea:

**Definición 2.6.** Sea  $X$  una variable aleatoria con función de distribución  $F_X$  y  $(x_1, \dots, x_n)$  una muestra suya. Entonces, la **función de distribución empírica** es la función  $\hat{F}_X : \mathbb{R} \rightarrow [0, 1]$ , dada por:

$$\hat{F}_X(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}(x_i \leq x).$$

A continuación se enuncia el Teorema de Glivenko-Cantelli, que nos indica que la función de distribución empírica no es solamente un estimador natural de la función de distribución teórica sino que, además, converge a ella. Este resultado será de vital importancia para el desarrollo del trabajo pues nos permite estudiar la dominancia estocástica a partir de funciones de distribución empíricas.

**Teorema 2.1** (Glivenko-Cantelli). *Sean  $X$  una variable aleatoria con función de distribución  $F$  y  $(X_1, \dots, X_n)$  una muestra aleatoria simple suya. Se considera la función:*

$$\Delta_n : (X_1, \dots, X_n) \longrightarrow \mathbb{R}$$

$$(x_1, \dots, x_n) \longrightarrow \sup_{x \in \mathbb{R}} \left| \hat{F}(x) - F(x) \right|,$$

entonces se verifica que

$$\Delta_n \xrightarrow{c.s.} 0.$$

## 2.2. Distribuciones

A continuación se exponen las distribuciones que son utilizadas en el desarrollo de los tests o en la experimentación; así como algunas de sus propiedades.

En primer lugar, se introduce la distribución multinomial  $k$ -dimensional de parámetros  $n$  y  $\vec{p}$ .

**Definición 2.7.** *Se considera el espacio de probabilidad  $(\Omega, \mathcal{A}, P)$ , y sean  $k$  sucesos  $\mathcal{A}_1, \dots, \mathcal{A}_k \in \mathcal{A}$  que forman una partición de  $\Omega$ , es decir:*

$$\bigcup_{i=1}^k \mathcal{A}_i = \Omega \quad \text{y} \quad \mathcal{A}_m \cap \mathcal{A}_n = \emptyset \text{ si } m \neq n.$$

Se define  $p_i = P(\mathcal{A}_i)$  para todo  $i = \{1, \dots, k\}$ , teniéndose que, por ser una partición,  $\sum_{i=1}^k p_i = 1$ . Se dice que  $\vec{X}$  sigue una **distribución multinomial  $k$ -dimensional de parámetros  $n$  y  $\vec{p}' := (p_1, \dots, p_k)$**  ( $\vec{X} \equiv \mathcal{M}_k(n, \vec{p})$ ) si  $\vec{X}$  es la suma de  $n$  variables  $Y : \Omega \rightarrow \mathbb{R}^k$  independientes de la forma:

$$Y(w) = \begin{cases} e_1, & \text{si } w \in \mathcal{A}_1, \\ \vdots & \\ e_k, & \text{si } w \in \mathcal{A}_k, \end{cases}$$

siendo  $\mathcal{B} = \{e_1, \dots, e_k\}$  la base canónica de  $\mathbb{R}^k$ .

Su función de probabilidad es:

$$P(\vec{X} = (x_1, \dots, x_k)) = \frac{n!}{x_1! \cdot \dots \cdot x_k!} \cdot p_1^{x_1} \cdot \dots \cdot p_k^{x_k}.$$

**Propiedades:**

- $E(\vec{X}) = \vec{\mu} = n\vec{p}$ .

- $Var(\vec{X}) = \Sigma = n \begin{pmatrix} p_1(1-p_1) & -p_1p_2 & \dots & -p_1p_k \\ -p_2p_1 & p_2(1-p_2) & \dots & -p_2p_k \\ \vdots & \vdots & \ddots & \vdots \\ -p_kp_1 & -p_kp_2 & \dots & p_k(1-p_k) \end{pmatrix}$ .

A continuación se presenta la distribución normal de parámetros  $\mu$  y  $\sigma$ , así como su versión multidimensional: la distribución normal  $k$ -dimensional.

**Definición 2.8.** Sea  $X$  una variable aleatoria. Se dice que  $X$  sigue una **distribución normal** de parámetros  $\mu \in \mathbb{R}$  y  $\sigma \in \mathbb{R}^+$  ( $X \equiv \mathcal{N}(\mu, \sigma)$ ) si es una variable continua con función de densidad:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}.$$

**Propiedades:**

- $E(X) = \mu$ .

- $Var(X) = \sigma^2$ .

**Definición 2.9.** Un vector aleatorio  $k$ -dimensional  $\vec{X}$  sigue una **distribución normal  $k$ -dimensional** ( $\vec{X} \equiv \mathcal{N}_k(\vec{\mu}, \Sigma)$ ), con  $\vec{\mu} \in \mathbb{R}^k$  y  $\Sigma$  una matriz semidefinida positiva de dimensión  $k \times k$ , si para cada  $\vec{a} \in \mathbb{R}^k$ ,  $\vec{a}'\vec{X}$  sigue una distribución normal (suponiendo que una distribución degenerada en  $c$  es una normal de media  $c$  y varianza 0).

Su función de densidad es:

$$f_{\vec{X}}(\vec{x}) = \frac{1}{|2\pi\Sigma|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}(\vec{x}-\vec{\mu})'\Sigma^{-1}(\vec{x}-\vec{\mu})}.$$

### Propiedades:

- Si  $\vec{X} \equiv \mathcal{N}_k(\vec{\mu}, \Sigma)$  y  $A \in \mathcal{M}_{q,k}(\mathbb{R})$ , entonces  $A\vec{X} \equiv \mathcal{N}_q(A\vec{\mu}, A\Sigma A')$ .
- Si  $X \equiv \mathcal{N}_1(\mu, \Sigma)$ , entonces  $X \equiv \mathcal{N}(\mu, \sqrt{\Sigma})$ .

Por último, se exponen las distribuciones: uniforme en el intervalo  $[a, b]$  y exponencial de parámetro  $\lambda$ , que se utilizarán en la experimentación.

**Definición 2.10.** Sea  $X$  una variable aleatoria. Se dice que  $X$  sigue una **distribución uniforme en el intervalo**  $[a, b]$  ( $X \equiv \mathcal{U}(a, b)$ ) si es una variable continua con función de densidad:

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{si } x \in [a, b]. \\ 0, & \text{si } x \notin [a, b]. \end{cases}$$

### Propiedades

- $E(X) = \frac{a+b}{2}$ .
- $Var(X) = \frac{(b-a)^2}{12}$ .
- Si  $X \equiv \mathcal{U}(a, b)$ ,  $c > 0$  y  $d \in \mathbb{R}$  constantes, entonces se tiene que:  $Y = cX + d \equiv \mathcal{U}(ca + d, cb + d)$ .

**Definición 2.11.** Sea  $X$  una variable aleatoria. Se dice que  $X$  sigue una **distribución exponencial de parámetro**  $\lambda$  ( $X \equiv Exp(\lambda)$ ) si es una variable continua con función de densidad:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{si } x > 0. \\ 0, & \text{en otro caso.} \end{cases}$$

### Propiedades

- $E(X) = \frac{1}{\lambda}$ .
- $Var(X) = \frac{1}{\lambda^2}$ .
- Si  $X \equiv Exp(\lambda)$  y  $a > 0$  constante, entonces:  $Y = aX \equiv Exp(\frac{\lambda}{a})$ .

## 2.3. Resultados auxiliares

A continuación se introduce el módulo máximo estudentizado con  $k$  y  $m$  grados de libertad, que nos servirá para realizar los contrastes múltiples [6] necesarios en ciertos tests.

**Definición 2.12.** [18] Sean  $(X_1, \dots, X_k)$  una muestra aleatoria simple proveniente de una distribución  $\mathcal{N}(0, \sigma)$  y  $S^2$  un estadístico para estimar  $\sigma^2$  que sea independiente de la muestra y tal que  $mS^2 \equiv \sigma^2 \chi_m^2$ . Entonces, el **módulo máximo estudentizado con  $k$  y  $m$  grados de libertad** se define como:

$$M(k, m) = \max_{1 \leq i \leq k} \frac{X_i}{S}.$$

### Propiedades

- Si  $\sigma^2$  es conocido, se considera que  $m = \infty$ .
- En el trabajo se considera  $M_{\infty, \alpha}^k := F_{M(k, \infty)}^{-1}(1 - \alpha)$
- En [6] se expone el método GT2, que utiliza el módulo máximo estudentizado, como extensión del contraste múltiple de Tukey para casos donde no hay igualdad de varianzas o covarianzas.

**Observación 1.** Notar que las variables que siguen una distribución  $\mathcal{N}(0, \sigma)$  deben ser independientes. Si no se cumple esta hipótesis, el test GT2 es conservador [6, 16].

El test de Anderson, presentado más adelante, requiere de dos nociones particulares, por ello, se enuncian a continuación.

El Teorema del Límite Central en dimensión 1 es un resultado importantísimo en la Estadística y ampliamente conocido, que además puede ser extendido a cualquier dimensión. La extensión al caso multivariante es usada en el desarrollo de dicho test, por ello, se enuncia a continuación.

**Teorema 2.2.** [12, Cap. 2][Teorema del Límite Central multivariante] Sea  $\vec{X}$  un vector aleatorio  $p$ -dimensional cuya distribución tiene vector de medias  $\vec{\mu}$  y matriz de varianzas-covarianzas  $\Sigma$ . Sea  $(\vec{X}_1, \vec{X}_2, \dots, \vec{X}_n)$  una muestra aleatoria simple de  $\vec{X}$ , entonces:

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n (\vec{X}_i - \vec{\mu}) \xrightarrow{\mathcal{L}} \mathcal{N}_p(\vec{0}_p, \Sigma).$$

La última definición de esta subsección corresponde a la de vector de frecuencias empíricas, un concepto clave para desarrollar más adelante dicho test.

**Definición 2.13.** Sea  $X$  una variable aleatoria y  $(x_1, \dots, x_n)$  una muestra suya. Se considera una partición disjunta de los datos en  $I_1, \dots, I_k$  intervalos. Entonces, el **vector de frecuencias empíricas** es  $\vec{z} = (z_1, \dots, z_k)'$ , siendo  $z_i = \text{card}(\{x_j \in I_i : j = 1, \dots, n\})$ . Notar que si  $p_j := P(x \in I_j)$ , entonces se tiene que:

$$\vec{z} \equiv \mathcal{M} \left( n, \begin{pmatrix} p_1 \\ \vdots \\ p_k \end{pmatrix} \right).$$

**Observación 2.** Por el Teorema (2.2), se tiene:

$$\frac{1}{\sqrt{n}}(\vec{z} - n\vec{p}) \xrightarrow{\mathcal{L}} \mathcal{N}_k(\vec{0}_k, \Sigma). \quad (2.1)$$

## 2.4. Procesos estocásticos

Debido a que los estadísticos de algunos de los tests presentados más adelante no son de distribución conocida, se recurrirá a distribuciones asintóticas utilizando composiciones con procesos estocásticos. Para ello, se introducen las siguientes definiciones a continuación.

**Definición 2.14.** Un **proceso estocástico** es cualquier familia de variables aleatorias,  $\{X_t\}_{t \in I}$ , definidas en el mismo espacio de probabilidad y con llegada en un espacio medible. Es decir, para todo  $t \in I$ :

$$\begin{aligned} X_t : (\Omega, \sigma, P) &\longrightarrow (\Omega_t, \sigma_t) \\ \omega &\longrightarrow X_t(\omega). \end{aligned}$$

Para cada  $\omega \in \Omega$ ,  $\{X_t(\omega)\}_{t \in I}$  se llama **trayectoria del proceso**.

Si  $(\Omega_t, \sigma_t) = (\mathbb{R}, \mathcal{B}) \forall t \in I$ , se dice que el proceso estocástico es **real**.

**Definición 2.15.** Un proceso estocástico real  $\{W_t\}_{t \in I}$ , con  $I = [0, T]$  se dice **proceso de Wiener** si cumple:

1.  $W_0 = 0$ .

2. Si  $t, s \geq 0$  y  $t + s \in I$ , entonces  $W_{t+s} - W_t \equiv \mathcal{N}(0, s)$ .
3. Si  $0 \leq s_1 < t_1 \leq s_2 < t_2$ , todos ellos en  $I$ , entonces  $W_{t_1} - W_{s_1}$  y  $W_{t_2} - W_{s_2}$  son variables aleatorias independientes.

**Definición 2.16.** Un **proceso gaussiano** es un proceso estocástico tal que cada colección finita de variables aleatorias sigue una distribución normal multidimensional.

**Definición 2.17.** Un **punte browniano** es un proceso estocástico formado con variables aleatorias cuyas distribuciones de probabilidad son las de un proceso de Wiener condicionadas al estado final 0, es decir,  $\{Y_t\}_{t \in I}$ , con  $I = [0, T]$ , tal que:

$$Y_t = W_t \mid W_T = 0, \quad t \in I.$$

siendo  $\{W_t\}_{t \in I}$  un proceso de Wiener.

**Teorema 2.3.** [5][Donsker] Sea  $X$  una variable aleatoria con función de distribución  $F$ . Dada  $(X_1, \dots, X_n)$  una muestra aleatoria simple de  $X$ , se define  $G_n(w) = \sqrt{n} \left( \hat{F}_Y(w) - \hat{F}_X(w) \right)$ . Entonces,  $G_n \xrightarrow{\mathcal{L}} G$ , siendo  $G$  un proceso gaussiano que puede ser expresado como  $G(w) = \mathcal{B} \circ F(w)$ , con  $\mathcal{B}$  un puente browniano.

## 2.5. Teoría de cópulas

En esta sección se desarrollan los conceptos relativos a la dependencia entre variables aleatorias, ya que en la experimentación se trabajará con los casos aquí presentados.

**Definición 2.18.** Una función  $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$  es una **cópula** si cumple las siguientes propiedades:

1. Condiciones de frontera:

$$C(x, 0) = C(0, x) = 0, \quad \forall x \in [0, 1],$$

$$C(x, 1) = C(1, x) = x, \quad \forall x \in [0, 1].$$

2.  $C$  es 2-creciente.

El resultado presentado a continuación implica que toda función de distribución de dos dimensiones se puede descomponer en las funciones de distribución de las marginales y una cópula.

**Teorema 2.4.** [17][Sklar] Sea  $(X, Y)$  un vector aleatorio bidimensional con función de distribución  $F_{(X,Y)}$  y marginales  $F_X, F_Y$ . Entonces existe una cópula  $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$  que satisfice:

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

A continuación se presentan tres de las cópulas más populares.

**Definición 2.19.** Se llama **cópula del producto** a la cópula  $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$  definida por:

$$C_\pi(x, y) = x \cdot y, \quad \forall (x, y) \in [0, 1]^2.$$

Sean  $X$  e  $Y$  dos variables aleatorias. Se dice que  $X$  e  $Y$  son **independientes** si:

$$F_{(X,Y)}(x, y) = C_\pi(F_X(x), F_Y(y)), \quad \forall (x, y) \in \mathbb{R}^2.$$

De manera intuitiva, dos variables aleatorias son independientes si no existe una relación entre ellas, es decir, que el conocimiento de los valores de una variable no proporciona ninguna información sobre los valores de la otra.

**Definición 2.20.** Se llama **cópula del mínimo** a la cópula  $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$  definida por:

$$C_M(x, y) = \min\{x, y\}, \quad \forall (x, y) \in [0, 1]^2.$$

Sean  $X$  e  $Y$  dos variables aleatorias. Se dice que  $X$  e  $Y$  son **comonótonas** si:

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

De manera intuitiva, dos variables aleatorias son comonótonas si existe una dependencia positiva perfecta entre las variables, es decir, ambas variables crecen y decrecen de manera conjunta.

**Definición 2.21.** Se llama **Cópula de Lukasiwicz** a la cópula  $C : [0, 1] \times [0, 1] \rightarrow [0, 1]$  definida por:

$$C_L(x, y) = \max\{x + y - 1, 0\}, \quad \forall (x, y) \in [0, 1]^2.$$

Sean  $X$  e  $Y$  dos variables aleatorias. Se dice que  $X$  e  $Y$  son **contramonótonas** si:

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

De manera intuitiva, dos variables aleatorias son contramonótonas si existe una dependencia negativa perfecta entre las variables, es decir, si una variable crece, la otra decrece y viceversa.

**Teorema 2.5** (Desigualdad de Frechét-Hoeffding). *Toda cópula  $C$  verifica:*

$$C_L(x, y) \leq C(x, y) \leq C_M(x, y), \quad \forall (x, y) \in [0, 1]^2.$$

Como consecuencia del Teorema 2.4 se tiene que cualquier relación de dependencia se puede expresar como una cópula, y más aún, por el Teorema 2.5 se tiene que dicha cópula toma valores entre la de Lukasiwicz y la del mínimo.

## Capítulo 3

# Dominancia estocástica

Retomando lo visto en la introducción, el concepto de dominancia estocástica surge en economía para elegir entre dos activos de riesgo. Como es natural, se elegirá el activo que mayores beneficios reporte. Sin embargo, la certeza de cuál de los dos activos será más favorable no se suele tener, por este motivo entra en juego la probabilidad, cambiando así la elección al activo que tienda a tomar valores más altos, es decir, que sea más probable que tome valores altos. Mediante la identificación de los activos como variables aleatorias, la definición de dominancia estocástica no es nada más que la formalización de dicha idea.

### 3.1. Definiciones básicas

En primer lugar, presentamos la definición de dominancia estocástica entre dos variables aleatorias.

**Definición 3.1.** [15, Cap. 1] Sean  $X$  e  $Y$  dos variables aleatorias. Se dice que  $X$  es más pequeño que  $Y$  en el orden estocástico usual o que  $Y$  **domina estocásticamente a  $X$** , que se denota  $X \leq_{st} Y$ , si:

$$P(X > x) \leq P(Y > x), \quad \forall x \in \mathbb{R}.$$

o, equivalentemente, si:

$$P(X \leq x) \geq P(Y \leq x), \quad \forall x \in \mathbb{R}. \quad (3.1)$$

Intuitivamente,  $Y$  domina estocásticamente a  $X$  si  $Y$  tiende a tomar valores más grandes que  $X$ . Gráficamente, gracias a la Ecuación (3.1), se puede determinar que  $Y$  domina estocásticamente a  $X$  si la función de distribución de  $X$  es mayor o igual que la de  $Y$  en todo punto. La Figura 3.1 es un ejemplo claro de esta idea para  $X \equiv U(0, 1)$  e  $Y \equiv U(2, 3)$ .

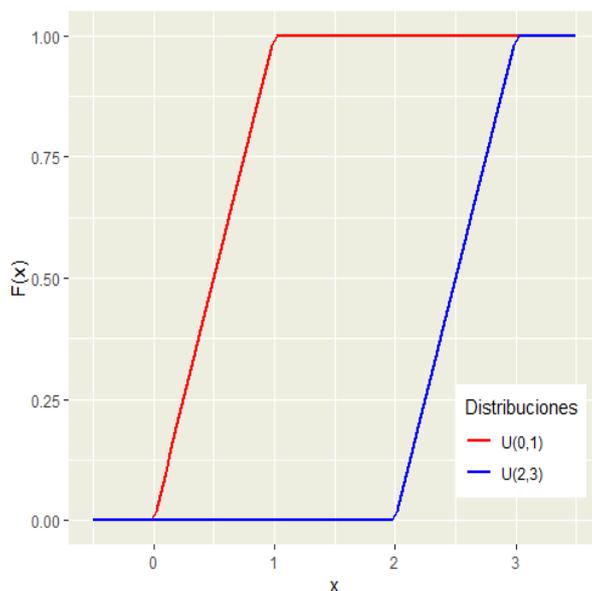


Figura 3.1: Ejemplo de dominancia estocástica.

Notar que, debido a esto y al Teorema 2.1, se puede analizar la dominancia estocástica entre dos variables aleatorias cuando no se conozcan las funciones de distribución, estudiando en su lugar las funciones de distribución empíricas. En este hecho se basan los tests que se desarrollan en el siguiente capítulo.

Resaltar que también hay distintas caracterizaciones de la dominancia estocástica como las que se presentan a continuación.

**Proposición 3.1.** [15, Cap. 1] Sean  $X$  e  $Y$  dos variables aleatorias. Entonces, se cumple que:  $X \leq_{st} Y \iff E(\phi(X)) \leq E(\phi(Y))$  para cualquier función creciente  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  para la cual existen ambas esperanzas.

La siguiente caracterización es también muy importante, ya que si una variable aleatoria  $X$  domina estocásticamente a otra variable aleatoria  $Y$ , entonces existen

copias  $\hat{X}$  y  $\hat{Y}$  de  $X$  e  $Y$ , con una relación de dependencia entre ellas, en general, distinta a la de  $X$  e  $Y$ , pero para las que  $P(\hat{X} \geq \hat{Y}) = 1$ .

**Teorema 3.2.** [15, Cap. 1] Sean  $X$  e  $Y$  dos variables aleatorias. Entonces, se tiene que:  $X \leq_{st} Y$  si y sólo si existen dos variables aleatorias,  $\hat{X}$  e  $\hat{Y}$ , definidas en el mismo espacio de probabilidad tales que:

$$\begin{aligned}\hat{X} &\stackrel{st}{=} X, \\ \hat{Y} &\stackrel{st}{=} Y, \\ P(\hat{X} \leq \hat{Y}) &= 1.\end{aligned}$$

Usando la Ecuación (3.1) y la Proposición 3.1 se obtiene el siguiente resultado.

**Proposición 3.3.** Sean  $X$  e  $Y$  dos variables aleatorias. Entonces:

$$X \leq_{st} Y \iff F_X^{-1}(u) \leq F_Y^{-1}(u), \quad \forall u \in (0, 1).$$

También es importante mencionar que hay más órdenes estocásticos y cada uno de ellos tiene propiedades interesantes para ser estudiadas. Un ejemplo es el de la dominancia estocástica estricta.

**Definición 3.2.** Sean  $X$  e  $Y$  dos variables aleatorias. Se dice que  $X$  es más pequeño que  $Y$  en el orden estocástico estricto o que  $Y$  **domina estocásticamente a  $X$  de manera estricta** si se cumplen, a la vez, las siguientes condiciones:

- $P(X > x) \leq P(Y > x), \forall x \in \mathbb{R}$ .
- $\exists x \in \mathbb{R} : P(X > x) < P(Y > x)$ .

**Observación 3.** Algunos tests se desarrollan tomando esta definición como la de dominancia estocástica usual. No obstante, este orden claramente es más restrictivo y por ello dichos tests se pueden adaptar de manera sencilla a la dominancia estocástica usual. Por ello, a lo largo del trabajo se tratan solamente los tests para el orden estocástico usual por ser el más importante entre todos ellos.

## 3.2. Dominancia estocástica en familias de localización-escala

En esta sección se introduce la noción de familia de localización-escala, así como se presentan distintos resultados concernientes a la dominancia estocástica dentro de una familia de localización-escala.

**Definición 3.3.** *Una familia de localización-escala es un conjunto de distribuciones tales que, dada  $X$  en ella,  $Y$  pertenece a dicha familia de distribuciones si y sólo si  $Y$  tiene la misma distribución que  $aX + b$  para algún  $a, b \in \mathbb{R}$ ,  $a > 0$ .*

*De manera similar, una familia de localización es un conjunto de distribuciones tales que, dada  $X$  en ella,  $Y$  pertenece a dicha familia de distribuciones si y sólo si  $Y$  es igual a  $X + b$  en distribución para algún  $b \in \mathbb{R}$ .*

A partir de la definición es inmediato ver que cualquier variable aleatoria,  $X$ , genera una familia de localización-escala de la forma:

$$\{aX + b \mid a, b \in \mathbb{R}, a > 0\}.$$

Por ello, es común hablar de que una cierta distribución estándar es utilizada para generar la familia de localización-escala a partir de ella. Por ejemplo:

- La familia de localización-escala normal es generada a partir de la distribución normal estándar,  $N(0, 1)$ . Por las propiedades de la distribución normal se tiene que los valores  $a$  y  $b$  utilizados en la generación de la familia se corresponden con la media ( $b$ ) y con la desviación típica ( $a$ ).
- La familia de localización-escala uniforme es generada a partir de la distribución uniforme estándar,  $\mathcal{U}(0, 1)$ . Por las propiedades de la distribución uniforme se tiene que los valores  $a$  y  $b$  utilizados en la generación de la familia se corresponden con el extremo izquierdo del soporte ( $b$ ) y con la longitud del intervalo ( $a$ ).

Para facilitar la notación en la caracterización de la dominancia estocástica en una familia de localización-escala, se realizan las siguientes observaciones.

**Observación 4.** Por definición de función de distribución, dado  $y \in \mathbb{R}$ :

$$F_{aX+b}(y) = P(aX + b \leq y) = P\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right). \quad (3.2)$$

Luego, el siguiente resultado es inmediato:

**Lema 3.4.** Sean  $X$  una variable aleatoria y  $a, b \in \mathbb{R}$  con  $a > 0$ , entonces:

$$F_{aX+b}^{-1}(p) = aF_X^{-1}(p) + b, \quad \forall p \in [0, 1]. \quad (3.3)$$

*Demostración.* Dado  $p \in [0, 1]$ , por definición de función cuantil:

$$\begin{aligned} F_{aX+b}^{-1}(p) &= \inf\{x \in \mathbb{R} \mid F_{aX+b}(x) \geq p\} \stackrel{3.2}{=} \inf\left\{x \in \mathbb{R} \mid F_X\left(\frac{x-b}{a}\right) \geq p\right\} \\ &= \inf\{ax + b \in \mathbb{R} \mid F_X(x) \geq p\} \stackrel{a>0}{=} a \cdot \inf\{x \in \mathbb{R} \mid F_X(x) \geq p\} + b \\ &= aF_X^{-1}(p) + b. \end{aligned}$$

□

La dominancia estocástica no tiene en cuenta la dependencia entre las variables, pues por Ecuación (3.1) es equivalente estudiar  $X \leq_{st} aX + b$  a estudiar  $X \leq_{st} Y$ , siendo  $Y$  una variable aleatoria con la misma distribución que  $aX + b$ . Es por ello que se usarán distribuciones dentro de la misma familia de localización-escala en la experimentación.

Para obtener un marco teórico claro y consistente, se enuncia y se demuestra el siguiente resultado.

**Proposición 3.5.** Sean  $X$  una variable aleatoria y  $a, b \in \mathbb{R}$  con  $a > 0$ .

1. Si  $S_X$  no está acotado, entonces:

$$X \leq_{st} aX + b \iff a = 1 \wedge b \geq 0.$$

2. Si  $S_X$  está acotado sólo inferiormente y  $c := \inf S_X$ , entonces:

$$X \leq_{st} aX + b \iff a \geq 1 \wedge b \geq c(1 - a).$$

3. Si  $S_X$  está acotado sólo superiormente y  $c := \sup S_X$ , entonces:

$$X \leq_{st} aX + b \iff a \leq 1 \wedge b \geq c(1 - a).$$

4. Si  $S_X$  está acotado y  $c := \inf S_X$  y  $d := \sup S_X$ , entonces:

$$X \leq_{st} aX + b \iff (a < 1 \wedge b \geq d(1 - a)) \vee (a \geq 1 \wedge b \geq c(1 - a)).$$

*Demostración.* En virtud de la Ecuación (3.1) y la definición de función de distribución, se tiene que:

$$X \leq_{st} aX + b \iff F_{aX+b}(y) \leq F_X(y), \quad \forall y \in \mathbb{R}.$$

Luego, por la Ecuación (3.2):

$$X \leq_{st} aX + b \iff F_X\left(\frac{y-b}{a}\right) \leq F_X(y), \quad \forall y \in \mathbb{R}. \quad (3.4)$$

Caso 1:  $S_X$  no está acotado.

Por tanto, como la función de distribución es monótona no decreciente y es distinta de 0 para todo  $y \in \mathbb{R}$ , usando la Ecuación (3.4), se deduce que:

$$\begin{aligned} X \leq_{st} aX + b &\iff \frac{y-b}{a} \leq y, \quad \forall y \in \mathbb{R} \\ &\iff y(1-a) \leq b, \quad \forall y \in \mathbb{R} \\ &\iff a = 1 \text{ y } b \geq 0. \end{aligned}$$

Pues:

1. Si  $a \neq 1$ , entonces es claro que:

$$\exists y_0 \in \mathbb{R} \text{ tal que: } \frac{b}{1-a} < y_0, \quad \forall b \in \mathbb{R}.$$

2. Si  $a = 1$  necesariamente  $b \geq 0$ , ya que si  $b < 0$  se da una contradicción.

Caso 2: Si  $S_X$  está acotado sólo inferiormente y  $c := \inf S_X$ .

Por tanto, como la función de distribución es no decreciente y  $F_X(y) = 0$  para todo  $y \notin [c, +\infty)$ , usando la Ecuación (3.4), se deduce que:

$$\begin{aligned} X \leq_{st} aX + b &\iff \frac{y-b}{a} \leq y, \quad \forall y \in [c, +\infty) \\ &\iff y(1-a) \leq b, \quad \forall y \in [c, +\infty) \\ &\iff a = 1 \text{ y } b \geq 0 \text{ ó } a > 1 \text{ y } b \geq c(1-a). \end{aligned}$$

Pues:

1. Si  $a < 1$ , definiendo  $f(x) := x(1-a) - b$  entonces, debido a que:  $\lim_{x \rightarrow +\infty} f(x) = +\infty$ , se tiene que:

$$\exists y_0 \in [c, +\infty) \text{ tal que: } y_0(1-a) - b > 0, \quad \forall b \in \mathbb{R}.$$

2. Si  $a = 1$  necesariamente  $b \geq 0$ , ya que si  $b < 0$  se da una contradicción.
3. Si  $a > 1$  definiendo  $f(x) := x(1-a) - b$  entonces, como  $f$  es estrictamente decreciente:

$$y(1-a) \leq b, \quad \forall y \in [c, +\infty) \iff c(1-a) \leq b.$$

Caso 3: Si  $S_X$  está acotado sólo superiormente y  $c := \sup S_X$ .

Por tanto, como la función de distribución es no decreciente y  $F_X(y) = 0$  para todo  $y \notin (-\infty, c]$ , usando la Ecuación (3.4), se deduce que:

$$\begin{aligned} X \leq_{st} aX + b &\iff \frac{y-b}{a} \leq y, \quad \forall y \in (-\infty, c] \\ &\iff y(1-a) \leq b, \quad \forall y \in (-\infty, c] \\ &\iff a = 1 \text{ y } b \geq 0 \text{ ó } a < 1 \text{ y } b \geq c(1-a). \end{aligned}$$

Pues:

1. Si  $a < 1$ , definiendo  $f(x) := x(1-a) - b$  entonces, como es estrictamente creciente:

$$y(1-a) - b \leq 0, \quad \forall y \in (-\infty, c] \iff c(1-a) \leq b.$$

2. Si  $a = 1$  necesariamente  $b \geq 0$ , ya que si  $b < 0$  se da una contradicción.
3. Si  $a > 1$ , definiendo  $f(x) := x(1-a) - b$  entonces, debido a que:  $\lim_{x \rightarrow -\infty} f(x) = +\infty$ , se tiene:

$$\exists y_0 \in \infty, c] \text{ tal que: } y_0(1-a) - b > 0, \quad \forall b \in \mathbb{R}.$$

Caso 4:  $S_X$  está acotado y  $c := \inf S_X$  y  $d := \sup S_X$ .

Por tanto, como la función de distribución es no decreciente y  $F_X(y) = 0$  para todo  $y \notin [c, d]$ , usando la Ecuación (3.4), se deduce que:

$$\begin{aligned} X \leq_{st} aX + b &\iff \frac{y-b}{a} \leq y, \quad \forall y \in [c, d] \\ &\iff y(1-a) \leq b, \quad \forall y \in [c, d] \\ &a = 1 \text{ y } b \geq 0 \text{ ó } a < 1 \text{ y } b \geq d(1-a) \text{ ó } a > 1 \text{ y } b \geq c(1-a). \end{aligned}$$

Pues:

1. Si  $a < 1$ , definiendo  $f(x) := x(1-a) - b$  entonces, como estrictamente creciente:

$$y(1-a) - b \leq 0, \quad \forall y \in [c, d] \iff d(1-a) \leq b.$$

2. Si  $a = 1$  necesariamente  $b \geq 0$ , ya que si  $b < 0$  se da una contradicción.

3. Si  $a > 1$  definiendo  $f(x) := x(1-a) - b$  entonces, como  $f$  es estrictamente decreciente:

$$y(1-a) \leq b, \quad \forall y \in [c, d] \iff c(1-a) \leq b.$$

□

Claramente, la anterior proposición satisface nuestra motivación, pues gracias a él se tienen infinitos ejemplos y la certeza de qué se debería obtener en cada uno. Sin embargo, en las condiciones de la proposición, no se sabe cuando se tiene que  $aX + b \leq_{st} X$ , ni qué implicaciones tiene sobre los parámetros; no obstante, mediante el siguiente resultado, la caracterización de  $aX + b \leq_{st} X$  es inmediata a partir de dicha proposición.

**Corolario 3.6.** Sean  $X$  una variable aleatoria y  $a, b \in \mathbb{R}$  con  $a > 0$ .

1. Si  $S_X$  no está acotado, entonces:

$$aX + b \leq_{st} X \iff a = 1 \wedge b \leq 0.$$

2. Si  $S_X$  está acotado sólo inferiormente y  $c := \inf S_X$ , entonces:

$$aX + b \leq_{st} X \iff a \leq 1 \wedge b \leq ca(1-a).$$

3. Si  $S_X$  está acotado sólo superiormente y  $c := \sup S_X$ , entonces:

$$aX + b \leq_{st} X \iff a \geq 1 \wedge b \leq ca(1 - a).$$

4. Si  $S_X$  está acotado y  $c := \inf S_X$  y  $d := \sup S_X$ , entonces:

$$aX + b \leq_{st} X \iff (a \leq 1 \wedge b \leq ca(1 - a)) \vee (a > 1 \wedge b \leq da(1 - a)).$$

*Demostración.* Basta notar que  $X = \frac{1}{a}(aX + b) - \frac{b}{a}$ . Así, considerando  $Y = (aX + b)$ , se busca caracterizar:  $Y \leq \frac{1}{a}Y - \frac{b}{a}$ , para lo cual basta aplicar la caracterización anterior.

Debido a que  $S_Y = aS_X + b$ , se puede seguir caracterizando respecto a cómo es  $S_X$ . □

Por simplicidad, durante la experimentación se considerará  $a = 1$ , es decir, distribuciones en la misma familia de localización. Para dejar claro cuándo se tiene dominancia estocástica dentro de dicha familia, se enuncia el siguiente resultado.

**Corolario 3.7.** Sean  $X$  una variable aleatoria y  $b \in \mathbb{R}$ . Entonces:

- $X \leq_{st} X + b \iff b \geq 0$ .
- $X + b \leq_{st} X \iff b \leq 0$ .

Los siguientes resultados y sus respectivos ejemplos se proponen para visualizar cómo se comportan las distribuciones que se utilizarán en la experimentación y ver gráficamente los resultados anteriores.

En la familia de localización-escala normal, se tiene que una variable aleatoria domina estocásticamente a otra si la desviación típica es la misma y la media es más grande en la primera de las variables aleatorias.

**Corolario 3.8.** Sean  $X, Y$  dos variables aleatorias tal que  $X \equiv \mathcal{N}(\mu_X, \sigma_X)$  e  $Y \equiv \mathcal{N}(\mu_Y, \sigma_Y)$ , entonces  $X \leq_{st} Y$  si y sólo si  $\mu_X \leq \mu_Y$  y  $\sigma_X = \sigma_Y$ .

*Demostración.* Si  $Y = a \cdot X + b$  en distribución, entonces, por la Proposición 3.5.1), se tiene que  $X \leq_{st} Y$  si y sólo si  $a = 1$  y  $b \geq 0$ , y esto es equivalente a  $\mu_X \leq \mu_Y$  y  $\sigma_X = \sigma_Y$  usando:

$$\mathcal{N}(\mu_Y, \sigma_Y) = a \cdot \mathcal{N}(\mu_X, \sigma_X) + b = \mathcal{N}(a\mu_X + b, |a|\sigma_X).$$

□

**Ejemplo 3.1.** Si  $A \equiv \mathcal{N}(0, 1)$  y se tiene que:  $B \equiv A - 1$ ,  $C \equiv A + 1$  y  $D \equiv 2A$ , en la Figura (3.2) se puede ver fácilmente que:

- $B \leq_{st} A \leq_{st} C$ .
- $D$  no domina, ni es dominada, estocásticamente por  $A$  (ni por  $B$  o  $C$ ).

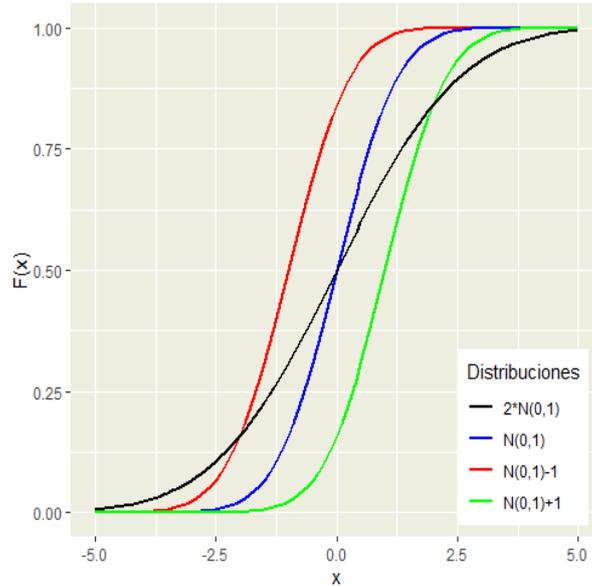


Figura 3.2: Comparación de funciones de distribución dentro de la familia de localización-escala normal.

En la familia de localización-escala uniforme, se tiene que una variable aleatoria domina estocásticamente a otra si tanto el extremo inferior del soporte como el superior son mayores en la primera variable aleatoria que en la segunda.

**Corolario 3.9.** Sean  $X, Y$  dos variables aleatorias tal que  $X \equiv \mathcal{U}(a, b)$  e  $Y \equiv \mathcal{U}(c, d)$ , entonces  $X \leq_{st} Y$  si y sólo si  $a \leq c$  y  $b \leq d$ .

*Demostración.* Si  $Y = \alpha \cdot X + \beta$  en distribución, entonces por Proposición 3.5.4), con  $a = \inf S_X$  y  $b = \sup S_X$ , se tiene que:  $X \leq_{st} Y$  si y sólo si  $\alpha < 1 \wedge \beta \geq b(1 - \alpha)$  ó  $a \geq 1 \wedge \beta \geq a(1 - \alpha)$ , y esto es equivalente a  $a \leq c$  y  $b \leq d$  usando:

$$\mathcal{U}(c, d) = \alpha \cdot \mathcal{U}(a, b) + \beta = \mathcal{U}(\alpha \cdot a + \beta, \alpha \cdot b + \beta).$$

□

**Ejemplo 3.2.** Si  $A \equiv \mathcal{U}(0, 1)$  y se tiene que:  $B \equiv A - 1$ ,  $C \equiv A + 1$ ,  $D \equiv 2A$ ,  $E \equiv 2A + 1$  y  $G \equiv \frac{1}{2}A$  en la Figura (3.3) se puede ver fácilmente que:

- $B \leq_{st} A \leq_{st} C$ .
- $A \leq_{st} D$ .
- $A \leq_{st} E$ .
- $G \leq_{st} A$ .

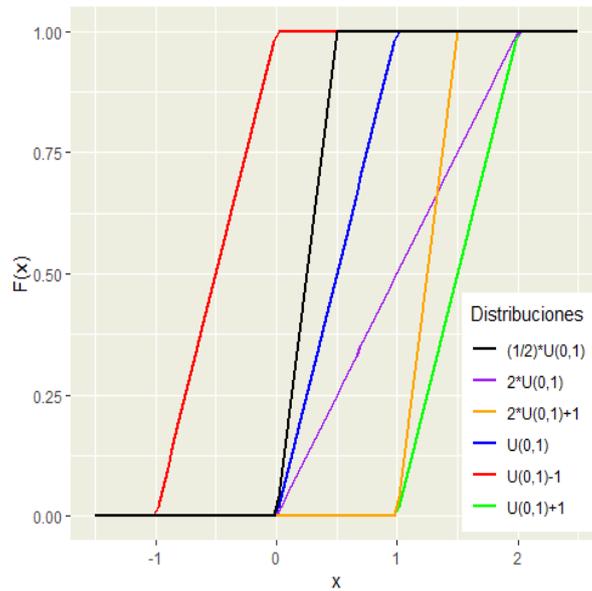


Figura 3.3: Comparación de funciones de distribución dentro de la familia de localización-escala uniforme.

En la familia de localización-escala exponencial, se tiene que una variable aleatoria domina estocásticamente a otra si el parámetro de escala es mayor en la primera de las variables aleatorias y, además, se cumple una condición adicional concerniente a los parámetros de localización.

**Corolario 3.10.** Sean  $X, Y$  dos variables aleatorias tal que  $X \equiv \text{Exp}(\lambda_X) + a$  e  $Y \equiv \text{Exp}(\lambda_Y) + b$ , entonces  $X \leq_{st} Y$  si y sólo si  $\lambda_X \leq \lambda_Y$  y  $b \geq \frac{\lambda_X}{\lambda_Y} \cdot a \left(1 - \frac{\lambda_X}{\lambda_Y}\right)$ .

*Demostración.* Si  $Y = c \cdot X + d$  en distribución, entonces por Proposición 3.5.2),  $a = \inf S_X$ , se tiene que:  $X \leq_{st} Y$  si y sólo si  $c \geq 1$  y  $d \geq a(1 - c)$ , y esto es equivalente a  $\lambda_X \leq \lambda_Y$  y  $b \geq \frac{\lambda_X}{\lambda_Y} \cdot a \left(1 - \frac{\lambda_X}{\lambda_Y}\right)$  usando:

$$Exp(\lambda_Y) + b = c * Exp(\lambda_X) + d = Exp\left(\frac{\lambda_X}{c}\right) + c \cdot d.$$

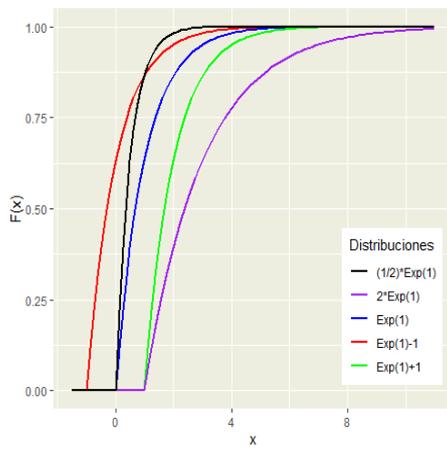
□

**Ejemplo 3.3.** *Se considera en primer lugar una distribución exponencial  $A \equiv Exp(1)$  y se estudia el comportamiento para distintos miembros dentro de la familia de localización-escala generada:  $B \equiv A - 1$ ,  $C \equiv A + 1$ ,  $D \equiv 2A$  y  $E \equiv \frac{1}{2}A$ , en la Figura (3.4a) se puede ver fácilmente que:*

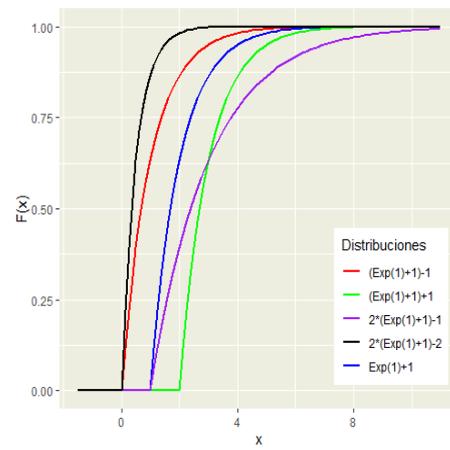
- $B \leq_{st} A \leq_{st} C$ .
- $A \leq_{st} D$ .
- $E \leq_{st} A$ .

*Por otro lado, se considera en segundo lugar una distribución exponencial desplazada  $A \equiv Exp(1) + 1$  y se analiza el comportamiento para distintos miembros dentro de la familia de localización-escala generada:  $B \equiv A - 1$ ,  $C \equiv A + 1$ ,  $D \equiv 2A - 1$  y  $E \equiv \frac{1}{2}A - 2$  que en la Figura (3.4b) se observa que:*

- $B \leq_{st} A \leq_{st} C$ .
- $A \leq_{st} D$ .
- $E \leq_{st} A$ .



(a)  $\inf S_A = c = 0$ .



(b)  $\inf S_A = c \neq 0$ .

Figura 3.4: Comparación de funciones de distribución dentro de la familia de localización-escala exponencial, generada a partir de una distribución exponencial (izquierda) y de una distribución exponencial desplazada (derecha).

## Capítulo 4

# Tests de dominancia estocástica

A lo largo de este capítulo se considerarán  $X$  e  $Y$  dos variables aleatorias con funciones de distribución  $F_X$  y  $F_Y$  respectivamente. Es lógico suponer que dichas funciones de distribución son desconocidas, pues si fuesen conocidas la dominancia estocástica sería fácil de comprobar. Sean  $\vec{x} = (x_1, \dots, x_{N_X})$  e  $\vec{y} = (y_1, \dots, y_{N_Y})$  muestras de  $X$  e  $Y$  respectivamente. Las hipótesis que interesan son:

$$\begin{cases} H_0 \equiv X \leq_{st} Y \stackrel{Eq. (3.1)}{\equiv} F_X(x) \geq F_Y(x), & \forall x \in \mathbb{R}. \\ H_1 \equiv X \not\leq_{st} Y \stackrel{Eq. (3.1)}{\equiv} F_X(x) < F_Y(x), & \text{para algún } x \in \mathbb{R}. \end{cases} \quad (4.1)$$

Por tanto, se busca que los tests estén desarrollados para dichas hipótesis o, en su defecto, poder modificarlos para conseguir el contraste deseado.

A lo largo de toda la sección, se considerarán las muestras  $\vec{x}$  e  $\vec{y}$  mostradas en la Tabla 4.1 para ilustrar los distintos tests de dominancia estocástica. Estas muestras fueron obtenidas por simulación de manera independiente a partir de  $X \equiv \mathcal{N}(0, 1)$  e  $Y \equiv \mathcal{N}(1, 1)$ , luego  $X \leq_{st} Y$ . Además, en este ejemplo se resalta la importancia de los tests de dominancia estocástica ya que, a pesar de que la haya, las funciones de distribución empíricas se cruzan, por lo que a priori es difícil establecer dicha relación.

$\vec{x}$	$\vec{y}$
0.55	-0.09
-0.28	1.46
1.78	-0.36
0.19	-0.86
1.14	0.56
0.42	0.81
1.23	2.40
0.24	1.10
-0.37	0.89
1.11	1.70

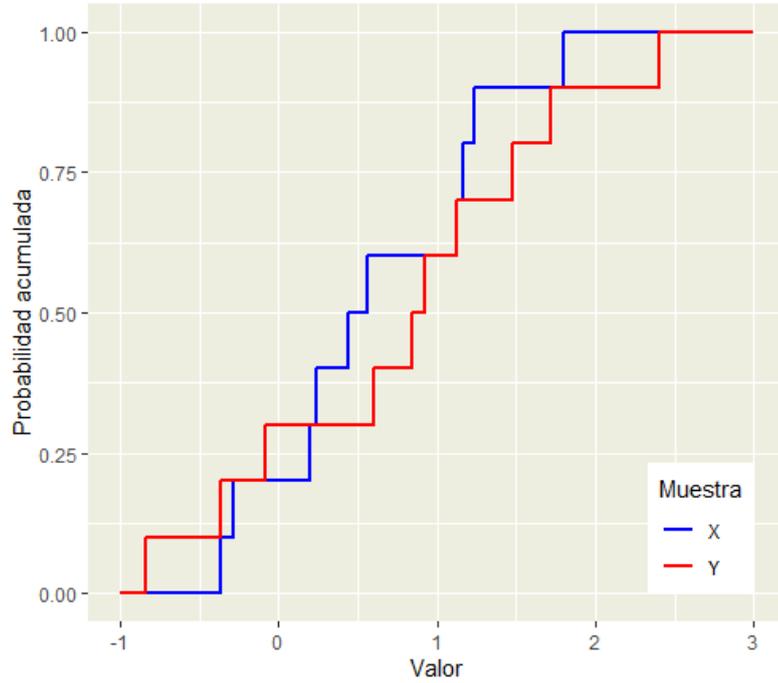


Tabla 4.1: Ejemplo recurrente utilizado para ilustrar los distintos tests de dominancia estocástica.

## 4.1. Test de Anderson

Este test fue propuesto por G. Anderson en [1].

Sea  $I = [a, b]$  el soporte de  $X$  e  $Y$ . Se considera una partición disjunta de  $I$  en  $k$  intervalos, y se definen  $p^X := \frac{1}{N_X} \vec{z}_X$  y  $p^Y := \frac{1}{N_Y} \vec{z}_Y$ , siendo  $\vec{z}_X$  y  $\vec{z}_Y$  los vectores de frecuencias empíricas de  $X$  e  $Y$  respectivamente. Se define:

$$I_f := \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & 1 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 1 & 1 & 1 & 1 & \dots & 0 \\ 1 & 1 & 1 & 1 & \dots & 1 \end{pmatrix},$$

que es un matriz de dimensiones  $k \times (k + 1)$ . Notar que:

$$\begin{cases} H_0 \equiv X \leq_{st} Y \equiv (I_f(p^Y - p^X))_i \leq 0, \quad \forall i \in \{1, \dots, k\}. \\ H_1 \equiv X \not\leq_{st} Y \equiv \exists i \in \{1, \dots, k\}, / (I_f(p^Y - p^X))_i > 0. \end{cases}$$

Debido a que  $\frac{1}{\sqrt{N}}(\vec{z} - N\vec{\mu}) \xrightarrow{\mathcal{L}} \mathcal{N}_k(\vec{0}_k, \Sigma)$ , se tiene que:

$$\begin{aligned} \sqrt{N_X} \cdot p^X - \frac{\vec{\mu}_X}{\sqrt{N_X}} &= \frac{1}{\sqrt{N_X}}(\vec{z}_X - \vec{\mu}_X) \xrightarrow{\mathcal{L}} \mathcal{N}_k(\vec{0}_k, \Sigma_X), \\ \sqrt{N_Y} \cdot p^Y - \frac{\vec{\mu}_Y}{\sqrt{N_Y}} &= \frac{1}{\sqrt{N_Y}}(\vec{z}_Y - \vec{\mu}_Y) \xrightarrow{\mathcal{L}} \mathcal{N}_k(\vec{0}_k, \Sigma_Y). \end{aligned}$$

Además, en el caso más desfavorable de  $H_0$ , es decir, cuando se tiene la igualdad, se verifica:  $p^X \equiv p^Y$ , si  $m = \frac{N_X + N_Y}{N_X \cdot N_Y}$ , y si se supone independencia entre las variables, se tiene que:

$$\begin{aligned} \frac{1}{\sqrt{m}}(p^Y - p^X) &\xrightarrow{\mathcal{L}} \mathcal{N}_k(0, \Sigma). \\ \frac{1}{\sqrt{m}}(I_f(p^Y - p^X)) &\xrightarrow{\mathcal{L}} \mathcal{N}_k(0, I_f \Sigma I_f'). \end{aligned}$$

Por tanto, un estadístico natural es:

$$A_i = \frac{[I_f(p^Y - p^X)]_i}{\sqrt{[m I_f \Sigma I_f']_{i,i}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1) \quad \forall i = 1, \dots, k. \quad (4.2)$$

Notar que, en el caso frontera de  $H_0$ ,  $\vec{p}$  puede aproximarse por:  $\frac{\vec{z}_X + \vec{z}_Y}{N_X + N_Y}$ .

En conclusión, se puede aplicar un test de unión-intersección para hacer múltiples comparaciones [19]. Por tanto, el nivel de significación vendrá dado asintóticamente por el valor crítico del módulo máximo estudentizado con  $k$  e  $\infty$  grados de libertad. Además, la región crítica es:

$$RC_\alpha = \left\{ \vec{x}, \vec{y} / -A_i \leq M_{\infty, \alpha}^k \quad \forall i \in \{1, \dots, k\} \wedge \exists j \in \{1, \dots, k\} : |A_j| \geq M_{\infty, \alpha}^k \right\}. \quad (4.3)$$

**Observación 5.** *Se puede notar que:*

- *Se asume independencia en el muestreo.*
- *Se recomienda usar entre 10 y 15 puntos en la malla [20] ya que, aunque si se usasen más puntos se obtendría más información de las distribuciones, se incumpliría la hipótesis de independencia del test GT2 [18].*

- Aunque la partición se suele hacer usando los deciles para estar en sintonía con las publicaciones usuales, si una de las distribuciones fuese conocida, el test ganaría potencia si los puntos de la partición se ubican en la intersección de la otra distribución bajo  $H_0$  y bajo  $H_1$ . En general, basta imponer que  $np_i > 5$  para todo  $i \in \{1, \dots, k\}$  y para  $n = N_X, N_Y$  [1].

**Ejemplo 4.1.** En este caso, el valor crítico para  $\alpha = 0.05$  es  $M_{\infty, \alpha}^k = 2.8$ . Por otro lado, los valores observados del estadístico del test de Anderson para las muestras de la Tabla 4.1 son:  $A_1 = 0$ ,  $A_2 = 0$ ,  $A_3 = 0$ ,  $A_4 = -0.91$ ,  $A_5 = -0.81$ ,  $A_6 = 0$ ,  $A_7 = 0$ ,  $A_8 = -1.12$ ,  $A_9 = 0$ ,  $A_{10} = 1$ , por lo que  $-A_i \leq 2.8$  y  $|A_i| < 2.8$  para todo  $i \in \{1, \dots, k\}$ . Por ello, la muestra no está en la región crítica y no hay evidencias significativas para rechazar  $H_0$ .

## 4.2. Test de Davidson-Duclos

Este test fue propuesto por R. Davidson y J. Duclos en [4].

Para definir el estadístico se consideran:

$$\hat{V}_{F_H}(w) := \frac{1}{N_H} \left( \frac{1}{N_H} \sum_{i=1}^{N_H} \mathbb{I}(h_i \leq w) - \hat{F}_H(w) \right)^2, \quad \text{con } H = X, Y.$$

$$\hat{V}_{F_X F_Y}(w) := \frac{1}{n} \left( \frac{1}{n} \sum_{i=1}^n \mathbb{I}(x_i \leq w) \mathbb{I}(y_i \leq w) - \hat{F}_X(w) \hat{F}_Y(w) \right), \quad \text{si } n := N_X = N_Y.$$

En [8] se ve que el test es válido tanto si se supone independencia como dependencia ( $N_X = N_Y = n$  y, de esta manera,  $(x_i, y_i)$  es un par observado) de las variables, sin embargo, el estadístico cambia en cada caso:

- **Independencia**

$$T(w) := \frac{\hat{F}_X(w) - \hat{F}_Y(w)}{\sqrt{\hat{V}_{F_X}(w) + \hat{V}_{F_Y}(w)}}. \quad (4.4)$$

- **Dependencia**

$$T(w) := \frac{\hat{F}_X(w) - \hat{F}_Y(w)}{\sqrt{\hat{V}_{F_X}(w) + \hat{V}_{F_Y}(w) - 2\hat{V}_{F_X F_Y}(w)}}. \quad (4.5)$$

En ambos casos, bajo  $H_0$  se tiene que:

$$T \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

Notar que  $H_0$  se tiene que examinar en todo el soporte de  $F_X$  y  $F_Y$ , lo cual es imposible empíricamente. Para solucionar esto, en [3] se propone aplicar un test de unión-intersección para hacer múltiples comparaciones y, por tanto, nuevamente el nivel de significación vendrá dado asintóticamente por el valor crítico del módulo máximo estudentizado con  $k$  e  $\infty$  grados de libertad. Además, la región crítica es:

$$RC_\alpha = \left\{ \bar{x}, \bar{y} / T(z_i) \leq M_{\infty, \alpha}^k \quad \forall i \in \{1, \dots, k\} \wedge \exists j \in \{1, \dots, k\} : |T(z_j)| \geq M_{\infty, \alpha}^k \right\}. \quad (4.6)$$

**Observación 6.** *Se puede notar que:*

- *Al admitirse distintos tamaños de muestra y ya que en ningún momento se supone independencia a la hora de construir el estadístico, el test es válido tanto si se supone independencia como dependencia ( $N_X = N_Y = n$  y, de esta manera,  $(x_i, y_i)$  es un par observado) de las variables.*
- *Se recomienda usar entre 10 y 15 puntos en la malla [20] ya que, aunque si se usasen más puntos se obtendría más información de las distribuciones, se incumpliría la hipótesis de independencia del test GT2 [18].*
- *El número de puntos de la malla se suele elegir en base al sentido común. Debido a esta falta de un criterio claro, muchas publicaciones usan los deciles para hacer la partición [19].*

**Ejemplo 4.2.** *En este caso, el valor crítico para  $\alpha = 0.05$  es  $M_{\infty, \alpha}^k = 2.8$ . Por otro lado, los valores observados del estadístico del test de Davidson-Duclos para las muestras de la Tabla 4.1 son:  $T_1 = 0$ ,  $T_2 = 0$ ,  $T_3 = 0$ ,  $T_4 = 0.93$ ,  $T_5 = 0.91$ ,  $T_6 = 0$ ,  $T_7 = 0$ ,  $T_8 = 1.15$ ,  $T_9 = 0$ ,  $T_{10} = 0$ , por lo que  $T_i \leq 2.8$  y  $|T_i| < 2.8$  para todo  $i \in \{1, \dots, k\}$ . Por ello, la muestra no está en la región crítica y no hay evidencias significativas para rechazar  $H_0$ .*

### 4.3. Test de Barrett-Donald

Este test fue propuesto por G.-F. Barrett y S.G. Donald en [2] como generalización del test propuesto por F. McFadden en [13].

Se supondrá que:

- (i)  $F_X$  y  $F_Y$  tienen soporte común  $I = [a, b]$ .
- (ii)  $F_X$  y  $F_Y$  son funciones continuas en  $I$ .
- (iii) Se supone que  $\vec{x}$  e  $\vec{y}$  son muestras independientes.
- (iv) Si  $N_X, N_Y \rightarrow \infty$ , entonces  $\frac{N_X}{N_X + N_Y} \rightarrow \phi \in (0, 1)$ .

Se define el estadístico:

$$BD := \left( \frac{N_X \cdot N_Y}{N_X + N_Y} \right)^{\frac{1}{2}} \sup_{w \in \mathbb{R}} \left( \hat{F}_Y(w) - \hat{F}_X(w) \right). \quad (4.7)$$

Si  $\mathcal{B}_X, \mathcal{B}_Y$  son puentes Brownianos estándar independientes, en el caso más desfavorable de  $H_0$ , es decir, cuando se tiene la igualdad; se verifica, por el Teorema 2.3, que:

$$\begin{aligned} BD &= \left( \frac{N_Y \cdot N_X}{N_Y + N_X} \right)^{\frac{1}{2}} \sup_{w \in \mathbb{R}} \left[ (\hat{F}_Y(w) - F_Y(w)) - (\hat{F}_X(w) - F_X(w)) \right] \\ &\xrightarrow{\mathcal{L}} \left( \frac{1}{N_Y + N_X} \right)^{\frac{1}{2}} \sup_{w \in \mathbb{R}} \left[ \sqrt{N_X} (\mathcal{B}_Y \circ F_Y(w)) - \sqrt{N_Y} (\mathcal{B}_X \circ F_X(w)) \right]. \end{aligned}$$

Además, la región crítica es:

$$RC_\alpha = \{ \vec{x}, \vec{y} / BD > c_\alpha \}. \quad (4.8)$$

Como la distribución asintótica de  $BD$  depende de las funciones de distribución, en [2] se proponen tres procedimientos para hallar el p-valor mediante simulación, de los cuales ninguno resalta por su eficiencia. En este trabajo se considera, por su fácil implementación, el método de los multiplicadores: Se considera  $(U_1^X, \dots, U_{N_X}^X)$  una muestra aleatorias simple de  $\mathcal{N}(0, 1)$ , entonces:

$$\left( \mathcal{B}_X^* \circ \hat{F}_X \right) (w) := \frac{1}{\sqrt{N_X}} \sum_{i=1}^{N_X} \left( \mathbb{I}(x_i \leq w) - \hat{F}_X(w) \right) U_{X,i}.$$

simula,  $(\mathcal{B}_X \circ F_X) (w)$ . De esta manera, el p-valor es:

$$p = P_U \left( \sup_{w \in I} \left( \mathcal{B}_X^* \circ \hat{F}_X \right) > BD \right).$$

siendo  $P_U$  la función de probabilidad asociada a la variable aleatoria  $U_{X,i}$ , que depende de la muestra realizada. Para aproximar dicha probabilidad se utiliza simulación de Monte Carlo. Si  $\{U_{i,r}^X\}_{i=1}^{N_X}$  es la  $r$ -ésima muestra de  $U_i^X$ , con  $r = 1, 2, \dots, R$ ;

siendo  $R$  el número de iteraciones usadas en la simulación, se define:

$$S_r^X = \frac{1}{\sqrt{N_X}} \sup_{w \in I} \sum_{i=1}^{N_X} \left( \mathcal{B}_X^* \circ \hat{F}_X \right) (w) \cdot U_{i,r}^X.$$

Entonces, el p-valor se puede aproximar por:

$$p \approx \frac{1}{R} \sum_{r=1}^R \mathbb{I} (S_r^X \leq BD).$$

**Observación 7.** *Se puede notar que:*

- *En las hipótesis del test se asume independencia entre las variables.*
- *Es costoso computacionalmente. Sin embargo, hay que destacar que en [2] se afirma que no se suele necesitar muchos cálculos para llegar a una buena aproximación del p-valor.*

**Ejemplo 4.3.** *En este caso, el p-valor del test de Barrett-Donald para las muestras de la Tabla 4.1 es 0.61. Como el nivel de significación considerado es  $\alpha = 0.05$ , la muestra no está en la región crítica y no hay evidencias significativas para rechazar  $H_0$ .*

#### 4.4. Test de Linton-Maasoumi-Whang

Este test fue propuesto por O. Linton, E. Maasoumi y Y.-J. Whang en [11].

Sea  $I = [a, b]$  es el soporte de  $X$  e  $Y$  y se supone que  $n = N_X = N_Y$ . Se considera el siguiente contraste:

$$\begin{cases} H'_0 \equiv \min\{\sup_{w \in I}(F_Y(w) - F_X(w)), \sup_{w \in I}(F_X(w) - F_Y(w))\} \leq 0, \\ H'_1 \equiv \min\{\sup_{w \in I}(F_Y(w) - F_X(w)), \sup_{w \in I}(F_X(w) - F_Y(w))\} > 0. \end{cases}$$

La interpretación de la hipótesis nula es que una de las dos variables domina estocásticamente a la otra. En caso de no tener suficiente evidencia para rechazar la hipótesis nula, se puede representar las funciones de distribución empírica y analizar la dominancia estocástica gráficamente.

Volviendo con el contraste, se define el estadístico:

$$LMW := \sqrt{n} \cdot \min \left\{ \sup_{w \in I} \left( \hat{F}_Y(w) - \hat{F}_X(w) \right), \sup_{w \in I} \left( \hat{F}_X(w) - \hat{F}_Y(w) \right) \right\}. \quad (4.9)$$

Si  $\mathcal{B}_X, \mathcal{B}_Y$  son puentes Brownianos estándar independientes, en el caso más desfavorable de  $H_0$ , es decir, cuando se tiene la igualdad; argumentando como en el test de Barret-Donald, si  $\mathcal{C}_0 := \{w \in I / \hat{F}_Y(w) - \hat{F}_X(w) = 0\}$ , se tiene que:

$$LMW \xrightarrow{\mathcal{L}} \min \left\{ \sup_{w \in \mathcal{C}_0} (\mathcal{B}_Y \circ F_Y(w) - \mathcal{B}_X \circ F_X(w)), \sup_{w \in \mathcal{C}_0} (\mathcal{B}_X \circ F_X(w) - \mathcal{B}_Y \circ F_Y(w)) \right\}.$$

Además, la región crítica es:

$$RC_\alpha = \{\vec{x}, \vec{y} / LMW > c_\alpha\}. \quad (4.10)$$

Al igual que en el test de Barret-Donald, la distribución asintótica del estadístico depende de las funciones de distribución. Sin embargo, en [11] se propone un procedimiento de submuestreo, que no sólo unifica la distribución simulada del estadístico, sino que también es válido para cualquier esquema de muestreo.

Considerando  $Z_i$  un par observado de  $X$  e  $Y$ , se define:

$$t_m(Z_1, \dots, Z_m) = \min \left\{ \sup_{w \in I} (\hat{F}_Y(w) - \hat{F}_X(w)), \sup_{w \in I} (\hat{F}_X(w) - \hat{F}_Y(w)) \right\}.$$

Fijado  $b \in \{1, \dots, n\}$ , el procedimiento de submuestreo consiste en:

1. Sea  $\mathcal{S} := \{W_1, \dots, W_n\}$  la muestra original, con  $W_i := (x_i, y_i)$  para todo  $i \in \{1, \dots, n\}$ . Notar que:

$$LMW = \sqrt{n} \cdot t_n(\mathcal{S}). \quad (4.11)$$

2. Generar submuestras  $\mathcal{S}_{i,b} = \{W_i, \dots, W_{i+b-1}\}$  para  $i = 1, \dots, n - b + 1$ .
3. Calcular  $t_{b,i} := t_b(W_i, \dots, W_{i+b-1})$  para  $i = 1, \dots, n - b + 1$  usando las submuestras.
4. Calcular el valor crítico  $c_\alpha$ , teniendo en cuenta la Ecuación (4.11), es el cuantil  $(1 - \alpha)$  de:

$$\hat{G}_b(w) = \frac{1}{n - b + 1} \sum_{i=1}^{n-b+1} \mathbb{I}(\sqrt{b} \cdot t_{b,i} \leq w).$$

Ahora bien, una pregunta natural es si la elección de  $b$  puede afectar al resultado del test. En [11][página 748] se afirma que « $b$  no afecta a la distribución del estadístico de primer orden bajo la hipótesis nula». Así, en este trabajo se considera para la experimentación  $b = \lceil \frac{n}{2} \rceil$ , por estar definido para cualquier  $n \geq 2$  y por simplicidad.

Dado que se buscan tests que contrasten las Hipótesis de la forma mostrada en la Ecuación (4.1), se harán pequeños ajustes en la implementación del test, los cuales se basan en la idea de comparar las funciones de distribución gráficamente en caso de haber dominancia estocástica. También hay que tener en cuenta que en el desarrollo de [11] se considera la dominancia estocástica estricta, por tanto, quizás se rechaza  $H'_0$  pero hay dominancia estocástica. El pseudocódigo utilizado es el Algoritmo 1.

**Observación 8.** *Se puede notar que:*

- *En ningún momento se asume independencia entre las variables. Por tanto, el test es válido tanto si se supone independencia como dependencia ( $N_X = N_Y = n$  y, de esta manera,  $(x_i, y_i)$  es un par observado) de las variables.*
- *El test se puede extender para cualquier número de variables aleatorias.*

---

**Algoritmo 1:** Pseudocódigo para la implementación del test LMW

---

**Data:** Muestras de X e Y

```

1 if Se rechaza  $H'_0$  then
2   if la probabilidad de que  $|\hat{F}_Y - \hat{F}_X|$  en  $I$  sea pequeño es más de 0.5 then
3     no hay evidencias significativas para rechazar  $H_0$ ;
4   else
5     se rechaza  $H_0$ ;
6 else
7   if la probabilidad de que  $\hat{F}_Y - \hat{F}_X$  en  $I$  sea negativo es alta then
8     no hay evidencias significativas para rechazar  $H_0$ ;
9   else
10    se rechaza  $H_0$ ;

```

---

**Ejemplo 4.4.** *En este caso, el valor crítico para  $\alpha = 0.05$  es  $\hat{G}_5^{-1}(1 - \alpha) = 0.2$ . Por otro lado, el valor observado del estadístico del test de Linton-Maasoumi-Whang para las muestras de la Tabla 4.1 es  $LMW = 0.32$ , por lo que  $LMW > 0.2$ . Por ello, la muestra está en la región crítica y hay evidencias significativas para rechazar  $H_0$ .*

## 4.5. Test de Kaur-Rao-Singh

Este test fue propuesto por A. Kaur, B.L.S.P. Rao y H. Singh en [7]. Pese a que originalmente el test KRS se introdujo para dominancia estocástica de segundo orden, se ha extendido a otros órdenes [19]. Dicha extensión es debida a que el estadístico que se considera para dominancia estocástica de primer orden es asintóticamente equivalente al estadístico propuesto en el test original.

Originalmente, se considera el siguiente contraste:

$$\begin{cases} H'_0 \equiv X \not\leq_{st} Y \stackrel{Eq. (3.1)}{\equiv} F_X(x) < F_Y(x) & \text{para algún } x \in \mathbb{R}. \\ H'_1 \equiv X \leq_{st} Y \stackrel{(3.1)}{\equiv} F_X(x) \geq F_Y(x) & \forall x \in \mathbb{R}. \end{cases}$$

No obstante, como la distribución del estadístico bajo  $H_0$  se calcula en el caso frontera, y por tanto el marco teórico no cambia, se puede considerar las Hipótesis mostradas en la Ecuación (4.1).

Suponiendo que  $T$  el estadístico definido en las Ecuaciones (4.4) y (4.5) en su correspondiente caso e  $I = [a, b]$  es el soporte de  $X$  e  $Y$ , se define el estadístico:

$$T_M := \inf_{x \in I} \{T(x)\}. \quad (4.12)$$

El cual sigue una distribución  $\mathcal{N}(0, 1)$  bajo  $H_0$ . Además, la región crítica es:

$$RC_\alpha = \{\vec{x}, \vec{y} / T_M \leq c_\alpha\}. \quad (4.13)$$

**Observación 9.** *Al usar el estadístico del test de Davidson-Duclos y asumir independencia entre las variables en ningún momento, se tiene que el test es válido tanto si se supone independencia como dependencia ( $N_X = N_Y = n$  y, de esta manera,  $(x_i, y_i)$  es un par observado) de las variables.*

**Ejemplo 4.5.** *En este caso, el valor crítico para  $\alpha = 0.05$  es  $F_{\mathcal{N}(0,1)}^{-1}(\alpha) = -1.64$ . Por otro lado, el valor observado del estadístico del test de Kaur-Rao-Singh para las muestras de la Tabla 4.1 es  $T_M = -1.05$ , por lo que  $T_M \leq -1.64$ . Por ello, la muestra no está en la región crítica y no hay evidencias significativas para rechazar  $H_0$ .*

## 4.6. Test de Kolmogorov-Smirnov

Como se ha visto en los epígrafes dedicados a los tests: «Barret-Donald», «Linton-Maasoumi-Whang» y «Kaur-Rao-Singh», el estadístico de Kolmogorov-Smirnov:

$$KS := \sup_{w \in I} \left| \hat{F}_Y(w) - \hat{F}_X(w) \right|,$$

con algunas variaciones puede ser usado con el fin de realizar el contraste con las Hipótesis mostradas en la Ecuación (4.1). Por lo tanto, es natural preguntarse si dicho estadístico por sí sólo puede servir para estudiar la dominancia estocástica.

En RStudio, la función «ks.test» tiene el argumento de entrada «alternative», el cual, si se especifica que sea igual a «greater» y se introducen primero la muestra de  $Y$  y después la de  $X$ , nos permite realizar el contraste:

$$\left\{ \begin{array}{l} H'_0 \equiv F_Y(x) = F_X(x), \quad \forall x \in \mathbb{R}. \\ H'_1 \equiv F_Y(x) > F_X(x), \quad \forall x \in \mathbb{R}. \end{array} \right.$$

Así, teniendo en cuenta la caracterización de la Ecuación (4.1) según las funciones de distribución, si se rechaza  $H'_0$  lo natural es que se rechace  $H_0$  también, pues se acepta que:  $F_Y(x) > F_X(x)$ ,  $\forall x \in \mathbb{R}$ . En contraposición, si no hay evidencias suficientes para rechazar  $H'_0$ , lo lógico es que no haya evidencias suficientes para rechazar  $H_0$  pues no se rechaza que:  $F_Y(x) = F_X(x)$ ,  $\forall x \in \mathbb{R}$ , que es el caso frontera de  $H_0$ ; por ello también se mantiene el nivel de significación.

**Observación 10.** *Al usar el test de Kolmogorov-Smirnov, no se asume independencia entre las variables. Por tanto, el test se puede utilizar en cualquier situación de manera análoga a otros tests.*

**Ejemplo 4.6.** *En este caso, el  $p$ -valor del test de Kolmogorov-Smirnov para las muestras de la Tabla 4.1 es 0.42. Como el nivel de significación considerado es  $\alpha = 0.05$ , la muestra no está en la región crítica y no hay evidencias significativas para rechazar  $H_0$ .*

## Capítulo 5

# Experimentación

### 5.1. Metodología

Se estudiará el comportamiento de los tests presentados en el capítulo anterior como función de un factor de desplazamiento entre dos distribuciones pertenecientes a la misma familia de localización. En concreto, se estudian las familias de localización generadas por la distribución normal estándar,  $\mathcal{N}(0, 1)$ , la distribución exponencial estándar,  $\text{Exp}(1)$ , y la distribución uniforme estándar,  $\mathcal{U}(0, 1)$ ; con desplazamientos de paso 0.1 entre  $-1$  y  $1$ , es decir, se estudia la potencia de los tests para detectar  $X \leq_{st} Y$  siendo  $X$  una de esas distribuciones e  $Y$  la distribución de  $X + b$  con  $b \in [-1, 1]$ . Por el Teorema 3.5 y el Corolario 3.6, se tiene que  $H_0$  está asociada con los desplazamientos mayores o iguales que 0 ( $b \geq 0$ ) y  $H_1$  está asociada con los desplazamientos menores que 0 ( $b < 0$ ). Los tamaños muestrales considerados son  $m \in \{10, 65, 100\}$ .

Para dicho objetivo, se usará simulación por Monte Carlo, con  $10^3$  réplicas en cada caso. Se trabajará a nivel de significación  $\alpha = 0.05$ , motivo por el que la recta horizontal punteada en negro que aparece en las gráficas de la potencia es la recta  $y = 0.05$ . Además, como se dijo en la introducción, se consideran los siguientes casos de relación entre las variables: independencia, comonotonía y contramonotonía. Para el primer caso, simplemente se generan ambas muestras por separado. Para expresar las relaciones de dependencia en los casos de comonotonía y contramonotonía se seguirá el **método de la distribución condicionada** [14, Cap. 2], es decir, dadas

$X$  e  $Y$  dos variables aleatorias, se siguen estos pasos:

1. Generar dos valores aleatorios  $u, t$  que provienen de una  $U(0, 1)$ .
2. Sea  $v = c_u^{-1}(t)$ , donde  $c_u^{-1}$  es la pseudo-inversa de  $c_u$ , siendo  $c_u$  una función particular a cada cópula:

- Cópula del mínimo (comonotonía):

$$c_{u,M}(v) = \begin{cases} 1, & \text{si } v \leq u, \\ 0, & \text{si } v > u, \end{cases} \implies c_{u,M}^{-1}(t) = \inf[u, \infty) = u.$$

- Cópula de Lukaszewicz (contramonotonía):

$$c_{u,L}(v) = \begin{cases} 0, & \text{si } v \leq 1 - u, \\ 1, & \text{si } v > 1 - u, \end{cases} \implies c_{u,L}^{-1}(t) = \inf[1 - u, \infty) = 1 - u.$$

3. Para hallar un valor aleatorio de  $(X, Y)$ , se toma  $x := F_X^{-1}(u)$  e  $y := F_Y^{-1}(v)$ .

## 5.2. Independencia entre las variables

### 5.2.1. Variables aleatorias con distribución normal

Viendo las Tablas 5.1, 5.2 y 5.3, se puede observar que:

1. En este caso, los tests: KRS, LMW; no son capaces de mantener el nivel de significación bajo  $H_0$ .
2. Los tests: A y DD; son conservadores, en este caso.
3. En este caso, los tests: BD y KS; son los más potentes bajo  $H_1$  entre los que mantienen el nivel de significación bajo  $H_0$ .
4. Para los distintos tamaños muestrales no hay un cambio significativo en dichos comportamientos. Lo más relevante en este aspecto es que si  $m = 10$ , BD y KS no mantienen el nivel de significación bajo  $H_0$  por poco, pero esto tiene sentido por el escaso tamaño muestral.

Luego, el mejor test para esta situación es el de Kolmogorov-Smirnov, pues prácticamente mantiene el nivel de significación bajo  $H_0$  y, entre los que lo mantienen, es el más potente bajo  $H_1$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	0.28	0.63	0.49	0.85	1.00	0.71
-0.90	0.24	0.54	0.39	0.78	1.00	0.65
-0.80	0.22	0.49	0.34	0.74	1.00	0.57
-0.70	0.14	0.36	0.26	0.66	1.00	0.49
-0.60	0.11	0.32	0.21	0.62	0.99	0.43
-0.50	0.08	0.26	0.15	0.54	0.99	0.33
-0.40	0.06	0.23	0.13	0.47	0.98	0.29
-0.30	0.04	0.17	0.07	0.38	0.96	0.21
-0.20	0.02	0.11	0.05	0.32	0.95	0.16
-0.10	0.01	0.09	0.03	0.26	0.93	0.14
0.00	0.01	0.07	0.03	0.22	0.88	0.09
0.10	0.01	0.04	0.01	0.16	0.84	0.05
0.20	0.00	0.03	0.01	0.12	0.77	0.04
0.30	0.00	0.02	0.00	0.07	0.69	0.03
0.40	0.00	0.02	0.00	0.06	0.66	0.03
0.50	0.00	0.02	0.00	0.05	0.56	0.01
0.60	0.00	0.00	0.00	0.03	0.49	0.01
0.70	0.00	0.00	0.00	0.02	0.44	0.00
0.80	0.00	0.00	0.00	0.01	0.31	0.00
0.90	0.00	0.00	0.00	0.01	0.23	0.00
1.00	0.00	0.00	0.00	0.00	0.20	0.00

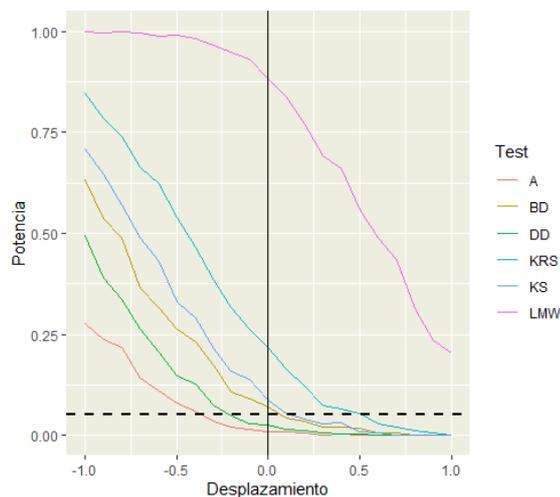


Tabla 5.1: Potencia para diferentes desplazamientos de los diferentes test bajo independencia con tamaño muestral  $m = 10$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	0.99
-0.90	0.99	1.00	0.99	1.00	1.00	0.97
-0.80	0.96	0.98	0.96	1.00	1.00	0.96
-0.70	0.88	0.96	0.88	1.00	1.00	0.93
-0.60	0.77	0.89	0.77	0.98	1.00	0.87
-0.50	0.62	0.79	0.62	0.95	0.98	0.74
-0.40	0.41	0.65	0.41	0.92	0.96	0.61
-0.30	0.25	0.45	0.25	0.77	0.92	0.42
-0.20	0.11	0.27	0.11	0.62	0.83	0.28
-0.10	0.05	0.14	0.05	0.44	0.76	0.14
0.00	0.01	0.07	0.01	0.29	0.71	0.06
0.10	0.00	0.03	0.00	0.16	0.62	0.02
0.20	0.00	0.00	0.00	0.09	0.58	0.01
0.30	0.00	0.00	0.00	0.04	0.50	0.00
0.40	0.00	0.00	0.00	0.02	0.37	0.00
0.50	0.00	0.00	0.00	0.01	0.26	0.00
0.60	0.00	0.00	0.00	0.00	0.18	0.00
0.70	0.00	0.00	0.00	0.00	0.12	0.00
0.80	0.00	0.00	0.00	0.00	0.08	0.00
0.90	0.00	0.00	0.00	0.00	0.06	0.00
1.00	0.00	0.00	0.00	0.00	0.04	0.00

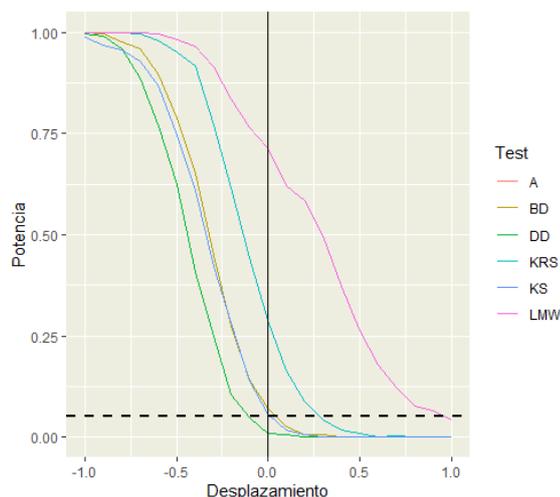


Tabla 5.2: Potencia para diferentes desplazamientos de los diferentes test bajo independencia con tamaño muestral  $m = 65$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.90	1.00	1.00	1.00	1.00	1.00	1.00
-0.80	1.00	1.00	1.00	1.00	1.00	0.99
-0.70	0.98	1.00	0.99	1.00	1.00	0.99
-0.60	0.95	0.97	0.96	1.00	1.00	0.97
-0.50	0.80	0.90	0.82	0.99	1.00	0.90
-0.40	0.63	0.77	0.66	0.95	0.98	0.83
-0.30	0.41	0.58	0.44	0.86	0.94	0.65
-0.20	0.20	0.36	0.22	0.71	0.87	0.44
-0.10	0.07	0.14	0.08	0.47	0.73	0.20
0.00	0.01	0.07	0.02	0.34	0.67	0.12
0.10	0.01	0.02	0.01	0.17	0.62	0.03
0.20	0.00	0.00	0.00	0.07	0.56	0.00
0.30	0.00	0.00	0.00	0.03	0.45	0.00
0.40	0.00	0.00	0.00	0.02	0.33	0.00
0.50	0.00	0.00	0.00	0.01	0.23	0.00
0.60	0.00	0.00	0.00	0.00	0.13	0.00
0.70	0.00	0.00	0.00	0.00	0.07	0.00
0.80	0.00	0.00	0.00	0.00	0.06	0.00
0.90	0.00	0.00	0.00	0.00	0.03	0.00
1.00	0.00	0.00	0.00	0.00	0.02	0.00

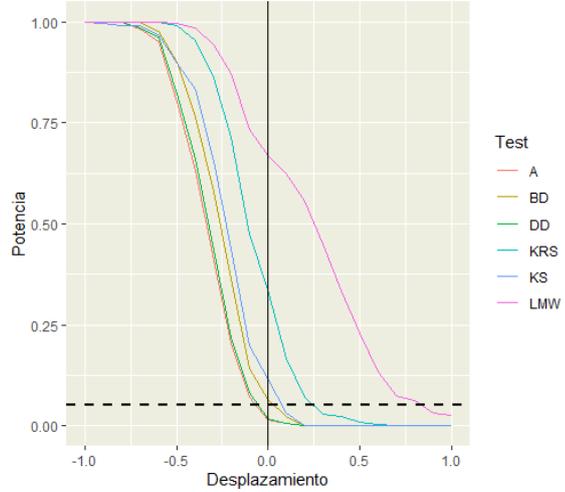


Tabla 5.3: Potencia para diferentes desplazamientos de los diferentes test bajo independencia con tamaño muestral  $m = 100$ .

### 5.2.2. Variables aleatorias con distribución uniforme

Viendo las Tablas 5.4, 5.5 y 5.6, se puede observar que:

1. En este caso, los tests: LMW, y en menor medida, KRS; no son capaces de mantener el nivel de significación bajo  $H_0$ .
2. Los tests: A y DD; son algo conservadores, en este caso.
3. En este caso, los tests: A, DD y KS; son los más potentes bajo  $H_1$  entre los que mantienen el nivel de significación bajo  $H_0$ .
4. De manera completamente análoga al caso anterior, para los distintos tamaños muestrales no hay un cambio significativo en dichos comportamientos, excepto si  $m = 10$  en cuyo caso BD y KS no mantienen el nivel de significación bajo  $H_0$  por poco, pero, nuevamente, esto tiene sentido por el escaso tamaño muestral.

Por tanto, el mejor test para esta situación es el de Kolmogorov-Smirnov por los mismos motivos que en el caso anterior: prácticamente mantiene el nivel de significación bajo  $H_0$  y, entre los que lo mantienen, es el más potente bajo  $H_1$ . Destacar

que el de Anderson y el de Davidson-Duclos también funcionan muy bien en este caso, pero son más conservadores.

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.90	1.00	1.00	1.00	1.00	1.00	1.00
-0.80	1.00	1.00	1.00	1.00	1.00	1.00
-0.70	0.99	0.99	0.99	1.00	1.00	1.00
-0.60	0.95	0.98	0.96	1.00	1.00	1.00
-0.50	0.79	0.89	0.84	0.99	1.00	0.98
-0.40	0.55	0.77	0.64	0.97	1.00	0.94
-0.30	0.31	0.55	0.42	0.86	1.00	0.81
-0.20	0.11	0.34	0.19	0.64	0.99	0.59
-0.10	0.04	0.16	0.08	0.41	0.97	0.32
0.00	0.01	0.07	0.03	0.22	0.86	0.08
0.10	0.00	0.02	0.00	0.09	0.61	0.01
0.20	0.00	0.01	0.00	0.03	0.36	0.00
0.30	0.00	0.00	0.00	0.00	0.16	0.00
0.40	0.00	0.00	0.00	0.00	0.05	0.00
0.50	0.00	0.00	0.00	0.00	0.01	0.00
0.60	0.00	0.00	0.00	0.00	0.00	0.00
0.70	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.00	0.00	0.00	0.00	0.01	0.00
0.90	0.00	0.00	0.00	0.00	0.01	0.00
1.00	0.00	0.00	0.00	0.00	0.01	0.00

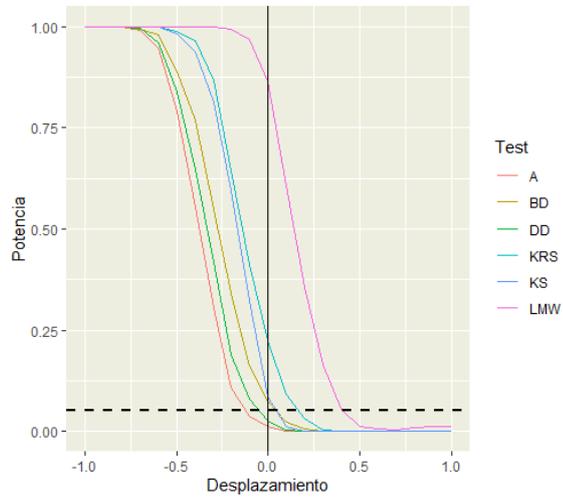


Tabla 5.4: Potencia para diferentes desplazamientos de los diferentes test bajo independencia con tamaño muestral  $m = 10$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.90	1.00	1.00	1.00	1.00	1.00	1.00
-0.80	1.00	1.00	1.00	1.00	1.00	1.00
-0.70	1.00	1.00	1.00	1.00	1.00	1.00
-0.60	1.00	1.00	1.00	1.00	1.00	1.00
-0.50	1.00	1.00	1.00	1.00	1.00	1.00
-0.40	1.00	1.00	1.00	1.00	1.00	1.00
-0.30	1.00	1.00	1.00	1.00	1.00	1.00
-0.20	0.98	0.91	0.98	0.99	1.00	0.97
-0.10	0.41	0.44	0.41	0.81	1.00	0.68
0.00	0.02	0.06	0.02	0.19	0.82	0.07
0.10	0.00	0.00	0.00	0.02	0.28	0.00
0.20	0.00	0.00	0.00	0.00	0.02	0.00
0.30	0.00	0.00	0.00	0.00	0.00	0.00
0.40	0.00	0.00	0.00	0.00	0.00	0.00
0.50	0.00	0.00	0.00	0.00	0.00	0.00
0.60	0.00	0.00	0.00	0.00	0.00	0.00
0.70	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00

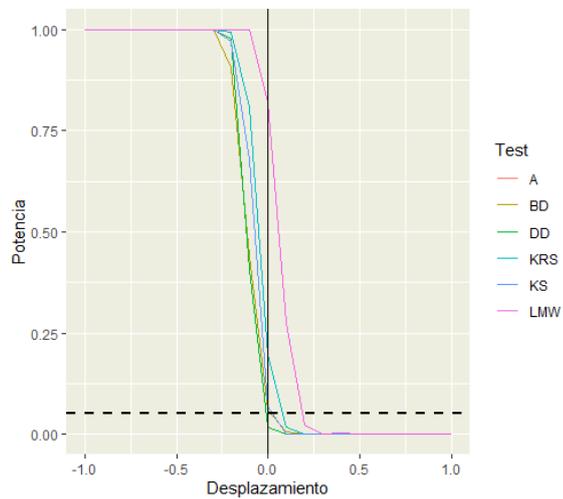


Tabla 5.5: Potencia para diferentes desplazamientos de los diferentes test bajo independencia con tamaño muestral  $m = 65$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.90	1.00	1.00	1.00	1.00	1.00	1.00
-0.80	1.00	1.00	1.00	1.00	1.00	1.00
-0.70	1.00	1.00	1.00	1.00	1.00	1.00
-0.60	1.00	1.00	1.00	1.00	1.00	1.00
-0.50	1.00	1.00	1.00	1.00	1.00	1.00
-0.40	1.00	1.00	1.00	1.00	1.00	1.00
-0.30	1.00	1.00	1.00	1.00	1.00	1.00
-0.20	1.00	0.98	1.00	1.00	1.00	1.00
-0.10	0.70	0.55	0.71	0.87	1.00	0.81
0.00	0.02	0.05	0.02	0.23	0.83	0.07
0.10	0.00	0.00	0.00	0.00	0.21	0.00
0.20	0.00	0.00	0.00	0.00	0.01	0.00
0.30	0.00	0.00	0.00	0.00	0.00	0.00
0.40	0.00	0.00	0.00	0.00	0.00	0.00
0.50	0.00	0.00	0.00	0.00	0.00	0.00
0.60	0.00	0.00	0.00	0.00	0.00	0.00
0.70	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00

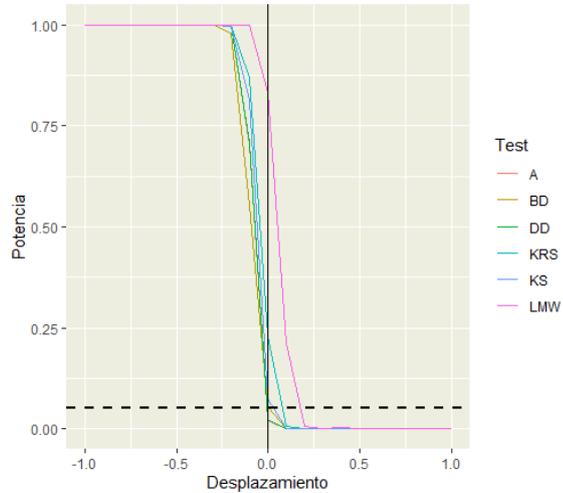


Tabla 5.6: Potencia para diferentes desplazamientos de los diferentes test bajo independencia con tamaño muestral  $m = 100$ .

### 5.2.3. Variables aleatorias con distribución exponencial

Viendo las Tablas 5.7, 5.8 y 5.9, se puede observar que:

1. En este caso, los tests: KRS y LMW; no son capaces de mantener el nivel de significación bajo  $H_0$ .
2. Los tests: A y DD; son un poco conservadores, en este caso.
3. En este caso, los tests: A y DD; son los más potentes bajo  $H_1$  entre los que mantienen el nivel de significación bajo  $H_0$ .
4. Dichos comportamientos son muy claros si  $m \in \{65, 100\}$ , pero si  $m = 10$  hay que añadir BD a los tests más potente bajo  $H_1$  entre los que mantienen el nivel de significación bajo. Además, KS mantiene el nivel de significación bajo  $H_0$  por poco, pero, como ya se comentó, esto tiene sentido por el escaso tamaño muestral.

Por tanto, los mejores tests para esta situación son el de Anderson y el Davidson-Duclos, ya que mantienen el nivel de significación bajo  $H_0$  (siendo un poco conservadores) y, entre los que lo mantienen, son los más potentes bajo  $H_1$ , independien-

temente del tamaño muestral, aunque el de Barrett-Donald funciona bastante bien para tamaños muestrales bajos.

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	0.78	0.78	0.81	0.85	1.00	0.67
-0.90	0.73	0.71	0.76	0.81	1.00	0.65
-0.80	0.63	0.60	0.67	0.73	1.00	0.57
-0.70	0.53	0.52	0.58	0.69	1.00	0.51
-0.60	0.40	0.44	0.46	0.61	0.99	0.52
-0.50	0.27	0.36	0.34	0.57	0.98	0.42
-0.40	0.17	0.25	0.24	0.44	0.97	0.39
-0.30	0.09	0.18	0.15	0.36	0.95	0.34
-0.20	0.05	0.11	0.09	0.27	0.91	0.22
-0.10	0.03	0.08	0.05	0.22	0.88	0.18
0.00	0.01	0.04	0.02	0.14	0.80	0.08
0.10	0.00	0.04	0.02	0.13	0.79	0.04
0.20	0.00	0.02	0.01	0.09	0.72	0.02
0.30	0.00	0.01	0.00	0.07	0.65	0.00
0.40	0.00	0.01	0.00	0.07	0.58	0.00
0.50	0.00	0.01	0.00	0.03	0.54	0.00
0.60	0.00	0.01	0.00	0.04	0.48	0.00
0.70	0.00	0.00	0.00	0.02	0.45	0.00
0.80	0.00	0.00	0.00	0.02	0.40	0.00
0.90	0.00	0.00	0.00	0.01	0.36	0.00
1.00	0.00	0.00	0.00	0.00	0.31	0.00

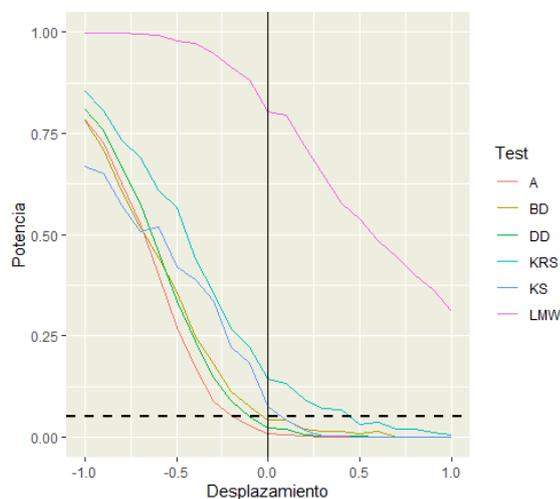


Tabla 5.7: Potencia para diferentes desplazamientos de los diferentes test bajo independencia con tamaño muestral  $m = 10$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	0.97	1.00	0.97	0.96	0.88
-0.90	1.00	0.97	1.00	0.95	0.96	0.85
-0.80	1.00	0.93	1.00	0.93	0.93	0.80
-0.70	1.00	0.88	1.00	0.89	0.93	0.77
-0.60	1.00	0.81	1.00	0.84	0.89	0.73
-0.50	1.00	0.73	1.00	0.80	0.88	0.64
-0.40	1.00	0.60	1.00	0.70	0.84	0.57
-0.30	0.98	0.46	0.98	0.58	0.81	0.50
-0.20	0.72	0.28	0.72	0.47	0.76	0.37
-0.10	0.16	0.13	0.16	0.30	0.66	0.23
0.00	0.01	0.05	0.01	0.20	0.60	0.05
0.10	0.00	0.02	0.00	0.13	0.50	0.00
0.20	0.00	0.01	0.00	0.10	0.48	0.00
0.30	0.00	0.00	0.00	0.07	0.45	0.00
0.40	0.00	0.00	0.00	0.05	0.41	0.00
0.50	0.00	0.00	0.00	0.05	0.41	0.00
0.60	0.00	0.00	0.00	0.03	0.38	0.00
0.70	0.00	0.00	0.00	0.02	0.36	0.00
0.80	0.00	0.00	0.00	0.02	0.31	0.00
0.90	0.00	0.00	0.00	0.01	0.28	0.00
1.00	0.00	0.00	0.00	0.01	0.27	0.00

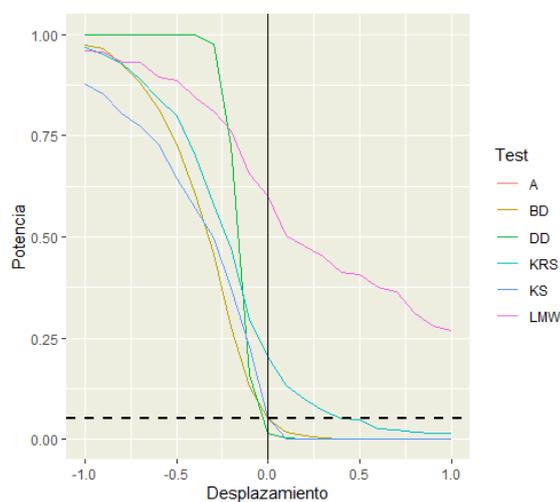


Tabla 5.8: Potencia para diferentes desplazamientos de los diferentes test bajo independencia con tamaño muestral  $m = 65$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	0.98	1.00	0.98	0.98	0.96
-0.90	1.00	0.97	1.00	0.97	0.96	0.96
-0.80	1.00	0.95	1.00	0.96	0.95	0.91
-0.70	1.00	0.93	1.00	0.94	0.93	0.88
-0.60	1.00	0.86	1.00	0.89	0.92	0.84
-0.50	1.00	0.80	1.00	0.84	0.89	0.79
-0.40	1.00	0.70	1.00	0.77	0.88	0.75
-0.30	1.00	0.53	1.00	0.67	0.82	0.62
-0.20	0.94	0.32	0.94	0.47	0.78	0.51
-0.10	0.36	0.14	0.36	0.34	0.70	0.32
0.00	0.02	0.04	0.02	0.20	0.60	0.08
0.10	0.00	0.01	0.00	0.15	0.56	0.00
0.20	0.00	0.00	0.00	0.11	0.52	0.00
0.30	0.00	0.00	0.00	0.07	0.48	0.00
0.40	0.00	0.00	0.00	0.08	0.47	0.00
0.50	0.00	0.00	0.00	0.06	0.44	0.00
0.60	0.00	0.00	0.00	0.02	0.38	0.00
0.70	0.00	0.00	0.00	0.02	0.34	0.00
0.80	0.00	0.00	0.00	0.02	0.34	0.00
0.90	0.00	0.00	0.00	0.01	0.26	0.00
1.00	0.00	0.00	0.00	0.01	0.25	0.00

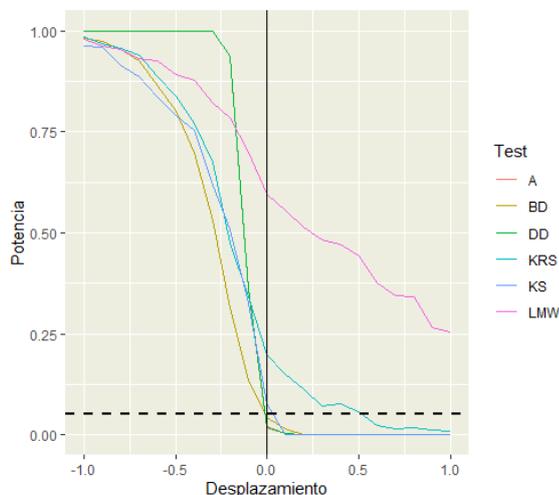


Tabla 5.9: Potencia para diferentes desplazamientos de los diferentes test bajo independencia con tamaño muestral  $m = 100$ .

### 5.3. Variables aleatorias comonótonas

Pese a que los tests de Anderson y Barrett-Donald suponen independencia, se consideran para esta situación de dependencia con el objetivo de analizar si es una condición muy relevante para la experimentación.

#### 5.3.1. Variables aleatorias con distribución normal

Viendo las Tablas 5.10, 5.11 y 5.12, se puede observar que:

1. En este caso, los tests: LMW y KS; no son capaces de mantener el nivel de significación bajo  $H_0$ .
2. Los tests: A, BD, DD y KRS; son muy conservadores, en este caso.
3. En este caso, el test KRS es el más potentes bajo  $H_1$ , de entre los que mantienen el nivel de significación bajo  $H_0$ .
4. Independientemente del tamaño muestral se pueden apreciar dichos comportamientos.

Luego, el mejor test para esta situación es el de Kaur-Rao-Singh, pues, pese a ser muy conservador, es el más potente bajo  $H_1$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	0.26	0.68	0.51	0.96	1.00	1.00
-0.90	0.17	0.62	0.40	0.92	1.00	1.00
-0.80	0.14	0.50	0.32	0.84	1.00	1.00
-0.70	0.06	0.37	0.18	0.76	1.00	1.00
-0.60	0.05	0.28	0.12	0.63	1.00	1.00
-0.50	0.02	0.19	0.06	0.49	1.00	1.00
-0.40	0.00	0.11	0.02	0.32	1.00	1.00
-0.30	0.00	0.05	0.01	0.17	1.00	1.00
-0.20	0.00	0.02	0.00	0.07	1.00	1.00
-0.10	0.00	0.00	0.00	0.01	1.00	1.00
0.00	0.00	0.00	0.00	0.00	1.00	1.00
0.10	0.00	0.00	0.00	0.00	1.00	0.00
0.20	0.00	0.00	0.00	0.00	0.99	0.00
0.30	0.00	0.00	0.00	0.00	0.88	0.00
0.40	0.00	0.00	0.00	0.00	0.61	0.00
0.50	0.00	0.00	0.00	0.00	0.40	0.00
0.60	0.00	0.00	0.00	0.00	0.24	0.00
0.70	0.00	0.00	0.00	0.00	0.13	0.00
0.80	0.00	0.00	0.00	0.00	0.08	0.00
0.90	0.00	0.00	0.00	0.00	0.04	0.00
1.00	0.00	0.00	0.00	0.00	0.02	0.00

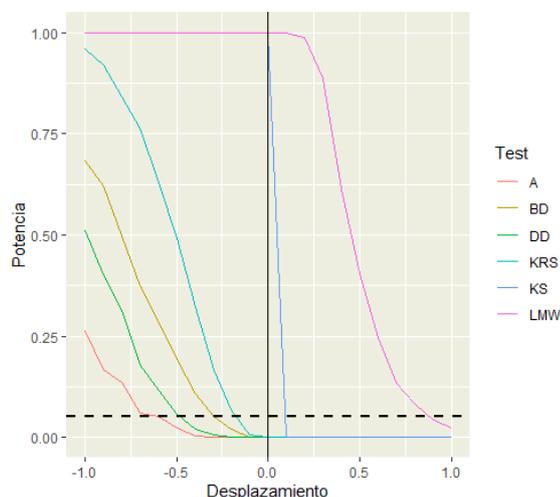


Tabla 5.10: Potencia para diferentes desplazamientos de los diferentes test bajo comonotonía con tamaño muestral  $m = 10$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.90	1.00	1.00	1.00	1.00	1.00	1.00
-0.80	1.00	1.00	1.00	1.00	1.00	1.00
-0.70	1.00	1.00	1.00	1.00	1.00	1.00
-0.60	0.95	0.99	0.95	1.00	1.00	1.00
-0.50	0.66	0.91	0.66	1.00	1.00	1.00
-0.40	0.22	0.67	0.22	0.97	1.00	1.00
-0.30	0.02	0.27	0.02	0.81	1.00	1.00
-0.20	0.00	0.03	0.00	0.30	1.00	1.00
-0.10	0.00	0.00	0.00	0.01	1.00	1.00
0.00	0.00	0.00	0.00	0.00	1.00	1.00
0.10	0.00	0.00	0.00	0.00	0.95	0.00
0.20	0.00	0.00	0.00	0.00	0.32	0.00
0.30	0.00	0.00	0.00	0.00	0.10	0.00
0.40	0.00	0.00	0.00	0.00	0.04	0.00
0.50	0.00	0.00	0.00	0.00	0.02	0.00
0.60	0.00	0.00	0.00	0.00	0.01	0.00
0.70	0.00	0.00	0.00	0.00	0.01	0.00
0.80	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00

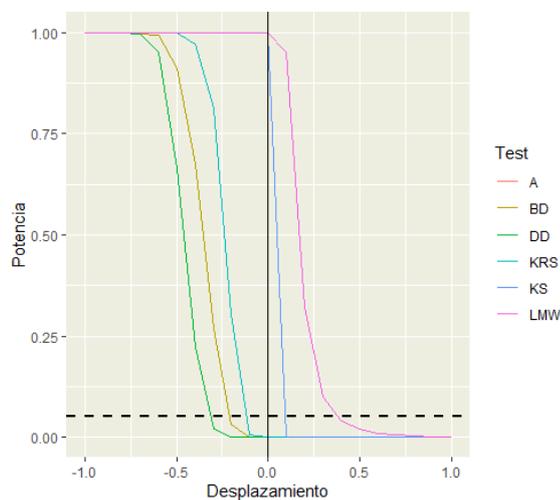


Tabla 5.11: Potencia para diferentes desplazamientos de los diferentes test bajo comonotonía con tamaño muestral  $m = 65$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.90	1.00	1.00	1.00	1.00	1.00	1.00
-0.80	1.00	1.00	1.00	1.00	1.00	1.00
-0.70	1.00	1.00	1.00	1.00	1.00	1.00
-0.60	1.00	1.00	1.00	1.00	1.00	1.00
-0.50	0.98	1.00	0.99	1.00	1.00	1.00
-0.40	0.65	0.92	0.74	1.00	1.00	1.00
-0.30	0.11	0.55	0.14	0.96	1.00	1.00
-0.20	0.00	0.07	0.01	0.55	1.00	1.00
-0.10	0.00	0.00	0.00	0.01	1.00	1.00
0.00	0.00	0.00	0.00	0.00	1.00	1.00
0.10	0.00	0.00	0.00	0.00	0.80	0.00
0.20	0.00	0.00	0.00	0.00	0.19	0.00
0.30	0.00	0.00	0.00	0.00	0.06	0.00
0.40	0.00	0.00	0.00	0.00	0.03	0.00
0.50	0.00	0.00	0.00	0.00	0.01	0.00
0.60	0.00	0.00	0.00	0.00	0.00	0.00
0.70	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00

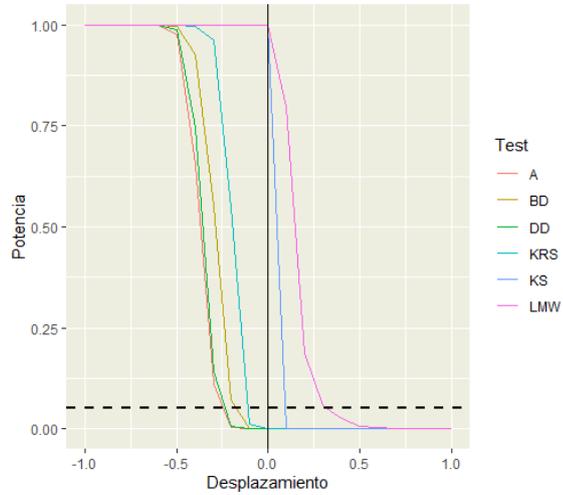


Tabla 5.12: Potencia para diferentes desplazamientos de los diferentes test bajo comonotonía con tamaño muestral  $m = 100$ .

### 5.3.2. Variables aleatorias con distribución uniforme

Viendo las Tablas 5.13, 5.14 y 5.15, se puede observar que:

1. En este caso, los tests: LMW y KS; no son capaces de mantener el nivel de significación bajo  $H_0$ .
2. Los tests: A, BD, DD y KRS; son muy conservadores, en este caso.
3. En este caso, el test KRS es el más potentes bajo  $H_1$ , de entre los que mantienen el nivel de significación bajo  $H_0$ .
4. De manera análoga al caso anterior, para los distintos tamaños muestrales no hay un cambio significativo en dichos comportamientos.

Luego, el mejor test para esta situación es el de Kaur-Rao-Singh por el mismo motivo que en el caso anterior: pese a ser muy conservador, es el más potente bajo  $H_1$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.90	1.00	1.00	1.00	1.00	1.00	1.00
-0.80	1.00	1.00	1.00	1.00	1.00	1.00
-0.70	1.00	1.00	1.00	1.00	1.00	1.00
-0.60	1.00	1.00	1.00	1.00	1.00	1.00
-0.50	0.93	0.99	0.96	1.00	1.00	1.00
-0.40	0.67	0.89	0.77	1.00	1.00	1.00
-0.30	0.29	0.64	0.41	0.97	1.00	1.00
-0.20	0.06	0.26	0.12	0.67	1.00	1.00
-0.10	0.00	0.04	0.01	0.17	1.00	1.00
0.00	0.00	0.00	0.00	0.00	1.00	1.00
0.10	0.00	0.00	0.00	0.00	0.81	0.00
0.20	0.00	0.00	0.00	0.00	0.05	0.00
0.30	0.00	0.00	0.00	0.00	0.00	0.00
0.40	0.00	0.00	0.00	0.00	0.00	0.00
0.50	0.00	0.00	0.00	0.00	0.00	0.00
0.60	0.00	0.00	0.00	0.00	0.00	0.00
0.70	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00

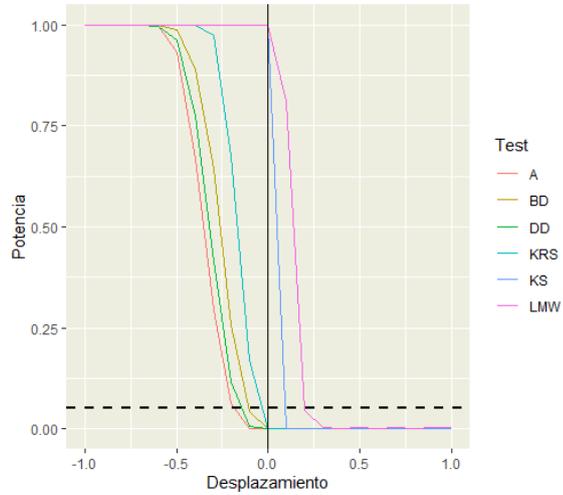


Tabla 5.13: Potencia para diferentes desplazamientos de los diferentes test bajo comonotonía con tamaño muestral  $m = 10$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.90	1.00	1.00	1.00	1.00	1.00	1.00
-0.80	1.00	1.00	1.00	1.00	1.00	1.00
-0.70	1.00	1.00	1.00	1.00	1.00	1.00
-0.60	1.00	1.00	1.00	1.00	1.00	1.00
-0.50	1.00	1.00	1.00	1.00	1.00	1.00
-0.40	1.00	1.00	1.00	1.00	1.00	1.00
-0.30	1.00	1.00	1.00	1.00	1.00	1.00
-0.20	1.00	0.99	1.00	1.00	1.00	1.00
-0.10	0.22	0.19	0.22	0.85	1.00	1.00
0.00	0.00	0.00	0.00	0.00	1.00	1.00
0.10	0.00	0.00	0.00	0.00	0.00	0.00
0.20	0.00	0.00	0.00	0.00	0.00	0.00
0.30	0.00	0.00	0.00	0.00	0.00	0.00
0.40	0.00	0.00	0.00	0.00	0.00	0.00
0.50	0.00	0.00	0.00	0.00	0.00	0.00
0.60	0.00	0.00	0.00	0.00	0.00	0.00
0.70	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00

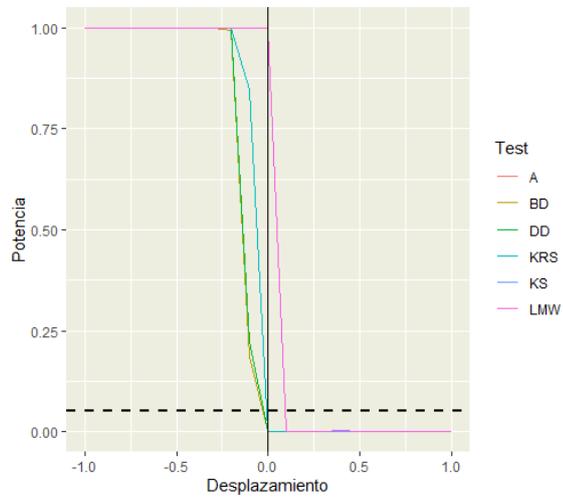


Tabla 5.14: Potencia para diferentes desplazamientos de los diferentes test bajo comonotonía con tamaño muestral  $m = 65$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.90	1.00	1.00	1.00	1.00	1.00	1.00
-0.80	1.00	1.00	1.00	1.00	1.00	1.00
-0.70	1.00	1.00	1.00	1.00	1.00	1.00
-0.60	1.00	1.00	1.00	1.00	1.00	1.00
-0.50	1.00	1.00	1.00	1.00	1.00	1.00
-0.40	1.00	1.00	1.00	1.00	1.00	1.00
-0.30	1.00	1.00	1.00	1.00	1.00	1.00
-0.20	1.00	1.00	1.00	1.00	1.00	1.00
-0.10	0.71	0.44	0.72	0.96	1.00	1.00
0.00	0.00	0.00	0.00	0.00	1.00	1.00
0.10	0.00	0.00	0.00	0.00	0.00	0.00
0.20	0.00	0.00	0.00	0.00	0.00	0.00
0.30	0.00	0.00	0.00	0.00	0.00	0.00
0.40	0.00	0.00	0.00	0.00	0.00	0.00
0.50	0.00	0.00	0.00	0.00	0.00	0.00
0.60	0.00	0.00	0.00	0.00	0.00	0.00
0.70	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00

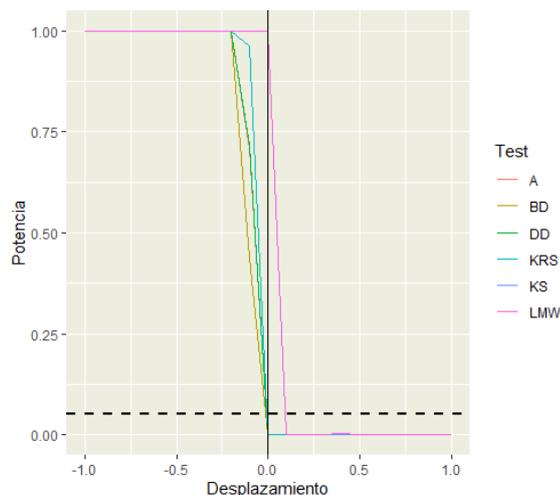


Tabla 5.15: Potencia para diferentes desplazamientos de los diferentes test bajo comonotonía con tamaño muestral  $m = 100$ .

### 5.3.3. Variables aleatorias con distribución exponencial

Viendo las Tablas 5.16, 5.17 y 5.18, se puede observar que:

1. En este caso, los tests: LMW y KS; no son capaces de mantener el nivel de significación bajo  $H_0$ .
2. Para  $m \in \{65, 100\}$ , los tests: A, BD, DD y KRS; son muy conservadores, en este caso.
3. En este caso, los tests: A y DD; son los más potentes bajo  $H_1$  entre los que mantienen el nivel de significación bajo  $H_0$ .
4. Para  $m = 10$ , KRS es el test más potente bajo  $H_1$  entre los que mantienen el nivel de significación bajo  $H_0$ .

Por tanto, los mejores tests para esta situación son: para tamaños muestrales grandes, el de Anderson y el Davidson-Duclos, ya que mantienen el nivel de significación bajo  $H_0$  (siendo muy conservadores) y, entre los que lo mantienen, son los más potente bajo  $H_1$ , y para tamaños muestrales pequeños el de Kaur-Rao-Singh por el mismo motivo.

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	0.82	0.80	0.86	0.89	1.00	1.00
-0.90	0.73	0.73	0.79	0.84	1.00	1.00
-0.80	0.64	0.66	0.70	0.79	1.00	1.00
-0.70	0.50	0.56	0.56	0.73	1.00	1.00
-0.60	0.38	0.41	0.44	0.62	1.00	1.00
-0.50	0.25	0.30	0.30	0.52	1.00	1.00
-0.40	0.14	0.17	0.17	0.39	1.00	1.00
-0.30	0.07	0.09	0.09	0.24	1.00	1.00
-0.20	0.02	0.02	0.02	0.08	1.00	1.00
-0.10	0.00	0.00	0.00	0.01	1.00	1.00
0.00	0.00	0.00	0.00	0.00	1.00	1.00
0.10	0.00	0.00	0.00	0.00	1.00	0.00
0.20	0.00	0.00	0.00	0.00	0.92	0.00
0.30	0.00	0.00	0.00	0.00	0.80	0.00
0.40	0.00	0.00	0.00	0.00	0.64	0.00
0.50	0.00	0.00	0.00	0.00	0.51	0.00
0.60	0.00	0.00	0.00	0.00	0.42	0.00
0.70	0.00	0.00	0.00	0.00	0.32	0.00
0.80	0.00	0.00	0.00	0.00	0.27	0.00
0.90	0.00	0.00	0.00	0.00	0.21	0.00
1.00	0.00	0.00	0.00	0.00	0.18	0.00

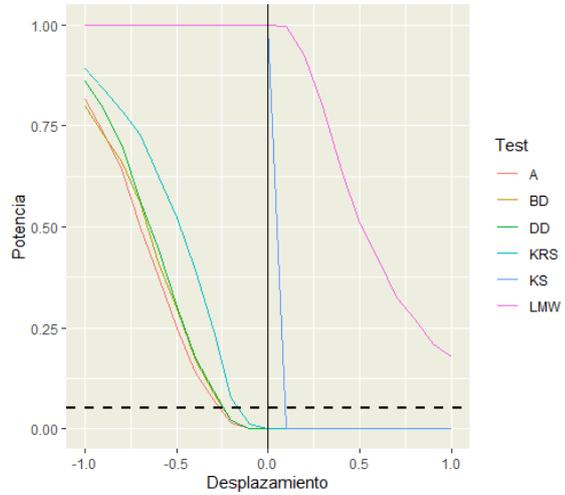


Tabla 5.16: Potencia para diferentes desplazamientos de los diferentes test bajo comonotonía con tamaño muestral  $m = 10$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	0.99	1.00	0.99	1.00	1.00
-0.90	1.00	0.98	1.00	0.98	1.00	1.00
-0.80	1.00	0.97	1.00	0.97	1.00	1.00
-0.70	1.00	0.94	1.00	0.95	1.00	1.00
-0.60	1.00	0.92	1.00	0.93	1.00	1.00
-0.50	1.00	0.86	1.00	0.87	1.00	1.00
-0.40	1.00	0.75	1.00	0.76	1.00	1.00
-0.30	0.99	0.56	0.99	0.60	1.00	1.00
-0.20	0.78	0.25	0.78	0.28	1.00	1.00
-0.10	0.08	0.01	0.08	0.03	1.00	1.00
0.00	0.00	0.00	0.00	0.00	1.00	1.00
0.10	0.00	0.00	0.00	0.00	0.98	0.00
0.20	0.00	0.00	0.00	0.00	0.78	0.00
0.30	0.00	0.00	0.00	0.00	0.56	0.00
0.40	0.00	0.00	0.00	0.00	0.44	0.00
0.50	0.00	0.00	0.00	0.00	0.30	0.00
0.60	0.00	0.00	0.00	0.00	0.27	0.00
0.70	0.00	0.00	0.00	0.00	0.22	0.00
0.80	0.00	0.00	0.00	0.00	0.16	0.00
0.90	0.00	0.00	0.00	0.00	0.14	0.00
1.00	0.00	0.00	0.00	0.00	0.12	0.00

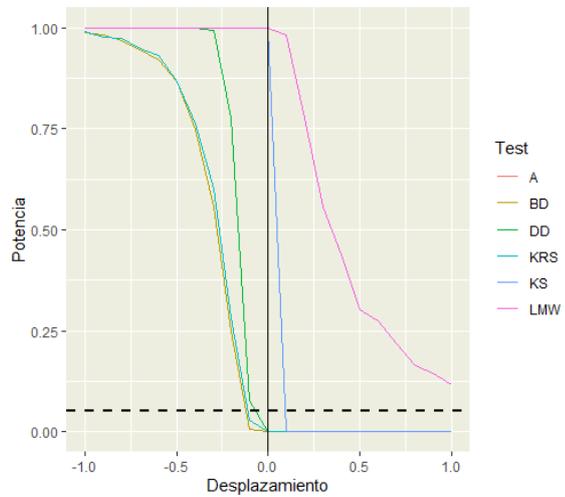


Tabla 5.17: Potencia para diferentes desplazamientos de los diferentes test bajo comonotonía con tamaño muestral  $m = 65$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	0.99	1.00	1.00
-0.90	1.00	0.99	1.00	0.99	1.00	1.00
-0.80	1.00	0.99	1.00	0.99	1.00	1.00
-0.70	1.00	0.98	1.00	0.97	1.00	1.00
-0.60	1.00	0.96	1.00	0.95	1.00	1.00
-0.50	1.00	0.95	1.00	0.91	1.00	1.00
-0.40	1.00	0.85	1.00	0.84	1.00	1.00
-0.30	1.00	0.74	1.00	0.69	1.00	1.00
-0.20	0.98	0.46	0.98	0.43	1.00	1.00
-0.10	0.29	0.03	0.29	0.09	1.00	1.00
0.00	0.00	0.00	0.00	0.00	1.00	1.00
0.10	0.00	0.00	0.00	0.00	0.97	0.00
0.20	0.00	0.00	0.00	0.00	0.70	0.00
0.30	0.00	0.00	0.00	0.00	0.53	0.00
0.40	0.00	0.00	0.00	0.00	0.38	0.00
0.50	0.00	0.00	0.00	0.00	0.30	0.00
0.60	0.00	0.00	0.00	0.00	0.23	0.00
0.70	0.00	0.00	0.00	0.00	0.18	0.00
0.80	0.00	0.00	0.00	0.00	0.14	0.00
0.90	0.00	0.00	0.00	0.00	0.12	0.00
1.00	0.00	0.00	0.00	0.00	0.09	0.00

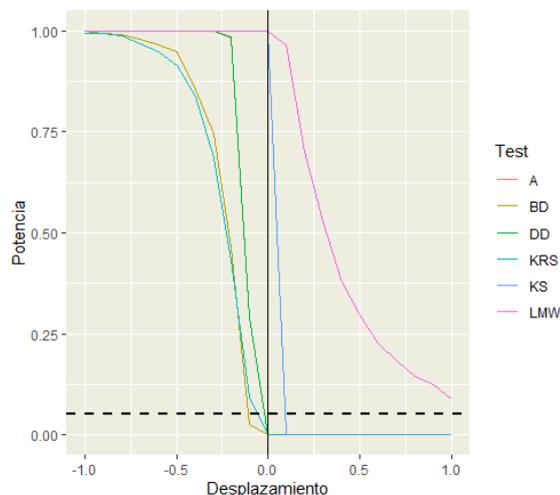


Tabla 5.18: Potencia para diferentes desplazamientos de los diferentes test bajo comonotonía con tamaño muestral  $m = 100$ .

## 5.4. Variables aleatorias contramonótonas

Análogamente a la situación de dependencia anterior, se consideran los tests de Anderson y Barrett-Donald, a pesar de suponer independencia en sus desarrollos, con el fin de estudiar si es una condición muy relevante en la práctica.

### 5.4.1. Variables aleatorias con distribución normal

Viendo las Tablas 5.19, 5.20 y 5.21, se puede observar que:

1. En este caso, los tests: KRS, LMW, KS, y BD por poco; no son capaces de mantener el nivel de significación bajo  $H_0$ .
2. El test A es bastante conservador. También lo es DD si  $m \in \{65, 100\}$ .
3. En este caso, el test DD es el más potente bajo  $H_1$  entre los que mantienen el nivel de significación bajo  $H_0$ .
4. Lo más relevante en este aspecto es que si  $m = 10$ , DD no mantiene el nivel de significación bajo  $H_0$  por poco, pero, de nuevo, esto tiene sentido por el escaso tamaño muestral.

Por tanto, el mejor test para esta situación es el de Davidson-Duclos.

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	0.26	0.50	0.42	0.73	0.98	0.78
-0.90	0.21	0.47	0.38	0.68	0.98	0.75
-0.80	0.17	0.40	0.32	0.64	0.97	0.68
-0.70	0.14	0.36	0.27	0.59	0.95	0.64
-0.60	0.12	0.32	0.25	0.53	0.95	0.58
-0.50	0.07	0.26	0.18	0.47	0.92	0.50
-0.40	0.06	0.22	0.15	0.40	0.91	0.46
-0.30	0.06	0.20	0.14	0.40	0.87	0.41
-0.20	0.03	0.15	0.09	0.32	0.86	0.35
-0.10	0.02	0.11	0.07	0.27	0.81	0.29
0.00	0.02	0.09	0.07	0.22	0.77	0.24
0.10	0.01	0.07	0.05	0.20	0.69	0.20
0.20	0.02	0.06	0.04	0.15	0.67	0.16
0.30	0.01	0.04	0.03	0.12	0.63	0.14
0.40	0.01	0.04	0.02	0.10	0.58	0.12
0.50	0.00	0.02	0.01	0.06	0.51	0.08
0.60	0.00	0.02	0.01	0.06	0.42	0.04
0.70	0.00	0.02	0.01	0.05	0.38	0.05
0.80	0.00	0.01	0.01	0.02	0.33	0.04
0.90	0.00	0.01	0.01	0.03	0.30	0.02
1.00	0.00	0.00	0.00	0.01	0.23	0.01

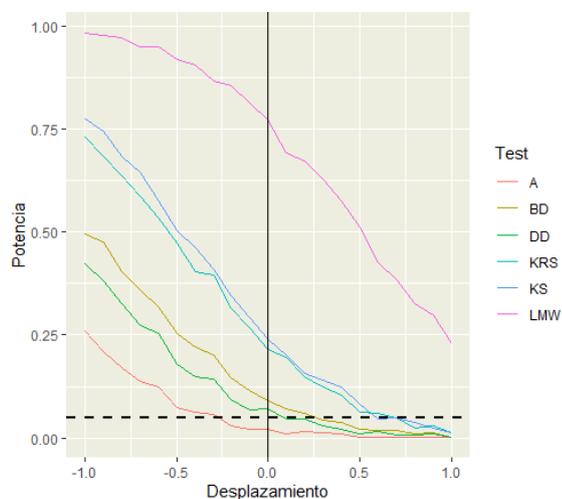


Tabla 5.19: Potencia para diferentes desplazamientos de los diferentes test bajo contramonotonía con tamaño muestral  $m = 10$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	0.26	0.50	0.42	0.73	0.98	0.78
-0.90	0.21	0.47	0.38	0.68	0.98	0.75
-0.80	0.17	0.40	0.32	0.64	0.97	0.68
-0.70	0.14	0.36	0.27	0.59	0.95	0.64
-0.60	0.12	0.32	0.25	0.53	0.95	0.58
-0.50	0.07	0.26	0.18	0.47	0.92	0.50
-0.40	0.06	0.22	0.15	0.40	0.91	0.46
-0.30	0.06	0.20	0.14	0.40	0.87	0.41
-0.20	0.03	0.15	0.09	0.32	0.86	0.35
-0.10	0.02	0.11	0.07	0.27	0.81	0.29
0.00	0.02	0.09	0.07	0.22	0.77	0.24
0.10	0.01	0.07	0.05	0.20	0.69	0.20
0.20	0.02	0.06	0.04	0.15	0.67	0.16
0.30	0.01	0.04	0.03	0.12	0.63	0.14
0.40	0.01	0.04	0.02	0.10	0.58	0.12
0.50	0.00	0.02	0.01	0.06	0.51	0.08
0.60	0.00	0.02	0.01	0.06	0.42	0.04
0.70	0.00	0.02	0.01	0.05	0.38	0.05
0.80	0.00	0.01	0.01	0.02	0.33	0.04
0.90	0.00	0.01	0.01	0.03	0.30	0.02
1.00	0.00	0.00	0.00	0.01	0.23	0.01

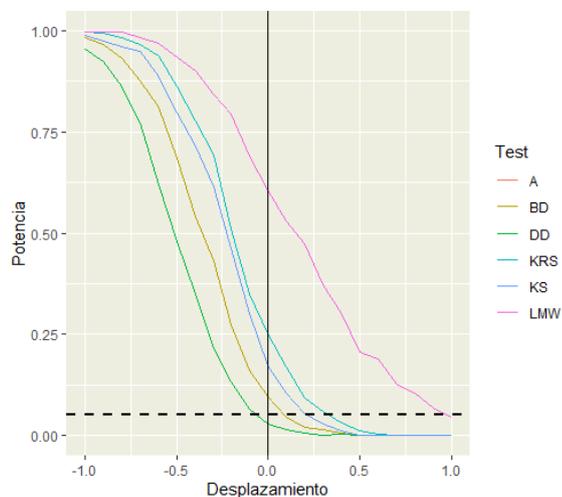


Tabla 5.20: Potencia para diferentes desplazamientos de los diferentes test bajo contramonotonía con tamaño muestral  $m = 65$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	0.26	0.50	0.42	0.73	0.98	0.78
-0.90	0.21	0.47	0.38	0.68	0.98	0.75
-0.80	0.17	0.40	0.32	0.64	0.97	0.68
-0.70	0.14	0.36	0.27	0.59	0.95	0.64
-0.60	0.12	0.32	0.25	0.53	0.95	0.58
-0.50	0.07	0.26	0.18	0.47	0.92	0.50
-0.40	0.06	0.22	0.15	0.40	0.91	0.46
-0.30	0.06	0.20	0.14	0.40	0.87	0.41
-0.20	0.03	0.15	0.09	0.32	0.86	0.35
-0.10	0.02	0.11	0.07	0.27	0.81	0.29
0.00	0.02	0.09	0.07	0.22	0.77	0.24
0.10	0.01	0.07	0.05	0.20	0.69	0.20
0.20	0.02	0.06	0.04	0.15	0.67	0.16
0.30	0.01	0.04	0.03	0.12	0.63	0.14
0.40	0.01	0.04	0.02	0.10	0.58	0.12
0.50	0.00	0.02	0.01	0.06	0.51	0.08
0.60	0.00	0.02	0.01	0.06	0.42	0.04
0.70	0.00	0.02	0.01	0.05	0.38	0.05
0.80	0.00	0.01	0.01	0.02	0.33	0.04
0.90	0.00	0.01	0.01	0.03	0.30	0.02
1.00	0.00	0.00	0.00	0.01	0.23	0.01

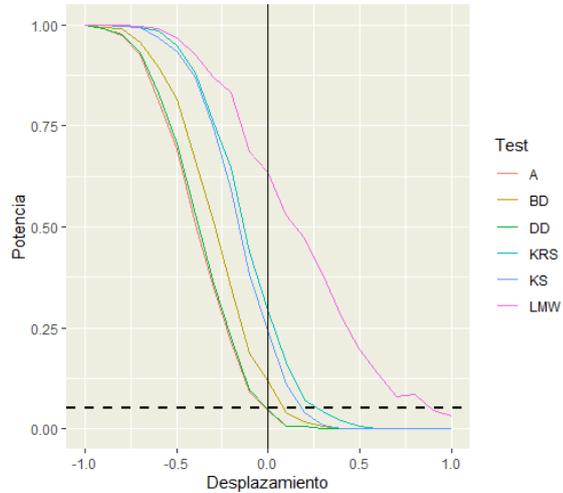


Tabla 5.21: Potencia para diferentes desplazamientos de los diferentes test bajo contramonotonía con tamaño muestral  $m = 100$ .

#### 5.4.2. Variables aleatorias con distribución uniforme

Viendo las Tablas 5.22, 5.23 y 5.24, se puede observar que:

1. En este caso, los tests: KRS, LMW, KS, y BD por poco; no son capaces de mantener el nivel de significación bajo  $H_0$ .
2. El test A es algo conservador, en este caso.
3. En este caso, el test DD es el más potente bajo  $H_1$  entre los que mantienen el nivel de significación bajo  $H_0$ .
4. El único cambio significativo en dichos comportamientos para los distintos tamaños muestrales es que si  $m = 10$ , BD y DD no mantienen el nivel de significación bajo  $H_0$  por poco, pero esto tiene sentido por el escaso tamaño muestral, como ya se ha visto en anteriores ejemplos.

Por tanto, el mejor test para esta situación es el de Davidson-Duclos.

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.90	1.00	1.00	1.00	1.00	1.00	1.00
-0.80	0.98	0.99	0.99	1.00	1.00	1.00
-0.70	0.92	0.98	0.94	1.00	1.00	1.00
-0.60	0.79	0.93	0.85	0.99	1.00	0.99
-0.50	0.56	0.78	0.66	0.94	1.00	0.96
-0.40	0.38	0.62	0.50	0.87	1.00	0.91
-0.30	0.21	0.45	0.34	0.77	0.99	0.81
-0.20	0.11	0.31	0.22	0.59	0.97	0.65
-0.10	0.05	0.17	0.12	0.39	0.89	0.47
0.00	0.02	0.10	0.07	0.22	0.76	0.24
0.10	0.01	0.05	0.03	0.13	0.52	0.08
0.20	0.00	0.03	0.01	0.08	0.35	0.03
0.30	0.00	0.02	0.01	0.03	0.22	0.01
0.40	0.00	0.00	0.00	0.01	0.10	0.01
0.50	0.00	0.00	0.00	0.00	0.03	0.00
0.60	0.00	0.00	0.00	0.00	0.01	0.00
0.70	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00	0.00	0.01	0.00

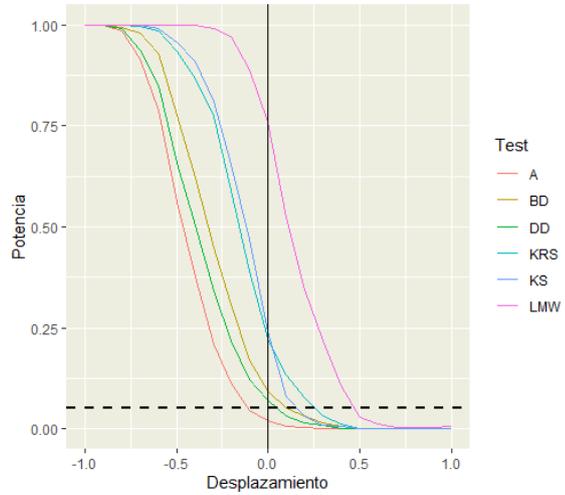


Tabla 5.22: Potencia para diferentes desplazamientos de los diferentes test bajo contramonotonía con tamaño muestral  $m = 10$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.90	1.00	1.00	1.00	1.00	1.00	1.00
-0.80	1.00	1.00	1.00	1.00	1.00	1.00
-0.70	1.00	1.00	1.00	1.00	1.00	1.00
-0.60	1.00	1.00	1.00	1.00	1.00	1.00
-0.50	1.00	1.00	1.00	1.00	1.00	1.00
-0.40	1.00	1.00	1.00	1.00	1.00	1.00
-0.30	1.00	1.00	1.00	1.00	1.00	1.00
-0.20	0.88	0.83	0.88	0.99	1.00	0.94
-0.10	0.28	0.36	0.28	0.79	0.99	0.71
0.00	0.03	0.10	0.03	0.22	0.71	0.17
0.10	0.00	0.01	0.00	0.03	0.26	0.00
0.20	0.00	0.00	0.00	0.00	0.05	0.00
0.30	0.00	0.00	0.00	0.00	0.01	0.00
0.40	0.00	0.00	0.00	0.00	0.00	0.00
0.50	0.00	0.00	0.00	0.00	0.00	0.00
0.60	0.00	0.00	0.00	0.00	0.00	0.00
0.70	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00

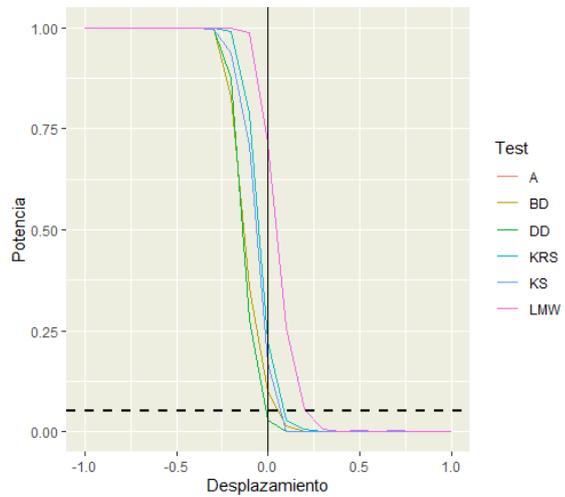


Tabla 5.23: Potencia para diferentes desplazamientos de los diferentes test bajo contramonotonía con tamaño muestral  $m = 65$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	1.00	1.00	1.00	1.00	1.00
-0.90	1.00	1.00	1.00	1.00	1.00	1.00
-0.80	1.00	1.00	1.00	1.00	1.00	1.00
-0.70	1.00	1.00	1.00	1.00	1.00	1.00
-0.60	1.00	1.00	1.00	1.00	1.00	1.00
-0.50	1.00	1.00	1.00	1.00	1.00	1.00
-0.40	1.00	1.00	1.00	1.00	1.00	1.00
-0.30	1.00	1.00	1.00	1.00	1.00	1.00
-0.20	0.98	0.96	0.98	1.00	1.00	0.98
-0.10	0.50	0.47	0.51	0.88	0.99	0.81
0.00	0.04	0.11	0.05	0.26	0.74	0.24
0.10	0.00	0.01	0.00	0.02	0.22	0.00
0.20	0.00	0.00	0.00	0.00	0.03	0.00
0.30	0.00	0.00	0.00	0.00	0.00	0.00
0.40	0.00	0.00	0.00	0.00	0.00	0.00
0.50	0.00	0.00	0.00	0.00	0.00	0.00
0.60	0.00	0.00	0.00	0.00	0.00	0.00
0.70	0.00	0.00	0.00	0.00	0.00	0.00
0.80	0.00	0.00	0.00	0.00	0.00	0.00
0.90	0.00	0.00	0.00	0.00	0.00	0.00
1.00	0.00	0.00	0.00	0.00	0.00	0.00

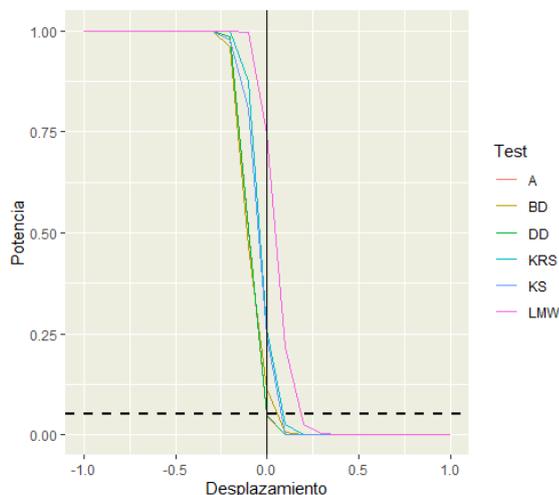


Tabla 5.24: Potencia para diferentes desplazamientos de los diferentes test bajo contramonotonía con tamaño muestral  $m = 100$ .

### 5.4.3. Variables aleatorias con distribución exponencial

Viendo las Tablas 5.25, 5.26 y 5.27, se se puede observar que:

1. En este caso, los tests: KRS, LMW, KS, y BD por poco; no son capaces de mantener el nivel de significación bajo  $H_0$ .
2. El test A es algo conservador, en este caso.
3. En este caso, el test DD es el más potente bajo  $H_1$  entre los que mantienen el nivel de significación bajo  $H_0$ .
4. El comportamiento para distintos tamaños muestrales es idéntico al caso anterior: no hay un cambio significativo en dichos comportamientos, a excepción de que si  $m = 10$ , BD y DD no mantienen el nivel de significación bajo  $H_0$  por poco, pero ya se comentó que esto tiene sentido por el escaso tamaño muestral.

Por tanto, el mejor test para esta situación es el de Davidson-Duclos.

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	0.76	0.76	0.77	0.83	0.99	0.63
-0.90	0.67	0.68	0.70	0.77	0.99	0.61
-0.80	0.59	0.60	0.63	0.73	0.99	0.54
-0.70	0.50	0.51	0.55	0.67	0.98	0.55
-0.60	0.39	0.44	0.45	0.59	0.97	0.49
-0.50	0.27	0.32	0.34	0.53	0.94	0.44
-0.40	0.19	0.24	0.26	0.40	0.92	0.40
-0.30	0.13	0.20	0.20	0.36	0.90	0.38
-0.20	0.06	0.14	0.13	0.28	0.86	0.33
-0.10	0.03	0.10	0.08	0.23	0.82	0.29
0.00	0.02	0.09	0.07	0.18	0.78	0.24
0.10	0.01	0.07	0.04	0.18	0.70	0.10
0.20	0.02	0.06	0.04	0.13	0.68	0.04
0.30	0.01	0.04	0.02	0.11	0.65	0.02
0.40	0.01	0.03	0.02	0.09	0.62	0.01
0.50	0.00	0.02	0.00	0.07	0.57	0.00
0.60	0.00	0.02	0.01	0.05	0.49	0.00
0.70	0.00	0.01	0.00	0.05	0.45	0.00
0.80	0.00	0.01	0.01	0.02	0.43	0.00
0.90	0.00	0.01	0.00	0.03	0.43	0.00
1.00	0.00	0.00	0.00	0.02	0.36	0.00

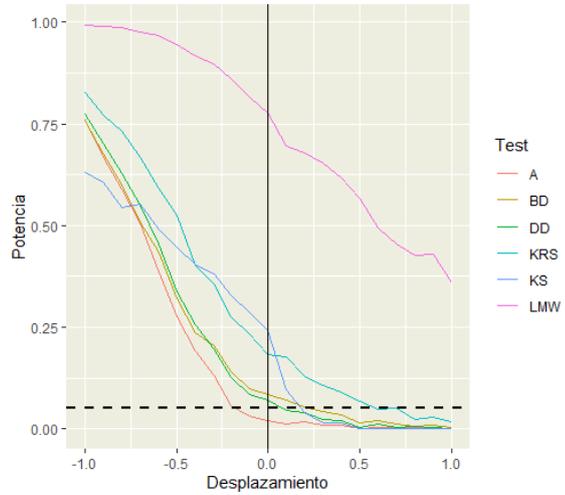


Tabla 5.25: Potencia para diferentes desplazamientos de los diferentes test bajo contramonotonía con tamaño muestral  $m = 10$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	0.96	1.00	0.96	0.97	0.86
-0.90	1.00	0.95	1.00	0.95	0.95	0.84
-0.80	1.00	0.92	1.00	0.92	0.94	0.81
-0.70	1.00	0.85	1.00	0.88	0.90	0.75
-0.60	1.00	0.80	1.00	0.84	0.90	0.69
-0.50	1.00	0.70	1.00	0.74	0.86	0.62
-0.40	1.00	0.56	1.00	0.67	0.84	0.56
-0.30	0.98	0.46	0.98	0.58	0.81	0.49
-0.20	0.72	0.30	0.72	0.43	0.75	0.37
-0.10	0.19	0.14	0.19	0.29	0.65	0.28
0.00	0.03	0.09	0.03	0.23	0.57	0.17
0.10	0.01	0.04	0.01	0.17	0.53	0.00
0.20	0.01	0.02	0.01	0.12	0.51	0.00
0.30	0.00	0.01	0.00	0.10	0.46	0.00
0.40	0.00	0.00	0.00	0.08	0.44	0.00
0.50	0.00	0.00	0.00	0.03	0.39	0.00
0.60	0.00	0.00	0.00	0.03	0.38	0.00
0.70	0.00	0.00	0.00	0.03	0.35	0.00
0.80	0.00	0.00	0.00	0.02	0.32	0.00
0.90	0.00	0.00	0.00	0.02	0.32	0.00
1.00	0.00	0.00	0.00	0.01	0.25	0.00

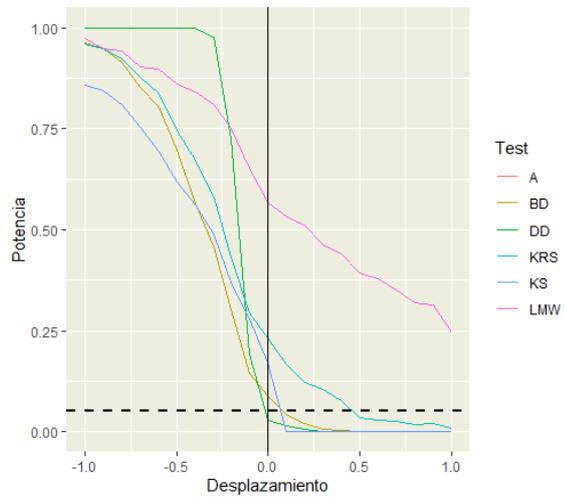


Tabla 5.26: Potencia para diferentes desplazamientos de los diferentes test bajo contramonotonía con tamaño muestral  $m = 65$ .

Desplazamiento	A	BD	DD	KRS	LMW	KS
-1.00	1.00	0.98	1.00	0.98	0.97	0.94
-0.90	1.00	0.97	1.00	0.97	0.96	0.93
-0.80	1.00	0.94	1.00	0.95	0.96	0.91
-0.70	1.00	0.89	1.00	0.91	0.94	0.88
-0.60	1.00	0.84	1.00	0.85	0.91	0.81
-0.50	1.00	0.76	1.00	0.81	0.90	0.76
-0.40	1.00	0.64	1.00	0.73	0.85	0.70
-0.30	1.00	0.53	1.00	0.63	0.80	0.61
-0.20	0.94	0.37	0.94	0.51	0.79	0.52
-0.10	0.38	0.18	0.39	0.35	0.68	0.35
0.00	0.04	0.09	0.05	0.25	0.63	0.24
0.10	0.01	0.04	0.01	0.17	0.56	0.00
0.20	0.00	0.01	0.00	0.11	0.52	0.00
0.30	0.00	0.00	0.00	0.09	0.49	0.00
0.40	0.00	0.00	0.00	0.06	0.46	0.00
0.50	0.00	0.00	0.00	0.04	0.40	0.00
0.60	0.00	0.00	0.00	0.03	0.38	0.00
0.70	0.00	0.00	0.00	0.02	0.32	0.00
0.80	0.00	0.00	0.00	0.02	0.33	0.00
0.90	0.00	0.00	0.00	0.02	0.30	0.00
1.00	0.00	0.00	0.00	0.01	0.26	0.00

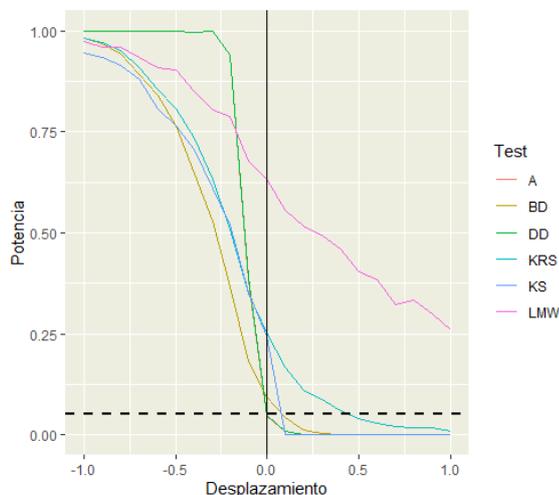


Tabla 5.27: Potencia para diferentes desplazamientos de los diferentes test bajo contramonotonía con tamaño muestral  $m = 100$ .

## 5.5. Síntesis

Finalmente, podemos concluir que:

- Bajo condiciones de **independencia entre variables aleatorias**, los tests que mejor funcionan son el de Kolmogorov-Smirnov, el de Anderson y el de Davidson-Duclos.
- Bajo condiciones de **comonotonía entre variables aleatorias**, los tests que mejor funcionan son el de Kaur-Rao-Singh, y el de Davidson-Duclos.
- Bajo condiciones de **contramonotonía entre variables aleatorias**, el test que mejor funciona es el de Davidson-Duclos.

Además, para variables aleatorias con distribución exponencial, el test de Davidson-Duclos funciona especialmente bien en cualquier caso de dependencia entre las variables aleatorias.

Resaltar que dichos comportamientos se verifican en general, independientemente del tamaño muestral. Por supuesto, a mayor tamaño muestral, más claro será que dichos tests se desempeñan mejor que el resto para cada caso en particular.

También es destacable el buen desempeño del test de Anderson en situaciones de dependencia entre variables aleatorias, en contraposición a lo que ocurre con el de Barrett-Donald, tal y como cabría esperar en un principio por asumir independencia entre las variables. Sin embargo, en estos casos, su comportamiento es algo peor que el de Davidson-Duclos, por ello, dicho test es mejor que ambos en el supuesto de dependencia entre variables al alcanzar potencias más altas.

Por último, destacar el pobre desempeño del test de Linton-Maasoumi-Whang y lo costoso computacionalmente que es el test de Barret-Donald, siendo este el test que más tiempo tardaba en ejecutarse con bastante diferencia. Señalar que el rendimiento del test de Linton-Maasoumi-Whang puede ser debido a la modificación hecha para contrastar las Hipótesis mostradas en la Ecuación (4.1).

## Capítulo 6

# Conclusiones

En resumen, se ha hecho una revisión de la literatura de los tests más usuales para contrastar la dominancia estocástica entre dos distribuciones. Concretamente, los tests seleccionados fueron el test de Anderson, el test de Davidson-Duclos, el test de Barrett-Donald, el test de Linton-Maasoumi-Whang, el test de Kaur-Rao-Singh y el test de Kolmogorov-Smirnov. Tras desarrollar estos seis tests teóricamente, se han implementado en R los cinco primeros, habida cuenta de que el último de ellos ya se encuentra implementado en el paquete base de R. Nótese que el test de Linton-Maasoumi-Whang se ha adaptado para el contraste considerado. Posteriormente, se ha comparado empíricamente la potencia de todos los tests presentados.

En base a la experimentación realizada, se recomienda utilizar el test Davidson-Duclos, pese a que los tests de Kolmogorov-Smirnov y de Anderson funcionan bien si hay independencia entre variables aleatorias, y el de Kaur-Rao-Singh si las variables aleatorias involucradas son contramonótonas.

Futuras líneas de investigación conciernen la realización de una experimentación más detallada para analizar:

- la potencia de los tests en otras familias de distribuciones, no necesariamente restringiendo la experimentación dentro de familias de localización-escala;
- la potencia de los tests en otras situaciones de dependencia entre variables aleatorias;
- el rol de la asimetría de la variable aleatoria en la potencia de los tests.

También sería interesante extender la presente experimentación en el contexto de otros órdenes estocásticos estudiados en la literatura [15], como por ejemplo:

- Dominancia estocástica de segundo orden.
- Orden monótono convexo.
- Orden de cociente de riesgo.
- Orden de razón de verosimilitudes.

# Bibliografía

- [1] G. Anderson, Nonparametric tests of stochastic dominance in income distributions, *Econometrica*, Vol. 64 (1996), 1183–1193.
- [2] G.F. Barrett, S.G. Donald, Consistent tests for stochastic dominance, *Econometrica*, Vol. 71 (2003), 71—104.
- [3] J.A. Bishop, J.P. Formby, P.D. Thistle, Convergence of the South and non-South income distributions, *American Economic Review*, Vol. 82 (1992), 262—272.
- [4] R. Davidson, J.-Y. Duclos, Statistical inference for stochastic dominance and for the measurement of poverty and inequality, *Econometrica*, Vol. 68 (2000), 1435–1464.
- [5] M. D. Donsker, An invariance principle for certain probability limit theorems, *Memoirs of the American Mathematical Society*, Vol. 6 (1951), 1–12.
- [6] Y. Hochberg, Some generalizations of the T-method in simultaneous inference, *Journal of Multivariate Analysis*, Vol. 4 (1974), 224–234.
- [7] A. Kaur, B.L.S.P. Rao, H. Singh, Testing for second-order stochastic dominance of two distributions, *Econometric Theory*, Vol. 10 (1994), 849–866.
- [8] H.-H. Lean, W.-K. Wong, X. Zhang, The sizes and powers of some stochastic dominance tests: A Monte Carlo study for correlated and heteroskedastic distributions, *Mathematics and Computers in Simulation*, Vol. 79 (2008), 30–48.
- [9] K. Lee, Y.-J. Whang, PySDTest: a Python package for stochastic dominance tests, arXiv:2307.10694 [econ.EM](2023).

- [10] H. Levy, *Stochastic Dominance (Third Edition)*, Springer International Publishing Switzerland, Cham, 2016.
- [11] O. Linton, E. Maasoumi, Y.-J. Whang, Consistent testing for stochastic dominance under general sampling schemes, *Review of Economic Studies*, Vol. 72 (2005), 735–765.
- [12] K.V. Mardia, J.T. Kent, J.M. Bibby, *Multivariate Analysis*, Academic Press, San Diego, 1979.
- [13] F. McFadden, Testing for stochastic dominance, en: T.B. Fomby, T.K. Seo (Eds.), *Studies in the Economics of Uncertainty*, Springer-Verlag, New York, 1989, 113—132.
- [14] R.B. Nelsen, *An Introduction to Copulas (Second Edition)*, Springer Science+Business Media, New York, 2006.
- [15] M. Shaked, J. G. Shanthikumar, *Stochastic Orders*, Springer Science+Business Media, New York, 2007.
- [16] Z. Sidak, Rectangular confidence regions for the means of multivariate normal distributions, *Journal of the American Statistical Association*, Vol. 62 (1967), 626–633.
- [17] H. Sklar, Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Université de Paris*, Vol. 8 (1959), 229–231.
- [18] M.R. Stoline, H.K. Ury, Tables of the Studentized Maximum Modulus distribution and an application to multiple comparisons among means, *Technometrics*, Vol. 21 (1979), 87–93.
- [19] Y.K. Tse, X. Zhang, A Monte Carlo investigation of some tests for stochastic dominance, *Journal of Statistical Computation Simulation*, Vol. 74 (2004), 361—378.
- [20] B. Zheng, J.P. Formby, W.J. Smith, V.K. Chow, Inequality orderings, normalized stochastic dominance, and statistical inference, *Journal of Business & Economic Statistics*, Vol. 18 (2000), 479–488.