NOTAS SOBRE ÓRDENES ESTOCÁSTICOS

DANIEL FELIPE LOAIZA CORREA

INFORME DE PRÁCTICA INVESTIGATIVA I

ASESOR: DR. GERARDO ARANGO OSPINA

ESCUELA DE CIENCIAS Y HUMANIDADES

DEPARTAMENTO DE CIENCIAS BÁSICAS

MEDELLÍN

JUNIO 2009

Índice general

Introducción

Estas notas constituyen las memorias del curso práctica investigativa I en 2009-I, formadas por una primera parte de un trabajo monoráfico sobre Ordenes estocásticos. El objeto fundamental es presentar un material que sirva de apoyo para estudiantes que aborden estudios sobre ordenes estocásticos.

En el primer capítulo se presentan tanto conceptos como resultados preliminares y en el resto se presenta el material estudiado, sobre ordenes estocásticos, según el texto [1]. En la medida de lo posible, se mostrarán algunos desarrollos en los cálculos que los autores sólo mencionan.

Capítulo 1

Preliminares

1.1. Algunos elementos de probabilidad

En esta sección se presenta, en forma abreviada, conceptos básicos de teoría de la probabilidad según la teoría de la medida. Iniciemos con conceptos tomados del libro *Introducción a la teoría avanzada de la probabilidad* [2], en el cual se pueden consultar las definiciones y pruebas de esta sección.

Un **espacio medible** es una estructura $(\Omega, \mathcal{A})$, donde Ω es un conjunto no vacío y $\mathcal{A}$ es una σ-algebra en Ω; en tal caso a los subconjuntos de Ω que sean elementos de la σ-álgebra, se les llama **conjuntos $\mathcal{A}$-medibles**.

Se le llama **espacio de medida** a la terna $(\Omega, \mathcal{A}, \mu)$, donde μ es una **medida** sobre $\mathcal{A}$, esto es: μ es una función de la σ-álgebra en los reales extendidos no negativos $\mu : \mathcal{A} :\longrightarrow [0, \infty]$, tal que $\mu(\Phi) = 0$, si $E_i \in \mathcal{A}$ para $i = 1, 2, ...$ con $E_i \cap E_j = \Phi$ para $i \neq j$ entonces, $\mu(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} \mu(E_i)$.

Recordemos que un **espacio de probabilidad** es un espacio de medida $(\Omega, \mathcal{A}, P)$ tal que $P(\Omega) = 1$. En tal caso: a la medida P se le llama función de probabilidad, a los conjuntos $\mathcal{A}$-medibles (es decir, a los miembros de $\mathcal{A}$) se les llama eventos y a Ω se le llama espacio muestral.

Dados dos espacios medibles $(\Omega_1, \mathcal{A}_1,)$ y $(\Omega_2, \mathcal{A}_2)$ a una función $f : \Omega_1 \longrightarrow \Omega_2$ se le llama **función $\mathcal{A}_1 - \mathcal{A}_2$-medible** si

$$E \in \mathcal{A}_2 \implies f^{-1}(E) \in \mathcal{A}_1,$$

esto es, f devuelve conjuntos $\mathcal{A}_2$ medibles en conjuntos $\mathcal{A}_1$ medibles; análogo a lo que sucede con el concepto de continuidad entre espacios topológicos. En particular, se dirá que f es función medible cuando $(\Omega_2, \mathcal{A}_2) = (\mathbb{R}, \mathcal{B})$, donde $\mathcal{B}$ es la σ-álgebra de Borel en $\mathbb{R}$, que es la menor σ-álgebra que contiene a los conjuntos abiertos en $\mathbb{R}$.

Recordemos también la definición de variable aleatoria con valor real. Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y consideremos el espacio medible $(\mathbb{R}, \mathcal{B})$. Una **variable aleatoria** con valor real es una función medible $X : \Omega \longrightarrow \mathbb{R}$, o sea que

$$E \in \mathcal{B} \implies X^{-1}(E) \in \mathcal{A}.$$

Podría definirse el concepto de variable aleatoria con valor en un espacio diferente a los reales, pero en estas notas siempre se tratarán variables aleatorias con valor real, que llamaremos simplemente variable aleatoria. Muchas veces se escribirá **v.a.** para abreviar variable aleatoria.

Para ver como una v.a permite transportar la probabilidad de su dominio Ω a los reales, se presenta el siguiente concepto: Si $(\Omega_1, \mathcal{A}_1, \mu)$ es un espacio de medida, $(\Omega_2, \mathcal{A}_2)$ es un espacio medible y $T : \Omega_1 \longrightarrow \Omega_2$ es función $\mathcal{A}_1 - \mathcal{A}_2$-medible entonces, la función $\mu_T : \mathcal{A}_2 \longrightarrow [0, \infty]$, definida por $\mu_T(E) = \mu(T^{-1}(E))$ es una medida en el espacio medible $(\Omega_2, \mathcal{A}_2)$. A la medida μ_T se le llama **medida transportada** por T.

Ahora bien, al considerar una v. a. $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B})$, por ser X función medible, la medida de probabilidad P se puede transportar de $(\Omega, \mathcal{A})$ a $(\mathbb{R}, \mathcal{B})$, mediante la medida transportada por X. Se define entonces P_X en B

$$P_X(E) = P(X^{-1}(E)) = P(\{\omega \ : \ X(\omega) \in E\}), \quad \forall E \in \mathcal{B}$$

Es fácil ver que P_X es una probabilidad definida sobre $(\mathbb{R}, \mathcal{B})$. La probabilidad P_X se llama **distribución de probabilidad** de la v.a X.

Lo anterior permite definir la llamada **función de distribución de probabilidad** F_X asociada a la v. a. X, como

$$F_X : \mathbb{R} \longrightarrow \mathbb{R}, \quad \text{tal que para cada } t \in \mathbb{R},\ F_X(t) = P_X((-\infty, t]) = P(X^{-1}((-\infty, t]))$$

De manera que

$$\begin{aligned} F_X(t) &= P(X^{-1}((-\infty, t])) \\ &= P(\{w \in \Omega :\ X(w) \in (-\infty, t]\}) \\ &= P(\{w \in \Omega :\ X(w) \leq t\}). \end{aligned}$$

En forma abreviada, la anterior igualdad se suele denotar por $F_X(t) = P[X \leq t]$.

Si F es la función de distribución de una variable aleatoria, se la llama **función de supervivencia** a la función

$$\overline{F} = 1 - F.$$

Sean un espacio de probabilidad $(\Omega, \mathcal{A}, P)$ y una v.a $X : \Omega \longrightarrow \mathbb{R}$, con X P-integrable (considerando la integral de Lebesgue según la definición 2.2.10, sección 2.2 de [2]). Se definen:

a) el **valor esperado** de X como

$$EX := E(X) := \int_\Omega X dP$$

b) la **varianza** de X como

$$Var(X) := E(X - E(X))^2$$

c) la **desviación estándar** de X como el número

$$\sigma(X) = \sqrt{Var(X)}$$

A continuación se resaltan algunas propiedades de los conceptos definidos, las cuales se proponen como ejercicio.

1. $\forall E \in \mathcal{A} \quad P(E) = E(\chi_E)$, siendo χ_E la llamada función característica definida como $\chi_E(x) = 0$ si $x \notin E$ y $\chi_E(x) = 1$ si $x \in E$.
2. $\forall c \in \mathbb{R} \quad E(c) = c$
3. $X_1 \leq X_2$ implica $EX_1 \leq EX_2$
4. El valor esperado es aplicación lineal sobre el conjunto de v.a en $\mathbb{R}$, es decir, para dos reales a, b y para dos variables aleatorias X e Y cualesquiera, se cumple $E(aX + bY) = aE(X) + bE(Y)$
5. $|E(X)| \leq E(|X|)$
6. Para $a, b \in \mathbb{R}$, $Var(aX + b) = a^2 Var(X)$.

Dado un espacio de medida $(\Omega, \mathcal{A}, \mu)$ y $W \subset \Omega$, se dice que una propiedad H_x se cumple **casi en todas partes de** W (o abreviando c.t.p de W) si existe $M \subset \Omega$, con $\mu(M) = 0$ tal que H_x es cierta para todo $x \in (W - M)$. Cuando se habla de un espacio de probabilidad, en lugar de decir casi en todas partes, se dice **casi seguro** o abreviadamente, **c.s.**

En la página 84 de [2] se prueba el siguiente lema.

Lema 1.1.1. *Si X es v.a P-integrable entonces* $Var(X) = E(X^2) - (E(X))^2$

Se puede demostrara fácilmente que $E(X^2) < \infty$ implica $E(X) < \infty$. En consecuencia, del lema se deduce que $Var(X) < \infty \iff E(X^2) < \infty$.

Además, $Var(X) = 0 \iff X = E(X)$ casi en todas partes de Ω, es decir, $X(t) = E(X)$ para todo $t \in (\Omega - A)$, donde A es algún subconjunto de Ω con $P(A) = 0$.

En la página 5 de [1] se presenta una forma de representar la esperanza para variables aleatorias de esperanza finita, y es la siguiente: si X es una v.a con función de distribución de probabilidad F_X y esperanza finita, entonces

$$EY = \int_0^\infty [1 - F_Y(t)]dt - \int_{-\infty}^0 F_Y(t)dt \tag{1.1}$$

Muchas expresiones son muy útiles para representar integrales de v.a. y se pueden obtener a partir del siguiente teorema. Su prueba se puede consultar en la página 90 de [2].

Teorema 1.1.1. *(Teorema de transformación para integrales) Sea $(\Omega_1, \mathcal{A}_1, \mu)$ un espacio de medida y $(\Omega_2, \mathcal{A}_2)$ un espacio medible. Sean $g : \Omega_1 \longrightarrow \Omega_2$ una función $\mathcal{A}_1 - \mathcal{A}_2$-medible y μ_g la medida transportada por g. Si una función $f : \Omega_2 \longrightarrow \mathbb{R}$ es $\mathcal{A}_2 - \overline{\mathcal{B}}$-medible (siendo $\overline{\mathcal{B}}$ la σ-álgebra de Borel en los reales ampliados) y $E \in \mathcal{A}_2$ entonces,*

$$\int_E f d\mu_g \ \text{ existe } \iff \int_{g^{-1}(E)} f \circ g d\mu \ \text{ existe}$$

y en ese caso

$$\int_E f d\mu_g = \int_{g^{-1}(E)} f \circ g d\mu.$$

Del teorema anterior se tiene lo siguiente: considérese una v.a. $g = X$ definida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, P)$; con lo cual $(\Omega_2, \mathcal{A}_2) = (\mathbb{R}, \mathcal{B})$. Sea $f : \mathbb{R} \longrightarrow \overline{\mathbb{R}}$ función $\mathcal{B} - \overline{\mathcal{B}}$-medible entonces, para $E \in \mathbb{R}$ se tiene

$$\int_E f dP_X \quad \text{existe} \quad \Longleftrightarrow \quad \int_{X^{-1}(E)} f \circ X dP \quad \text{existe}$$

y en ese caso

$$\int_E f dP_X = \int_{X^{-1}(E)} f \circ X dP.$$

En particular, considerando $E = \mathbb{R}$ se tiene $X^{-1}(E) = \Omega$ y en consecuencia los siguientes hechos:

1. si $f(x) = x$ entonces
$$E(X) = \int_{-\infty}^{\infty} x dP_X = \int_\Omega X dP.$$
2. si $f(x) = |x|$ entonces
$$E(|X|) = \int_{-\infty}^{\infty} |x| dP_X = \int_\Omega |X| dP.$$
3. si $f(x) = x^2$ entonces
$$E(X^2) = \int_{-\infty}^{\infty} x^2 dP_X = \int_\Omega X^2 dP.$$
De lo anterior $E((X - EX)^2) = \int_\Omega (X - EX)^2 dP = \int_{-\infty}^{\infty} (x - EX)^2 dP_X$.
Los momentos de orden $k > 0$ se definen como
$$E(X^k) = \int_{-\infty}^{\infty} x^k dP_X = \int_{-\infty}^{\infty} x^k dF_X(x).$$
4. Si X es una v.a. discreta con valores $\{x_i\}$ y con $P_i = P(X = x_i)$ y además $f : \mathbb{R} \longrightarrow \mathbb{R}$ es una función medible, entonces
$$E(f \circ X) = \int_\Omega (f \circ X) dP = \int_{-\infty}^{\infty} f dP_x = \sum_i f(x_i) P(X = x_i) = \sum_i f(x_i) P_i$$

Nos concentramos ahora en definir las funciones de densidad mediante la aplicación del teorema de Radon Nikodym que se verá más adelante. Para ello se necesitan unos conceptos y resultados previos.

Sea $(\Omega, \mathcal{A}, \mu)$ es un espacio medida. Una función medible no negativa $f : \Omega \longrightarrow \mathbb{R}$ define una medida ν en $\mathcal{A}$ mediante:

$$\nu(A) = \int_A f d\mu \tag{1.2}$$

A tal medida ν se le denota por $\nu := f \cdot \mu$, para indicar que ν depende de f y de μ. A la medida ν se le llama **medida con densidad** f.

Teorema 1.1.2. *Sean un espacio de medida* $(\Omega, \mathcal{A}, \mu)$ *y una función medible no negativa* $f : \Omega \longrightarrow \mathbb{R}$. *Considérese* ν, *una medida con densidad* f.

1. *Si* $g : \Omega \longrightarrow \mathbb{R}$ *es función medible no negativa, se satisface que*

$$\int g d\nu = \int g f d\mu, \qquad \text{esto es,} \qquad \int g d(f.\mu) = \int g f d\mu \tag{1.3}$$

2. *Una función* $g : \Omega \longrightarrow \mathbb{R}$ *es* ν*-integrable si y sólo si* fg *es* μ*-integrable. En tal caso también se satisface 1.3.*

Definición 1.1.1. *Se dice que una medida* ν *sobre* $(\Omega, \mathcal{A})$ *es absolutamente continua respecto a una medida* μ *(y se escribe* $\nu << \mu$*), si y sólo si, para cada* $A \in \mathcal{A}$, $\quad \mu(A) = 0 \Longrightarrow \nu(A) = 0$.

Teorema 1.1.3. *Si* ν *es una medida finita en* $(\Omega, \mathcal{A})$, *(en particular, una medida de probabilidad* $\nu = P$*), y* μ *es una medida en* $(\Omega, \mathcal{A})$, *entonces*

$$\nu << \mu \iff \forall \epsilon > 0, \exists \delta > 0 : \ \forall A \in \mathcal{A} \quad \text{se cumple } \mu(A) < \delta \longrightarrow \nu(A) < \epsilon$$

A continuación se presenta el teorema de Radon Nikodym, del cual se desprende la existencia y unicidad c.t.p de $\mathbb{R}$ de las funciones de densidad de probabilidad. Por la importancia del teorema de Radon Nikodym, el libro [2] le dedica una sección completa; la 3.2.

Teorema 1.1.4. *(Teorema de Radon Nikodym) Sea* $(\Omega, \mathcal{A})$ *un espacio medible. Dadas* μ *y* ν *medidas en* $\mathcal{A}$, *si* μ *es* σ*-finita (esto es,* Ω *es unión numerable de subconjuntos de medida finita) entonces, las siguientes condiciones equivalen:*

1. $\nu << \mu$

2. *existe una función medible no negativa* $g : \Omega \longrightarrow \mathbb{R}$ *tal que para cada* $A \in \mathcal{A}$ *se cumple que*

$$\nu(A) = \int_A g d\mu$$

Si h *es otra función que satisface 1.2, entonces* $g = h$ *c.t.p de* Ω.

A la función g del teorema anterior, se le llama función de **densidad** de ν y se denota $g = \frac{d\nu}{d\mu}$.

Como aplicación del teorema de Radon Nikodym se pueden considerar las distribuciones de las v.a y sus densidades.

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y sea X una v.a, con distribución de probabilidad P_X y función de distribución F_X. Sean $\nu = P_X$ y μ la medida de Lebesgue (ver capítulo [2], para la medida de Lebesgue) en $\mathbb{R}$, la cual es σ-finita. Por ser la distribución P_X una medida de probabilidad, es una medida finita y, por el teorema 1.1.3, se tiene que $P_X << \mu$. Ahora bien, por el teorema de Radon Nikodym se sigue la existencia de la función de densidad de P_X, digamos $f_X : \Omega \longrightarrow \mathbb{R}$. La función f_X es llamada función **de densidad de probabilidad** de X.

Además, si f es una función no negativa en $\mathbb{R}$ y

$$P_X(A) = P(X \in A) = \int_A f(x)dx \qquad A \in \mathbb{B},$$

donde la integral se toma respecto a la medida de Lebesgue.
Si la variable aleatoria X tiene densidad f y la función de distribución F, la relación entre F y f es

$$F(x) = \int_{-\infty}^{x} f(t)dt,$$

donde la integral es la de Lebesgue.

Si F tiene derivada continua, su derivada es igual a f c.s. Si la función de densidad f de una variable aleatoria es continua y si $g : \mathbb{R} \to \mathbb{R}$ es una función estrictamente creciente, con $T = g^{-1}$, la función de distribución de probabilidad de $g(X)$ está dada por

$$F_{g(X)} = P(g(X) \leq x) = P(X \leq T(x)) = F(T(x)).$$

Si T es diferenciable y F tiene derivada continua, $\frac{d}{dx}F_{g(X)} = f(T(x))T'(x)$, lo cual es una densidad de $g(X)$.

A continuación se define el concepto de mixtura entre variables aleatorias[5].

Consideremos una v.a X que depende de un parámetro Θ que asume valores θ. Adoptamos la notación $[X|\Theta = \theta]$ para la v.a X cuando θ es fijado y, en tal caso, se escribirá $F_X(x|\theta)$ o $f_X(x|\theta)$ para indicar las funciones de distribución de probabilidad y de densidad de probabilidad respectivamente.

Se puede suponer que Θ se distribuye de acuerdo con una función de distribución $U(\Theta)$ y función de densidad $u(\Theta)$. Se define la función de distribución de **mixtura** de la forma $F_X(x|\theta)$ como

$$F_X(x) = \int_{\Theta} F_X(x|\theta)dU(\theta) = E(\Theta)F_X(x|\Theta). \tag{1.4}$$

La ecuación (1.4) define la distribución incondicional de X en términos de la distribución condicional de X dado $\Theta = \theta$. La función de densidad incondicional de X es entonces:

$$f_X(x) = \int_{\Theta} f_X(x|\theta)dU(\theta) = E(\Theta)f_X(x|\Theta)$$

El momento incondicional de orden k (ver [5]) esta dado por:

$$E[X^k] = \int_{\Theta} E[X^k|\theta]dU(\theta) = E(\Theta)E[X^k|\Theta] \quad k = 1, 2, ...,$$

1.2. Herramientas para análisis de datos

A continuación se introducen algunos conceptos de bastante utilidad.

Función cuantíl

Primero establecemos la siguiente notación: dados $0 \leq u \leq 1$ y una función de distribución de probabilidad F, se define la **función de distribución inversa** (también llamada **función cuantíl**) mediante

$$F^{-1}(u) = \text{ínf}\{x \ : \ F(x) \geq u\}.$$

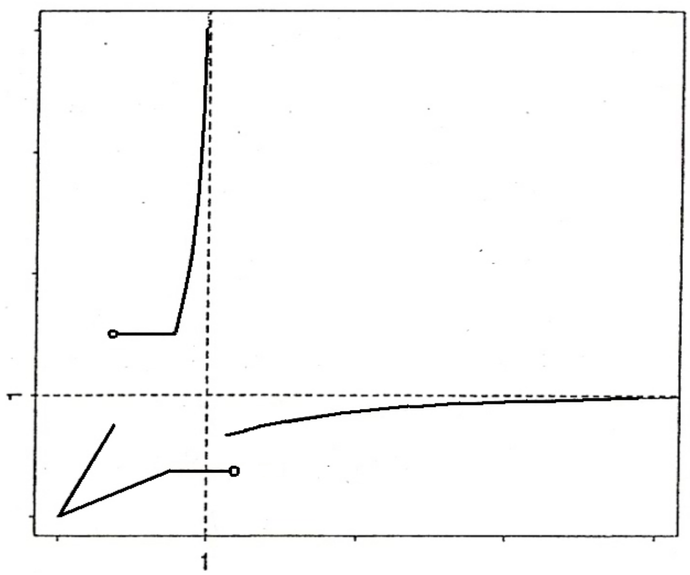

Figura 1.1: Una función de distribución F en $[0, \infty)$ y su función cuantíl F^{-1}.

La Figura resalta que el gráfico de F^{-1} es la reflexión del gráfico de F con respecto a la recta $y = x$.

F^{-1} no es precisamente la función inversa de F, por ejemplo $F \circ F^{-1}$ y $F^{-1} \circ F$ no tienen que ser la función identidad. Pero F^{-1} si cumple las siguientes propiedades:

$$F(F^{-1}(u)) = F(\text{ínf}\{y \ : \ F(y) \geq u\}) \geq u$$

$$F^{-1}(F(x)) = \text{ínf}\{y \ : \ F(y) \geq F(x)\} \leq x$$

Si F es monótona creciente (tal como la función de distribución ϕ de la distribución normal estándar) observamos que F^{-1} es la función inversa conocida de F. Una ilustración de la función cuantíl esta dada en la figura 1.1. Nótese que los intervalos donde F es constante corresponde a un salto de F^{-1}, y un salto de F corresponde a un segmento horizontal en F^{-1}. De esta manera podemos definir la inversa generalizada de la función de distribución empírica F_n, por ejemplo $x_1, ..., x_n$, es decir

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{(-\infty, x]}(x_i), \quad \forall x \in \mathbb{R}.$$

Es fácil verificar que F_n tiene todas las propiedades de una función de distribución

- $\lim_{x\to-\infty} F_n(x) = 0$ y $\lim_{x\to\infty} F_n(x) = 1$
- F_n es no decreciente: $F_n(x) \leq F_n(y)$ para $x \leq y$
- F_n es continua a derecha: $\lim_{y\to x} F_n(y) = F_n(x), \quad \forall x \in \mathbb{R}$

El siguiente resultado es de gran importancia, por lo cual se incluye su demostración.

Proposición 1.2.1. *Si U es ua v.a. que distribuye uniforme en $(0,1)$ y F es una función de distribución, entonces $X = F^{-1}(U)$ es ua v.a. que tiene función de distribución F.*

Demostración. En primer lugar, veamos que $\{(u,x)\ /\ u \leq F(x)\} = \{(u,x)\ /\ F^{-1}(u) \leq x\}$.

Sea (u,x) un elemento del primer conjunto, entonces

$$u \leq F(x) \quad \to \quad F^{-1}(u) = \inf\{y\ /\ F(y) \geq u\} \leq x$$

Luego (u,x) está en el segundo conjunto.

Sea ahora (u,x) un elemento del segundo conjunto.

$$F^{-1}(u) \leq x \quad \to \quad u \leq F(F^{-1}(u)) \leq F(x)$$

Por lo tanto, (u,x) pertenece al primer conjunto.

De otro lado, consideremos $(\mathbb{R}, \mathcal{A}, P)$ y en $(\mathbb{R}^2, \mathcal{A}^2)$ sea $\overline{p}$ la medida producto, como se presenta por ejemplo en el capítulo 3 de [2], que según el teorema 3.1.20 de dicho texto, resulta ser invariante por proyecciones. En el contexto de la presente demostración, la propiedad de invarianza por proyección permite escribir las dos siguientes igualdades:

$$P(\{u\ /\ F^{-1}(u) \leq x\}) = \overline{P}(\{(u,x)\ /\ F^{-1}(u) \leq x\}) \quad \text{y}$$
$$P(\{u\ /\ u \leq F(x)\}) = \overline{P}(\{(u,x)\ /\ u \leq F(x)\}).$$

Ahora bien, partiendo de la igualdad conjuntista probada, se tiene lo siguiente:

$$\begin{aligned} P(X \leq x) &= P(F^{-1}(U) \leq x) \\ &= P(\{u\ /\ F^{-1}(u) \leq x\}) \\ &= \overline{P}(\{(u,x)\ /\ F^{-1}(u) \leq x\}) \\ &= \overline{P}(\{(u,x)\ /\ u \leq F(x)\}) \\ &= P(\{u\ /\ u \leq F(x)\}) \\ &= P(U \leq F(x)) \\ &= F(x). \end{aligned}$$

La última igualdad es por que U es una v.a. uniformemente distribuida en (0,1). □

1.3. P-P plots

Una p-p plot es una técnica gráfica para determinar si un conjunto de datos sigue una distribución, como la normal o Weibull o cualquier otra, y se usa para estimar visualmente la ubicación a escala de los parámetros de la distribución elegida.

Para construir la gráfica p-p plot, los datos ordenados de menor a mayor se representan gráficamente confrontados con los que puede asumir una distribución teórica. Precisamente, si se tienen los datos muestrales $x_1, x_2, ...x_n$ y una función de distribución teórica F, la gráfica p-p plot será la correspondiente a los puntos $((i-0.5)/n, F(x_i))$ para $i = 1, 2, ...n$. Según (Hazen A) en 1930. (para mas información a cerca de como se obtienen los puntos de la gráfica p-p plot y como fue cambiando la manera de gráficar estos a través del tiempo ver articulo [6]).

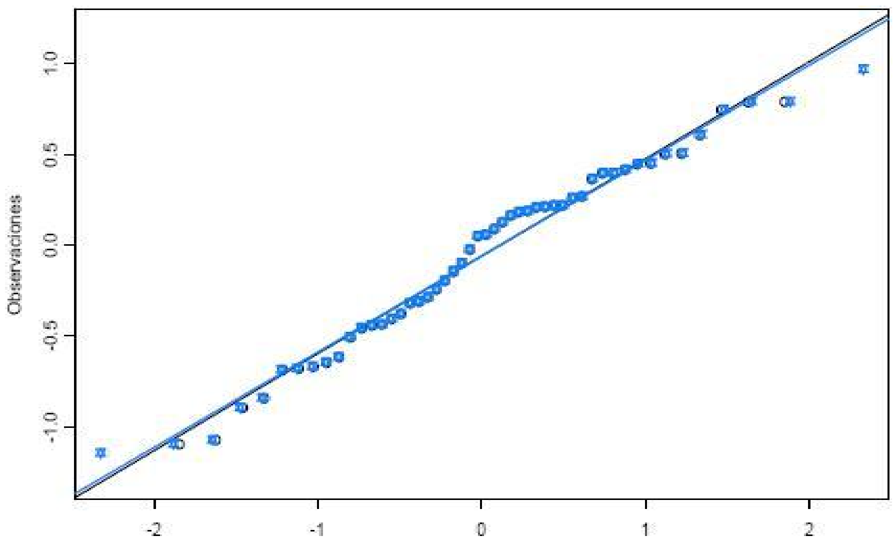

Figura 1.2: Distribución Normal considerando $((i-0.5)/n, F(x_i))$

Si la gráfica p-p plot (1.2) se ajusta a una recta, entonces se asume que los datos distribuyen según F. Si los puntos forman una línea recta aproximadamente, las desviaciones de los puntos respecto de esta línea recta indican las desviaciones de la distribución especificada.

El coeficiente de correlación de la p-p plot, es definido como el coeficiente de correlación lineal asociado con el ajuste de los datos en el gráfico de probabilidad. Dicho coeficiente es una medida de la bondad de ajuste. Las estimaciones de la ubicación a escala de los parámetros de la distribución se da por el intercepto y la pendiente. Varias gráficas P-P plot pueden ser generadas para varias distribuciones que compiten para ver cuál ofrece el mejor ajuste, y la p-p plot que genere la mayor correlación entre los coeficientes es la mejor opción.

1.4. Q-Q plots

Consideramos ahora otra herramienta sencilla para el análisis exploratorio estadístico y aplicable a situaciones simuladas en la vida real (ver aplicaciones en el libro [4]), precisamente la herramienta llamada Q-Q plot o Cuantíl, que corresponde a la inversa generalizada de una función de distribución.

Sea $X_{(1)} \leq, ... \leq X_{(n)}$ una muestra ordenada de $X_1, ...X_n$ y supongamos que la muestra no tiene lazos, es decir, $X_{(1)} <, ... < X_{(n)}$ casi-seguramente (c.s.) por ejemplo, si las x_i son independientes e idénticamente distribuidas(iid) con una densidad la muestra no tiene lazos. Desde entonces la función de distribución empírica de una muestra desde que sea función de distribución, se puede calcular la función cuantíl F_n^{-1} la cual nosotros llamamos la función cuantíl empírica, si la muestra no tiene lazos se puede ver que:

$$F_n(X_{(k)}) = \frac{k}{n}, \qquad k = 1, ...n$$

es decir F_n salta por $\frac{1}{n}$ en todos los valores $X_{(k)}$ y es constante en $[X_{(k)}, X_{(k+1)})$ para $k < n$. esto quiere decir que la función cuantíl empírica F_n^{-1} tiene saltos en los valores $\frac{k}{n}$ por $X_{(k)} - X_{(k-1)}$ y permanece constante en $(\frac{(k-1)}{n}, \frac{k}{n}]$:

$$F_n^{\leftarrow}(t) = \begin{cases} X_{(k)} & t \in (\frac{(k-1)}{n}, \frac{k}{n}], \quad k = 1, ...n-1 \\ \\ X_{(k)} & t \in (\frac{(n-1)}{n}, 1) \end{cases}$$

Un resultado fundamental de la teoría de probabilidad, es lema de Glivenko-Cantelli(ver libro Billingsley [13]pág.275) nos dice lo siguiente: si $X_1, X_2, ...$ es una secuencia (iid) con función de distribución F, entonces

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \overrightarrow{c.s.} 0$$

esto implica que $F_n(x) \approx F(x)$ uniformemente para toda x. Uno puede mostrar que el lema de Glivenko-Cantelli implica $F_n^{-1}(t)$ c.s. como $n \to \infty$ para t continuos en F^{-1} (ver Resnick[64], pág 5) esta observación es la idea básica para las Q-Q plot: si $X_1, ...X_n$ es una muestra con una función de distribución F conocida, nosotros esperaríamos que $F_n^{-1}(t)$ esta cerca de $F^{-1}(t)$ para todo $t \in (0, 1)$, siempre que n sea grande.Por lo tanto, si nosotros trazamos $F_n^{-1}(t)$ contra $F^{-1}(t)$ para todo t en $(0, 1)$ veremos una línea recta aproximadamente. Esto es común al trazar la gráfica.

$$\{(X_k, F^{-1}(\frac{k}{n+1})), \qquad k = 1, ..., n\}$$

Para una función de distribución F dada. Las modificaciones de la posición de las trazas también han sido bien usadas. Chambers [21] da las siguientes propiedades de una Q-Q plot:

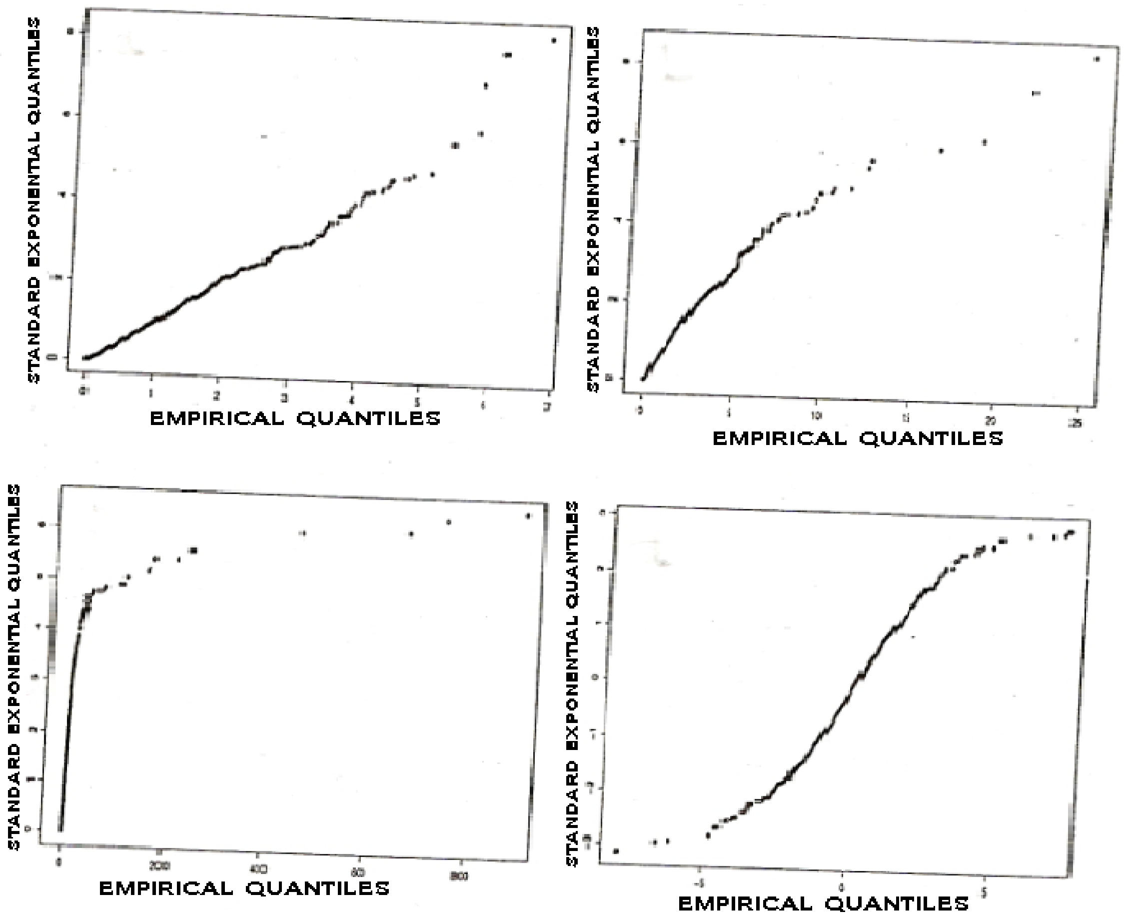

Figura 1.3: QQ-plots para muestras de tamaño 1.000.exponencial estandarizada(arriba derecha), log-normal estandarizada(arriba izquierda), datos distribuidos Pareto con índice de cola 4(abajo derecha) versus función cuantíl de la exponencial estandarizada. datos distribuidos t-student con índice de cola 4 (abajo izquierda) versus función cuantíl de la distribución normal estandarizada.

(a) **Comparación de distribuciones**. Si los datos son generados de una muestra aleatoria De la distribución de referencia, la traza debería verse bruscamente lineal. Esto seria verdadero si los datos provienen de una transformación lineal de la distribución.

(b) **Outliers**. Si uno o unos pocos de los valores de los datos son contaminados por un error grosero o por alguna razón son notablemente diferentes en el valor del resto de valores, siendo estos últimos más o menos distribuidos como la distribución de referencia, los OUTLIERS pueden ser fácilmente identificados en la gráfica.

(c) **Ubicación y escala**. Porque un cambio de una de las distribuciones por una transformación lineal simplemente transforma la gráfica por la misma transformación, uno puede estimar gráficamente(a través del intercepto y de la pendiente)la localización y escala de los parámetros para una muestra de datos, en el supuesto de que los datos vienen de la distribución de referencia.

(d) **Forma**.Algunas diferencias en la forma de la distribución de la gráfica puede ser deducida.

Capítulo 2

Ordenes estocásticos

La teoría acá presentada es tomada básicamente del texto [1]. Como un orden estocástico es un caso particular de un orden parcial, naturalmente iniciemos con la definición de este último.

Una relación binaria $\preceq$ definida en un conjunto no vacío W es llamado **orden parcial** en W si cumple las siguientes propiedades.

1. Reflexiva: para cada $x \in W,\ x \preceq x$.
2. Transitiva : $x \preceq y \quad \text{y} \quad y \preceq z \ \longrightarrow x \preceq z$.
3. Antisimétrica: $x \preceq y \quad \text{y} \quad y \preceq x \ \longrightarrow x = y$

Muchas veces es conveniente usar la notación $Y \succeq X$ en lugar de $X \preceq Y$.

Un **orden estocástico** es un orden parcial en el conjunto S de todas (o algunas convenientemente escogidas) las funciones de distribución de probabilidad.

El orden estocástico usual (o fuerte) se define en S como sigue: dadas dos v. a. X e Y (con funciones de distribución F_X y F_Y, respectivamente) se dice que

$$X \leq_{st} Y, \ \text{si y sólo si,} \quad F_X(t) \geq F_Y(t), \text{ para todo } t \in \mathbb{R}.$$

Idénticamente, en términos de la función de sobrevivencia se puede escribir:

$$X \leq_{st} Y, \ \text{si y sólo si,} \quad \overline{F_X}(t) \leq \overline{F_Y}(t), \text{ para todo } t \in \mathbb{R}.$$

De tal forma, $X \leq_{st} Y$ significa que X asume pequeños valores con una probabilidad mas alta que Y.

Nótese que pueden existir variables aleatorias diferentes con igual distribución, entonces la relación $\leq_{st}$ es antisimétrica como una relación entre distribuciones, pero es antisimétrica como una relación para variables aleatorias.

El siguiente es un teorema de caracterización del orden $\leq_{st}$ en términos de la comparación punto a punto (en algún sentido) entre las v.a.

Teorema 2.0.1. *Sean X, Y variables aleatorias con funciones de distribución F_X y F_Y, respectivamente. Las siguientes afirmaciones equivalen.*

1. $X \leq_{st} Y$
2. *Existen un espacio de probabilidad* $(\Omega, \mathcal{A}, p)$ *y variables aleatorias* $\widehat{X}, \widehat{Y}$ *tales que* $F_{\widehat{X}} = F_X$, *y* $F_{\widehat{Y}} = F_Y$ *con* $\widehat{X}(w) \leq \widehat{Y}(w)$ *para todo* $w \in \Omega$.

Demostración. (1) $\Longrightarrow$ (2). Sean X, Y variables aleatorias con funciones de distribución F_X y F_Y, respectivamente y supongamos que $X \leq_{st} Y$.

En general, la función de distribución inversa para una función de distribución F está dada por

$$F^{-1}(u) = \text{ínf}\{x \ : \ F(x) \geq u\}.$$

Sea U una variable aleatoria uniforme en $(0,1)$. Consideremos $\widehat{X} = F_X^{-1}(U)$ y $\widehat{Y} = F_Y^{-1}(U)$ definidas respecto el espacio de probabilidad $((0,1), \mathcal{B}, P_X)$. Así, por la proposición 1.2.1 se tiene que $F_{\widehat{X}} = F_X$ y $F_{\widehat{Y}} = F_Y$.

De otro lado, como $X \leq_{st} Y$ entonces $F_X(t) \geq F_Y(t)$ para toda $t \in \mathbb{R}$ y así $F_X^{-1}(t) \leq F_Y^{-1}(t)$. Al hacer $t = U(w)$ se tiene $F_X^{-1}(U(w)) \leq F_Y^{-1}(U(w))$, es decir, $\widehat{X}(w) \leq \widehat{Y}(w)$ para todo $w \in \Omega$.

(2) $\Longrightarrow$ (1) Supongamos ahora dos v.a $\widehat{X}$ e $\widehat{Y}$ tales que $F_{\widehat{X}} = F_X$, y $F_{\widehat{Y}} = F_Y$ con $\widehat{X}(w) \leq \widehat{Y}(w)$ para todo $w \in \Omega$.

Resulta verdadera la inclusión

$$\{w \ : \ \widehat{Y}(w) < t\} \subset \{w \ : \ \widehat{X}(w) < t\};$$

porque si $w \in \{w \ : \ \widehat{Y}(w) < t\}$ entonces $\widehat{Y}(w) < t$ y por transitividad $\widehat{X}(w) < t$, quedando que $w \in \{w \ : \ \widehat{Y}(w) < t\}$.

Como la probabilidad es monótona por ser una medida, de la inclusión anterior se sigue que

$$P(\{w \ : \ \widehat{Y}(w) < t\}) \leq P(\{w \ : \ \widehat{X}(w) < t\})$$

y la anterior desigualdad es $F_{\widehat{X}}(t) \leq F_{\widehat{Y}}(t)$, que por la hipótesis equivale a $F_X(t) \leq F_Y(t)$ y en consecuencia $X \leq_{st} Y$. $\square$

Un candidato natural para un orden parcial que compare el tamaño de las variables aleatorias podría ser la relación $\leq_{c.s.}$ dada por $X \leq_{c.s} Y$. Es decir, que existe $E \in \mathcal{A}$, con $P(E) = 0$ tal que $\{w \ : \ X(w) > Y(w)\} \subset E$.

Muchas veces esta relación de orden no sólo depende de la distribución, por ejemplo, $X \leq_{c.s.} X$ se da siempre, pero $X \leq_{c.s.} Y$ no se da si X e Y son independientes e idénticamente distribuidas, con la misma distribución no degenerada (puede suceder por ejemplo, si X representa los tiempos de arribo de camiones a una empresa de carga en un país e Y representa los tiempos de arribo de aviones de una aerolínea a cierto aeropuerto en otro país, de manera que ambas v.a distribuyan gama y con iguales parámetros). Por lo tanto, la definición $\leq_{st}$ es mucho más útil como lo muestra el teorema anterior. De hecho $X \leq_{c.s.} Y$ implica $F_X \leq F_Y$, como se muestra a continuación.

Proposición 2.0.1. $X \leq_{c.s.} Y$ *implica* $X \leq_{st} Y$

Demostración. Sean dos v.a. X e Y definidas sobre $(\Omega, \mathcal{A}, P)$ tales que $X \leq_{c.s.} Y$. Por tanto si $M = \{w \in \Omega : \ X(w) > Y(w)\}$ entonces $P(M) = 0$. Para cada real t denotemos

$$A_t := \{w \in \Omega : \ Y(w) \leq t\} \ \text{ y } \ B_t := \{w \in \Omega : \ X(w) \leq t\}.$$

Veamos que para cada real t se cumple

$$A_t \subset B_t \cup M.$$

Sea $u \in A_t$, así $Y(u) \leq t$. En caso que $u \notin M$ se tiene $X(u) \leq Y(u)$ con lo cual $X(u) \leq t$, esto es $x \in B_t$. Por tanto $u \in B_t \cup M$, lo que prueba la inclusión propuesta.

Ahora, como la probabilidad es monótona,

$$P(A_t) \leq P(B_t \cup M) \leq P(B_t) + P(M) = P(B_t)$$

y $P(A_t) \leq P(B_t)$. En resumen, para cada $t \in \mathbb{R}$ $F_Y(t) \leq F_X(t)$ y en consecuencia $X \leq_{st} Y$. $\square$

Nótese que el orden $\leq_{st}$ es el orden más fuerte implicado por el orden $\leq_{c.s.}$, esto es, si para dos v.a X e Y cualesquiera, $X \leq_{c.s.} Y$ implica $X \leq_0 Y$, donde $\leq_0$ es un orden estocástico, entonces $X \leq_{st} Y$ implica $X \leq_0 Y$. En efecto, si $X \leq_{st} Y$ entonces por el teorema 2.0.1 sin perdida de generalidad se puede suponer que $X \leq_{c.s.} Y$ y por tanto $X \leq_0 Y$; así $X \leq_{st} Y$ implica $X \leq_0 Y$.

Un orden estocástico también puede ser caracterizado por la función de distribución inversa relativa Φ_{FG} definida como

$$\Phi_{FG}(t) := G^{-1}(F(t))$$

El siguiente teorema proporciona un procedimiento gráfico simple para verificar ordenes estocásticos entre dos funciones de distribución particulares F y G. Informalmente, dicho teorema afirma que $G \leq_{st} F$ se da si y solo si el $Q - Q$ plot de G^{-1} contra F^{-1} está por debajo de la gráfica de la función identidad.

Teorema 2.0.2. *Sea F y G funciones de distribución arbitrarias. Entonces $G \leq_{st} F$ si y solo si, $\Phi_{FG}(x) = G^{-1}(F(x)) \leq x$ para toda x.*

Demostración. Lo primero es que por definición $G^{-1}(G(x)) \leq x$ y $x \leq G(G^{-1}(x))$ para toda x.

$\Longrightarrow$) Si $G \leq_{st} F$ entonces $F(x) \leq G(x)$ para toda x.

De lo anterior, $G^{-1}(F(x)) \leq G^{-1}(G(x)) \ \ \forall x \in \mathbb{R}$. Por tanto $G^{-1}(F(x)) \leq x, \ \ \forall x \in \mathbb{R}$.

$\Longleftarrow$) Si $G^{-1}(F(x)) \leq x, \ \ \forall x \in \mathbb{R}$ entonces $G(G^{-1}(F(x))) \leq G(x)$ y como $F(x) \leq G(G^{-1}(F(x)))$ se tiene que $F(x) \leq G(x) \forall x \in \mathbb{R}$. Así $G \leq_{st} F$.

Teorema 2.0.3. *$F \leq_{st} G$, si y solo si, la correspondiente $P - P$ plot esta por debajo de la gráfica de la función identidad.*

El siguiente resultado plantea una importante caracterización del orden usual.

Teorema 2.0.4. *Las siguientes afirmaciones son equivalentes.*
1 $X \leq_{st} Y$.
2 la desigualdad $Ef(X) \leq Ef(Y)$ se cumple para todas las funciones no decrecientes f, para las cuales ambas esperanzas existan.

Comentario: Más aún, si para una función f se da la desigualdad de (2), para todas las v.a. X e Y tales que $X \leq_{st} Y$ entonces, f es no decreciente.

Demostración. □

$\Longrightarrow$) De acuerdo con el teorema 2.0.1, se puede asumir sin pérdida de generalidad, que $X \leq Y$ c.s. Ahora, al ser f no decreciente $f(X) \leq f(Y)$ c.s. y por la monotonía de la esperanza, $Ef(X) \leq Ef(Y)$.

$\Longleftarrow$) Sean X e Y tales que para cada función f no decreciente, se cumpla $Ef(X) \leq Ef(Y)$, siempre que ambas esperanzas existan.

Para cada t considérese la función f_t definida para cada x por

$$f_t(x) = \begin{cases} 1, & \text{si } x > t, \\ 0, & \text{en otro caso.} \end{cases}$$

Dicha función es no decreciente y además, $P(X > t) = Ef_t(X)$ y $P(Y > t) = Ef_t(Y)$. De tal forma la desigualdad $Ef_t(X) \leq Ef_t(Y)$ se transforma en $P(X > t) \leq P(Y > t)$, o bien $\overline{F_X} \leq \overline{F_Y}$. Así que, $X \leq_{st} Y$.

□

El siguiente resultado plantea que el orden usual implica el ordenamiento de esperanzas y que diferentes distribuciones con la misma esperanza, no pueden ser ordenadas con el orden usual.

Teorema 2.0.5. *Sean X e Y variables aleatorias con esperanza finita.*
a) Si $X \leq_{st} Y$ entonces $EX \leq EY$;

b) Si $X \leq_{st} Y$ y $EX = EY$ entonces X e Y tienen la misma distribución.

Demostración. a) Es claro, partiendo del teorema 2.0.4 para el caso particular $f(x) = x$.

b) Supongamos que $X \leq_{st} Y$ y $EX = EY$.

Por tener X e Y esperanza finita y según 1.1, se puede representar EY y EX por

$$EY = \int_0^\infty [1 - F_Y(t)]dt - \int_{-\infty}^0 F_Y(t)dt \quad \text{y} \quad EX = \int_0^\infty [1 - F_X(t)]dt - \int_{-\infty}^0 F_X(t)dt$$

Dada la convergencia de cada una de las integrales mencionadas se tiene:

$$\begin{aligned} EY - EX &= \int_0^\infty [1 - F_Y(t)]dt - \int_{-\infty}^0 F_Y(t)dt - \left(\int_0^\infty [1 - F_X(t)]dt - \int_{-\infty}^0 F_X(t)dt\right) \\ &= \int_0^\infty [F_X(t) - F_y(t)]dt + \int_{-\infty}^0 [F_X(t) - F_Y(t)]dt \\ &= \int_{-\infty}^\infty [F_X(t) - F_Y(t)]dt. \end{aligned}$$

Resumiendo,

$$EY - EX = \int_{-\infty}^\infty [F_X(t) - F_Y(t)]dt$$

Ahora, como $EX = EY$ entonces $EY - EX = 0$ y por tanto

$$0 = \int_{-\infty}^{\infty} [F_X(t) - F_Y(t)]dt$$

Como $X \leq_{st} Y$ entonces $F_X(t) \geq F_y(t)$ para todo t y en consecuencia el integrando de la anterior integral es no negativo, así para que la integral sea cero es necesario que $F_X(t) - F_Y(t) = 0$, o bien $F_X = F_Y$ casi seguramente. □

Teorema 2.0.6. *Si $X \leq_{st} Y$ entonces $EX^n \leq EY^n$, para todo $n = 1, 3, 5, 7, ...$ y siempre que las esperanzas existan. Si además, X e Y son no negativas entonces también se cumple $EX^n \leq EY^n$, para todo $n \in \mathbb{N}$.*

Demostración. Las funciones $f_n(x) = x^n$ son no decrecientes en $\mathbb{R}$ para todo $n = 1, 3, 5, 7, ...$ y son no decrecientes en $[0, \infty)$ para todo $n \in \mathbb{N}$. En esos casos, respectivamente, se puede aplicar el teorema 2.0.4 y se obtiene $Ef_n(X) \leq Ef_n(Y)$, o bien $EX^n \leq EY^n$, válido para los casos mencionados en el enunciado. □

Teorema 2.0.7. *Si $X \leq_{st} Y$ entonces para toda función f, no decreciente, se cumple que $f(X) \leq_{st} f(Y)$.*

Demostración. Supongamos que $X \leq_{st} Y$. Dada cualquier función g no decreciente, resulta que $g \circ f$ es no decreciente y por tanto $Eg(f(x)) = E(g \circ f)_(x) \leq E(g \circ f)_(y) = Eg(f(y))$. De lo anterior, por el teorema 2.0.4, se deduce que $f(X) \leq_{st} f(Y)$. □

En la sección 4.4 de [2], se presenta el concepto de convergencia débil para sucesiones de variables aleatorias. Al respecto, el siguiente teorema muestra que el orden $\leq_{st}$ respeta la convergencia débil entre sucesiones de variables aleatorias.

Teorema 2.0.8. *En un espacio de probabilidad $(\Omega, \mathcal{A}, P)$ sean (X_n), (Y_n) sucesiones de variables aleatorias, y X, Y variables aleatorias tales que X_n y Y_n convergen debilmente a X y Y respectivamente. Si $X_n \leq_{st} Y_n$ para toda $n \in \mathbb{N}$ entonces $X \leq_{st} Y$.*

Demostración. Sean F y G las funciones de distribución de probabilidad de X e Y repectivamente y para cada n natural, consideremos F_n y G_n las respectivas funciones de distribución de probabilidad correspondientes a X_n e Y_n. Como para cada natural n se tiene $X_n \leq_{st} Y_n$, entonces $F_n(t) \geq_{st} G_n(t)$ para todo real t. Ahora, por la convergencia débil de (X_n) y (Y_n) a X e Y se sigue que $F(t) \geq G(t)$, siempre que t sea un punto de continuidad de ambas funciones, F y G. Dado que el conjunto de puntos de discontinuidad de F y G es contable numerable, los puntos de continuidad son densos en $\mathbb{R}$. De lo anterior se obtiene que $X \leq_{st} Y$. □

A continuación se mostrará que el orden estocástico usual tiene la propiedad de ser cerrado bajo mixturas.

Teorema 2.0.9. *La relación de orden $\leq_{st}$ es cerrada bajo mixturas, es decir, si X, Y y Θ son variables aleatorias tales que $[X|\Theta = \theta] \leq_{st} [Y|\Theta = \theta]$ para todo θ en el soporte de Θ, entonces $X \leq_{st} Y$.*

Demostración. Esto es inmediato del teorema 2.0.4 dado que para cualquier función f no decreciente,

$$Ef(X) = E_\Theta E[f(X)|\Theta] \leq E_\Theta E[f(Y)|\Theta] = Ef(Y).$$

□

Teorema 2.0.10. *Sea $X_1, ..., X_n$ y $Y_1, ..., Y_n$ variables aleatorias independientes con $X_i \leq_{st} Y_i$ para $i = 1, ..., n$, y se supone que $\psi : \mathbb{R}^n \to \mathbb{R}$ es una función no decreciente (en el sentido: $x_i \leq y_i$ para algún $i = 1, 2, ..., n$ implica $\psi(a_1, a_2, ...x_i, ..., a_n) \leq \psi(a_1, a_2, ...y_i, ..., a_n)$), entonces*

$$\psi(X_1, ..., X_n) \leq_{st} \psi(Y_1, ..., Y_n)$$

Demostración. Procedamos por inducción matemática.

Para $n = 1$ se tiene $X_1 \leq_{st} Y_1$. Por el teorema 2.0.7, se tiene que $\psi(X_1) \leq_{st} \psi(X_2)$.

Como hipótesis inductiva, considérese cierto el enunciado para $n - 1$. Sea f una función real de variable real, no decreciente.

Definamos $g(x) = Ef(\psi(X_1, ..., X_{n-1}, x))$ y $h(x) = Ef(\psi(Y_1, ..., Y_{n-1}, x))$ para todo real x. De acuerdo con la hipótesis de inducción, $g(x) \leq h(x)$ para todo $x \in \mathbb{R}$.

Ambas g y h son funciones no decrecientes, dado que f y son ψ no decrecientes. Entonces, como $X_n \leq_{st} Y_n$

$$Ef(\psi(X_1, ..., X_n)) = Eg(X_n) \leq Eh(Y_n) = Ef(\psi(Y_1, ..., Y_n))$$

y por tanto,

$$\psi(X_1, ..., X_n) \leq_{st} \psi(Y_1, ..., X_n).$$

□

Si en el teorema anterior se toma el caso especial $\psi(x_1, ..., x_n) = \sum_{i=1}^n x_i$ se tiene que $\leq_{st}$ es preservado bajo adición para v.a independientes. Mas precisamente se tiene el siguiente resultado.

Teorema 2.0.11. *Sea $X_1, ..., X_n$ y $Y_1, ..., Y_n$ v.a. independientes con $X_i \leq_{st} Y_i$ para $i = 1, 2, ..., n$. Entonces*

$$X_1 + ... + X_n \leq_{st} Y_1 + ... + Y_n.$$

Bibliografía

[1] Müller. A. y Stoyan. D. : *Comparison methods for stochastic models and risks*. John Wiley & Sons, LTD. England, 2002.

[2] Muñoz, Myriam. y Blanco, Liliana. R.: *Intruducción a la teoría avanzada de la probabilidad*. Primera edición. Universidad Nacional de Colombia-UNILIBOS. Bogotá. 2002.

[3] Royden, H. L.: *Real Amalysis*. Segunda edición. The Macmillan Company. New York. 1968.

[4] Thomas Mikosch.: *Non-life Insurance Mathematics*. Springer-Verlag 2004.

[5] Willmot, Gordon. y Panjer, Harry H.: *Insurance Risk Models*. Society of Actuaries. Estados Unidos 1992

[6] universidad de JAÉN.: *Q-Q Plot Normal. Los puntos de posición gráfica*. http://virtual.ujaen.es/ininv/images/ficheros/art_investigacion/num2_0607/2a9.pdf consultada en mayo del 2009