1er Coloquio del Departamento de Matemáticas

Probabilidad, Esperanza, Métricas Probabilísticas y Estabilidad de Algunos Modelos en Seguros y Finanzas

Evgeni Gordienko

Casa abierta al tiempo

Probabilidad, Esperanza, Métricas Probabilísticas y Estabilidad de Algunos Modelos en Seguros y Finanzas

Evgeni Gordienko

Casa abierta al tiempo
Universidad Autónoma Metropolitana

Lista de abreviaciones y de notaciones estándares

v.a.	-	variable aleatoria;
f.d.	-	función de distribución;
a.c.	-	absolutamente continua;
i.i.d	-	independientes e idénticamente distribuidas;
L.F.G.N.	-	ley fuerte de los grandes números;
T.L.C.	-	teorema del límite central;
$\mathbb{R}$	-	el conjunto de los números reales;
Ω	-	espacio muestral;
P	-	probabilidad;
F_X	-	f.d. de la v.a. X;
f_X	-	densidad de la v.a. a.c. X;
EX	-	esperanza de la v.a. X;
$Var(X)$	-	varianza de la v.a. X.

Introducción

La finalidad del minicurso es recordar los conceptos importantes de la teoría de probabilidad y dar una corta introducción a la teoría moderna de métricas probabilísticas y sus aplicaciones en el estudio de estabilidad (robustez) de modelos particulares de riesgo y de optimización de inversiones. Con relación a los últimos se incluyen algunos resultados recientemente publicados. Al resolver los ejercicios incluídos, el alumno podrá ampliar su conocimiento en cuanto al objeto.

1. Espacio de probabilidad, variable aleatoria y su distribución

DEFINICIÓN 1.1. Un *espacio de probabilidad* es la tripleta $(\Omega, \mathfrak{F}, P)$, donde:

1. Ω es un conjunto, llamado *espacio muestral* ;
2. $\mathfrak{F} = \{A, B, C, \dots\}$ es una familia de subconjuntos de Ω, donde A, B, $C, \dots$ se denominan *eventos* (con la propiedad de que $A \cup B \cup C \cup \dots$, $A \cap B \cap C \cap \dots$, $\overline{A} = \Omega \setminus A$ también son eventos);
3. P es la *probabilidad*, i.e. una función (regla de correspondencia) que asigana a cada evento $A \in \mathfrak{F}$ el número $P(A) \in [0, 1]$, que se llama *la probabilidad del evento A*. La probabilidad P debe satisfacer lo siguiente:

$$P(\Omega) = 1;$$
$$P(A \cup B \cup C \cup \dots) = P(A) + P(B) + P(C) + \dots$$

siempre que los eventos A, B, C,... son *disjuntos* (en la figura 1 por ejemplo, A y D no son disjuntos pero A y C sí).

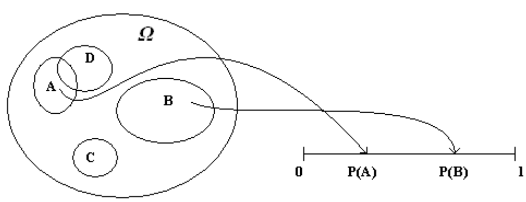

FIGURE 1. Conjuntos disjuntos.

EJEMPLO 1.1. Un dado simétrico se lanza dos veces. A este "experimento" corresponde el espacio de probabilidad:

$$\Omega = \{\omega = (i, j); \quad i, j = 1, 2, ..., 6\};$$
$$\mathfrak{F} = \{\textit{todos los subconjuntos}\};$$
$$P(A) = \frac{\#\ de\ \omega \in A}{36}$$

Aquí, i y j son los puntajes correspondientes al primer y segundo lanzamiento respectivamente. Por ejemplo, si $A :=$ "*la suma del puntaje es* $> 10 = \{(5,6),(6,5),(6,6)\}$, entonces $P(A) = \frac{3}{36} = \frac{1}{12}$ y $P(\overline{A}) = P($*la suma del puntaje es* $\leqslant 10) = 1 - P(A) = \frac{11}{12}$.

> **Nota 1.1:** (a) Cuando $\omega \in A$ nos estamos refiriendo a la **ocurrencia del evento A**. En el ejemplo anterior, si se tiene $\omega = (6,5)$, se dice que ocurre A, pero si se tiene $\omega = (4,1)$, entonces ocurre $\overline{A}$ (el complemento de A).
> (b) En cualquier espacio de probabilidad, $P(\overline{A}) = 1 - P(A)$.

DEFINICIÓN 1.2. Una *variable aleatoria* (v.a.) X es una función ("medible") que asigna a cada $\omega \in \Omega$ un número real $X(\omega)$ llamado *el valor de la v.a. X*.

EJEMPLO 1.2. Teniendo en cuenta el Ejemplo 1.1 consideremos que $X(\omega) = X((i,j)) := i + j$ (la suma de los puntajes). En este caso la v.a. X puede tomar valores en $\{2, 3, \ldots, 12\}$, y

$$\begin{aligned} P(X = 2) &= P((1,1)) = \frac{1}{36}, \\ P(X = 3) &= P((1,2)\ \acute{o}\ (2,1)) = \frac{2}{36} = \frac{1}{18},\ \text{etc.} \end{aligned} \tag{1.1}$$

DEFINICIÓN 1.3. (a) Una v.a. X se llama *discreta* si el conjunto:

$$\{x_1, x_2, ..., x_k, \ldots\} \tag{1.2}$$

de sus valores, es finito o numerable (en el último caso, sus elementos se pueden enumerar pero sin llegar a terminar).

(b) La *distribución* de una v.a. discreta X es el conjunto (1.2) y las probabilidades:

$$\boxed{P(X = x_k), \qquad k = 1, 2, \ldots;} \tag{1.3}$$

que se definen de forma semejante al caso particular 1.1. Los números $P(X = x_k)$ representan las probabilidades de que la v.a. X toma los valores x_k $(k = 1, 2, \ldots)$.

EJEMPLO 1.3. (a) La v.a. *de Bernoulli* con el parámetro $p \in [0,1]$ tiene sus valores en el conjunto: $\{0,1\}$ y $P(X = 0) = 1 - p$; $P(X = 1) = p$. Se escribe: $X \sim Bern(p)$.

(b) Si $S_n = X_1 + X_2 + \cdots + X_n$, $(n \geqslant 1)$ y $X_1, X_2, \ldots$ son independientes (vea õ2) y $X_i \sim Bern(p)$ $(i = 1, 2, ...)$, entonces se demuestra que:

$$P(S_n = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n,$$

y la v.a. S_n se llama *Binomial* (con parámetros n y p). Se escribe: $S_n \sim Bin(n, p)$.

DEFINICIÓN 1.4. La función $F_X \colon \mathbb{R} \to [0, 1]$ que toma los siguientes valores:

$$\boxed{F_X(x) := P(X \leqslant x), \qquad x \in \mathbb{R} \equiv (-\infty, \infty)} \tag{1.4}$$

se llama la *función de distribución* (f.d.) de la v.a. X.

De 1.4 y de que $P(A) \leqslant P(B)$ si $A \subset B$ sigue que $F_X(x)$ es no decreciente y sus valores se aproximan a 1 cuando $x \to \infty$ y se aproximan a 0 cuando $x \to -\infty$.

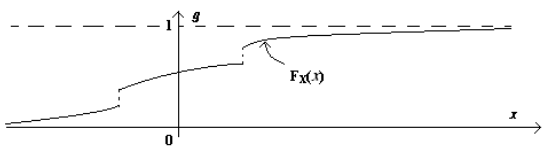

FIGURE 2. Función distribución de $F_X(x)$.

DEFINICIÓN 1.5. Una v.a. X se denomina *absolutamente continua* (a.c.) si:

$$\boxed{F_X = \int_{-\infty}^{x} f_X(t)dt, \qquad x \in \mathbb{R}} \tag{1.5}$$

donde la función $f_X \colon \mathbb{R} \to [0, \infty)$ se llama *la densidad de la v.a. X*.

La igualdad 1.5 implica que F_X es continua (no tiene saltos como en la imagen anterior) y además es derivable:

$$\boxed{F_X'(x) = f_X(x)} \tag{1.6}$$

en “casi todos" los puntos $x \in \mathbb{R}$.

PROPOSICIÓN 1.1. *Sea $I = [a, b] \subset \mathbb{R}$ un intervalo (o un subconjunto más general). Entonces se cumple:*

(a)

$$\boxed{P(X \in I) = \sum_{x_k \in I} P(X = x_k),} \tag{1.7}$$

si X es una v.a. discreta;

(b)

$$\boxed{P(X \in I) = \int_a^b f_X(x)dx.} \tag{1.8}$$

si X es a.c. con densidad $f_X(x)$.

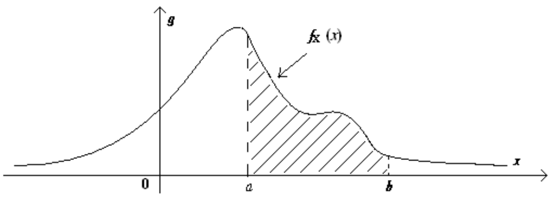

FIGURE 3. El área rayada = $\int_a^b f_X(x)dx = P(X \in [a,b])$.

Nota 1.2: A diferencia de 1.3, para una v.a. X a.c., por 1.8 se tiene que:

$$P(X = a) = \int_a^a f_X(x)dx = 0 \text{ para cada } a \in \mathbb{R}$$

EJEMPLO 1.4. La v.a. **exponencial** con parámetro $\lambda > 0$ tiene la siguiente función de densidad:

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{si } x > 0, \\ 0, & \text{si } x \leqslant 0. \end{cases} \tag{1.9}$$

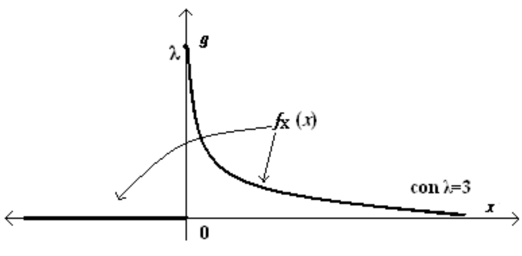

FIGURE 4. Densidad de la v.a. exponencial.

Por 1.5, tenemos que su f.d. es

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{si } x > 0, \\ 0, & \text{si } x \leqslant 0. \end{cases} \tag{1.10}$$

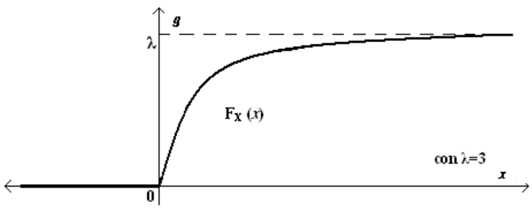

FIGURE 5. Distribución de la v.a. exponencial

En este caso escribimos: $X \sim Exp(\lambda)$.

Por ejemplo, de física sabemos que el isótopo (radiactivo) 283 de Uranio tiene una duración de "vida" (hasta su desintegración) que se da por la v.a. $X \sim Exp(\lambda)$ con $\lambda \approx 1.53896 \cdot 10^{-10}$ (1/año). Tendremos entonces lo siguiente:

$P(X > 5 \cdot 10^9$ años$) = 1 - P(X \leqslant 5 \cdot 10^9$ años$) =$ *(por 1.4)*$=1 - F_X(5 \cdot 10^9) =$ *(por 1.10)* $=$

$$e^{-1.53896 \cdot 10^{-10} \cdot 5 \cdot 10^9} \approx 0.46325 \tag{1.11}$$

(Alrededor de "50% de chances" de que un átomo "viva" más de $5 \cdot 10^9$ años.)

Nota 1.3: El tiempo de existencia de la tierra se estima entre los $4.5 \cdot 10^9$ y $4.8 \cdot 10^9$ años. Por 1.11 podemos ver que una cantidad considerable de los átomos de Uranio-283 existían antes del origen de la Tierra.

EJEMPLO 1.5. La v.a. a.c. que tiene densidad

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right), \quad x \in \mathbb{R} \tag{1.12}$$

se llama **Normal (Gaussiana)** con parámetros: $a \in \mathbb{R}$ ($a = EX$ es el promedio de X, vease õ3) y $\sigma > 0$ (éste se llama "la desviación estándar", ver õ3). Para una v.a. X con la densidad 1.12 escribimos: $X \sim Norm(a, \sigma)$. La v.a. $\eta \sim Norm(0, 1)$ se denomina **Normal estándar**.

Por 1.5 tenemos que

$$F_\eta(x) \equiv \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

Los valores de esta integral no se calculan explícitamente, y por eso es necesario auxiliarse de las tablas que contienen los valores de $\Phi(x)$.

En la década (1970–1980) estuvo de moda medir el coeficiente intelectual (I.Q.) de las personas por medio de pruebas (con escala de 0

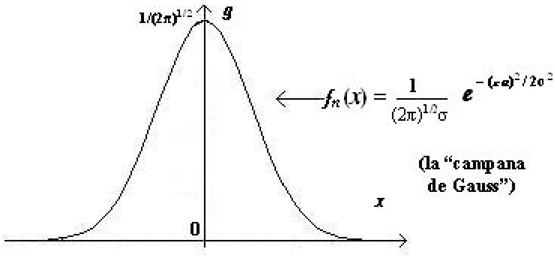

FIGURE 6. La Campana de Gauss.

a 200 puntos). Por estadísticas se sabe que en cierto país, el I.Q. de una persona escogida al azar se da por la v.a. $X \sim Norm(a = 100, \sigma = 17.2)$. Calculemos entonces la probabilidad de que "una persona sea muy lista" i.e. (es decir) $P(X > 140) = P(\frac{X-a}{\sigma} > \frac{140-a}{\sigma}) \approx 1 - P(\frac{X-a}{\sigma} \leqslant 23.3)$. Es fácil ver que la v.a. $\eta = \frac{X-a}{\sigma}$ es Normal estándar, por tanto: $P(X > 140) \approx 1 - P(\eta \leqslant 23.3) = 1 - \Phi(23.3) \approx$ (por tablas) ≈ 0.0099 (i.e. alrededor de 1 persona de cada 100; véase la ley de los grandes números en õ6 para la justificación de la relación entre probabilidad y "frecuencia").

Ejercicios 1

1.1 Considere en el Ejemplo 1.1 el evento A:= "en el primer lanzamiento sale un puntaje mayor que en el segundo". Calcule $P(A)$.

$\underline{Resp.}: P(A) = 1/3.$

1.2 ¿Cuál es la probabilidad de ganar el primer premio en el "Melate" comprando un boleto?

$\underline{Resp.}: 1/\frac{44!}{6!(44-6)!} = \frac{1}{7059052}$; *donde* $m! = m(m-1)(m-2) \cdot \ldots \cdot 2 \cdot 1$.

(Según la interpretación dada en el Ejemplo 6.1 de õ6, ud. tendría casi el 100% de chances de ganar el primer premio comprando alrededor de 7 millones de boletos.)

1.3 Un "punto" se "lanza al azar" dentro del intervalo [0.1].

(a) Construya un espacio de probabilidad adecuado para este "experimento".

(b) Sea X el punto de caída en la figura 7 Explique el hecho de que X es una v.a. y para el espacio construído calcule $P(X \leqslant 1/2)$ y muestre que $P(X = 1/2) = P(X \in \mathbb{Q}) = 0$ (a pesar de que los conjuntos $\{1/2\}$ y $\mathbb{Q}$:= *todos los racionales de* $[0, 1]$ no son vacíos y son muy distintos).

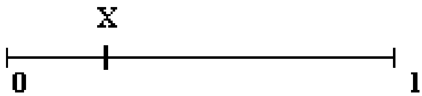

FIGURE 7. Punto de caída: X.

1.4 Una nueva familia está planeando el número n de hijos que desean tener . Ésta familia quiere tener no menos de dos varones. ¿Cuál es el número mínimo n para que se cumpla que: P(*nacerán por lo menos 2 varones*)$\geqslant$ 0.99?.
Resp. : $n = 11$.
1.5 Dos personas (digamos I y II) tienen la misma capacidad para jugar ajedrez. ¿Qué es más probable:
(a) que I gana (a II) 2 partidas de 3?
(b) que I gana 3 partidas de 6?
Resp. : Es más probable la ocurrencia de (a) que la ocurrencia de (b) (¿Contradice ésto a su intuición?).
1.6 Considerando el Ejemplo 1.5 (con X=I.Q.) calcule $P(80 \leqslant X \leqslant 120)$.
Resp. : 0.754.
1.7 Sean $X \sim Exp(\lambda = 1)$ y $[X]$ la parte entera de X. Calcule $P([x]$ *sea un número par*).
Resp. : $e(1+e)^{-1} \approx 0.73106$.
1.8 Sea $X \geqslant 0$ una v.a. a.c. Demuestre que para cada $t, s \geqslant 0$, $P(X > t+s|X > s) = P(X > t)$ (“ausencia de la memoria") sys (si y sólo si) existe $\lambda > 0$ tal que $X \sim Exp(\lambda)$. (Consulte probabilidad condicional en õ2.)

2. Independencia

La **probabilidad condicional** $P(A|B)$ **del evento** A **dado el evento** B se define como:

$$\boxed{P(A|B) := \frac{P(A \cap B)}{P(B)}.} \tag{2.13}$$

Los eventos A y B se llaman **independientes** si $P(A|B) = P(A)$ o por 2.13

$$\boxed{P(A \cap B) = P(A)P(B).} \tag{2.14}$$

Las v.a's. X y Y se llaman **independientes** si (análogamente a 2.14) se tiene que:

$$\boxed{P(X \in I, Y \in I) = P(X \in I)P(Y \in I),} \tag{2.15}$$

para **cada** par de intervalos $I, J \subset \mathbb{R}$. O bien:

$$\boxed{P(X \in I | Y \in J) = P(X \in I);} \tag{2.16}$$

i.e. la información sobre los valores de Y no afecta a las probabilidades de los valores de X.

Nota 2.1: La coma (,) en 2.15 representa la intersección (i.e. la ocurrencia simultánea de $X \in I$ y $Y \in J$).

Escogiendo en 2.15 a $I = (-\infty, x]$, $J = (-\infty, y]$, $x, y \in \mathbb{R}$ obtendremos la definición equi-valente:

DEFINICIÓN 2.1. Las v.a's X y Y son *independientes* si para cada $x, y \in \mathbb{R}$

$$\boxed{F_{X,Y} := P(X \leqslant x, Y \leqslant y) = P(X \leqslant x)P(Y \leqslant y) = F_X(x)F_Y(y).} \tag{2.17}$$

La función $F_{X,Y}$ (de dos variables reales) en la parte izquierda de 2.17 se denomina *la f.d. conjunta de X y Y*.

De forma semejante a 1.5 tenemos:

DEFINICIÓN 2.2. . El par de v.a's. X, Y se llama *absolutamente continuo* (a.c.) con la *densidad conjunta* $f_{X,Y}$, si:

$$\boxed{F_{X,Y}(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(s,t)\,ds\,dt \quad x, y \in \mathbb{R}.}$$

En este caso, para "casi todos" los puntos (x, y) en el plano:

$$\boxed{f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y}F_{X,Y}(x, y);} \tag{2.18}$$

y de 2.17, 2.18

$$\boxed{f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y), \quad x, y \in \mathbb{R};} \tag{2.19}$$

para las v.a's. a.c. e **independientes** X y Y.

EJEMPLO 2.1. Sean $D \subset \mathbb{R}^2$ un conjunto acotado con el área: $ar(D)$ bien definida, y

$$f_{X,Y}(x, y) := \begin{cases} 1/ar(D), & \text{si } (x, y) \in D, \\ 0, & \text{si } (x, y) \notin D. \end{cases}$$

(La densidad uniforme en D). Entonces (consulte el Ejercicio 2.2):

(a) X y Y son independientes (es decir, se cumplen 2.17 o 2.19) si D es como en la figura 8;

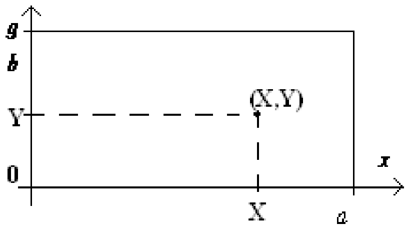

FIGURE 8. D = el rectángulo.

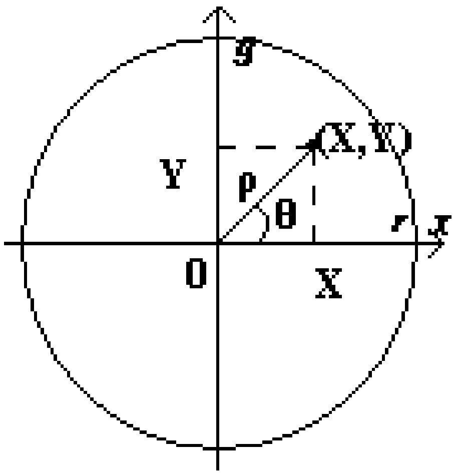

FIGURE 9. D = el círculo de radio r.

(b) X y Y son dependientes si D es como en la figura 9

Sin embargo, (vea el Ejercicio 2.3) las v.a's. ρ y θ (coordenadas polares de (X, Y)) son independientes. Es decir, las funciones de v.a's. dependientes podrían, a veces, ser independientes.

Por otro lado, de 2.16 tendremos que:

$$\boxed{\text{Las funciones de v.a's. independientes siempre son independientes.}} \tag{2.20}$$

DEFINICIÓN 2.3. . Sea $\{X_n\} = X_1, X_2, \ldots, X_n, \ldots$ una suceción de v.a's. Se dice que las v.a's. son *independientes e idénticamente distribuídas* (i.i.d. en lo que sigue) si:

a) $F_{X_1} \equiv F_{X_2} \equiv \ldots$;

b) Para cada $n = 1.2, \ldots$ análogamente a 2.17, la f.d. conjunta de $X_1, X_2, \ldots, X_n$ se factoriza en el producto de las f.d's. (marginales) $F_{X_1}, F_{X_2}, \ldots, F_{X_n}$ de X_k $(1 \leqslant k \leqslant n)$.

Sea $(\Omega, \mathfrak{F}, P)$ un espacio de probabilidad, A un evento y $B_1, B_2, \ldots$ eventos disjuntos que parten a Ω, i.e. $\Omega = B_1 \cup B_2 \cup \ldots$ De la Definición 1.1 y de 2.13 se obtiene la **fórmula de probabilidad total**:

$$\boxed{P(A) = \sum_{k=1}^{\infty} P(A|B_k)P(B_k).} \tag{2.21}$$

Una generalización de 2.21 son las siguientes igualdades útiles. Para v.a's. X y Y y $B \subset \mathbb{R}$ se tiene que:

$$\boxed{P(X \in B) = \sum_{k=1}^{\infty} P(Z \in B|Y = y_k)P(Y = y_k);} \tag{2.22}$$

si Y es discreta, y

$$\boxed{P(Z \in B) = \int_{-\infty}^{\infty} P(Z \in B|Y = y) f_Y(y) dy;} \tag{2.23}$$

si Y es a.c. con densidad f_Y.

Nota 2.2: (a) La probabilidad condicional $P(Z \in B|Y = y)$ de 2.23, en estos casos simples podría interpretarse como $\lim_{\Delta y \to 0} P(Z \in B|Y \in (y - \Delta y, y + \Delta y))$. El lector puede encontrar la definición general el [2].

(b)De 2.16 sigue que si en 2.22 o en 2.23 $Z = \varphi(X, Y)$ y las v.a's. X y Y son independientes, entonces:

$$\boxed{\begin{aligned} P(Z \in B|Y = y) &= P[\varphi(X, y) \in B|Y = y] \\ &= (\textit{por independencia}) = P[\varphi(X, y) \in B]. \end{aligned}} \tag{2.24}$$

(c) Según 2.16, $P(Z \in B|Y = y) = P(Z \in B)$ si las v.a's. Z e Y son independientes.

Aplicando 2.23 y 2.24 a $Z = X + Y$ y $P(X + Y \leqslant x)$, donde las v.a's. X y Y son idependientes con densidades f_X, f_Y respectivamente, podemos observar que $P(X + Y \leqslant x|Y = y) = P(X \leqslant x - y|Y = y) =$ (*por independencia*) $= P(X \leqslant x - y) = F_X(x - y)$ y llegar al siguiente teorema:

TEOREMA 2.1. *Para v.a's. X, Y independientes y absolutamente continuas se tiene (la fórmula de convolución):*

$$\boxed{f_{X+Y}(x) = \int_{-\infty}^{\infty} f_X(x-y) f_Y(y) dy.} \tag{2.25}$$

Luego, si aplicamos 6 y 2.25 obtendremos el:

TEOREMA 2.2. *Sean (con $n \geqslant 1$) $\eta_1, \eta_2, ..., \eta_n$ v.a's. i.i.d. y $\eta_k \sim Norm(0.1)$, ($k = 1, 2, ..., n$). Entonces:*

$$\boxed{\frac{\eta_1 + \eta_2 + \cdots + \eta_n}{\sqrt{n}} \sim Norm(0.1).} \tag{2.26}$$

Ejercicios 2

2.1 Sean A y B eventos, y $P(B) > 0$, $P(\overline{B}) > 0$. Muestre que si

$$P(A|B) > P(A), \tag{2.27}$$

entonces $P(A|\overline{B}) < P(A)$.

(En el caso de 2.27 se dice que B es favorable para A).

2.2 Muestre que en el Ejemplo 2.1 (a), X y Y son independientes, pero en el Ejemplo 2.1 (b), X y Y son dependientes (por ejemplo: si $Y = r$ resultará que $X = 0$).

2.3 Encuentre $F_{\rho,\theta}$ y $f_{\rho,\theta}$ para las v.a's. ρ, θ en el Ejemplo 2.1 (b) y establezca que éstas son independientes.

2.4 (a) Sean X y Y v.a's. i.i.d. con valores en el conjunto $\{0.1.2.3, ...\}$. Demuestre que $P(X+Y$ *sea un número par*$)\geqslant 1/2$.

(b) Supongamos que X y Y son v.a's. independientes con la distribución de Poissón, i.e. $P(X = k) = P(Y = k) = \frac{\lambda^k}{k!}e^{-\lambda}$, $k = 0, 1, 2, ...$ $(0! := 1)$. Para $\lambda = 1/2$ calcule $P(X + Y$ *sea un número par*).

SUGERENCIA: Para el inciso (b), deduzca de 2.21 que $X + Y$ tiene la distribución de Poissón con parámetro $\lambda = 1$. Para calcular el valor de la serie $\sum_{k=0}^{\infty} \frac{1}{(2k)!}$ evalúe el desarrollo de Taylor en el punto $x = 1$ para $f(x) = e^x + e^{-x}$.

Resp. : (b) ≈ 0.56767.

2.5 Un dado simétrico se lanza n veces. Considere el evento $A_n := \{$*sale por lo menos un "6"*$\}$ para demostrar que $P(A_n) \to 1$ cuando $n \to \infty$.

2.6 Una moneda simétrica se lanza 100 veces. Sea A:= "sólo salen 'soles'". Establezca que $P(A) = \frac{1}{2^{100}}$, i.e. A es un "evento raro". Sin embargo, cualquier otro resultado particular de 100 lanzamientos tiene la misma probabilidad de $\frac{1}{2^{100}}$.

2.7 Sean $X_1,\ X_2,\dots,X_n$ v.a's. independientes y $X_k \sim Exp(\lambda_k)$, $k = 1,\ 2,\dots,n$. Demuestre que $\min(X_1, X_2,\dots,X_n) \sim Exp(\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_n)$.

2.8 Sean $r > 0$ un número arbitrario pero fijo y $S_r := \{\overline{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$ t.q. $(x_1^2+x_2^2+\cdots+x_n^2)^{1/2} \leqslant r\}$; sean también $X_1,\ X_2,\dots,X_n$ v.a's. i.i.d. con la distribución $Norm(0.1)$. Demuestre que $\lim_{n\to\infty} P((X_1, X_2, ..., X_n) \in S_r) = 0$.

2.9 El número Y de partículas cósmicas que golpean al contador de Geiger (en una unidad pequeña de tiempo) es una v.a. con la distribución de Poissón con párametro $\lambda = 1$ (véase el Ejercicio 2.4 (b)). Cada partícula se registra por el contador con la probabilidad $2/3$ (independiente para distintas partículas). Calcule $P(X = 0)$, donde X es el número de partículas registradas.

$\underline{Resp.}: e^{-2/3} \approx 0.51342$.

2.10 Sean η_1, η_2 v.a's. i.i.d. con la distribución $Norm(0.1)$. Demuestre que la densidad de la v.a. $X = \frac{\eta_1}{\eta_2}$ está dada por la fórmula (la densidad de Cauchy):

$$\boxed{f_X(x) = \frac{1}{\pi(x^2+1)}, \qquad x \in \mathbb{R}.} \tag{2.28}$$

3. Esperanza y Varianza

Sea X una v.a. tal que

$$\int_\Omega |X(\omega)| dP(\omega) < \infty \tag{3.29}$$

(la integral de Lebesgue); el **número real** $a \equiv EX := \int_\Omega X(\omega) dP(\omega)$ se llama **la esperanza de** X (o **el promedio de** X). Haciendo el cambio de variables: $x = X(\omega)$ obtenemos:

$$\boxed{EX = \sum_{k=1}^{\infty} x_k P(X = x_k);\ \text{si } X \text{ es discreta.}} \tag{3.30}$$

$$\boxed{EX = \int_{-\infty}^{\infty} x f_X(x) dx;\ \text{si } X \text{ es a.c.}} \tag{3.31}$$

Ejemplo 3.1. Se lanza un dado simétrico (bien balanceado). Sea X el puntaje que sale. Por 3.30 tenemos que $EX = \sum_{k=1}^{6} kP(X = k) = \sum_{k=1}^{6} k(\frac{1}{6}) = 3.5$ (que es fracción a pesar de que los valores de X son enteros).

Nota 3.1: Por 3.29 la suma o la integral en 3.30 y 3.31 debe converger **absolutamente**. Más adelante veremos que no todas las v.a. tienen esperanza. Por otro lado, cuando $X \geqslant 0$ (toma valores no negativos) y en 3.30 o en 3.31 la suma o la integral diverge (i.e. si las sumas parciales tienden a infinito), convendremos en definir: $\boxed{EX = \infty}$.

Sean X una v.a. y $g : \mathbb{R} \to \mathbb{R}$ una función, tales que $Y = g(X)$ es una v.a. (con g continua, como ejemplo particular). Es claro que Y toma el valor $g(x)$ cuando $X = x$. Por 3.30 y 3.31 (suponiendo que $Eg(X)$ existe) tendremos que:

$$\boxed{Eg(X) = \sum_{k=1}^{\infty} g(x_k)P(X = x_k)} \tag{3.32}$$

$$\boxed{Eg(X) = \int_{-\infty}^{\infty} g(x)f_X(x)dx} \tag{3.33}$$

(para X discreta y a.c. respectivamente).

Un caso particular importante es cuando $g(X) = (X - EX)^2$.

La **varianza**, $Var(X)$ de la v.a. X es el número no negativo que se define por:

$$\boxed{Var(X) := E(X - E)^2 \equiv E(X^2) - (EX)^2} \tag{3.34}$$

(suponiendo que $E(X - EX)^2 < \infty$).

Las propiedades de la esperanza (de las cuales la más importante es la linealidad) y la varianza son consecuencia de las propiedades comunes de las integrales.

Proposición 3.1. *Sean X y Y v.a's. para las cuales $Var(X)$ y $Var(Y)$ (y por consecuencia EX y EY) existen y $c \in \mathbb{R}$ una constante (interpretada como la v.a. con el único valor c). Entonces:*

E	Var
$Ec = c$	$Var(c) = 0$
$E(cX) = cEX$	$Var(cX) = c^2Var(X)$
	$Var(X + c) = Var(X)$
$E(X + Y) = EX + EY$	$Var(X + Y) = Var(X) + Var(Y)$*(*)*
	(si X e Y son independientes)
$EX \geqslant 0$ *si* $X \geqslant 0$	$Var(X) = 0$ *sys* $P(X = c) = 1$
$\lvert EX\rvert \leqslant E\lvert X\rvert$	*(para alguna constante c)*

Nota 3.2: (a) En el caso general (con X, Y dependientes), puede ser que $Var(X + Y) \neq Var(X) + Var(Y)$ (Véase el Ejercicio 3.1). En efecto, la propiedad (*) se obtiene de forma simple del

siguiente teorema que a su vez puede demostrarse usando 2.17 o 2.19.

: (b) $\sigma(x) = \sqrt{Var(X)}$ se denomina: la **desviación estándar de** X.

TEOREMA 3.1. *Sean X, Y v.a's. independientes para las cuales existen EX y EY. Entonces:*

$$\boxed{E(X \cdot Y) = (EX)(EY).}$$

EJEMPLO 3.2. (a) Para $X \sim Bern(p)$ tenemos por 3.30 que $EX = 0(1-p) + 1 \cdot p = p$, y de 3.34 y 3.32, $Var(X) = p(1-p)$.

(b) Por la Proposición 3.1 para $Y_n = X_1 + ... + X_n \sim Bin(n,p)$, $EY_n = np$ y $Var(Y_n) = np(1-p)$.

EJEMPLO 3.3. (a) Para $X \sim Exp(\lambda)$ por 3.31, 3.34 y 3.33, $EX = 1/\lambda$ y $Var(X) = 1/\lambda^2$.

(b) Para $X \sim Norm(a, \sigma)$, $EX = a$ y $Var(X) = \sigma^2$

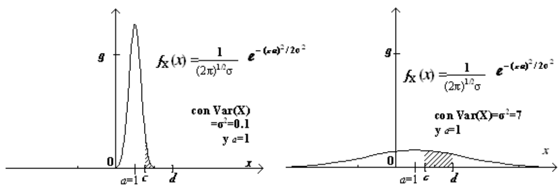

FIGURE 10. Esperanza y varianza en la v.a. Normal.

Nota 3.3: Mientras que la esperanza EX es una indicadora del "centro de los valores de la v.a. X", la varianza $Var(X)$ es una medida de "disperción de los valores". Las áreas resaltadas en la figura 10 representan (por 1.8) las probabilidades $P(X \in [c,d])$ para un intervalo $[c,d]$. Si este intervalo es distante de la esperanza $a = 1$, entonces $P(X \in [c,d])$ es casi cero cuando la varianza es pequeña (imagen de la izquierda en la figura10); la misma probabilidad $P(X \in [c,d])$ es bastante distinta de cero cuando la varianza es mayor (imagen de la derecha).

Es importante acentuar que para cualquier $\epsilon > 0$ (puede ser pequeña) se tiene que:

$$\boxed{\begin{array}{c} Cuando \quad Var(X) \to 0, \\ P(EX - \varepsilon \leqslant X \leqslant EX + \varepsilon) = \\ \textit{el área resaltada en la figura siguiente}\to \\ \int_{-\infty}^{\infty} f_X(x)dx = (por \quad 1.8) = P(-\infty < X < \infty) = 1 \end{array}} \tag{3.35}$$

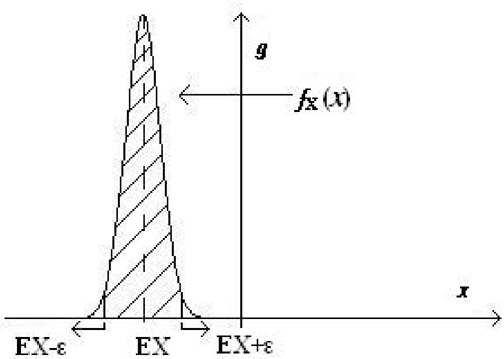

FIGURE 11. La esperanza con pequeña varianza.

Es decir, anulando la $Var(X)$, la v.a. X con probabilidad casi igual a 1 toma los valores muy cercanos al valor promedio EX.

EJEMPLO 3.4. Supongamos que hay que medir una magnitud física a desconocida. Para disminuir el error tal magnitud se mide n veces obteniendo los resultados: $X_1, X_2, ..., X_n$. Frecuentemente es razonable considerar que son v.a's. i.i.d. con $EX_k = a$, $Var(X_k) = \sigma^2 < \infty$, $(k = 1, ..., n)$. (Notemos que de la identidad de distribuciones resulta la igualdad de esperanzas y varianzas). Como estimación del valor desconocido de a se usa la v.a. S_n/n, donde

$$\boxed{S_n = X_1 + X_2 + \cdots + X_n.} \tag{3.36}$$

Con respecto al error de la estimación: $\delta_n := \frac{S_n}{n} - a$, se tiene que (consulte la Proposición 3.1):

$$\boxed{E\delta_n = \frac{1}{n}E(X_1 + \cdots + X_n) - Ea = \frac{na}{n} - a = 0\,,} \tag{3.37}$$

y $Var(\delta_n) = Var(\frac{S_n}{n} - a) = Var(\frac{S_n}{n}) = \frac{1}{n^2}Var(X_1 + \cdots + X_n) =$ (por independencia) $= \frac{1}{n^2}n\sigma^2 = \frac{\sigma^2}{n}$, i.e.

$$\boxed{Var(\delta_n) = \frac{\sigma^2}{n}.} \tag{3.38}$$

Por consiguiente, al aumentar el número de las mediciones resulta que $Var(\delta_n) \to 0$, y por esto los valores del estimador $\frac{S_n}{n}$ son cercanos al valor desconocido a (con probabilidad igual a 1, como lo veremos al estudiar la ley de los grandes números en õ6).

EJEMPLO 3.5. *(Nos muestra la importancia del uso de procesos aleatorios en algunos modelos.)*

Un señor maneja de su casa a la oficina (en días hábiles) con una velocidad aleatoria V que tiene (por suposición) la **densidad uniforme**:

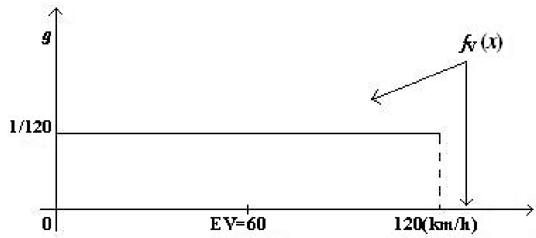

FIGURE 12. Densidad uniforme.

Usando 3.31 es fácil obtener que la velocidad promedio es $EV = 60(km/h)$. Supongamos que la distancia entre la casa y la oficina es $S = 30km$ ¿Cuál es el tiempo promedio por viaje?

Denotemos por T la v.a. que representa el tiempo (aleatorio, puesto que la velocidad es aleatoria) de un viaje. Entonces la respuesta a la pregunta se tendrá al estimar ET. Como

$$S = VT, T \quad = S/V, \tag{3.39}$$

intuitivamente se pienza que $ET = 0.5(hr.)$, que también se obtiene de 3.39 si aplicamos la igualdad:

$$ET = E(\frac{S}{V}) = \frac{ES}{EV} = \frac{30}{60}. \tag{3.40}$$

Sin embargo, vamos a ver que la segunda igualdad en 3.40 es falsa. Considerando 3.39 y 3.33 y tomando $g(x) = 1/x$ tenemos que $ET = 30$

$$\begin{aligned} E(1/V) &= 30 \int_0^{120} \frac{1}{x}\frac{1}{120}dx = \frac{1}{4}\int_0^{120}\frac{dx}{x} \\ &= \frac{1}{4}\lim_{\varepsilon\to 0}\int_\varepsilon^{120}\frac{dx}{x} = \frac{1}{4}\lim_{\varepsilon\to 0}[\ln(120) - \ln(\varepsilon)] = \infty \end{aligned}$$

(pues $\ln(x) \to -\infty$ cuando $x \to 0^+$).

Es necesario entender por qué en lugar de 0.5 hr hemos obtenido el resultado absurdo $ET = \infty$. La explicación es la siguiente: nuestro modelo del manejo con velocidad aleatoria no es correcto pues la aplicación de las ecuaciones 3.39 es válida cuando la velocidad V es **constante** durante todo el viaje. Se ha obtenido el resultado $ET = \infty$ por la posibilidad de valores pequeños para velocidades constantes durante todo el viaje, por ejemplo, una velocidad constante de 0.1 mm/h.

En realidad la velocidad V es una v.a. que depende del tiempo que transcurre durante el viaje: t (nula frente a un semáforo en rojo y grande en otros instantes). Es decir, $V = V(t, \omega)$ es un **proceso aleatorio (estocástico)**. Al cambiar 3.39 por una ecuación con velocidad variable:

$$\boxed{S = \int_0^T V(\omega, t)dt,} \tag{3.41}$$

podemos ver que $T \to \infty$ cuando $S \to \infty$, y bajo ciertas condiciones, el promedio en el tiempo $\frac{1}{T}\int_0^T V(\omega, t)dt$ se acerca al promedio en ω: $EV(\omega, t) = 60(km/h)$ cuando $T \to \infty$. O bien, por 3.41, $\frac{S}{T} \approx EV = 60$, o $T \approx \frac{S}{60}$ que resulta del hecho de que la v.a. T toma valores cercanos al valor constante $S/60$. Por lo tanto $ET \approx E(S/60) = 0.5$ (hr). Ésto recupera nuestra suposición intuitiva.

EJERCICIOS 3

3.1 Sea $X = \eta \sim Norm(0.1)$ y $Y = -X$. Muestre que $Var(X + Y) = 0$, pero $Var(X) + Var(Y) = 2$.

3.2 Sea X v.a. a.c. **simétrica** i.e. $f_X(-x) = f_X(x)$, $x \in \mathbb{R}$.
  (a) Muestre que si $E|X| < \infty$ (i.e. la esperanza existe y es finita), entonces $EX = 0$.
  (b) Para la densidad simétrica de Cauchy en 2.28, muestre que $\int_{-c}^{c} xf_X(x)dx = 0$ para cada $c > 0$, pero $E|X| = \infty$ y por lo tanto EX no existe.

3.3 Encuentre un ejemplo de dos v.a's. no negativas X, Y tales que $P(0 \leqslant X \leqslant Y) = 1 - 10^{-100}$, pero $EX = \infty$ y $EY < \infty$.
  SUGERENCIA: Considere $X = |\xi|$, donde ξ tiene la densidad de Cauchy 2.28.

3.4 Un dado simétrico se lanza 6 veces. ¿Cuál es el número esperado de caras que no salgan ni una vez?
  $\underline{Resp.}$: $5^6/6^5 \approx 2.000939$

3.5 Sea X la v.a. con la distribución de Poissón con $\lambda = 1$ (véase el Ejercicio 2.4 (b)). Calcule $E(\frac{1}{1+X})$.
  $\underline{Resp.}$: $1 - e^{-1} \approx 0.632121$.

3.6 Cierto río tiene crecimientos anuales. Supongamos que la marca del nivel bajo se sitúa en $1m$ y que la marca del crecimiento X es la v.a. con la siguiente f.d.:

$$F_X(x) = \begin{cases} 1 - 1/x^3, & \text{si } x \geqslant 1, \\ 0, & \text{si } x < 1. \end{cases}$$

Los daños materiales (representados en millones de pesos) por crecida X, se dan por la v.a. $Y = g(X)$ con:

$$Y = \begin{cases} 0, & \text{si } x \leqslant 2, \\ 0.7(X-2)^2, & \text{si } x > 2. \end{cases}$$

Calcule los daños promedio EY.

<u>*Resp.*</u> : 0.875 millones de pesos.

3.7 Sean X y Y v.a's. i.i.d. con densidad:

$$f_X(x) = \begin{cases} \frac{1}{2\sqrt{2}\sqrt{x}}, & \text{si } 0 < x < 2, \\ 0, & \text{si } x \notin (0.2). \end{cases}$$

Y sea $T = \max(X, Y)$. Encuentre F_T, f_T y calcule ET.

<u>*Resp.*</u> : $ET = 0.666667$

3.8 En el movimiento térmico de un gas en equilibrio, el módulo de la velocidad de cada molécula es la v.a. V con la densidad de Maxwell:

$$f_V(x) = \begin{cases} \frac{4}{\sqrt{\pi}}\lambda^{3/2}x^2 e^{-\lambda x^2}, & \text{si } x > 0, \\ 0, & \text{si } x \leqslant 0, \end{cases}$$

donde $\lambda = \frac{m}{2\varkappa T}$, m es la masa de la molécula, $\varkappa$ es la constante de Bolzmann y T es la **temperatura del gas**. Calcule la energía cinética promedio: $E(\frac{mV^2}{2})$.

<u>*Resp.*</u> : $\frac{3}{2}\varkappa T$ (Es decir, el calor o frío que sentimos es proporcional a la energía promedio de las moléculas del aire.)

3.9 Un "optimista" juega diariamente el juego de azar que se describe en el Ejemplo 5.2 de õ5 (pagando 100 pesos en cada ocasión que participa). Su finalidad es ganar en un día más de 1 millón de pesos. Sea $N := \{$el número de días hasta la primera vez que gana una cantidad mayor a 1 millón de pesos$\}$. Calcule EN.

Sugerencia: Primero, estime $p = P$(ganar más de un millón en un solo juego). Después muestre que:

$$P(N = n) = p(1-p)^{n-1}, \quad n = 1.2, \dots \tag{3.42}$$

(que es la **distribución geométrica**). Finalmente, usando 3.30 verifique que $EN = 1/p$.

<u>*Resp.*</u> : $EN = 524288$ *días* ≈ 1436 *años.* (Hay pocas chances de sobrevivir hasta este afortunado día, pero suponiendo que sucede, hay que pagar la participación en todos los juegos hasta ese día que es en promedio 52 millones de pesos. Sin embargo, el juego es "extremadamente favorable" de forma que con un capital de 100 pesos, puede "ganarse en promedio" un capital infinito (vea el Ejemplo 5.2).

De un teorema límite de la teoría de probabilidad seguirá que con $n = 524288$ repeticiones del juego pueden, en total, ganarse alrededor de 10 millones de pesos.)

4. Condicionamiento en esperanza y proceso de riesgo

Sean $B \subset \mathbb{R}$ un subconjunto y X una v.a. dados. Para v.a's. $Z = Z(X)$ de la forma:

$$Z = I_B := \begin{cases} 1, & \text{si } X \in B, \\ 0, & \text{si } X \notin B, \end{cases} \tag{4.43}$$

se tiene que $EZ = 1P(Z = 1) + 0P(Z = 0) = P(Z = 1) = P(X \in B)$. Si usamos la fórmula 2.22 para tales v.a's. y luego nos extendemos a v.a's. generales (aproximando éstas últimas por funciones indicadoras del tipo 4.43), llegaremos a las siguientes versiones de 2.22 y 2.23 en las cuales las probabilidades han sido cambiadas por esperanzas:

$$\boxed{EZ = \sum_{k=1}^{\infty} E(Z|Y = y_k)P(Y = y_k);} \tag{4.44}$$

para el caso de una v.a. Y discreta. Y:

$$\boxed{EZ = \int_{-\infty}^{\infty} E(Z|Y = y) f_Y(y)dy,} \tag{4.45}$$

cuando Y es una v.a. a.c. con densidad f_Y.

Nota 4.1: No hemos definido las esperanzas condicionales que usamos en 4.44 y 4.45. En ejemplos próximos veremos cómo calcularlas en algunos casos particulares.

EJEMPLO 4.1. (¿***Cuánto tiempo hay que esperar para recibir una oferta mejor que la primera***?).

Una persona quiere vender su coche y recibe las ofertas sucesivas: $X_0, X_1, X_{2,\dots}$ que se supone son v.a's. i.i.d. (**no negativas**) con una densidad común f_X y f.d. común F_X. Supongamos que el vendedor ha rechazado la primera oferta X_0 y ha decidido esperar hasta una oferta mejor, i.e. hasta el primer n tal que $X_n > X_0$. Sea $N := \min\{n \geqslant 1 \textit{ t.q. } X_n > X_0\}$. Calcularemos el promedio EN del número de ofertas que hay que esperar. Como $N \in \{1.2, \dots\}$, por 3.30 tenemos que $EN = \sum_{k=1}^{\infty} kP(N = k) =$ *(ver Ejercicio 4.1)* $= \sum_{n=0}^{\infty} P(N > n)$. De igual forma, se obtiene que

$$\boxed{E(N|Y = y) = \sum_{n=0}^{\infty} P(N > n|Y = y).} \tag{4.46}$$

Escogemos $Z = N, Y = X_0$ en 4.46. Para $n = 1, 2, \dots$ $P(N > n|X_0 = y) = P(x_1 \leqslant y, X_2 \leqslant y, \dots, X_n \leqslant y|X_0 = y) =$ (por ser X_0 independiente de $X_1, X_2, \dots, X_n) = P(X_1 \leqslant y, X_2 \leqslant y, ..., X_n \leqslant y) =$ (por la independencia de $X_1, X_2, \dots, X_n) = P(X_1 \leqslant y)P(X_2 \leqslant y) \dots P(X_n \leqslant y) =$ (por 1.4) $= [F_X(y)]^n$ es decir:

$$\boxed{P(N > n|X_0 = y) = [F_X(y)]^n.} \tag{4.47}$$

Entonces, de 4.46 y 4.47 tenemos que

$$E(N|Y = y) = 1 + \sum_{n=1}^{\infty}[F_X(y)]^n = \sum_{n=0}^{\infty}[F_X(y)]^n = \frac{1}{1 - F_X(y)}, \text{ i.e.:}$$

$$E(N|Y = y) = \frac{1}{1 - F_X(y)}, \tag{4.48}$$

(suponiendo que $F_X(y) < 1$, para $y > 0$). Luego, por 4.45 y 4.48 (tomando en cuenta que $f_X(x) = 0$ para $x < 0$ ya que $X_0 > 0$),

$$\begin{aligned} EN &= \int_0^{\infty} \frac{1}{1 - F_X(y)} f_Y(y)dy = \int_0^{\infty} \frac{1}{1 - F_X(y)} d[1 - F_X(y)] \\ &= -\int_1^0 \frac{dz}{z} = \int_0^1 \frac{dz}{z} \end{aligned} \tag{4.49}$$

(pues $F_X(0) = 0$ y $F_X(1) \to 1$ cuando $x \to \infty$).

Los cálculos hechos en el Ejemplo 3.5 muestran que la última integral en 4.49 es igual a ∞. Por lo tanto, $EN = \infty$ y el número **promedio** de las ofertas que debe esperar el vendedor para llegar a una mejor que la primera es **infinito**. Este resultado no es optimista para el vendedor, pero en realidad, la hipótesis de independencia entre las ofertas no se cumple.

EJEMPLO 4.2. ***(Caminata aleatoria simple.)***

Un jugador con un capital inicial de $C_0 = m \geqslant 1$ pesos, apuesta sucesivamente (decimos: en los “instantes" $n = 1.2, ...$) en una serie de lanzamientos de una moneda simétrica, un peso por lanzamiento. Después de n lanzamientos su capital será

$$C_n = m + X_1 + X_2 + ... + X_n, \tag{4.50}$$

donde $X_1, X_2, ...$ son v.a's. i.i.d. con la distribución:

$$X_k = \begin{cases} 1, & \text{con probabilidad } 1/2, \\ -1 & \text{con probabilidad } 1/2. \end{cases}$$

Para C_n en 4.50 sea $T_m := \min\{n : C_n = 0\}$, el instante de la ruina del jugador. En el momento T_m la serie de apuestas se detiene. No es difícil demostrar que $P(T_m < \infty) = 1$ (i.e. que la serie de apuestas

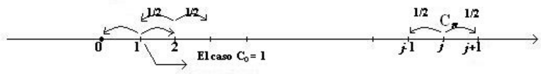

FIGURE 13. Caminata aleatoria simple.

seguramente terminará). Sin embargo mostraremos ahora que para cualquier $m \geqslant 1$, $E(T_m) = \infty$.

Si consideramos que $Z = T_m$ y $Y = X_1$ en 4.44 entonces $ET_m = E(T_m|X_1 = 1/2)1/2 + E(T_m|X_1 = 1)1/2 = 1/2[(1 + ET_{m+1}) + (1 + ET_{m-1})]$, es decir:

$$ET_m = \frac{1}{2}[(1 + ET_{m+1}) + (1 + ET_{m-1})]. \tag{4.51}$$

Para obtener 4.51 hemos usado las igualdades $E(T_m|X_1 = 1) = 1 + ET_{m+1}$ y $E(T_m|X_1 = -1) = 1 + ET_{m-1}$ que se cumplen puesto que bajo la condición de que, por ejemplo, $X_1 = 1$ pasaríamos al "nuevo capital inicial" $S_1 = m + 1$ y habríamos hecho una apuesta. Denotando $\varphi(m) := ET_m - ET_{m-1}$, para $m \geqslant 1$ (con $ET_0 = 0$), tenemos por 4.51 que

$$\varphi(m + 1) = \varphi(m) - 2, \quad m = 1, 2, ... \tag{4.52}$$

Iterando estas igualdades, vemos que $ET_m - ET_{m-1} = \varphi(m) = \varphi(1) - 2(m - 1) < 0$ para una m suficientemente grande. Por otro lado, como $T_m \geqslant T_{m-1}$, se cumple que $ET_m \geqslant ET_{m-1}$. La contradicción obtenida muestra que para el sistema de ecuaciones (4.10) y (4.9) no existe una solución tal que $ET_m < \infty$ para todo $m = 1, 2, ...$. Pero si para algún $m_0 \geqslant 1$ $ET_{m_0} = \infty$, entonces por 4.51 resultará que $ET_m = \infty$ para toda $m \geqslant 1$.

EJEMPLO 4.3. ***(El modelo clásico de riesgo.)***

Al denotar por $X(t)$, con $t \geqslant 0$ el capital corriente de una compañía de seguros, el modelo correspondiente se representa mediante la ecuación:

$$\boxed{X(t) = c + \gamma t - \sum_{n=1}^{N(t)} \xi_n, \quad t \geqslant 0,} \tag{4.53}$$

donde:

- $c = X(0)$ es el capital inicial de la compañía;
- $\gamma > 0$ es la prima acumulada (de todos los clientes) por unidad de tiempo;
- $N(t)$ es el número de reclamos en el intervalo $[0, t]$. En este modelo se supone que es un proceso de Poissón con parámetro $\lambda > 0$;
- $\xi_1, \xi_2, ...$ son los tamaños sucesivos de los reclamos que por suposición son v.a's. i.i.d. no negativas tales que $E\xi_k = a$.

Supondremos también que en este modelo, $N(t)$ no depende de $\xi_1, \xi_2, \ldots$(*)

Nota 4.2: El Proceso de Poissón con un parámetro $\lambda > 0$ es la familia de las v.a's. $N(t)$ que dependen del tiempo $t \geqslant 0$, tales que:

(a) $N(0) = 0$;

(b) $N(t) \sim Poisson(\lambda t)$ (la distribución de Poissón con parámetro λt, véase el Ejercicio 2.4 (b));

(c) para cualquier $0 < t_1 < t_2 \leqslant t_3 < t_4$, los **incrementos** de $N(t)$: $N(t_2) - N(t_1)$ y $N(t_4) - N(t_3)$ son independientes.

De (b) y 3.30 sigue que $EN(t) = \lambda t$, $t \geqslant 0$.

Calcularemos el capital corriente promedio:

$$EX(t) = c + \gamma t + E(\sum_{n=1}^{N(t)} \xi_n). \tag{4.54}$$

Por 4.44, con $Z = \sum_{n=1}^{N(t)} \xi_n$ y $Y = N(t)$ se tendrá:

$$\begin{aligned}
E\left(\sum_{n=1}^{N(t)} \xi_n\right) &= \sum_{k=1}^{\infty} E\left(\sum_{n=1}^{N(t)} \xi_n | N(t) = k\right) P(N(t) = k) \\
&= \sum_{k=1}^{\infty} E(\sum_{n=1}^{k} \xi_n | N(t) = k) P(N(t) = k) && \text{(por (*))} \\
&= \sum_{k=1}^{\infty} E(\sum_{n=1}^{k} \xi_n) P(N(t) = k) && \text{(por linealidad de la esperanza)} \\
&= a \sum_{k=1}^{\infty} k P(N(t) = k) && \text{(por 3.30)} \\
&= aEN(t) = a\lambda t
\end{aligned}$$

Usando ésto último y 4.54, resultará que

$$EX(t) = c + (\gamma - a\lambda)t; \tag{4.55}$$

o bien, que $EX(t)$ crece (linealmente) si

$$\boxed{\gamma > a\lambda,} \tag{4.56}$$

La condición 4.56 significa que por unidad de tiempo, la compañía gana en promedio γ más de lo que gasta, $a\lambda$ = el pago promedio por un reclamo $\times$ el número promedio de reclamos por unidad de tiempo.

A pesar de que en el caso 4.56, $EX(t) \to \infty$, hay chances no nulas de que una trayectoria de $X(t)$ en 4.53 bajara de tal modo que en algún

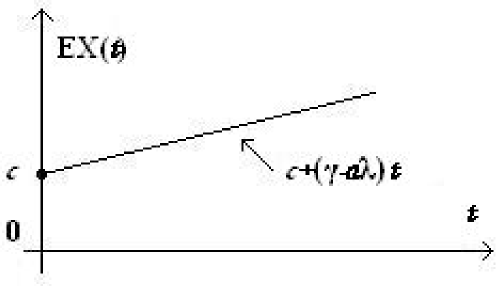

FIGURE 14. Ganancias esperadas.

instante ocurriera que $X(t) < 0$. En este caso estaríamos hablando de la ruina de la compañía.

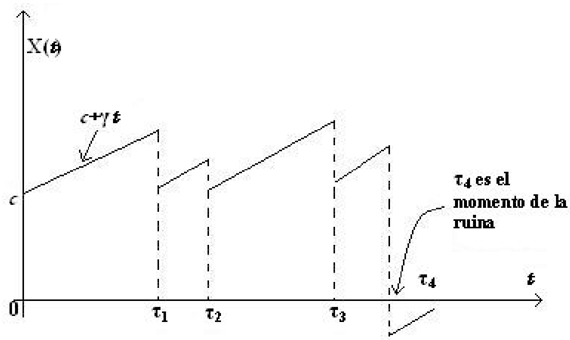

FIGURE 15. Instantes en que se reciben los reclamos.

En ésta imágen, $\tau_1, \tau_2, \ldots$ son los instantes (aleatorios) en que se reciben los reclamos sucesivos.

Entonces, a pesar de que con 4.56 la mayoría de las trayectorias de $X(t)$ nunca cruzan el eje $\overline{0t}$, $P(de\ ruina) := P(\inf_{t>0} X(t) < 0)$ regularmente es mayor que cero. En õ8 se discutirá un poco el problema (no fácil) de la estimación de $P(de\ ruina)$.

EJERCICIOS 4

4.1 Sea N una v.a. con valores en $\{0.1.2, \ldots\}$ demuestre que $EN = \sum_{k=0}^{\infty} P(N > k)$.

4.2 Sean $X_1, X_2, \ldots$ v.a's. i.i.d. con la distribución $Exp(\lambda)$, N una v.a. independiente de $X_1, X_2, \ldots$. Supongamos que $P(N = n) = 1/2^n$, $n = 1, 2, \ldots$y definimos la función $R(\lambda) := E(X_1 \cdot X_2 \cdot \ldots \cdot X_N)$, $\lambda > 0$. Determine los valores de $R(\lambda)$.

<u>*Resp.*</u> :

$$R(\lambda) = \begin{cases} \frac{1}{2\lambda-1}, & \text{si } \lambda > 1/2, \\ \infty, & \text{si } \lambda \leqslant 1/2. \end{cases}$$

4.3 Sean X, Y v.a's. idénticamente distribuídas.
(a) Encuentre un ejemplo que muestre que en general, $E(\frac{X}{X+Y}) \neq E(\frac{Y}{X+Y})$.
(b) Muestre que si además X y Y son independientes, entonces $E(\frac{X}{X+Y}) = E(\frac{Y}{X+Y})$.

4.4 En un juego de azar, un dado simétrico se lanza hasta la primera salida del "6". La ganancia se estima mediante $S_N = X_1 + X_2 + ... + X_N$, donde N es el número (aleatorio) de lanzamientos. Calcule la ganancia promedio $E(S_N)$.

<u>*Resp.*</u> : $E(S_N) = 21$.

4.5 Imaginemos a una persona de 50 años de edad la cual tiene n órganos principales que pueden ser transplantados (como el corazón, hígado, córnea, etc.). Actualmente n se considera 20 aproximadamente. Supongamos que cada año (después de la edad de 50 años) deja de funcionar un órgano y se hace un transplante. Supongamos también que los órganos dejan de funcionar de forma independiente unos de otros y que no importa si son propios o transplantados.

Sea N la v.a. (con valores enteros) tal que a la edad de $50 + N$ años, por primera vez, los n órganos de la persona mencionada se han cambiado por órganos transplantados. Calcule EN.

<u>*Resp.*</u> : $EN = n(1 + \frac{1}{2} + \frac{1}{3} + ... + \frac{1}{n})$, $EN \approx 72$ *años* para $n = 20$.

4.6 En el Ejemplo 4.3 supongamos (no realistamente) que (en miles de pesos):

$$\xi_1 = \begin{cases} 1 & \text{con probabilidad} 1/2, \\ 2 & \text{con probabilidad} 1/2. \end{cases}$$

Encuentre la f.d. F_{X_t} de la v.a. X_t (para cada $t > 0$).

5. Esperanza geométrica

Definición 5.1. Sea $X > 0$ una v.a. (con valores positivos) tal que $E|\ln X| < \infty$. La *esperanza geométrica* E_gX de X se define como:

$$\boxed{E_gX := e^{E\ln X}.} \tag{5.57}$$

EJEMPLO 5.1. Sean $x_1, x_2, \ldots, x_m$ números positivos dados y $P(X = x_k) = 1/m$, $k = 1, 2, \ldots, m$. Entonces por 3.30,

$$EX = \frac{1}{m}(x_1 + x_2 + \cdots + x_m)$$

(el promedio aritmético) y por (5.1) y (3.4),

$$E_g X = \exp\{\frac{1}{m}\sum_{k=1}^{m} \ln x_k\} = \prod_{k=1}^{m} e^{\frac{1}{m}\ln x_k} = (x_1 \cdot x_2 \cdot \cdots \cdot x_m)^{1/m}$$

(el promedio geométrico).

EJEMPLO 5.2. *(La paradoja de Petersburgo)*

Consideremos el sigiente juego de azar. Una moneda simétrica se lanza hasta la primera salida del "sol". Sea X el número de lanzamientos. Es fácil calcular que para $k = 1, 2, ...$ $P(X = k) = \frac{1}{2^k}$. Cuando $X = k$ un jugador gana 2^k pesos. Según 3.32 la ganancia promedio es

$$EX = \sum_{k=1}^{\infty} 2^k \frac{1}{2^k} = \infty. \tag{5.58}$$

Por otro lado, por 5.57 y 3.32 $E_g X = \exp\{\sum_{k=1}^{\infty} \frac{k \ln 2}{2^k}\} = 2^{\sum_{k=1}^{\infty} \frac{k}{2^k}} = 2^2 = 4.$

Nota 5.1: Según la definición común, un juego de azar se denomina "justo" si el pago por la participación coincide con la ganacia promedio. La paradoja surge porque 5.58 supone que cualquier "pago por participación" finito es "injusto". Por otro lado, no muchas personas están de acuerdo en pagar 100 pesos para participar en tal juego. (Véa el Ejercicio 3.9 para ms información.)

La siguiente afirmación es la contraparte para la linealidad de la esperanza habitual.

TEOREMA 5.1. *Sean $n \geqslant 1$ y $X_1, X_2, ..., X_n$ v.a. positivas con $E|\ln X_k| < \infty$, $k = 1, 2, ..., n$. Entonces:*

$$\boxed{E_g(X_1 \cdot X_2 \cdot ... \cdot X_n) = E_g(X_1)E_g(X_2) \cdot ... \cdot E_g(X_n).} \tag{5.59}$$

La demostración es evidende, pues $\ln(X_1 \cdot X_2 \cdot ... \cdot X_n) = \ln X_1 + \ln X_2 + ... + \ln X_n$ y $E(Y_1 + Y_2 + ...) = EY_1 + EY_2 + ...$

EJERCICIOS 5

5.1 Sean X, Y v.a's. positivas e identicamente distribuidas. Muestre que

(a) $1 = \frac{E_g X}{E_g Y} = E_g(\frac{X}{Y})$.

(b) Si $X, Y \sim Exp(\lambda)$ y son independientes, entonces $1 = \frac{EX}{EY} \neq E(\frac{X}{Y}) = \infty$.

5.2 Sean $X_1, X_2, \ldots$ v.a's. a.c. e i.i.d. con densidad:

$$f_{X_1} = \begin{cases} \frac{1}{x\ln(2)}, & \text{si } 1 < x < 2, \\ 0, & \text{si } x\ notin (1.2). \end{cases}$$

Demuestre que con probabilidad 1: $(X_1 \cdot X_2 \cdot \ldots \cdot X_n)^{1/n} \to \sqrt{2}$ cuando $n \to \infty$.

5.3 (a) Encuentre un ejemplo de una v.a. positiva X tal que $EX < \infty$, pero $EX^2 = \infty$.

(b) Muestre que si $E_g X$ existe, entonces para cada $\alpha > 0$, $E_g(X^\alpha)$ existe y además $[E_g(X^\alpha)] = [E_g(X)]^\alpha$.

5.4 Encuentre un ejemplo de v.a. $X > 0$ y la constante c tales que $Y = X + c > 0$, $E_g X$ existe y es finita, pero $E_g(Y) = \infty$.

SUGERENCIA: Use la densidad del Ejercicio 5.2.

6. Ley fuerte de los grandes números y modelos simples de inversión óptima

Sean $X_1, X_2, \ldots$ v.a's. i.i.d. con esperanza común $\boxed{a = EX_k}$. Como en 3.36, denotamos:

$$\boxed{S_n = X_1 + X_2 + \ldots + X_n, \quad n = 1, 2, \ldots}$$

Si suponemos ademas que $Var(X_k) < \infty$, por 3.37 y 3.38 tendremos que $E\left(\frac{S_n}{n}\right) = a$, $Var\left(\frac{S_n}{n}\right) = \frac{Var(X_1)}{n} \to 0$ cuando $n \to \infty$. Es decir (consulte la Nota 3.3 y 3.35) para grandes n, casi todos los valores de la v.a. S_n/n están concentrados cerca del valor esperado a. Éste es el sentido de un grupo de teoremas conocidos como "leyes de los grandes números". Un resultado muy profundo y dificil de demostrar es el siguiente teorema (que no requiere la existencia de la varianza).

TEOREMA 6.1. *(L.F.G.N.) (Se recomienda consultar [2].) Sean $X_1, X_2, \ldots$ v.a's. i.i.d.. Tendremos que:*

(a) *Si $E|X_1| < \infty$, entonces*

$$P\left(\frac{S_n}{n} \to a \text{ cuando } n \to \infty\right) = 1; \tag{6.60}$$

aquí, $a = EX_1 (= EX_2 = EX_3 = \ldots)$.

(b) *Si $E|X_1| = \infty$, entonces:*

$$P\left(\text{la sucesión } \{S_n/n\} \text{ diverge}\right) = 1. \tag{6.61}$$

Nota 6.1: (a) El tipo de convergencia que se menciona en 6.60 se denomina: **convergencia con probabilidad 1** o también: **convergencia casi segura** (i.e. para todo $\omega \in \Omega_0$ con $P(\Omega_0) = 1$)

(b) En el Teorema 6.1 la independencia de sumandos X_k en $S_n = X_1 + X_2 + ... + X_n$ es escencial. (Véase el Ejercicio 6.1.)

EJEMPLO 6.1. Supongamos que se lanza sucesivamente un dado bien balanceado. Para $k = 1, 2, ...$ sean

$$X_k = \begin{cases} 1, & \textit{si en el k-ésimo lanzamiento sale ``6"}, \\ 0, & \textit{si en el k-ésimo lanzamiento sale otro puntaje.} \end{cases}$$

Entonces la v.a. $S_n = X_1 + X_2 + ... + X_n$ representa el número de salidas de "6" en los n primeros lanzamientos. Podemos suponer que $X_1, X_2, ...$ son i.i.d. con $EX_1 = 1 \cdot P(\textit{salga} \text{ ``6"}) + 0 \cdot P(\textit{no salga} \text{ ``6"}) = 1/6$. Por 6.60, con probabilidad uno, **la frecuencia del "6"** S_n/n se aproxima a $1/6 = P(\textit{salga} \text{ ``6"})$ (cuando $n \to \infty$). Se enfoca a este sentido la interpretación frecuentista de la probabilidad: al repetir "las pruebas" muchas veces, la frecuencia de un evento aproxima la probabilidad de éste evento.

EJEMPLO 6.2. Sean $X_1, X_2, ...$ v.a's. i.i.d. con la densidad de Cauchy: $f_X(x) = \dfrac{1}{\pi(1 + x^2)}$, $x \in \mathbb{R}$.

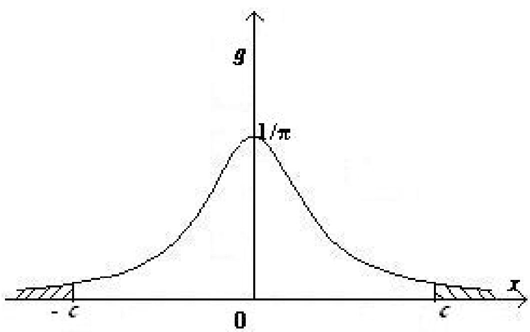

FIGURE 16. Densidad de Cauchy.

Por la simetría de ésta densidad, se podría pensar que $\frac{S_n}{n} = \frac{X_1+X_2+...+X_n}{n} \to 0$ cuando $n \to \infty$ (cada sumando X_k, $k = 1, 2, ..., n$ tiene las mismas chances de ser positivo o negativo). Sin embargo, (consulte el Ejercicio 6.2) $E|X_1| = \displaystyle\int_{-\infty}^{\infty} |x| \frac{dx}{\pi(1 + x^2)} = \infty$ y por 6.61 la sucesión $\{S_n/n, n = 1, 2, ...\}$ diverge con probabilidad 1. En particular, para cada $m \geqslant 1$

$$P\left(\sup_{n \geqslant m} |S_n/n| = \infty\right) = 1. \qquad (6.62)$$

Ahora bien usando el hecho de que

$$\int_{-\infty}^{\infty} f_X(x)dx = \textit{el área bajo la gráfica de la densidad} = 1$$

seleccionemos c tal que

$$P(|X| > c) = \int_{|x|>c} f_X(x)dx = \textit{el área resaltada} = 10^{-1000},$$

e introducimos las v.a's. $\widetilde{X}_1, \widetilde{X}_2, \ldots$ definidas de la forma:

$$\widetilde{X}_k = \begin{cases} X_k, & \text{si } |X_k| \leqslant c, \\ 0, & \text{si } |X_k| > c. \end{cases} \tag{6.63}$$

Las v.a's. $\widetilde{X}_1, \widetilde{X}_2, \ldots$ son i.i.d. (como las funciones de v.a's. i.i.d. $X_1, X_2, \ldots$, vea 2.20) y por la Proposición 3.1, $|E\widetilde{X}_k| \leqslant E|\widetilde{X}_k| \leqslant Ec = c < \infty$. Además, por simetría, $E\widetilde{X}_k = 0$. Por lo tanto el Teorema 6.1 (a) se cumple para las v.a's. $\widetilde{X}_1, \widetilde{X}_2, \ldots$ con $a = 0$. Aplicando 6.60 tendremos que $\frac{\widetilde{S}_n}{n} = \frac{\widetilde{X}_1+\ldots+\widetilde{X}_n}{n} \to 0$ cuando $n \to \infty$, con probabilidad 1. Comparando 6.62 con la ecuación anterior, podemos observar que los comportamientos asintóticos de S_n y $\widetilde{S}_n$ son muy diferentes. Por otro lado, por 6.63 se tiene que $P(\widetilde{X}_k \neq X_k) = P(|X_k| > c) = 10^{-1000}$, i.e. las v.a's. X_k y $\widetilde{X}_k$ son "prácticamente indistinguibles".

Proposición 6.1. *Sean $X_1, X_2, \ldots$ v.a's. positivas i.i.d. para las que existe $b := E_g X_1$. Entonces: $P((X_1 \cdot X_2 \cdot \ldots \cdot X_n)^{1/n} \to b) = 1$.*

Para hacer la demostracion es suficiente tomar el logaritmo y usar 6.60 y 5.57.

Ejemplo 6.3. ***(Un modelo simple de optimización de inversiones.)***

Al principio de un año ($t = 0$) una persona tiene el capital inicial $Y_0 > 0$ y planea usarlo dividiendo Y_0 en dos tipos de inversiones:

- La no riesgosa (como un depósito en el banco), donde por cada peso invertido obtendrá al final del año $\boxed{\alpha > 1}$ pesos;

- La riesgosa (como comprar acciones), donde por cada peso invertido resulta al final del año la cantidad aleatoria $X_1 > 0$ (que es ganancia si $X_1 > 1$ y es pérdida si $X_1 < 1$).

Supongamos que el inversionista escoge cierto número $q \in [0.1]$ (un parámetro "controlable") y gasta qY_0 por la inversión no riesgosa y $(1-q)Y_0$ por la riesgosa. Entonces al final del año (o al principio del siguiente ($t = 1$)) su capital (aleatorio) será:

$$Y_1 = \alpha q Y_0 + X_1(1-q)Y_0 = Y_0[q\alpha + (1-q)X_1]. \tag{6.64}$$

Al principio del siguiente año ($t = 1$) el inversionista usa el **mismo** valor de q para ahora repartir la cantidad Y_1 en los dos tipos de inversión y por lo tanto al inicio del año $t = 2$ su capital será:

$$Y_2 = Y_1[q\alpha + (1-q)X_2] = Y_0[q\alpha + (1-q)X_1][q\alpha + (1-q)X_2], \tag{6.65}$$

donde X_2 es la v.a. que representa el movimiento de precios de acciones en el segundo año (y se supone que α permanece constante en todos los años).

Suponiendo que el inversionista maneja su capital de la misma forma en los años $t = 2, 3, \ldots$, por 6.64 y 6.65 tendremos que al principio del $t-$*ésimo* año su capital será:

$$\boxed{Y_t = Y_0 \prod_{k=1}^{t}[q\alpha + (1-q)X_k], \quad t = 1, 2, \ldots} \tag{6.66}$$

SUPOSICIÓN 6.1. *(a) Las v.a's. $X_1, X_2, \ldots$ son i.i.d.;*
(b) Existe $a = EX_k$ y

$$\boxed{a > \alpha > 1;} \tag{6.67}$$

(i.e. en promedio, la inversión riesgosa es más lucrativa).

La idea general del problema de opitmización de inversiones en este modelo es la elección de una q que provoque un rápido crecimiento de Y_t cuando t aumente.

La aleatoriedad de Y_t no nos permite hacer esto para todas las trayectorias de Y_t simultaneamente. Es por ésto que necesitamos un criterio "promedio" de optimización que podemos escoger de diferentes maneras.

Pl. I $\equiv$ Planteamiento I. *(Maximización del capital promedio EY_t).*

Por 6.66 y el Teorema 3.1:

$$\boxed{EY_t = Y_0 \prod_{k=1}^{t}[q\alpha + (1-q)EX_k] = Y_0[q\alpha + (1-q)a]^t.} \tag{6.68}$$

Puesto que para $t = 1, 2, \ldots$ la función $g(x) = x^t$ con $x \geqslant 0$ es creciente, el máximo de la esperanza en 6.68 (¡para cada $t = 1, 2, \ldots$!) se alcanza en un punto donde la función lineal $\psi(q) := q\alpha + (1-q)a$ con $q \in [0.1]$, es maximal. La gráfica de $\psi(q)$ es la siguiente (véase 6.67):

El parámetro óptimo, en el sentido del Pl.I, resulta de elegir $q = q_0 = 0$, o bien, usar todo el capital en una inversión riesgosa. Al usar $q_0 = 0$, por 6.68 se tiene:

$$\boxed{EY_t = Y_0 a^t, \quad t = 1, 2, \ldots,} \tag{6.69}$$

lo que garantiza el crecimiento exponencial (con el máximo exponente posible) de la **riqueza promedio**. Veamos que en algunas situaciones dicho procedimiento de inversión podría ser "demasiado riesgoso".

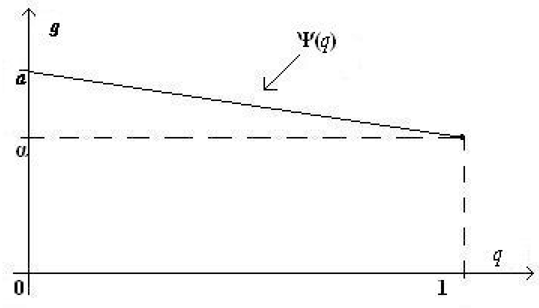

FIGURE 17. Gráfica de $\psi(q)$.

Sea, por ejemplo,

$$X_1 = \begin{cases} 3, & \textit{con probabilidad 0.95,} \\ 10^{-12}, & \textit{con probabilidad 0.05.} \end{cases}$$

Entonces $a = EX_1 \approx 2.85(>> \alpha \approx 1.03 - 1.1)$, y por 6.69,

$$\boxed{EY_t = Y_0(2.85)^t, \quad (\uparrow \infty).} \tag{6.70}$$

Por ejemplo:

$$\boxed{EY_{50} \approx 5.52 \cdot 10^{22} \quad pesos.} \tag{6.71}$$

Por otro lado, de 6.66 con $q = q_0 = 0$, $Y_t = Y_0 \prod_{k=1}^t X_k$. De aquí y de la Proposición 6.1 obtenemos que cuando $t \to \infty$, con probabilidad 1: $(Y_0)^{-1/t}(Y_t)^{1/t} \to E_g X_1$, donde (vease 5.57 y 3.32) $E_g X_1 = \exp\{\ln 3 \cdot 0.95 + \ln(10^{-12}) \cdot 0.05\} \approx e^{-0.33787}$. De aquí, para las t suficientemente grandes, se tiene con probabilidad 1 que:

$$\boxed{Y_t \approx Y_0 \ e^{-0.33787 \cdot t}.} \tag{6.72}$$

Consiguientemente, por 6.70 el **capital promedio** crece sin cota mientras que el **capital real** Y_t casi seguro (c.s.) (i.e. con probabilidad 1) **se anula** cuando $t \to \infty$. Por ejemplo, (considerando una vaga estimación) $Y_{50} \approx 4.63 \cdot 10^{-8} \approx 0$ pesos (compare con 6.71 y con 6.78 más adelante).

Pl. II $\equiv$ **Planteamiento II**. *(Maximización de la esperanza geométrica $E_g Y_t$ del capital.)*

De 6.66 obtenemos que $E_g Y_t = \exp\{E \ln(Y_0 \prod_{k=1}^{t} [q\alpha + (1-q)X_k])\}$

$= \exp\{\ln Y_0 + \sum_{k=1}^{t} E\ln[q\alpha + (1-q)X_k]\}$ =*(pues* $X_1, X_2, \ldots$ *tienen la misma distribución)* $= Y_0 \exp\{nE\ln[a\alpha + (1-q)X_1]\}$ i.e.

$$E_g Y_t = Y_0 \exp\{nE\ln[q\alpha + (1-q)X_1]\}. \tag{6.73}$$

Como e^x es una función creciente, para encontrar $\max_q E_g Y_t$ es suficiente buscar

$$\max_q\{E\ln[q\alpha + (1-q)X_1]\}. \tag{6.74}$$

SUPOSICIÓN 6.2.

$$E\left(\frac{1}{X_1^2}\right) < \infty. \tag{6.75}$$

La condición 6.75 garantiza que $E|\ln[q\alpha+(1-q)X_1]| < \infty$, $q \in [0.1]$ (i.e. las esperanzas involucradas en 6.73 y 6.74 existen) y que la función $\varphi(q) := E\ln[q\alpha+(1-q)X_1]$, $q \in [0.1]$ tiene la segunda derivada negativa como se muestra en la figura 18

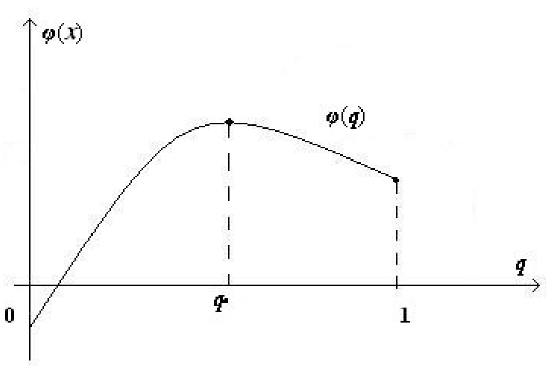

FIGURE 18. Máximo global.

Entonces existe un único $q_* \in [0.1]$ en el cual la función $\varphi(q)$ alcanza el máximo global, y como $\varphi(1) = \ln\alpha > 0$, entoces:

$$\mu := \varphi(q_*) = \max_{q\in[0.1]} E\ln[q\alpha + (1-q)X_1] > 0. \tag{6.76}$$

¿Cuáles serán los resultados al usar q_* en el procedimiento de inversión?

Sustituyendo q_* en 6.66 y aplicando la Proposicion 6.1 obtenemos que con probabilidad 1 $Y_0^{-1/t}(Y_t)^{1/t} \rightarrow E_g[q_*\alpha + (1-q_*)X_1] =$

(véase 6.76) $= e^{\mu}$, cuando $t \to \infty$. Por lo tanto, para toda t suficientemente grande, con probabilidad 1 tendremos:

$$\boxed{Y_t \approx Y_0 e^{\mu t} \ \ (\uparrow \infty).} \tag{6.77}$$

Además, debido a 6.76 (¡casi seguro!) la tasa del crecimiento de los capitales Y_t en 6.77 es máxima entre todas las posibles.

Ahora bien, calculando la esperanza en 6.66 con $q = q_*$ y tomando en cuenta que por 6.67 $q_*\alpha + (1 - q_*)a = \lambda > 1$, tendremos el siguiente crecimiento exponencial para los capitales promedio (al utilizar q_* para las inversiones planteadas): $EY_t = Y_0\lambda^t \uparrow$ cuando $t \to \infty$.

Regresando al ejemplo de arriba (con los valores 3 y 10^{-12} para X_k), podemos calcular que $q_* \approx 1/13$ (se guarda como 8% de capital en el banco, consulte el Ejercicio 6.7), $EY_{50} \sim 8.6 \cdot 10^{21}$ (menos que en 6.71). Sin embargo, con probabilidad 1:

$$Y_{50} \sim 2.35 \cdot 10^{17} \ \ pesos. \tag{6.78}$$

Nota 6.2: Utilizando un valor $q \in [0.1]$ para realizar el proceso de inversión descrito arriba, el término $G := q\alpha + (1 - q)X_k$ representa la ganancia (si $G > 1$) o la pérdida (si $G < 1$) durante un año (por cada peso invertido). En el Pl.I se busca $\max_q EG$ (véase 6.68) y en el Pl.II, $\max_q E\ln(G)$ (véase 6.74 y 6.76). Esto significa que en el segundo caso se usa la **función de utilidad**: $f(x) = \ln x$. Es decir, el "valor" de la ganancia de x pesos se estima no como x, sino como $\ln x$

EJERCICIOS 6

6.1 Sean $X_1 \sim Exp(\lambda = 1), X_1 = X_2 = ...$ Muestre que $\frac{S_n}{n} \nrightarrow EX_1 = 1$ cuando $n \to \infty$.

6.2 En el Ejemplo 6.2 muestre que $E|X_1| = \infty$.

6.3 Sean $X_1, X_2, ..., X_n$ los valores observados de v.a's. i.i.d. con la densidad exponencial con parámetro λ que se supone desconocido. Demuestre que para n grandes, $(X_1 + X_2 + ... + X_n)^{-1} \approx \lambda$, i.e., la parte izquierda de la igualdad nos sirve para hacer una estimación estadística de λ.

6.4 Sean $X_1, X_2, ...$ v.a's. i.i.d. con distribución de Bernoulli con $p = 1/2$. Muestre que

(a) $P(S_n/n \to 1/2) = 1$;

(b) $P(S_n/n = 1/2) \approx 1/\sqrt{2\pi n} \to 0$ cuando $n \to \infty$.

SUGERENCIA: Para (b) use la fórmula de Stirling: $n! \sim \sqrt{2\pi n}(n/e)^n$ cuando $n \to \infty$.

6.5 Para el Ejemplo 4.3 demuestre que con probabilidad 1, $N_t/t \to \lambda$ cuando $t \to \infty$.

6.6 Considerando los valores $X_1, X_2, ..., X_n$ de v.a's. i.i.d.con f.d. F_X, se define la f.d. empírica:

$\widehat{F}_n(x) := 1/n$ (*# de* X_k *tales que* $X_k \leqslant x$), $x \in \mathbb{R}$.

(a) Para los valores de $X_1, X_2, ..., X_n$ fijos, trace la gráfica de $\widehat{F}_n(x)$, $x \in \mathbb{R}$.

(b) Demuestre que para cada $x \in \mathbb{R}$ con probabilidad 1, $\widehat{F}_n(x) \to F_X(x)$ cuando $n \to \infty$.

6.7 En el Ejemplo 6.3 para la v.a.

$$X_1 = \begin{cases} 3, & \textit{con probabilidad 0.95,} \\ 10^{-12}, & \textit{con probabilidad 0.05;} \end{cases}$$

encuentre el valor (“la política") óptimo q_* (ver Pl.II).

Resp. : $q_* \approx 1/13$.

6.8 Para el Ejemplo 6.3 muestre que 6.75 garantiza la existencia de $E_g[q\alpha + (1 - q)X_1]$ para cada $q \in [0.1]$ y que la función $\varphi(q) = E \ln[q\alpha + (1 - q)X_1]$ tiene la segunda derivada negativa.

6.9 Sean $\eta_1, \eta_2, ...$ v.a's. i.i.d. con la distribución $Norm(0.1)$. Para los vectores aleatorios $\overline{X}_n = (\eta_1, \eta_2, ..., \eta_n)$ demuestre que con probabilidad 1,

$$|\overline{X}_n|/\sqrt{n} \to 1 \ \ cuando \ \ n \to \infty. \tag{6.79}$$

aquí $|\overline{X}_n| := \sqrt{\eta_1^2 + \eta_2^2 + ... + \eta_n^2}$.

Nota 6.3: En 6.79 vemos que para grandes n los valores del vector X_n están concentrados en un anillo estrecho de radio $\sqrt{n}$.

7. Métricas probabilísticas y teorema del límite central

DEFINICIÓN 7.1. Un par (X, d) se llama *espacio métrico* si:

- X es un conjunto;

- $d : X \times X \to [0, \infty]$ es una función denominada *métrica* que asigna a cada par de puntos $x, y \in X$, la *distancia* $d(x, y)$ entre x e y.

La métrica debe satisfacer los siguientes axiomas (para cualesquiera $x, y, z \in X$):

(1) $d(x, y) = 0$ sys $x = y$;

(2) $d(x, y) = d(y, x)$;

(3) $d(x, y) \leqslant d(x, z) + d(z, y)$.

Nota 7.1: Según la definición de arriba, puede ser que para algunos $x, y \in X$, $d(x, y) = \infty$.

DEFINICIÓN 7.2. Sean (X, d) un espacio métrico y $x, x_1, x_2, ... \in X$. Se dice que la sucesión $\{x_n\}$ *converge a* x si $d(x_n, x) \to 0$ cuando $n \to \infty$. En este caso escribimos: $\boxed{x_n \to^d x}$.

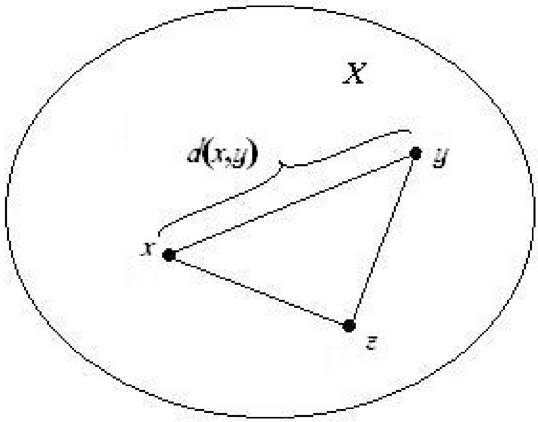

FIGURE 19. Espacio métrico.

Por ejemplo, si $X = \mathbb{R}$ y $d(x, y) = |x - y|$, entonces $x_n \to^d x$ equivale a la convergencia que se usa en el cálculo. Pero si

$$d_1(x, y) = \begin{cases} 1, & \text{si } x \neq y, \\ 0, & \text{si } x = y, \end{cases}$$

entonces $x_n \to^{d_1} x$ significa que a partir de un N, $x_n = x$.

DEFINICIÓN 7.3. Una métrica d en el espacio de todas las f.d's. (de todas las v.a's. con valores en $\mathbb{R}$) se llama *métrica probabilistica.*

Nota 7.2: (a) Cuando las v.a's. X e Y tienen las f.d's. F_X y F_Y respectivamente y d es una métrica probabilística, a partir de la igualdad $\boxed{d(X, Y) := d(F_X, F_Y)}$, se define "la distancia d entre las v.a's. X e Y".

(b) La Definición 7.3 es un caso particular de la definición general (puede consultar [11] y [15]) en la cual en general, la distancia $d(X, Y)$ se determina por la f.d. conjunta de X e Y: $F_{X,Y}$ (véase 2.17). Un ejemplo muy simple es: $d(X, Y) = E|X - Y|$.

EJEMPLO 7.1. ***(La métrica ρ uniforme o de Kolmogorov.)***

$$\boxed{\rho(X, Y) = \rho(F_X, F_Y) := \sup_{x\in\mathbb{R}} |F_X(x) - F_Y(x)|.} \tag{7.80}$$

La convergencia $X_n \to^{\rho} X$ significa que las f.d's. F_{X_n} aproximan uniformemente sobre x a la f.d. F_X.

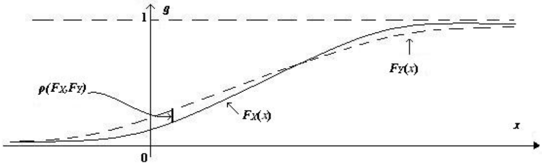

FIGURE 20. Aproximación uniforme.

DEFINICIÓN 7.4. Se dice que la sucesión de v.a's. X_n *converge a la v.a. X debilmente*, si $F_{X_n}(x) \to F_X(x)$ cuando $n \to \infty$ en cada punto x donde F_X es continua. En este caso escribimos: $\boxed{X_n \Rightarrow X}$.

El primer inciso de la siguiente afirmación se sigue directamente de la Definición 7.4. El segundo es transparente en el sentido intuitivo.

PROPOSICIÓN 7.1. (a) *Si $X_n \to^{\rho} X$, entonces $X_n \Rightarrow X$.*
(b) *Si F_X es continua, entonces $X_n \to^{\rho} X$ sys $X_n \Rightarrow X$.*

EJEMPLO 7.2. (Se usa en la teoría de confiabilidad.)
Sean $\xi_1, \xi_2, ...$ v.a's. no negativas i.i.d. tales que $F_{\xi_1}(x) = \lambda x + o(x)$ cuando $x \to 0$, donde $\lambda > 0$, $\frac{o(x)}{x} \to 0$ cuando $x \to \infty$. Para $n = 1, 2, ...$ sean $X_n := n \min(\xi_1, \xi_2, ..., \xi_n)$.
Ahora bien, para cualquier $x > 0$,

$$\begin{aligned}
P(X_n > x) &= P(\min\{\xi_1, \xi_2, \dots, \xi_n\} > x/n) \\
&= P(\xi_1 > \frac{x}{n}, \dots, \xi_n > \frac{x}{n}) && \text{(por independencia)} \\
&= P(\xi_1 > \frac{x}{n}) \cdot ... \cdot P(\xi_n > \frac{x}{n}) && \text{(por la identidad de f.d's.)} \\
&= [P(\xi_1 > \frac{x}{n})]^n = [1 - F_{\xi_1}(\frac{x}{n})]^n \\
&= [1 - (\frac{\lambda x}{n} + o(\frac{x}{n}))]^n \to e^{-\lambda x}
\end{aligned}$$

cuando $n \to \infty$. Por lo tanto, $F_{X_n}(x) = [1 - P(X_n > x)] \to (1 - e^{-\lambda x})$ cuando $n \to \infty$, o bien $X_n \Rightarrow X \sim Exp(\lambda)$.

EJEMPLO 7.3. ***(Métricas de Zolotarev [15].)*** Introduciremos ahora dos métricas probabilísiticas importantes para la teoría moderna de probabilidad:

$$\boxed{\zeta_2(X, Y) := \sup_{\varphi \in \mathcal{D}_2} |E\varphi(X) - E\varphi(Y)|} \tag{7.81}$$

$$\boxed{\zeta_3(X, Y) := \sup_{\varphi \in \mathcal{D}_3} |E\varphi(X) - E\varphi(Y)|,} \tag{7.82}$$

donde:

$$\mathcal{D}_2 := \{\varphi : \mathbb{R} \to \mathbb{R} \ t.q. \ |\varphi''(x)| \leqslant 1, x \in \mathbb{R}\} \tag{7.83}$$

$$\mathcal{D}_3 := \{\varphi : \mathbb{R} \to \mathbb{R} \ t.q. \ |\varphi^{(3)}(x)| \leqslant 1, x \in \mathbb{R}\} \tag{7.84}$$

Nota 7.3: $E\varphi(X)$ y $E\varphi(Y)$ se determinan por F_X y F_Y respectivamente (véase 3.32 y 3.33), por lo cual $\zeta_m(X,Y) \equiv \zeta_m(F_X, F_Y)$, $m = 2, 3$.

Proposición 7.2.

a) *Si* $EX = EY$, $EX^2 < \infty$ *y* $EY^2 < \infty$, *entonces*

$$\boxed{\zeta_2(X,Y) < \infty.} \tag{7.85}$$

b) *Si* $EX = EY$, $EX^2 = EY^2$, $E|X|^3 < \infty$ *y* $E|Y|^3 < \infty$, *entonces*

$$\boxed{\zeta_3(X,Y) < \infty.} \tag{7.86}$$

La demostración sigue del desarrollo de Taylor para φ. Por ejemplo en *b*):

$$\varphi(X) = \varphi(0) + \varphi'(0)X + \frac{\varphi''(0)}{2}X^2 + \frac{\varphi^{(3)}(\tau_X)}{6}X^3,$$
$$\varphi(Y) = \varphi(0) + \varphi'(0)Y + \frac{\varphi''(0)}{2}Y^2 + \frac{\varphi^{(3)}(\tau_Y)}{6}Y^3.$$

Al sustituir esto en 7.82 y aplicando 7.86 se obtiene:

$$\zeta_2(X,Y) \leqslant \sup_{\varphi \in \mathcal{D}_3} \left| E\frac{\varphi^{(3)}(\tau_X)}{6}X^3 - E\frac{\varphi^{(3)}(\tau_Y)}{6}Y^3 \right|. \tag{7.87}$$

Pero $|\varphi^{(3)}| \leqslant 1$ y $E|X|^3, E|Y|^3 < \infty$.

Nota 7.4: (a) Si $EX \neq EY$, entonces $\zeta_2(X,Y) = \infty$.

(b) Si $EX = EY$ o $EX^2 \neq EY^2$, entonces $\zeta_3(X,Y) = \infty$ (consulte el Ejercicio 7.2).

(c) Se conocen las siguientes cotas superiores para las métricas de Zolotarev (puede consultar [11]):

$$\boxed{\zeta_2(X,Y) \leqslant \int_{-\infty}^{\infty} |x||F_X(x) - F_Y(x)|dx,}$$

$$\boxed{\zeta_3(X,Y) \leqslant 1/2 \int_{-\infty}^{\infty} x^2|F_X(x) - F_Y(x)|dx.}$$

La utilidad de la métrica ζ_3 en el estudio de los teoremas del límite para la teoría de probabilidad, se explica con las siguientes propiedades de ζ_3.

PROPOSICIÓN 7.3. *Para cualquier $\alpha \geqslant 0$ se tiene:*

$$(a) \quad \boxed{\zeta_3(\alpha X, \alpha Y) \leqslant \alpha^3 \zeta(X, Y).} \tag{7.88}$$

$$(b) \quad \boxed{\zeta_3(X+Y, Z+Y) \leqslant \zeta_3(X, Z),} \tag{7.89}$$

para cualesquiera v.a's. X, Y, Z tales que Z no depende de X e Y.

Para demostrar (a) es suficiente observar que si $\varphi \in \mathcal{D}_3$ (vease 7.84), entonces
$\psi(x) := \frac{1}{\alpha^3}\varphi(\alpha x) \in \mathcal{D}_3$, por lo cual, $\sup\limits_{\varphi \in \mathcal{D}_3} |E\varphi(\alpha X) - E\varphi(\alpha Y)| =$
$\alpha^3 \sup\limits_{\varphi \in \mathcal{D}_3} |E[\dfrac{1}{a^3}\varphi(\alpha X)] - E[\dfrac{1}{a^3}\varphi(\alpha Y)]| = \alpha^3 \sup\limits_{\psi} |E\psi(X) - E\psi(Y)| \leqslant$
$\alpha^3 \zeta_3(X, Y)$.

Y en el caso de (b), fijamos un arbitrario $\varphi \in \mathcal{D}_3$ y por 4.45 (suponiendo, por ejemplo, que Y es a.c.)

$$|E\varphi(X+Y) - E\varphi(Z+Y)| = \left| \int_{-\infty}^{\infty} (E[\varphi(X+Y) - \varphi(Z+Y)]|Y = y) f_Y(y) dy \right|$$

(por la independencia de Y con X y Z)

$$= \left| \int_{-\infty}^{\infty} [E\varphi(X+y) - E\varphi(Z+y)] f_Y(y) dy \right|$$

$$\leqslant \int_{-\infty}^{\infty} |E\varphi(X+y) - E\varphi(Z+y)| f_Y(y) dy.$$

Pero para cada $y \in \mathbb{R}$ fijo, $\psi(x) := \varphi(x+y) \in \mathcal{D}_3$. Entonces $|E\varphi(X+y) - E\varphi(Z+y)| \leqslant \zeta_3(X, Z)$ y además $\int_{-\infty}^{\infty} f_Y(y) dy = 1$.

Uniendo 7.88 y 7.89 en la Proposición 7.3 y aplicando la inducción, llegamos al:

TEOREMA 7.1. *Sean $n \geqslant 1$, $\alpha \geqslant 0$, $X_1, X_2, ..., X_n$; $Y_1, Y_2, ..., Y_n$ v.a's. independientes. Entonces:*

$$\boxed{\zeta_3\left(\sum_{k=1}^{n} X_k, \sum_{k=1}^{n} Y_k\right) \leqslant \alpha^3 \sum_{k=1}^{n} \zeta_3(X_k, Y_k).} \tag{7.90}$$

TEOREMA 7.2. *(Consulte [11], [15].) Si $X_n \to^{\zeta_3} X$, entonces $X_n \Rightarrow X$.*

La demostración está basada en los hechos siguientes: que $F_{X_n}(x) = EI_{\{X_n \leqslant x\}}$ y que la función $g(y) := I_{\{y \leqslant x\}}$ (para x fijo) podría ser aproximada por funciones φ que pertenecen a $\mathcal{D}_3$.

De entre la gran clase de teoremas denominados "del límite central", el siguiente es una variante de la versión clásica del teorema del límite central de A. Lyapunov.

Teorema 7.3. *(TLC) Sean $X_1, X_2, \ldots$ v.a's. i.i.d. para las cuales:*

$$E|X_k|^3 < \infty, \tag{7.91}$$

$$y \quad a = EX_k, \quad 0 < \sigma^2 = Var(X_k).$$

Para $n = 1, 2, \ldots$ se definen las sumas normadas:

$$\boxed{Z_n := \frac{S_n - na}{\sigma\sqrt{n}}.} \tag{7.92}$$

donde $S_n = X_1 + X_2 + \ldots + X_n$. Entonces:

$$\boxed{\zeta_3(X_n, \eta) \leqslant \frac{\zeta_3(\frac{X_1 - a}{\sigma}, \eta)}{\sqrt{n}} \quad n = 1, 2, \ldots} \tag{7.93}$$

donde $\eta \sim Norm(0.1)$ y $\zeta_3(\frac{X_1-a}{\sigma}, \eta) < \infty$.

De la Proposición 7.1 y del hecho de que

$$\boxed{F_\eta(x) \equiv \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{t^2/2} dt, \quad x \in \mathbb{R}} \tag{7.94}$$

es una f.d. continua se siguen los incisos (a) y (b) de la siguiente afirmación (el inciso (c) requiere una demostración aparte y no es sencilla, puede consultar [11]).

Corolario 7.1. *Bajo las hipótesis del Teorema 7.3:*

(a) $\rho(Z_n, \eta) \to 0$ *cuando* $n \to \infty$ *(véase 7.80).*

(b) $Z_n \Rightarrow \eta$*, es decir, para cada* $x \in \mathbb{R}$

$$\boxed{F_{Z_n}(x) \to \Phi(x) \text{ cuando } n \to \infty.} \tag{7.95}$$

(c)

$$\boxed{\rho(Z_n, \eta) \leqslant \frac{c}{\sqrt{n}} \max\left\{\rho(\frac{X_1 - a}{\sigma}, \eta), \zeta_3(\frac{X_1 - a}{\sigma}, \eta)\right\},} \tag{7.96}$$

$n = 1, 2, 3, \ldots$ *donde c es alguna constante.*

Nota 7.5: (a) El objetivo de normalizanción en 7.92 es obtener $EZ_n = 0$, $Var(Z_n) = 1$ en correspondencia con $E\eta = 0$, $Var(\eta) = 1$ para la v.a. $\eta \sim Norm(0.1)$ del límite.

(b) Las afirmaciones (a) y (b) en el Corolario 7.1 son ciertas sin la condición 7.91 (pues es suficiente suponer que $Var(X_k) < \infty$). Sin embargo, para obtener 7.93 y 7.96 la hipótesis de (7.12) es esencial. Además hay casos particulares en que $\rho(Z_n, \eta) \geqslant \frac{\widetilde{c}}{\sqrt{n}}$, $n = 1, 2, \ldots$ $\widetilde{c} > 0$, o bien, en general no es posible mejorar la estimación de la **taza de convergencia** dada en 7.96 en el TLC.

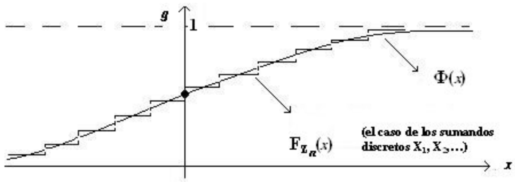

FIGURE 21. Las funciones $F_{Z_n}(x)$ y $\Phi(x)$.

(c) De 7.95 sigue que la f.d. de $S_n = X_1 + \ldots + X_n$ para grandes n es cercana a la Normal con la esperanza an y la varianza $\sigma^2 n$. Este hecho explica la amplia difusión de v.a's. Normales en la naturaleza y en las ciencias (por ejemplo, la estatura de una persona escogida al azar, la variación de precios por acciones en un mercado, los componentes de la velocidad en el movimiento de calor de las moléculas de un gas en equilibrio, etc.).

DEMOSTRACIÓN DEL TEOREMA 7.3.
De (7.13) resulta:

$$Z_n = \frac{X_1 + \cdots + X_n - na}{\sigma\sqrt{n}} = \frac{1}{\sqrt{n}}(Y_1 + \cdots + Y_n),$$

donde $Y_k = \frac{X_k - a}{\sigma}$, $k = 1, 2, \ldots, n$. Observemos que $EY_k = 0$, $Var(Y_k) = \frac{1}{\sigma^2} Var(X_k) = \sigma^2/\sigma^2 = 1$, y por 7.91, $E|Y_k|^3 < \infty$.

Sean $\eta_1, \ldots, \eta_n$ v.a's.$\sim$ $Norm(0.1)$, i.i.d. e independientes de $X_1, \ldots, X_n$, para las cuales tenemos que $E\eta_k = 0$, $Var(\eta_k) = 1$ y $E|\eta_k|^3 < \infty$.

Entonces por el Teorema 2.2, $\eta = \dfrac{1}{\sqrt{n}}(\eta_1 + \ldots + \eta_n) \sim Norm(0.1)$.

Por 7.90,

$$\begin{aligned}
\zeta_3(Z_n, \eta) &= \zeta_3 \left[\frac{1}{\sqrt{n}}(Y_1 + \cdots + Y_n), \frac{1}{\sqrt{n}}(\eta_1 + \cdots + \eta_n) \right] \\
&\leqslant (\frac{1}{\sqrt{n}})^3 \sum_{k=1}^{n} \zeta_3(Y_k, \eta_k) \quad \text{(por la identidad de las f.d's.)} \\
&= \frac{1}{\sqrt{n}n} n \zeta_3(Y_1, \eta_1) = \frac{1}{\sqrt{n}} \zeta_3(Y_1, \eta_1).
\end{aligned}$$

Finalmente, por la Proposición 7.2 b), $\zeta_3(Y_1, \eta_1) < \infty$ (pues 7.86 se cumple bajo las hipótesis del teorema). □

Nota 7.6: Podemos ver que por 7.93 y 7.96 (a diferencia del TLC clásico) la f.d. F_{Z_n} es muy parecida a la Normal no solamente para un número n grande de sumandos X_k, pero también cuando la f.d. de cada sumando X_k aproxima a la Normal.

El siguiente resultado explica por qué en el Teorema 7.3 (y en otros teoremas que se estudian distribuciones de sumas de v.a's.) la distribución del límite es **Normal**. Para simplificar las denotaciones consideraremos en el Teorema 7.3, que $a = 0$ y $\sigma = 1$. Entonces $Z_n = \frac{1}{\sqrt{n}}(X_1 + \ldots + X_n)$ y para $n = 2m$ se tendrá que

$$Z_n = \frac{X_1 + \ldots + X_n}{\sqrt{n}} \equiv Z_{2m} = \frac{X_1 + \ldots + X_m}{\sqrt{2}\sqrt{m}} + \frac{X_{m+1} + \ldots + X_n}{\sqrt{2}\sqrt{m}} \tag{7.97}$$

Supongamos que $Z_n \Rightarrow \xi$, y de esto establezcamos que $\xi \sim Norm(0.1)$. Por 7.97 resultará por un lado que $Z_n \Rightarrow \xi$ y por otro lado (ya que $m \to \infty$ cuando $n \to \infty$) $Z_n \Rightarrow \frac{1}{\sqrt{2}}(\xi_1 + \xi_2)$, donde las v.a's. ξ_1 y ξ_2 son independietes (como los límites de las v.a's. independientes $\frac{1}{\sqrt{2m}}(X_1 + \ldots + X_m)$ y $\frac{1}{\sqrt{2m}}(X_{m+1} + \ldots + X_n)$) y tienen la misma f.d. como la v.a. ξ. El resultado deseado surge de la siguiente afirmación.

TEOREMA 7.4. *Sean ξ, ξ_1, ξ_2 v.a's. tales que:*

a) $E\xi = E\xi_1 = E\xi_2 = 0$, $Var(\xi) = Var(\xi_1) = Var(\xi_2) = 1$, $E|\xi|^3 < \infty$, $E|\xi_1|^3 < \infty$, $E|\xi_2|^3 < \infty$;
b) *ξ_1 y ξ_2 son independientes;*
c) *las f.d's. F_ξ, F_{ξ_1}, F_{ξ_2} y $F_{\frac{\xi_1+\xi_2}{\sqrt{2}}}$ son idénticas.*

Entonces $\xi \sim Norm(0,1)$.

Para la demostración consideremos las v.a's. i.i.d. $\eta_1, \eta_2 \sim Norm(0,1)$ (independientes de ξ_i). Por el Teorema 2.2, $\eta = \frac{\eta_1+\eta_2}{\sqrt{2}} \sim Norm(0.1)$. Entonces por 7.90 y por la condición c), $\zeta_3(\xi,\eta) = \zeta_3(\frac{\xi_1+\xi_2}{\sqrt{2}}, \frac{\eta_1+\eta_2}{\sqrt{2}}) \leqslant (1/\sqrt{2})^2[\zeta_3(\xi_1,\eta_1) + \zeta_3(\xi_2,\eta_2)] = \frac{1}{\sqrt{2}}\zeta_3(\xi_1,\eta_1)$, i.e.

$$\zeta_3(\xi,\eta) \leqslant \frac{1}{\sqrt{2}}\zeta_3(\xi_1,\eta_1) \tag{7.98}$$

Por la igualdad de las f.d's. $\zeta_3(\xi,\eta) = \zeta_3(\xi_1,\eta_1)$, y por lo tanto la desigualdad 7.98 se cumple sólo cuando $\zeta_3(\xi,\eta) = 0$, o bien, cuando $F_\xi \equiv F_\eta$.

EJEMPLO 7.4. ***(Ruleta Americana.)***

Dicha ruleta está dividida en 38 sectores iguales de los cuales 18 son rojos, 18 son negros y 2 son verdes. En un juego, la bola tiene la misma probabilidad (1/38) de pararse en cada sector. El hecho de apostar $x > 0$ dólares por "rojo" significa que el jugador gana x dólares si "sale" rojo y pierde x dólares en el caso contrario. Es decir, la ganancia-pérdida se representa por la v.a.:

$$X = \begin{cases} x, & \textit{con probabilidad 18/38=9/19,} \\ -x, & \textit{con probabilidad 20/38=10/19.} \end{cases} \tag{7.99}$$

Notemos que el juego (como todos los juegos en un casino) es injusto en el sentido de que

$$EX = x\frac{9}{19} + (-x)\frac{10}{19} = -x/19 < 0, \tag{7.100}$$

y al apostar un número grande de veces, la serie de juegos seguramente (con probabilidad 1) se terminará por la ruina del jugador.

En efecto, sea $C > 0$ el capital inicial del jugador (que podría ser arbitrariamente grande) y $X_1, X_2, \ldots$ v.a's. i.i.d. con la distribución dada en 7.99. Aplicando la ley fuerte de los grandes números (Teorema 6.1) para el capital corriente (luego de n apuestas),

$$C_n = C + X_1 + \ldots + X_n = C + S_n, \tag{7.101}$$

obtendremos que con probabilidad 1, $\frac{C_n}{n} = \frac{C}{n} + \frac{S_n}{n} \to EX_1 < 0$ lo cual significa que a partir de un n, $C_n < 0$.

Supongamos que un jugador apuesta 5 dólares (que en muchos casinos de Las Vegas es la mínima cantidad para apostar) por "rojo", $n = (19)^2 = 361$ veces. Como en 7.101, su ganancia-pérdida neta es $S_n = X_1 + \ldots + X_n$, con $x = 5$ en 7.99. Estimemos ahora a partir del TLC las siguientes probabilidades:

(1) $P(S_n > 0)$ (de ganar algo).
(2) $P(S_n \geqslant 100)$ (de ganar al menos 100 dólares).
(3) $P(S_n \leqslant -100)$ (de perder por lo menos 100 dólares).

Primero, por 7.100 $a = EX_1 = -5/19$ y por 3.34 $\sigma^2 = Var(X_1) = E(X_1^2) - (EX_1)^2 =$ *(ya que por 7.99 $X^2 = x^2$ con probabilidad 1)* $= 25 - (5/19)^2 \approx 25$, o bien $\sigma = \sqrt{Var(X_1)} \approx 5$.

Para (2) (de arriba), tenemos: $P(S_n \geqslant 100) = P(\frac{S_n - na}{\sigma\sqrt{n}} \geqslant \frac{100 - na}{\sigma\sqrt{n}}) =$ *(véase 7.92)* $= P(Z_n > \frac{100+19^2(5/19)}{5\cdot 19}) \approx P(X_n \geqslant 2.05) = 1 - P(X_n \leqslant 2.05) =$ *(por 1.4)* $= 1 - F_{Z_n}(2.05) \approx$ *(por 7.95)* $\approx 1 - \Phi(2.05) \approx$ *(por valores de Φ en tablas)* ≈ 0.0202.

Con un procedimiento semejante para los otros incisos, tendremos finalmente que:

1. $P(S_n > 0) \approx 0.1587$.
2. $P(S_n \geqslant 100) \approx 0.0202$.
3. $P(S_n \leqslant -100) \approx 0.4801$.

Nota 7.7: (a) La estrategia "precavida" de juego que realiza el jugador en el ejemplo de arriba, da resultados poco promisorios pues tiene tan sólo 16% (aproximadamente) de chances de ganar algo, pero casi la mitad de probabilidades de perder más de 100 dólares y muy pocas chances (2%) de ganar más de 100 dólares. Por otro lado, una estrategia "arriesgada" como apostar 100 dólares por rojo una sola vez nos da una probabilidad de $18/38 \approx 0.4737$ de ganar 100 dólares y se perdería esta misma cantidad con una probabilidad de $20/38 \approx 0.5263$. Por eso esta estrategia es mucho mejor que la primera. Aunque usándola, el jugador se pierde la oportunidad de gozar del juego durante 3 horas.

(b) Para alguien que planea visitar Las Vegas se sugiere la siguiente estrategia como opción, para ganar mil dólares con probabilidad 1 jugando en la ruleta:
   - Primero: Apostar mil dólares por rojo.
   - Segundo: Si gana, entonces salga.
     Si pierde, entonces apueste dos mil dólares por "rojo".
   - Tercero: Si gana, salga.
     Si pierde, apueste cuatro mil dólares por "rojo".
   - Siga duplicando sus apuestas hasta que salga por primera vez el "rojo".
     En ese instante deje de jugar y salga con la ganancia de mil dólares.

   Sin embargo existen dos obstáculos para la realización de tal estrategia:
   (a) Para el desarrollo del juego necesitará (con probabilidad positiva) poder adquirir como préstamo una cantidad de dólares arbitrariamente granade.
   (b) En los casinos no se permiten apuestas mayores a una cantidad establecida.

EJERCICIOS 7

7.1 Encuentre una sucesión de v.a's. tal que $X_n \Rightarrow X$, pero $X_n \not\to^{\rho} X$.

7.2 Muestre que si $EX \neq EY$ o $EX^2 \neq EY^2$, entonces $\zeta_3(X,Y) = \infty$.

7.3 Sean X una v.a. a.c. con la densidad **uniforme**:

$$f_X(x) = \begin{cases} 1, & \text{si } x \in (0.1), \\ 0, & \text{si } x \notin (0.1), \end{cases}$$

y para $n = 1, 2, \ldots$ la v.a. $X_n \in \{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}\}$ con $P(X_n = k/n) = 1/n$, $k = 0, 1, 2, \ldots, n-1$. Demuestre que $\rho(X_n, X) = 1/n$, $n = 1, 2, \ldots$ y por consecuencia, que $X_n \Rightarrow X$.

7.4 Sean $X \equiv 0$, $X_n \sim Norm(a = 0, \sigma = 1/n)$, $n = 1, 2, \ldots$. Muestre que $\zeta_2(X, X_n) \leqslant 1/n^2$, $n = 1, 2, \ldots$ En particular: $X_n \Rightarrow X$.

7.5 La probabilidad de que un recién nacido sea un varón es aproximadamente $p = 0.512$. Suponiendo que los sexos de los recién nacidos son independientes, estime la probabilidad de que entre 1000 seres que van a nacer el próximo mes en el D.F., el número de mujeres será mayor que el número de varones.

Resp. : ≈ 0.0082.

7.6 ([2]) Sean $X_1, X_2, \ldots$ v.a's. no negativas i.i.d. para las que $EX_1 = 1$ y $Var(X_1) = 1$. Muestre que $2(\sqrt{S_n} - \sqrt{n}) \Rightarrow \eta \sim Norm(0.1)$

SUGERENCIA: Multiplique y divida por $(\sqrt{S_n} + \sqrt{n})$.

7.7 Un dado simétrico se lanza 36 veces. Sea S_n el número de salidas del "6". Estime $P(S_n \leqslant 2)$:

(a) de forma precisa (a partir de la distribución Binomial), y

(b) aproximadamente (a partir del TLC).

Resp. : (a) 0.04712177; (b) ≈ 0.0367.

Nota 7.8: Las respuestas indican que el número de sumandos $n = 36$ (de v.a's. discretas, en este caso) en $S_n = X_1 + \ldots + X_n$ (X_k es la indicadora de salidas de "6") no es suficientemente grande para obtener una buena aproximación. En efecto, el error relativo es $\frac{0.04712177 - 0.0367}{0.04712177} \approx 0.22117 (\approx 22\%)$.

7.8 Al llegar a una oficina, un cliente se encuentra con que hay 40 clientes en fila esperando servicio. Supongamos que los tiempos de servicio son las v.a's. a.c. i.i.d. $X_1, X_2, \ldots$ con promedio $a = EX_1 = 1.4$ min. y con desviación estándar $\sigma = 0.7$. Estime la probabilidad de que el cliente recién llegado tendrá que esperar más de una hora para ser atendido.

Resp. : ≈ 0.184.

Nota 7.9: A diferencia del caso de v.a's. discretas (ver Ejercicio 7.7), para un número de sumandos a.c. como $n = 40$ el TLC nos da una bastante buena aproximación.

7.9 Sean $X_1, X_2, \ldots; Y_1, Y_2, \ldots$ v.a's. i.i.d. con densidad exponencial, con parámetro $\lambda = 1$ y $S_n = (X_1 - Y_1) + (X_2 - Y_2) + \ldots + (X_n - Y_n)$, $n = 1, 2, \ldots$. Calcule los siguientes límites:

(a)

$$\lim_{n\to\infty} P(S_n < \sqrt{n});$$

(b)

$$\lim_{n\to\infty} P(S_n < n^{1/3});$$

(c)

$$\lim_{n\to\infty} P(S_n < n^{2/3}).$$

Resp. : (a)$\approx$ 0.7611; (b) =0.5; (c) =1.

7.10 Sea ρ la métrica de Kolmogorov en (7.1). Demuestre que:
a) $\rho(X+c, Y+c) = \rho(X,Y)$, para cualquier constante $c \in \mathbb{R}$;
b) $\rho(\alpha X, \alpha Y) = \rho(X,Y)$, para cualquier $\alpha > 0$.

8. Comparación de distribuciones de sumas de variables aleatorias y estabilidad de modelos estocásticos aplicados

Las sumas de v.a's. independientes entran como elementos importantes en muchos modelos aplicados. Para mencionar algunos, señalamos los procesos de riesgo (véase el Ejemplo 4.3), procesos de almacenamiento, procesos de regularización del nivel del agua en presas, modelos de colas (filas de espera), procesos de optimización en el reemplazo de equipo, etc. (consulte por ejemplo [1], [9], [10], [11], [12], [14]). El siguiente problema surge en el estudio de la estabilidad (robustez) de tales modelos.

Sean $X_1, X_2, \ldots$ y $\widetilde{X_1}, \widetilde{X_2}, \ldots$ dos sucesiones de v.a's. i.i.d. con f.d's. F y $\widetilde{F}$ respectivamente; y sea

$$\boxed{S_N = X_1 + \ldots + X_N, \quad \widetilde{S}_N = \widetilde{X}_1 + \ldots + \widetilde{X}_N,} \tag{8.102}$$

donde N es una v.a. con valores en $\{1, 2, 3, \ldots\}$. El problema consiste en establecer desigualdades de la forma:

$$\boxed{\rho(S_N, \widetilde{S}_N) \leqslant g[\mu(F, \widetilde{F})]} \tag{8.103}$$

donde ρ es la métrica de Kolmogorov (consulte 7.80), μ es una métrica probabilística y $g(x) \downarrow 0$ cuando $x \to 0$ es una función que o bien no depende de N o bien se anula cuando $EN \to \infty$.

Consideremos el caso particular: $N = n$. Es fácil ver que cuando $EX_1 \neq E\widetilde{X}_1$ no es posible llegar a una desigualdad razonable como en 8.103 con las propiedades de $g(x)$ que se han mencionado (debido a que $ES_n - E\widetilde{S}_n = n(EX_1 - E\widetilde{X}_1) \to \infty$ cuando $n \to \infty$ y por ejemplo, $EX_1 > E\widetilde{X}_1$). Por eso, en el resto del texto, aplicamos la siguiente condición.

Suposición 8.1. *Existen las esperanzas $EX_1, E\widetilde{X}_1$, y*

$$\boxed{a := EX_1 = E\widetilde{X}_1} \tag{8.104}$$

No es tan fácil pero es factible dar algunos ejemplos en los que se cumple 8.104; $\mu(F, \widetilde{F}) \to 0$ en cualquier métrica "razonable", no obstante, $\rho(S_N, \widetilde{S}_N) \geqslant \Delta > 0$. Estos ejemplos muestran que para obtener 8.103 hay que buscar algunas restricciones adicionales a las distribuciones de las v.a's. X_1 y $\widetilde{X}_1$.

Primero intentaremos usar el TLC (para el caso particular, cuando $N = n$ no es aleatorio). Supongamos que $0 < \sigma^2 := Var(X_1) = Var(\widetilde{X}_1)$ y de que $E|X_1|^3 < \infty$, $E|\widetilde{X}_1| < \infty$. Por el Ejercicio 7.10, la desigualdad del triángulo para la métrica ρ y por 7.96 tenemos que

$$\begin{aligned}\rho(S_n, \widetilde{S}_n) &= \rho\left(\frac{S_n - na}{\sigma\sqrt{n}}, \frac{\widetilde{S}_n - na}{\sigma\sqrt{n}}\right)\\ &\leqslant \rho\left(\frac{S_n - na}{\sigma\sqrt{n}}, \eta\right) + \rho\left(\frac{\widetilde{S}_n - na}{\sigma\sqrt{n}}, \eta\right)\\ &\leqslant \frac{C}{\sqrt{n}} \max\left\{\rho\left(\frac{X_1 - a}{\sigma}, \eta\right), \zeta_3\left(\frac{X_1 - a}{\sigma}, \eta\right)\right\}\\ &+ \frac{C}{\sqrt{n}} \max\left\{\rho\left(\frac{\widetilde{X}_1 - a}{\sigma}, \eta\right), \zeta_3\left(\frac{\widetilde{X}_1 - a}{\sigma}, \eta\right)\right\},\end{aligned}$$

es decir,

$$\begin{aligned}\rho(S_n, \widetilde{S}_n) \leqslant \frac{C}{\sqrt{n}} \max\{\rho(\frac{X_1 - a}{\sigma}, \eta), \zeta_3(\frac{X_1 - a}{\sigma}, \eta)\}\\ + \frac{C}{\sqrt{n}} \max\{\rho(\frac{\widetilde{X}_1 - a}{\sigma}, \eta), \zeta_3(\frac{\widetilde{X}_1 - a}{\sigma}, \eta)\},\end{aligned} \tag{8.105}$$

donde ζ_3 es la métrica de Zolotov definida en 7.82.

La parte derecha de 8.105 se anula cuando $n \to \infty$ (que es una buena propiedad), sin embargo, al comparar con 8.103 podemos ver que en 8.105 se ha perdido el factor esencial de "cercanía" entre los modelos: $\mu(F, \widetilde{F}) \equiv \mu(X_1, \widetilde{X}_1)$. Para obtener las desigualdades (de "estabilidad") como en 8.103, necesitamos de las siguientes hipótesis:

SUPOSICIÓN 8.2. (a) *En 8.102 la v.a. N no depende de $X_1, X_2, \ldots$ ni de $\widetilde{X}_1, \widetilde{X}_2, \ldots$.*

(b) *Existe un entero $m \geqslant 1$ tal que las v.a's. $X_1 + X_2 + \ldots + X_m$ y $\widetilde{X}_1 + \widetilde{X}_2 + \ldots + \widetilde{X}_m$ tienen densidades derivables: f_X y $f_{\widetilde{X}}$ respectivamente, tales que:*

$$\int_{-\infty}^{\infty} |f'_X(x)|dx < \infty, \int_{-\infty}^{\infty} |f'_{\widetilde{X}}(x)|dx \quad < \infty \tag{8.106}$$

Nota 8.1: La condición (b) se satisface para la mayoría de densidades de uso habitual, por ejemplo la densidad exponencial que podemos ver en la figura 22 no es derivable en el punto $x = 0$, sin embargo para $X_1 \sim Exp(\lambda)$ la v.a $X_1 + X_2 + X_3$ tiene densidad derivable y ésta derivada satisface (8.5).

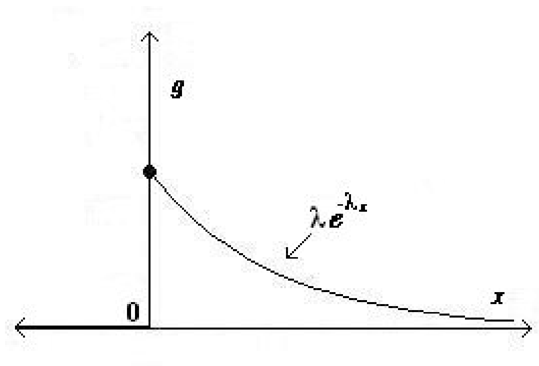

FIGURE 22. La densidad exponencial.

Adelante discutiremos brevemente los métodos para la demostración de las siguientes desigualdades de estabilidad.

TEOREMA 8.1. *([4]) Admitamos que las Suposiciones 8.1 y 8.2 se cumplen, y que* $EX_1^2 < \infty$, $E\widetilde{X}_1^2 < \infty$. *Entonces existe una constante* C_1 *tal que para cualquier v.a. N*

$$\boxed{\rho(S_N, \widetilde{S}_N) \leqslant C_1 \max\{\rho(X_1, \widetilde{X}_1), \zeta_2(X_1, \widetilde{X}_1)\}.} \tag{8.107}$$

TEOREMA 8.2. *([4]) Admitamos que las Suposiciones 8.1 y 8.2 se cumplen, y adicionalmente supongamos que*

$$\boxed{E(X_1^2) = E(\widetilde{X}_1^2),} \tag{8.108}$$

$E|X_1|^3 < \infty$, $E|\widetilde{X}_1|^3 < \infty$. *Entonces existe una constante* C_2 *tal que para cualquier v.a. N:*

$$\boxed{\rho(S_N, \widetilde{S}_N) \leqslant \frac{C_2}{(EN)^{1/2}} \max\{\rho(X_1, \widetilde{X}_1), \zeta_3(X_1, \widetilde{X}_1)\}.} \tag{8.109}$$

Nota 8.2: (a) Las constantes C_1 y C_2 en 8.107 y 8.109 se calculan explícitamente en términos de ciertos parámetros de las funciones características de las v.a's. X_1 y $\widetilde{X}_1$.

(b) En el caso general, no es posible quitar los términos $\zeta_2(X_1, \widetilde{X}_1)$ y $\zeta_3(X_1, \widetilde{X}_1)$ en las partes derechas de 8.107 y 8.109 respectivamente (vea el Ejercicio 8.3). Si se cumplen las condiciones de los Teoremas 8.1 y 8.2, entonces por la Proposición 7.2 dichos términos son finitos.

(c) Cuando $N = n$ (no aleatorio), 8.109 nos da la desigualdad:

$$\boxed{\rho(S_n, \widetilde{S}_n) \leqslant \tfrac{C_2}{\sqrt{n}} \max\{\rho(X_1, \widetilde{X}_1), \zeta_3(X_1, \widetilde{X}_1)\},} \tag{8.110}$$

que asemeja el TLC en la forma 7.96 (donde $\widetilde{X}_1 = \eta_1 \sim Normal$). Es necesario destacar que para obtener el factor $1/\sqrt{n}$ en 8.110 la condición 8.108 es esencial. Tal factor implica que para grandes n las v.a's. S_n y $\widetilde{S}_n$ tienen distribuciones muy parecidas (a pesar de que las distribuciones de sumandos podrian ser muy distintas).

Para demostrar por ejemplo 8.109, primero se verifica la siguiente proposición que es bastante sencilla.

PROPOSICIÓN 8.1. *Sean X, Y, Z v.a's independientes, y Z a.c. con una densidad f_Z para la cual existe la segunda derivada. Entoces:*

$$\rho(X + Z, Y + Z) \leqslant \sup_{x\in\mathbb{R}}\{|f_Z''(x)|\}\zeta_3(X, Y). \tag{8.111}$$

Luego, cuando Z es una suma de v.a's. i.i.d. es posible acotar el valor de $|f_Z''|$ (usando la Suposición 8.2) en términos de algunos parámetros de las funciones características de los sumandos. Para la segunda parte de la demostración se usa la desigualdad del triángulo (para la métrica ρ), las propiedades básicas 7.90 de la métrica de Zolotarev ζ_3 y la fórmula de probabilidad total de la forma 2.22.

EJEMPLO 8.1. ***(La estimación de estabilidad del capital en el modelo clásico de riesgo.)***

Regresando al Ejemplo 4.3, supongamos que la f.d. común F de las v.a's. i.i.d.: $\xi_1, \xi_2, \ldots$ (tamaños de reclamos) involucradas en la ecuación del modelo (vease 4.53):

$$\boxed{X(t) = c + \gamma t - \sum_{n=1}^{N(t)} \xi_n, \quad t \geqslant 0} \tag{8.112}$$

es desconocida. Ésta situación es típica en la práctica y en tal caso se usa una aproximación de F: $\widetilde{F}$, obtenida por estimaciones estadísticas y simiplificaciones teóricas.

Sean $\widetilde{\xi}_1, \widetilde{\xi}_2$ v.a's. i.i.d. con la f.d. $\widetilde{F}$. Entonces, en realidad, un investigador no puede estudiar el modelo "real" 8.112 sino el modelo aproximado:

$$\boxed{\widetilde{X}(t) = c + \gamma t - \sum_{n=1}^{N(t)} \widetilde{\xi}_n, \quad t \geqslant 0.} \tag{8.113}$$

(Aquí no estamos considerando lo referente a la aproximación de $N(t)$.)

La aplicación directa de los Teoremas 8.1 y 8.2 con $S_N = \sum_{n=1}^{N(t)} \xi_n$ y $\widetilde{S}_N = \sum_{n=1}^{N(t)} \widetilde{\xi}_n$ (tomando en cuenta al Ejercicio 7.10) nos da las estimaciones de estabilidad en el modelo de riesgo.

TEOREMA 8.3. *Admitamos que las v.a's. $\xi_1, \xi_2, \ldots$ y $\widetilde{\xi}_1, \widetilde{\xi}_2, \ldots$ satisfacen la Suposición 8.2.*

(a) *Si* $\boxed{E\xi_1 = E\widetilde{\xi}_1}$, *y* $E(\xi_1^2) < \infty$, $E(\widetilde{\xi}_1^2) < \infty$, *entonces para toda* $t \geqslant 0$*:*

$$\boxed{\rho(X_t, \widetilde{X}_t) \leqslant C_1 \max\{\rho(\xi_1, \widetilde{\xi}_1), \zeta_2(\xi_1, \widetilde{\xi}_1)\}.} \tag{8.114}$$

(b) *Si* $\boxed{E\xi_1 = E\widetilde{\xi}_1}$, $\boxed{Var(\xi_1) = Var(\widetilde{\xi}_1)}$ *y* $E|\xi_1|^3 < \infty$, $E|\widetilde{\xi}_1|^3 < \infty$, *entonces para toda* $t \geqslant 0$*:*

$$\boxed{\rho(X_t, \widetilde{X}_t) \leqslant \frac{C_2}{\lambda^{1/2}\sqrt{t}} \max\{\rho(\xi_1, \widetilde{\xi}_1), \zeta_3(\xi_1, \widetilde{\xi}_1)\}.} \tag{8.115}$$

Nota 8.3: (a) La desigualdad 8.114 afirma que si del modelo 8.113 se obtiene una buena aproximación para el modelo real 8.112 de forma que $E\xi_1 = E\widetilde{\xi}_1$ y la distancia $\max\{\rho(\xi_1, \widetilde{\xi}_1), \zeta_2(\xi_1, \widetilde{\xi}_1)\} \equiv \max\{\rho(F, \widetilde{F}), \zeta_2(F, \widetilde{F})\}$ es pequeña, enotonces la distancia unforme entre los capitales corrientes: $\rho(X(t), \widetilde{X}(t)) = \sup_{x\in\mathbb{R}} |P(X(t) \leqslant x) - P(\widetilde{X}(t) \leqslant x)|$ también será pequeña para toda $t \geqslant 0$.

(b) La desigualdad 8.115 se cumple cuando los promedios y las varianzas de la v.a. ξ_1 y su aproximación $\widetilde{\xi}_1$ son iguales. Ésto nos da la ventaja de que la parte derecha de 8.115 se anula cuando $t \to \infty$. Por lo tanto, las f.d's. de $X(t)$ y $\widetilde{X}(t)$ son cercanas para grandes t incluso si las f.d's. de ξ_1 y $\widetilde{\xi}_1$ son bastante diferentes (es decir, cuando no hay una buena aproximación de F por $\widetilde{F}$). Sin embargo, cuando $\max\{\rho(F, \widetilde{F}), \zeta_3(F, \widetilde{F})\}$ se aproxima a cero y $t \to \infty$, la desigualdad 8.115 garantiza "un efecto doble de estabilidad".

(c) La desigualdad 8.115 mejora una cota de estabilidad del proceso de riesgo dada en [11].

EJEMPLO 8.2. ***(Estabilidad de la probabilidad de ruina en el modelo clásico de riesgo.)***

Al recordar las definiciones del Ejemplo 4.3, la **probabilidad de ruina** para el proceso 8.112 con el capital inicial $c > 0$, se denota aquí por $q(c)$, y se define como:

$$\boxed{q(c) := P(\inf_{t>0} X(t) < 0);} \tag{8.116}$$

(la probabilidad de que en un instante $\tau \in (0, \infty)$, el capital de la compañía se anule).

Del mismo modo, para el proceso aproximado $\widetilde{X}(t)$ en 8.113 la probabilidad de ruina $\widetilde{q}(c)$ se define como:

$$\boxed{\widetilde{q}(c) := P(\inf_{t>0} \widetilde{X}(t) < 0).} \tag{8.117}$$

De nuevo, subrayamos que como el investigador no sabe la f.d. F de las v.a's. $\xi_1, \xi_2, ...$ en 8.112, entonces no puede calcular $q(c)$ en 8.116. Por otro lado, teoricamente la probabilidad de ruina $\widetilde{q}(c)$ en 8.117 podría calcularse usando la aproximación (admisible) $\widetilde{F}$. (Sin embargo, exceptuando una muy pequeña cantidad de casos particulares, el cálculo de la probabilidad de ruina es un problema difícil.) En vista de que $q(c)$ es el valor de interés principal (la probabilidad de ruina el el modelo "real"), surge la cuestión de estimación de la cercanía de $\widetilde{q}(c)$ a $q(c)$ (o bien, la estimación de estabilidad).

Usando las denotaciones del Ejemplo 4.3, introducimos la constante:

$$\boxed{\rho = \frac{\gamma}{\lambda a} - 1,} \tag{8.118}$$

que se denomina: la carga de seguridad relativa (the relative safety loading). De la condición 4.56 sigue que $\rho > 0$ y que

$$\boxed{q := \frac{\rho}{1+\rho} < 1.} \tag{8.119}$$

Es un hecho bien conocido (consulte por ejemplo [9], [12]) que

$$\boxed{q(c) = (1-q)P\left(\sum_{k=1}^{\nu} X_k > c\right),} \tag{8.120}$$

donde la v.a. ν es independiente de $X_1, X_2, ...$; y tiene la distribución geométrica con el parámetro q (véase 3.42). Por otro lado, las v.a's. i.i.d.: $X_1, X_2, ...$ tienen la siguiente f.d. común:

$$\boxed{F_{X_1}(x) = 1/a \int_0^x [1 - F(t)]dt, \quad x > 0,} \tag{8.121}$$

donde F es la f.d. de $\xi_1, \xi_2, ...$ en 8.112 y $a = E\xi_1$.

Ahora, bajo la hipótesis:

$$\boxed{E\widetilde{\xi}_1 = E\xi_1 = a,} \tag{8.122}$$

vemos que le modelo aproximado 8.113 tiene la misma ρ como en 8.118 y la misma q como en 8.119. Por consiguiente, analogamente a 8.120 y 8.121 se obtiene que:

$$\widetilde{q}(c) = (1-q)P\left(\sum_{k=1}^{\nu} \widetilde{X}_k > c\right), \tag{8.123}$$

donde las v.a's. i.i.d. $\widetilde{X}_1, \widetilde{X}_2, \ldots$ tienen la f.d.

$$F_{X_1}(x) = 1/a \int_0^x [1-\widetilde{F}(t)]dt, \quad x > 0; \tag{8.124}$$

(y $\widetilde{F}$ es la f.d. de $\widetilde{\xi}_1, \widetilde{\xi}_2, \ldots$ en 8.113). Por lo tanto, para $S_\nu := \sum_{k=1}^{\nu} X_k$ y $\widetilde{S}_\nu := \sum_{k=1}^{\nu} \widetilde{X}_k$, usando las formas de 8.120 y 8.123 y el hecho de que

$\sup_{c>0} |q(c) - q(c)| = (1-q)\sup_{c>0} |P(S_\nu > c) - P(\widetilde{S}_\nu > c)|$

$= (1-q)\sup_{c>0} |P(S_\nu \leqslant c) - P(\widetilde{S}_\nu \leqslant c)| =$ *(véase 7.80)* $= (1-q)\rho(S_\nu, \widetilde{S}_\nu)$, o bien,

$$\sup_{c>0} |q(c) - q(c)| = (1-q)\rho(S_\nu, \widetilde{S}_\nu), \tag{8.125}$$

podemos aplicar el Teorema 8.1.

Para garantizar que se cumple la Suposición 8.1 de este teorema, i.e. que $EX_1 = E\widetilde{X}_1$, necesitamos (desafortunadamente) una restricción adicional:

$$\boxed{Var(\widetilde{\xi}_1) = Var(\xi_1), \quad E|\widetilde{\xi}_1|^3, E|\xi_1|^3 < \infty.} \tag{8.126}$$

Teorema 8.4. *Supongamos que se cumplen las condiciones 8.122 y 8.126. Entonces:*

$$\boxed{\sup_{c>0} |q(c) - \widetilde{q}(c)| \leqslant \tfrac{C_1}{a}(1-q)\max\{\overline{\mu}(\xi_1, \widetilde{\xi}_1), \tfrac{1}{2}\kappa_3(\xi_1, \widetilde{\xi}_1)\};} \tag{8.127}$$

donde C_1 es la constante que aparece en 8.107,

$$\overline{\mu}(\xi_1, \widetilde{\xi}_1) := \overline{\mu}(F, \widetilde{F}) := sup_{x>0}\left|\int_0^x [F(t) - \widetilde{F}(t)]dt\right|, \ y \tag{8.128}$$

$$\kappa_3(\xi_1, \widetilde{\xi}_1) \equiv \kappa_3(F, \widetilde{F}) := \int_0^\infty x^2|[F(x) - \widetilde{F}(x)]|dx. \tag{8.129}$$

Nota 8.4: (a) Las métricas probabilísticas $\overline{\mu}$ y κ_3 en 8.128 y 8.129, aparecen como las cotas para las distancias $\rho(X_1, \widetilde{X}_1)$, $\zeta_2(X_1, \widetilde{X}_1)$ en 8.107, donde las v.a's. X_1 y $\widetilde{X}_1$ tienen las f.d's. escritas en 8.121 y 8.123.

(b) La Suposición 8.2 (b) que usamos en el Teorema 8.1, se ha cambiado aquí (i.e. en el Teorema 8.4) por una condición análoga, que se cumple para cualesquiera f.d's. de la forma 8.121 y 8.124.

(c) Supongamos que las v.a's. $\widetilde{\xi}_1, \widetilde{\xi}_2, \ldots$ en el modelo aproximado 8.113, tienen la f.d. $\widetilde{F}(x) = 1 - e^{-\beta x}$, $x > 0$, exponencial con el parámetro $\beta > 0$. Entonces (véase el Ejercicio 8.1) la probabilidad de ruina $\widetilde{q}(c)$ se calcula explícitamente a partir de 8.123 como:

$$\boxed{\widetilde{q}(c) = \frac{\lambda}{\gamma\beta} e^{-c\rho\frac{\lambda}{\gamma}}\,;} \tag{8.130}$$

donde ρ está definida en 8.118.

Imaginemos ahora que en el modelo real 8.112, la f.d. F de $\widetilde{\xi}_1, \widetilde{\xi}_2, \ldots$ es cercana, con respecto a la métrica $\mu = \max\{\overline{\mu}, \frac{1}{2}\kappa_3\}$, a la f.d. exponencial $\widetilde{F}$ (por ejemplo, supongamos que la parte derecha de 8.127 es menor que ϵ, con un $\epsilon > 0$ pequeño). Entonces, por 8.127 se tendrá que el valor desconocido $q(c) \in [\widetilde{q}(c) - \epsilon, \widetilde{q}(c) + \epsilon]$, con $\widetilde{q}(c)$ dado en 8.130.

(d) Para conocer otras aplicaciones de las desigualdades 8.107 y 8.110, (y de sus verciones con otras métricas) en el estudio de la estabilidad de los modelos estocásticos, el lector puede consultar [4] y [6]. En el artículo [5] se dan las desigualdades de estabilidad para sumas de vectores aleatorios.

EJERCICIOS 8

8.1 Supongamos que en el modelo 8.113 $\widetilde{\xi}_1, \widetilde{\xi}_2, \ldots$ son v.a's. i.i.d. $\sim Exp(\beta)$. Muestre la fórmula 8.130 para la probabilidad de ruina.

SUGERENCIAS: Derivando la equación 8.123 con respecto a $c \equiv x$, y aplicando una versión de la fórmula de probabilidad total 2.22, se obtiene que

$$-\widetilde{q}'(x) = (1-q)\sum_{k=1}^{\infty} p(1-q)^{k-1} f_{S_k}(x)\,; \tag{8.131}$$

(ya que ν es independiente de $\widetilde{\xi}_1, \widetilde{\xi}_2, \ldots$ y tiene la distribución geométrica: $P(\nu = k) = q(1-q)^{k-1}$, $k = 1, 2, \ldots$), En 8.131 f_{S_k}

denota la densidad de $\widetilde{\xi}_1 + \widetilde{\xi}_2 + ... + \widetilde{\xi}_k$. Usando la inducción y la fórmula de convolució 2.25, no es difícil establecer que $f_{S_k}(x) = \beta \frac{(\beta x)^{k-1}}{(\kappa-1)!} e^{-\beta x}$, $x > 0$ (la densidad gama).

Resta calcular la suma en 8.131 e integrar el resultado desde c hasta ∞.

8.2 Usando las hipótesis del Ejercicio 8.1, sean $\lambda = 1000$ ($reclamos/mes$), $a = 1/\beta = 10000$ ($pesos$) y $\gamma = 1.05 \cdot 10^7$ ($pesos$). Calcule la probabilidad de ruina $\widetilde{q}(c)$ para $c = 10^6$ ($pesos$) y $c = 10^7$ ($pesos$).

<u>*Resp.*</u> : $\widetilde{q}(10^6) \approx 0.0081422$;

$$\widetilde{q}(10^7) \approx 1.98665 \cdot 10^{-21} \tag{8.132}$$

Nota 8.5: La última respuesta señala que en el caso de reclamos representados por v.a's. exponcenciales, para un capital inicial "razonable" (como 10 millones de pesos en el último ejemplo) resulta una probabilidad de ruina "excesivamente pequeña" (practicamente cero). En efecto, las estimaciónes de las probabilidades de ruina se usan en la práctica de seguros para recalcular (en algunos periodos) las primas que se cobran a los clientes de la compañía. La razón para hacerlo es que por el crecimiento de precios el parámetro $a = E\xi_k$ aumenta y de ahí, la carga de seguridad relativa ρ en 8.118 se aproxima a cero. Esto podría aumentar la probabilidad de ruina hasta "un nivel peligroso" (para la compañía de seguro).

Por ejemplo, con un capital inicial $c = 10^7$ en el Ejercicio 8.2, sea $a = 10490$ (el crecimiento de precios menor al 5%). Nuevamente, aplicando 8.130, obtenemos que el valor de la probabilidad de ruina es $\boxed{\widetilde{q}(10^7) \approx 0.40314}$ en lugar de 8.132. El aumento en la cantidad de primas γ en 8.118 evita tal aumento de riesgo. Por ejemplo, en el mismo ejercicio con $a = 10490$, al aumentar γ de $1.05 \cdot 10^7$ hasta $1.09 \cdot 10^7$ (el 3.8%), disminuye la probabilidad de ruina (calculada con 8.130) hasta el valor insubsistente $\boxed{\widetilde{q}(10^7) \approx 2.57376 \cdot 10^{-16}}$.

8.3 Sean $X_1, X_2, ...$ v.a's. i.i.d. con la distribución uniforme en (-1,1), es decir

$$f_{X_1}(x) = \begin{cases} \frac{1}{2}, & \text{si } x \in (-1,1), \\ 0, & \text{si } x \notin (-1,1), \end{cases}$$

y $\widetilde{X}_1, \widetilde{X}_2, ...$ v.a's. i.i.d. definidas como sigue:

$$\widetilde{X}_1 = \begin{cases} X_1 & \textit{con probabilidad} \quad 1-\epsilon; \\ \frac{1}{\epsilon} & \textit{con probabilidad} \quad \frac{\epsilon}{2}; \\ \frac{-1}{\epsilon} & \textit{con probabilidad} \quad \frac{\epsilon}{2}, \end{cases}$$

donde $\epsilon \in (0, 1)$ es un número arbitrario.

(a) Muestre que para cualquier ϵ:
$a = EX_1 = E\widetilde{X}_1 = 0$;
$Var(X_1) = 1/3$;

$$\sigma_\epsilon^2 := Var(\widetilde{X}_1) = [\frac{1}{3}(1-\epsilon) + \frac{1}{\epsilon}] \to \infty \quad cuando\ \epsilon \to 0. \tag{8.133}$$

(b) Compruebe que

$$\rho(X_1, \widetilde{X}_1) \to 0 \quad cuando\ \epsilon \to 0. \tag{8.134}$$

(c) Sean $S_n = X_1 + ... + X_n$; $\widetilde{S}_n = \widetilde{X}_1 + ... + \widetilde{X}_n$. Verifique que para $N = n$ se satisfacen todas las hipótesis del Teorema 8.1.

(d) Muestre que existe una constante $\Delta > 0$ tal que para cualquier $\epsilon > 0$ y para toda n suficientemente grande,

$$\rho(S_n, \widetilde{S}_n) \geqslant \Delta. \tag{8.135}$$

Nota 8.6: (a) En el caso anterior, de 8.134 y 8.135 vemos que no hay estabilidad para la sumas con respecto a la métrica ρ de Kolmogorov. En particular, esto muestra que no es posible suprimir la métrica ζ_2 de la parte derecha de 8.107.

(b) A pesar de que no hay estabilidad, la desigualdad 8.107 se cumple. La dificultad que se presenta es que por 8.133, $\zeta_2(X_1, \widetilde{X}_1) \nrightarrow 0$ cuando $\epsilon \to 0$ (¡Verifiquelo!).

SUGERENCIAS: Del Ejercicio 7.10:

$$\rho(S_n, \widetilde{S}_n) = \rho\left(\frac{S_n}{\sqrt{1/3}\sqrt{n}}, \frac{\widetilde{S}_n}{\sqrt{1/3}\sqrt{n}}\right). \tag{8.136}$$

Según el TLC (Corolario 7.1 (a)), para $\eta \sim N(0.1)$ se tiene que

$$\rho\left(\frac{S_n}{\sqrt{1/3}\sqrt{n}}, \eta\right) \to 0\ cuando\ n \to \infty, \tag{8.137}$$

y $\rho\left(\frac{\widetilde{S}_n}{\sigma_\epsilon\sqrt{n}}, \eta\right) \to 0$, o bien (consulte el Ejercicio 7.10)

$$\rho\left(\frac{\widetilde{S}_n}{\sigma_\epsilon\sqrt{n}}, \eta\right) = \rho\left(\frac{\widetilde{S}_n}{\sqrt{1/3}\sqrt{n}}, \frac{\sigma_\epsilon}{\sqrt{1/3}}\eta\right) \to 0\ cuando\ n \to \infty. \tag{8.138}$$

Falta comparar 8.136, 8.137 y 8.138, y observar que $\sigma_\epsilon \neq \sqrt{1/3}$ y por lo tanto las v.a's. η y $\frac{\sigma_\epsilon}{\sqrt{1/3}}\eta$ tienen diferentes f.d's.

Nota 8.7: (De clausura.) El lector que necesite formación en la teoría de probabilidad puede usar el curso introductorio [7]. El que desea entender muchas más cosas interesantes y útiles de probabilidad y paralelamente divertirse en buen grado (sin una preparación seria en matemáticas), puede leer el libro [8].

Bibliografía

[1] S. Asmussen. *Applied Probability and Queues.* Wiley, Chichester, 1987.

[2] R. Durrett. *Probability. Theory and Examples.* Wadsworth & Brooks, 1991.

[3] E. Gordienko. *Estimates of stability of geometric convolutions. Applied Mathematics.* Letters, 12 (1999), 103-106.

[4] E. Gordienko. *Stability estimates of generalized geometric sums and their applications.* Kybernetika, 40 (2004), 257-272.

[5] E. Gordienko. *Comparing the distributions of sums of independient random vectors.* Kybernetika, 41 (2005), 519-529.

[6] E. Gordienko, J. Ruiz de Chávez. *Comarison between distributions of sums of random variables and robustness of some applied stochastic models.* International Journal of Pure and Applied Mathematics, 13 (2004), 293-319.

[7] P. Hoel, S. Port, C. Stone. *Introduction to Probability Theory.* Howghton Miffin, Boston, 1971.

[8] Isaac, R. *The Pleasures of Probability.* Springer-Verlag, N. Y., 1995.

[9] V. V. Kalashnikov. *Geometric Sums: Bounds for Rare Events with Applications.* Kluwer Academic Publishers, Dordrecht, 1997.

[10] R. Kaas, M. Goovaerts, J. Dhaene and M. Denuit. *Modern Actuarial Risk theory.* Kluwer Academic Publishers, Boston, 2001.

[11] S. T. Rachev. *Probability Metrics and the Stability of Stochastic Models.* Wiley, Chichester, 1991.

[12] T. Rolski, H. Shmidli, V. Schmidt and J. Teugels. *Stochastic Processes for Insurance and Finance.* Wiley, Chichester, 1999.

[13] B. Ross, D. Pfeifer. *On the distance between the distributions of random sums.* Journal of Applied Probability, 40 (2003), 87-106.

[14] H. C. Tijms. *Stochastic Models: an Algorithmic Approach.* Wiley, New York, 1994.

[15] V. Zolotarev. *Probability Metrics.* Theory Probability and its Applications, 28 (1983), 264-287.