00365
9

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

POSGRADO EN CIENCIAS MATEMATICAS

FACULTAD DE CIENCIAS

CARACTERIZACIÓN DE DISTRIBUCIONES POR MEDIO DE ESPERANZAS CONDICIONALES

T E S I S

QUE PARA OBTENER EL GRADO ACADEMICO DE:

M A E S T R O E N C I E N C I A S

P R E S E N T A :

HUGO VILLASEÑOR HERNÁNDEZ

DIRECTOR DE TESIS:
DR. JOSÉ MARÍA GONZÁLEZ BARRIOS MURGUÍA

MÉXICO, D. F. **JUNIO DE 2003.**

A

PAGINACIÓN DISCONTINUA

Caracterización de Distribuciones por medio de Esperanzas Condicionales

Tesis de Maestría en Ciencias Matemáticas

Hugo Villaseñor Hernández [1]

[1]**Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas**

B

I

Agradecimientos

A José María González-Barrios por dirigir este trabajo.

A Raúl Rueda, Jorge León, Ana Meda y Mogens Bladt por la revisión de esta tesis.

Prefacio

En este trabajo se reúne una amplia cantidad de material sobre esperanzas condicionales, martingalas y caracterización de distribuciones mediante esperanzas condicionales.
El primer capítulo contiene las principales propiedades de esperanzas condicionales que han sido desglosadas de los textos de probabilidad empleados en cursos de maestría, lo mismo el capítulo dos que desarrolla la parte de martingalas, convergencia de éstas, tiempos de paro e integrabilidad uniforme, esta parte es bastante extensa.
En ambos capítulos se hacen las pruebas de los teoremas lo más detalladas posibles y se exponen algunos ejemplos.
El capítulo tres trata sobre la caracterizaciones de distribuciones que se hacen usando esperanzas condicionales, este material se ha tomado de algunos artículos, todos ellos recientes que en el trabajo son mencionados.

III

Índice general

Capítulo 1

Probabilidad y Esperanza Condicional

1.1. Introducción

La mayoría de los resultados presentados en este capítulo se extrajeron de Ash (1972), Laha y Rohatgi (1974) y Dudley (1989).
Cuando se tienen variables aleatorias X y Y definidas sobre el espacio de probabilidad $(\Omega, \mathcal{F}, P)$ frecuentemente se tienen problemas con la interpretación y el cálculo de probabilidades del tipo: $P[X \geqslant a|Y = b]$, o en general $P[X \in A|Y = y]$, donde $A \in \mathcal{B}(\mathbb{R})$ el σ-álgebra de Borel en $\mathbb{R}$; si las variables aleatorias fueran discretas se podría usar la definición de probabilidad condicional

$$P(C|D) = \frac{P(C \cap D)}{P(D)} \text{ con } C,\, D \in \mathcal{F} \text{ si } P(D) > 0.$$

Las probabilidades de arriba se calculan entonces como

$$P[X \geqslant a|Y = b] = \frac{P[X^{-1}([a,\infty)) \cap Y^{-1}(\{b\})]}{P[Y^{-1}(\{b\})]} \qquad \text{si} \quad P[Y^{-1}(\{b\})] > 0$$

y

$$P[X \in A|Y = y] = \frac{P[X^{-1}(A) \cap Y^{-1}(\{y\})]}{P[Y^{-1}(\{y\})]} \qquad \text{si} \quad P[Y^{-1}(\{y\})] > 0$$

aunque es más usual la notación $P[X \in A]$ que significa $P[X^{-1}(A)]$ entonces podemos escribir

$$P[X \geqslant a|Y = b] = \frac{P[X \in ([a,\infty)), Y \in \{b\}]}{P[Y \in \{b\}]} \qquad \text{si} \quad P[Y \in \{b\}] > 0$$

y

$$P[X \in A|Y = y] = \frac{P[X \in A, Y \in \{y\}]}{P[Y \in \{y\}]} \qquad \text{si} \quad P[Y \in \{y\}] > 0.$$

Notemos que no se tendría ningún problema si en vez de $\{b\}$ o de $\{y\}$ se escribe B, para algún $B \in \mathcal{B}(\mathbb{R})$.

Cuando las variables aleatorias son continuas (basta con que Y sea continua) el evento $[Y = y] = [Y \in \{y\}]$ tiene probabilidad cero y por tanto no tienen sentido las expresiones anteriores (queda una indeterminación del tipo $\frac{0}{0}$ porque $[X \in A, Y \in \{y\}]$ también tiene probabilidad cero).

Un caso en donde se tiene que calcular $P(X \in A|Y = y)$ con $[Y = y]$ un evento con probabilidad cero, es el siguiente:

Se tiene una variable aleatoria Y con densidad uniforme en el intervalo $(0,1)$ $(Y \sim U_{(0,1)})$.

Cuando $Y = y$ se lanza una moneda con $P(\text{águila}) = y$ (suponiendo que esto fuera posible) n veces, entonces X (que depende de y) se define como el número de águilas obtenidas.

Buscamos un modelo para encontrar $P(y; A) = P(X \in A | Y = y)$ puesto que $P(X \in A, Y = y) = 0$ y $P(Y = y) = 0$ no se puede aplicar $\dfrac{P(X \in A, Y = y)}{P(Y = y)}$.

Una manera de modelar esta situación es como sigue:

Sean $\Omega = \mathbb{R}^2$, $\mathcal{F} = \mathcal{B}(\mathbb{R}^2)$ y las variables aleatorias X y Y se definen como $X(x, y) = x$, $Y(x, y) = y$, la medida de probabilidad P_Y está especificada (es Uniforme en $(0, 1)$).

Para cada $y \in \mathbb{R}$ hay una medida de probabilidad $P(y, \cdot) : \mathcal{B}(\mathbb{R}) \longrightarrow \mathbb{R}$ que se interpreta como la probabilidad condicional dado que Y toma el valor y, es decir, $P(y, A) = P(X \in A | Y = y)$.

Se necesita también calcular la probabilidad de eventos de la forma $[(X, Y) \in C]$ con $C \in \mathcal{B}(\mathbb{R}^2)$. Dado que $Y = y$ la pareja (X, Y) pertenece a C si y sólo si X pertenece a la sección $C(y) = \{x : (x, y) \in C\}$ la probabilidad de este evento es $P(y, C(y)) = P(X \in C(y) | Y = y)$.

Para ilustrar el concepto de sección $C(y) = \{x : (x, y) \in C\}$ con $C \subseteq \mathbb{R}^2$ notemos que si C es la circunferencia de radio 2 con centro en el origen entonces $C(1) = \{-\sqrt{3}, \sqrt{3}\}$ pero si C es la bola cerrada de radio 2 con centro en el origen entonces $C(1) = [-\sqrt{3}, \sqrt{3}]$.

Entonces la probabilidad de que $(X, Y) \in C$ que denotamos por $P(C)$ se calcula como:

$$P(C) = \int_{-\infty}^{\infty} P(y, C(y)) dP_Y \text{ que es una fórmula de probabilidad total.}$$

Para conjuntos $C \in \mathcal{B}(\mathbb{R}^2)$ de la forma $A \times B$, es decir,

$$C = \{(x, y) : x \in A, y \in B\} \text{ se tiene que } \quad C(y) = \begin{cases} A & \text{si } y \in B \\ \varnothing & \text{si } y \notin B, \end{cases}$$

y entonces $P(C) = P(A \times B) = \int\limits_B P(y, A) dP_Y$

Cuando $P(y, A)$ es Borel-medible como función de y para cada A Boreliano fijo por el Teorema de la medida producto existe una única medida de probabilidad que satisface:

$$P(C) = \int\limits_B P(y, A) dP_Y \text{ con } C = A \times B \text{ para todos } A \text{ y } B \text{ borelianos,}$$

entonces tenemos una medida P sobre $\mathcal{F} = \mathcal{B}(\mathbb{R}^2)$ que es la única medida de probabilidad determinada por P_Y y la familia de medidas $P(y, \cdot)$, $y \in \mathbb{R}$ definidos como antes.

Ejemplo 1.1.1. Considerar la ya mencionada situación en donde $\Omega = \mathbb{R}^2$, $\mathcal{F} = \mathcal{B}(\mathbb{R}^2)$, $X(x, y) = x$, $Y(x, y) = y$, la medida de probabilidad P_Y es Uniforme en $[0, 1]$ y X es el número de águilas obtenidas al lanzar n veces la moneda en donde $P(\text{águila}) = y$.

Por tanto, para cada $y \in \mathbb{R}$ hay una medida de probabilidad $P(y, \cdot)$: $\mathcal{B}(\mathbb{R}) \longrightarrow \mathbb{R}$ que es la probabilidad condicional X dado que Y toma el valor de y, esto es $P(y, A) = P(X \in A | Y = y)$.

Ahora calcularemos $P[X = k]$ para $k = 0, 1, \ldots, n$.

Sean $\Omega_2 = [0, 1]$, $\mathcal{F}_2 = \mathcal{B}([0, 1])$, sabemos que $P_Y(B) = \int\limits_B dy$ para $B \in$

$\mathcal{B}([0,1])$.

Si $y \in [0,1]$ se toma como la probabilidad condicional de que $X \in A$ dado que $Y = y$ a $P(y, A) = P[X \in A | Y = y]$. Tomamos ahora $\Omega_1 = \{0, 1, 2, \dots, n\}$ y $\mathcal{F}_1 = 2^{\Omega_1}$

$$P(y, \{k\}) = \binom{n}{k} y^k (1-y)^{n-k} = P[X = k | Y = y]$$

es la probabilidad de obtener k águilas dado que la probabilidad de águila es y con $0 \leqslant y \leqslant 1$, notemos que $P(y, k)$ es una función Borel medible en y puesto que es un polinomio de grado n. Si hacemos $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$ el σ-álgebra producto que es el mínimo σ-álgebra que contiene a todos los conjuntos de la forma $D_1 \times D_2$ con $D_1 \in \mathcal{F}_1$ y $D_2 \in \mathcal{F}_2$, P la medida de probabilidad determinada por P_Y y $P(y, \cdot)$ definida como antes, entonces

$$P(C) = \int_0^1 P(y, C(y)) dP_Y \quad \text{para } C \in \mathcal{F}.$$

Ahora si $X(x, y) = x$, $Y(x, y) = y$ tenemos que $P[X = k] = P[\{k\} \times \Omega_2]$, es decir con la notación anterior $C = \{k\} \times \Omega_2$ por lo que cada sección $C(y) = \{x : (x, y) \in \{k\} \times \Omega_2\} = \{k\}$

$$\begin{aligned}
P[X = k] &= P[\{k\} \times \Omega_2] = \int_0^1 P(y, \{k\}) dy = \int_0^1 \binom{n}{k} y^k (1-y)^{n-k} dy \\
&= \binom{n}{k} \int_0^1 y^k (1-y)^{n-k} dy = \binom{n}{k} B(k+1, n-k+1) \\
&= \frac{n!}{k!(n-k)!} \frac{\Gamma(k+1)\Gamma(n-k+1)}{\Gamma(n+2)} = \frac{n!}{k!(n-k)!} \frac{k!(n-k!)}{(n+1)!} \\
&= \frac{1}{n+1} \quad k = 0, 1, \dots, n,
\end{aligned}$$

que es Uniforme Discreta en el conjunto $\{0, 1, 2, \ldots, n\}$, donde $B(\alpha, \beta)$ es la función beta evaluada en $\alpha > 0$ y $\beta > 0$ y $\Gamma(\alpha)$ es la función gamma evaluada en $\alpha > 0$. ■

Ejemplo 1.1.2. Y_1 es una v.a. Uniforme en $(0, 1)$, si $Y_1 = y_1$ definimos Y_2 Uniforme en $(0, y_1)$ y así sucesivamente, esto es, si $Y_1 = y_1$, $Y_2 = y_2, \ldots$, $Y_k = y_k$ hacemos Y_{k+1} Uniforme en $(0, y_k)$ para $k = 1, \ldots, n-1$.

El problema consiste en modelar esta situación para poder calcular $E(Y_n)$.

Ahora tenemos un caso n-variado, definimos $\Omega_j = \mathbb{R}$, $\mathcal{F}_j = \mathcal{B}(\mathbb{R})$, $j = 1, 2, \ldots, n$, $\Omega = \Omega_1 \times \cdots \times \Omega_n$, $\mathcal{F} = \mathcal{F}_1 \times \cdots \times \mathcal{F}_n$ el σ-álgebra producto, que es el mínimo σ-álgebra en Ω que contiene a los conjuntos de la forma $D_1 \times \cdots \times D_n$ con $D_j \in \mathcal{F}_j$, $Y_j(y_1, \ldots, y_n) = y_j \quad j = 1, \ldots, n$.

P_1 es la medida de Lebesgue en $(0, 1)$, para cada $y_1 \in (0, 1)$ y $A \in \mathcal{B}(\mathbb{R})$ tenemos que

$P(y_1, A) = \int\limits_{A \cap (0, y_1)} \frac{1}{y_1} dy_2$ igual que antes $P(y_1, A) = P[Y_2 \in A | Y_1 = y_1]$ que es uniforme en $(0, y_1)$.

En general si son dados $y_1, y_2, \ldots, y_k$ entonces

$$P(y_1, \ldots, y_k, A) = \int\limits_{A \cap (0, y_k)} \frac{1}{y_k} dy_{k+1}.$$

Si P es la única medida de probabilidad sobre $\mathcal{F}$ determinada por P_1 y la medida $P(y_1, ..., y_k, \cdot)$ se puede encontrar la esperanza de una función Borel medible, $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ por Fubini:

$$E(g) = \int\limits_{\Omega} g dP = \int\limits_{\Omega_1} \int\limits_{\Omega_2} \cdots \int\limits_{\Omega_n} g(y_1, \ldots, y_n) P(y_1, \ldots, y_{n-1}, dy_n) \cdots P(y_1, dy_2) P_1(dy_1)$$

sustituyendo $g(y_1,\dots,y_n)=Y_n(y_1,\dots,y_n)=y_n$ se tiene:

$$\begin{aligned}
E(Y_n) &= \int_0^1\int_0^{y_1}\int_0^{y_2}\cdots\int_0^{y_{n-2}}\int_0^{y_{n-1}} y_n\frac{1}{y_{n-1}}\frac{1}{y_{n-2}}\cdots\frac{1}{y_2}\cdot\frac{1}{y_1}\cdot 1dy_ndy_{n-1}\cdots dy_3dy_2dy_1\\
&= \int_0^1\int_0^{y_1}\int_0^{y_2}\cdots\int_0^{y_{n-2}}\left(\frac{y_{n-1}^2}{2y_{n-1}}\right)\frac{1}{y_{n-2}}\cdots\frac{1}{y_2}\cdot\frac{1}{y_1}\cdot dy_{n-1}\cdots dy_3dy_2dy_1\\
&= \int_0^1\int_0^{y_1}\int_0^{y_2}\cdots\int_0^{y_{n-2}}\frac{y_{n-1}}{2}\frac{1}{y_{n-2}}\cdots\frac{1}{y_2}\cdot\frac{1}{y_1}\cdot dy_{n-1}\cdots dy_3dy_2dy_1\\
&= \int_0^1\int_0^{y_1}\int_0^{y_2}\cdots\frac{y_{n-2}}{4}\frac{1}{y_{n-3}}\cdots\frac{1}{y_2}\cdot\frac{1}{y_1}\cdots dy_3dy_2dy_1=\cdots\\
&= \int_0^1\int_0^{y_1}\frac{y_2}{2^{n-2}}\frac{1}{y_1}dy_1dy_2=\int_0^1\frac{1}{y_1}\left(\frac{y_2^2}{2^{n-1}}\bigg|_0^{y_1}\right)dy_1\\
&= \int_0^1\frac{y_1}{2^{n-1}}dy_1=\frac{y_1^2}{2^n}\bigg|_0^1=\frac{1}{2^n}. \quad \blacksquare
\end{aligned}$$

Ejemplo 1.1.3. Supongamos que Y es una variable aleatoria discreta que toma valores en $\mathbb{N}=\{1,2,\dots\}$ con probabilidades $\{p_n : n\geqslant 1\}$.
Si $Y=n$ se selecciona un número $X\geqslant 0$ con densidad f_n.
Nuestro problema ahora consiste en calcular $P[2\leqslant X+Y\leqslant 5]$.
Definamos $\Omega_2=\{1,2,\dots\}$, $\mathcal{F}_2=2^{\Omega_2}$ y $P_2(\{k\})=p_k$ el modelo para Y, mientras que el modelo para X dado Y.es:

$$\Omega_1=\mathbb{R},\quad \mathcal{F}_1=\mathcal{B}(\mathbb{R}) \text{ y } P(n,A)=\int_A f_n(y)dy,$$

es decir $P(n,A)=P[X\in A|Y=n]$, $\quad P_2(\{k\})=P[Y=k]$.

Se definen nuevamente $X(\omega_1, \omega_2) = \omega_1$, $Y(\omega_1, \omega_2) = \omega_2$ y la medida de probabilidad determinada por P_2 y $P(n, \cdot)$ es:

$$P(C) = \int_{\Omega_2} P(\omega_2, C(\omega_2))P_2(d\omega_2) \qquad \text{con } C \in \mathcal{F}_1 \times \mathcal{F}_2.$$

Como P_2 es discreta entonces: $P(C) = \sum_{n=1}^{\infty} P(n, C(n))p_n$ pero aquí consideraremos $C = \{(\omega_1, \omega_2) : 2 \leqslant \omega_1 + \omega_2 \leqslant 5\}$ entonces

$$C(n) = \{\omega_1 : 2 \leqslant \omega_1 + n \leqslant 5\} = \{\omega_1 : 2 - n \leqslant \omega_1 \leqslant 5 - n\},$$

de este modo

$$\begin{aligned} P(C) &= P[2 \leqslant X + Y \leqslant 5] = \sum_{n=1}^{\infty} p_n \int_{2-n}^{5-n} f_n(y)dy \\ &= p_1 \int_1^4 f_1(y)dy + p_2 \int_0^3 f_2(y)dy + p_3 \int_0^2 f_3(y)dy + \\ &\quad p_4 \int_0^1 f_4(y)dy \end{aligned}$$

puesto que $X \geqslant 0$. ■

1.2. Probabilidad y Esperanza Condicional

Iniciamos esta sección enunciando el Teorema de Radon-Nikodym, recordemos que si λ y P son dos medidas en $(\Omega, \mathcal{F})$, λ es absolutamente continua con respecto a P si y sólo si $P(A) = 0$ implica que $\lambda(A) = 0$ (se denota $\lambda \ll P$), entonces el teorema de Radon-Nikodym dice: "Sea μ una medida σ-finita y λ una medida signada sobre $\mathcal{F}$ (σ-álgebra de subconjuntos de Ω). Supongamos que λ es absolutamente continua con respecto a μ, entonces hay una función Borel-medible $g : \Omega \longrightarrow \overline{\mathbb{R}}$ tal que $\lambda(B) = \int\limits_B g d\mu$ para todo $B \in \mathcal{F}$. Si h es otra función con esta propiedad entonces $h = g$ casi seguramente (μ)."

Una construcción muy detallada del Teorema Radon-Nikodym aparece en Billingsley (1986). Se ha interpretado $P(y, A)$ como la probabilidad condicional de que X tome valores en $A \in \mathcal{B}(\mathbb{R})$ dado que la variable aleatoria Y vale y; $P(y, A)$ define una medida de probabilidad única sobre $\mathcal{B}(\mathbb{R}^2)$.

El problema ahora es ver que se haría si Y es un objeto o elemento aleatorio, es decir, $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$ donde $(\Omega', \mathcal{F}')$ es cualquier espacio medible y no necesariamente $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ entonces $P(y, A) = P[A|Y = y]$ con $A \in \mathcal{F}$, $y \in \Omega'$, con $[Y = y]$ un evento de probabilidad cero.

Si tomamos $B \in \mathcal{F}'$ y $[Y \in B] \in \mathcal{F}$ entonces $[Y \in B] \cap A \in \mathcal{F}$, por lo que: $P[A \cap [Y \in B]]$ se puede escribir como $P[A \cap [Y \in B]] = \int\limits_B P(y, A) dP_Y(y)$ en donde P_Y es la medida de probabilidad inducida por Y, esto es,

$$P_Y(B) = P[Y \in B] = P[\{\omega \in \Omega : Y(\omega) \in B\}]$$

Teorema 1.2.1. *Sean* $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$ *un objeto aleatorio sobre*

$(\Omega, \mathcal{F}, P)$ *y* $A \in \mathcal{F}$ *fijo. Existe una función* $g : (\Omega', \mathcal{F}') \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *que cumple con que para todo* $B \in \mathcal{F}'$ *se tiene:*
$P[A \cap [Y \in B]] = \int_B g(y) dP_Y(y)$, *además si* h *es otra función con la misma propiedad entonces* $g = h$ *casi seguramente con la medida* P_Y.

(Definimos $P(A|Y = y) = g(y)$ porque A es fijo)

Demostración: Sea $\lambda(B) = P[A \cap [Y \in B]]$ entonces $\lambda : \mathcal{F}' \longrightarrow [0, 1]$ es una medida tal que $\lambda \ll P_Y$ porque $\lambda(B) = P[A \cap [Y \in B]] \leqslant P[Y \in B]$.
Por tanto si $P[Y \in B]=0$ se tiene que $P[A \cap [Y \in B]]=\lambda(B)=0$.
La conclusión incluyendo unicidad casi seguramente con respecto a P_Y se sigue del Teorema de Radon-Nikodym. □

Ejemplo 1.2.2. Sea Y una variable aleatoria discreta que toma el valor y_i con probabilidad p_i, $i = 1, 2, \ldots$ esto es, $p_i = P[Y = y_i]$, $\sum_{i=1}^{\infty} p_i = 1$ entonces

$$g(y_i) = P(A|Y = y_i) = \frac{P[A \cap [Y = y_i]]}{p_i} \quad i = 1, 2, \ldots$$

se puede definir también $g(y) = 0$ para todo $y \notin \{y_1, y_2, \ldots\}$.
Para ver esto hacemos $\Omega' = \{y_1, y_2, \ldots\}$, $\mathcal{F}' = 2^{\Omega'}$, tomemos $B \in \mathcal{F}'$ y de este modo

$$\int_B g(y) dP_Y(y) = \int_{\Omega'} g(y) I_B(y) dP_Y(y) = \sum_{i=1}^{\infty} g(y) I_B(y_i) P_Y(y_i)$$

$$= \sum_{y_i \in B} g(y_i) p_i = \sum_{y_i \in B} P[A \cap [Y = y_i]] = P[A \cap [Y \in B]].$$

$$\text{Por tanto } \int_B g(y) dP_Y(y) = P[A \cap [Y \in B]]. \quad \blacksquare$$

Ejemplo 1.2.3. Sean X y Y variables aleatorias con densidad conjunta f (Recordemos que

$\Omega=\mathbb{R}^2$, $\mathcal{F}=\mathcal{B}(\mathbb{R}^2)$, $X(x,y)=x$, $Y(x,y)=y$, $P(D)=\iint_D f(x,y)dxdy$, $D\in\mathcal{F}\Big)$.

Sabemos que si $A=[Y=y]$ entonces $P(A)=0$, pero una aproximación a $P[X\in C|Y=y]$ se puede hace así:

$$P[X\in C|y-h\leqslant Y\leqslant y+h]=\frac{P[X\in C, y-h\leqslant Y\leqslant y+h]}{P[y-h\leqslant Y\leqslant y+h]}$$

$$=\frac{\int\limits_{y-h}^{y+h}\int\limits_{C} f(x,v)dxdv}{\int\limits_{y-h}^{y+h} f_2(v)dv},$$

donde f_2 es la densidad de Y, es decir, $f_2(y)=\int\limits_{-\infty}^{\infty} f(x,y)dx$. Si h es pequeña, $P[X\in C|y-h\leqslant Y\leqslant y+h]$ se puede aproximar por:

$$\frac{2h\int\limits_{C} f(x,y)dx}{2hf_2(y)} \text{ para casi toda } y\in\mathbb{R}$$

como consecuencia del teorema del valor intermedio para integrales, pero

$$\frac{2h\int\limits_{C} f(x,y)dx}{2hf_2(y)}=\frac{\int\limits_{C} f(x,y)dx}{f_2(y)}=\int\limits_{C}\frac{f(x,y)}{f_2(y)}dx,$$

por lo que se define $h(x|y)=\dfrac{f(x,y)}{f_2(y)}$ que es la densidad condicional de X dado que Y toma el valor y, ésta función se define así cuando $f_2(y)\neq 0$. Esperamos ahora que $P[X\in C|Y=y]=\int\limits_{C} h(x|y)dx$.

Más generalmente, si $A\in\mathcal{F}$ y $Y=y$ entonces A ocurre si y sólo si $X\in A(y)$, y para encontrar

$P[X \in A(y)|Y = y]$ integramos $h(x|y)$ sobre $x \in A(y)$.

Por tanto, proponemos $g(y) = \int\limits_{A(y)} h(x|y)dx$ con $A \in \mathcal{F} = \mathcal{B}(\mathbb{R}^2)$, $y \in \mathbb{R}$ como la probabilidad condicional de A dado que $Y = y$.

Notemos que $g(y) = \int\limits_{-\infty}^{\infty} I_A(x,y)h(x|y)dx$ es Borel medible.

Si $B \in \mathcal{B}(\mathbb{R})$ entonces:

$$\begin{aligned}
P[A \cap [Y \in B]] &= \iint\limits_{\substack{y\in B \\ (x,y)\in A}} f(x,y)dxdy = \iint\limits_{\substack{y\in B \\ (x,y)\in A}} h(x|y)f_2(y)dxdy \\
&= \int\limits_{-\infty}^{\infty}\left(\int\limits_{-\infty}^{\infty} I_A(x,y)h(x|y)dx\right)I_B(y)f_2(y)dy \\
&= \int\limits_{y\in B}\left(\int\limits_{A(y)} h(x|y)dx\right)f_2(y)dy = \int\limits_{B} g(y)f_2(y)dy \\
&= \int\limits_{B} g(y)dP_Y(y).
\end{aligned}$$

De este modo $g(y) = P[A|Y = y]$

Del mismo modo se puede trabajar en dimensiones mayores que 2, por ejemplo si X, Y, Z, y W tienen densidad conjunta f. la densidad condicional de (Z, W) dado (X, Y) es:

$$h(z,\omega|x,y) = \frac{f(x,y,z,\omega)}{f_{X,Y}(x,y)} \text{ donde } f_{X,Y}(x,y) = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f(x,y,z,\omega)dzd\omega.$$

Si $A \in \mathcal{B}(\mathbb{R}^4)$ entonces $P(A|X = x, Y = y) = \iint\limits_{A(x,y)} h(z,\omega|x.y)dzd\omega$ se puede comprobar tomando $B \in \mathcal{B}(\mathbb{R}^2)$ y verificando que

$$P[[(X,Y) \in B] \cap A] = \iint_B P[A|X = x, Y = y]f_{X,Y}(x,y)dxdy. \quad \blacksquare$$

Ejemplo 1.2.4. Sean $(\Omega_1, \mathcal{F}_1)$ y $(\Omega_2, \mathcal{F}_2)$ dos espacios medibles. Hacemos

$\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$, $X(\omega_1, \omega_2) = \omega_1$, $Y(\omega_1, \omega_2) = \omega_2$.

Supongamos que tenemos P_Y una medida de probabilidad sobre $(\Omega_2, \mathcal{F}_2)$ y también tenemos $P(y, A)$ para $y \in \Omega_2$ y $A \in \mathcal{F}_1$, una medida de probabilidad en A para cada y fijo (Es decir, tenemos la distribución de Y y la distribución condicional de X dado que $Y = y$).

Sabemos entonces que hay una única medida P sobre $\mathcal{F}$ tal que $P[X \in A, Y \in B] = P(A \times B) = \int_B P(y, A) dP_Y(y)$.

Calcularemos ahora la esperanza condicional de X dado que $Y = y$ que se denota por $E(X|Y = y)$ y que se entiende como el valor promedio de X cuando Y toma el valor de y.

Si X y Y son discretas y $Y = y$ entonces $p(x|y) = P[X = x|Y = y]$ es la probabilidad condicional de que $X = x$ dado que $Y = y$, entonces $E(X|Y = y) = \sum_x xp(x|y)$.

De la misma manera, en el caso continuo si $h(x|y)$ es la densidad condicional correspondiente entonces $E(X|Y = y) = \int_{-\infty}^{\infty} xh(x|y)dx$.

Es necesario definir este concepto en un caso general.

Sea X una variable aleatoria sobre $(\Omega, \mathcal{F}, P)$, esto es, $X : (\Omega, \mathcal{F}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, Y un objeto o elemento aleatorio $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$. Sabemos que

$$P[A \cap [Y \in B]] = \int_B P(A|Y = y) dP_Y(y) \text{ con } A \in \mathcal{F} \text{ y } B \in \mathcal{F}'.$$

La esperanza de X es el promedio de las contribuciones $E(X|Y = y)$, es

decir, $E(X) = \int_{\Omega'} E(X|Y = y)dP_Y(y)$.
Si reemplazamos X por $XI_{[Y\in B]}$ con $B \in \mathcal{F}'$ se cumple que cuando $y \in B$ obtenemos la igualdad:
$E[XI_{[Y\in B]}|Y = y] = E(X|Y = y)$. Haciendo la sustitución se tiene:

$$E[XI_{[Y\in B]}] = \int_{\Omega'} E[XI_{[Y\in B]}|Y = y]dP_Y(y).$$

O equivalentemente:

$$\int_{Y\in B} XdP = \int_{B} E[X|Y = y]dP_Y(y). \quad \blacksquare$$

Teorema 1.2.5. *Sean X una variable aleatoria extendida definida sobre $(\Omega, \mathcal{F}, P)$ y $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$ un objeto aleatorio. Si $E(X)$ existe (en $\overline{\mathbb{R}}$, los reales extendidos), hay una función $g : (\Omega', \mathcal{F}') \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ tal que para todo $B \in \mathcal{F}'$ se cumple que $\int_{[Y\in B]} XdP = \int_B g(y)dP_Y(y) \quad B \in \mathcal{F}'$.*
Más aún, si h es otra función que cumple con la igualdad anterior entonces $g = h$ casi seguramente (P_Y). Se define $E(X|Y = y) = g(y)$.

Demostración: Sea $\lambda(B) = \int_{Y\in B} XdP$ con $B \in \mathcal{F}'$, entonces λ es una función de $\mathcal{F}'$ en $\overline{\mathbb{R}}$ numerablemente aditiva porque

$$\lambda(\bigcup_{j\in\mathbb{N}} B_j) = \int_{[Y\in \bigcup_{j\in\mathbb{N}} B_j]} XdP = \int_{\bigcup_{j\in\mathbb{N}}[Y\in B_j]} XdP = \sum_{j\in\mathbb{N}} \int_{[Y\in B_j]} XdP,$$

cuando las B_j son ajenas en $\mathcal{F}'$, además $\lambda \ll P_Y$ porque si $P[Y \in B] = P_Y(B) = 0$ entonces $\lambda(B) = \int_{[Y\in B]} XdP = 0$ ya que estaríamos integrando sobre un conjunto con probabilidad cero.

Luego, la conclusión incluyendo unicidad es válida por el Teorema de Radon-Nikodym. □

Corolario 1.2.6. *Sean* $Y : (\Omega, \mathcal{F}, P) \longrightarrow (\Omega', \mathcal{F}')$ *un objeto aleatorio* $A \in \mathcal{F}$, *entonces* $E(I_A|Y = y) = P(A|Y = y)$ *c.s.* (P_Y).

Demostración:

En el Teorema 1.2.5 si hacemos $X = I_A$ obtenemos

$$\int_{[Y\in B]} I_A dP = \int_B g(y) dP_Y(y),$$

donde $g(y) = E(I_A|Y = y)$ pero

$$\int_{[Y\in B]} I_A dP = \int_\Omega I_{A\cap[Y\in B]} dP = P[A \cap [Y \in B]].$$

Por el Teorema 1.2.1, $P[A \cap [Y \in B]] = \int_B g_1(y) dP_Y(y)$ donde $g_1(y) = P(A|Y = y)$.

Por tanto $P[A \cap [Y \in B]] = \int_B E(I_A|Y = y) dP_Y(y)$ y de este modo $P(A|Y = y) = E(I_A|Y = y)$ c.s. (P_Y).□

Ejemplo 1.2.7. Sea Y la variable aleatoria del Ejemplo 1.2.2.

Vimos que $P(A|Y = y_i) = \frac{P[A\cap[Y=y_i]]}{P[Y=y_i]}$ con $A \in \mathcal{F}$.

Esperaríamos que $E(I_A|Y = y_i) = \frac{1}{P[Y=y_i]} \int_{[Y=y_i]} I_A dP$ pero por el Corolario $E(I_A|Y = y_i) = P[A|Y = y_i]$ mientras que $\int_{[Y=y_i]} I_A dP = P[A \cap [Y = y_i]]$ por lo que

$$\frac{1}{P[Y = y_i]} \int_{[Y=y_i]} I_A dP = \frac{P(A \cap [Y = y_i])}{P[Y = y_i]} = P(A|Y = y_i),$$

que indica que $E(I_A|Y=y_i)$ en efecto es $\frac{1}{P[Y=y_i]}\int\limits_{[Y=y_i]} I_A dP$.
En general se debe cumplir

$$E(X|Y=y_i)=\frac{1}{P(Y=y_i)}\int\limits_{[Y=y_i]} XdP \quad \text{toda vez que } E(X) \text{ exista.}$$

Esta última igualdad la verificamos de la siguiente manera:
Sean $\Omega'=\{y_1;y_2,\ldots\}$, $\mathcal{F}'=2^{\Omega'}$ y $g(y_i)=\frac{1}{P[Y=y_i]}\int\limits_{[Y=y_i]} XdP$ entonces debemos comprobar que $\int\limits_{[Y\in B]} XdP=\int\limits_B g(y)dP_Y(y)$ con $B\in\mathcal{F}'$ estamos tomando $Y:(\Omega,\mathcal{F},P)\longrightarrow(\Omega',\mathcal{F}')$ así, $[Y\in B]\in\mathcal{F}$. Pero

$$\begin{aligned}\int\limits_{[Y\in B]} XdP &= \sum_{y_i\in B}\int\limits_{[Y=y_i]} XdP=\sum_{y_i\in B}P[Y=y_i]\frac{1}{P[Y=y_i]}\int\limits_{[Y=y_i]} XdP\\ &=\sum_{y_i\in B}P[Y=y_i]g(y_i)=\int\limits_B g(y)dP_Y(y).\end{aligned}$$

Por tanto $E(X|Y=y_i)=g(y_i)=\frac{1}{P[Y=y_i]}\int\limits_{[Y=y_i]} XdP$.

También se puede comprobar que $E(\sum\limits_{k=1}^{n} X_k=|Y=y_i)=\sum\limits_{k=1}^{n}E(X_k|Y=y_i)$, esto se hace mostrando que $\int\limits_{[Y\in B]}(\sum\limits_{k=1}^{n}X_k)dP=\int\limits_B(\sum\limits_{k=1}^{n}g_k(y))dP_Y(y)$ con $B\in\mathcal{F}'$.

Entonces $\int\limits_{[Y\in B]}(\sum\limits_{k=1}^{n}X_k)dP=\sum\limits_{k=1}^{n}\int\limits_{[Y\in B]}X_k dP=\sum\limits_{k=1}^{n}\int\limits_B g_k(y)dP_Y(y)$
$=\int\limits_B(\sum\limits_{k=1}^{n}g_k(y))dP_Y(y)$. Por lo que $E(\sum\limits_{k=1}^{n}X_k|Y=y_i)=\sum\limits_{k=1}^{n}E(X_k|Y=y_i)$ indica que $E(X|Y=y_i)$ es lineal cuando X es discreta y toma los valores $x_1,x_2,\ldots$ entonces $E(X|Y=y_i)=\sum\limits_{j=1}^{\infty}x_jP[X=x_j|Y=y_i]$.
Por el Teorema 1.2.5 debemos comprobar que si $B\in\mathcal{F}'$

$$\begin{aligned}\int_{[Y\in B]} XdP &= \int_B E(X|Y=y_i)dP_Y(y), \text{ es decir,}\\ \int_{[Y\in B]} XdP &= \int_B \left(\sum_{j=1}^{\infty} x_j P[X=x_j|Y=y_i]\right)dP_Y(y).\end{aligned}$$

Notemos que $X = \sum_{j=1}^{\infty} x_j I_{[X=x_j]}$ entonces

$$\begin{aligned}\int_{[Y\in B]} XdP &= \sum_{y_i\in B}\int_{[Y=y_i]} XdP = \sum_{y_i\in B}\int_{[Y=y_i]}\left(\sum_{j=1}^{\infty} x_j I_{[X=x_j]}\right)dP\\ &= \sum_{y_i\in B}\sum_{j=1}^{\infty} x_j \int_{[Y=y_i]} I_{[X=x_j]}dP = \sum_{y_i\in B}\sum_{j=1}^{\infty} x_j P[X=x_j, Y=y_i]\\ &= \sum_{y_i\in B}\sum_{j=1}^{\infty} P[Y=y_j]x_j P[X=x_j|Y=y_i]\\ &= \sum_{y_i\in B} P[Y=y_j]\left(\sum_{j=1}^{\infty} x_j P[X=x_j|Y=y_i]\right)\\ &= \int_B \left(\sum_{j=1}^{\infty} x_j P[X=x_j|Y=y_i]\right)dP_Y(y)\end{aligned}$$

que es el resultado esperado. ■

Ejemplo 1.2.8. Sea $A \in \mathcal{F}$ tal que $P(A) > 0$. Si $E(X)$ existe se puede definir $E(X|A)$ como:

$$E(X|A) = E(X|Y=1) = \frac{1}{P(A)}\int_A XdP,$$

donde $Y = I_A$ que es consecuencia del ejemplo 1.2.7. También se puede escribir como:

$$E(X|A) = \frac{1}{P(A)}\int_\Omega XI_A dP = \frac{E(XI_A)}{P(A)}. \quad ■$$

Ejemplo 1.2.9. Supongamos que las variables aleatorias X y Y tienen densidad conjunta f y densidad condicional de X dado Y, $h(x|y)$.
Afirmamos que $E(X|Y=y)=\int\limits_{-\infty}^{\infty} xh(x|y)dx$ cuando $E(X)$ existe.
De nuevo por el Teorema 1.2.5 debemos mostrar que $\int\limits_{[Y\in B]} XdP=\int\limits_{B} g(y)dP_Y(y)$ con $B\in\mathcal{B}(\mathbb{R})$ donde $g(y)=\int\limits_{-\infty}^{\infty} xh(x|y)dx$.

$$\begin{aligned}\int\limits_{[Y\in B]} XdP &= \iint\limits_{\{(x,y):y\in B\}} xf(x,y)dxdy=\iint\limits_{\mathbb{R}\times[y\in B]} xf(x,y)dxdy\\ &= \int\limits_{y\in B}\int\limits_{-\infty}^{\infty} xh(x|y)f_2(y)dxdy=\int\limits_{y\in B}\left(\int\limits_{-\infty}^{\infty} xh(x|y)dx\right)f_2(y)dy\\ &= \int\limits_{B} g(y)dP_Y(y).\end{aligned}$$

Aquí $f_2(y)$ es la densidad marginal de Y. Ahora supongamos que $q:\mathbb{R}\longrightarrow\mathbb{R}$ es Borel medible y veremos que $E[q(X)|Y=y]=\int\limits_{-\infty}^{\infty} q(x)h(x|y)dx$ toda vez que $E\left(q(X)\right)$ exista.
Debemos ver que si $B\in\mathcal{B}(\mathbb{R})$ se cumple que

$$\int\limits_{[Y\in B]} q(X)dP=\int\limits_{B} g(y)dP_Y(y)\quad\text{donde}\quad g(y)=\int\limits_{-\infty}^{\infty} q(x)h(x|y)dx.$$

Así

$$\begin{aligned}\int\limits_{[Y\in B]} q(X)dP &= \iint\limits_{\{(x,y):y\in B\}} q(x)f(x,y)dxdy=\int\limits_{y\in B}\int\limits_{-\infty}^{\infty} q(x)h(x|y)f_2(y)dxdy\\ &= \int\limits_{y\in B}\left(\int\limits_{-\infty}^{\infty} q(x)h(x|y)dx\right)f_2(y)dy=\int\limits_{B} g(y)dP_Y(y)\end{aligned}$$

que es lo que debíamos probar.

Por último en el caso discreto se comprueba que si $E\,(q(X))$ existe entonces $E(q(X)|Y=y_i) = \sum_{j=1}^{\infty} q(x_j)P[X=x_j|Y=y_i]$. ■

Ejemplo 1.2.10. Sean $(\Omega_1, \mathcal{F}_1)$ y $(\Omega_2, \mathcal{F}_2)$ espacios medibles $\Omega{=}\Omega_1 \times \Omega_2$, $\mathcal{F}{=}\mathcal{F}_1 \times \mathcal{F}_2$ el mínimo σ-álgebra que contiene a los conjuntos de la forma $C \times D$ con $C \in \mathcal{F}_1$ y $D \in \mathcal{F}_2$, $X(x,y) = x$, $Y(x,y) = y$ se sabe que hay una medida P_Y sobre $\mathcal{F}_2$ y también una medida $P(y, A)$ para $y \in \Omega_2$, $A \in \mathcal{F}_1$ como antes.

Si $f : (\Omega_1, \mathcal{F}_1) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ y $E\,(f(X))$ existe entonces afirmamos que

$$E(f(X)|Y=y) = \int_{\Omega_1} f(x)P(dx,y).$$

Tenemos que ver otra vez que

$$\int_{[Y\in B]} f(X)dP = \int_B g(y)dP_Y(y) \text{ donde } B \in \mathcal{B}(\mathbb{R}) \text{ y } g(y) = \int_{\Omega_1} f(x)P(dx,y).$$

$$\begin{aligned}\int_{[Y\in B]} f(X)dP &= \int_\Omega f(X)I_{[Y\in B]}dP = \int_{\Omega_1}\int_{\Omega_2} f\,(X(x,y))\,I_B(y)P(dx,y)dP_Y(y)\\ &= \int_B \left(\int_{\Omega_1} f(x)P(dx,y)\right) dP_Y(y). \quad ■\end{aligned}$$

Notemos que en todos los ejemplos el procedimiento seguido es proponer alguna g, ver que funcione como esperanza condicional y utilizar el teorema de Radon-Nikodym para asegurar que es única casi seguramente.

1.3. Esperanza Condicional dado un σ-álgebra

Si X es una variable aleatoria extendida sobre $(\Omega, \mathcal{F}, P)$ tal que $E(X) = \int_\Omega XdP$ existe y $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$ es un objeto aleatorio, entonces $g(y) = E(X|Y = y)$ es la única función c.s. (P_Y) que va de $(\Omega', \mathcal{F}')$ en $(\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ que satisface

$$\int_{[Y\in B]} XdP = \int_B g(y)dP_Y(y) = \int_B E[X|Y = y]dP_Y \quad \text{con} B \in \mathcal{F}'.$$

Si hacemos $h(\omega) = g(Y(\omega))$ tenemos una función $h : (\Omega, \mathcal{F}) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ que es medible porque es composición de medibles, además se tiene el esquema siguiente

$$\begin{tikzcd}
(\Omega, \mathcal{F}, P) \arrow[r, "Y"] \arrow[rr, bend left, "X"] \arrow[rr, bend right, "h = g \circ Y"'] & (\Omega', \mathcal{F}', P_Y) \arrow[r, "g"] & (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))
\end{tikzcd}$$

De este modo h es una variable aleatoria extendida, en donde $h(\omega)$ es la esperanza condicional de X dado que Y tomó el valor y en ω, es decir $Y(\omega) = y$.

h mide el valor promedio de X dado Y pero h está definida en Ω no en Ω'.
Además se cumple que $\int_{[Y\in B]} XdP = \int_{[Y\in B]} hdP$ para $B \in \mathcal{F}'$ que es equivalente a la expresión

$$\int_{[Y\in B]} XdP = \int_B g(y)dP_Y(y).$$

Probaremos ahora la igualdad $\int_{[Y\in B]} hdP = \int_{[Y\in B]} XdP$.

$$\begin{aligned}\int_{[Y\in B]} hdP &= \int_{\Omega} g\left(Y(\omega)\right) I_B\left(Y(\omega)\right) dP(\omega) = \int_{\Omega'} g(y) I_B(y) dP_Y(y) \\ &= \int_B g(y)dP_Y(y) = \int_{[Y\in B]} XdP.\end{aligned}$$

Como $[Y\in B] = Y^{-1}(B) = \{\omega\in\Omega : Y(\omega)\in B\}$ la igualdad $\int_{[Y\in B]} XdP = \int_{[Y\in B]} hdP$ se puede expresar simplemente

$$\int_C hdP = \int_C XdP \quad \text{para cada } C\in Y^{-1}(\mathcal{F}'),$$

donde $Y^{-1}(\mathcal{F}') = \{Y^{-1}(B) : B\in\mathcal{F}'\}$.

Si $Y:(\Omega,\mathcal{F}) \longrightarrow (\Omega',\mathcal{F}')$ es un objeto aleatorio se define el σ-álgebra inducido por Y como $\mathcal{F}(Y) = Y^{-1}(\mathcal{F}')$ y si $Y=(Y_1,\ldots,Y_n)$ es un vector de objetos aleatorios entonces $\mathcal{F}(Y)$ consiste de todos los conjuntos $[Y\in D]$ con $D\in\mathcal{B}(\mathbb{R}^n)$.

$\mathcal{F}(Y)$ es el mínimo σ-álgebra que hace medible a Y como veremos adelante.

Ejemplo 1.3.1. a) Si $Y:(\mathbb{R},\mathcal{B}(\mathbb{R})) \longrightarrow (\mathbb{R},\mathcal{B}(\mathbb{R}))$ es tal que $Y(\omega)=2$ para todo $\omega\in\mathbb{R}$ entonces $\mathcal{F}(Y) = Y^{-1}(\mathcal{F}') = Y^{-1}(\mathcal{B}(\mathbb{R})) = \{Y^{-1}(B) : B\in\mathcal{B}(\mathbb{R})\}$ pero

$$Y^{-1}(B) = \begin{cases}\mathbb{R} & \text{si } 2\in B\\ \varnothing & \text{si } 2\notin B.\end{cases}$$

Por lo tanto $\mathcal{F}(Y) = \{\varnothing, \mathbb{R}\}$

b)Si $Y : (\Omega, 2^{\Omega}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ es tal que

$$Y(\omega) = \begin{cases} 0 & \text{si } \omega \in A_1 \\ 1 & \text{si } \omega \in A_2 \\ 2 & \text{si } \omega \in A_3, \end{cases}$$

con $\{A_1, A_2, A_3\}$ partición de Ω (estamos suponiendo que Ω tiene al menos 3 elementos) entonces $\mathcal{F}(Y) = Y^{-1}(\mathcal{F}')$ pero

$$Y^{-1}(B) = \begin{cases} \varnothing & \text{si } 0 \notin B, 1 \notin B, 2 \notin B \\ A_1 & \text{si } 0 \in B, 1 \notin B, 2 \notin B \\ A_2 & \text{si } 0 \notin B, 1 \in B, 2 \notin B \\ A_3 & \text{si } 0 \notin B, 1 \notin B, 2 \in B \\ A_1 \cup A_2 & \text{si } 0 \in B, 1 \in B, 2 \notin B \\ A_1 \cup A_3 & \text{si } 0 \in B, 1 \notin B, 2 \in B \\ A_2 \cup A_3 & \text{si } 0 \notin B, 1 \in B, 2 \in B \\ \Omega & \text{si } 0 \in B, 1 \in B, 2 \in B \end{cases}$$

que es un σ-álgebra contenido en $\mathcal{F} = 2^{\Omega}$. ■

El siguiente teorema nos indica algunas propiedades del σ-álgebra inducido.

Teorema 1.3.2. *Sea $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$ un objeto aleatorio, entonces:*

a) $\mathcal{F}(Y)$ es el σ-álgebra más pequeño en Ω que hace medible a Y.

b) Si $\Omega' = \prod_{j=1}^{\infty} \Omega_j$ $\mathcal{F}' = \prod_{j=1}^{\infty} \mathcal{F}_j$ y $Y = (Y_1, Y_2, \ldots)$ donde $Y_j : (\Omega, \mathcal{F}) \longrightarrow (\Omega_j, \mathcal{F}_j)$ entonces $\mathcal{F}$ es el más pequeño σ-álgebra en Ω que hace medible a cada Y_j.

c) Si $Z : (\Omega, \mathcal{F}(Y)) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ *entonces* $Z = f \circ Y$ *para alguna* $f : (\Omega', \mathcal{F}') \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$.

Recíprocamente, si $Z = f \circ Y$ *y* $f : (\Omega', \mathcal{F}') \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ *entonces* $Z : (\Omega, \mathcal{F}(Y)) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$

Demostración:

a) $\mathcal{F}(Y) = Y^{-1}(\mathcal{F}')$ es σ-álgebra sobre Ω ya que

i) $Y^{-1}(\varnothing) = \varnothing$, $Y^{-1}(\Omega') = \Omega$.

ii) $Y^{-1}(\Omega' - B) = Y^{-1}(\Omega') - Y^{-1}(B) = \Omega - Y^{-1}(B) \qquad B \in \mathcal{F}'$.

iii) $Y^{-1}(\bigcup_{i=1}^{\infty} B_i) = \bigcup_{i=1}^{\infty} Y^{-1}(B_i) \qquad B_i \in \mathcal{F}'$.

Además si $B \in \mathcal{F}'$ entonces $Y^{-1}(B) \in \mathcal{F}(Y)$ por lo que $\mathcal{F}(Y)$ hace medible a Y vista como función de $(\Omega, \mathcal{F}(Y))$ en $(\Omega', \mathcal{F}')$.

Supongamos que $\mathcal{G}$ es un σ-álgebra sobre Ω que haga medible a Y, es decir, que para todo $B \in \mathcal{F}'$ se tiene que $Y^{-1}(B) \in \mathcal{G}$, entonces $\mathcal{F}(Y) = \{Y^{-1}(B) : B \in \mathcal{F}'\} \subseteq \mathcal{G}$

b) Un vector o sucesión aleatoria Y es medible si y sólo si cada Y_j es medible y luego se aplica a) a cada coordenada

c) Supongamos que $Z : (\Omega, \mathcal{F}(Y)) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ y que Z es de la forma $Z = I_C$ para algún $C \in \mathcal{F}(Y)$ por tanto existe un $B \in \mathcal{F}'$ tal que $C = Y^{-1}(B) = [Y \in B]$.

Si $f = I_B$ entonces $f \circ Y = I_{[Y \in B]} = I_C = Z$.

Ahora supongamos que $Z = \sum_{k=1}^{n} \alpha_k I_{C_k}$ con $\alpha_k \in \mathbb{R}$ y $C_k \in \mathcal{F}(Y)$ para todo $k = 1, 2, \ldots, n$ entonces $I_{C_k} = f_k \circ Y$ igual que antes y luego $Z = f \circ Y$ donde $f = \sum_{k=1}^{n} \alpha_k f_k$.
En general si $Z_1, Z_2, \ldots$ son simples tales que $Z_n \longrightarrow Z$ en donde $Z_n = f_n \circ Y$ para todo $n \in \mathbb{N}$ hacemos

$$f = \begin{cases} \lim\limits_{n\to\infty} f_n & \text{donde exista} \\ 0 & \text{si no existe.} \end{cases}$$

Notemos que f así definida es medible (ver Dudley, 1989, sección 4.2), entonces $Z(\omega) = \lim\limits_{n\to\infty} Z_n(\omega) = \lim\limits_{n\to\infty} (f_n \circ Y)(\omega) = f(Y(\omega))$ c.s. (P_Y). Ahora, si $Z = f \circ Y$ con $f: (\Omega', \mathcal{F}') \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ y $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$ donde Y y f son medibles, entonces $Z : (\Omega, \mathcal{F}(Y)) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ es medible porque la composición de funciones medibles es medible. $\square$

La relación $\int_C h dP = \int_C X dP$ para todo $C \in \mathcal{F}(Y)$ cuando se tiene que $h = g \circ Y$ y $g(y) = E(X|Y = y)$ tiene importancia en resultados posteriores, aquí $g : (\Omega', \mathcal{F}') \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ y por tanto $h : (\Omega, \mathcal{F}(Y)) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$.

Teorema 1.3.3. *Sean $X : (\Omega, \mathcal{F}, P) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ tal que $E(X)$ existe y $\mathcal{G} \subseteq \mathcal{F}$ un $\sigma-$álgebra. Entonces hay una función $E(X|\mathcal{G})$ que va de $(\Omega, \mathcal{G})$ en $(\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ que es la esperanza condicional de X dado $\mathcal{G}$ y que cumple con $\int_C X dP = \int_C E(X|\mathcal{G}) dP$ para todo $C \in \mathcal{G}$. Además si hay otra función que cumpla con la igualdad de arriba entonces es igual a $E(X|\mathcal{G})$ c.s. (P).*

Demostración:

Sea $\lambda(C) = \int_C XdP$ con $C \in \mathcal{G}$ entonces $\lambda : \mathcal{G} \longrightarrow \overline{\mathbb{R}}$ es numerablemente aditiva ya que $\lambda\left(\bigcup_{j=1}^{\infty} C_j\right) = \int_{\bigcup_{j=1}^{\infty} C_j} XdP = \sum_{j=1}^{\infty} \int_{C_j} XdP$ si $C_1, C_2, \ldots$ son ajenos.

Luego $\lambda\left(\bigcup_{j=1}^{\infty} C_j\right) = \sum_{j=1}^{\infty} \lambda(C_j)$, además $\lambda \ll P$ porque si $P(C) = 0$ entonces $\lambda(C) = 0$ ya que estaríamos integrando sobre un conjunto de medida cero.

Luego por el Teorema de Radon-Nikodym existe $h : (\Omega, \mathcal{G}) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ tal que para todo $C \in \mathcal{G}$, $\lambda(C) = \int_C hdP$, h es única c.s. (P) y definimos por lo tanto $E(X|\mathcal{G}) = h$. $\square$

Notemos que $E(X|\mathcal{G})$ es una variable aleatoria $\mathcal{G}$-medible que en general no es determinista y que X es medible con respecto a $\mathcal{F}$ pero NO sabemos si lo es con respecto a $\mathcal{G}$ porque se podría pensar que $X = E(X|\mathcal{G})$ pero esto sólo es posible si X es $\mathcal{G}$-medible.

Observación. Si $g(y) = E(X|Y = y)$ y $h(\omega) = g(Y(\omega))$ se tiene por el Teorema 1.3.3 que $h = E(X|\mathcal{G})$ donde $\mathcal{G} = \mathcal{F}(Y)$ porque ya vimos que $h : (\Omega, \mathcal{F}(Y)) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ se puede escribir como $h = E(X|Y)$, así por ejemplo si $E(X|Y = y) = y^3$ entonces $E(X|Y) = Y^3$.

Hemos visto que $g(y) = E(X|Y = y)$ con $y \in \Omega'$, pero haciendo la composición con Y, es decir, $h = g \circ Y$ se transfiere a una función cuyo dominio es Ω, recíprocamente cualquier $E(X|\mathcal{G})$ con $\mathcal{G} \subseteq \mathcal{F}$ surge de un objeto aleatorio Y haciendo $\mathcal{G} = \mathcal{F}(Y)$, esto es, podemos hacer $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega, \mathcal{G})$ definiendo $Y(\omega) = \omega$ y por tanto $Y^{-1}(\mathcal{F}') = Y^{-1}(\mathcal{G}) = \mathcal{G}$, así $g(y) = E(X|Y = y)$ entonces $h = E(X|\mathcal{F}(Y)) = E(X|\mathcal{G})$.

Ahora $E(X|\mathcal{G}) = E(X|Y)$ es el promedio de X dado que Y es conocido. Pero, ¿Qué significa que Y una función que va de $(\Omega, \mathcal{F})$ en $(\Omega, \mathcal{G})$ sea conocido? Los eventos en Y son de la forma $[Y \in D]$ con $D \in \mathcal{G}$ pero como Y es un mapeo identidad restringido entonces $[Y \in D] = D$ y de este modo para cada evento, se tienen dos respuestas: ocurre D o no ocurre D, $E(X|\mathcal{G})$ se puede interpretar como el valor promedio de $X(\omega)$ dado que sabemos que para todo $D \in \mathcal{G}$, ω está o no está en D.

Ejemplo 1.3.4. Sea Y discreta que toma los valores $y_1, y_2, \ldots$, $\Omega' = \{y_1, y_2, \ldots\}$, $\mathcal{F}' = 2^{\Omega'}$, $p_i = P[Y = y_i] > 0 \quad i = 1, 2, \ldots$.

Se mostró antes que $g(y_i) = E(X|Y = y_i) = \dfrac{1}{P[Y = y_i]} \displaystyle\int_{[Y=y_i]} X dP$.

Sea $h = E(X|Y)$, esto es, $h(\omega) = g(Y(\omega))$ entonces h tiene el valor constante $g(y_i)$ sobre el conjunto $[Y = y_i] = \{\omega \in \Omega : Y(\omega) = y_i\} \in \mathcal{F}$.

Además $\mathcal{G} = Y^{-1}(\mathcal{F}')$ con $[Y = y_i] \in \mathcal{G}$, es más, $\mathcal{G}$ consiste de todas las uniones de conjuntos de la forma $[Y = y_i]$. El conocimiento del valor de $Y(\omega)$ es equivalente al conocimiento para cada $D \in \mathcal{G}$ de la pertenencia o no de ω a D. ∎

Ejemplo 1.3.5. Sean X y Y variables aleatorias con densidad conjunta f, $\Omega = \mathbb{R}^2$. $\mathcal{F} = \mathcal{B}(\mathbb{R}^2)$, $P(A) = \iint_A f(x, y)dxdy$ para $A \in \mathcal{F}$. $X(x, y) = x$, $Y(x, y) = y$.

Tomamos $\Omega' = \mathbb{R}$, $\mathcal{F}' = \mathcal{B}(\mathbb{R})$, ya hemos visto que $g(y) = E(X|Y = y) = \displaystyle\int_{-\infty}^{\infty} xf(x|y)dx$ donde $f(x|y)$ es la densidad condicional de X dado que Y toma el valor y.

Sea $h = E(X|Y)$, esto es, $h(\omega) = g(Y(\omega))$ pero como los elementos $\omega \in \Omega$

son de la forma (x, y) entonces $h(x, y) = g(Y(x, y)) = g(y)$.
Así $E(X|Y)$ es constante sobre líneas horizontales, también $Y^{-1}(\mathcal{F}')$ consta de todos los conjuntos $\mathbb{R} \times A$ con $A \in \mathcal{B}(\mathbb{R})$.
Sucede que $y \in A$ si y sólo si $(x, y) \in \mathbb{R} \times A$, la información acerca de $Y(w)$ es equivalente a la información acerca de la pertenencia de ω en conjuntos de $\mathcal{G}$. ■

Teorema 1.3.6. *Sean $(\Omega, \mathcal{F}, P)$ espacio de probabilidad $\mathcal{G} \subseteq \mathcal{F}$ un sub σ álgebra de $\mathcal{F}$, $A \in \mathcal{F}$ fijo. Entonces hay una función $P(A|\mathcal{G}) : (\Omega, \mathcal{G}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ que se llama la probabilidad condicional de A dado $\mathcal{G}$ tal que $P(C \cap A) = \int_C P(A|\mathcal{G})dP$ pcra todo $C \in \mathcal{G}$. Si hubiera otra función que cumpla con esta igualdad entonces es igual a $P(A|\mathcal{G})$ c.s. (P), de hecho se tiene que $E(I_A|\mathcal{G}) = P(A|\mathcal{G})$ c.s. (P).*

Demostración:
Sea $\lambda(C) = P(C \cap A)$, $\lambda : \mathcal{G} \longrightarrow [0, 1]$ es numerablemente aditiva porque P lo es.
$\lambda(\varnothing) = P(\varnothing \cap A) = P(\varnothing) = 0$ además $\lambda \ll P$ porque $C \cap A \subseteq C$.
Luego la existencia de $h : (\Omega, \mathcal{G}) \longrightarrow (\mathbb{R}. \mathcal{B}(\mathbb{R}))$ con $\lambda(C) = \int_C hdP$ para $C \in \mathcal{F}$ se obtiene del Teorema de Radon-Nikodym, igual que la unicidad.
Entonces $h = P(A|\mathcal{G})$ y la conexión entre la esperanza condicional y la probabilidad condicional es como sigue

$$\lambda(C) = P(C \cap A) = \int_C I_A dP = \int_C E(I_A|\mathcal{G})dP \quad C \in \mathcal{G},$$

donde $I_A : (\Omega, \mathcal{F}, P) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ es medible y $E(I_A)$ existe puesto que

$E(I_A) = P(A)$.
También $\int_C I_A dP = \int_C P(A|\mathcal{G})dP = P(C \cap A)$ por lo que $P(A|\mathcal{G}) = E(I_A|\mathcal{G})$. $\square$

Observación. Si $g(y) = P(A|Y = y)$ con $A \in \mathcal{F}$, $Y:(\Omega, \mathcal{F}, P) \longrightarrow (\Omega', \mathcal{F}', P_Y)$ entonces $g(y) = E(I_A|Y = y)$ c.s. (P_Y).
Pero si $h(\omega) = g(Y(\omega))$ entonces $h = E(I_A|\mathcal{F}(Y))$ por tanto $h = P(B|\mathcal{F}(Y))$, es decir, si $g(y) = P(A|Y = y)$ y $h(\omega) = g(Y(\omega))$ entonces $h = P(A|\mathcal{G})$, $\mathcal{G} = \mathcal{F}(Y)$ y se puede denotar que $h = P(A|Y)$, para algunos cálculos $E(X|Y = y)$ es más importante que $E(X|Y) = E(X|\mathcal{F}(Y))$.

1.4. Propiedades de la Esperanza Condicional

Ahora veremos las propiedades de $E(X|Y)$ y$E(X|Y=y)$ con la notación y los términos en que se trataron estos conceptos en las páginas anteriores, cada vez que aparezca $E(X|Y)$ en realidad es $E(X|\mathcal{F}(Y))$, consideramos también $X_1, X_2, X_3, \ldots$ una sucesión de variables aleatorias extendidas definidas en $(\Omega, \mathcal{F}, P)$. $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$ un objeto aleatorio y cuando no se especifíque con respecto a qué medida de probabilidad se establece una igualdad casi segura nos referimos a la medida P y $\mathcal{G}$ se considera un sub σ-álgebra de $\mathcal{F}$.

Teorema 1.4.1. *a) Sea $X = k$ c.s. entonces $E(X|\mathcal{G}) = k$ c.s.*

b) Si $X = k$ c.s. entonces $E(X|Y = y) = k$ c.s. (P_Y).

c) Si $X_1 \leqslant X_2$ c.s. entonces $E(X_1|\mathcal{G}) \leqslant E(X_2|\mathcal{G})$ c.s.

d) Si $X_1 \leqslant X_2$ c.s. entonces $E(X_1|Y = y) \leqslant E(X_2|Y = y)$ c.s. (P_Y)

e) $\left|E(X|\mathcal{G})\right| \leqslant E(|X||\mathcal{G})$ c.s.

f) $\left|E(X|Y = y)\right| \leqslant E(|X||Y = y)$ c.s. (P_Y)

Demostración:

a) Notemos primero que si $X = k$ entonces X es $\mathcal{G}$-medible para todo $\mathcal{G} \subseteq \mathcal{F}$ sub σ-álgebra, porque las imágenes inversas de cualquier conjunto son o bien $\varnothing$ o bien Ω, que pertenecen a $\mathcal{G}$ siempre.

Luego por el Teorema 1.3.3 se cumple la igualdad

$$\int_C X dP = \int_C E(X|\mathcal{G}) dP$$

para todo $C \in \mathcal{G}$ en donde $E(X|\mathcal{G})$ es $\mathcal{G}$-medible, por tanto sólo tenemos que ver si $\int_C X dP = \int_C k dP$ para todo $C \in \mathcal{G}$. Pero $X = k$ c.s. es $\mathcal{G}$-medible por la observación de arriba, entonces $\int_C X dP = \int_C k dP$ para $C \in \mathcal{G}$ y por tanto $k = E(X|\mathcal{G})$ c.s.

b) Si definimos $g(y) = k$, $g : (\Omega', \mathcal{F}') \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ entonces g es $\mathcal{F}'$–medible por la misma razón que antes.

$E(X|Y = y)$ cumple por el Teorema 1.2.5 con

$$\int_{[Y \in B]} X dP = \int_B E(X|Y = y) dP_Y(y) \quad \text{para todo} \quad B \in \mathcal{F}',$$

donde $E(X|Y = y)$ es $\mathcal{F}'$–medible, $y \in \Omega'$.

Entonces sólo tenemos que ver si se cumple

$$\int_{[Y \in B]} X dP = \int_B k dP_Y(y) \quad \text{para todo} \quad B \in \mathcal{F}'.$$

Pero $X = k$ c.s. y k es $\mathcal{F}'$–medible entonces

$$\int_{[Y \in B]} X dP = \int_{[Y \in B]} k dP = kP[Y \in B] \quad \text{y}$$

$$\int_B k dP_Y(y) = kP_Y(B) = kP[Y \in B] \quad \text{por tanto } k = E(X|Y = y) \text{ c.s. } (P_Y).$$

c) Como $X_1 \leqslant X_2$ c.s. entonces $\int_C X_1 dP \leqslant \int_C X_2 dP$ para todo $C \in \mathcal{G}$, pero $\int_C X_1 dP = \int_C E(X_1|\mathcal{G})dP$ y $\int_C X_2 dP = \int_C E(X_2|\mathcal{G})dP$.
De aqui que $\int_C E(X_1|\mathcal{G}) \leqslant \int_C E(X_2|\mathcal{G})$ y como esta desigualdad se cumple para todo $C \in \mathcal{G}$ entonces $E(X_1|\mathcal{G}) \leqslant E(X_2|\mathcal{G})$ c.s.

d) Como $X_1 \leqslant X_2$ c.s. entonces $\int_{[Y\in B]} X_1 dP \leqslant \int_{[Y\in B]} X_2 dP$ para todo $B \in \mathcal{F}'$ pero $\int_{[Y\in B]} X_1 dP = \int_B E(X_1|Y = y)dP_Y(y)$ y $\int_{[Y\in B]} X_2 dP = \int_B E(X_2|Y = y)dP_Y(y)$ por lo que

$$\int_B E(X_1|Y = y)dP_Y(y) \leqslant \int_B E(X_2|Y = y)dP_Y(y) \quad \text{para todo} \quad B \in \mathcal{F}',$$

y por la misma razón que antes se tiene que $E(X_1|Y = y) \leqslant E(X_2|Y = y)$ c.s. (P_Y).

e) Sabemos que $-|X| \leqslant X$, de este modo $-|X| \leqslant X \leqslant |X|$ y aplicando c) obtenemos

$$E\left(-|X||\mathcal{G}\right) \leqslant E(X|\mathcal{G}) \leqslant E\left(|X||\mathcal{G}\right) \quad \text{c.s.}$$

Para llegar al resultado sólo falta mostrar que $E\left(-|X||\mathcal{G}\right) = -E\left(|X||\mathcal{G}\right)$ c.s.
Sea $C \in \mathcal{G}$ entonces

$$\int_C E\left(-|X||\mathcal{G}\right) = \int_C -|X|dP = -\int_C |X|dP = -\int_C E\left(|X||\mathcal{G}\right)dP =$$

$$\int_C -E\left(|X| \middle| \mathcal{G}\right) dP \quad \text{por lo que} \quad E\left(|X| \middle| \mathcal{G}\right) = -E\left(|X| \middle| \mathcal{G}\right) \text{ c.s.}$$

De esta manera

$$-E\left(|X| \middle| \mathcal{G}\right) \leqslant E\left(X \middle| \mathcal{G}\right) \leqslant E\left(|X| \middle| \mathcal{G}\right) \quad \text{c.s.}$$

y luego

$$\left|E\left(X \middle| \mathcal{G}\right)\right| \leqslant E\left(|X| \middle| \mathcal{G}\right) \quad \text{c.s.}$$

f) De nuevo $-|X| \leqslant X \leqslant |X|$ y por d) se tiene que

$$E\left(-|X| \middle| Y = y\right) \leqslant E\left(X \middle| Y = y\right) \leqslant E\left(|X| \middle| Y = y\right) \quad \text{c.s.} \quad (P_Y)$$

de la misma manera que en e) se probaría que

$$E\left(-|X| \middle| Y = y\right) = -E\left(|X| \middle| Y = y\right) \quad \text{c.s.} \quad (P_Y)$$

por lo que

$$\left|E\left(X \middle| Y = y\right)\right| \leqslant E\left(|X| \middle| Y = y\right) \quad \text{c.s.} \quad (P_Y). \; \square$$

Veremos que la esperanza condicional es lineal.

Teorema 1.4.2. *a) Si $a, b \in \mathbb{R}$, $E(X_1)$ y $E(X_2)$ existen entonces $E(aX_1 + bX_2|\mathcal{G}) = aE(X_1|\mathcal{G}) + bE(X_2|\mathcal{G})$ c.s.*

b) Si $a, b \in \mathbb{R}$, $E(X_1)$ y $E(X_2)$ existen entonces $E(aX_1 + bX_2|Y = y) = aE(X_1|Y = y) + bE(X_2|Y = y)$.

Demostración:

a) Tomemos $C \in \mathcal{G}$ entonces

$$\int_C E(aX_1 + bX_2|\mathcal{G})dP = \int_C (aX_1 + bX_2)dP = a\int_C X_1 dP + b\int_C X_2 dP$$

$$= a\int_C E(X_1|\mathcal{G})dP + b\int_C E(X_2|\mathcal{G})dP = \int_C (aE(X_1|\mathcal{G}) + bE(X_2|\mathcal{G}))dP,$$

que se cumple para cualquier $C \in \mathcal{G}$, por tanto

$$E(aX_1 + bX_2|\mathcal{G}) = aE(X_1|\mathcal{G}) + bE(X_2|\mathcal{G}) \quad \text{c.s.}$$

b) Tomemos $B \in \mathcal{F}'$ entonces

$$\int_B E(aX_1 + bX_2|Y = y)dP_Y(y) = \int_{[Y\in B]} (aX_1 + bX_2)dP = a\int_{[Y\in B]} X_1 dP$$

$$+b\int_{[Y\in B]} X_2 dP = a\int_B E(X_1|Y = y)dP_Y(y) + b\int_B E(X_2|Y = y)dP_Y(y)$$

$$= \int_B (aE(X_1|Y = y) + bE(X_2|Y = y))dP_Y(y) \text{ y luego}$$

$$E(aX_1 + bX_2|Y = y) = aE(X_1|Y = y) + bE(X_2|Y = y). \ \square$$

Se tiene un teorema de convergencia monótona para esperanzas condicionales.

Teorema 1.4.3. *Supongamos que* $\{X_n : n \in \mathbb{N}\}$ *es una sucesión de variables aleatorias tales que* $X_n \geqslant 0$ *c.s. para todo* $n \in \mathbb{N}$ *y* $X_n \uparrow X$ *c.s. con* X *integrable entonces:*

a) $E(X_n|\mathcal{G}) \uparrow E(X|\mathcal{G})$ *c.s.*

b) $E(X_n|Y = y) \uparrow E(X|Y = y)$ *c.s.* (P_Y).

c) $E\left(\sum_{n=1}^{\infty} X_n \Big| \mathcal{G}\right) = \sum_{n=1}^{\infty} E(X_n|\mathcal{G})$ *c.s.*

d) $E\left(\sum_{n=1}^{\infty} X_n \Big| Y = y\right) = \sum_{n=1}^{\infty} E(X_n|Y = y)$ *c.s.* (P_Y).

Y cuando $A_1, A_2 \ldots \in \mathcal{F}$ *son ajenos entonces:*

e) $P\left(\bigcup_{n=1}^{\infty} A_n \Big| \mathcal{G}\right) = \sum_{n=1}^{\infty} P(A_n|\mathcal{G})$ *c.s.*

f) $P\left(\bigcup_{n=1}^{\infty} A_n \Big| Y = y\right) = \sum_{n=1}^{\infty} P(A_n|Y = y)$ *c.s.* (P_Y).

Demostración:

a) Se tiene para todo $C \in \mathcal{G}$ que $\int_C X_n dP = \int_C E(X_n|\mathcal{G})dP$ pero por hipótesis $0 \leqslant X_1 \leqslant X_2 \leqslant X_3 \leqslant \cdots$ c.s. entonces $0 \leqslant E(X_1|\mathcal{G}) \leqslant E(X_2|\mathcal{G}) \leqslant \cdots$ c.s. por Teorema 1.4.1.
Esto quiere decir que para casi todo $\omega \in \Omega$ tenemos
$0 \leqslant E(X_1|\mathcal{G})(\omega) \leqslant E(X_2|\mathcal{G})(\omega) \leqslant \cdots$
Sabemos que $E(X_n|\mathcal{G})$ es $\mathcal{G}-$medible para todo $n \in \mathbb{N}$ y que el límite de funciones medibles es medible, así que definimos

$$h(\omega) = \begin{cases} \lim\limits_{n \to \infty} E(X_n|\mathcal{G})(\omega) & \text{si existe el límite} \\ 0 & \text{si no existe.} \end{cases}$$

Sólo faltaría ver que $h = E(X|\mathcal{G})$ c.s.

Por definición de h se tiene que $E(X_n|\mathcal{G}) \uparrow h$ c.s. y por el Teorema de Convergencia Monótona Clásico sucede que $\int_C X_n dP \uparrow \int_C X dP$ cuando $C \in \mathcal{G}$ porque $\int_C X_n dP = \int_\Omega X_n I_C dP$

Pero $X_n I_C \uparrow X I_C$ c.s.

$$\text{entonces} \quad \int_\Omega X_n I_C dP \uparrow \int_\Omega X I_C dP = \int_C X dP,$$

$$\text{además} \quad \int_C X_n dP = \int_C E(X_n|\mathcal{G}) dP \quad \text{y} \quad \int_C X dP = \int_C E(X|\mathcal{G}) dP$$

$$\text{por tanto} \quad \int_C E(X_n|\mathcal{G}) \uparrow \int_C E(X|\mathcal{G}) \quad \text{c.s.}$$

Y como $E(X_n|\mathcal{G}) \uparrow h$ c.s. que implica $\int_C E(X_n|\mathcal{G}) dP \uparrow \int_C h dP$ entonces $h = E(X|\mathcal{G})$ c.s. y de este modo $E(X_n|\mathcal{G}) \uparrow E(X|\mathcal{G})$ c.s.

b) Igual que a).

c) Se sabe por el Teorema 1.4.1 que si $n \in \mathbb{N}$ fija entonces

$$E\left(\sum_{k=1}^{n} X_k \Big| \mathcal{G}\right) = \sum_{k=1}^{n} E(X_k|\mathcal{G}) \quad \text{c.s.}$$

y si $Z_n = \sum_{k=1}^{n} X_k$ y $Z = \sum_{k=1}^{\infty} X_k$ sucede que $Z_n \uparrow Z$ c.s. porque $X_n \geqslant 0$ c.s. para todo $n \in \mathbb{N}$ y aplicando a) $E(Z_n|\mathcal{G}) \uparrow E(Z|\mathcal{G})$ c.s.

$$\text{Esto es} \quad E\left(\sum_{n=1}^{n} X_k \Big| \mathcal{G}\right) \uparrow E\left(\sum_{k=1}^{\infty} X_k \Big| \mathcal{G}\right) \quad \text{c.s.}$$

$$\text{Que equivale a} \quad \sum_{n=1}^{n} E(X_k \big| \mathcal{G}) \uparrow E\left(\sum_{k=1}^{\infty} X_k \Big| \mathcal{G}\right) \quad \text{c.s.}$$

$$\text{Pero} \quad \sum_{k=1}^{\infty} E(X_k|\mathcal{G}) = \lim_{n\to\infty} \sum_{k=1}^{n} E(X_k|\mathcal{G}) = E(\sum_{k=1}^{\infty} X_k|\mathcal{G}) \quad \text{c.s.}$$

está última igualdad es por a).

$$\text{Por tanto} \quad \sum_{k=1}^{\infty} E(X_k|\mathcal{G}) = E(\sum_{k=1}^{\infty} X_k|\mathcal{G}).$$

d) Igual que c).

e) Tomemos $X_k = I_{A_k} \geqslant 0 \quad k = 1, 2, \ldots$ funciones medibles entonces $\sum_{k=1}^{\infty} X_k = I_{\bigcup_{k=1}^{\infty} A_k}$ porque la sucesión de eventos en $A_1, A_2, \ldots$ es ajena.

$$\text{Luego} \quad E\left(\sum_{k=1}^{\infty} X_k \Big| \mathcal{G}\right) = E(I_{\bigcup_{k=1}^{\infty} A_k}|\mathcal{G}) = P\left(\bigcup_{k=1}^{\infty} A_k \Big| \mathcal{G}\right) \quad \text{c.s.}$$

por Teorema 1.3.6.

$$\text{Ahora aplicamos c),} \quad E(\sum_{k=1}^{\infty} X_k|\mathcal{G}) = \sum_{k=1}^{\infty} E(X_k|\mathcal{G}) \quad \text{es decir,}$$

$$\sum_{k=1}^{\infty} P(A_k|\mathcal{G}) = P(\bigcup_{k=1}^{\infty} A_k|\mathcal{G}) \quad \text{c.s.} \ \square$$

f) Igual que e).

Teorema 1.4.4. *Supongamos que $E(X)$ existe. Entonces:*

a) $E\Big(E(X|\mathcal{G})\Big) = E(X)$.

b) $\displaystyle\int_{\Omega'} E(X|Y = y)dP_Y(y) = E(X)$.

Demostración:

a) Por definición $\int_C XdP = \int_C E(X|\mathcal{G})dP$ para todo $C \in \mathcal{G}$.

Luego $\int_\Omega XdP = \int_\Omega E(X|\mathcal{G})dP$ pero $E(X) = \int_\Omega XdP$ y

$E[E(X|\mathcal{G})] = \int_\Omega E(X|\mathcal{G})dP$ por tanto $E(X) = E(E(X|\mathcal{G}))$.

b) Por definición $\int_{[Y\in B]} XdP = \int_B E(X|Y=y)dP_Y(y)$ para todo $B \in \mathcal{F}'$.

Como $\mathcal{F}'$ es $\sigma-$álgebra en Ω' entonces $\Omega' \in \mathcal{F}'$ y así

$$\int_{[Y\in\Omega']} XdP = \int_{\Omega'} E(X|Y=y)dP_Y(Y),$$

pero $[Y \in \Omega'] = \{\omega \in \Omega : Y(\omega) \in \Omega'\} = \Omega$.

De este modo $\int_\Omega XdP = \int_{\Omega'} E(X|Y=y)dP_Y(y)$.

Concluyendo que $E(X) = \int_{\Omega'} E(X|Y=y)dP_Y(y)$. $\Box$

Una versión del teorema de Convergencia Dominada para esperanzas condicionales.

Teorema 1.4.5. *Si $|X_n| \leqslant Z$ c.s. para todo $n \in \mathbb{N}$ con $E(Z) < \infty$ y $X_n \xrightarrow[n\to\infty]{} X$ c.s. entonces*

a) $E(X_n|\mathcal{G}) \longrightarrow E(X|\mathcal{G})$ c.s. y en L^1

b) $E(X_n|Y=y) \longrightarrow E(X|Y=y)$ *c.s.* (P_Y).

Demostración:

a) Sea $Z_n = \sup\limits_{k\geqslant n}|X_k - X|$ es decir $Z_1 = \sup\limits_{k\geqslant 1}|X_k - X| \geqslant \sup\limits_{k\geqslant 2}|X_k - X| = Z_2 \geqslant \cdots$ es una sucesión decreciente c.s. de funciones medibles porque el supremo de medibles es medible y como $X_n \xrightarrow[n\to\infty]{} X$ c.s. entonces $Z_n \downarrow 0$ c.s.

Tenemos que X_n y X son integrables porque $E(|X_n|) \leqslant E(Z) < \infty$ y X es límite de integrables, así, $E(X_n|\mathcal{G})$ y $E(X|\mathcal{G})$ son finitos c.s. y además $E(|X_n||\mathcal{G}) \leqslant E(Z|\mathcal{G})$ por Teorema 1.4.1.

Por otro lado

$$|E(X_n|\mathcal{G}) - E(X|\mathcal{G})| = |E(X_n - X|\mathcal{G})| \leqslant E(|X_n - X||\mathcal{G}) \leqslant E(Z_n|\mathcal{G}),$$

porque $Z_n = \sup\limits_{k\geqslant n}|X_k - X| = \sup\{|X_n - X|, |X_{n+1} - X|, \ldots\} \geqslant |X_n - X|$ y por Teorema 1.4.1.

Supongamos que $E(Z_n|\mathcal{G}) \downarrow h$ c.s. ya que la sucesión $\{E(Z_n|\mathcal{G}) : n\in\mathbb{N}\}$ es decreciente casi seguramente puesto que $Z_1 \geqslant Z_2 \geqslant \cdots \geqslant 0$ y entonces $E(Z_1|\mathcal{G}) \geqslant E(Z_2|\mathcal{G}) \geqslant \cdots \geqslant E(0|\mathcal{G}) = 0$.

Basta con mostrar que $h = 0$ c.s.

Para esto notemos que

$$|X_n - X| \leqslant |X_n| + |X| \leqslant Z + Z = 2Z \text{ para todo } n \in \mathbb{N}$$

$$\text{entonces} \quad Z_n = \sup\{|X_n - X|, |X_{n+1} - X|, \ldots\} \leqslant 2Z.$$

De este modo $0 \leqslant Z_n \leqslant 2Z$.

Luego $h \geqslant 0$ c.s. por ser límite de funciones mayores o iguales que cero,

$$\text{así} \quad 0 \leqslant \int_{\Omega} hdP \leqslant \int_{\Omega} E(Z_n|\mathcal{G})dP = \int_{\Omega} Z_n dP$$

por definición de esperanza condicional.

Por tanto $0 \leqslant \int_{\Omega} hdP \leqslant \int_{\Omega} Z_n dP$.

Pero $Z_n \downarrow 0$ c.s. entonces $\int_{\Omega} Z_n dP \downarrow 0$.

Así $\int_{\Omega} hdP = 0$ siendo $h \geqslant 0$ c.s. Concluimos que $h = 0$ c.s. Por otro lado se tenía la desigualdad

$$0 \leqslant \left|E(X_n|\mathcal{G}) - E(X|\mathcal{G})\right| \leqslant E(Z_n|\mathcal{G}.)$$

$$\text{Pero} \quad E(Z_n|\mathcal{G}) \downarrow 0 \quad \text{c.s.}$$

$$\text{por lo que} \quad \left|E(X_n|\mathcal{G}) - E(X|\mathcal{G})\right| \xrightarrow[n\to\infty]{} 0 \quad \text{c.s.}$$

$$\text{esto es} \quad E(X_n|\mathcal{G}) \xrightarrow[n\to\infty]{} E(X|\mathcal{G})$$

b) Igual que a). □

En el siguiente resultado a) y b) son una versión más general del Teorema de Convergencia Monótona para Esperanzas Condicionales, c) y d) son versiones del Lema de Fatou para Esperanzas Condicionales.

Teorema 1.4.6. *Si $\{X_n : n \in \mathbb{N}\}$ es una sucesión de variables aleatorias tales que $X_n \geqslant Z$ para todo $n \in \mathbb{N}$ donde $E(Z) > -\infty$ entonces:*

a) Si $X_n \uparrow X$ *c.s. entonces* $E(X_n|\mathcal{G}) \uparrow E(X|\mathcal{G})$ *c.s.*

b) Si $X_n \uparrow X$ *c.s. entonces* $E(X_n|Y = y) \uparrow E(X|Y = y)$ *c.s.* (P_Y).

c) $\liminf\limits_{n\to\infty} E(X_n|\mathcal{G}) \geqslant E\big(\liminf\limits_{n\to\infty} X_n|\mathcal{G}\big)$ *c.s.*

d) $\liminf\limits_{n\to\infty} E(X_n|Y = y) \geqslant E\big(\liminf\limits_{n\to\infty} X_n|Y = y\big)$ *c.s.* (P_Y).

Demostración:

a) Se tiene que $X_n = X_n^+ - X_n^-$ en donde $X_n^+ \geqslant 0$, $X_n^- \geqslant 0$.
Cuando $X_n \uparrow X$ se cumple que $X_n^+ \uparrow X^+$, porque $X_n \leqslant X_{n+1} \leqslant \cdots \leqslant X$.
Entonces $X_n^+ = \max\{X_n, 0\} \leqslant \max\{X_{n+1}, 0\} = X_{n+1}^+ \leqslant \cdots \leqslant X^+$ y por el Teorema 1.4.3 $E(X_n^+|\mathcal{G}) \uparrow E(X^+|\mathcal{G})$ c.s.
También $0 \leqslant X_n^- \leqslant Z^-$ donde Z^- es integrable ya que $Z = Z^+ - Z^-$ y $E(Z) > -\infty$.
Pero $E(Z) = E(Z^+) - E(Z^-)$ si Z^- no fuera integrable o $E(Z^-) = \infty$ entonces Z no sería integrable con $E(Z) > -\infty$.
La desigualdad $X_n^- \leqslant Z^-$ se verifica de la siguiente manera

$$X_n^-(\omega) = -\min\{X_n(\omega), 0\} = \begin{cases} -X_n(\omega) & \text{si } X_n(\omega) < 0 \\ 0 & \text{si } X_n(\omega) \geqslant 0, \end{cases}$$

y por hipótesis $X_n \geqslant Z$, es decir, $X_n(\omega) \geqslant Z(\omega)$ entonces si $Z(\omega) \geqslant 0$ tenemos que $X_n(\omega) \geqslant 0$.

Y si $Z^-(\omega) = 0$ entonces $X_n^-(\omega) = 0$ por lo que $X_n^-(w) \leqslant Z^-(\omega)$.
Ahora si $Z(\omega) < 0$ entonces $Z^-(\omega) = -Z(\omega) \geqslant -X_n(\omega)$ pero $X_n(\omega)$ puede ser mayor o igual que cero o menor que cero.
Si $X_n(\omega) \geqslant 0$ entonces $X_n^-(\omega) = 0$ y $-X_n(\omega) = 0$ pero $Z^-(\omega) > 0 = -X_n(\omega) = X_n^-(\omega)$.
Y si $X_n(\omega) < 0$ obtenemos que $X_n^-(\omega) = -X_n(\omega)$ pero $Z_n^-(\omega) \geqslant -X_n(\omega) = X_n^-(\omega)$ por tanto $X_n^- \leqslant Z^-$ que es integrable.
Como se había dicho antes $E(Z^-) < \infty$ porque de otro modo $E(Z)$ no sería mayor que $-\infty$.
Además $X_n^- \longrightarrow X^-$ y por el Teorema 1.4.5, $E(X_n^-|\mathcal{G}) \longrightarrow E(X^-|\mathcal{G})$ por lo que $E(X_n|\mathcal{G}) = E(X_n^+ - X_n^-|\mathcal{G}) = E(X_n^+|\mathcal{G}) - E(X_n^-|\mathcal{G})$ y así

$$E(X_n|\mathcal{G}) \longrightarrow E(X^+|\mathcal{G}) - E(X^-|\mathcal{G}) = E(X^+ - X^-|\mathcal{G}) = E(X|\mathcal{G}) \quad \text{c.s.}$$

b) Igual que a).

c) Sea $X_n' = \inf_{k \geqslant n} X_k = \inf\{X_n, X_{n+1}, \ldots\}$ que es medible porque el ínfimo de medibles es medible.
Se tiene una sucesión creciente debido a que $X_1' = \inf\{X_1, X_2, \ldots\} \leqslant \inf\{X_2, X_3, \ldots\} = X_2' \leqslant \ldots$
De este modo si $\lim_{n\to\infty} X_n' = \liminf_{n\to\infty} X_n = X'$ entonces $X_n' \uparrow X'$.
Esto resulta porque $\liminf X_n = \sup_n \inf_{k \geqslant n} X_k$.
Luego por a) se obtiene que $E(X_n'|\mathcal{G}) \uparrow E(X'|\mathcal{G})$ c.s.

Pero $X_n' \leqslant X_n$ entonces

$$E(X'|\mathcal{G}) = \lim_{n\to\infty} E(X_n'|\mathcal{G}) = \liminf_{n\to\infty} E(X_n'|\mathcal{G}) \leqslant \liminf_{n\to\infty} E(X_n|\mathcal{G}) \quad \text{c.s.}$$

$$\text{Así} \quad E(\liminf_{n\to\infty} X_n|\mathcal{G}) \leqslant \liminf_{n\to\infty} E(X_n|\mathcal{G}) \quad \text{c.s.}$$

d) Igual que c). $\square$

Tenemos una contraparte del teorema 1.4.6.

Teorema 1.4.7. *Si $\{X_n : n \in \mathbb{N}\}$ es una sucesión de variables aleatorias tales que $X_n \leqslant Z$ para toda $n \in \mathbb{N}$ donde $E(|Z|) < \infty$ entonces:*

a) Si $X_n \downarrow X$ c.s. entonces $E(X_n|\mathcal{G}) \downarrow E(X|\mathcal{G})$ c.s. y en L^1.

b) Si $X_n \downarrow X$ c.s. entonces $E(X_n|Y=y) \downarrow E(X|Y=y)$ c.s. (P_Y).

c) $\limsup_{n\to\infty} E(X_n|\mathcal{G}) \leqslant E(\limsup_{n\to\infty} X_n|\mathcal{G})$ c.s.

d) $\limsup_{n\to\infty} E(X_n|Y=y) \leqslant E(\limsup_{n\to\infty} X_n|Y=y)$ c.s. (P_Y).

Demostración:

a) $X_n \leqslant Z$ implica que $-X_n \geqslant -Z$ y $E(Z) < \infty$ hace que $E(-Z) > -\infty$.
Si $X_n \downarrow X$ entonces $-X_n \uparrow -X$.
Luego por a) del Teorema 1.4.6 $E(-X_n|\mathcal{G}) \uparrow E(-X|\mathcal{G})$ es decir, $-E(X_n|\mathcal{G}) \uparrow -E(X|\mathcal{G})$ o equivalentemente $E(X_n|\mathcal{G}) \downarrow E(X|\mathcal{G})$.

b) Igual que a).

c) $E(\limsup_{n\to\infty} X_n|\mathcal{G}) = E[\inf_n \sup_{k\geqslant n} X_k|\mathcal{G}] = -E[-\inf_n \sup_{k\geqslant n} X_k|\mathcal{G}] =$

$- E[\sup_n(-\sup_{k\geqslant n} X_k)|\mathcal{G}] = -E[\sup_n \inf_{k\geqslant n}(-X_k)|\mathcal{G}] = -E[\liminf_{n\to\infty}(-X_n)|\mathcal{G}].$

Pero por el Teorema 1.4.6 $E(\liminf_{n\to\infty} X_n|\mathcal{G})] \leqslant \liminf_{n\to\infty} E(X_n|\mathcal{G})$ entonces

$-E(\liminf_{n\to\infty}(-X_n)|\mathcal{G}) \geqslant -\liminf_{n\to\infty} E(-X_n|\mathcal{G}) = -\liminf_{n\to\infty}[-E(X_n|\mathcal{G})] =$

$\limsup_{n\to\infty} E(X_n|\mathcal{G})$.

De este modo se llega a que

$$E(\limsup_{n\to\infty} E(X_n|\mathcal{G})) \geqslant \limsup_{n\to\infty} E(X_n|\mathcal{G}).$$

d) ·Igual que c). □

Ahora se verán algunos resultados en donde las condiciones que se piden no son sobre X o la sucesión de variables aleatorias $\{X_n : n \in \mathbb{N}\}$, sino en el $\sigma-$álgebra $\mathcal{G}$ o en el objeto aleatorio Y.

Teorema 1.4.8. *a)* $E(X|\mathcal{G}) = E(X)$ *c.s. si* $\mathcal{G}$ *es el* $\sigma-$*álgebra trivial, es decir,* $\mathcal{G} = \{\varnothing, \Omega\}$.

b) Si Y *es una constante* b *c.s. entonces* $E(X|Y = y) = E(X)$ *c.s.* (P_Y).

c) $E(X|\mathcal{F}) = X$ *c.s.*

d) Si $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega, \mathcal{F})$ *es la identidad entonces* $E(X|Y = \omega) = Y(\omega)$ *c.s.* (P_Y).

Demostración:

a) Sea $\mathcal{G} = \{\varnothing, \Omega\}$ entonces si tomamos $C \in \mathcal{G}$ se debe cumplir que $\int\limits_C X dP = \int\limits_C E(X|\mathcal{G}) dP$.

$$\int\limits_C X dP = \begin{cases} 0 & \text{si } C = \varnothing \\ E(X) & \text{si } C = \Omega. \end{cases}$$

Por otro lado si suponemos que $E(X|\mathcal{G}) = E(X)$ que es una constante y por lo tanto $\mathcal{G}$—medible entonces

$$\int\limits_C E(X|\mathcal{G}) dP = \int\limits_C E(X) dP = E(X) \cdot \begin{cases} P(\varnothing) = 0 & \text{si } C = \varnothing \\ P(\Omega) = 1 & \text{si } C = \Omega \end{cases} =$$

$$\begin{cases} 0 & \text{si } C = \varnothing \\ E(X) & \text{si } C = \Omega \end{cases}$$

Por lo que $E(X|\mathcal{G}) = E(X)$ c.s.

b) Tomemos $B \in \mathcal{F}'$,

por definición se cumple que $\int\limits_{[Y \in B]} X dP = \int\limits_B E(X|Y = y) dP_Y(y)$,

entonces debemos comprobar que $\int\limits_{[Y \in B]} X dP = \int\limits_B E(X) dP_Y(y)$.

Sabemos que Y es constante, es decir, $Y(\omega) = b \in \Omega'$ para todo $\omega \in \Omega$ puesto que $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$.

Hay dos casos:

1. $b \in B$ entonces $\int\limits_{[Y\in B]} XdP = \int\limits_{\Omega} XdP = E(X) = E(X)\cdot 1 =$

$= E(X)P[Y \in B] = E(X)\int\limits_{B} dP_Y(y) = \int\limits_{B} E(X)dP_Y(y).$

2. $b \notin B$ entonces $\int\limits_{[Y\in B]} XdP = \int\limits_{\varnothing} XdP = 0 = E(X)\cdot 0 =$

$= E(X)P[Y \in B] = E(X)\int\limits_{B} dP_Y(y) = \int\limits_{B} E(X)dP_Y(y).$

De 1) y 2) se concluye que

$$E(X|Y = y) = E(X) \quad \text{c.s.} \quad (P_Y).$$

c) Sea $\mathcal{G} = \mathcal{F}$ se quiere probar que $E(X|\mathcal{F}) = X$ c.s.
Recordemos que $E(X|\mathcal{G})$ es $\mathcal{G}-$medible, en este caso, $E(X|\mathcal{F})$ debe ser $\mathcal{F}-$medible. Pero X es $\mathcal{F}-$medible por lo que sólo debemos comprobar que para todo $C \in \mathcal{F}$ se cumple la igualdad $\int\limits_{C} XdP = \int\limits_{C} E(X|\mathcal{F})dP$.
Si $E(X|\mathcal{F}) = X$ la igualdad se tiene, por tanto $E(X|\mathcal{F}) = X$.

d) Igual que c). □

Definimos ahora el concepto de átomo de un $\sigma-$álgebra relativo a una medida.

Si μ es una medida $\mathcal{G}$ es un $\sigma-$álgebra $B \in \mathcal{G}$ es un átomo de $\mathcal{G}$ relativo a μ, si $\mu(B) > 0$ y cuando $A \in \mathcal{G}$ con $A \subseteq B$ entonces $\mu(A) = 0$ ó $\mu(B - A) = 0$. En nuestro caso μ será una medida de probabilidad y la definición de átomo establece que no hay conjuntos medibles más pequeños que B con medida mayor que cero.

Lema 1.4.9. *Si $f : (\Omega, \mathcal{G}) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$, μ es una medida, B un átomo de $\mathcal{G}$ relativo a μ entonces f es constante sobre B c.s. (μ).*

Demostración:

Si $f(\omega) = \infty$ para todo $\omega \in B$ entonces la conclusión es cierta y no hay nada que probar,

y si $f(\omega) = -\infty$ para todo $\omega \in B$, lo mismo.

Supongamos por tanto que $f(\omega) \in \mathbb{R}$ para todo $\omega \in B$.
Sea $x \in \overline{\mathbb{R}}$, nos fijamos en $\mu(\{\omega \in B : f(\omega) < x\})$.
Aquí hay que notar lo siguiente
$\{\omega \in B : f(\omega) < x\} = B \cap f^{-1}[-\infty, x) \in \mathcal{G}$ porque $B \in \mathcal{G}$ y f es medible.
Entonces tiene sentido calcular

$$\mu(\{\omega \in B : f(\omega) < x\}) = \mu(B \cap f^{-1}[-\infty, x)).$$

Nos tomamos $x \in \overline{\mathbb{R}}$ tal que $\mu(B \cap f^{-1}[-\infty, x)) = 0$.
Ahora tomemos $y \leqslant x$, entonces $f^{-1}[-\infty, y) \subseteq f^{-1}[-\infty, x)$ y $B \cap f^{-1}[-\infty, y) \subseteq B \cap f^{-1}[-\infty, x)$.
Como μ es medida tenemos

$$0 \leqslant \mu(B \cap f^{-1}[-\infty, y)) \leqslant \mu(B \cap f^{-1}[-\infty, x)) = 0.$$

$$\text{Por lo que} \quad \mu(B \cap f^{-1}[-\infty, y)) = 0 \text{ para todo } y \leqslant x.$$

Ahora sea $k = \sup\{x \in \overline{\mathbb{R}} : \mu(B \cap f^{-1}[-\infty, x)) = 0\}$, el supremo existe porque $x \in \overline{\mathbb{R}}$.

Además el conjunto es no vacío porque si $\mu(B \cap f^{-1}[-\infty, x)) > 0$ para todo $x \in \overline{\mathbb{R}}$ (el conjunto es vacío) entonces B no sería átomo: $B \cap f^{-1}[-\infty, x) \subseteq B$ y $\mu(B \cap f^{-1}[-\infty, x)) > 0$ entonces

$$\mu(B \cap f^{-1}[-\infty, k)) = \mu\Big(\bigcup_{\substack{r \in \mathbb{Q} \\ r<k}} (B \cap f^{-1}[-\infty, r))\Big)$$

$$\text{porque } f^{-1}[-\infty, k) = \bigcup_{\substack{r \in \mathbb{Q} \\ r<k}} f^{-1}[-\infty, r), \text{ entonces}$$

$$\mu\Big(\bigcup_{\substack{r \in \mathbb{Q} \\ r<k}} (B \cap f^{-1}[-\infty, r))\Big) \leqslant \sum_{\substack{r \in \mathbb{Q} \\ r<k}} \mu\Big((B \cap f^{-1}[-\infty, r))\Big) = \sum_{\substack{r \in \mathbb{Q} \\ r<k}} 0 = 0.$$

Por definición de k se tiene que para todo $r < k$, $\mu(B \cap f^{-1}[\infty, r)) = 0$.
Ahora nos fijamos en $x > k$, entonces $x > \sup\{y \in \mathbb{R} : \mu(B \cap f^{-1}[-\infty, y)) = 0\}$ esto implica que $\mu(B \cap f^{-1}[-\infty, x)) \neq 0$, es decir, $\mu(B \cap f^{-1}[-\infty, x)) > 0$.
Pero $B = \left[B \cap f^{-1}[-\infty, x)\right] \cup \left[B \cap f^{-1}[x, \infty]\right]$.
Entonces $0 < \mu(B) = \mu(B \cap f^{-1}[-\infty, x)) + \mu(B \cap f^{-1}[x, \infty])$.
$\mu(B \cap f^{-1}[x, \infty]) = 0$ porque B es átomo.
Si $A = B \cap f^{-1}[x, \infty]$ con $x > k$ entonces $A^C = B^C \cup f^{-1}[-\infty, x)$.
Al ser B átomo sucede que o bien $\mu(B \cap f^{-1}[x, \infty]) = 0$ que *no* lo es, o bien
$\mu(B \cap A^C) = \mu\left(B \cap (B^C \cup f^{-1}[-\infty, x))\right)$
$= \mu(B \cap f^{-1}[-\infty, x)) = 0$ que tendrá que ser.
Ahora $\mu(B \cap f^{-1}[-\infty, k)) = 0$ entonces
$0 < \mu(B) = \mu(B \cap f^{-1}[-\infty, k)) + \mu(B \cap f^{-1}(\{k\})) + \mu(B \cap f^{-1}(k, \infty]) =$
$0 + \mu(B \cap f^{-1}(\{k\})) + 0 = \mu(B \cap f^{-1}(\{k\}))$.
Por tanto $B = B \cap f^{-1}(\{k\})$ c.s., esto es, casi todo punto de B está en $f^{-1}(\{k\})$ esto significa que $f(\omega) = k$ para casi todo $\omega \in B$.

Entonces $f = k$ c.s. (μ) sobre B. □

Si $B_1, B_2, \ldots$ son subconjuntos de Ω, el mínimo $\sigma-$álgebra que contiene a $B_1, B_2, \ldots$ es denotado por $\sigma(B_1, B_2, \ldots)$.

Teorema 1.4.10. *a) Sea B un átomo de $\mathcal{G}$ relativo a P, entonces*

$$E(X|\mathcal{G}) = \frac{1}{P(B)} \int_B X dP = \frac{E(XI_B)}{P(B)} \quad c.s. \quad en\ B,$$

y como un caso especial, sean $B_1, B_2, \ldots$ ajenos en $\mathcal{F}$ con $\Omega = \bigcup_{i=1}^{\infty} B_i$ y $P(B_i) > 0$ para todo $i = 1, 2, \ldots$ con $\mathcal{G} = \sigma(B_1, B_2, \ldots)$ entonces

$$E(X|\mathcal{G}) = \frac{1}{P(B_n)} \int_{B_n} X dP = \frac{E(XI_{B_n})}{P(B_n)} \quad c.s. \quad en\ B_n \quad n = 1, 2, \ldots$$

b) Si $B = [Y = y_0]$ y $P(B) > 0$ entonces

$$E(X|Y = y_0) = \frac{1}{P(B)} \int_B X dP.$$

Demostración:

a) Por el Lema 1.4.9, $E(X|\mathcal{G})(\omega) = k$ c.s. para $\omega \in B$ puesto que $E(X|\mathcal{G}) : (\Omega, \mathcal{G}) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ es medible y B es átomo.

Para cada $A \in \mathcal{G}$ se cumple que $\int_A X dP = \int_A E(X|\mathcal{G}) dP$ por definición de esperanza condicional dado un $\sigma-$álgebra.

Si tomamos $A = B \in \mathcal{G}$ se tiene

$$\int_B X dP = \int_B k dP = kP(B) \quad \text{por lo que} \quad k = \frac{1}{P(B)} \int_B X dP$$

y sustituyendo $E(X|\mathcal{G})(\omega) = \dfrac{\int_B XdP}{P(B)} = \dfrac{\int_\Omega XI_B dP}{P(B)} = \dfrac{E(XI_B)}{P(B)}.$

b) Mostraremos primero que $B = [Y = y_0]$ es un átomo de $\mathcal{G} = \mathcal{F}(Y)$.

Sea $A \in \mathcal{F}'$ con $Y^{-1}(A) \subseteq B = [Y = y_0] = Y^{-1}(\{y_0\})$.

Recordemos que $y_0 \in \Omega'$ puesto que $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$ es un objeto aleatorio.

Si $y_0 \in A$ entonces $Y^{-1}(\{y_0\}) \subseteq Y^{-1}(A) \subseteq Y^{-1}(\{y_0\})$ por tanto $Y^{-1}(\{y_0\}) = Y^{-1}(A) = B$.

Si $y_0 \notin A$ se tiene que $Y^{-1}(A) \cap Y^{-1}(\{y_0\}) = \varnothing$ entonces $Y^{-1}(A) = \varnothing$ puesto que $Y^{-1}(A) \subseteq Y^{-1}(\{y_0\})$.

Ahora sea $g(y) = E(X|Y = y) \quad y \in \Omega'$.

$h(\omega) = g(Y(\omega)) = E(X|\mathcal{G})(\omega)$

$\omega \in \Omega$ por a) de este Teorema, $h(\omega) = k$ c.s. si $\omega \in B$, entonces $g(Y(\omega)) = \dfrac{\int_B XdP}{P(B)}$ $\omega \in B$ y así $g(y_0) = \dfrac{\int_B XdP}{P(B)}$ c.s. (P_Y).

Pero

$$g(y_0) = E(X|Y = y_0) \quad \text{luego} \quad E(X|Y = y_0) = \frac{1}{P[Y = y_0]} \int_{[Y=y_0]} XdP. \ \square$$

Este último resultado es muy importante porque nos indica la manera de calcular esperanzas condicionales cuando $\mathcal{G}$ está generada por átomos. Veamos algunos ejemplos.

Ejemplo 1.4.11. Sean $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}[0,1]$, $P(C) = \int_C dx$ la medida de Lebesgue en $[0,1]$, $C \in \mathcal{B}[0,1]$, $\mathcal{G} = \sigma([0, \frac{1}{4}])$.

Sea $X : \Omega \longrightarrow \mathbb{R}$ tal que $X(\omega) = \text{sen}(2\pi\omega)$.

Calcular $E(X|\mathcal{G})$ para todo $\omega \in [0,1]$.

Notemos primero que $\mathcal{G} = \{\varnothing, [0,1], [0,\frac{1}{4}], (\frac{1}{4},1]\}$ que es una sub $\sigma-$álgebra de $\mathcal{F} = \mathcal{B}[0,1]$.

Los átomos de $\mathcal{G}$ son $B_1 = [0,\frac{1}{4}]$ y $B_2 = (\frac{1}{4},1]$.

Además $P(B_1) = \frac{1}{4}$ y $P(B_2) = \frac{3}{4}$.

Como B_1 y B_2 son ajenos y su unión es $\Omega = [0,1]$, al definir $E(X|\mathcal{G})$ en los átomos queda definida para todo $\omega \in \Omega$.

Si $\omega \in B_1$ entonces $E(X|\mathcal{G})(\omega) = \dfrac{1}{P(B_1)} \displaystyle\int_{B_1} X dP = 4 \int_0^{\frac{1}{4}} \operatorname{sen}(2\pi\omega) d\omega =$
$\dfrac{4}{2\pi}\left(\cos(0) - \cos\left(\dfrac{\pi}{2}\right)\right) = \dfrac{2}{\pi}$.

Si $\omega \in B_2$ entonces $E(X|\mathcal{G})(\omega) = \dfrac{1}{P(B_2)} \displaystyle\int_{B_2} X dP = \dfrac{4}{3} \int_{\frac{1}{4}}^{1} \operatorname{sen}(2\pi\omega) d\omega =$
$\dfrac{4}{6\pi}\left(\cos\left(\dfrac{\pi}{2}\right) - \cos(2\pi)\right) = -\dfrac{2}{3\pi}$.

Esto es:

$$E(X|\mathcal{G})(\omega) = \begin{cases} \dfrac{2}{\pi} & \text{si } 0 \leqslant \omega \leqslant \frac{1}{4} \\ \\ -\dfrac{2}{3\pi} & \text{si } \frac{1}{4} < \omega \leqslant 1. \end{cases}$$

Notemos que $X(\omega)$=sen$(2\pi\omega)$ no es $\mathcal{G}-$medible, es por esto que $E(X|\mathcal{G})(\omega) \neq X(\omega)$. ■

Ejemplo 1.4.12. Con los datos del ejemplo 1.4.11, sea $Y : (\Omega, \mathcal{F}, P) \longrightarrow (\Omega', \mathcal{F}')$ con $\Omega' = \mathbb{R}$, $\mathcal{F}' = \mathcal{B}(\mathbb{R})$, $Y(\omega) = I_{[0,\frac{1}{3}]}(\omega)$.

Calcular $E(X|Y)$ es decir, $E(X|\mathcal{F}(Y))$.

Primero recordemos que $\mathcal{F}(Y) = \{Y^{-1}(C) : C \in \mathcal{B}(\mathbb{R})\}$.

La variable aleatoria Y sólo toma dos valores: 0 y 1 entonces

$$Y^{-1}(C) = \begin{cases} \varnothing & \text{si } 0 \notin C, \quad 1 \notin C \\ [0, \frac{1}{3}] & \text{si } 0 \notin C, \quad 1 \in C \\ (\frac{1}{3}, 1] & \text{si } 0 \in C, \quad 1 \notin C \\ \Omega & \text{si } 0 \in C, \quad 1 \in C. \end{cases}$$

Por lo tanto $\mathcal{F}(Y) = \{\varnothing, [0,1], [0, \frac{1}{3}], (\frac{1}{3}, 1]\}$ que tiene dos átomos: $B_1 = [0, \frac{1}{3}]$ y $B_2 = (\frac{1}{3}, 1]$ con $P(B_1) = \frac{1}{3}$, $P(B_2) = \frac{2}{3}$.

Si $\omega \in B_1$ entonces $E(X|Y)(\omega) = \dfrac{1}{P(B_1)} \displaystyle\int_{B_1} X dP = 3 \int_0^{\frac{1}{3}} \operatorname{sen}(2\pi\omega) d\omega =$
$\dfrac{3}{2\pi}\left(1 - \cos\left(\dfrac{2\pi}{3}\right)\right) = \dfrac{3}{2\pi}\left(1 + \dfrac{1}{2}\right) = \dfrac{9}{4\pi}$.

Si $\omega \in B_2$ entonces $E(X|Y)(\omega) = \dfrac{1}{P(B_2)} \displaystyle\int_{B_2} X dP = \dfrac{3}{2} \int_{\frac{1}{3}}^{1} \operatorname{sen}(2\pi\omega) d\omega =$
$\dfrac{3}{2}\left(\dfrac{\cos\left(\frac{2\pi}{3}\right) - \cos(2\pi)}{2\pi}\right) = \dfrac{3}{2}\left(\dfrac{-\frac{1}{2} - 1}{2\pi}\right) = -\dfrac{9}{8\pi}$.

Entonces

$$E(X|Y)(\omega) = \begin{cases} \dfrac{9}{4\pi} & \text{si } 0 \leqslant \omega \leqslant \frac{1}{3} \\ \\ -\dfrac{9}{8\pi} & \text{si } \frac{1}{3} < \omega \leqslant 1. \end{cases} \qquad \blacksquare$$

Teorema 1.4.13. *a) Si $\mathcal{G}_1 \subseteq \mathcal{G}_2$ entonces $E\left[E(X|\mathcal{G}_2)|\mathcal{G}_1\right] = E(X|\mathcal{G}_1)$ c.s.*

a') Si $f : (\Omega', \mathcal{F}') \longrightarrow (\Omega'', \mathcal{F}'')$ entonces $E\left[E(X|Y)|f \circ Y\right] = E(X|f \circ Y)$ c.s.

b) Si $\mathcal{G}_1 \subseteq \mathcal{G}_2$ entonces $E\left[E(X|\mathcal{G}_1)|\mathcal{G}_2\right] = E(X|\mathcal{G}_2)$ c.s.

b') Si $f : (\Omega', \mathcal{F}') \longrightarrow (\Omega'', \mathcal{F}'')$ *entonces* $E\left[E(X|f\circ Y)|Y\right] = E(X|f\circ Y)$ *c.s.*

Demostración:

a) Sea $C \in \mathcal{G}_1$ entonces $C \in \mathcal{G}_2$ luego $\int\limits_C XdP = \int\limits_C E(X|\mathcal{G}_1)dP$ porque $C \in \mathcal{G}_1$ pero también $\int\limits_C XdP = \int\limits_C E(X|\mathcal{G}_2)dP$ pues $C \in \mathcal{G}_2$.
De este modo se cumple que para todo $C \in \mathcal{G}_1$
$\int\limits_C E\left[E(X|\mathcal{G}_2)|\mathcal{G}_1\right] dP = \int\limits_C E(X|\mathcal{G}_2)dP = \int\limits_C XdP = \int\limits_C E(X|\mathcal{G}_1)dP.$
Por lo tanto $E(X|\mathcal{G}_1) = E\left[E(X|\mathcal{G}_2)|\mathcal{G}_1\right]$ c.s.

a') $Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$ es medible y
$f : (\Omega', \mathcal{F}') \longrightarrow (\Omega'', \mathcal{F}'')$ es medible entonces
$f\circ Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega'', \mathcal{F}'')$ es medible porque es la composición de medibles.
Sean $\mathcal{G}_1 = (f\circ Y)^{-1}(\mathcal{F}'') = Y^{-1}(f^{-1}(\mathcal{F}''))$ donde $f^{-1}(\mathcal{F}'')$ es sub $\sigma-$álgebra de $\mathcal{F}'$, $\mathcal{G}_2 = Y^{-1}(\mathcal{F}')$ sub $\sigma-$álgebra de $\mathcal{F}$.
Entonces $\mathcal{G}_1 \subseteq \mathcal{G}_2$ porque $f^{-1}(\mathcal{F}'') \subseteq \mathcal{F}'$.
Así por a) se tiene que $E\left[E(X|\mathcal{G}_2)|\mathcal{G}_1\right] = E(X|\mathcal{G}_1)$ c.s.
Entonces $E\left[E(X|Y^{-1}(\mathcal{F}'))|(f\circ Y)^{-1}(\mathcal{F}'')\right] = E[X|(f\circ Y)^{-1}(\mathcal{F}'')]$
Por tanto $E\left[E(X|Y)|(f\circ Y)\right] = E[X|(f\circ Y)]$ c.s.

b) $E(X|\mathcal{G}_1)$ es $\mathcal{G}_1-$medible y como $\mathcal{G}_1 \subseteq \mathcal{G}_2$ entonces $E(X|\mathcal{G}_1)$ también es $\mathcal{G}_2$-medible, de este modo, para todo $C \in \mathcal{G}_2$

$$\int_C E(X|\mathcal{G}_1)dP = \int_C E(X|\mathcal{G}_1)dP \text{ implica que } E(X|\mathcal{G}_1) = E[E(X|\mathcal{G}_1)|\mathcal{G}_2] \text{ c.s.}$$

b′) Igual que b). □

Teorema 1.4.14. *a) Si Z es $\mathcal{G}-$ medible y X junto con XZ son integrables entonces $E[XZ|\mathcal{G}] = ZE(X|\mathcal{G})$ c.s. En particular $E(Z|\mathcal{G}) = Z$ c.s.*

a′) Si $f : (\Omega', \mathcal{F}') \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ y X junto con $X(f \circ Y)$ son integrables entonces $E[X(f \circ Y)|Y = y] = f(y)E[X|Y = y]$ c.s. (P_Y).

Demostración:

a) Primero supongamos que $Z = I_A$ con $A \in \mathcal{G}$ y tomemos $C \in \mathcal{G}$ entonces

$E(XZ|\mathcal{G}) = E(XI_A|\mathcal{G})$ cumple con $\int_C XZdP = \int_C E(XZ|\mathcal{G})dP$.

Pero $\int_C XZdP = \int_C XI_AdP = \int_{A\cap C} XdP = \int_{A\cap C} E(X|\mathcal{G})dP = \int_C I_AE(X|\mathcal{G})dP$.

Por lo tanto $\int_C E(XZ|\mathcal{G})dP = \int_C I_AE(X|\mathcal{G})dP$ para todo $C \in \mathcal{G}$.

De este modo $E(XI_A|\mathcal{G}) = I_AE(X|\mathcal{G})$ c.s.

Supongamos ahora que Z es una función simple $\mathcal{G}-$medible, es decir $Z=\sum_{j=1}^{n} \alpha_j I_{A_j}$ donde $A_1, \ldots, A_n$ es partición de Ω y $A_j \in \mathcal{G}$ para todo $j = 1, \ldots, n$.

Entonces $XZ = \sum_{j=1}^{n} \alpha_j I_{A_j} X$, y

$$E(XZ|\mathcal{G}) = E\left[\sum_{j=1}^{n} \alpha_j I_{A_j} X|\mathcal{G}\right] = \sum_{j=1}^{n} \alpha_j I_{A_j} E(X|\mathcal{G}).$$
$$= ZE(X|\mathcal{G}). \text{ Pues } A_j \in \mathcal{G} \text{ para todo } j.$$

Por último si Z es cualquier función $\mathcal{G}$ medible tomemos una sucesión $Z_1, Z_2, \ldots,$ de funciones simples $\mathcal{G}-$medibles tales que $|Z_n| \leqslant Z$ y $Z_n \uparrow Z$. Entonces $E(XZ_n|\mathcal{G}) = Z_n E(X|\mathcal{G})$ por ser Z_n simple, pero $XZ_n \xrightarrow{c.s.} XZ$, también $E(XZ_n|\mathcal{G}) \xrightarrow{c.s.} E(XZ|\mathcal{G})$.

Como X es integrable entonces $E(X|\mathcal{G})$ también lo es.

Por tanto $E(X|\mathcal{G})$ es finita c.s. y entonces $Z_n E(X|\mathcal{G}) \xrightarrow{c.s.} ZE(X|\mathcal{G})$.

Por lo tanto $E(XZ|\mathcal{G}) = ZE(X|\mathcal{G})$ c.s.

Notemos que es indispensable el hecho de que $E(X|\mathcal{G})$ sea finita puesto que $\frac{1}{n} \longrightarrow 0$ pero $\frac{1}{n} \cdot \infty \not\rightarrow 0 \cdot \infty = 0$.

Ahora si $X = 1$ c.s. entonces $E(XZ|\mathcal{G}) = E(Z \cdot 1|\mathcal{G}) = ZE(1|\mathcal{G}) = Z \cdot 1 = Z$.

a′) Sea $f = I_A$ con $A \in \mathcal{F}'$ entonces $\int\limits_B E[X(f \circ Y)|Y = y]dP_Y(y) = \int\limits_{[Y\in B]} X(f\circ Y)dP = \int\limits_{[Y\in B]} XI_{[Y\in A]}dP = \int\limits_{[Y\in A]\cap[Y\in B]} XdP = \int\limits_{A\cap B} E(X|Y = y)dP_Y(y) = \int\limits_B I_A E(X|Y = y)dP_Y(y) = \int\limits_B f(y)E(X|Y = y)dP_Y(y)$.

Por lo tanto $E(X(f \circ Y)|Y = y) = f(y)E(X|Y = y)$.

Los pasos siguientes son iguales a los de la demostración de a). $\square$

Corolario 1.4.15. *Sean $\mathcal{G}_1$ y $\mathcal{G}_2$ sub $\sigma-$álgebras de $\mathcal{F}$ tales que $\mathcal{G}_1 \subseteq \mathcal{G}_2$, X y Y variables aleatorias tal que $E(XY)$ y $E(Y)$ existen, X es $\mathcal{G}_2-$medible. Entonces $E(XY|\mathcal{G}_1) = E\left[XE(Y|\mathcal{G}_2)|\mathcal{G}_1\right]$ c.s.*

Demostración:

Sea $C \in \mathcal{G}_1$ entonces $C \in \mathcal{G}_2$, luego por definición

$$\int_C E\left[XE(Y|\mathcal{G}_2)|\mathcal{G}_1\right] dP = \int_C XE(Y|\mathcal{G}_2)dP.$$

Como X es $\mathcal{G}_2-$medible

$XE(Y|\mathcal{G}_2) = E(XY|\mathcal{G}_2)$.

Así $\int_C XE(Y|\mathcal{G}_2)dP = \int_C E(XY|\mathcal{G}_2)dP = \int_C XYdP$ porque $C \in \mathcal{G}_2$.

Pero $\int_C XYdP = \int_C E(XY|\mathcal{G}_1)dP$ porque $C \in \mathcal{G}_1$.

Ahora por unicidad casi seguramente obtenemos que $E\left[XE(Y|\mathcal{G}_2)|\mathcal{G}_1\right] = E(XY|\mathcal{G}_1)$ c.s. $\square$

Corolario 1.4.16. *Sean $\mathcal{G}_1 \subseteq \mathcal{G}_2$ sub $\sigma-$álgebras de $\mathcal{F}$.*
Si $E(Y)$ existe entonces $E(Y|\mathcal{G}_1) = E\left[E(Y|\mathcal{G}_1)|\mathcal{G}_2\right] = E\left[E(Y|\mathcal{G}_2)|\mathcal{G}_1\right]$.

Demostración:

En el corolario 1.4.15 notemos que $E(XY|\mathcal{G}_1)$ es $\mathcal{G}_1-$medible por definición.
Como $\mathcal{G}_1 \subseteq \mathcal{G}_2$ entonces también es $\mathcal{G}_2-$medible.
Luego $E\left[E(XY|\mathcal{G}_1)|\mathcal{G}_2\right] = E(XY|\mathcal{G}_1) = E\left[XE(Y|\mathcal{G}_2)|\mathcal{G}_1\right]$ la última igualdad es por el resultado anterior.
Si tomamos $X = 1$ c.s. obtenemos
$E\left[E(Y|\mathcal{G}_1)|\mathcal{G}_2\right] = E(Y|\mathcal{G}_1) = E\left[E(Y|\mathcal{G}_2)|\mathcal{G}_1\right]$. $\square$

Proposición 1.4.17. *Sean $\mathcal{G}$ sub $\sigma-$álgebra de $\mathcal{F}$, Y una variable aleatoria tal que $E(Y)$ existe. Supongamos que $\sigma(Y) = \mathcal{F}(Y)$ el $\sigma-$álgebra generado por Y y $\mathcal{G}$ son independientes, entonces $E(Y|\mathcal{G}) = E(Y)$.*

Demostración:

Recordemos primero que dos sub $\sigma-$álgebras de $\mathcal{F}$, $\mathcal{G}_1$ y $\mathcal{G}_2$ son independientes si para todos $A_1 \in \mathcal{G}_1$ y $A_2 \in \mathcal{G}_2$ se tiene que $P(A_1 \cap A_2) = P(A_1)P(A_2)$. El $\sigma-$álgebra trivial $\{\varnothing, \Omega\}$ es independiente de cualquier otro sub $\sigma-$álgebra de $\mathcal{F}$.

Ahora bien, tomemos $C \in \mathcal{G}$ entonces las variables aleatorias I_C y Y son independientes.

Si $A \in \sigma(Y)$ tenemos que $P(A \cap C) = P(A)P(C) = P(A)P[I_C = 1]$ pues los eventos C y $[I_C = 1]$ son el mismo.

$A \in \sigma(Y)$ si y sólo si existe $B \in \mathcal{B}(\mathbb{R})$ tal que $A = Y^{-1}(B) = [Y \in B]$ entonces $P[A \cap C] = P[[Y \in B] \cap [I_C = 1]] = P[Y \in B, I_C = 1] = P[Y \in B]P[I_C = 1]$.

También $P(A \cap C^c) = P[Y \in B]P[I_C = 0]$.

Entonces $\int_C E(Y|\mathcal{G})dP = \int_C YdP = \int_\Omega YI_CdP = E(YI_C) = E(Y)E(I_C) = P(C)E(Y) = E(Y)\int_C dP = \int_C E(Y)dP$.

Por lo tanto $E(Y|\mathcal{G}) = E(Y)$ c.s. □

Proposición 1.4.18. *Sean X una variable aleatoria integrable, $\mathcal{G}_1$ y $\mathcal{G}_2$ sub $\sigma-$álgebras de $\mathcal{F}$ tales que $\mathcal{F}(X)$ y $\mathcal{G}_1$ son independientes de $\mathcal{G}_2$.*
Entonces $E[X|\mathcal{G}_1 \curlyvee \mathcal{G}_2] = E(X|\mathcal{G}_1)$ c.s. donde $\mathcal{G}_1 \curlyvee \mathcal{G}_2 = \sigma(\mathcal{G}_1 \cup \mathcal{G}_2)$.

Demostración:

Sea

$$\mathcal{G} = \{A \in \mathcal{F} : A = \bigcup_{i=1}^{n} C_i \text{ con } C_i = D_{i1} \cap D_{i2}, D_{i1} \in \mathcal{G}_1, D_{i2} \in \mathcal{G}_2, C_i \cap C_j = \varnothing\}$$

Veremos que $\mathcal{G} = \mathcal{G}_1 \curlyvee \mathcal{G}_2$ de la siguiente manera:

a) $\mathcal{G}$ es $\sigma-$álgebra.

b) $\mathcal{G}_1 \subseteq \mathcal{G}$.

c) $\mathcal{G}_2 \subseteq \mathcal{G}$.

d) $\mathcal{G}_1 \cup \mathcal{G}_2 \subseteq \mathcal{G}$.

e) $\mathcal{G}_1 \curlyvee \mathcal{G}_2 \subseteq \mathcal{G}$.

f) $\mathcal{G} \subseteq \mathcal{G}_1 \curlyvee \mathcal{G}_2$.

a) i) $\varnothing \in \mathcal{G}$ y $\Omega \in \mathcal{G}$.

ii) Sea $C = D_1 \cap D_2$ con $D_1 \in \mathcal{G}_1$ y $D_2 \in \mathcal{G}_2$ entonces $C^c = (D_1 \cap D_2^c) \cup (D_1^c \cap D_2) \cup (D_1^c \cap D_2^c)$ en donde $D_1, D_1^c \in \mathcal{G}_1$ y $D_2, D_2^c \in \mathcal{G}_2$.
Así $C^c = \bigcup_{k=1}^{3} C_k$ con $C_k = D_{k1} \cap D_{k2}, D_{k1} \in \mathcal{G}_1$, $D_{k2} \in \mathcal{G}_2, C_i \cap C_j = \varnothing,\ i \neq j$.
Por lo tanto $C^c \in \mathcal{F}$.
Ahora si $C = \bigcup_{i=1}^{n} C_i = \bigcup_{i=1}^{n} (D_{i1} \cap D_{i2})$ entonces
$C^c = \bigcap_{i=1}^{n} (D_{i1} \cap D_{i2})^c = \bigcap_{i=1}^{n} [(D_{i1} \cap D_{i2}^c) \cup (D_{i1}^c \cap D_{i2}) \cup (D_{i1}^c \cap D_{i2}^c)] =$

$$\left[\left(\bigcap_{i=1}^{n} D_{i1}\right) \cap \left(\bigcap_{i=1}^{n} D_{i2}^c\right)\right] \cup \left[\left(\bigcap_{i=1}^{n} D_{i1}^c\right) \cap \left(\bigcap_{i=1}^{n} D_{i2}\right)\right] \cup \left[\left(\bigcap_{i=1}^{n} D_{i1}^c\right) \cap \left(\bigcap_{i=1}^{n} D_{i2}^c\right)\right]$$

iii) Sean $A_1, A_2, \ldots \in \mathcal{G}$ entonces $A_j = \bigcup_{i=1}^{n_j} C_{ij}$ con $C_{ij} = D_{ij1} \cap D_{ij2}$, $D_{ij1} \in \mathcal{G}_1, D_{ij2} \in \mathcal{G}_2$.

De este modo $\bigcup_{j=1}^{\infty} A_j = \bigcup_{j=1}^{\infty}\left(\bigcup_{i=1}^{n_j} C_{ij}\right) = \bigcup_{i=1}^{n_j}\left(\bigcup_{j=1}^{\infty} C_{ij}\right)$.

Por lo tanto $\bigcup_{j=1}^{\infty} A_j \in \mathcal{G}$ así $\mathcal{G}$ es $\sigma-$álgebra.

b) $\mathcal{G}_1 \subseteq \mathcal{G}$ porque si $A \in \mathcal{G}_1$ entonces $A = A \cap \Omega$ con $A \in \mathcal{G}_1, \Omega \in \mathcal{G}_2$.

Por lo tanto $A \in \mathcal{G}$ y luego $\mathcal{G}_1 \subseteq \mathcal{G}$.

c) $\mathcal{G}_2 \subseteq \mathcal{G}$ por la misma razón expuesta en b).

d) La unión de subconjuntos de $\mathcal{G}$ es subconjunto de $\mathcal{G}$.

Por lo tanto $\mathcal{G}_1 \cup \mathcal{G}_2 \subseteq \mathcal{G}$

e) Como $\mathcal{G}_1 \cup \mathcal{G}_2 \subseteq \mathcal{G}$ entonces $\sigma(\mathcal{G}_1 \cup \mathcal{G}_2) \subseteq \sigma(\mathcal{G})$.

Pero $\sigma(\mathcal{G}_1 \cup \mathcal{G}_2) = \mathcal{G}_1 \curlyvee \mathcal{G}_2$ y $\sigma(\mathcal{G}) = \mathcal{G}$.

Por lo tanto $\mathcal{G}_1 \curlyvee \mathcal{G}_2 \subseteq \mathcal{G}$

f) $\mathcal{G} \subseteq \mathcal{G}_1 \curlyvee \mathcal{G}_2$ se cumple porque cada generador C_i de $\mathcal{G}$ pertenece a $\mathcal{G}_1 \curlyvee \mathcal{G}_2$.

Por lo tanto $\mathcal{G} = \mathcal{G}_1 \curlyvee \mathcal{G}_2$.

La variable aleatoria $E(X|\mathcal{G}_1)$ es $\mathcal{G}_1-$medible y como $\mathcal{G}_1 \subseteq \mathcal{G}$ también es $\mathcal{G}$ medible, es suficiente mostrar que es una versión de $E(X|\mathcal{G})$ con $\mathcal{G} = \mathcal{G}_1 \curlyvee \mathcal{G}_2$.

Como vimos antes, necesitamos comprobar que para todo $D \in \mathcal{G}$ se cumple

$\int_D XdP = \int_D E(X|\mathcal{G}_1)dP.$

Nos bastará ver que para todos $D_1 \in \mathcal{G}_1$ y $D_2 \in \mathcal{G}_2$ se tiene

$\int_{D_1\cap D_2} XdP = \int_{D_1\cap D_2} E(X|\mathcal{G}_1)dP$ porque cualquier $D \in \mathcal{G}$ se genera con elementos que tienen la forma $D_1 \cap D_2$.

Así $\int_{D_1\cap D_2} XdP = \int_\Omega XI_{D_1\cap D_2}dP = \int_\Omega XI_{D_1}I_{D_2}dP = E(XI_{D_1}I_{D_2}) = E(XI_{D_1})E(I_{D_2})$ debido a que $\mathcal{F}(X)$ y $\mathcal{G}_1$ son independientes de $\mathcal{G}_2$.

Entonces $\int_{D_1\cap D_2} XdP = E(XI_{D_1})E(I_{D_2}) = E(I_{D_2}) \int_\Omega E(XI_{D_1}|\mathcal{G}_1)dP_{\mathcal{G}_1} =$

$E(I_{D_2}) \int_\Omega E(XI_{D_1}|\mathcal{G}_1)dP =$

$E(I_{D_2}) \int_\Omega I_{D_1}E(X|\mathcal{G}_1)dP.$

Ya que $E(XI_{D_1}) = E[E(XI_{D_1}|\mathcal{G}_1)]$ y porque I_{D_1} es $\mathcal{G}_1-$medible.

Además $I_{D_1}E(X|\mathcal{G}_1)$ es $\mathcal{G}_1-$medible e I_{D_2} es $\mathcal{G}_2-$medible.

Por independencia:

$\int_{D_1\cap D_2} XdP = E(I_{D_2}) \int_\Omega I_{D_1}E(X|\mathcal{G}_1)dP = \int_\Omega I_{D_2}dP \int_\Omega I_{D_1}E(X|\mathcal{G}_1)dP =$

$\int_\Omega I_{D_2}I_{D_1}E(X|\mathcal{G}_1)dP = \int_{D_1\cap D_2} E(X|\mathcal{G}_1)dP.$ $\square$

Una función $g : \mathbb{R} \longrightarrow \mathbb{R}$ continua es convexa si $g(\alpha x + (1-\alpha)y) \leqslant \alpha g(x) + (1-\alpha)g(y)$ para $0 < \alpha < 1$, $x, y \in \mathbb{R}$.

Si $z = \alpha x + (1-\alpha)y$ es claro que z está entre x y y.

Suponiendo que $x < y$ entonces $x < z < y$ luego $z = \alpha x+(1-\alpha)y$ implica que $z = \alpha(x-y)+y$, asi $\alpha = \frac{z-y}{x-y}$, $1-\alpha = 1 - \frac{z-y}{x-y} = \frac{x-z}{x-y} = \frac{-(z-x)}{x-y}$. La defínción de convexidad es:

$$g(z) \leqslant \frac{z-y}{x-y}g(x) - \frac{(z-x)}{x-y}g(y) \text{ es decir}$$
$$g(z) - g(y) \leqslant \frac{z-y}{x-y}g(x) - \frac{(z-x)}{x-y}g(y) - g(y) \text{ esto es}$$
$$g(z) - g(y) \leqslant \frac{z-y}{x-y}g(x) - \left[\frac{z-x+x-y}{x-y}\right] g(y) \text{ entonces}$$
$$g(z) - g(y) \leqslant \frac{z-y}{x-y}g(x) - \frac{z-y}{x-y}g(y) \text{ por último}$$
$$\frac{g(z)-g(y)}{z-y} \geqslant \frac{g(x)-g(y)}{x-y} \text{ porque } z-y<0.$$

Supongamos y fija, sea $h(x) = \dfrac{g(x)-g(y)}{x-y}$ entonces $h(x) \leqslant h(z)$ (recordar que $x < z < y$).

h es creciente y continua porque g es continua.

Si definimos $k(y) = \lim_{x \uparrow y} h(x) = \lim_{x \uparrow y} \dfrac{g(x)-g(y)}{x-y}$, tenemos que $k(y)$ existe porque es el valor de la pendiente de la recta tangente a g en y por izquierda.

Además $g(x) - g(y) \geqslant k(y)(x-y)$ porque h es creciente.

Entonces $k(y) \geqslant h(x)$ implica que $k(y) \geqslant \dfrac{g(x)-g(y)}{x-y}$.

Luego $k(y)(x-y) \leqslant g(x) - g(y)$ puesto que $x < y$.

Esta desigualdad también se cumple cuando $x > y$ haciendo una argumentación similar.

Así, si g es convexa entonces para todo $x, y \in \mathbb{R}$ con $x \neq y$ se vale $g(x) - g(y) \geqslant k(y)(x-y)$.

Este resultado de funciones reales convexas nos servirá para establecer la siguiente proposición.

Proposición 1.4.19 (Desigualdad de Jensen). *Sean Y una variable aleatoria definida en $(\Omega, \mathcal{F}, P)$ tal que $E(Y)$ existe, g continua y convexa en $\mathbb{R}$ en donde $E[g(Y)]$ existe. Consideremos $\mathcal{G}$ un sub $\sigma-$álgebra de $\mathcal{F}$.*

Entonces $E[g(Y)|\mathcal{G}] \geqslant g(E(Y|\mathcal{G}))$ *c.s.*

Demostración:

Sabemos que $g(x) - g(y) \geqslant k(y)(x-y)$ entonces
$g(Y) - g(E(Y|\mathcal{G})) \geqslant k(E(Y|\mathcal{G})(Y - E(Y|\mathcal{G})))$
Como g es continua y k es creciente tenemos que $g(E(Y|\mathcal{G}))$ y $k(E(Y|\mathcal{G}))$ son $\mathcal{G}$−medibles.
Pero $E[g(Y)|\mathcal{G}] - g(E(Y|\mathcal{G})) = E\left[g(Y) - g(E(Y|\mathcal{G}))\Big|\mathcal{G}\right] \geqslant$
$E\left[k(E(Y|\mathcal{G}))(Y - E(Y|\mathcal{G}))|\mathcal{G}\right]$
$= k(E(Y|\mathcal{G}))E[Y - E(Y|\mathcal{G})\Big|\mathcal{G}]$
$= k(E(Y|\mathcal{G}))\Big\{E(Y|\mathcal{G}) - E[E(Y|\mathcal{G})|\mathcal{G}]\Big\}$
$= k(E(Y|\mathcal{G}))\Big\{E(Y|\mathcal{G}) - E[Y|\mathcal{G}]\Big\} = 0.$
Por lo tanto $E[g(Y)|\mathcal{G}] - g(E(Y|\mathcal{G})) \geqslant 0.$
Asi $E[g(Y)|\mathcal{G}] \geqslant g(E(Y|\mathcal{G})).$ $\square$

Corolario 1.4.20. *Si g es convexa, Y variable aleatoria tal que $E(Y)$ y $E[g(Y)]$ existen entonces $E[g(Y)] \geqslant g(E(Y))$.*

Demostración:

Si $\mathcal{G} = \{\varnothing, \Omega\}$ entonces $\sigma(Y)$ y $\mathcal{G}$ son independientes.
De este modo $E(Y|\mathcal{G}) = E(Y)$ c.s. y $E[g(Y)|\mathcal{G}] = E(g(Y))$ y por la proposición 1.4.19 (Jensen) $E[g(Y)|\mathcal{G}] \geqslant g(E(Y|\mathcal{G}))$ se cumple que $E[g(Y)] \geqslant g(E(Y))$. $\square$

Corolario 1.4.21. *Sean $p \geqslant 1$, Y una variable aleatoria tal que $E(|Y|^p) < \infty$, entonces $|E(Y|\mathcal{G})|^p \leqslant E[|Y|^p|\mathcal{G}]$ c.s.*

Demostración:

Sea $g(x) = |x|^p$ entonces por la proposición 1.4.19 (Jensen) tenemos $E[|Y|^p|\mathcal{G}] \geqslant |E(Y|\mathcal{G})|^p$. $\square$

Ejemplo 1.4.22. Sean $X_1, X_2, \ldots, X_n$ v.a.i.i.d. continuas con densidad f, $S_n = \sum_{i=1}^{n} X_i$.

Si $B \in \mathcal{B}(\mathbb{R})$ tenemos que

$\int_{[S_n \in B]} X_j dP = E\left[X_j I_{[S_n \in B]}\right] = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_j I_B(x_1 + \cdots + x_n) \prod_{i=1}^{n} dF(x_i)$ donde F es la distribución de cada una de éstas v.a.

Entonces
$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_j I_B(x_1 + \cdots + x_n) \prod_{i=1}^{n} dF(x_i) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_k I_B(x_1 + \cdots + x_n) \prod_{i=1}^{n} dF(x_i) = \int_{[S_n \in B]} X_k dP$ donde la penúltima igualdad se debe al teorema de cambio de variables haciendo la transformación de $\mathbb{R}^n$ en $\mathbb{R}^n$, T, definida por $T(x_1, \ldots, x_j, \ldots, x_k, \ldots, x_n) = (x_1, \ldots, x_k, \ldots, x_j, \ldots, x_n)$ esto es, una permutación de las variables que sólo intercambia a X_k con X_j y que por tanto tiene Jacobiano igual a 1 en valor absoluto.

Como $\int_{[S_n \in B]} X_j dP = \int_{[S_n \in B]} X_k dP$ se concluye que $E(X_j|S_n) = E(X_k|S_n)$ para todas $j, k \in \{1, \ldots, n\}$.

Entonces $S_n = E(S_n|S_n) = E(X_1 + \cdots + X_n|S_n) = \sum_{i=1}^{n} E(X_i|S_n) = E(X_1|S_n) + \cdots + E(X_n|S_n) = E(X_j|S_n) + E(X_j|S_n) + \cdots + E(X_j|S_n) = nE(X_j|S_n)$.

Entonces $E(X_j|S_n) = \dfrac{S_n}{n}$ c.s. $\blacksquare$

Un texto que desarrolla las propiedades de esperanza condicional de manera completa es Ibarrola et. al (1997).

Capítulo 2

Martingalas

La mayor parte de los resultados de este capítulo fueron extraídos de Laha-Rotahgi (1979), se completaron los detalles omitidos en las pruebas de este libro.

2.1. Definiciones Básicas

Trataremos con colecciones de variables aleatorias numerables o no, en donde hay una estructura especial de dependencia, veremos algunos resultados límite que comúnmente se estudian con variables aleatorias independientes, ahora con la propiedad de dependencia que estudiaremos en este capítulo.
Un conjunto $(T, \prec)$ es un conjunto con orden parcial si:

a) $t \prec t$ para todo $t \in T$ (reflexividad).

b) Si $t \prec s$ y $s \prec t$ entonces $t = s$ (antisimetría).

c) Si $t \prec s$ y $s \prec r$ entonces $t \prec r$ (transitividad).

Algunos ejemplos de conjuntos con orden parcial son:
$(\mathbb{R}, \leqslant)$, $(\mathbb{N}, \leqslant)$, $(2^A, \subseteq)$ con A cualquier conjunto.
Consideremos $(\Omega, \mathcal{F}, P)$ espacio de probabilidad y sea $\{\mathcal{F}_t : t \in T\}$ una colección de $\sigma-$álgebras tal que $\mathcal{F}_t \subseteq \mathcal{F}$ y para todo $s < t$ tenemos que $\mathcal{F}_s \subseteq \mathcal{F}_t$.
Cuando $T = \mathbb{N}$ y $\mathcal{F}_n = \sigma(X_1, \ldots, X_n) = \sigma\left(\bigcup_{k=1}^{n} \sigma(X_k)\right)$ si $n < m$ es claro que $\mathcal{F}_n \subseteq \mathcal{F}_m$. Sea $\underline{X} = \{X_t : t \in T\}$ una familia de variables aleatorias definidas en $(\Omega, \mathcal{F}, P)$ con $E(|X_t|)$ finita, $\underline{X}$ es una **martingala** con respecto a $\{\mathcal{F}_\mathbf{t} : \mathbf{t} \in \mathbf{T}\}$ si:

***i*) $\mathbf{X_t}$ es $\mathcal{F}_\mathbf{t}-$medible.**

***ii*) $\mathbf{E(X_t|\mathcal{F}_s) = X_s}$ c.s. si $\mathbf{s < t}$.**

$\underline{X}$ es una **submartingala** si cambiamos $=$ por $\geqslant$ en *ii*).
$\underline{X}$ es una **supermartingala** si cambiamos $=$ por $\leqslant$ en *ii*).
No necesariamente se cumple una de las tres posibilidades puesto que dadas dos funciones medibles h_1 y h_2 no se cumple que $h_1 = h_2$, $h_1 \geqslant h_2$ o $h_1 \leqslant h_2$ c.s.
Cuando $T = \mathbb{N}$, $\prec$ es $\leqslant$ y $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ diremos simplemente que $\{X_n : n \in \mathbb{N}\}$ es martingala.
La teoría de martingalas tiene su origen en los juegos de azar, tiene múltiples aplicaciones en Matemáticas tales como:

Teoría de la Medida, Análisis Funcional, Problema Dirichlet en Ecuaciones Diferenciales Parciales. Estas aplicaciones aparecen en Bojdecki (1985) con todo detalle.
Además la teoría de martingalas es una herramienta básica en teoría de probabilidad, está intimamente relacionada con esperanza condicional, por lo que consideramos importante incluirla en este trabajo.

Proposición 2.1.1. *$\{X_n, \mathcal{F}_n\}$ es martingala (submartingala o supermartingala) si y sólo si para todo $n > 1$ $E(X_n|\mathcal{F}_{n-1}) = X_{n-1}$ ($\geqslant$ si es submartingala, $\leqslant$ si es supermartingala).*

Demostración:
Por definición de martingala sabemos que para toda $m < n$ tenemos que $E(X_n|\mathcal{F}_m) = X_m$ c.s.
En particular si $m = n-1$ obtenemos $E(X_n|\mathcal{F}_{n-1}) = X_{n-1}$ c.s.
Para la otra implicación, sea $m < n$ entonces $m \leqslant n-1$, por tanto $\mathcal{F}_m \subseteq \mathcal{F}_{n-1}$.
Sabemos que $E(X_n|\mathcal{F}_{n-1}) = X_{n-1}$ c.s.
Ahora $E(X_n|\mathcal{F}_m) = E\left[E(X_n|\mathcal{F}_{n-1})|\mathcal{F}_m\right] = E(X_{n-1}|\mathcal{F}_m) = \cdots = E(X_{m+1}|\mathcal{F}_m) = X_m$ c.s. □

Ejemplo 2.1.2. Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de variables aleatorias i.i.d. con $P[X_n = 1] = \frac{1}{2} = P[X_n = -1]$ para todo $n \in \mathbb{N}$.
Se gana si sale águila y se pierde si sale sol.
Se apuesta la cantidad b_{i-1} en el $i-$ésimo juego, la ganancia esperada en

cualquier juego es $b_{i-1}P[X_i = 1] + (-b_{i-1})P[X_i = -1] = 0$.

Si la estrategia es apostar $b_0 > 0$ en el primer lanzamiento y $b_n = b_n(X_1, \ldots, X_n)$ en el $(n+1)-$ésimo lanzamiento, entonces b_n es $\sigma(X_1, \ldots, X_n)$ medible.

Si S_0 es el capital inicial, $S_0 > b_0$ y S_n es el capital después del $n-$ésimo juego, entonces $S_n = S_n(b_0, b_1, \ldots, b_{n-1}, X_1, X_2, \ldots, X_n)$ por lo que S_n es $\mathcal{F}_n-$medible.

Además $S_{n+1} = S_0 + \sum\limits_{k=0}^{n} b_k X_{k+1} = S_n + b_n X_{n+1}$.

Como $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ y $E(X_n) = 0$ para todo $n \geqslant 1$ entonces $E(S_{n+1}|\mathcal{F}_n) = E[S_n + b_n X_{n+1}|\mathcal{F}_n] = E(S_n|\mathcal{F}_n) + E(b_n X_{n+1}|\mathcal{F}_n) = S_n + b_n E(X_{n+1}|\mathcal{F}_n) = S_n + b_n E(X_{n+1}) = S_n + b_n \cdot 0 = S_n$ pues $E(X_{n+1}|\mathcal{F}_n) = E(X_{n+1})$.

Por lo tanto $E(S_{n+1}|\mathcal{F}_n) = S_n$ asi $\{S_n, \mathcal{F}_n\}$ es martingala.

Ahora consideremos una nueva estrategia:

Se apuesta $b_0 = b > 0$ inicialmente, doblamos la apuesta hasta ganar y en ese momento nos retiramos.

La probabilidad de ganar al menos un juego es 1, pues definamos Y como el número del juego en que se gana por primera vez, entonces $Y \sim G(\frac{1}{2})$, esto es, $P[Y = k] = \frac{1}{2}(\frac{1}{2})^{k-1}$, $k = 1, 2, \ldots$ por tanto $\sum\limits_{k=1}^{\infty} P[Y = k] = \sum\limits_{k=1}^{\infty} (\frac{1}{2})^k = 1$ así que es seguro ganar alguna vez.

Pero $b_{n-1} = \begin{cases} 2^{n-1}b & \text{si } X_i = -1 \text{ para } i = 1, 2, \ldots, n-1, \\ 0 & \text{si no.} \end{cases}$

es decir, la apuesta en el $n-$ésimo juego es $b \cdot 2^{n-1}$ si se han perdido los $(n-1)$ juegos anteriores.

Si el jugador gana por primera vez en el juego $(n+1)$, para ese entonces ya ha perdido $b(1 + 2 + 2^2 + \cdots + 2^{n-1}) = b\sum\limits_{k=0}^{n-1} 2^k = b(2^n - 1)$ y por tanto apuesta

$b \cdot 2^n$ por lo que la ganancia es $b \cdot 2^n - b(2^n - 1) = b$ pero se necesita una fortuna muy grande para seguir esta estrategia (fortuna ilimitada) aunque en realidad con una fortuna menor conviene esta estrategia.

Finalmente supongamos que el jugador puede saltarse juegos, $\delta_n = \delta_n(X_1, \ldots, X_{n-1})$ una función $\mathcal{F}_{n-1}$ medible tal que

$$\delta_n = \begin{cases} 0 & \text{si se salta el } n\text{−ésimo juego,} \\ 1 & \text{si no.} \end{cases}$$

Ahora S_n^* es el capital después de n juegos, entonces $S_{n+1}^* = S_n^* + \delta_{n+1}(X_1, \ldots, X_n)b_n(X_1, \ldots, X_n)X_{n+1}$, $|S_{n+1}^*| \leqslant |S_n^*| + |b_n|$ por lo que $|S_n^*| \leqslant |S_0^*| + |b_0| + |b_1| + \cdots + |b_{n-1}|$.

De donde $E(|S_n^*|) < \infty$, luego

$E(S_{n+1}^*|\mathcal{F}_n) = E(S_n^* + \delta_{n+1}(X_1, \ldots, X_n)b_n(X_1, \ldots, X_n)X_{n+1}|\mathcal{F}_n) = S_n^* + \delta_{n+1}(X_1, \ldots, X_n)b_n(X_1, \ldots, X_n)E(X_{n+1}|\mathcal{F}_n)$ porque δ_{n+1} y b_n son $\mathcal{F}_n$ medibles.

Así $E(S_{n+1}^*|\mathcal{F}_n) = S_n^* + \delta_{n+1}(X_1, \ldots, X_n)b_n(X_1, \ldots, X_n)E(X_{n+1}) = S_n^*$ porque $E(X_{n+1}|\mathcal{F}_n) = E(X_{n+1}) = 0$.

Por lo tanto $(S_n^*, \mathcal{F}_n)$ es martingala. ■

Ejemplo 2.1.3. Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de v.a.i. con $E(X_n) = 0$ para todo $n \in \mathbb{N}$, $S_n = \sum_{k=1}^{n} X_k$.

Entonces $E(|S_n|) \leqslant \sum_{k=1}^{n} E(|X_k|) < \infty$ porque $E(|X_k|) < \infty$ para todo $k \in \mathbb{N}$.

Así $E(S_{n+1}|S_1, \ldots, S_n) = E(S_n + X_{n+1}|X_1, \ldots, X_n)$, ya que $S_{n+1} = S_n + X_{n+1}$ y $\sigma(S_1, \ldots, S_n) = \sigma(X_1, \ldots, X_n)$.

Luego $E(S_{n+1}|S_1,\ldots,S_n) = S_n + E(X_{n+1}|X_1,\ldots,X_n) = S_n + E(X_{n+1}) = S_n + 0 = S_n$ porque S_n es $\sigma(X_1,\ldots,X_n)$ medible y por independencia.

De este modo $\{S_n,\mathcal{F}_n\}$ es martingala, aquí $\mathcal{F}_n$ puede ser tanto $\sigma(S_1,\ldots,S_n)$ como $\sigma(X_1,\ldots,X_n)$. ■

Ejemplo 2.1.4. Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de variables aleatorias independientes con $E(X_n) = 1$ para todo $n \in \mathbb{N}$, definimos Z_n como $Z_n = \prod_{i=1}^{n} X_i$ para cada $n \in \mathbb{N}$.

Entonces $E(|Z_n|) = E(|\prod_{i=1}^{n} X_i|) = E(\prod_{i=1}^{n} |X_i|) = \prod_{i=1}^{n} E(|X_i|) < \infty$ porque $E(|X_i|) < \infty$ y por independencia.

$E(Z_{n+1}|X_1,\ldots,X_n) = E(X_{n+1}Z_n|X_1,\ldots,X_n) = Z_nE(X_{n+1}|X_1,\ldots,X_n)$ porque $Z_n = \prod_{i=1}^{n} X_i$ es $\sigma(X_1,\ldots,X_n)$ medible

$E(Z_{n+1}|X_1,\ldots,X_n) = Z_nE(X_{n+1}) = Z_n$ porque $E(X_{n+1}) = 1$ y por independencia.

Por tanto $\{Z_n,\mathcal{F}_n\}$ es martingala donde $\mathcal{F}_n = \sigma(X_1,\ldots,X_n)$. ■

Ejemplo 2.1.5. Sean $\{X_n : n \in \mathbb{N}\}$ una sucesión de variables aleatorias i.i.d. con $P[X_k = 1] = p$, $P[X_k = -1] = q = 1 - p$ para todo $k \in \mathbb{N}$, $0 < p < 1$ y $S_n = \sum_{k=1}^{n} X_k$.

Definimos $Z_n = \left(\frac{q}{p}\right)^{S_n}$ entonces $E(|Z_n|) = E\left(\left|\frac{q}{p}\right|^{S_n}\right) = E\left[\left|\frac{q}{p}\right|^{\sum_{k=1}^{n} X_k}\right] = E\left[\prod_{k=1}^{n}\left|\frac{q}{p}\right|^{X_k}\right] = \prod_{k=1}^{n} E\left(\left|\frac{q}{p}\right|^{X_k}\right) = \left(E\left(\left|\frac{q}{p}\right|^{X_1}\right)\right)^n < \infty$ porque $E\left[\left|\frac{q}{p}\right|^{X_1}\right] < \infty$ y por independencia.

Ahora $E(Z_{n+1}|\mathcal{F}_n) = E(Z_{n+1}|X_1,\ldots,X_n) = E\left(\left(\frac{q}{p}\right)^{S_{n+1}}\middle|X_1,\ldots,X_n\right) = E\left(\left(\frac{q}{p}\right)^{S_n}\left(\frac{q}{p}\right)^{X_{n+1}}\middle|X_1,\ldots,X_n\right) = \left(\frac{q}{p}\right)^{S_n} E\left(\left(\frac{q}{p}\right)^{X_{n+1}}\middle|X_1,\ldots,X_n\right)$ porque

$\left(\frac{q}{p}\right)^{S_n}$ es $\sigma(X_1,\ldots,X_n)$ medible.

$$E(Z_{n+1}|\mathcal{F}_n) = \left(\tfrac{q}{p}\right)^{S_n} E\left(\left(\tfrac{q}{p}\right)^{X_{n+1}}\right) \text{ por independencia}$$
$$= \left(\tfrac{q}{p}\right)^{S_n}\left[\left(\tfrac{q}{p}\right)^{1}\cdot p + \left(\tfrac{q}{p}\right)^{-1}\cdot q\right] = \left(\tfrac{q}{p}\right)^{S_n}[q+p] = \left(\tfrac{q}{p}\right)^{S_n} = Z_n.$$

Por lo tanto $\{Z_n, \mathcal{F}_n\}$ es martingala. ■

Ejemplo 2.1.6. Una urna tiene b bolas negras y w blancas.

Se extrae una bola, se anota su color y se devuelve junto con $c \geqslant 1$ bolas del mismo color, el esquema de la urna de Polya.

Sean $X_0 = \dfrac{b}{b+w}$ la proporción inicial de bolas negras, y X_n = Proporción de bolas negras después de la $n-$ésima extracción.

Veremos que $\{X_n : n \in \mathbb{N}\}$ es martingala.

Sean $Y_0 = 1,\ Y_n = \begin{cases} 1 & \text{si la } n\text{−ésima bola extraída es negra,} \\ 0 & \text{si no,} \end{cases}$

b_n y w_n son el número de bolas negras y blancas después de la $n-$ésima extracción respectivamente.

Asi $b_0 = b$, $w_0 = w$, $X_n = \dfrac{b_n}{b_n + w_n}$.

Además $b_{n+1} = b_n + cY_{n+1}$, $w_{n+1} = w_n + c(1 - Y_{n+1})$ entonces

$P(Y_{n+1} = 1|Y_0, Y_1, \ldots, Y_n) = X_n$ =Proporción de bolas negras después de la $n-$ésima extracción, para todo $n \in \mathbb{N}$ c.s.

Por otro lado

$E(X_{n+1}|Y_0,Y_1,\ldots,Y_n) = E\left[\frac{b_{n+1}}{b_{n+1}+w_{n+1}}\Big|Y_0,Y_1,\ldots,Y_n\right]$

$= E\left[\frac{b_n+cY_{n+1}}{b_{n+1}+w_{n+1}}\Big|Y_0,Y_1,\ldots,Y_n\right] = E\left[\frac{b_n+cY_{n+1}}{b_n+w_n+c}\Big|Y_0,Y_1,\ldots,Y_n\right]$

$= \frac{b_n}{b_n+w_n+c} + \frac{c}{b_n+w_n+c}E(Y_{n+1}|Y_0,Y_1,\ldots,Y_n).$

porque $\frac{b_n}{b_n+w_n+c}$ es $\sigma(Y_0,Y_1,\ldots,Y_n)$ medible.

$E(X_{n+1}|Y_0,Y_1,\ldots,Y_n) = \frac{b_n}{b_n+w_n+c} + \frac{c}{b_n+w_n+c}P(Y_{n+1}=1|Y_0,Y_1,\ldots,Y_n)$

$= \frac{b_n}{b_n+w_n+c} + \frac{c}{b_n+w_n+c}X_n = \frac{b_n}{b_n+w_n+c} + \frac{c}{b_n+w_n+c}\cdot\frac{b_n}{b_n+w_n}$

$= \frac{b_n(b_n+w_n)+cb_n}{(b_n+w_n+c)(b_n+w_n)} = \frac{b_n(b_n+w_n+c)}{(b_n+w_n+c)(b_n+w_n)} = \frac{b_n}{b_n+w_n} = X_n$ c.s.

Por lo tanto $\{X_n, \sigma(Y_0,Y_1,\ldots,Y_n)\}$ es martingala. ■

Ejemplo 2.1.7. Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de v.a. con densidad conjunta p_n ó bien q_n.

El radio de verosimilitud se define como $\lambda_n(x_1,\ldots,x_n) = \frac{q_n(x_1,\ldots,x_n)}{p_n(x_1,\ldots,x_n)}$.

Cuando λ_n es "pequeño"se tiene evidencia de que p_n es cierta, y cuando es "grande"se tiene evidencia de que q_n es cierta.

Supongamos que p_n es continua para todo n y que $p_n = 0$ si y sólo si $q_n = 0$.

También q_n es continua, y λ_n es $\sigma(X_1,\ldots,X_n)$ medible.

Veremos que $\{\lambda_n : n \in \mathbb{N}\}$ es martingala.

Si p_n es la verdadera densidad del vector $(X_1,\ldots,X_n)$ entonces la densidad de X_{n+1} dados $X_1,\ldots,X_n$ es $\frac{p_{n+1}}{p_n}$ por lo que

$$E(\lambda_{n+1}|X_1 = x_1, \ldots, X_n = x_n) = \int_{-\infty}^{\infty} \lambda_{n+1}(x_1, \ldots, x_n, y) \frac{p_{n+1}(x_1, \ldots, x_n, y)}{p_n(x_1, \ldots, x_n)} dy$$

$$= \int_{-\infty}^{\infty} \frac{q_{n+1}(x_1, \ldots, x_n, y)}{p_{n+1}(x_1, \ldots, x_n, y)} \frac{p_{n+1}(x_1, \ldots, x_n, y)}{p_n(x_1, \ldots, x_n)} dy = \int_{-\infty}^{\infty} \frac{q_{n+1}(x_1, \ldots, x_n, y)}{p_n(x_1, \ldots, x_n)} dy$$

$$= \frac{q_n(x_1, \ldots, x_n)}{p_n(x_1, \ldots, x_n)} = \lambda_n \text{ porque } \int_{-\infty}^{\infty} q_{n+1}(x_1, \ldots, x_n, y) dy = q_n(x_1, \ldots, x_n).$$

Por lo tanto $E(\lambda_{n+1}|X_1, \ldots, X_n) = \lambda_n$ y como $\sigma(\lambda_1, \ldots, \lambda_n) \subseteq \sigma(X_1, \ldots, X_n,)$ porque las λ^s son funciones de las X^s entonces $E(\lambda_{n+1}|\lambda_1, \ldots, \lambda_n) = \lambda_n$.

Por lo tanto $\{\lambda_n, \sigma(\lambda_1, \ldots, \lambda_n)\}$ también es martingala. ■

El siguiente resultado es útil para construir martingalas cuando se tiene una colección de $\sigma-$álgebras y cualquier variable aleatoria.

Proposición 2.1.8. *Sean $\{\mathcal{F}_n : n \in \mathbb{N}\}$ tal que $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}$ es una colección de $\sigma-$álgebras creciente, X cualquier variable aleatoria tal que $E(|X|) < \infty$, entonces si definimos $X_n = E(X|\mathcal{F}_n)$ c.s. es una martingala.*

Demostración: X_n es $\mathcal{F}_n-$medible y si $\mathcal{G}_n = \sigma(X_1, \ldots, X_n)$ tenemos que $\mathcal{G}_n \subseteq \mathcal{F}_n$ entonces

$E[X_{n+1}|X_1, \ldots, X_n] = E(X_{n+1}|\mathcal{G}_n) = E[E(X_{n+1}|\mathcal{F}_n)|\mathcal{G}_n]$.

Por otro lado $E(X_{n+1}|\mathcal{F}_n) = E[E(X|\mathcal{F}_{n+1})|\mathcal{F}_n] = E[E(X|\mathcal{F}_n)|\mathcal{F}_{n+1}]$ porque $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$.

Esto es, $E(X_{n+1}|\mathcal{F}_n) = E(X_n|\mathcal{F}_{n+1})$.

De este modo

$$\begin{aligned} E(X_{n+1}|X_1,\ldots,X_n) &= E[E(X_n|\mathcal{F}_{n+1})|\mathcal{G}_n] \\ &= E[X_n|\mathcal{G}_n] \text{ porque } \mathcal{G}_n = \sigma(X_1,\ldots,X_n) \subseteq \mathcal{F}_n \subseteq \mathcal{F}_{n+1} \\ &= E(X_n|X_1,\ldots,X_n) = X_n \text{ c.s.} \end{aligned}$$

puesto que X_n es $\sigma(X_1,\ldots,X_n)$ medible, de este modo $\{X_n : n \in \mathbb{N}\}$ es martingala. $\square$

Observación. No toda martingala es de esta forma, es decir, si $\{X_n : n \in \mathbb{N}\}$ es martingala, no necesariamente existen una v.a. X con $E(|X|) < \infty$ y $\{\mathcal{F}_n : n \in \mathbb{N}\}$ una sucesión creciente de $\sigma-$álgebras tal que $X_n = E(X|\mathcal{F}_n)$, veamos el siguiente ejemplo.

Ejemplo 2.1.9. Sea $\{Z_n : n \in \mathbb{N}\}$ sucesión de v.a.i.i.d. tal que $P[Z_i = 0] = \frac{1}{2} = P[Z_i = 2]$ para todo $i \in \mathbb{N}$ asi $E(Z_i) = 0 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = 1$ para todo $i \in \mathbb{N}$ entonces $X_n = \prod_{i=1}^{n} Z_i$ es martingala por el Ejemplo 2.1.4.

Supongamos que existen X v.a. con $E(|X|) < \infty$ además de $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}$ $\sigma-$álgebras tal que $X_n = E(X|\mathcal{F}_n)$.

Sea $A_n = [X_n = 0] = X_n^{-1}(\{0\}) \in \mathcal{F}_n$ porque X_n es $\mathcal{F}_n-$medible.

Entonces por definición de esperanza condicional

$$0 = \int_{A_n} X_n dP = \int_{A_n} E(X|\mathcal{F}_n) dP = \int_{A_n} X dP = E(XI_{A_n}),$$

pero $I_{A_n} \xrightarrow[n\to\infty]{\text{c.s.}} 1$ porque $A_n^c = [X_n > 0] = \bigcap_{i=1}^{n} [Z_i > 0] \xrightarrow[n\to\infty]{} \varnothing$ además $P[X_n > 0] = \prod_{i=1}^{n} P[Z_i > 0] = (\frac{1}{2})^n \xrightarrow[n\to\infty]{} 0$.

Por lo tanto $P[X_n = 0] \xrightarrow[n\to\infty]{} 1$.

Luego $|XI_{A_n}| \leqslant |X|$ y por el Teorema de Convergencia Dominada se tiene que $E(XI_{A_n}) \xrightarrow[n\to\infty]{} E(X) = E(E(X|\mathcal{F}_n)) = E(X_n) = E[\prod_{i=1}^{n} Z_i] = \prod_{i=1}^{n} E(Z_i) =$

$\prod_{i=1}^{n} 1 = 1$ que es una contradicción, porque se tenía que $E(XI_{A_n}) = 0$.
Por lo tanto, no toda martingala es de la forma a la que se refiere la proposición 2.1.8. ■

Proposición 2.1.10. *Sea $\{X_n : n \in \mathbb{N}\}$ sucesión de v.a. con $E(|X_n|) < \infty$ para todo $n \in \mathbb{N}$. $\{X_n : n \in \mathbb{N}\}$ es martingala (submartingala o supermatingala) si y sólo si para todo $m \geqslant n$ y para todo $A \in \sigma(X_1, \ldots, X_n)$ se cumple que $\int_A X_m dP = \int_A X_n dP$ (cambiando $=$ por $\geqslant$ en el caso de submartingala e $=$ por $\leqslant$ en el caso de supermartingala).*

Demostración:
Primero supongamos que $\{X_n : n \in \mathbb{N}\}$ es martingala entonces para $n \leqslant m$ se tiene que $E(X_m|\mathcal{F}_n) = X_n$ con $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$.
Notemos que $\mathcal{F}_n \subseteq \mathcal{F}_m$.
Tenemos que para todo $A \in \mathcal{F}_n$, $A \in \mathcal{F}_m$ y por definición de esperanza condicional

$$\int_A X_n dP = \int_A E(X_m|\mathcal{F}_n) dP = \int_A X_m dP.$$

Si ahora suponemos que se vale la igualdad $\int_A X_m dP = \int_A X_n dP$ para cada $A \in \mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ donde $n \leqslant m$ entonces por la definición de esperanza condicional $X_n = E(X_m|\mathcal{F}_n)$, por lo tanto $\{X_n : n \in \mathbb{N}\}$ es martingala.
Los casos de submartingalas o supermartingalas son claros. □

2.2. Descomposición de Submartingalas y Martingalas Reversas

El siguiente resultado es el Teorema de Descomposición de Submartingalas.

Proposición 2.2.1. *Sea $\{X_n, \mathcal{F}_n : n \in \mathbb{N}\}$ una submartingala. Entonces X_n se descompone como $X_n = X'_n + X''_n$ c.s. donde $\{X'_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala y $\{X''_n : n \in \mathbb{N}\}$ es una sucesión creciente de v.a. mayores o iguales a 0 c.s. tal que X''_n es $\mathcal{F}_{n-1}-$medible para $n \geqslant 2$.*

Demostración:

X_n se puede escribir como $X_n = X_1 + (X_2 - X_1) + (X_3 - X_2) + \cdots + (X_n - X_{n-1}) = X_1 + \sum\limits_{j=1}^{n-1}(X_{j+1} - X_j) = X_1 + \sum\limits_{j=1}^{n-1}[X_{j+1} - E(X_{j+1}|\mathcal{F}_j) + E(X_{j+1}|\mathcal{F}_j) - X_j]$.

Sea $X'_1 = X_1$ y para $n \geqslant 2$ definimos X'_n como

$X'_n = X_1 + \sum\limits_{j=1}^{n-1}[X_{j+1} - E(X_{j+1}|\mathcal{F}_j)]$.

También hacemos $X''_1 = 0$ y para $n \geqslant 2:$ $X''_n = \sum\limits_{j=1}^{n-1}[E(X_{j+1}|\mathcal{F}_j) - X_j]$ por lo que $X_n = X'_n + X''_n$ para $n \geqslant 1$.

Sólo falta ver que $\{X'_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala y que $\{X''_n : n \in \mathbb{N}\}$ es una sucesión creciente no negativa de v.a., además X''_n es $\mathcal{F}_{n-1}$ medible para $n \geqslant 2$, entonces probaremos cada una de estas afirmaciones a continuación.

Como $\{X_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es submartingala se se cumple $E(X_{j+1}|\mathcal{F}_j) \geqslant X_j$ entonces $E(X_{j+1}|\mathcal{F}_j) - X_j \geqslant 0$.

Por tanto $X''_n = \sum\limits_{j=1}^{n-1}[E(X_{j+1}|\mathcal{F}_j) - X_j] \geqslant 0$ para todo $n \in \mathbb{N}$.

Además $X''_{n+1} - X''_n = E(X_{n+1}|\mathcal{F}_n) - X_n \geqslant 0$ implica que $X''_{n+1} \geqslant X''_n$ para todo $n \in \mathbb{N}$.

Luego $\{X''_n : n \in \mathbb{N}\}$ es creciente.

$X''_n = \sum_{j=1}^{n-1}[E(X_{j+1}|\mathcal{F}_j) - X_j]$ pero $E(X_{j+1}|\mathcal{F}_j)$ y X_j son $\mathcal{F}_j-$ medibles.

Como $\mathcal{F}_j \subseteq \mathcal{F}_{n-1}$ para todo $j = 1, \ldots, n-1$ entonces $E(X_{j+1}|\mathcal{F}_j) - X_j$ es $\mathcal{F}_{n-1}-$ medible y por tanto X''_n es $\mathcal{F}_{n-1}-$ medible.

Ahora veremos que $\{X'_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala.

$E(X'_{n+1}|\mathcal{F}_n) = E[X_1 + \sum_{j=1}^{n}(X_{j+1} - E(X_{j+1}|\mathcal{F}_j))|\mathcal{F}_n]$.

Notemos que $X_1, X_2, \ldots, X_n$ son $\mathcal{F}_n-$ medibles.

Así $E(X'_{n+1}|\mathcal{F}_n) = X_1 + \sum_{j=1}^{n} E(X_{j+1}|\mathcal{F}_n) - \sum_{j=1}^{n} E[E(X_{j+1}|\mathcal{F}_j)|\mathcal{F}_n] = X_1 + \sum_{j=1}^{n-1} X_{j+1} + E(X_{n+1}|\mathcal{F}_n) - \sum_{j=1}^{n-1} E[E(X_{j+1}|\mathcal{F}_j)|\mathcal{F}_n] - E[E(X_{n+1}|\mathcal{F}_n)|\mathcal{F}_n]$.

De nuevo como $\mathcal{F}_j \subseteq \mathcal{F}_n$ $j = 1, \ldots, n$ tenemos que

$E(X'_{n+1}|\mathcal{F}_n) = \sum_{j=1}^{n} X_j + E(X_{n+1}|\mathcal{F}_n) - \sum_{j=1}^{n-1} E(X_{j+1}|\mathcal{F}_j) - E(X_{n+1}|\mathcal{F}_n) = \sum_{j=1}^{n} X_j - \sum_{j=1}^{n-1} E(X_{j+1}|\mathcal{F}_j) = X_1 + \sum_{j=1}^{n-1}[X_{j+1} - E(X_{j+1}|\mathcal{F}_j)] = X'_n$.

Por lo tanto $\{X'_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala. □

Observación. Si $X_n = X'_n + X''_n$ con X''_n no negativa, X''_n es creciente y $\mathcal{F}_{n-1}$−medible y además $\{X'_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala, entonces $\{X_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es submartingala.

Es decir, la proposición anterior puede ser un si y sólo sí.

Demostración:

Basta ver que $E(X_{n+1}|\mathcal{F}_n) \geqslant X_n$.

$E(X_{n+1}|\mathcal{F}_n) = E(X'_{n+1} + X''_{n+1}|\mathcal{F}_n) = E(X'_{n+1}|\mathcal{F}_n) + E(X''_{n+1}|\mathcal{F}_n) = X'_n + X''_{n+1} \geqslant X'_n + X''_n = X_n$ porque $\{X''_n : n \in \mathbb{N}\}$ es creciente. □

Observación. Si $\{X_n : n \in \mathbb{N}\}$ es submartingala entonces $X_n = X'_n + X''_n$ implica que $X'_n = X_n - X''_n$.
Luego $|X'_n| \leqslant |X_n| + |X''_n| = |X_n| + X''_n$ porque $X''_n \geqslant 0$.
Así $E(|X'_n|) \leqslant E(|X_n|) + E(X''_n)$, pero la sucesión $\{E(X''_n) : n \in \mathbb{N}\}$ es creciente y no negativa.
De aqui que si $\sup\limits_{n\geqslant 1} E(|X_n|) < \infty$ entonces $E(X'_n)$ y $E(X''_n)$ existen porque $E(X_n) = E(X'_n) + E(X''_n)$ y además $\sup\limits_{n\geqslant 1} E(|X'_n|) < \infty$ y $\sup\limits_{n\geqslant 1} E(|X''_n|) < \infty$.
De este modo $X''_n \uparrow X''$ c.s. porque es una sucesión creciente con esperanzas acotadas.
El estudio de la convergencia de X_n se reduce al estudio de X'_n que es martingala, siempre y cuando $\sup\limits_{n\geqslant 1} E(|X_n|) < \infty$. $\square$

Proposición 2.2.2. *a) Sean $\{X_n : n \in \mathbb{N}\}$ martingala y $g : \mathbb{R} \longrightarrow \mathbb{R}$ convexa, entonces $\{g(X_n) : n \in \mathbb{N}\}$ es submartingala cuando $E(|g(X_n)|) < \infty$ para todo $n \in \mathbb{N}$.*

b) Si $\{X_n : n \in \mathbb{N}\}$ es submartingala, $g : \mathbb{R} \longrightarrow \mathbb{R}$ convexa y creciente tal que $E(|g(X_n)|) < \infty$ para todo $n \in \mathbb{N}$ entonces $\{g(X_n) : n \in \mathbb{N}\}$ también es submartingala.

Demostración:

a) $E(X_n|\mathcal{F}_{n-1}) = X_{n-1}$ pero g es convexa y por la desigualdad de Jensen condicional
$E[g(X_n)|\mathcal{F}_{n-1}] \geqslant g(E(X_n|\mathcal{F}_{n-1})) = g(X_{n-1})$, por lo tanto $\{g(X_n) : n \in \mathbb{N}\}$ es submartingala.

b) Como $\{X_n : n \in \mathbb{N}\}$ es submartingala sabemos que $E(X_n|\mathcal{F}_{n-1}) \geqslant X_{n-1}$.

$E[g(X_n)|\mathcal{F}_{n-1}] \geqslant g(E(X_n|\mathcal{F}_{n-1})) \geqslant g(X_{n-1})$ por ser g creciente. □

Corolario 2.2.3. *Si* $\{X_n : n \in \mathbb{N}\}$ *es submartingala entonces* $\{Y_n : n \in \mathbb{N}\}$ *donde* $Y_n = \text{máx}\{X_n, a\}$, $a \in \mathbb{R}$ *también es submartingala.*

Demostración:

Sabemos que $g(x) = x^+ = \text{máx}\{x, 0\}$ es una función convexa no negativa. Además es creciente.

Por lo tanto por la proposición 2.2.2 $\{X_n^+ : n \in \mathbb{N}\}$ es submartingala. (es claro que $E(|X_n^+|) = E(X_n^+)$ y existe porque X_n es submartingala) Además $\{X_n - a : n \in \mathbb{N}\}$ es submartingala por lo que $(X_n - a)^+$ es submartingala. También $\{(X_n - a)^+ + a : n \in \mathbb{N}\}$ es submartingala, pero es claro que $(X_n - a)^+ + a = \text{máx}\{X_n, a\} = Y_n$.

Por lo tanto $\{Y_n : n \in \mathbb{N}\}$ es submartingala. □

Una sucesión $\{X_n : n \in \mathbb{N}\}$ de variables aleatorias con $E(|X_n|) < \infty$ para todo $n \in \mathbb{N}$ es martingala reversa si para todo $n \in \mathbb{N}$ tenemos que $E(X_n|X_{n+1}, X_{n+2}, \ldots) = X_{n+1}$ c.s., submartingala reversa si $E(X_n|X_{n+1}, X_{n+2}, \ldots) \geqslant X_{n+1}$ c.s. para todo $n \in \mathbb{N}$, supermartingala reversa si $E(X_n|X_{n+1}, X_{n+2}, \ldots) \leqslant X_{n+1}$ c.s.

Proposición 2.2.4. $\{X_n : n \in \mathbb{N}\}$ *es martingala reversa si y sólo si para todo* $m > n$ *se cumple que* $E(X_n|X_m, X_{m+1}, \ldots) = X_m$ *c.s.*

Demostración:

Definamos $\mathcal{F}_m = \sigma(X_m, X_{m+1}, \ldots)$.

Sea $m > n$, esto es, $m \geqslant n+1$ entonces
$\mathcal{F}_m = \sigma(X_m, X_{m+1}, \ldots) \subseteq \sigma(X_{n+1}, \ldots, X_m, X_{m+1}, \ldots) = \mathcal{F}_{n+1}$.
Supongamos que $E(X_n|\mathcal{F}_m) = X_m$ para todo $m > n$, en particular si $m = n+1$ tenemos $E(X_n|\mathcal{F}_{n+1}) = X_{n+1}$, esto es, $E(X_n|X_{n+1}, X_{n+2}, \ldots) = X_{n+1}$ c.s. Entonces $\{X_n : n \in \mathbb{N}\}$ es martingala reversa.
Ahora supongamos que $\{X_n : n \in \mathbb{N}\}$ es martingala reversa, esto es, $E(X_n|\mathcal{F}_{n+1}) = X_{n+1}$ c.s. donde $\mathcal{F}_{n+1} = \sigma(X_{n+1}, X_{n+2}, \ldots)$ y tomemos $m > n$ cualquiera. Debemos mostrar que $E(X_n|X_m, X_{m+1}, \ldots) = E(X_n|\mathcal{F}_m) = X_m$ c.s.
Pero $E(X_n|\mathcal{F}_m) = E[E(X_n|\mathcal{F}_{n+1})|\mathcal{F}_m]$ porque $\mathcal{F}_m \subseteq \mathcal{F}_{n+1}$ debido a que $\mathcal{F}_m = \sigma(X_m, X_{m+1}, \ldots) \subseteq \sigma(X_{n+1}, \ldots, X_m, X_{m+1}, \ldots) = \mathcal{F}_{n+1}$ y por hipótesis $E(X_n|\mathcal{F}_{n+1}) = X_{n+1}$.
Por lo que $E(X_n|\mathcal{F}_m) = E(X_{n+1}|\mathcal{F}_m) = \cdots = E(X_{m-1}|\mathcal{F}_m) = X_m$ c.s. $\square$
Varias de las propiedades de martingalas se cumplen con martingalas reversas haciendo las modificaciones correspondientes, aunque no todas.

Proposición 2.2.5. *Si $\{X_n : n \in \mathbb{N}\}$ es martingala reversa entonces existen:*

i) Una v.a. X con $E(|X|) < \infty$ y

ii) Una sucesión $\mathcal{F}_n$ de $\sigma-$álgebras decreciente tal que $E(X|\mathcal{F}_n) = X_n$.

Demostración:

i) Sea $X = X_1$, que como $\{X_n : n \in \mathbb{N}\}$ es martingala reversa se cumple que $E(|X_n|) < \infty$ para todo $n \in \mathbb{N}$ en particular si $n = 1$, $E(|X_1|) < \infty$, por lo tanto $E(|X|) < \infty$.

ii) Tomemos $\mathcal{F}_n = \sigma(X_n, X_{n+1}, \ldots)$ y veamos que esta colección de $\sigma-$álgebras nos sirve.

$\mathcal{F}_n = \sigma(X_n, X_{n+1}, \ldots) \supseteq \sigma(X_{n+1}, X_{n+2}, \ldots) = \mathcal{F}_{n+1}$ por tanto la sucesión de $\sigma-$álgebras es decreciente.

Sólo nos falta comprobar que $X_n = E(X|\mathcal{F}_n)$ pero

$E(X|\mathcal{F}_n) = E(X_1|X_n, X_{n+1}, \ldots) = X_n$ c.s. aplicando directamente la proposición 2.2.4. □

Cuando se tiene una sucesión de variables aleatorias $\{X_n : n \in \mathbb{N}\}$ con $E(X_n) = 0$ para todo $n \in \mathbb{N}$ y $S_n = \sum_{k=1}^{n} X_k$, en el caso en que las v.a. X_i sean independientes hay algunos resultados sobre $\overline{X}_n = \frac{S_n}{n}$ en cuanto a la convergencia o no de esta nueva sucesión.

Esto se debe a que $E(X_j X_k) = 0$ si $j \neq k$, esta propiedad es cierta en martingalas con incrementos que tienen varianza finita.

En próximos resultados consideraremos $\{S_n, \mathcal{F}_n : n \in \mathbb{N}\}$ martingala con $E(S_n) = 0$ para todo $n \in \mathbb{N}$ tal que $X_n = S_n - S_{n-1}$, $S_0 = 0$, $E(X_n^2) < \infty$, $Var(S_n - S_{n-1}) = E(X_n^2)$ para todo $n \in \mathbb{N}$.

Proposición 2.2.6. *Sea $\{S_n, \mathcal{F}_n : n \in \mathbb{N}\}$ una martingala con $E(S_n) = 0$ para todo $n \in \mathbb{N}$ tal que $X_n = S_n - S_{n-1}$, $S_0 = 0$, , $E(X_n^2) < \infty$, $Var(S_n - S_{n-1}) = E(X_n^2)$ para todo $n \in \mathbb{N}$. Entonces $P[|S_n| \geqslant \epsilon] \leqslant \sum_{j=1}^{n} \frac{E(X_j^2)}{\epsilon^2}$ con $\epsilon > 0$.*

Demostración:

La desigualdad de Chebyshev establece que $P[|Y - E(Y)| \geqslant k] \leqslant \dfrac{Var(Y)}{k^2}$ para Y v.a. con media $E(Y)$ y varianza $Var(Y)$ y $k > 0$ cualquiera.

Tomando $Y = S_n$ y $k = \epsilon$ obtenemos

$$P[|S_n - E(S_n)| \geqslant \epsilon] \leqslant \frac{Var(S_n)}{\epsilon^2}$$

pero $E(S_n) = 0$ y $Var(S_n) = E(S_n^2) - E^2(S_n) = E(S_n^2)$.

Por lo que $P[|S_n)| \geqslant \epsilon] \leqslant \dfrac{E(S_n^2)}{\epsilon^2}$.

Ahora calcularemos $E(S_n^2)$ y veremos que es igual a $\sum\limits_{j=1}^{n} E(X_j^2)$.

$E(S_n^2) = E[(X_1 + \cdots + X_n)^2] = E[\sum\limits_{j=1}^{n} X_j^2 + 2 \sum\limits_{1 \leqslant i < j \leqslant n} X_i X_j] =$
$\sum\limits_{j=1}^{n} E(X_j^2) + 2 \sum\limits_{1 \leqslant i < j \leqslant n} E(X_i X_j)$.

Por otro lado si $i \leqslant j - 1$ y como:

X_i es $\mathcal{F}_i$−medible y $\mathcal{F}_i \subseteq \mathcal{F}_{j-1}$ entonces X_i es $\mathcal{F}_{j-1}$−medible.

$E(X_i X_j) = E[E(X_i X_j)|\mathcal{F}_{j-1}] = E[X_i E(X_j|\mathcal{F}_{j-1})] = E[X_i E(S_j - S_{j-1}|\mathcal{F}_{j-1})] =$
$E[X_i[E(S_j|\mathcal{F}_{j-1}) - E(S_{j-1}|\mathcal{F}_{j-1})]] = E[X_i[E(S_j|\mathcal{F}_{j-1}) - S_{j-1}]]$
$= E[X_i(S_{j-1} - S_{j-1})] = 0$.

Porque $\{S_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala y S_{j-1} es $\mathcal{F}_{j-1}$−medible.

Por lo tanto $E(X_i X_j) = 0$ si $i < j$ entonces $E(S_n^2) = \sum\limits_{j=1}^{n} E(X_j^2)$.

Por lo tanto $P[|S_n)| \geqslant \epsilon] \leqslant \displaystyle\sum_{j=1}^{n} \frac{E(X_j^2)}{\epsilon^2}$. □

Corolario 2.2.7. *Si* $\displaystyle\sum_{j=1}^{n} \frac{E(X_j^2)}{n^2} \xrightarrow[n\to\infty]{} 0$ *entonces se cumple la ley débil de los grandes números, es decir,* $\dfrac{S_n}{n} \xrightarrow[n\to\infty]{P} 0$ *o equivalentemente* $P[|\frac{S_n}{n}| \geqslant \epsilon] \xrightarrow[n\to\infty]{} 0$ *para todo* $\epsilon > 0$.

Demostración:

Sea $\epsilon > 0$ entonces $P[|\frac{S_n}{n}| \geqslant \epsilon] = P[|S_n| \geqslant n\epsilon] \leqslant \sum_{j=1}^{n} \frac{E(X_j^2)}{n^2\epsilon^2} = \frac{1}{\epsilon^2} \sum_{j=1}^{n} \frac{E(X_j^2)}{n^2} \xrightarrow[n\to\infty]{}$ 0, por lo tanto $P[|\frac{S_n}{n}| \geqslant \epsilon] \xrightarrow[n\to\infty]{} 0$. □

Veremos ahora algunas definiciones y resultados no condicionales que nos permitirán seguir adelante en la prueba de los resultados condicionales que nos interesan.

Si X es integrable $(E(|X|) < \infty)$ y elegimos $c = E(X)$ decimos que X está centrada en su esperanza si $E(X) = 0$.

Si $c > 0$ se define la variable aleatoria truncada en c, X^c como

$$X^c = \begin{cases} X & \text{si } |X| < c \\ 0 & \text{si } |X| \geqslant c \end{cases}.$$

Notemos que X^c es una v.a. acotada por lo que tiene momentos de todos los órdenes.

Dos sucesiones de variables aleatorias $\{X_n : n \in \mathbb{N}\}$ y $\{X'_n : n \in \mathbb{N}\}$ tienen colas equivalentes si para casi todo $\omega \in \Omega$ existe $n(\omega) \in \mathbb{N}$ tal que para todo $n \geqslant n(\omega)$, $X_n(\omega) = X'_n(\omega)$, es decir, $P[X_n \neq X'_n \ i.o.] = 0$.

Las iniciales *i.o.* significan "infinitamente seguido".

Una sucesión de eventos E_n que tiene límite superior se acostumbra escribir $\limsup_{n\to\infty} E_n$ ó bien E_n *i.o.*.

E_n *i.o.* es de hecho el evento $\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k$, por lo que $[X_n \neq X'_n \ i.o.] = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} [X_k \neq X'_k]$.

Las sucesiones $\{X_n : n \in \mathbb{N}\}$ y $\{X'_n : n \in \mathbb{N}\}$ son equivalentemente convergentes si convergen en el mismo evento excepto por un conjunto de probabilidad cero.

Estos últimos conceptos aplicados a teoremas límites se encuentran en Neuts (1973).

Proposición 2.2.8. *Sean* $\{X_n : n \in \mathbb{N}\}$ *y* $\{X'_n : n \in \mathbb{N}\}$ *dos sucesiones de variables aleatorias tal que* $\sum_{n=1}^{\infty} P[X_n \neq X'_n] < \infty$, *entonces:*

a) $\{X_n : n \in \mathbb{N}\}$ *y* $\{X'_n : n \in \mathbb{N}\}$ *tienen colas equivalentes.*

b) $\sum_{n=1}^{\infty} X_n$ *y* $\sum_{n=1}^{\infty} X'_n$ *son equivalentemente convergentes.*

c) $\frac{\sum_{k=1}^{n} X_k}{b_n}$ *y* $\frac{\sum_{k=1}^{n} X'_k}{b_n}$ *donde* $b_n \uparrow \infty$ *converge en el mismo evento y al mismo límite excepto en un conjunto nulo (de probabilidad cero)*

Demostración:

a) Debemos mostrar que $P[X_n \neq X'_n \ i.o.] = 0$.
Sea E_n el evento $[X_n \neq X'_n]$ entonces $P[X_n \neq X'_n \ i.o.] = P[\limsup E_n] = P[\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} E_k]$, ahora si hacemos $B_n = \bigcup_{k=n}^{\infty} E_k$ notamos que $B_{n+1} = \bigcup_{k=n+1}^{\infty} E_k \subseteq \bigcup_{k=n}^{\infty} E_k = B_n$ por lo que la sucesión de eventos B_n es decreciente.
$P[X_n \neq X'_n \ i.o] = P[\bigcap_{n=1}^{\infty} B_n] = \lim_{n\to\infty} P(B_n) = \lim_{n\to\infty} P[\bigcup_{k=n}^{\infty} E_k]$
$\leqslant \lim_{n\to\infty} \sum_{k=n}^{\infty} P(E_k) = \lim_{n\to\infty} \sum_{k=n}^{\infty} P[X_k \neq X'_k] = 0$ por ser la cola de una serie convergente.

b) Por a) sabemos que $\{X_n : n \in \mathbb{N}\}$ y $\{X'_n : n \in \mathbb{N}\}$ tienen colas equivalentes, esto es, para casi todo $\omega \in \Omega$ existe $n(\omega) \in \mathbb{N}$ tal que

para todo $n \geqslant n(\omega)$ tenemos que $X_n(\omega) = X'_n(\omega)$.
$\sum\limits_{n=n(\omega)}^{\infty} X_n(\omega) = \sum\limits_{n=n(\omega)}^{\infty} X'_n(\omega)$
entonces para casi todo $\omega \in \Omega$, $\sum\limits_{n=1}^{\infty} X_n(\omega)$ y $\sum\limits_{n=1}^{\infty} X_n(\omega)$ difieren a lo más en un número finito de términos.
Asi para casi todo $\omega \in \Omega$, $\sum\limits_{n=1}^{\infty} X_n(\omega)$ converge si y sólo si $\sum\limits_{n=1}^{\infty} X'_n(\omega)$ converge.
Por lo tanto $\sum\limits_{n=1}^{\infty} X_n$ y $\sum\limits_{n=1}^{\infty} X'_n$ son equivalentemente convergentes.

c) por a) de nuevo tenemos que para casi todo $\omega \in \Omega$ existe $n(\omega)$ tal que para todo $k \geqslant n(\omega)$ $\quad X_k(\omega) = X'_k(\omega)$.
Luego si n es grande $\sum\limits_{k=1}^{n}(X_k(\omega) - X'_k(\omega)) = \sum\limits_{k=1}^{n(\omega)-1}(X_k(\omega) - X'_k(\omega)) + \sum\limits_{k=n(\omega)}^{n}(X_k(\omega) - X'_k(\omega)) = \sum\limits_{k=1}^{n(\omega)-1}(X_k(\omega) - X'_k(\omega))$.
Asi $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \frac{(X_k(\omega) - X'_k(\omega))}{b_n} = \lim\limits_{n\to\infty} \sum\limits_{k=1}^{n(\omega)-1} \frac{(X_k(\omega) - X'_k(\omega))}{b_n} =$
$\lim\limits_{n\to\infty} \frac{d(\omega)}{b_n} = d(\omega) \lim\limits_{n\to\infty} \frac{1}{b_n} = d(\omega)\cdot 0 = 0 \quad$ donde $d(\omega) = \sum\limits_{k=1}^{n(\omega)-1}(X_k(\omega) - X'_k(\omega))$ es una cantidad finita que no depende de n sólo de ω.
Por lo que para casi todo $\omega \in \Omega$, $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \frac{(X_k(\omega) - X'_k(\omega))}{b_n} = 0$.
Entonces para casi todo $\omega \in \Omega$, $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \frac{X_k(\omega)}{b_n} = \lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} \frac{X'_k(\omega)}{b_n}$.
Por lo tanto $\sum\limits_{k=1}^{n} \frac{X_k}{b_n}$ y $\sum\limits_{k=1}^{n} \frac{X'_k}{b_n}$ convergen en el mismo evento y al mismo límite excepto en un conjunto de probabilidad cero. □

Proposición 2.2.9. *Sean $X_1, X_2, \ldots, X_n$ variables aleatorias independientes con $E(X_k) = 0$ y $Var(X_k) = E(X_k^2) = \sigma_k^2 < \infty$ para $k = 1, 2, \ldots, n$.*

Hacemos $S_k = \sum_{j=1}^{k} X_j$ *para* $k = 1, 2, \ldots, n$
Sea $\epsilon > 0$ *entonces*

$$P[\max_{1\leqslant k\leqslant n} |S_k| \geqslant \epsilon] \leqslant \frac{1}{\epsilon^2}\sum_{k=1}^{n}\sigma_k^2$$

Demostración:

Para $\epsilon > 0$ fija definimos los eventos $E_1 = [|S_1| \geqslant \epsilon]$ y para $k = 2, 3, \ldots, n$ hacemos $E_k = [|S_k| \geqslant \epsilon] \cap [\bigcap_{j=1}^{k-1}[|S_j| < \epsilon]]$ es decir, el evento E_k ocurre si y sólo si la primera vez que la suma S_j en valor absoluto iguala o excede a ϵ es para $j = k$.

Los E_k son ajenos, porque E_k es de la forma $E_k = B_k \cap B_{k-1}^c \cap \cdots \cap B_1^c$ donde $B_j = [|S_j| \geqslant \epsilon]$.

Ahora veremos que $E = [\max_{1\leqslant k\leqslant n} |S_k| \geqslant \epsilon] = \bigcup_{k=1}^{n} E_k$.
Sea $\omega \in E$ esto es, $\omega \in [\max_{1\leqslant k\leqslant n} |S_k| \geqslant \epsilon]$ entonces $\max_{1\leqslant k\leqslant n} |S_k(\omega)| \geqslant \epsilon$.
Supongamos sin pérdida de generalidad que $\max_{1\leqslant k\leqslant n} |S_k(\omega)| = |S_m(\omega)|$.
Entonces$|S_m(\omega)| \geqslant \epsilon$ implica que $\omega \in E_m$ o en un E_k con $k \leqslant m$.
Asi $\omega \in \bigcup_{k=1}^{n} E_k$.
Ahora sea $\omega \in \bigcup_{k=1}^{n} E_k$ entonces existe $1 \leqslant m \leqslant n$ tal que $\omega \in E_m$ implica que $|S_m(\omega)| \geqslant \epsilon$.
Pero $\max_{1\leqslant k\leqslant n} |S_k(\omega)| \geqslant |S_m(\omega)| \geqslant \epsilon$.
Por lo que $\max_{1\leqslant k\leqslant n} |S_k(\omega)| \geqslant \epsilon$.
Por lo tanto $\omega \in [\max_{1\leqslant k\leqslant n} |S_k| \geqslant \epsilon]$.
Por otro lado $\sum_{k=1}^{n} \sigma_k^2 = \sum_{k=1}^{n} Var(X_k) = Var(\sum_{k=1}^{n} X_k)$ por independencia.
$\sum_{k=1}^{n} \sigma_k^2 = Var(S_n) = E(S_n^2) - E^2(S_n) = E(S_n^2)$ porque $E(S_n) = 0$.

$\sum\limits_{k=1}^{n}\sigma_k^2=\int\limits_{\Omega}S_n^2dP\geqslant\int\limits_{E}S_n^2dP=\sum\limits_{k=1}^{n}\int\limits_{E_k}S_n^2dP$ porque $S_n^2\geqslant 0$ donde $E=\bigcup\limits_{k=1}^{n}E_k$.
Ahora para $1\leqslant k\leqslant n$ $\int\limits_{E_k}S_n^2dP=\int\limits_{E_k}(S_k+X_{k+1}+\cdots+X_n)^2dP=\int\limits_{E_k}S_k^2dP+$
$\sum\limits_{j=k+1}^{n}\int\limits_{E_k}X_j^2dP+2\sum\limits_{j=k+1}^{n}\int\limits_{E_k}S_kX_jdP+\sum\limits_{\substack{j\neq j'\\ j,j'=k+1}}^{n}\int\limits_{E_k}X_jX_{j'}dP$

Calcularemos cada uno de los términos anteriores.
Si $j=1,\ldots,n-k$ $\int\limits_{E_k}S_kX_{k+j}dP=E[S_kI_{E_k}X_{k+j}]$ pero
$S_kI_{E_k}$ es $\sigma(X_1,\ldots,X_k)$ medible.
Mientras que X_{k+j} es $\sigma(X_{k+j})$ medible y por independencia se tiene que $E[S_kI_{E_k}X_{k+j}]=E[S_kI_{E_k}]E(X_{k+j})=0$.
Ahora si $j\neq j'$
$$\int\limits_{E_k}X_{k+j}X_{k+j'}dP=E[I_{E_k}X_{k+j}X_{k+j'}]=E(I_{E_k})E(X_{k+j})E(X_{k+j'})$$
$$=P(E_k)\cdot 0=0 \text{ por independencia.}$$
Luego $\int\limits_{E_k}S_n^2dP=\int\limits_{E_k}S_k^2dP+\sum\limits_{j=k+1}^{n}\int\limits_{E_k}X_j^2dP\geqslant\int\limits_{E_k}S_k^2dP\geqslant\int\limits_{E_k}\epsilon^2dP$
la última desigualdad es por la definición de E_k.

Así
$$\int\limits_{E_k}S_n^2dP\geqslant\int\limits_{E_k}\epsilon^2dP=\epsilon^2P(E_k).$$

Por todo lo anterior
$$\sum_{k=1}^{n}\sigma_k^2\geqslant\sum_{k=1}^{n}\int\limits_{E_k}S_n^2dP\geqslant\epsilon^2\sum_{k=1}^{n}P(E_k)=\epsilon^2P(E).$$
$$\text{Por lo tanto } \sum_{k=1}^{n}\sigma_k^2\geqslant\epsilon^2P[\max_{1\leqslant k\leqslant n}|S_k|\geqslant\epsilon]$$
$$\text{implica que } P[\max_{1\leqslant k\leqslant n}|S_k|\geqslant\epsilon]\leqslant\frac{1}{\epsilon^2}\sum_{k=1}^{n}\sigma_k^2.\ \square$$

Corolario 2.2.10. *Sean $X_1, \dots, X_n$ v.a.i. con $Var(X_k) = \sigma_k^2 < \infty$ para $k = 1, \dots, n$. Si $\epsilon > 0$ $\quad P[\max\limits_{1 \leqslant k \leqslant n} |S_k - E(S_k)| \geqslant \epsilon] \leqslant \frac{1}{\epsilon^2} \sum\limits_{k=1}^{n} \sigma_k^2$.*

Demostración:

Sea $Y_k = X_k - E(X_k)$ entonces $E(Y_k) = 0$ por lo que $Var(Y_k) = E(Y_k^2) = Var(X_k) = \sigma_k^2$, además $Y_1, Y_2, \dots, Y_n$ son independientes por la proposición 2.2.9 se tiene

$$P[\max_{1 \leqslant k \leqslant n} |Y_1 + \cdots + Y_k| \geqslant \epsilon] \leqslant \frac{1}{\epsilon^2} \sum_{k=1}^{n} Var(Y_k)$$

$$\text{Entonces } P[\max_{1 \leqslant k \leqslant n} |S_k - E(S_k)| \geqslant \epsilon] \leqslant \frac{1}{\epsilon^2} \sum_{k=1}^{n} \sigma_k^2$$

porque $\sum\limits_{j=1}^{k} Y_j = (X_1 - E(X_1)) + \cdots + (X_k - E(X_k)) = (X_1 + \cdots + X_k) - (E(X_1) + \cdots + E(X_k)) = S_k - E(S_k)$. $\square$

En el corolario anterior si $n = 1$ se tiene la desigualdad de Chebyshev.

Proposición 2.2.11. *Sean $X_1, \dots, X_n$ variables aleatorias independientes, $\gamma > 0$ tal que $P[|X_k| \leqslant \gamma] = 1$ para todo $k = 1, \dots, n$.*
Si $\epsilon > 0$ entonces $P[\max\limits_{1 \leqslant k \leqslant n} |S_k - E(S_k)| \geqslant \epsilon] \geqslant 1 - \frac{(\epsilon + 2\gamma)^2}{Var(S_n)}$.

Demostración:

Se tiene que con probabilidad 1 $-\gamma \leqslant X_k \leqslant \gamma$.

Entonces $-\gamma \leqslant E(X_k) \leqslant \gamma$ implica que $-\gamma \leqslant -E(X_k) \leqslant \gamma$.

Así $-2\gamma \leqslant X_k - E(X_k) \leqslant 2\gamma$.

Luego $|X_k - E(X_k)| \leqslant 2\gamma$.

Podemos suponer que X_k está centrado en $E(X_k)$ y que satisface $P[|X_k| \leqslant 2\gamma] = 1$.

Sean ahora $\epsilon > 0$ y $S_0 = 0$.
Definimos $F_k = \bigcap_{i=1}^{k} [|S_i| < \epsilon]$ para $k = 1, \ldots, n$ con $F_0 = \Omega$.
Notemos que $F_0 \supseteq F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n$.
Además $F_n^c = \left[\bigcap_{i=1}^{n} [|S_i| < \epsilon]\right]^c = \bigcup_{i=1}^{n} [|S_i| \geqslant \epsilon] = [\max_{1 \leqslant k \leqslant n} |S_k| \geqslant \epsilon]$.

Como en la proposición 2.2.9, $E_k = [|S_1| < \epsilon, \ldots, |S_{k-1}| < \epsilon, |S_k| \geqslant \epsilon]$
$E = [\max_{1 \leqslant k \leqslant n} |S_k| \geqslant \epsilon] = F_n^c$.
Pero como $F_k = [|S_1| < \epsilon, \ldots, |S_k| < \epsilon]$ tenemos que $E_k \cup F_k = [|S_1| < \epsilon, \ldots, |S_{k-1}| < \epsilon] = F_{k-1}$, $\quad E_k \cap F_k = \varnothing$.

Ahora calcularemos la cantidad $I_k = \int_{F_k} S_k^2 dP - \int_{F_{k-1}} S_{k-1}^2 dP$.

$\int_{F_k} S_k^2 dP = \int_{F_{k-1}} S_k^2 dP - \int_{E_k} S_k^2 dP$ por lo que

$I_k = \int_{F_{k-1}} S_k^2 dP - \int_{F_{k-1}} S_{k-1}^2 dP - \int_{E_k} S_k^2 dP = \int_{F_{k-1}} (S_k^2 - S_{k-1}^2) dP - \int_{E_k} S_k^2 dP$,

pero como $S_k^2 = (X_1 + \cdots + X_k)^2 = (S_{k-1} + X_k)^2 = S_{k-1}^2 + 2X_k S_{k-1} + X_k^2$ entonces

$I_k = \int_{F_{k-1}} (X_k^2 + 2X_k S_{k-1}) dP - \int_{E_k} S_k^2 dP = \int_{F_{k-1}} X_k^2 dP + 2 \int_{F_{k-1}} X_k S_{k-1} dP - \int_{E_k} S_k^2 dP$.

Pero $\int_{F_{k-1}} X_k^2 dP = \int_{\Omega} I_{F_{k-1}} X_k^2 dP = E[I_{F_{k-1}} X_k^2] = E[I_{F_{k-1}}] E[X_k^2]$
$= P(F_{k-1}) E(X_k^2) \geqslant P(F_n) E(X_k^2)$ porque $F_n \subseteq F_{k-1}$.

Por otro lado

$\int_{F_{k-1}} X_k S_{k-1} dP = \int_{\Omega} I_{F_{k-1}} S_{k-1} X_k dP = E(I_{F_{k-1}} S_{k-1} X_k)$
$= E(I_{F_{k-1}} S_{k-1}) E(X_k) = 0$ por independencia y porque $E(X_k) = 0$.

Sobre el conjunto $E_k = [|S_1| < \epsilon, \ldots, |S_{k-1}| < \epsilon, |S_k| \geqslant \epsilon]$
tenemos que $\quad |S_k| = |S_{k-1} + X_k| \leqslant |S_{k-1}| + |X_k| \leqslant \epsilon + 2\gamma.$

Así $\int\limits_{E_k} S_k^2 dp \leqslant \int\limits_{E_k} (\epsilon + 2\gamma)^2 dP = (\epsilon + 2\gamma)^2 P(E_k)$

por lo que $-\int\limits_{E_k} S_k^2 dP \geqslant -(\epsilon + 2\gamma)^2 P(E_k)$.
Entonces
$I_k = \int\limits_{F_{k-1}} X_k^2 dP + 2\int\limits_{F_{k-1}} X_k S_{k-1} dP - \int\limits_{F_k} S_k^2 dP \geqslant P(F_n)E(X_k^2) - (\epsilon+2\gamma)^2 P(E_k)$
$k = 1, \ldots, n$
$\sum\limits_{k=1}^{n} I_k = \sum\limits_{k=1}^{n} \left(\int\limits_{F_k} S_k^2 dP - \int\limits_{F_{k-1}} S_{k-1}^2 dP \right) = \left(\int\limits_{F_1} S_1^2 dP - \int\limits_{F_0} S_0^2 dP \right)$
$+ \left(\int\limits_{F_2} S_2^2 dP - \int\limits_{F_1} S_1^2 dP \right) + \cdots + \left(\int\limits_{F_n} S_n^2 dP - \int\limits_{F_{n-1}} S_{n-1}^2 dP \right) = \int\limits_{F_n} S^2_- n dP \geqslant$
$\sum\limits_{k=1}^{n} [P(F_n)E(X_k^2) - (\epsilon+2\gamma)^2 P(E_k)] = P(F_n) \sum\limits_{k=1}^{n} E(X_k^2) - (\epsilon+2\gamma)^2 \sum\limits_{k=1}^{n} P(E_k) =$
$P(F_n) \sum\limits_{k=1}^{n} E(X_k^2) - (\epsilon + 2\gamma)^2 P(E)$ porque $E = \bigcup\limits_{k=1}^{n} E_k$ con los E_k ajenos.

Por lo tanto $\int\limits_{F_n} S_n^2 dP \geqslant P(F_n) Var(S_n) - (\epsilon+2\gamma)^2 P(E)$ porque $\sum\limits_{k=1}^{n} E(X_k^2) =$
$Var(S_n)$.
Por otro lado $\int\limits_{F_n} S_n^2 dP \leqslant \epsilon^2 P(F_n)$ porque $F_n = [|S_1| < \epsilon, \ldots, |S_n| < \epsilon]$.
Así $P(F_n)Var(S_n) - (\epsilon + 2\gamma)^2 P(E) \leqslant \int\limits_{F_n} S_n^2 dP \leqslant \epsilon^2 P(F_n)$,
pero $E = F_n^c$ entonces $P(E) = 1 - P(F_n)$.

De este modo $P(F_n)Var(S_n) - (\epsilon + 2\gamma)^2(1 - P(F_n)) \leqslant \epsilon^2 P(F_n)$, es decir,
$P(F_n)Var(S_n) \leqslant P(F_n)[Var(S_n) + (\epsilon + 2\gamma)^2 - \epsilon^2] \leqslant (\epsilon + 2\gamma)^2$
Asi $P(F_n) \leqslant \frac{(\epsilon+2\gamma)^2}{Var(S_n)}$ entonces $1 - P(E) \leqslant \frac{(\epsilon+2\gamma)^2}{Var(S_n)}$. Luego $P(E) \geqslant 1 - \frac{(\epsilon+2\gamma)^2}{Var(S_n)}$.

Por lo tanto $P[\max\limits_{1 \leqslant k \leqslant n} |S_k - E(S_k)| \geqslant \epsilon] \geqslant 1 - \frac{(\epsilon+2\gamma)^2}{Var(S_n)}$. $\square$

2.3. Criterios de Convergencia para Martingalas

Proposición 2.3.1. *Sea* $\{X_n : n \in \mathbb{N}\}$ *una sucesión de v.a.i. con* $Var(X_n) = \sigma_n^2 < \infty$ *para todo* $n \in \mathbb{N}$.
Si $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$ *entonces* $\sum_{n=1}^{\infty}(X_n - E(X_n))$ *converge casi seguramente.*

Demostración:
Sea $S_n = \sum_{k=1}^{n} X_k$ y consideremos el conjunto de m variables aleatorias $Y_1 = X_{n+1} - E(X_{n+1}), \ldots, Y_m = X_{n+m} - E(X_{n+m})$ a las que aplicamos para $\epsilon > 0$ el resultado de la proposición 2.2.9, esto es

$$P[\max_{1 \leqslant k \leqslant m} |\sum_{j=1}^{k} Y_j| \geqslant \epsilon] \leqslant \frac{1}{\epsilon^2} \sum_{k=1}^{m} Var(Y_k).$$

Pero $Y_1 + Y_2 + \cdots + Y_k = (X_{n+1} - E(X_{n+1})) + \cdots + (X_{n+k} - E(X_{n+k})) = (S_{n+k} - S_n) - E(S_{n+k} - S_n)$.
$Var(Y_k) = Var(X_{n+k} - E(X_{n+k})) = Var(X_{n+k}) = \sigma_{n+k}^2$.

Por lo tanto $P[\max_{1 \leqslant k \leqslant m} |(S_{n+k} - S_n) - E(S_{n+k}) + E(S_n)| \geqslant \epsilon] \leqslant \frac{1}{\epsilon^2} \sum_{k=1}^{m} \sigma_{n+k}^2$.
Definimos ahora $T_k = S_k - E(S_k)$, $\Delta_k = \sup_{r \geqslant 1} |T_{k+r} - T_k|$ y
$\Delta = \inf_{k \geqslant 1} \Delta_k = \inf_{k \geqslant 1} \sup_{r \geqslant 1} |T_{k+r} - T_k|$.

Entonces $\lim_{m \to \infty} P[\max_{1 \leqslant k \leqslant m} |(S_{n+k} - S_n) - E(S_{n+k}) + E(S_n)| \geqslant \epsilon] \leqslant \frac{1}{\epsilon^2} \sum_{k=1}^{\infty} \sigma_{n+k}^2$.
Notemos que $T_{k+r} - T_k = S_{k+r} - E(S_{k+r}) - S_k + E(S_k) = S_{k+r} - S_k - E(S_{k+r}) + E(S_k)$.

Luego $\Delta_n = \sup_{r \geqslant 1} |T_{n+r} - T_n| = \max_{r \geqslant 1} |S_{n+r} - S_n - E(S_{n+r}) + E(S_n)|$.
Por lo tanto $\lim_{m \to \infty} P[\max_{1 \leqslant k \leqslant m} |(S_{n+k} - S_n) - E(S_{n+k}) + E(S_n)| \geqslant \epsilon] = P[\Delta_n \geqslant \epsilon]$ porque se está calculando la probabilidad de la unión de eventos monótonos

crecientes que tienden a Δ_n.

De esta manera $P[\Delta_n \geqslant \epsilon] \leqslant \frac{1}{\epsilon^2} \sum_{k=n+1}^{\infty} \sigma_k^2$.

Ahora

$\Delta_n = \sup_{r\geqslant 1} |T_{n+r} - T_n|$, entonces $P[\limsup |T_{n+r} - T_n| \geqslant \epsilon] = P[\inf_{n\geqslant 1} \sup_{r\geqslant 1} |T_{n+r} - T_n| \geqslant \epsilon] = P[\inf_{n\geqslant 1} \Delta_n \geqslant \epsilon] = P[\Delta \geqslant \epsilon] \leqslant \frac{1}{\epsilon^2} \sum_{k=n+1}^{\infty} \sigma_k^2$ para todo $n \in \mathbb{N}$.

Y como $\sum_{k=1}^{\infty} \sigma_k^2$ se sigue que $P[\Delta \geqslant \epsilon] = 0$.

Por lo tanto $\{S_n - E(S_n) : n \in \mathbb{N}\}$ es de Cauchy casi seguramente.

Asi $\{S_n : n \in \mathbb{N}\}$ es de Cauchy casi seguramente.

Concluyendo que S_n converge c.s. $\square$

Proposición 2.3.2. *Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de v.a.i. tales que $|X_n| \leqslant \gamma$ c.s. con $\gamma > 0$ y $Var(X_n) = \sigma_n^2$ para todo $n \in \mathbb{N}$. Si $\sum_{n=1}^{\infty}(X_n - E(X_n))$ converge casi seguramente entonces $\sum_{n=1}^{\infty} \sigma_n^2 < \infty$.*

Demostración:

Usaremos la misma notación que en la proposición anterior.

Por la proposición 2.2.11 tenemos que $P[\Delta_n \geqslant \epsilon] \geqslant 1 - \dfrac{(\epsilon + 2\gamma)^2}{\sum_{k=n+1}^{\infty} \sigma_k^2}$ donde

$\Delta_n = \sup_{r\geqslant 1} |T_{n+r} - T_n|$ y $T_k = S_k - E(S_k)$.

Porque $P[\sup_{r\geqslant 1} |T_{n+r} - T_n| \geqslant \epsilon] = P[\sup_{r\geqslant 1} |S_{n+r} - E(S_{n+r}) - S_n + E(S_n)| \geqslant \epsilon] = P[\sup_{r\geqslant 1} |X_{n+1} + X_{n+2} + \cdots + X_{n+r} - E(X_{n+1} + \cdots + X_{n+r})| \geqslant \epsilon] \geqslant$

$$1 - \frac{(\epsilon + 2\gamma)^2}{Var(X_{n+1} + X_{n+2} + \cdots)} = 1 - \frac{(\epsilon + 2\gamma)^2}{\sum_{k=n+1}^{\infty} \sigma_k^2}.$$

Supongamos ahora que $\sum_{k=1}^{\infty} \sigma_k^2 = \infty$ entonces para cada $n \geqslant 1$ obtenemos

$P[\Delta_n \geqslant \epsilon] \geqslant 1 - \frac{(\epsilon+2\gamma)^2}{\infty} = 1$, por lo tanto $P(\Delta_n \geqslant \epsilon) = 1$.
Así
$\Delta = \inf_{k\geqslant 1} \Delta_k = \inf_{k\geqslant 1} \sup_{r\geqslant 1} |T_{k+r} - T_k| = \limsup |T_{k+r} - T_k|$ y $\lim_{n\to\infty} P(\Delta_n \geqslant \epsilon) = P(\Delta \geqslant \epsilon) = 1$.
Por lo tanto $\sum_{n=1}^{\infty}(X_n - E(X_n))$ diverge casi seguramente de acuerdo a la argumentación de la proposición 2.3.1. □

Observación. Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de v.a.i. con $|X_n| \leqslant \gamma$ c.s. Sea $Var(X_n) = \sigma_n^2$ entonces por las proposiciones 2.3.1 y 2.3.2 $\sum_{n=1}^{\infty}(X_n - E(X_n))$ converge o diverge casi seguramente de acuerdo a si $\sum_{n=1}^{\infty} \sigma_n^2$ converge o diverge respectivamente.

Proposición 2.3.3. *Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de variables aleatorias independientes tal que $|X_n| < \gamma$ c.s. para todo $n \in \mathbb{N}$. $\sum_{n=1}^{\infty} X_n$ converge casi seguramente si y sólo si se cumplen:*

i) $|\sum_{n=1}^{\infty} E(X_n)| < \infty$.

ii) $\sum_{n=1}^{\infty} Var(X_n) < \infty$.

Demostración:

Primero supongamos i) y ii).
Como $\sum_{n=1}^{\infty} \sigma_n^2$ por la proposición 2.3.1, $\sum_{n=1}^{\infty}(X_n - E(X_n))$ converge casi seguramente.
Esto significa que para casi todo $\omega \in \Omega$: $\sum_{n=1}^{\infty}(X_n(\omega) - E(X_n))$ tiene límte.
Asi $\sum_{n=1}^{\infty} X_n(\omega) - \sum_{n=1}^{\infty} E(X_n)$ converge, pero como $\sum_{n=1}^{\infty} E(X_n)$ converge entonces

$\sum_{n=1}^{\infty} X_n(\omega)$ tiene límte para casi todo $\omega \in \Omega$.

Por lo tanto $\sum_{n=1}^{\infty} X_n$ converge casi seguramente.

Ahora supongamos que $\sum_{n=1}^{\infty} X_n$ converge casi seguramente.

Introducimos una sucesión $\{X'_n : n \in \mathbb{N}\}$ de v.a.i. que cumple:

a) Para n fija, X_n y X'_n son idénticamente distribuidas.

b) La sucesión combinada $\{X_n, X'_n : n \in \mathbb{N}\}$ es una sucesión de variables aleatorias independientes.

Formamos la sucesión simetrizada $X^s_n = X_n - X'_n \quad n = 1, 2, \ldots$

$\{X^s_n : n \in \mathbb{N}\}$ es una sucesión de variables aleatorias independientes pues $X^s_n = X_n - X'_n$ es independiente de $X^s_m = X_m - X'_m$.

$|X^s_n| = |X_n - X'_n| \leqslant |X_n| + |X'_n| \leqslant \gamma + \gamma = 2\gamma$ c.s.

$E(X^s) = E(X_n - X'_n) = E(X_n) - E(X'_n) = 0.$

$Var(X^s_n) = Var(X_n - X'_n) = Var(X_n) + Var(X'_n) = 2Var(X_n).$

Luego $\sum_{n=1}^{\infty} X^s_n = \sum_{n=1}^{\infty}(X_n - X'_n)$ por lo que si $\sum_{n=1}^{\infty} X_n$ y $\sum_{n=1}^{\infty} X'_n$ convergen c.s. entonces $\sum_{n=1}^{\infty} X^s_n$ converge c.s.

Por la proposición 2.3.2 tenemos que $\sum_{n=1}^{\infty} Var(X^s_n) < \infty$ entonces $2\sum_{n=1}^{\infty} Var(X_n) < \infty$.Asi $\sum_{n=1}^{\infty} Var(X_n) < \infty$.

Por lo tanto ii) se cumple.

Ahora por la proposición 2.3.1 $\sum_{n=1}^{\infty}(X_n - E(X_n))$ converge casi seguramente y por tanto $\sum_{n=1}^{\infty} E(X_n) = \sum_{n=1}^{\infty} X_n - \sum_{n=1}^{\infty}(X_n - E(X_n))$ converge, por tanto i) también se cumple. $\square$

Proposición 2.3.4 (Criterio de las tres series). *Considerar una sucesión* $\{X_n : n \in \mathbb{N}\}$ *de variables aleatorias independientes.* $\sum_{n=1}^{\infty} X_n$ *converge casi seguramente si y sólo si para* $c > 0$ *fija cada una de las siguientes tres series converge:*

a) $\sum_{n=1}^{\infty} P(|X_n| \geqslant c)$.

b) $\sum_{n=1}^{\infty} E(X_n^c)$ *donde* $X_n^c = \begin{cases} X_n & si\ |X_n| < c \\ 0 & si\ |X_n| \geqslant c. \end{cases}$

c) $\sum_{n=1}^{\infty} Var(X_n^c)$.

esto es, $\{X_n^c : n \in \mathbb{N}\}$ *es la sucesión de las variables aleatorias truncadas. Además si las series de a), b), o c) convergen para alguna* $c > 0$, *convergen para todo* $c > 0$.

Demostración:

Supongamos que $\sum_{n=1}^{\infty} X_n$ converge c.s. entonces $X_n \xrightarrow[n\to 0]{c.s.} 0$.

Esto significa que $P[\lim_{n\to\infty} X_n = 0] = 1$ y que para cualquier $c > 0$

$P[\limsup |X_n| \geqslant c] = 0$.

Esto nos lleva a que $P[|X_n| \geqslant c \quad i.o.] = 0$.

El lema de Borel-Cantelli dice: "$\{E_n : n \in \mathbb{N}\}$ sucesión de eventos entonces:

i) Si $\sum_{n=1}^{\infty} P(E_n) < \infty$ entonces $P(\limsup E_n) = 0$.

ii) Si además los eventos son independientes entonces $P(\limsup E_n) = 0$ ó 1 de acuerdo a si $\sum_{n=1}^{\infty} P(E_n)$ converge o diverge casi seguramente respectivamente'.'

Y un corolario de Borel-Cantelli establece que:
"$P(\limsup E_n) = 0$ si y sólo si $\sum_{n=1}^{\infty} P(E_n) < \infty$ cuando E_n son independientes" entonces en nuestro caso tenemos $P[\limsup |X_n| \geqslant c] = 0$.
Por lo tanto $\sum_{n=1}^{\infty} P[|X_n| \geqslant c] < \infty$ es decir a) se cumple.
Ahora por la proposición 2.2.8. $\sum_{n=1}^{\infty} X_n$ y $\sum_{n=1}^{\infty} X_n^c$ son equivalentemente convergentes porque $P[X_n \neq X_n^c] = P[|X_n| \geqslant c]$.
Como $\sum_{n=1}^{\infty} P[X_n \neq X_n^c] = \sum_{n=1}^{\infty} P[|X_n| \geqslant c] < \infty$ entonces por b) de la proposición 2.2.8 $\sum_{n=1}^{\infty} X_n$ y $\sum_{n=1}^{\infty} X_n^c$ son equivalentemente convergentes.
Ahora por la proposición 2.3.3 como $|X_n^c| \leqslant c$ tenemos que
$\sum_{n=1}^{\infty} X_n^c$ converge c.s. si y sólo si *i)* $\sum_{n=1}^{\infty} E(X_n^c)$ converge y *ii)* $\sum_{n=1}^{\infty} Var(X_n^c) < \infty$
Por lo tanto b) y c) también se cumplen.
Ahora supongamos que $\sum_{n=1}^{\infty} P[|X_n| \geqslant c]$ converge entonces como decíamos antes, $\sum_{n=1}^{\infty} X_n$ y $\sum_{n=1}^{\infty} X_n^c$ son equivalentemente convergentes.
Si además
$\sum_{n=1}^{\infty} E(X_n^c) < \infty$ y $\sum_{n=1}^{\infty} Var(X_n^c) < \infty$ tenemos por la proposición 2.3.3 que $\sum_{n=1}^{\infty} X_n^c$ converge casi seguramente y por tanto $\sum_{n=1}^{\infty} X_n$ converge casi seguramente.

Además la c que hace que se cumpla este resultado es cualquiera mayor que 0. $\square$

Lema 2.3.5 (Toeplitz). *Sea* $\{a_n : n \in \mathbb{N}\}$ *una sucesión de números reales tales que* $a_n \xrightarrow[n\to\infty]{} a$. *Entonces* $\frac{S_n}{n} = \frac{a_1 + \ldots + a_n}{n} \xrightarrow[n\to\infty]{} a$

Demostración. Sea $\epsilon > 0$ y como $a_n \xrightarrow[n\to\infty]{} a$ existe $n_0(\epsilon) = n_0$ tal que para

todo $n \geqslant n_0$ se cumple que $|a_n - a| < \frac{\epsilon}{2}$.
Luego elegimos $n_0^* > n_0$ tal que $\frac{1}{n_0^*}\sum\limits_{k=1}^{n_0}|a_k - a| < \frac{\epsilon}{2}$ esto es posible porque n_0^* es lo suficientemente grande para que al dividir una cantidad fija por n_0^* se haga pequeña.
Ahora tomemos n grande que cumpla que $n > n_0^* > n_0$ entonces ambas cosas se cumplen

$$|a_n - a| < \tfrac{\epsilon}{2} \text{ y } \tfrac{1}{n_0^*}\sum_{k=1}^{n_0}|a_k - a| < \tfrac{\epsilon}{2}, \text{ luego}$$

$$\left|\frac{1}{n}\sum_{k=1}^{n} a_k - a\right| = \left|\frac{1}{n}\sum_{k=1}^{n_0} a_k + \frac{1}{n}\sum_{k=n_0+1}^{n} a_k - a\right| = \left|\frac{1}{n}\sum_{k=1}^{n_0}(a_k - a) + \frac{1}{n}\sum_{k=n_0+1}^{n}(a_k - a)\right|$$

por la desigualdad del triángulo

$|\frac{S_n}{n} - a| \leqslant \frac{1}{n}\sum\limits_{k=1}^{n_0}|a_k - a| + \frac{1}{n}\sum\limits_{k=n_0+1}^{n}|a_k - a| \leqslant \frac{1}{n_0^*}\sum\limits_{k=1}^{n_0}|a_k - a| + \frac{1}{n}\sum\limits_{k=n_0+1}^{n}|a_k - a|$
porque $\frac{1}{n} < \frac{1}{n_0^*}$.

Así $|\frac{S_n}{n} - a| \leqslant \frac{\epsilon}{2} + \frac{1}{n}(n - n_0)\frac{\epsilon}{2} = \frac{\epsilon}{2} + \frac{\epsilon}{2} - \frac{n_0}{n}\frac{\epsilon}{2} < \epsilon$ porque $\frac{n_0}{n} < 1$.

Por lo tanto $\dfrac{S_n}{n} \xrightarrow[n\to\infty]{} a$. $\square$

Lema 2.3.6 (Kronecker). *Sea $\{a_n : n \in \mathbb{N}\}$ una sucesión de números reales tal que $\sum\limits_{n=1}^{\infty} a_n$ converge, entonces $\sum\limits_{k=1}^{\infty}\dfrac{ka_k}{n} \xrightarrow[n\to\infty]{} 0$.*

Demostración:
Sabemos que $S_n = a_1 + \cdots + a_n$ con $S_0 = 0$ por lo que $\dfrac{1}{n}\sum\limits_{k=1}^{n} ka_k =$ $\dfrac{1}{n}\sum\limits_{k=1}^{n} k(S_k - S_{k-1}) = \dfrac{1}{n}[(S_1 - S_0) + 2(S_2 - S_1) + \cdots + (n-1)(S_{n-1} - S_{n-2}) +$ $n(S_n - S_{n-1})] = \dfrac{1}{n}[-S_1 - S_2 - \cdots - S_{n-1} + nS_n] = S_n - \dfrac{1}{n}\sum\limits_{k=1}^{n-1} S_k$.
La sucesión $\{S_n : n \in \mathbb{N}\}$ converge a un número $s \in \mathbb{R}$ y por el lema 2.3.5

$\frac{1}{n-1}\sum_{k=1}^{n-1} S_k \xrightarrow[n\to\infty]{} s.$

Pero también $\frac{n}{n-1}\frac{1}{n}\sum_{k=1}^{n-1} S_k \xrightarrow[n\to\infty]{} s$ por lo que $\frac{1}{n}\sum_{k=1}^{n-1} S_k \xrightarrow[n\to\infty]{} s.$

De este modo $S_n - \frac{1}{n}\sum_{k=1}^{n-1} S_k \xrightarrow[n\to\infty]{} s - s \;=\; 0.$

Por lo tanto $\sum_{k=1}^{\infty}\frac{ka_k}{n} \xrightarrow[n\to\infty]{} 0.$ □

Proposición 2.3.7. *Sea* $\{X_n : n \in \mathbb{N}\}$ *una sucesión de v.a.i. con* $Var(X_n) = \sigma_n^2 < \infty \quad n = 1, 2, \ldots$
Supongamos que $\sum_{n=1}^{\infty}\frac{\sigma_n^2}{n^2} < \infty$ *entonces* $\frac{S_n - E(S_n)}{n} \xrightarrow[n\to 0]{c.s.} 0$ *donde* $S_n = \sum_{k=1}^{n} X_k.$

Demostración:

Definimos $Y_n = \frac{X_n - E(X_n)}{n} \quad n = 1, 2, \ldots.$

Entonces $E(Y_n) = \; 0$ y $Var(Y_n) = \frac{1}{n^2}\, Var(X_n) = \frac{\sigma_n^2}{n^2}.$

La sucesión $\{Y_n : n = 1, 2, \ldots\}$ es de v.a.i., además $\sum_{n=1}^{\infty} Var(Y_n) = \sum_{n=1}^{\infty}\frac{\sigma_n^2}{n^2} < \infty$ por hipótesis.

Se sigue de la proposición 2.3.1 que $\sum_{n=1}^{\infty}(Y_n - E(Y_n))$ converge casi seguramente pero $E(Y_n) = 0$.

Por lo tanto $\sum_{n=1}^{\infty}(Y_n - E(Y_n)) \;=\; \sum_{n=1}^{\infty} Y_n$ converge c.s. y aplicando el Lema 2.3.6 tenemos: $\frac{1}{n}\sum_{k=1}^{n} kY_k \xrightarrow[n\to 0]{c.s.} 0$, pero $Y_k = \frac{X_k - E(X_k)}{k}$ entonces

$\frac{1}{n}\sum_{k=1}^{n} k\left(\frac{X_k - E(X_k)}{k}\right) = \frac{1}{n}\sum_{k=1}^{n}(X_k - E(X_k)) = \frac{1}{n}[S_n - E(S_n)] \xrightarrow[n\to 0]{c.s.} 0.$ □

Observación. Si $\{X_n : n \in \mathbb{N}\}$ es una sucesión de v.a. y $\{A_n : n \in \mathbb{N}\}$ es una sucesión de constantes, una afirmación del tipo $\frac{S_n - A_n}{n} \xrightarrow[n\to 0]{\text{c.s.}} 0$ es conocida como Ley Fuerte de los Grandes Números, a veces se establece con sucesiones $\{A_n : n \in \mathbb{N}\}$ y $\{B_n : n \in \mathbb{N}\}$ donde $B_n \xrightarrow[n\to\infty]{} \infty$ tal que $\frac{S_n - A_n}{B_n} \xrightarrow[n\to 0]{\text{c.s.}} 0$.

Teorema 2.3.8 (Kolmogorov). *Sea $\{X_n : n \in \mathbb{N}\}$ sucesión de variables aleatorias independientes e idénticamente distribuidas. $S_n = \sum_{k=i}^{n} X_k$ $n = 1, 2, \ldots$ entonces $\frac{S_n}{n} \xrightarrow[n\to\infty]{\text{c.s.}} \alpha \in \mathbb{R}$ si y sólo si $E(|X_n|) < \infty$. Además en este caso $E(X_n) = \alpha$.*

Demostración:

Supongamos que $E(|X_1|) < \infty$ (las v.a. son i.d.).

Mostraremos que $\frac{S_n}{n} \xrightarrow[n\to 0]{\text{c.s.}} E(X_1)$.

Introducimos una v.a. X idénticamente distribuida con X_1.

Sean $E_0 = \Omega$,

$E_n = [|X| \geqslant n]$ $n = 1, 2, \ldots$

$E_{n+1} \subseteq E_n$ porque si $|X| \geqslant n+1$ entonces $|X| \geqslant n$.

Asi $E_n \downarrow$, el límite es el $\varnothing$ porque $\bigcap_{n=1}^{\infty} E_n = \bigcap_{n=1}^{\infty} [|X| \geqslant n] = \varnothing$.

Más aún, $E_k - E_{k+1} = [|X| \geqslant k, |X| < k+1] = [k \leqslant |X| < k+1]$.

Por lo que $\bigcup_{k=n}^{\infty} (E_k - E_{k+1}) = \bigcup_{k=n}^{\infty} [k \leqslant |X| < k+1] = [|X| \geqslant n] = E_n$.

Es decir E_n ha sido representado como la unión de eventos ajenos, por lo que $P(E_n) = P[\bigcup_{k=n}^{\infty} (E_k - E_{k+1})] = \sum_{k=n}^{\infty} P(E_k - E_{k+1})$.

Por lo tanto $\sum_{n=1}^{\infty} P(E_n) = \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} P(E_k - E_{k+1})$.

Las sumas dobles se trabajan como integrales dobles, en esta suma tenemos

$1 \leqslant n \leqslant k < \infty$ tomando valores enteros.

Asi que $\sum\limits_{n=1}^{\infty} P(E_n) = \sum\limits_{k=1}^{\infty}\sum\limits_{n=1}^{k} P(E_k - E_{k+1}) = \sum\limits_{k=1}^{\infty} kP(E_k - E_{k+1})$.

Luego $\sum\limits_{n=1}^{\infty} P(E_n) = \sum\limits_{n=1}^{\infty} nP(E_n - E_{n+1})$.

Por otro lado $\sum\limits_{n=0}^{\infty} P(E_n) = P(E_0) + \sum\limits_{n=1}^{\infty} P(E_n) = P(\Omega) + \sum\limits_{n=1}^{\infty} P(E_n) = 1 + \sum\limits_{n=1}^{\infty} P(E_n)$.

También $\sum\limits_{n=0}^{\infty} P(E_n) = \sum\limits_{n=0}^{\infty}\sum\limits_{k=n}^{\infty} P(E_k - E_{k+1}) = \sum\limits_{k=0}^{\infty}\sum\limits_{n=0}^{k} P(E_k - E_{k+1}) = \sum\limits_{k=0}^{\infty}(k+1)P(E_k - E_{k+1}) = \sum\limits_{n=1}^{\infty} nP(E_{n-1} - E_n)$ si $n = k+1$.

Cualquier $\omega \in \Omega$ con $X(\omega) \geqslant 1$ pertenece a uno y sólo uno de los conjuntos $[n \leqslant |X| < n+1] = E_n - E_{n+1}$ entonces $I_{[E_n - E_{n+1}]}(\omega) = 1$.

Pero $nI_{[E_n - E_{n+1}]}(\omega) = n \leqslant |X|$ luego $\sum\limits_{n=1}^{\infty} nI_{[E_n - E_{n+1}]} \leqslant |X|$.

Del mismo modo cualquier $\omega \in \Omega$ pertenece a exactamente uno de los conjuntos $[n-1 \leqslant |X| < n] = E_{n-1} - E_n$

entonces $I_{[E_{n-1} - E_n]}(\omega) = 1$ asi $nI_{[E_{n-1} - E_n]}(\omega) = n > |X|$ implica que $\sum\limits_{n=1}^{\infty} nI_{[E_{n-1} - E_n]} > |X|$.

Por lo tanto $\sum\limits_{n=1}^{\infty} nI_{[E_n - E_{n+1}]} \leqslant |X| < \sum\limits_{n=1}^{\infty} nI_{[E_{n-1} - E_n]}$.

Tomando esperanza

$\sum\limits_{n=1}^{\infty} nP(E_n - E_{n+1}) \leqslant E(|X|) < \sum\limits_{n=1}^{\infty} nP(E_{n-1} - E_n)$.

Luego $\sum\limits_{n=1}^{\infty} P(E_n) \leqslant E(|X|) < 1 + \sum\limits_{n=1}^{\infty} P(E_n)$, y como $E(|X|) < \infty$ entonces $\sum_{n=1}^{\infty} P(E_n) < \infty$.

Ahora para cada $n \in \mathbb{N}$ sea $X_n^* = X_n^n = \begin{cases} X_n & \text{si } |X_n| < n \\ 0 & \text{si } |X_n| \geqslant n \end{cases}$ la truncación en n.

Sabemos que $Var(X_n^*) = E(X_n^{*2}) - E^2(X_n^*) \leqslant E(X_n^{*2}) = \int_\Omega X_n^{*2} dP = \int_{[|X|<n]} X_n^2 dP + \int_{[|X|\geqslant n]} 0^2 dP = \int_{[|X|<n]} X_n^2 dP = \int_{E_n^c} X^2 dP$ porque $E_n = [|X| \geqslant n]$.

Luego $Var(X_n^*) = \int_{E_n^c} X^2 dP \leqslant \sum_{k=1}^{n} k^2 P(E_{k-1} - E_k)$ porque

$X^2 I_{E_n^c} \leqslant \sum_{k=1}^{n} k^2 I_{[E_{k-1}-E_k]}$.

Para mostrar esta afirmación notamos que $[E_{k-1} - E_k] = [k-1 \leqslant |X| < k]$ y que $E_n^c = [|X| < n] = [E_0 - E_1] \cup [E_1 - E_2] \cup \cdots \cup [E_{n-1} - E_n] = \bigcup_{k=1}^{n} (E_{k-1} - E_k)$.

Asi $I_{E_n^c} = I_{[E_0-E_1]} + I_{[E_1-E_2]} + \cdots + I_{[E_{n-1}-E_n]}$ por lo que si $\omega \in E_n^c$ entonces existe una única $1 \leqslant k \leqslant n$ tal que $\omega \in [E_{k-1} - E_k] = [k-1 \leqslant |X| < k] = [(k-1)^2 \leqslant |X|^2 < k^2]$.

Pero $X^2(\omega) I_{E_n^c}(\omega) \leqslant k^2 I_{[E_{k-1}-E_k]}$ asi $X^2 I_{E_n^c}(\omega) \leqslant \sum_{k=1}^{n} k^2 I_{[E_{k-1}-E_k]}$.

Tenemos que

$Var(X_n^*) \leqslant \sum_{k=1}^{n} k^2 P(E_{k-1} - E_k)$ y $\sum_{n=1}^{\infty} \frac{Var(X_n^*)}{n^2} \leqslant \sum_{n=1}^{\infty} \sum_{k=1}^{n} \frac{k^2}{n^2} P(E_{k-1} - E_k) = \sum_{k=1}^{\infty} \sum_{n=k}^{\infty} \frac{k^2}{n^2} P(E_{k-1} - E_k) = \sum_{k=1}^{\infty} k^2 P(E_{k-1} - E_k) \sum_{n=k}^{\infty} \frac{1}{n^2}$.

Por otro lado $\sum_{n=k}^{\infty} \frac{1}{n^2} = \frac{1}{k^2} + \frac{1}{(k+1)^2} + \frac{1}{(k+2)^2} + \cdots \leqslant \frac{1}{k^2} + \int_k^\infty \frac{1}{x^2} dx = \frac{1}{k^2} - \frac{1}{x}\Big|_k^\infty = \frac{1}{k^2} + \frac{1}{k} \leqslant \frac{1}{k} + \frac{1}{k} = \frac{2}{k}$.

Por lo tanto $\sum_{n=1}^{\infty} \frac{Var(X_n^*)}{n^2} \leqslant \sum_{k=1}^{\infty} k^2 P(E_{k-1}-E_k)\frac{2}{k} = 2\sum_{k=1}^{\infty} kP(E_{k-1}-E_k) = 2\left[1+\sum_{n=1}^{\infty} P(E_n)\right] < \infty.$

Ahora sea $S_n^* = \sum_{k=1}^{n} X_n^*$ $n = 1, 2, \ldots$ por la proposición 2.3.7 aplicada a la sucesión $\{X_n^* : n \in \mathbb{N}\}$ se obtiene que $\frac{(S_n^* - E(S_n^*))}{n} \xrightarrow[n\to 0]{\text{c.s.}} 0$, se verá entonces que $\frac{1}{n}E(S_n^*) \xrightarrow[n\to\infty]{} E(X)$.

Sea n fija, $E(X_n^*) = \int\limits_{\Omega} X_n^* dP = \int\limits_{[|X_n|<n]} X_n + \int\limits_{[|X|\geqslant n]} 0dP = \int\limits_{[|X|<n]} X_n dP = \int\limits_{E_n^c} X_n dP = \int\limits_{E_n^c} XP = E[XI_{E_n^c}]$ pero $E_n^c = [|X| < n] \xrightarrow[n\to\infty]{} \Omega$, por lo que $XI_{E_n} \xrightarrow{\text{c.s.}} XI_{\Omega} = X$ y por el Teorema de Convergencia Dominada $E(X_n^*) \longrightarrow E(X)$ (notar que $|XI_{E_n^c}| \leqslant |X|$.)

Aplicando el Lema 2.3.5:

$\frac{1}{n}[E(X_1^*) + \cdots + E(X_n^*)] = \frac{1}{n}E(S_n^*) \xrightarrow[n\to\infty]{} E(X)$.

Por último se sabe que $\{X_n : n \in \mathbb{N}\}$ y $\{X_n^* : n \in \mathbb{N}\}$ tienen colas equivalentes porque para casi todo $\omega \in \Omega$ existe $n(\omega)$ tal que para todo $n \geqslant n(\omega)$ tenemos

$$X_n(\omega) = X_n^*(\omega) = \begin{cases} X_n(\omega) & \text{si } |X_n(\omega)| < n \\ 0 & \text{si } |X_n(\omega)| \geqslant n. \end{cases}$$

Por c) de la proposición 2.2.8 sabemos que $\frac{1}{n}\sum_{k=1}^{n} X_k$ y $\frac{1}{n}\sum_{k=1}^{n} X_k^*$ convergen al mismo límite y sobre el mismo conjunto casi seguramente.

Por lo tanto $\frac{1}{n}\left(\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} X_k^*\right) \xrightarrow[n\to\infty]{\text{c.s.}} 0$.

De este modo $\frac{1}{n}(S_n - S_n^*) \xrightarrow{\text{c.s.}} 0$ y como $\frac{1}{n}(S_n^* - E(S_n^*)) \xrightarrow[n\to\infty]{\text{c.s.}} 0$ se obtiene que

$\frac{S_n}{n} - \frac{E(S_n^*)}{n} \xrightarrow{c.s.} 0$ pero $\frac{E(S_n^*)}{n} \longrightarrow E(X)$.
Asi $\frac{S_n}{n} \xrightarrow{c.s.} E(X)$.

Ahora supongamos que $\frac{S_n}{n} \xrightarrow{c.s.} \alpha \in \mathbb{R}$, veremos que $E(|X|) < \infty$ y además $E(X) = \alpha$.
Notemos que $\frac{X_n}{n} = \frac{S_n - S_{n-1}}{n} = \frac{S_n}{n} - \frac{S_{n-1}}{n} = \frac{S_n}{n} - \frac{n-1}{n}\frac{S_{n-1}}{n-1} \xrightarrow[n\to\infty]{c.s.} 0$.
Por un Corolario del Lema de Borel-Cantelli que dice
"$X_n \xrightarrow[n\to\infty]{c.s.} 0$ si y sólo si $\sum_{n=1}^{\infty} P[|X_n| \geqslant \epsilon] < \infty$ para todo $\epsilon > 0$ si $\{X_n : n \in \mathbb{N}\}$ son v.a.i."
tenemos que $\sum_{n=1}^{\infty} P[|\frac{X_n}{n}| \geqslant \epsilon] = \sum_{n=1}^{\infty} P[|X_n| \geqslant n\epsilon] < \infty$ para todo $\epsilon > 0$.
En particular si $\epsilon = 1$, $\sum_{n=1}^{\infty} P[|X_n| \geqslant n] < \infty$, es decir, $\sum_{n=1}^{\infty} P(E_n) < \infty$.
Por lo tanto $E(|X|) < \infty$, luego $\alpha = E(X)$ porque si $\frac{S_n}{n} \xrightarrow[n\to\infty]{c.s.} \alpha$ y $\frac{S_n}{n} \xrightarrow{c.s.} E(X)$ entonces $\alpha = E(X)$. □

El resultado anterior es una versión de la ley fuerte de los grandes números, es el caso en que se tiene independencia e identidad en la distribución de las variables aleatorias.

X está centrada en su esperanza condicional si $E(X|\mathcal{G}) = 0$.

Aqui $\mathcal{G}$ es sub$\sigma-$álgebra de $\mathcal{F}$.

En las aplicaciones tomaremos $X = X_n$ y $\mathcal{G} = \sigma(X_1, \ldots, X_{n-1})$ con $n \geqslant 2$.

Si $\{X_n : n \in \mathbb{N}\}$ es una sucesión de v.a. con $E(|X_n|) < \infty$ para todo $n \geqslant 1$, $\mu_1 = E(X_1)$ c.s. y para $n \geqslant 2$ $\mu_n = E(X_n|X_1, \ldots, X_{n-1})$ c.s. entonces $E(\mu_n) = E[E(X_n|X_1, \ldots, X_{n-1})] = E(X_n)$ para todo $n \in \mathbb{N}$.

$\{\mu_n, \mathcal{F}_n = \sigma(X_1, \ldots, X_n) : n \in \mathbb{N}\}$ no es martingala ya que $E(\mu_n|\mathcal{F}_{n-1}) =$

$E(\mu_n|X_1,\ldots,X_{n-1}) = E[E(X_n|X_1,\ldots,X_{n-1})\Big|X_1,\ldots,X_{n-1}]$
$= E(X_n|X_1,\ldots,X_{n-1}) = \mu_n \neq \mu_{n-1}$.
Ahora sea $m < n$ y como $(X_m - \mu_m)$ es $\sigma(X_1,\ldots,X_{n-1})$ medible:
$E[(X_m-\mu_m)(X_n-\mu_n)|X_1,\ldots,X_{n-1}] = (X_m-\mu_m)E[X_n-\mu_n|X_1,\ldots,X_{n-1}] = (X_m - \mu_m)\{E(X_n|X_1,\ldots,X_{n-1}) - E(\mu_n|X_1,\ldots,X_{n-1})\} = (X_m - \mu_m)[\mu_n - \mu_n] = 0$.
Por lo tanto $(X_m - \mu_m)$ y $(X_n - \mu_n)$ son independientes dado
$\mathcal{F}_{n-1} = \sigma(X_1,\ldots,X_{n-1})$
Por otro lado:
$E\left[\left(\sum_{k=1}^{n}(X_k - \mu_k)\right)^2\right] = E\Big\{[(X_1-\mu_1)+\cdots+(X_n-\mu_n)]^2\Big\} =$
$E\left[\sum_{k=1}^{n}(X_k-\mu_k)^2 + 2\sum_{1\leqslant k<j\leqslant n}(X_k-\mu_k)(X_j-\mu_j)\right] = \sum_{k=1}^{n}E\left[(X_k-\mu_k)^2\right] +$
$2\sum_{1\leqslant k<j\leqslant n}E\left[(X_k-\mu_k)(X_j-\mu_j)\right] = \sum_{k=1}^{n}E\left[(X_k-\mu_k)^2\right] +$
$2\sum_{1\leqslant k<j\leqslant n}E\left[E\left[(X_k-\mu_k)(X_j-\mu_j)\Big|X_1,\ldots,X_{j-1}\right]\right] = \sum_{k=1}^{n}E[(X_k-\mu_k)^2] +$
$0 = \sum_{k=1}^{n}E[(X_k-\mu_k)^2]$. Así si las variables aleatorias X_n están centradas en μ_n y $S_n = \sum_{k=1}^{n}X_k$ entonces $Var(S_n) = \sum_{k=1}^{n}Var(X_k)$.
A continuación estudiaremos algunos resultados sobre la convergencia de Martingalas.

Teorema 2.3.9. *Sea $\{S_n : n \in \mathbb{N}\}$ una martingala con $E(S_n^2) < c < \infty$ para todo $n \geqslant 1$, entonces existe una v.a. S tal que $S_n \xrightarrow[n\to\infty]{c.s.} S$ y $S_n \xrightarrow[n\to\infty]{L_2} S$, además $E(S_n) = E(S)$ para todo n.*

Demostración:

Como $\{S_n : n \in \mathbb{N}\}$ es martingala, se cumple que para todo $n > m$ $E(S_n|S_m) = S_m$ porque $\sigma(S_m) \subseteq \sigma(S_1, \ldots, S_m) = \mathcal{F}_m$.

Entonces $E[E(S_n|\mathcal{F}_m)|S_m] = E(S_n|S_m)$.

Pero por otro lado $E[E(S_n|\mathcal{F}_m)|S_m] =$

$E(S_m|S_m) = S_m$.

Por lo tanto $E(S_n|S_m) = S_m$ luego $E[E(S_n|S_m|)] = E(S_m)$.

Asi $E(S_n) = E(S_m)$ =constante.

También para $n > m$ $E(S_mS_n|S_m) = S_mE(S_n|S_m) = S_m^2$,

entonces $E[E(S_mS_n|S_m)] = E(S_m^2)$ de aquí $E(S_mS_n) = E(S_m^2)$.

De nuevo para todo $n > m$ $E[(S_n - S_m)^2] = E(S_n^2) + E(S_m^2) - 2E(S_nS_m) = E(S_n^2) + E(S_m^2) - 2E(S_m^2) = E(S_n^2) - E(S_m^2)$.

Por la proposición 2.2.2 $\{S_n^2 : n \in \mathbb{N}\}$ es submartingala.

Como $E(S_n^2) < c < \infty$ es una sucesión acotada y creciente de números reales ya que $E(S_{n+1}^2|S_n^2) \geqslant S_n^2$ entonces $E(S_n^2) \leqslant E(S_{n+1}^2) < c$, por tanto converge.

Ahora $\lim\limits_{n,m\to\infty} E[(S_n - S_m)^2] = \lim\limits_{\substack{n,m\to\infty \\ n>m}} [E(S_n^2) - E(S_m^2)] = L - L = 0$ donde $L = \lim\limits_{n\to\infty} E(S_n^2)$.

Sea $\epsilon > 0$ y definimos $D(\epsilon) = \bigcap\limits_{m=1}^{\infty} \bigcup\limits_{n=1}^{\infty} \Big[|S_{n+m} - S_m| \geqslant \epsilon\Big]$, asi

$\omega \in D(\epsilon)$ si y sólo si $\omega \in \bigcap\limits_{m=1}^{\infty} \bigcup\limits_{n=1}^{\infty} \Big[|S_{n+m} - S_m| \geqslant \epsilon\Big]$ si y sólo si

$\omega \in \bigcup\limits_{n=1}^{\infty} \Big[|S_{n+m} - S_m| \geqslant \epsilon\Big]$ para todo $m \in \mathbb{N}$ si y sólo si para todo $m \in \mathbb{N}$ y alguna $n \in \mathbb{N}$ $\Big|S_{n+m}(\omega) - S_m(\omega)\Big| \geqslant \epsilon$.

Si hacemos ahora $D = \bigcup\limits_{\epsilon>0} D(\epsilon)$ entonces D es el conjunto donde la sucesión

$\{S_n : n \in \mathbb{N}\}$ no es de Cauchy.

Veremos que $P(D) = 0$.

Si $0 < \epsilon_1 < \epsilon_2$, $\omega \in D(\epsilon_2)$ entonces para todo $m \geqslant 1$ existe $n \geqslant 1$ tal que $|S_{n+m}(\omega) - S_m(\omega)| \geqslant \epsilon_2 > \epsilon_1$ de aqui que $\omega \in D(\epsilon_1)$, por lo tanto $D(\epsilon_2) \subseteq D(\epsilon_1)$.

Por lo que bastaría ver que $P[D(\epsilon)] = 0$ para todo $\epsilon > 0$.

Sea $D_m(\epsilon) = \bigcup\limits_{n=1}^{\infty} \Big[|S_{n+m} - S_m| \geqslant \epsilon\Big]$ entonces $D(\epsilon) = \bigcap\limits_{m=1}^{\infty} D_m(\epsilon)$.

Tenemos que mostrar que $\lim\limits_{m\to\infty} P[D_m(\epsilon)] = 0$.

Sea $D_{m,n}(\epsilon) = \Big[\max\limits_{1\leqslant k\leqslant n} |S_{n+k} - S_m| \geqslant \epsilon\Big] = \bigcup\limits_{k=1}^{n} \Big[|S_{n+k} - S_m| \geqslant \epsilon\Big] \xrightarrow[n\to\infty]{} D_m(\epsilon) = \bigcup\limits_{n=1}^{\infty} \Big[|S_{n+m} - S_m| \geqslant \epsilon\Big]$

$P[D_{m,n}(\epsilon)] = P\Big[\max\limits_{1\leqslant k\leqslant n} |S_{n+k} - S_m| \geqslant \epsilon\Big] = P\Big[\bigcup\limits_{k=1}^{n}[|S_{n+k} - S_m| \geqslant \epsilon]\Big]$

Sea $W_k = S_{m+k} - S_m = (S_{m+k} - S_{m+k-1}) + \cdots + (S_{m+1} - S_m) = \sum\limits_{j=1}^{k}(S_{m+j} - S_{m+j-1})$.

Por lo tanto $P[D_{m,n}(\epsilon)] = P[\max\limits_{1\leqslant k\leqslant n} |W_k| \geqslant \epsilon] \leqslant \frac{1}{\epsilon^2}\sum\limits_{j=1}^{n} E\Big[(S_{m+j} - S_{m+j-1})^2\Big]$

pero $E(S_m|S_1, \ldots, S_{m-1}) = S_{m-1}$ por ser martingala.

Luego $S_m - S_{m-1} = S_m - E(S_m|S_1, \ldots, S_{m-1})$ y además $E[(S_m - S_{m-1})^2] = E(S_m^2) - E(S_{m-1}^2) < \infty$.

Se tiene entonces que $P[D_{m,n}(\epsilon)] \leqslant \frac{1}{\epsilon^2}\sum\limits_{j=1}^{n} E\Big[(S_{m+j} - S_{m+j-1})^2\Big] =$

$\frac{1}{\epsilon^2}\sum\limits_{j=1}^{n}\Big[E(S_{m+j}^2) - E(S_{m+j-1}^2)\Big] = \frac{1}{\epsilon^2}\Big[E(S_{m+n}^2) - E(S_m^2)\Big]$.

Por lo tanto $P[D_m(\epsilon)] = \lim\limits_{n\to\infty} P[D_{m,n}(\epsilon)] \leqslant \frac{1}{\epsilon^2}\lim\limits_{n\to\infty}\Big[E(S_{m+n}^2) - E(S_m^2)\Big]$.

De aquí $P[D(\epsilon)] = \lim\limits_{m\to\infty} P[D_m(\epsilon)] \leqslant \lim\limits_{m\to\infty}\lim\limits_{n\to\infty} \frac{1}{\epsilon^2}\Big[E(S^2_{m+n}) - E(S^2_m)\Big] = 0.$

Asi $P(D) = 0$ y $\{S_n(\omega) : n \in \mathbb{N}\}$ es de Cauchy c.s.

De este modo $\{S_n(\omega : n \in \mathbb{N})\}$ converge.

Consideramos puntualmente el límite de S_n, esto es, $S(\omega) = \lim\limits_{n\to\infty} S_n(\omega)$ que es v.a.

Más aún, $S_n \xrightarrow[n\to\infty]{c.s.} S$.

Veamos que $E(S^2) < \infty$ y que $S_n \xrightarrow{L^2} S$.

Sabemos que $S_n \xrightarrow[n\to\infty]{c.s.} S$ entonces $S_n^2 \xrightarrow[n\to\infty]{c.s.} S^2$ y además $\{E(S_n^2) : n \in \mathbb{N}\}$ es una sucesión creciente y acotada de números reales, entonces por el Lema de Fatou:

$E(S^2) = \int\limits_\Omega S^2 dP = \int\limits_\Omega \liminf S_n^2 dP \leqslant \liminf \int\limits_\Omega S_n^2 dP = \liminf E(S_n^2) < \infty.$

Asi $E(S^2) < \infty$.

Por otro lado $E[(S_n - S)^2] = \int\limits_\Omega (S_n - S)^2 dP = \int\limits_\Omega \liminf\limits_{m\to\infty} (S_n - S_m)^2 dP \leqslant \liminf\limits_{m\to\infty} \int\limits_\Omega (S_n - S_m)^2 dP = \lim\limits_{m\to\infty} E[(S_n - S_m)^2]$ para todo $n \geqslant 1$.

Por lo tanto $\lim\limits_{n\to\infty} E[(S_n - S)^2] \leqslant \lim\limits_{n\to\infty}\lim\limits_{m\to\infty} E[(S_n - S_m)^2] = \lim\limits_{n\to\infty}\lim\limits_{m\to\infty} \Big\{E(S_n^2) - E(S_m^2)\Big\} = 0.$

Finalmente $E(S_n)$ no depende de n y como $S_n \xrightarrow{c.s.} S$ entonces S converge en distribución a S y $E[(S_n - S)^2] \longrightarrow 0$.

Luego $E(S_n) \longrightarrow E(S)$. □

Corolario 2.3.10. *Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de v.a. con $E(X_n|X_1, \ldots, X_{n-1}) = 0$ c.s. para todo $n \in \mathbb{N}$.*
Si $\sum\limits_{k=1}^{\infty} \frac{E(X_k^2)}{k^2} < \infty$ entonces $\frac{S_n}{n} \xrightarrow{c.s.} 0$.

Demostración:

Sea $Y_n = \sum_{k-=1}^{n} \frac{X_k}{k}$ entonces $E(Y_n|Y_1,\ldots,Y_{n-1}) =$

$E\Big[Y_{n-1}+\frac{X_n}{n}\Big|Y_1,\ldots,Y_{n-1}\Big] = Y_{n-1}+\frac{1}{n}E(X_n|X_1,\ldots,X_{n-1}) = Y_{n-1}.$

Por lo tanto $\{Y_n : n \in \mathbb{N}\}$ es martingala.

Luego $E(Y_n^2) = \sum_{k=1}^{n} \frac{E(X_k^2)}{k^2} + \sum_{j\neq k} \frac{X_jX_k}{jk}$.

Si suponemos que $j<k$ podemos calcular $E(X_jX_k)$.

$E(X_jX_k) = E\Big[E(X_jX_k|X_1,\ldots,X_{k-1})\Big] = E\Big[X_jE(X_k|X_1,\ldots,X_{k-1})\Big] =$
$E(X_j\cdot 0) = 0.$

De este modo $E(Y_n^2) = \sum_{k=1}^{n} \frac{E(X_k^2)}{k^2} \leqslant \sum_{k=1}^{\infty} \frac{E(X_k^2)}{k^2} < \infty.$

Por lo que existe una v.a. Y tal que $Y_n \xrightarrow{c.s.} Y$.

Asi $\sum_{k=1}^{n} \frac{X_k}{k} \xrightarrow{c.s.} Y$ como consecuencia del Teorema 2.3.9,

y por el lema (Kronecker) 2.3.6 haciendo $a_k = \frac{X_k}{k}$ tenemos que

$\sum_{k=1}^{n} \frac{ka_k}{n} = \sum_{k=1}^{n} \frac{X_k}{n} = \frac{S_n}{n} \xrightarrow[n\to\infty]{c.s.} 0.$ □

Corolario 2.3.11 (Teicher). *Sea $\{X_n : n\in\mathbb{N}\}$ una sucesión de variables aleatorias independientes con $E(X_n)=0$ y $E(X_n^2)=\sigma_n^2$.*

Si se cumplen:

i) $\sum_{n=2}^{\infty} \frac{\sigma_n^2}{n^4} \sum_{k=1}^{n-1} \sigma_k^2 < \infty.$

ii) $\sum_{j=1}^{n} \frac{\sigma_j^2}{n^2} \xrightarrow[n\to\infty]{} 0.$

iii) $\sum_{n=1}^{\infty} P[|X_n| \geqslant a_n] < \infty$ *para alguna sucesión $\{a_n : n\in\mathbb{N}\}$ de números* *reales positivos con* $\sum_{n=1}^{\infty} \frac{a_n^2 \sigma_n^2}{n^4} < \infty$.

Entonces $\frac{S_n}{n} \xrightarrow{c.s.} 0$.

Demostración:

Notemos que $\left(\frac{S_n}{n}\right)^2 = \left(\frac{X_1 + \cdots + X_n}{n}\right)^2 = \frac{1}{n^2}\sum_{k=1}^{n} X_k^2 + \frac{2}{n^2}\sum_{j<k} X_j X_k$.

Sea $Y_n = \begin{cases} X_n^2 & \text{si } |X_n| < a_n \\ 0 & \text{si no} \end{cases}$ entonces $\sum_{n=1}^{\infty} P[Y_n \neq X_n^2] = \sum_{n=1}^{\infty} P[|X_n| \geqslant a_n] < \infty$ por $iii)$.

Asi $\sum_{k=1}^{n} \frac{X_k^2}{n^2}$ y $\sum_{k=1}^{n} \frac{Y_k}{n^2}$ convergen al mismo límite.

Luego $Var(Y_n) = E(Y_n^2) - E^2(Y_n) \leqslant E(Y_n^2) = \int\limits_{[|X_n|<a_n]} X_n^4 dP < a_n^2 \int\limits_{[|X_n|<a_n]} X_n^2 dP < a_n^2 \int\limits_{\Omega} X_n^2 dP = a_n^2 Var(X_n) = a_n^2 \sigma_n^2$ implica que $\sum_{n=1}^{\infty} \frac{Var(Y_n)}{n^4} < \sum_{n=1}^{\infty} \frac{a_n^2 \sigma_n^2}{n^4} < \infty$ por $iii)$.

Ahora sea $Z_n = \frac{Y_n - E(Y_n)}{n}$.

Y_n y Z_n son funciones de X_n, donde $E(Z_n | Z_1, \ldots, Z_{n-1}) = E(Z_n) = 0$ por independencia.

$\sum_{k=1}^{\infty} \frac{E(Z_k^2)}{k^2} = \sum_{k=1}^{\infty} \frac{Var(Y_k)}{k^4} < \infty$ y por el Corolario 2.3.10 $\sum_{k=1}^{n} \frac{(Y_k - E(X_k))}{n^2} \xrightarrow{c.s.} 0$.

Además $\sum_{k=1}^{n} \frac{E(Y_k)}{n^2} \leqslant \sum_{k=1}^{n} \frac{\sigma_k^2}{n^2}$ porque $E(Y_k) = E(X_k^2 I_{[|X_k| \leqslant a_k]}) \leqslant E(X_k^2) = \sigma_k^2$.

De este modo $\sum_{k=1}^{n} \frac{Y_k}{n^2} - \sum_{k=1}^{n} \frac{E(Y_k)}{n^2} \xrightarrow[n\to\infty]{c.s.} 0$ pero $\sum_{k=1}^{n} \frac{E(Y_k)}{n^2} \xrightarrow[n\to\infty]{} 0$ por $ii)$.

Por lo tanto $\sum_{k=1}^{n} \frac{Y_k}{n^2} \xrightarrow[n\to\infty]{c.s.} 0$.

Hagamos $W_n = \sum_{j=2}^{n} \frac{X_j S_{j-1}}{j^2}$ para $n \geqslant 2$.
Notemos que W_n es función de
$X_1, \ldots, X_{n-1}, X_n$.
$E(W_n|W_1, \ldots, W_{n-1}) = E\Big[\sum_{j=2}^{n-1} \frac{X_j S_{j-1}}{j^2} + \frac{X_n S_{n-1}}{n^2}\Big| W_1, \ldots, W_{n-1}\Big] = \sum_{j=2}^{n-1} \frac{X_j S_{j-1}}{j^2} +$
$\frac{S_{n-1}}{n^2} E(X_n|X_1, \ldots, X_{n-1}) = \sum_{j=2}^{n-1} \frac{X_j S_{j-1}}{j^2} + \frac{S_{n-1}}{n^2} E(X_n) = \sum_{j=2}^{n-1} \frac{X_j S_{j-1}}{j^2}$ porque
$E(X_n) = 0$.
Por lo tanto $E(W_n|W_1, \ldots, W_{n-1}) = W_{n-1}$.
Mientras que $E(W_n^2) = \sum_{j=2}^{n} \frac{E(X_j^2 S_{j-1}^2)}{j^4} + \sum_{i \neq j} \frac{E(X_i X_j S_{i-1} S_{j-1})}{i^2 j^2}$.
Supongamos ahora que $i < j$:
$E[X_i X_j S_{i-1} S_{j-1}] = E\Big[E[X_i X_j S_{i-1} S_{j-1}|X_1, \ldots, X_{j-1}]\Big]$
$= E\Big[X_i S_{i-1} S_{j-1} E(X_j|X_1, \ldots, X_{j-1})\Big] = E\Big[X_i S_{i-1} S_{j-1} E(X_j)\Big] =$
$E\Big[X_i S_{i-1} S_{j-1} \cdot 0\Big] = 0$.
Así por independencia

$$E[W_n^2] = \sum_{j=2}^{n} \frac{E(X_j^2 S_{j-1}^2)}{j^4} = \sum_{j=2}^{n} \frac{E(X_j^2) E(S_{j-1}^2)}{j^4} = \sum_{j=2}^{n} \sigma_j^2 \frac{\sum_{i=1}^{j-1} \sigma_i^2}{j^4} \leqslant$$

$\sum_{j=2}^{\infty} \sigma_j^2 \sum_{i=1}^{j-1} \frac{\sigma_i^2}{j^4} < \infty$ por i).
Por el Teorema 2.3.9 W_n converge a W y por el Lema 2.3.6 de Kronecker
$W_n = \sum_{j=2}^{n} \frac{X_j S_{j-1}}{j^2} = \sum_{j=2}^{n} a_j$ asi $\sum_{j=2}^{n} \frac{\frac{j X_j S_{j-1}}{j^2}}{n} \xrightarrow[n \to \infty]{c.s.} 0$.
Luego $\sum_{j=2}^{n} \frac{X_j S_{j-1}}{jn} \xrightarrow[n \to \infty]{c.s.} 0$ entonces $\sum_{j=2}^{n} \frac{X_j S_{j-1}}{n^2} \xrightarrow[n \to \infty]{c.s.} 0$,
porque $j \leqslant n$, $jn \leqslant n^2$, $\frac{1}{n^2} \leqslant \frac{1}{jn}$.

Ahora $\sum\sum_{j<k} X_jX_k = \sum_{j=2}^{n} X_jS_{j-1} = \sum_{j=2}^{n} X_j(X_1+\cdots+X_{j-1})$.
Por lo tanto $\sum_{j<k} \frac{X_jX_k}{n^2} \xrightarrow[n\to\infty]{c.s.} 0$.
Entonces $\frac{S_n^2}{n^2} = \sum_{k=1}^{n} \frac{X_k^2}{n^2} + 2\sum_{j<k} \frac{X_jX_k}{n^2} \xrightarrow[n\to\infty]{c.s.} 0$ porque ambos términos convergen a cero c.s.
Asi $\frac{S_n}{n} \xrightarrow[n\to\infty]{c.s.} 0$.
La igualdad $\sum\sum_{j<k} X_jX_k = \sum_{j=2}^{n} X_jS_{j-1}$ se cumple porque $\sum\sum_{j<k} X_jX_k = \sum_{k=2}^{n}\sum_{j=1}^{k-1} X_jX_k = \sum_{k=2}^{n} X_k(X_1+\cdots+X_{k-1}) = \sum_{k=2}^{n} X_kS_{k-1}$. □

Consideremos una sucesión de variables aleatorias $\{X_n : n \in \mathbb{N}\}$, esta sucesión tiene límite finito o infinito siempre y cuando el número de oscilaciones entre dos reales $a < b$ cualquiera es finito.
Dicho de otro modo, para que una sucesión no tenga límite, basta con la existencia de dos reales $a < b$ tal que el número de oscilaciones entre ellos sea infinita.
Tomemos dos números reales $a < b$, para cada $\omega \in \Omega$, $\{X_n(\omega) : n \in \mathbb{N}\}$ es una sucesión de números reales.
Definimos:
$T_1(\omega) = \min\{n : X_n(\omega) \leqslant a\}$ = Primera vez que la sucesión está por debajo de a.
$T_2(\omega) = \min\{n : n > T_1(\omega), X_n(\omega) \geqslant b\}$ = Primera vez que la sucesión está por arriba de b después de estar por debajo de a.

$T_3(\omega) = \min\{n : n > T_2(\omega), X_n(\omega) \leqslant a\}$

$T_4(\omega) = \text{mín}\{n : n > T_3(\omega), X_n(\omega) \geqslant b\}$

$\vdots$

$T_{2k-1}(\omega) = \text{mín}\{n : n > T_{2k-2}(\omega), X_n(\omega) \leqslant a\}$ = La $k-$ésima vez que la sucesión está por debajo de a.

$T_{2k}(\omega) = \text{mín}\{n : n > T_{2k-1}(\omega), X_n(\omega) \geqslant b\}$ = La $k-$ésima vez que la sucesión está por arriba de b.

En el caso en que algún conjunto sea vacío, el $T(\omega)$ correspondiente es infinito.

Además $T_1 < T_2 < T_3 < \cdots$.

Por otro lado, cada v.a. T_s es una variable extendida, con $T_s \geqslant s$ c.s.

se cumpliría la igualdad si en cada paso hay una oscilación.

Para $n \geqslant 1$ sea

$$U_n(a,b) = \begin{cases} \text{máx}\{s : T_{2s} \leqslant n\} & \text{si } T_2 \leqslant n \\ 0 & \text{si } T_2 > n. \end{cases}$$

$U_n(a,b)$ es el número de cruces u oscilaciones de a a b.

Notemos que $0 \leqslant U_n(a,b) \leqslant [\frac{n}{2}]$.

$\lim_{n\to\infty} U_n(a,b)(\omega)$ existe en $\mathbb{R}^*$, si vale ∞ entonces la sucesión $\{X_n(\omega) : n \in \mathbb{N}\}$ no tiene límite en los números reales.

Veremos que $U_n(a,b) = \sum_{s=1}^{[\frac{n}{2}]} s I_{[T_{2s} \leqslant n, T_{2s+2} > n]}$.

Si $T_2(\omega) > n$ entonces $U_n(a,b) = 0$ cumpliéndose la igualdad porque todas las indicadoras del lado derecho valen cero debido a que

$[T_{2s} \leqslant n, T_{2s+2} > n] = \varnothing$ para todo $s \geqslant 1$.

Supongamos ahora que $T_2(\omega) \leqslant n$.

Sea $r = \max\{s \geqslant 1 : T_{2s}(\omega) \leqslant n\}$ entonces

$U_n(a,b) = r$, $T_{2r}(\omega) \leqslant n$, $T_{2r+2}(\omega) > n$ asi $rI_{[T_{2r}\leqslant n, T_{2r+2}>n]} = r$ y también para todo $s \neq r$ tenemos que $I_{[T_{2s}\leqslant n, T_{2s+2}>n]} = 0$.

De este modo $\sum_{s=1}^{[\frac{n}{2}]} sI_{[T_{2s}\leqslant n, T_{2s+2}>n]} = r = U_n(a,b)$.

$0 \leqslant U_n(a,b) \leqslant \left[\frac{n}{2}\right]$ por lo que $U_n(a,b)$ es integrable.

Para $j \geqslant 2$ definimos $Y_j(\omega)$ como

$$Y_j(\omega) = \begin{cases} 0 & \text{si existe } s \text{ tal que } T_{2s+1}(\omega) < j \leqslant T_{2s+2}(\omega) \\ 1 & \text{en otro caso.} \end{cases}$$

$$Y_j(\omega) = \begin{cases} 0 & \text{si la } (s+1)\text{−ésima vez que la sucesión estuvo por debajo de } a \\ & \text{es menor que } j \text{ que a su vez es menor o igual que la} \\ & (s+1)\text{−ésima vez que la sucesión estuvo arriba de} \\ & b, \text{ para alguna } s \geqslant 0. \\ 1 & \text{en otro caso.} \end{cases}$$

$$Y_j(\omega) = \begin{cases} 0 & \text{si para alguna } s \geqslant 0, \text{ la sucesión está en una transición} \\ & \text{de por debajo de } a \text{ a por arriba de } b \text{ al tiempo } j \text{ o paso } j. \\ 1 & \text{si no.} \end{cases}$$

Veremos que $Y_2 = I_{[X_1>a]}$.

Si $X_1(\omega) > a$ entonces $T_1(\omega) \neq 1$.

Luego $T_1(\omega) \geqslant 2$ pero $2 \leqslant T_1(\omega) < T_2(\omega) < T_3(\omega) < \cdots$ asi que $T_l(\omega) > 2$ para todo l de aqui que $Y_2(\omega) = 1$.

Ahora supongamos que $X_1(\omega) \leqslant a$, entonces $T_1(\omega) = 1$.
Asi que
$1 = T_1(\omega) < T_2(\omega)$ de donde $T_2(\omega) \geqslant 2$.
Por lo tanto $T_1(\omega) < 2 \leqslant T_2(\omega)$ tomando $s = 0$.Luego $Y_2(\omega) = 0$.
De este modo $Y_2 = I_{[X_1 > a]}$.
Ahora mostraremos que para $j \geqslant 3$ se cumple la igualdad
$Y_j = I_{[Y_{j-1}=0, X_{j-1} \geqslant b] \cup [Y_{j-1}=1, X_{j-1} > a]}$.
Definamos los eventos:
$A = [Y_{j-1} = 0, X_{j-1} \geqslant b]$, $B = [Y_{j-1} = 1, X_{j-1} > a]$,
entonces mostraremos que $Y_j = I_{A \cup B}$, pero antes explicaremos qué significa la ocurrencia de A y B.
El evento A ocurre si al tiempo $(j-1)$ la sucesión está en una transición de a hacia b y ya llegó a b.
El evento B ocurre si al tiempo $(j-1)$ la sucesión no está en una transición de a hacia b y está por arriba de a.
Notemos que los eventos A y B son ajenos.
Entonces verificamos a continuación la igualdad $Y_j = I_{A \cup B}$ donde
$A = [Y_{j-1} = 0, X_{j-1} \geqslant b]$. $B = [Y_{j-1} = 1, X_{j-1} > a]$.

1. Supongamos que ocurre A, es decir $\omega \in A$ entonces $Y_{j-1}(\omega) = 0$, $X_{j-1}(\omega) \geqslant b$.
   Asi existe $s \in \{0, 1, 2, \ldots\}$ tal que $T_{2s+1}(\omega) < j - 1 \leqslant T_{2s+2}(\omega) = \min\{n : n > T(\omega)_{2s+1}, X_{2s+2}(\omega) \geqslant b\}$ pero

$X_{j-1}(\omega) \geqslant b$.

Luego $T_{2s+2}(\omega) \leqslant j-1$.

Por lo tanto $T_{2s+2}(\omega) = j-1$ esto es j está en la transición de b hacia a, por lo que $Y_j(\omega) = 1$.

2. Supongamos que ocurre B, es decir $\omega \in B$ entonces $Y_{j-1}(\omega) = 1$, $X_{j-1}(\omega) > a$.

   Si nunca ha cruzado por debajo de a entonces $Y_j(\omega) = 1$.

   Si ya cruzó por debajo de a alguna vez, está entre b y a pues $X_{j-1}(\omega) > a$ asi en el tiempo j de todos modos sigue en una transición entre b y a pues lo mejor que podía ocurrir es que $X_j(\omega) \leqslant a$ esto es $Y_j(\omega) = 1$.

   Por lo tanto $I_{A\cup B}(\omega) = 1$ implica que $Y_j(\omega) = 1$.

   Ahora supongamos que $A \cup B$ no ocurre y veremos que $Y_j = 0$.

   Supongamos $(A\cup B)^c = A^c \cap B^c = \Big\{[Y_{j-1} = 1] \cup [X_{j-1} < b]\Big\} \cap \Big\{[Y_{j-1} = 0] \cup [X_{j-1} \leqslant a]\Big\}$ ocurre.

   Sea $\omega \in \Big\{[Y_{j-1} = 1] \cup [X_{j-1} < b]\Big\} \cap \Big\{[Y_{j-1} = 0] \cup [X_{j-1} \leqslant a]\Big\}$ entonces $\Big\{Y_{j-1}(\omega) = 1 \text{ ó } X_{j-1}(\omega) < b\Big\}$ y $\Big\{Y_{j-1}(\omega) = 0 \text{ ó } X_{j-1}(\omega) \leqslant a\Big\}$.

   Si $Y_{j-1}(\omega) = 1$ implica que $X_{j-1}(\omega) \leqslant a$.

   Entonces, en el tiempo $(j-1)$ la sucesión está en una transición de b hacia a y acaba de llegar a a.

   Luego, en el tiempo j irá de regreso, es decir, la sucesión irá de a hacia b.

   Esto implica que $Y_j(\omega) = 0$.

Si $Y_{j-1}(\omega) = 0$ entonces $X_{j-1}(\omega) < b$.

Luego, en el tiempo $(j-1)$ la sucesión está en una transición de a hacia b pero aún no llega a b pues $X_{j-1}(\omega) < b$.

Por lo que, en el tiempo j seguirá en el mismo recorrido.

De este modo $Y_j(\omega) = 0$.

Por lo tanto $Y_j = I_{A \cup B}$.

Asi que Y_j es variable aleatoria $\sigma(X_1, \ldots, X_{j-1})$ medible.

Un tratamiento diferente de este problema se puede encontrar en Chow y Teicher (1978).

Lema 2.3.12. *Para toda sucesión* $\{X_n : n \in \mathbb{N}\}$ *se cumple que* $\sum_{s=2}^{n} Y_s(X_s - X_{s-1}) + (b-a)U_n(a,b) \leqslant (X_n - a)^+$ *c.s.*

Demostración:

Hay varios casos:

1. Si $U_n(a,b) = 0$, entonces $T_2 > n$.

   i) Si $T_1 = 1$ tenemos que $X_1 \leqslant a$, luego $Y_s = 0$ para $s = 2, \ldots, n$. Por lo tanto $\sum_{s=2}^{n} Y_s(X_s - X_{s-1}) + (b-a)U_n(a,b) = 0 \leqslant (X_n - a)^+$.

   ii) Si $1 < T_1 \leqslant n$ entonces $Y_s = 0$ cuando $s > T_1$ y $Y_s = 1$ para $s \leqslant T_1$.

   Asi que $\sum_{s=2}^{n} Y_s(X_s - X_{s-1}) = \sum_{s=2}^{T_1} Y_s(X_s - X_{s-1}) = X_{T_1} - X_1 < 0$ porque $X_{T_1} \leqslant a$ y $X_1 > a$.

   De aqui que

   $$\sum_{s=2}^{n} Y_s(X_s - X_{s-1}) + (b-a)U_n(a,b) = X_{T_1} - X_1 < 0 \leqslant (X_n - a)^+.$$

iii) Si $T_1 > n$ entonces la sucesión no ha estado por debajo de a.

Tenemos que $Y_s = 1$ para $s = 2, 3, \ldots, n$;
así $\sum_{s=2}^{n} Y_s(X_s - X_{s-1}) = X_n - X_1 < X_n - a \leqslant (X_n - a)^+ = \max\{0, X_n - a\}$ porque $X_1 > a$.
De este modo $\sum_{s=2}^{n} Y_s(X_s - X_{s-1}) + (b-a)U_n(a,b) \leqslant (X_n - a)^+$ c.s.

2. Si $U_n(a,b) > 0$, entonces $T_2 \leqslant n$ asi que $T_1 < T_2 \leqslant n$.

Si $T_1 \geqslant 1$ para todo j tal que $T_1 < j \leqslant T_2$ se tiene que $Y_j = 0$.
Por lo tanto $\sum_{s=2}^{n} Y_s(X_s - X_{s-1}) = (X_{T_1} - X_1) + \sum_{s=T_2+1}^{n} Y_s(X_s - X_{s-1})$
pero $X_{T_1} - X_1 \leqslant 0$ porque $X_{T_1} \leqslant a$ y $X_1 = X_{T_1}$ si $T_1 = 1$ de otro modo $X_1 > a$.
Luego $(X_{T_1} - X_1) + \sum_{s=T_2+1}^{n} Y_s(X_s - X_{s-1}) \leqslant \sum_{s=T_2+1}^{n} Y_s(X_s - X_{s-1})$ ya que si $j \leqslant T_1$ entonces $Y_j = 1$.
Aquí hay dos posibilidades:
Existe t tal que $T_{2t+1} < n \leqslant T_{2t+2}$ o bien existe t' tal que $T_{2t'} < n \leqslant T_{2t'+1}$.
Son ajenos ambos casos.
Si se da la primer posibilidad $\sum_{s}^{n} Y_s(X_s - X_{s-1}) = (X_{T_3} - X_{T_2}) + \cdots + (X_{T_{2t+1}} - X_{T_{2t}}) \leqslant (a-b)t = (a-b)U_n(a,b)$.
Asi $\sum_{s=2}^{n} Y_s(X_s - X_{s-1}) \leqslant (a-b)U_n(a,b)$.
Por lo que $\sum_{s=2}^{n} Y_s(X_s - X_{s-1}) + (b-a)U_n(a,b) \leqslant 0 < (X_n - a)^+$.
Si se da la segunda posibilidad entonces,
$\sum_{s=T_2+1}^{n} Y_s(X_s - X_{s-1}) = (X_{T_3} - X_{T_2}) + \cdots + (X_{T_{2t'-1}} - X_{T_{2t'-2}}) + (X_n -$

$X_{T_{2t'}}) \leqslant (a-b)t' + (a - X_{T_{2t'}}) + (X_n - a) \leqslant (a-b)t' + (X_n - a)$ porque $(a - X_{T_{2t'}}) < 0$ debido a que $X_{T_{2t'}} \geqslant b > a$.

Pero $(a-b)t' + (X_n - a) \leqslant (a-b)U_n(a,b) + (X_n - a)^+$.

Luego $\sum_{s=2}^{n} Y_s(X_s - X_{s-1}) + (b-a)U_n(a,b) \leqslant (X_n - a)^+$. □

Teorema 2.3.13 (Desigualdad de Cruces de Doob). *Sea* $\{X_1, \ldots, X_n\}$ *una submartingala con respecto a* $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_n \subseteq \mathcal{F}$ *entonces* $E[U_n(a,b)] \leqslant \dfrac{1}{b-a} E[(X_n - a)^+]$.

Demostración:

Y_s definida como antes es $\mathcal{F}_{s-1}$ medible entonces

$E\Big[Y_s(X_s - X_{s-1})\Big] = E\Big[E[Y_s(X_s - X_{s-1}) | \mathcal{F}_{s-1}]\Big] = E\Big[Y_s E[(X_s - X_{s-1}) | \mathcal{F}_{s-1}]\Big] = E\Big[Y_s[E(X_s | \mathcal{F}_{s-1}) - X_{s-1}]\Big] \geqslant 0$ porque $E[X_s | \mathcal{F}_{s-1}] \geqslant X_{s-1}$ (submartingala).

Asi $0 \leqslant E\Big[\sum_{s=2}^{n} Y_s(X_s - X_{s-1})\Big] \leqslant (a-b)E[U_n(a.b)] + E[(X_n - a)^+]$.

De este modo $(b-a)E[U_n(a,b)] \leqslant E[(X_n - a)^+]$.

Por lo tanto $\dfrac{E[(X_n - a)^+]}{b-a} \geqslant E[U_n(a,b)]$ que establece una cota superior para el número promedio de cruces de a hacia b. □

Teorema 2.3.14 (Convergencia de submartingalas). *Sea* $\{X_n, \mathcal{F}_n : n \in \mathbb{N}\}$ *una submartingala. Supongamos que* $\limsup E(|X_n|) < \infty$. *Existe una v.a.* X *que es* $\sigma\left(\bigcup_{n=1}^{\infty} \mathcal{F}_n\right)$ *—medible tal que* $X_n \xrightarrow[n\to\infty]{c.s.} X$ *y además* $E(|X|) \leqslant \limsup_{n\to\infty} E(|X_n|) < \infty$.

Demostración:

Sean $a < b$ con $a, b \in \mathbb{Q}$.

Definimos $D(a,b) = \{\omega : \liminf X_n(\omega) < a < b < \limsup X_n(\omega)\}$

Sea $D = \bigcup\limits_{\substack{a<b \\ a,b\in\mathbb{Q}}} D(a,b)$.

Notemos que D es el conjunto en donde la sucesión $\{X_n : n \in \mathbb{N}\}$ no converge, porque una sucesión a_n no converge si $\liminf a_n < \limsup a_n$ y por tanto introducimos en medio dos racionales.

$D(a,b)$ es medible porque $D(a,b) = [\limsup X_n > b] \cap [\liminf X_n < a]$ y además $\liminf X_n$ y $\limsup X_n$ son v.a.

Veremos que $P(D) = 0$.

Basta probar que $P[D(a,b)] = 0$ porque

$P(D) \leqslant \sum\limits_{a,b\in\mathbb{Q}} P[D(a,b)]$.

Sea $n \geqslant 1$, $U_n = U_n(a,b)$ es el número de cruces de $X_1, \ldots, X_n$ de a hacia b.

Entonces $\{U_n : n \in \mathbb{N}\}$ es una sucesión creciente de v.a. no negativas por tanto U_n converge o diverge a infinito.

Para $\omega \in \Omega$ hacemos $U(\omega) = \lim\limits_{n\to\infty} U_n(\omega)$.

Mostraremos que $P[\{\omega : U(\omega) = \infty\}] = 0 = P[D(a,b)]$.

$\lim\limits_{n\to\infty} E(U_n) = \limsup\limits_{n\to\infty} E(U_n) \leqslant \dfrac{1}{b-a} \limsup\limits_{n\to\infty} E[(X_n - a)^+]$ por la desigualdad de cruces de Doob.

$$\frac{1}{b-a} \limsup_{n\to\infty} E[(X_n - a)^+] \leqslant \frac{1}{b-a} \limsup_{n\to\infty} E[(|X_n| + |a|)] < \infty$$

ya que $0 \leqslant |X_n| + |a|$, $X_n - a \leqslant |X_n| + |a|$ entonces $\max\{0, X_n - a\} \leqslant |X_n| + |a|$.

Por lo tanto $(X_n - a)^+ \leqslant |X_n| + |a|$.

Así $\lim\limits_{n\to\infty} E(U_n) < \infty$, por el Teorema de Convergencia Monótona $U_n \uparrow U$.

De aqui que $E(U)$ existe, luego U es finita casi seguramente.

Por lo tanto $P[U = \infty] = 0$ pero $D(a, b) \subseteq \{\omega : U(\omega) = \infty\}$.

Entonces $P[D(a, b)] \leqslant P[\{\omega : U(\omega) = \infty\}] = P[U = \infty] = 0$.

De este modo $P(D) = 0$.

Entonces X_n converge casi seguramente a una v.a. X.

X_n es $\mathcal{F}_n-$medible para todo $n \geqslant 1$ implica que X es $\sigma\left(\bigcup_{n=1}^{\infty} \mathcal{F}_n\right)-$medible.

Además por el Lema de Fatou, $E(|X|) \leqslant E[\liminf_{n\to\infty} |X_n|] \leqslant \liminf_{n\to\infty} E(|X_n|) \leqslant \limsup_{n\to\infty} E(|X_n|) < \infty$. □

Observaciones:

1. El resultado anterior se cumple si reemplazamos $\limsup_{n\to\infty} E(|X_n|) < \infty$ por $\sup_{n\geqslant 1} E(|X_n|) < \infty$. En la demostración se usó que $\limsup_{n\to\infty} E(|X_n|) < \infty$ en los siguientes pasos:

$$\begin{aligned} \lim_{n\to\infty} E(U_n) &= \limsup_{n\to\infty} E(U_n) \text{ por monotonía} \\ &= \sup E(U_n) \text{ por monotonía} \\ &\leqslant \sup \frac{1}{b-a} E[(X_n - a)^+] \leqslant \frac{1}{b-a} \sup[E(|X_n|) + |a|]. \end{aligned}$$

   Otro lugar en donde aparece $\limsup_{n\to\infty} E(|X_n|)$ es:

   Por el Lema de Fatou: $E(|X|) \leqslant E[\liminf_{n\to\infty} |X_n|] \leqslant \liminf_{n\to\infty} E(|X_n|) = \sup_{n\geqslant 1} \inf_{k\geqslant r} E(|X_k|) \leqslant \sup_{n\geqslant 1} E(|X_n|) < \infty$.

2. Si X_n es submartingala entonces $-X_n$ es supermartingala por tanto el Teorema es cierto para supermartingalas con la misma hipótesis, porque $\limsup E[|-X_n|] = \limsup E[|X_n|]$, como una martingala siempre es submartingala el resultado es cierto para martingalas.

3. Notar que:

a) $\limsup E(|X_n|) < \infty$ si y sólo si $\limsup E(X_n^+) < \infty$.

b) $\sup E(|X_n|) < \infty$ si y sólo si $\sup E(X_n^+) < \infty$.

Demostración:

$|X_n| = X_n^+ + X_n^- = X_n^+ + X_n^- + X_n^+ - X_n^+ = 2X_n^+ - (X_n^+ - X_n^-) = 2X_n^+ - X_n$.

Entonces $E(|X_n|) = 2E(X_n^+) - E(X_n)$, pero como $\{X_n : n \in \mathbb{N}\}$ es submartingala tenemos que $E(X_n|\mathcal{F}_{n-1}) \geqslant X_{n-1}$. Luego $E[E(X_n|\mathcal{F}_{n-1})] = E(X_n) \geqslant E(X_{n-1}) \geqslant \cdots \geqslant E(X_1)$.

Por lo tanto $E(|X_n|) = 2E(X_n^+) - E(X_n) \leqslant 2E(X_n^+) - E(X_1)$.

Asi $\limsup\limits_{n\to\infty} E(|X_n|) \leqslant 2 \limsup\limits_{n\to\infty} E(X_n^+) - E(X_1)$.

De este modo si $\limsup\limits_{n\to\infty} E(X_n^+) < \infty$ entonces $\limsup\limits_{n\to\infty} E(|X_n|) < \infty$.

Lo mismo ocurre con supremo en vez de límite superior.

Mientras que por otro lado:

$|X_n| = X_n^+ + X_n^-$ con $X_n^+ \geqslant 0$ y $X_n^- \geqslant 0$.

Entonces $X_n^+ \leqslant |X_n|$ asi $E(X_n^+) \leqslant E(|X_n|)$ por lo que $\limsup\limits_{n\to\infty} E(X_n^+) \leqslant \limsup\limits_{n\to\infty} E(|X_n|)$.

Por otro lado, si $\limsup\limits_{n\to\infty} E(|X_n|) < \infty$ entonces $\limsup\limits_{n\to\infty} E(X_n^+) < \infty$.

Lo mismo ocurre con supremo en vez de límite superior.

También notemos que como $\{X_n : n \in \mathbb{N}\}$ es submartingala,

$\{X_n^+ : n \in \mathbb{N}\}$ también lo es. porque $h(x) = x^+$ es convexa y $E[X_n^+|\mathcal{F}_{n-1}] = E[h(X_n^+)|\mathcal{F}_{n-1}] \geqslant h(E(X_n|\mathcal{F}_{n-1})) = E(X_n|\mathcal{F}_{n-1})^+ \geqslant X_{n-1}^+$ luego $E(X_n^+) \geqslant E(X_{n-1}^+)$.

Por lo tanto $\{E(X_n^+) : n \in \mathbb{N}\}$ es una sucesión creciente.

Entonces $\limsup_{n\to\infty} E(X_n^+) = \infty$ ó $\limsup_{n\to\infty} E(X_n^+) = c < \infty$.

4. La condición $\limsup_{n\to\infty} E(X_n^+) < \infty$ ó $\sup E(X_n^+) < \infty$ no es suficiente para asegurar que $X_n \xrightarrow{L_1} X$.

Si tomamos $Z_1, Z_2, \ldots$ una sucesión de variables aleatorias i.i.d. tal que $P[Z_n = 0] = \frac{1}{2} = P[Z_n = 2]$ y definimos $X_n = \prod_{k=1}^{n} Z_k$ tenemos que X_n es $\sigma(Z_1, \ldots, Z_n)$ medible y $E(X_n|X_1, \ldots, X_{n-1}) = E(X_n|Z_1, \ldots, Z_{n-1}) =$
$E(Z_1 \cdots Z_{n-1} Z_n | Z_1, \ldots, Z_{n-1}) Z_1 \cdots Z_{n-1} E(Z_n|Z_1, \ldots, Z_{n-1})$
$= Z_1 \cdots Z_{n-1} E(Z_n) = Z_1 \cdots Z_{n-1}$
por independencia y porque $E(Z_n) = 0 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = 1$.
Por lo tanto $E(X_n|X_1, \ldots, X_{n-1}) = Z_1 \cdots Z_{n-1} = X_{n-1}$ asi que $\{X_n : n \in \mathbb{N}\}$ es martingala.

X_n tiene la forma: $X_n = \begin{cases} 2^n & \text{con probabilidad } (1/2)^n \\ 0 & \text{con probabilidad } 1 - (1/2)^n. \end{cases}$

Entonces $E(X_n) = 2^n \dfrac{1}{2^n} + 0 \cdot (1 - (1/2)^n) = 1$.

Además $|X_n| = X_n$ por lo que $\limsup_{n\to\infty} E(|X_n|) = \limsup_{n\to\infty} E(X_n) = 1 < \infty$.

Luego por el Teorema 2.3.14 existe X v.a. tal que $X_n \xrightarrow[n\to\infty]{c.s.} X$ donde $X = 0$ c.s.

Pues $P[X_n \neq 0] = (1/2)^n \xrightarrow[n\to\infty]{} 0$, luego $E[|X_n - X|] = E[|X_n|] = 1 \underset{n\to\infty}{\not\to} 0$, es decir, $E(X_n) \not\to E(X)$ de donde convergencia casi segura no implica convergencia en L^1.

Por lo tanto $\{X_1, X_2, \ldots, X\}$ no es submartingala ya que
$E(X|X_1, \ldots, X_n) = E(0|X_1, \ldots, X_n) = 0 \not\geq X_n$ c.s.

5. Toda submartingala uniformemente acotada (o supermartingala) converge casi seguramente.

Esto es, $\{X_n : n \in \mathbb{N}\}$ tal que $E(X_n|\mathcal{F}_{n-1}) \geqslant X_{n-1}$ c.s. en donde existe $M \geqslant 0$ tal que $|X_n| \leqslant M$ para todo $n \in \mathbb{N}$ converge c.s. porque $E(|X_n|) \leqslant M$ asi $\limsup E(|X_n|) \leqslant M < \infty$ y por la observación 2) también vale para supermartingalas.

Ahora si tenemos una martingala no negativa entonces converge c.s. ya que $\limsup E(|X_n|) = \limsup E(X_n) = E(X_1) < \infty$.

De hecho si X_n es supermartingala no negativa tenemos que $\limsup E(|X_n|) = \limsup E(X_n) \leqslant E(X_1) < \infty$ pues $E(X_n|\mathcal{F}_{n-1}) \leqslant X_{n-1}$ entonces $E(X_n) \leqslant E(X_{n-1}) \leqslant \cdots \leqslant E(X_1) < \infty$.

También si X_n es submartingala no negativa tenemos, $\limsup E(|X_n|) = \limsup E(-X_n) \leqslant -E(X_1) < \infty$ debido a que $E(X_n|\mathcal{F}_{n-1}) \geqslant X_{n-1}$ luego $E(X_n) \geqslant E(X_{n-1}) \geqslant \cdots \geqslant E(X_1)$.

Asi $-E(X_n) \leqslant -E(X_1)$.

2.4. Integrabilidad Uniforme

Definición 2.4.1. Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, $\{X_t : t \in T\}$ una familia de variables aleatorias sobre $(\Omega, \mathcal{F}, P)$ tal que $E(|X_t|) < \infty$ para todo $t \in T$, es uniformemente integrable con respecto a P si:

$$\sup_{t\in T} \int_{[|X_t|\geqslant a]} |X_t| dP \xrightarrow[a\to\infty]{} 0$$

Ejemplo 2.4.2. Sea $\{X_n : n \in \mathbb{N}\}$ tal que $X_n \sim N(0, n^2)$, esto es, $E(X_n) = 0$ y $\dfrac{X_n}{n} \sim N(0,1)$, entonces si $a > 0$:

$$\int_{[|X_n|\geqslant a]} |X_n| dP \geqslant \int_{[|X_n|\geqslant a]} a dP = aP[|X_n| \geqslant a] = a\{1 - P[|X_n| < a]\} =$$

$$a\{1 - P(-a < X_n < a)\} = a\left[1 - P\left(-\frac{a}{n} < \frac{X_n}{n} < \frac{a}{n}\right)\right] =$$

$$a\left[1 - \left(\phi\left(\frac{a}{n}\right) - \phi\left(-\frac{a}{n}\right)\right)\right] = a\left[1 - \left(2\phi\left(\frac{a}{n}\right) - 1\right)\right] = 2a\left[1 - \phi\left(\frac{a}{n}\right)\right]$$

donde $\phi(x) = P[Y \leqslant x]$ con $Y \sim N(0,1)$.

Luego $\sup_{n\in\mathbb{N}} \int_{[|X_n|\geqslant a]} |X_n| dP \geqslant 2a \sup_{n\in\mathbb{N}} \left[1 - \phi\left(\frac{a}{n}\right)\right] = 2a[1 - \frac{1}{2}] = a.$

Por lo tanto $\sup_{n\in\mathbb{N}} \int_{[|X_n|\geqslant a]} |X_n| dP \geqslant a$ entonces

$\lim_{a\to\infty} \sup_{n\in\mathbb{N}} \int_{[|X_n|\geqslant a]} |X_n| dP \geqslant \lim_{a\to\infty} a \neq 0.$

De este modo $\{X_n : n \in \mathbb{N}\}$ no es uniformemente integrable. ■

Ejemplo 2.4.3. Sea $\{X_n : n \in \mathbb{N}\}$ tal que $X_n \sim N(0, \frac{1}{n})$ por lo que $Y_n = \sqrt{n} X_n \sim N(0,1)$, entonces si $a > 0$

$$\int_{[|X_n|\geqslant a]} |X_n| dP = \frac{1}{\sqrt{n}} \int_{[|Y_n|\geqslant a\sqrt{n}]} |Y_n| dP = \frac{1}{\sqrt{n}} \left[\int_{\Omega} |Y_n| dP - \int_{[|Y_n|\geqslant a\sqrt{n}]} |Y_n| dP\right]$$

Calcularemos cada uno de los dos términos:

i) $\int\limits_{\Omega} |Y_n| dP,$

y

ii) $\int\limits_{[|Y_n|<a\sqrt{n}]} |Y_n| dP.$

i) $\int\limits_{\Omega} |Y_n| dP = \int\limits_{-\infty}^{\infty} |y| \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{\infty} y e^{-\frac{y^2}{2}} dy - \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{0} y e^{-\frac{y^2}{2}} dy =$
$\frac{1}{\sqrt{2\pi}} \left[e^{-\frac{y^2}{2}} \Big|_{\infty}^{0} + e^{-\frac{y^2}{2}} \Big|_{-\infty}^{0} \right] = \frac{1}{\sqrt{2\pi}} \cdot 2 = \sqrt{\frac{2}{\pi}}.$

ii) $\int\limits_{[|Y_n|<a\sqrt{n}]} |Y_n| dP = \int\limits_{-a\sqrt{n}}^{a\sqrt{n}} |y| \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \frac{1}{\sqrt{2\pi}} \int\limits_{0}^{a\sqrt{n}} y e^{-\frac{y^2}{2}} dy -$
$\frac{1}{\sqrt{2\pi}} \int\limits_{-a\sqrt{n}}^{0} y e^{-\frac{y^2}{2}} dy = \frac{1}{\sqrt{2\pi}} \left[e^{-\frac{y^2}{2}} \Big|_{a\sqrt{n}}^{0} + e^{-\frac{y^2}{2}} \Big|_{-a\sqrt{n}}^{0} \right] = \frac{1}{\sqrt{2\pi}} \left[2 - 2e^{-\frac{na^2}{2}} \right] =$
$\frac{2}{\sqrt{2\pi}} \left[1 - e^{-\frac{na^2}{2}} \right] = \sqrt{\frac{2}{\pi}} \left[1 - e^{-\frac{na^2}{2}} \right].$

Asi $\int\limits_{[|X_n| \geqslant a]} |X_n| dP = \frac{1}{\sqrt{n}} \left\{ \sqrt{\frac{2}{\pi}} - \sqrt{\frac{2}{\pi}} \left[1 - e^{-\frac{na^2}{2}} \right] \right\} = \sqrt{\frac{2}{n\pi}} e^{-\frac{na^2}{2}}.$

Luego $\sup\limits_{n\in\mathbb{N}} \int\limits_{[|X_n| \geqslant a]} |X_n| dP = \sup\limits_{n\in\mathbb{N}} \sqrt{\frac{2}{n\pi}} e^{-\frac{na^2}{2}} = \sqrt{\frac{2}{\pi}} e^{-\frac{a^2}{2}} \xrightarrow[a\to\infty]{} 0.$

Por lo tanto $\{X_n : n \in \mathbb{N}\}$ es uniformemente integrable. ■

Ejemplo 2.4.4. Si $X_n = \begin{cases} 2^n & \text{con probabilidad } (\frac{1}{2})^n \\ 0 & \text{con probabilidad } 1 - (\frac{1}{2})^n \end{cases}$ entonces

sea $a > 0$:

$$\int\limits_{[|X_n|\geqslant a]} |X_n|dP = \int\limits_{[X_n\geqslant a]} X_n dP = \begin{cases} \int\limits_{[X_n\geqslant a]} 0dP & \text{si } a > 2^n \\ \int\limits_{[X_n\geqslant a]} 2^n dP = 2^n \cdot \left(\frac{1}{2}\right)^n = 1 & \text{si } a \leqslant 2^n \end{cases}$$

Entonces $\sup\limits_{n\in\mathbb{N}} \int\limits_{[|X_n|\geqslant a]} |X_n|dP = 1 \underset{a\to\infty}{\nrightarrow} 0.$

Por lo tanto $\{X_n : n \in \mathbb{N}\}$ no es uniformemente integrable. ■

Proposición 2.4.5. *Sea $\{X_t : t \in T\}$ uniformemente integrable, entonces se cumplen:*

a) $\sup\limits_{t\in T} E(|X_t|) < \infty.$

b) $\sup\limits_{t\in T} P[|X_t| \geqslant a] \underset{a\to\infty}{\longrightarrow} 0.$

c) $h_t : \mathcal{F} \longrightarrow \mathbb{R}$ *tal que* $h_t(E) = \int\limits_E |X_t|dP$ *es uniformemente absolutamente continua, esto es, para todo $\epsilon > 0$ existe $\delta(\epsilon) > 0$ tal que si $P(E) < \delta$ entonces $h_t(E) < \epsilon$ para toda $t \in T$.*

Además si se cumple a) y c) ó b) y c) entonces $\{X_t : t \in T\}$ es uniformemente integrable.

Demostración:

a) Como $\lim\limits_{a\to\infty} \sup\limits_{t\in T} \int\limits_{[|X_t|\geqslant a]} |X_t|dP = 0$ entonces existe $a_0 > 0$ tal que $\sup\limits_{t\in T} \int\limits_{[|X_t|\geqslant a_0]} |X_t|dP < 1,$ luego

$$E(|X_t|) = \int\limits_{[|X_t|<a_0]} |X_t|dP + \int\limits_{[|X_t|\geqslant a_0]} |X_t|dP < a_0P[|X_t| < a_0]+1 \leqslant a_0+1.$$

Por lo tanto $\sup\limits_{t\in T} E(|X_t|) \leqslant a_0 + 1 < \infty$

b) Por a) tenemos que para todo $t \in T$ y $a > 0$

$$\infty > c = \sup_{t\in T} E(|X_t|) \geqslant E(|X_t|) = \int\limits_{[|X_t|<a]} |X_t|dP + \int\limits_{[|X_t|\geqslant a]} |X_t|dP \geqslant$$

$$\int\limits_{[|X_t|\geqslant a]} |X_t|dP \geqslant aP[|X_t| \geqslant a].$$

Asi que $P[|X_t| \geqslant a] \leqslant \dfrac{c}{a} \xrightarrow[a\to\infty]{} 0.$

Notemos que esta prueba equivale a a) implica b).

c) Sea $E \in \mathcal{F}$, entonces para todo $a > 0$ y para todo $t \in T$ tenemos que

$$\int\limits_E |X_t|dP = \int\limits_{E\cap[|X_t|\geqslant a]} |X_t|dP + \int\limits_{E\cap[|X_t|<a]} |X_t|dP.$$

Para $\epsilon > 0$ elegimos $a_0 > 0$ tal que $\int\limits_{[|X_t|\geqslant a_0]} |X_t|dP < \frac{\epsilon}{2}$ para todo $t \in T$

luego $\int\limits_E |X_t|dP \leqslant \int\limits_{[|X_t|\geqslant a_0]} |X_t|dP + \int\limits_E a_0dP,$

porque $E \cap [|X_t| \geqslant a_0] \subseteq [|X_t| \geqslant a_0]$ y $E \cap [|X_t| < a_0] \subseteq E,$

así $\int\limits_E |X_t|dP \leqslant \int\limits_{[|X_t|\geqslant a_0]} |X_t|dP + a_0dP(E) < \frac{\epsilon}{2} + a_0P(E).$

Por lo que si $P(E) < \dfrac{\epsilon}{2a_0}$ se cumple

$$\int\limits_E |X_t|dP < \frac{\epsilon}{2} + a_0\frac{\epsilon}{2a_0} = \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Ahora veremos que b) y c) implican integrabilidad uniforme y como a) implica b) entonces a) y c) automáticamente implicarán integrabilidad uniforme.

Sea $\epsilon > 0$, por c) existe $\delta > 0$ tal que si $P(E) < \delta$ entonces $\int_E |X_t| dP < \epsilon$ para todo $t \in T$.

Ahora por b) para $\delta > 0$ existe $a_0 > 0$ tal que para todo $a \geqslant a_0$, $P[|X_t| \geqslant a] < \delta$.

Sea $E = [|X_t| \geqslant a]$ con $a \geqslant a_0$, $P(E) < \delta$.

Entonces $\int_{[|X_t| \geqslant a]} |X_t| dP < \epsilon$ para todo $t \in T$.

Por lo tanto $\lim_{a \to \infty} \sup_{t \in T} \int_{[|X_t| \geqslant a]} |X_t| dP = 0$. □

El resultado anterior y los siguientes sirven para verificar integrabilidad uniforme.

Corolario 2.4.6. *i) Sea $|X_t| \leqslant Y$ c.s. para todo $t \in T$, con $E(Y) < \infty$ entonces $\{X_t : t \in T\}$ es uniformemente integrable.*

ii) Si T es finito y $E(|X_t|) < \infty$ para todo $t \in T$ entonces $\{X_t : t \in T\}$ es uniformemente integrable.

Demostración:

i) Como $|X_t| \leqslant Y$ entonces $E(|X_t|) \leqslant E(Y) < \infty$ para todo $t \in T$.

Por lo tanto $\sup_{t \in T} E(|X_t|) < \infty$.

Asi que a) del resultado anterior se cumple.

Luego si $E \in \mathcal{F}$, $\lambda(E) = \int_E Y dP \longrightarrow 0$ si $P(E) \longrightarrow 0$ por Radon-Nikodym debido a que $E(Y) < \infty$.

Por lo que $0 \leqslant \int_E |X_t| dP \leqslant \int_E Y dP \longrightarrow 0$ si $P(E) \longrightarrow 0$.

De este modo $c)$ del resultado anterior se cumple.

Y como $a), c)$ implican integrabilidad uniforme entonces $\{X_t : t \in T\}$ es uniformemente integrable (u.i.).

ii) Sea T finito y $E(|X_t|) < \infty$ para todo $t \in T$.

Tomamos $Y = \sum_{t \in T} |X_t|$ entonces $|X_{t_0}| \leqslant Y$ para todo $t_0 \in T$ y $E(Y) = \sum_{t \in T} E(|X_t|) < \infty$ y por $i)$ $\{X_t : t \in T\}$ es uniformemente integrable. □

Corolario 2.4.7. *Si $\{X_t : t \in T\}$ y $\{Y_t : t \in T\}$ son u.i. entonces $\{X_t + Y_t : t \in T\}$ y $\{X_t - Y_t : t \in T\}$ son u.i.*

Demostración:

$$E(X_t + Y_t) = E(X_t) + E(Y_t) < \infty,$$

$$\text{y } E(X_t - Y_t) = E(X_t) - E(Y_t) < \infty.$$

También

$$|X_t \pm Y_t| \leqslant |X_t| + |Y_t|$$

luego $\sup_{t \in T} E(X_t \pm Y_t) \leqslant \sup_{t \in T} E(|X_t|) + \sup_{t \in T} E(|Y_t|) < \infty$.

Por lo tanto $a)$ se cumple.

Si hacemos $h(E) = \int_E |X_t| dP$ y $g(E) = \int_E |Y_t| dP$,

entonces $L(E) = \int_E |X_t \pm Y_t| dP \leqslant \int_E |X_t| dP + \int_E |Y_t| dP = h(E) + g(E)$.

Sea $\epsilon > 0$, existen $\delta_1 > 0$ y $\delta_2 > 0$ tales que si $P(E) < \delta_1$ entonces $h(E) < \epsilon/2$

y si $P(E) < \delta_2$ entonces $g(E) < \epsilon/2$.
Haciendo $\delta = \text{mín}\{\delta_1, \delta_2\}$ sucede que cuando $P(E) < \delta$ tenemos que $h(E) < \epsilon/2$, $g(E) < \epsilon/2$ y por tanto $\int\limits_E |X_t \pm Y_t| dP \leqslant h(E) + g(E) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$.
De este modo $c)$ se cumple, ahora $a)$ y $c)$ implican u.i. □

Proposición 2.4.8. *Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de variables aleatorias tales que $X_n \xrightarrow[n\to\infty]{P} X$.*

1. *Si $\{X_n : n \in \mathbb{N}\}$ es u.i. entonces $E(|X|) < \infty$ y además $X_n \xrightarrow[n\to\infty]{L^1} X$*

2. *Si $E(|X_n|) < \infty$ para todo $n \geqslant 1$, $E(|X|) < \infty$ y $X_n \xrightarrow[n\to\infty]{L^1} X$ entonces $\{X_n : n \in \mathbb{N}\}$ es u.i.*

Demostración:

1. Sea $\{X_n : n \in \mathbb{N}\}$ u.i. entonces $\sup\limits_{n\in\mathbb{N}} E(|X_n|) < \infty$ por $a)$ de la proposición 2.4.5.

   Como $X_n \xrightarrow[n\to\infty]{P} X$ existe una subsucesión $\{X_{n_k} : k \in \mathbb{N}\}$ tal que $X_{n_k} \xrightarrow[n\to\infty]{c.s.} X$.

   Así, $|X| = \lim\limits_{k\to\infty} |X_{n_k}| = \liminf\limits_{k\to\infty} |X_{n_k}|$.

   Aplicando esperanzas

   $E(|X|) = E\left[\liminf\limits_{k\to\infty} |X_{n_k}|\right] \leqslant \liminf\limits_{k\to\infty} E(|X_{n_k}|) \leqslant \sup\limits_{k\in\mathbb{N}} E(|X_{n_k}|) \leqslant \sup\limits_{n\in\mathbb{N}} E(|X_n|) < \infty$ por el Lema de Fatou y porque el supremo de un subconjunto es menor o igual que el supremo del conjunto.

   Por lo tanto $E(|X|) < \infty$.

Sólo falta ver que $X_n \xrightarrow[n\to\infty]{L^1} X$.

Sea $\epsilon > 0$ entonces $E(|X_n - X|) = \int\limits_{\Omega} |X_n - X| dP =$

$\int\limits_{[|X_n - X| < \epsilon]} |X_n - X| dP + \int\limits_{[|X_n - X| \geqslant \epsilon]} |X_n - X| dP \leqslant \epsilon P[|X_n - X| < \epsilon] +$
$\int\limits_{[|X_n - X| \geqslant \epsilon]} |X_n - X| dP \leqslant \epsilon + \sup\limits_{n\in\mathbb{N}} \int\limits_{[|X_n - X| \geqslant \epsilon]} |X_n - X| dP$ pero como
$|X_n - X| \leqslant |X_n| + |X|$,

tenemos que

$E(|X_n - X|) \leqslant \epsilon + \sup\limits_{n\in\mathbb{N}} \int\limits_{[|X_n - X| \geqslant \epsilon]} |X_n| dP + \int\limits_{[|X_n - X| < \epsilon]} |X_n - X| dP.$

Como $X_n \xrightarrow{P} X$ entonces $P[|X_n - X| \geqslant \epsilon] \xrightarrow[n\to\infty]{} 0$.

Ahora sea $n_0 \in \mathbb{N}$ tal que para todo $n \geqslant n_0$ $P[|X_n - X| \geqslant \epsilon] < \delta$ y por c) de la proposición 2.4.6 $\sup\limits_{n\in\mathbb{N}} \int\limits_{[|X_n - X| \geqslant \epsilon]} |X_n| dP < \epsilon$.

Tomemos a su vez $n_1 \in \mathbb{N}$ tal que para toda $n \geqslant n_1$, $P[|X_n - X| \geqslant \epsilon] < \delta$, implica que $\int\limits_{[|X_n - X| \geqslant \epsilon]} |X_n| dP < \epsilon$.

Así si $n \geqslant \text{máx}\{n_0, n_1\}$ se cumple que $E(|X_n - X|) \leqslant \epsilon + \epsilon + \epsilon = 3\epsilon$.

Por lo tanto $X_n \xrightarrow[n\to\infty]{L^1} X$.

2. Notemos que $E(|X_n - X|) < \infty$ para todo $n \in \mathbb{N}$ porque $|X_n - X| \leqslant |X_n| + |X|$ y entonces $E(|X_n - X|) \leqslant E(|X_n|) + E(|X|) < \infty$.

   Como $X_n \xrightarrow[n\to\infty]{L^1} X$ existe $N \in \mathbb{N}$ tal que para todo $n \geqslant N$ $E(|X_n - X|) < 1$.

   Sea $L = \text{máx}\{E(|X_k - X|) : k = 1, \ldots, N-1\} < \infty$.

   Asi $\sup\limits_{n\in\mathbb{N}} E(|X_n - X|) \leqslant \text{máx}\{L, 1\} < \infty$.

Por lo tanto vale a) de la proposición 2.4.5 para $\{(X_n - X) : n \in \mathbb{N}\}$.

Ahora verificamos c) del mismo resultado con la misma sucesión.

Como $X_n \xrightarrow[n\to\infty]{L^1} X$ entonces $E(|X_n - X|) \xrightarrow[n\to\infty]{} 0$.

Por lo que si $\epsilon > 0$ existe $n_0 = n_0(\epsilon) \in \mathbb{N}$ tal que para todo $n \geqslant n_0$ y $a > 0$

$$\int_{[|X_n - X| \geqslant a]} |X_n - X| dP \leqslant \int_{\Omega} |X_n - X| dP = E(|X_n - X|) < \epsilon.$$

Consideremos ahora números $a_1, \ldots, a_{n_0 - 1}$ tales que

$\displaystyle\int_{[|X_k - X| \geqslant a_k]} |X_k - X| dP < \epsilon$, $k = 1, \ldots, n_0 - 1$ esto es posible debido a

que $Y_k = |X_k - X|$ es integrable y además $\displaystyle\int_{[Y_k \geqslant a_k]} Y_k dP < \epsilon$.

Sea $a = \max\{a_1, a_2, \ldots, a_{n_0 - 1}\}$ asi para todo j se cumple

$$\int_{[|X_j - X| \geqslant a]} |X_j - X| dP < \epsilon.$$

Luego $\{X_n - X : n \in \mathbb{N}\}$ es u.i. pero $\{X : n \in \mathbb{N}\}$ es u.i.,

por lo que $\{(X_n - X) + X : n \in \mathbb{N}\}$ es u.i.

De donde $\{X_n : n \in \mathbb{N}\}$ es u.i. $\square$

Proposición 2.4.9. *Sea* $\varphi \geqslant 0$ *una función Borel medible definida en* $[0, \infty)$ *tal que* $\dfrac{\varphi(x)}{x} \xrightarrow[x\to\infty]{} \infty$.

Sea $\{X_t : t \in T\}$ *una familia de variables aleatorias con* $E(|X_t|) < \infty$ *para toda* $t \in T$ *tal que* $\sup_{t \in T} E[\varphi(|X_t|)] < \infty$.

Entonces $\{X_t : t \in T\}$ *es u.i. (Es otro criterio para verificar u.i.)*

Demostración:

Sea $N \geqslant 1$ entonces existe $a_0 > 0$ tal que $\dfrac{\varphi(x)}{x} \geqslant N$

para todo $x \geqslant a_0$.

Luego $\varphi(|X_t|) \geqslant N|X_t| > 0$ es decir $|X_t| \leqslant \dfrac{\varphi(|X_t|)}{N}$.

Por lo que $\displaystyle\int_{[|X_t| \geqslant a_0]} |X_t| dP \leqslant \frac{1}{N} \int_{[|X_t| \geqslant a_0]} \varphi(|X_t|) dP \leqslant \frac{1}{N} \int_{\Omega} \varphi(|X_t|) dP$ porque $\varphi \geqslant 0$.

Por lo tanto $\displaystyle\int_{[|X_t| \geqslant a_0]} |X_t| dP \leqslant \frac{1}{N} E[\varphi(|X_t|)] \leqslant \frac{1}{N} \sup_{t \in T} E[\varphi(|X_t|)]$.

Notemos que se pide que φ sea Borel medible para que $\varphi(|X_t|)$ sea variable aleatoria.

Por lo tanto $\displaystyle\int_{[|X_t| \geqslant a_0]} |X_t| dP \leqslant \frac{1}{N} \sup_{t \in T} E[\varphi(|X_t|)]$.

Si hacemos tender a_0 a infinito entonces $N \longrightarrow \infty$ porque $N \leqslant \dfrac{\varphi(x)}{x} \xrightarrow[x \to \infty]{} \infty$.

Así $\displaystyle\lim_{a \to \infty} \sup_{t \in T} \int_{[|X_t| \geqslant a]} |X_t| dP \leqslant \lim_{N \to \infty} \frac{1}{N} \sup_{t \in T} E[\varphi(|X_t|)] = 0$,

porque $\displaystyle\sup_{t \in T} E[\varphi(|X_t|)] < \infty$.

De donde $\displaystyle\lim_{a \to \infty} \sup_{t \in T} \int_{[|X_t| \geqslant a]} |X_t| dP = 0$ implica que $\{X_t : t \in T\}$ es u.i. □

Corolario 2.4.10. *Sea $\{X_n : n \in \mathbb{N}\}$ sucesión de variables aleatorias tal que $\displaystyle\sup_{n \in \mathbb{N}} E(|X_n|^p) < \infty$ para $p > 1$. Entonces $\{X_n : n \in \mathbb{N}\}$ es u.i.*

Demostración:

Sea $\varphi(x) = |x|^p$ entonces $\displaystyle\lim_{x \to \infty} \frac{\varphi(x)}{x} = \lim_{x \to \infty} \frac{|x|^p}{x} = \infty$ pues $p > 1$.

$\displaystyle\sup_{n \in \mathbb{N}} E[\varphi(|X_n|)] = \sup_{n \in \mathbb{N}} E[|X_n|^p] < \infty$.

Si $X_n \in L^p$ entonces $X_n \in L^1$ porque $L_p \subseteq L_1$ cuando $1 < p$.

Por lo tanto $E(|X_n|) < \infty$ para todo $n \in \mathbb{N}$.

Por la proposición 2.4.9 $\{X_n : n \in \mathbb{N}\}$ es u.i.

Notar que este resultado vale con T cualquiera en vez de $\mathbb{N}$, es decir, $\{X_t : t \in T\}$ es u.i. $\square$

Teorema 2.4.11. *Sea $\{X_n, \mathcal{F}_n : n \in \mathbb{N}\}$ una submartingala u.i.. Existe una v.a. X con $E(|X|) < \infty$ tal que $X_n \xrightarrow[n\to\infty]{L^1} X$ y $X_n \xrightarrow[n\to\infty]{c.s.} X$. Además $\{X_1, X_2, \ldots, X\}$ es submartingala y en particular si $\{X_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala entonces $\{X_1, X_2, \ldots, X\}$ es martingala.*

Demostración:

Como $\{X_n : n \in \mathbb{N}\}$ es u.i. entonces $\sup\limits_{n\in\mathbb{N}} E(|X_n|) < \infty$ o equivalentemente $\limsup\limits_{n\to\infty} E(|X_n|) < \infty$.

Por el Teorema de Convergencia de Submartingalas existe una v.a. X con $E(|X|) < \infty$ tal que $X_n \xrightarrow{c.s.} X$.

Como convergencia casi segura implica convergencia en probabilidad tenemos que

$X_n \xrightarrow[n\to\infty]{P} X$.

Y por la proposición 2.4.8 $X_n \xrightarrow{L^1} X$ es decir $E(|X_n - X|) \xrightarrow[n\to\infty]{} 0$.

Sean $m \in \mathbb{N}$ y $E \in \mathcal{F}_m$, entonces por la proposición 2.1.8 la sucesión $\left\{ \int\limits_E X_n dP : n \in \mathbb{N} \right\}$ es creciente para $n \geqslant m$.

También $0 \leqslant \left| \int\limits_E X_n dP - \int\limits_E X dP \right| = \left| \int\limits_E (X_n - X) dP \right| \leqslant \int\limits_E |X_n - X| dP \leqslant \int\limits_\Omega |X_n - X| dP = E(|X_n - X|) \xrightarrow[n\to\infty]{} 0$ porque $X_n \xrightarrow[n\to\infty]{L^1} X$.

Por lo tanto $\int\limits_E X_n dP \xrightarrow[n\to\infty]{} \int\limits_E X dP$ para todo $E \in \mathcal{F}_m$ si $m \leqslant n$.

Se sigue que para cada $E \in \mathcal{F}_m$ y $m \leqslant n$ por ser sucesión creciente $\int_E X_m dP \leqslant \int_E X_n dP \leqslant \int_E X dP = \int_E E(X|\mathcal{F}_m)dP$ la última igualdad es por definición de esperanza condicional por lo que $X_m \leqslant E(X|\mathcal{F}_m)$ c.s.

Si existiera $A \in \mathcal{F}_m$ tal que $E(X|\mathcal{F}_m)(\omega) < X_m(\omega)$ con $\omega \in A$ y $P(A) > 0$ entonces $\int_A E(X|\mathcal{F}_m)dP < \int_A X_m dP$

La conclusión por tanto es que $\{X_1, X_2, \ldots, X\}$ es submartingala. □

Corolario 2.4.12. *Sea* $\{X_n, \mathcal{F}_n : n \in \mathbb{N}\}$ *martingala o submartingala no negativa con*

$\sup_{n \in \mathbb{N}} E(|X_n|^p) < \infty$ *para alguna* $p > 1$. *Entonces* X_n *converge casi seguramente y en* L^p *a una v.a.* X *con* $E(|X|^p) < \infty$.

Demostración:

Por el Corolario 2.4.10, $\{X_n : n \in \mathbb{N}\}$ es uniformemente integrable.

Por el Teorema 2.4.11 $X_n \xrightarrow[n\to\infty]{c.s.} X$ y $X_n \xrightarrow{L^1} X$.

Además $\{X_1, X_2, \ldots, X\}$ es martingala y $E(|X|^p) < \infty$ implica que $\{|X_n|^p, \mathcal{F}_n\}$ es submartingala no negativa.

El hecho de que $E(|X|^p) < \infty$ no es inmediato, veamos

$X_n \xrightarrow[n\to\infty]{c.s.} X$ entonces $X_n^p \xrightarrow[n\to\infty]{c.s.} X^p$ asi existe $A \in \mathcal{F}$ con $P(A) = 1$ tal que $X_n^p(\omega) \xrightarrow[n\to\infty]{} X^p(\omega)$ para todo $\omega \in A$.

Luego si $\epsilon > 0$ existe $N \in \mathbb{N}$ tal que $|X_n^p(\omega) - X^p(\omega)| < \epsilon$ para todo $n \geqslant N$.

Asi que $\int_\Omega |X_n^p - X^p|dP = \int_A |X_n^p - X^p|dP + \int_{A^c} |X_n^p - X^p|dP = \int_A |X_n^p - X^p|dP < \int_A \epsilon dP = \epsilon P(A) = \epsilon$.

Entonces $\int\limits_{\Omega} |X_n^p - X^p| dP \xrightarrow[n\to\infty]{} 0.$

Por lo que $E[|X_n^p - X^p|] \xrightarrow[n\to\infty]{} 0.$

Pero por otro lado $|X^p| = |X^p - X_n^p + X_n^p| \leqslant |X_n^p - X^p| + |X_n^p|$.

Asi que $E(|X^p|) \leqslant E(|X_n^p|) + E(|X_n^p - X^p|)$.

Luego $E(|X^p|) \leqslant \sup\limits_{n\in\mathbb{N}} E(|X_n^p|) + \sup\limits_{n\in\mathbb{N}} E(|X_n^p - X^p|) < \infty$.

De este modo $E(|X|^p) < \infty$.

Por otro lado por la desigualdad de Markov $P[|X_n| \geqslant a] \leqslant \dfrac{E(|X_n|)}{a} \leqslant \dfrac{E(|X|)}{a} \xrightarrow[a\to\infty]{} 0.$

Entonces $\sup\limits_{n\in\mathbb{N}} \int\limits_{[|X_n|\geqslant a]} |X_n|^p dP \xrightarrow[a\to\infty]{} 0.$

Asi $\{|X_n|^p : n \in \mathbb{N}\}$ es u.i. y usando un argumento similar al de la proposición 2.4.8 obtenemos que $X_n \xrightarrow[n\to\infty]{L^p} X$. □

El siguiente resultado es el Teorema de Continuidad de Levy para esperanzas condicionales.

Teorema 2.4.13. *Sea X en $(\Omega, \mathcal{F}, P)$ una variable aleatoria tal que $E(|X|) < \infty$ y $\{\mathcal{F}_n : n \in \mathbb{N}\}$ una sucesión de sub $\sigma-$álgebras de $\mathcal{F}$.*

a) Si $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots$ Sea $\mathcal{G} = \sigma\left(\bigcup\limits_{n=1}^{\infty} \mathcal{F}_n\right)$ entonces $E(X|\mathcal{F}_n) \xrightarrow[n\to\infty]{c.s.,L^1} E(X|\mathcal{G})$.

b) Si $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \cdots$ Sea $\mathcal{G} = \bigcap\limits_{n=1}^{\infty} \mathcal{F}_n \neq \varnothing$ entonces $E(X|\mathcal{F}_n) \xrightarrow{c.s.,L^1} E(X|\mathcal{G})$.

Demostración:

Si $\sigma\left(\bigcup\limits_{n=1}^{\infty} \mathcal{F}_n\right) = \mathcal{F}$ entonces $E(X|\mathcal{G}) = E(X|\mathcal{F}) = X$.

a) Sea $X_n = E(X|\mathcal{F}_n)$ $n = 1, 2, \ldots$

Como $E(|X|) < \infty$ entonces

$\{X_n : n \in \mathbb{N}\}$ es martingala por la proposición 2.1.8.

Veremos que $\{X_n : n \in \mathbb{N}\}$ es u.i.

Sea $a > 0$ entonces $P[|X_n| \geqslant a] = P\,[|E(X|\mathcal{F}_n)| \geqslant a]$ pero tenemos por otro lado que $a \leqslant |E(X|\mathcal{F}_n)| \leqslant E[|X||\mathcal{F}_n]$.

Asi

$[|E(X|\mathcal{F}_n)| \geqslant a] \subseteq \left[E(|X||\mathcal{F}_n) \geqslant a\right]$.

De este modo $P[|X_n| \geqslant a] = P\,[|E(X|\mathcal{F}_n)| \geqslant a] \leqslant P\,\left[|E(|X||\mathcal{F}_n)| \geqslant a\right] \leqslant$

$\dfrac{E\left(E(|X||\mathcal{F}_n)\right)}{a} = \dfrac{E(|X|)}{a}$, la última desigualdad es por Markov.

Por lo tanto $P[|X_n| \geqslant a] \leqslant \dfrac{E(|X|)}{a}$.

Entonces $\sup\limits_{n \in \mathbb{N}} P[|X_n| \geqslant a] \leqslant \dfrac{E(|X|)}{a} \xrightarrow[a\to\infty]{} 0$.

Por lo que se satisface $b)$ de la proposición 2.4.5

Ahora veremos que se cumple $c)$ de la proposición 2.4.5 y sabemos que $b)$ y $c)$ implican u.i.

Sea $\epsilon > 0$ como la función $h : \mathcal{F} \longrightarrow \mathbb{R}$ tal que $h(E) = \displaystyle\int_E |X| dP \ll P$ (es absolutamente continua con respecto a P) ya que si $\epsilon > 0$, existe $\delta > 0$ tal que si $P(E) < \delta$ entonces $h(E) = \displaystyle\int_E |X| dP < \epsilon$.

Para tal $\delta > 0$ existe $a_0 = a_0(\delta)$ que cumple con que para todo $a \geqslant a_0$ y para todo $n \geqslant 1$,

$P(|X_n| \geqslant a) \leqslant \dfrac{E(|X|)}{a} < \delta$.

Ahora $\displaystyle\int\limits_{[|X_n|\geqslant a]} |X_n|dP = \int\limits_{[|X_n|\geqslant a]} |E(X|\mathcal{F}_n)|dP \leqslant \int\limits_{[|X_n|\geqslant a]} E(|X||\mathcal{F}_n)dP =$ $\displaystyle\int\limits_{[|X_n|\geqslant a]} |X|dP$ esta última igualdad es por la definición de esperanza condicional.

De aquí $\displaystyle\int\limits_{[|X_n|\geqslant a]} |X_n|dP \leqslant h([|X_n| \geqslant a]) < \epsilon$ si $a \geqslant a_0$.

Entonces $\displaystyle\sup_{n\in\mathbb{N}} \int\limits_{[|X_n|\geqslant a]} |X_n|dP < \epsilon$ si $a \geqslant a_0$.

Por lo tanto $c)$ de la proposición 2.4.5 se cumple.

Asi $\{X_n : n \in \mathbb{N}\}$ es u.i.

Por el Teorema 2.4.11 existe una v.a. X_∞ tal que $E(|X_\infty|) < \infty$ y

$X_n = E(X|\mathcal{F}_n) \xrightarrow[n\to\infty]{L^1} X_\infty$,

$X_n = E(X|\mathcal{F}_n) \xrightarrow[n\to\infty]{c.s.} X_\infty$.

Sólo nos falta mostrar que $X_\infty = E(X|\mathcal{G})$ c.s. donde $\mathcal{G} = \sigma\left(\bigcup\limits_{n=1}^{\infty} \mathcal{F}_n\right)$.

Sea $E \in \bigcup\limits_{n=1}^{\infty} \mathcal{F}_n$ entonces existe $n_0 \geqslant 1$ tal que $E \in \mathcal{F}_{n_0} \subseteq \mathcal{F}_{n_0+1} \subseteq \cdots$ es decir $E \in \mathcal{F}_n$ para todo $n \geqslant n_0$.

Para cada $n \geqslant 1$ si $X_n = E(X|\mathcal{F}_n)$ entonces

$$\left|\int\limits_E X_\infty dP - \int\limits_E E(X|\mathcal{G})dP\right| =$$

$$\left|\int\limits_E X_\infty dP - \int\limits_E X_n dP + \int\limits_E X_n dP - \int\limits_E E(X|\mathcal{G})dP\right| \leqslant \left|\int\limits_E X_\infty dP - \int\limits_E X_n dP\right| +$$

$$\left|\int\limits_E X_n dP - \int\limits_E E(X|\mathcal{G})dP\right| = \left|\int\limits_E (X_\infty - X_n)dP\right| + \left|\int\limits_E X_n dP - \int\limits_E E(X|\mathcal{G})dP\right| \leqslant$$

$$\int_E |X_\infty - X_n| dP + \left| \int_E X_n dP - \int_E E(X|\mathcal{G}) dP \right| \leqslant \int_\Omega |X_\infty - X_n| dP +$$
$$\left| \int_E X_n dP - \int_E E(X|\mathcal{G}) dP \right|,$$
porque $|X_\infty - X_n| \geqslant 0$ y $E \subseteq \Omega$.

Pero para $n \geqslant n_0$ tenemos que
$\int_E X_n dP = \int_E E(X|\mathcal{F}_n) = \int_E X dP$ por definición de esperanza condicional.

Y $\int_E X dP = \int_E E(X|\mathcal{G}) dP$ porque $E \in \mathcal{G}$,

entonces $\int_\Omega |X_\infty - X_n| dP + \left| \int_E X_n dP - \int_E E(X|\mathcal{G}) dP \right| =$

$$\int_\Omega |X_\infty - X_n| dP + \left| \int_E (X_n - E(X|\mathcal{G}))\, dP \right| =$$

$$E(|X_\infty - X_n|) + \left| \int_E (X_n - E(X|\mathcal{G}) dP) \right| \xrightarrow[n \to \infty]{} 0.$$

Por lo tanto $\left| \int_E X_\infty dP - \int_E E(X|\mathcal{G}) dP \right| = 0.$

De aquí $\int_E X_\infty dP = \int_E E(X|\mathcal{G}) dP$ para todo $E \in \bigcup_{n=1}^{\infty} \mathcal{F}_n$.

Pero como $\mathcal{G} = \sigma\left(\bigcup_{n=1}^{\infty} \mathcal{F}_n \right)$ si $A \in \mathcal{G}$ tomamos $A_n \uparrow A$ tal que $A_n \in \bigcup_{n=1}^{\infty} \mathcal{F}_n$ que es un álgebra, por Teorema de extensión se cumple $\int_A X_\infty dP = \int_A E(X|\mathcal{G}) dP$ para $A \in \mathcal{G}$.

Por lo tanto $X_\infty = E(X|\mathcal{G})$ c.s.

b) $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \cdots$ y $\mathcal{G} = \bigcap_{n=1}^{\infty} \mathcal{F}_n \neq \varnothing$.

Si $X_n = E(X|\mathcal{F}_n)$ entonces X_n es $\mathcal{F}_n-$medible.

Luego

$E(X_n|\mathcal{F}_{n+1}, \mathcal{F}_{n+2}, \ldots) = E\left[E(X|\mathcal{F}_n)|\mathcal{F}_{n+1}, \mathcal{F}_{n+2}, \ldots\right]$

$= E\left[E(X|\mathcal{F}_n)|\mathcal{F}_{n+1}\right]$ porque $\sigma(\mathcal{F}_{n+1} \cup \mathcal{F}_{n+2} \cup \ldots) = \mathcal{F}_{n+1}$ ya que $\mathcal{F}_{n+1} \supseteq \mathcal{F}_{n+2} \supseteq \cdots$

Como $\mathcal{F}_n \supseteq \mathcal{F}_{n+1}$ entonces $E\left[E(X|\mathcal{F}_n)|\mathcal{F}_{n+1}\right] = E(X|\mathcal{F}_{n+1}) = X_{n+1}$.

Por lo tanto $E(X_n|\mathcal{F}_{n+1}, \mathcal{F}_{n+2}, \ldots) = X_{n+1}$, esto es, $\{X_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala reversa.

De la misma manera se prueban *b*) y *c*) de la proposición 2.4.5.

De este modo existe una v.a. X_∞ tal que $E(|X_\infty|) < \infty$ con

$$X_n = E(X|\mathcal{F}_n) \xrightarrow[n\to\infty]{\mathcal{L}^1} X_\infty$$

$$\text{y } X_n = E(X|\mathcal{F}_n) \xrightarrow[n\to\infty]{\text{c.s.}} X_\infty.$$

Por último veremos que $X_\infty = E(X|\mathcal{G})$ con $\mathcal{G} = \bigcap_{n=1}^{\infty} \mathcal{F}_n \neq \varnothing$.

Sea $E \in \mathcal{G} = \bigcap_{n=1}^{\infty} \mathcal{F}_n$ entonces $E \in \mathcal{F}_n$ para todo $n \in \mathbb{N}$.

$$\left|\int_E X_\infty dP - \int_E E(X|\mathcal{G})dP\right| \leqslant \int_\Omega |X_\infty - X_n|dP + \left|\int_E (X_n - E(X|\mathcal{G}))\, dP\right| \xrightarrow[n\to\infty]{} 0$$

pues $X_n \xrightarrow[n\to\infty]{\mathcal{L}^1} X_\infty$ y también $\int_E X_n dP = \int_E E(X|\mathcal{F}_n)dP = \int_E X dP = \int_E E(X|\mathcal{G})dP$ ya que $E \in \mathcal{G}$.

Por lo tanto $\int\limits_E X_\infty dP = \int\limits_E E(X|\mathcal{G})dP$ para cada $E \in \bigcap_{n=1}^{\infty} \mathcal{F}_n = \mathcal{G}$.

Asi $X_\infty = E(X|\mathcal{G})$. $\square$

Corolario 2.4.14. *$\{X_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala u.i. si y sólo si existe una v.a. X con $E(|X|) < \infty$ tal que $X_n = E(X|\mathcal{F}_n)$ c.s. $n = 1, 2, \ldots$ En este caso $X_n \xrightarrow[n\to\infty]{c.s., L^1} E(X|\mathcal{G})$ con $\mathcal{G} = \sigma\left(\bigcup_{n=1}^{\infty} \mathcal{F}_n\right)$.*

Demostración:

Si X es una v.a. con $E(|X|) < \infty$ y $X_n = E(X|\mathcal{F}_n)$ entonces

$\{X_n : n \in \mathbb{N}\}$ es martingala por la proposición 2.1.8.

Es decir

$\{X_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala y por el Teorema 2.4.13 se tiene que

$X_n \xrightarrow[n\to\infty]{L^1} E(X|\mathcal{G})$ y $X_n \xrightarrow[n\to\infty]{c.s.} E(X|\mathcal{G})$ con $\mathcal{G} = \sigma\left(\bigcup_{n=1}^{\infty} \mathcal{F}_n\right)$.

Ahora supongamos que $\{X_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala uniformemente integrable entonces por el Teorema 2.4.11 existe una v.a. X con $E(|X|) < \infty$ tal que $X_n \xrightarrow[n\to\infty]{c.s.} X$ y $X_n \xrightarrow[n\to\infty]{L^1} X$.

Además $\{X_1, X_2, \ldots, X\}$ es martingala u.i. entonces $E(X|\mathcal{F}_n) = X_n$. $\square$

Corolario 2.4.15. *La sucesión $\{X_n : n \in \mathbb{N}\}$ es martingala reversa si y sólo si*

$E(X_n|X_{n+k}, X_{n+k+1}, \ldots, X_{n+k+m}) = X_{n+k}$ c.s. para todas $n, m, k \in \mathbb{N}$.

Demostración:

Sea $\{X_n : n \in \mathbb{N}\}$ martingala reversa.

Por definición

$E(X_n|X_{n+k}, X_{n+k+1}, \ldots) = X_{n+k}$ c.s.

Luego $E\left[E(X_n|X_{n+k},X_{n+k+1},\ldots)|X_{n+k},X_{n+k+1},\ldots,X_{n+k+m}\right] =$
$E(X_{n+k}|X_{n+k},X_{n+k+1},\ldots,X_{n+k+m})$ que resulta de aplicar esperanzas condicionales con respecto al $\sigma-$álgebra $\sigma(X_{n+k},X_{n+k+1},\ldots,X_{n+k+m})$.
Pero, por un lado $E\left[E(X_n|X_{n+k},X_{n+k+1},\ldots)|X_{n+k},\ldots,X_{n+k+m}\right] =$
$E(X_n|X_{n+k},X_{n+k+1},\ldots,X_{n+k+m})$ debido a que
$\sigma(X_{n+k},\ldots,X_{n+k+m}) \subseteq \sigma(X_{n+k},X_{n+k+1},\ldots)$.
Mientras que por otro lado $E(X_{n+k}|X_{n+k},X_{n+k+1},\ldots,X_{n+k+m}) = X_{n+k}$ asi
$E(X_n|X_{n+k},X_{n+k+1},\ldots,X_{n+k+m}) = X_{n+k}$ c.s.
Para la otra implicación si $E(X_n|X_{n+k},\ldots,X_{n+k+m}) = X_{n+k}$ c.s., entonces por el Teorema 2.4.13 se tiene que
$\lim\limits_{m\to\infty} E(X_n|X_{n+k},\ldots,X_{n+k+m}) = E(X_n|X_{n+k},X_{n+k+1},\ldots) = X_{n+k}$ c.s.
Por lo tanto $\{X_n : n \in \mathbb{N}\}$ es martingala reversa.
Notemos que la última igualdad se obtendría más claramente si definimos
$\mathcal{F}_0 = \sigma(X_{n+k}), \mathcal{F}_1 = \sigma(X_{n+k},X_{n+k+1}),\ldots,$
$\mathcal{F}_m = \sigma(X_{n+k},X_{n+k+1},\ldots,X_{n+k+m})$, $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_m$ y por el Teorema 2.4.13 $E(X_n|\mathcal{F}_m) \xrightarrow[m\to\infty]{c.s.} E(X_n|\mathcal{F})$,
donde
$$\mathcal{F} = \sigma\left(\bigcup_{m=0}^{\infty} \mathcal{F}_m\right) = \sigma\left(\bigcup_{m=0}^{\infty} \sigma(X_{n+k},\ldots,X_{n+k+m})\right) = \sigma(X_{n+k},X_{n+k+1},\ldots). \qquad \square$$
Existen multiples aplicaciones de la Teoría de Martingalas en diversos campos y problemas de probabilidad, a continuación estudiaremos una Ley Fuerte para v.a.i.i.d. y una Ley $0-1$.

Proposición 2.4.16 (Ley 0-1). *El $\sigma-$álgebra cola de una sucesión $\{X_n : n \in \mathbb{N}\}$ de v.a.i. definidas en $(\Omega,\mathcal{F},P)$ es el $\sigma-$álgebra trivial $\{\varnothing,\Omega\}$.*

Demostración:

Por definición el $\sigma-$álgebra cola es $\tau = \bigcap_{n=0}^{\infty}(X_{n+1}, X_{n+2}, \ldots)$ entonces probaremos que $\tau = \{\varnothing, \Omega\}$.

Una contención ya está dada, por lo que comprobaremos sólo la otra contención.

Sea $A \in \tau$ entonces $A \in \sigma(X_{n+1}, X_{n+2}, \ldots)$ para todo $n \geqslant 0$.Luego I_A y $(X_1, \ldots, X_n)$ son independientes para todo $n \geqslant 1$ porque $\sigma(X_{n+1}, X_{n+2}, \ldots)$ y $\sigma(X_1, \ldots, X_n)$ son independientes. Asi para todo n tenemos que $E(I_A) = E(I_A|X_1, \ldots, X_n) = E(I_A|\mathcal{F}_n)$ donde $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$.

Como $E(|I_A|) = P(A) < \infty$ ocurre por el Teorema 2.4.13 que

$E(I_A|\mathcal{F}_n) \xrightarrow[n \to \infty]{c.s.} E(I_A|\mathcal{G})$ donde $\mathcal{G} = \sigma\left(\bigcup_{n=1}^{\infty} \mathcal{F}_n\right) = \sigma(X_1, X_2, \ldots)$.

Pero I_A es $\sigma(X_1, X_2, \ldots)$ medible porque $A \in \sigma(X_1, X_2, \ldots)$.

Por tanto $E(I_A|X_1, X_2, \ldots) = I_A$.

De este modo $E(I_A) \xrightarrow[n \to \infty]{c.s.} I_A$ por lo que

$E(I_A) = I_A$.

Asi $P(A) = I_A$ pero $I_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases}$

Se debe cumplir por tanto que I_A es constante.

Entonces $A = \varnothing$ ó $A = \Omega$. □

Proposición 2.4.17 (Ley Fuerte). *Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de v.a.i.i.d. con $E(X_n) = \mu$ para todo $n \in \mathbb{N}$, $S_n = \sum_{k=1}^{n} X_k$ entonces $\frac{S_n}{n} \xrightarrow{c.s., L^1} \mu$.*

Demostración:

Se sabe que $E(X_1|S_n, S_{n+1}, \ldots) = E(X_1|S_n) = \frac{S_n}{n}$ c.s.

Si τ es el $\sigma-$álgebra cola de $\{S_n : n \in \mathbb{N}\}$ entonces por el Teorema 2.4.13 tenemos

$E(X_1|S_n, S_{n+1}, \ldots) = \frac{S_n}{n} \xrightarrow{c.s.,\mathcal{L}^1} E(X_1|\tau)$ ya que $\tau = \bigcap_{n=1}^{\infty} \sigma(S_n, S_{n+1}, \ldots)$.

Pero $\lim_{n\to\infty} \frac{S_n}{n}$ es una función cola ya que es $\tau-$medible.

Por lo tanto $\lim_{n\to\infty} \frac{S_n}{n}$ es constante c.s.

Asi $E(X_1|\tau)$ es constante.

Digamos que $E(X_1|\tau) = c$ asi

$E\left(E(X_1|\tau)\right) = E(c) = c.$

Luego $E(X_1) = c$ implica $c = \mu$.

De este modo $\frac{S_n}{n} \xrightarrow[n\to\infty]{c.s.,\mathcal{L}^1} \mu$. $\square$

Proposición 2.4.18. *Sean $X, X_1, X_2, \ldots$ variables aleatorias definidas en $(\Omega, \mathcal{F}, P)$ tal que $E(|X|) < \infty$ entonces $E(X|X_1, \ldots, X_n) \xrightarrow[n\to\infty]{c.s.,\mathcal{L}^1} E(X|X_1, X_2, \ldots)$. Si para algún $p > 1$, $E(|X|^p) < \infty$ entonces $E(X|X_1, \ldots, X_n) \xrightarrow[n\to\infty]{c.s.,\mathcal{L}^p} E(X|X_1, X_2, \ldots)$. En particular si X es $\sigma(X_1, X_2, \ldots)-$medible entonces $E(X|X_1, X_2, \ldots) = X$ c.s.*

Demostración:

Definimos $Y_n = E(X|X_1, \ldots, X_n)$ $n = 1, 2, \ldots$

Si $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ entonces $\mathcal{F}_n \uparrow \mathcal{G} = \sigma\left(\bigcup_{n=1}^{\infty} \mathcal{F}_n\right) = \sigma(X_1, X_2, \ldots)$.

$Y_n \xrightarrow[n\to\infty]{c.s.,\mathcal{L}^1} E(X|\mathcal{G}) = E(X|X_1, X_2, \ldots)$ por el Teorema 2.4.13.

Ahora supongamos que para algún $p > 1$, $E(|X|^p) < \infty$.

Como $\{Y_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala debido a que

$E(Y_n|\mathcal{F}_{n-1}) = E\left[E(X|X_1, \ldots, X_n)|\mathcal{F}_{n-1}\right] = E(X|\mathcal{F}_{n-1}) = Y_{n-1}$ $(\mathcal{F}_{n-1} \subseteq$

$\mathcal{F}_n$) tenemos que $\{|Y_n|^p, \mathcal{F}_n : n \in \mathbb{N}\}$ es submartingala no negativa.

Debemos verificar que $E(|Y_n|^p) < \infty$.

$E(|Y_n|^p) = E\left[|E(X|X,\ldots,X_n)|^p\right] \leqslant E\left[E[|X|^p|X_1,\ldots,X_n]\right]$ por desigualdad de Jensen ya que $h(x) = |x|^p$ es convexa.

Por lo tanto $E[|Y_n|^p] \leqslant E(|X|^p) < \infty$ por hipótesis.

Asi $\{|Y_n|^p, \mathcal{F}_n : n \in \mathbb{N}\}$ es submartingala no negativa y por el Corolario 2.4.12 tenemos que $Y_n \xrightarrow[n\to\infty]{c.s., L^1} E(X|X_1, X_2, \ldots)$. $\square$

Corolario 2.4.19 (Ley 0-1 extendida). *Sea $E \in \sigma(X_1, X_2, \ldots)$. La sucesión de probabilidades condicionales $P(E|X_1,\ldots,X_n) \xrightarrow[n\to\infty]{c.s.} I_E$.*

Demostración:

Sea $X = I_E$ por la proposición 2.4.18,

$P(E|X_1,\ldots,X_n) = E(I_E|X_1,\ldots,X_n) = E(X|X_1,\ldots,X_n) \xrightarrow{c.s.} E(X|X_1,X_2,\ldots) = E(I_E|X_1,X_2,\ldots) = I_E$.

Notemos que $E(|X|) = P(E) < \infty$ y que I_E es $\sigma(X_1, X_2, \ldots)$ medible. $\square$

Proposición 2.4.20. *Sean $\{X_n : n \in \mathbb{N}\}$ una sucesión de v.a. c.s. no negativas y uniformemente acotadas sobre $(\Omega, \mathcal{F}, P)$, $\{\mathcal{F}_n : n \in \mathbb{N}\}$ una sucesión creciente de sub $\sigma-$álgebras de $\mathcal{F}$ tal que X_n es $\mathcal{F}_n-$medible para todo $n \in \mathbb{N}$ entonces $\sum_{n=1}^{\infty} X_n$ y $\sum_{n=1}^{\infty} E(X_n|\mathcal{F}_{n-1})$ ambas convergen o divergen casi seguramente. $\mathcal{F}_0 = \{\varnothing, \Omega\}$.*

Demostración:

Sean $Y_k = X_k - E(X_k|\mathcal{F}_{k-1})$ para $k \geqslant 1$.

$S_n = \sum_{k=1}^{n} Y_k$ entonces

$E(Y_k) = E(X_k) - E(X_k) = 0$ y $E(S_n) = \sum_{k=1}^{n} E(Y_k) = 0.$

Por otro lado $E(S_n|\mathcal{F}_{n-1}) = E\left[\sum_{k=1}^{n} Y_k|\mathcal{F}_{n-1}\right] = \sum_{k=1}^{n} E(Y_k|\mathcal{F}_{n-1}) =$

$\sum_{k=1}^{n} E\left[(X_k - E(X_k|\mathcal{F}_{k-1}))|\mathcal{F}_{n-1}\right] = \sum_{k=1}^{n}\{E(X_k|\mathcal{F}_{n-1}) - E\left[E(X_k|\mathcal{F}_{k-1})|\mathcal{F}_{n-1}\right]\} =$

$\sum_{k=1}^{n}\{E(X_k|\mathcal{F}_{n-1}) - E(X_k|\mathcal{F}_{k-1})\}$ porque $\mathcal{F}_{k-1} \subseteq \mathcal{F}_{n-1}$ para $k = 1, 2, \ldots, n.$

Y como $X_1, X_2, \ldots, X_{n-1}$ son $\mathcal{F}_{n-1}$ medibles tenemos que

$E(S_n|\mathcal{F}_{n-1}) = \sum_{k=1}^{n} E(X_k|\mathcal{F}_{n-1}) - \sum_{k=1}^{n} E(X_k|\mathcal{F}_{k-1}) = (X_1 + \cdots + X_{n-1}) +$

$E(X_n|\mathcal{F}_{n-1}) - [E(X_1|\mathcal{F}_0) + \cdots + E(X_n|\mathcal{F}_{n-1})] = \sum_{k=1}^{n-1}(X_k - E(X_k|\mathcal{F}_{k-1})) =$

$\sum_{k=1}^{n-1} Y_k = S_{n-1}.$

Por lo que $\{S_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala, por el Teorema 2.3.14 (Convergencia de Submartingalas).

Como $\limsup_{n\to\infty} E(|S_n|) < \infty$ existe una v.a. X que es $\sigma\left(\bigcup_{n=1}^{\infty} \sigma(S_n)\right)$ $-$medible tal que $S_n \xrightarrow[n\to\infty]{c.s.} X$.

Además $E(|X|) \leqslant \limsup_{n\to\infty} E(|S_n|) < \infty$ con $P[\lim_{n\to\infty} S_n = \infty] = 0$

$= P[\lim_{n\to\infty} S_n = -\infty].$

Pero notemos que

$S_n = \sum_{k=1}^{n} Y_k = \sum_{k=1}^{n}(X_k - E(X_k|\mathcal{F}_{k-1})) = \sum_{k=1}^{n} X_k - \sum_{k=1}^{n} E(X_k|\mathcal{F}_{k-1}).$

Por lo tanto $P\left[\lim_{n\to\infty}\left(\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} E(X_k|\mathcal{F}_{k-1})\right) = \infty\right] = 0 =$

$P\left[\lim_{n\to\infty}\left(\sum_{k=1}^{n} X_k - \sum_{k=1}^{n} E(X_k|\mathcal{F}_{k-1})\right) = -\infty\right].$

Luego

$$P\left[\sum_{n=1}^{\infty} X_n - \sum_{n=1}^{\infty} E(X_n|\mathcal{F}_{n-1}) = \infty\right] = 0 = P\left[\sum_{n=1}^{\infty} X_n - \sum_{n=1}^{\infty} E(X_n|\mathcal{F}_{n-1}) = -\infty\right]$$

Entonces

$$P\left[\sum_{n=1}^{\infty} X_n = \infty, \sum_{n=1}^{\infty} E(X_n|\mathcal{F}_{n-1}) < \infty\right] = 0 = P\left[\sum_{n=1}^{\infty} X_n < \infty, \sum_{n=1}^{\infty} E(X_n|\mathcal{F}_{n-1}) = \infty\right].$$

De este modo, las series $\sum_{n=1}^{\infty} X_n$ y $\sum_{n=1}^{\infty} E(X_n|\mathcal{F}_{n-1})$ ambas convergen o ambas divergen. □

Corolario 2.4.21 (Lema de Borel-Cantelli extendido). *Sea $\{\mathcal{F}_n : n \in \mathbb{N}\}$ una sucesión creciente de sub $\sigma-$álgebras de $\mathcal{F}$, $A_n \in \mathcal{F}_n$ $n = 1, 2, \ldots$ $p = P(A_1)$ y para $n \geqslant 2$ $p_n = P(A_n|\mathcal{F}_{n-1})$.*
Entonces $P\left[\left(\limsup_{n\to\infty} A_n\right) \triangle \left(\sum_{n=1}^{\infty} p_n = \infty\right)\right] = 0$, *es decir,*
$P\left[\left(\limsup_{n\to\infty} A_n\right)\right] = 1$ *si y sólo si* $P\left[\{\omega \in \Omega : \sum_{n=1}^{\infty} p_n = \infty\}\right] = 1$ *o equivalentemente, A_n ocurre infinitamente seguido si y sólo si* $\sum_{n=1}^{\infty} p_n = \infty$.

Demostración:

Sea $X_n = I_{A_n}$ para todo $n \geqslant 1$ entonces $\{X_n : n \in \mathbb{N}\}$ es una sucesión de v.a. no negativas en donde además $|X_n| \leqslant 1$ para todo $n \in \mathbb{N}$.
X_n es $\mathcal{F}_n-$medible, la sucesión es uniformemente acotada.
Por la proposición 2.4.20, $\sum_{n=1}^{\infty} X_n$ y $\sum_{n=1}^{\infty} E(X_n|\mathcal{F}_{n-1})$ ambas convergen o ambas divergen c.s.
Esto es, $\sum_{n=1}^{\infty} I_{A_n}$ y $\sum_{n=1}^{\infty} P(A_n|\mathcal{F}_{n-1}) = \sum_{n=1}^{\infty} p_n$ ambas convergen o ambas divergen c.s.
Pero, $\{\omega \in \Omega : \sum_{n=1}^{\infty} I_{A_n}(\omega) = \infty\} = \{\omega \in \Omega : \omega \in A_n \ i.o.\} = \limsup_{n\to\infty} A_n =$

$\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k$.

Por lo tanto $P\left[\sum_{n=1}^{\infty} p_n = \infty\right] = 1$ si y sólo si $P\left[\limsup_{n\to\infty} A_n\right] = 1$.

Así $P\left[\left(\limsup_{n\to\infty} A_n\right) \triangle \left(\sum_{n=1}^{\infty} p_n = \infty\right)\right] = 0.$ □

Proposición 2.4.22. *Sean $\{X_n : n \in \mathbb{N}\}$ una sucesión de v.a.i. $S_n = \sum_{k=1}^{n} X_k$ $n \geqslant 1$. La sucesión $\{S_n : n \in \mathbb{N}\}$ converge en ley si y sólo si $\{S_n : n \in \mathbb{N}\}$ convergen casi seguramente.*

Demostración:

Recordemos que convergencia en Ley significa convergencia de las funciones de distribución y que convergencia casi segura siempre implica convergencia en Ley.

Ahora supongamos que $S_n \xrightarrow[n\to\infty]{L} S$.

Veremos que $S_n \xrightarrow[n\to\infty]{c.s.} S$.

Sea $\varphi_j(t) = E\left(e^{itX_j}\right)$ la función característica de X_j entonces

$\prod_{j=1}^{n} \varphi_j(t) = E\left(e^{itS_n}\right)$ es la función característica de S_n.

Sabemos que

$\prod_{j=1}^{n} \varphi_j(t) \xrightarrow[n\to\infty]{} \varphi(t)$ donde φ es función característica.

Así, existe $t_0 > 0$ tal que $\varphi(t) \neq 0$ con $t \in (-t_0, t_0)$ pues $\varphi(0) = 1$ y φ es continua en cero.

Como $\varphi = \prod_{n=1}^{\infty} \varphi_n$ entonces $\varphi_n(t) \neq 0$ para $t \in (-t_0, t_0)$ para todo $n \geqslant 1$.

Definamos ahora $Z_n = \dfrac{e^{itS_n}}{\prod_{j=1}^{n} \varphi_j(t)}$ una sucesión de v.a. complejas.

$$E(Z_n) = \frac{1}{\prod_{j=1}^{n} \varphi_j(t)} E\left(e^{itS_n}\right) = \frac{\prod_{j=1}^{n} \varphi_j(t)}{\prod_{j=1}^{n} \varphi_j(t)} = 1.$$

$|Z_n| = \dfrac{\left|e^{itS_n}\right|}{\prod_{j=1}^{n} |\varphi_j(t)|} = \dfrac{1}{\prod_{j=1}^{n} |\varphi_j(t)|} \leqslant \dfrac{1}{|\varphi(t)|} \leqslant M$ para $t \in (-t_0, t_0)$ con $\varphi(t) = \prod_{j=1}^{\infty} \varphi_j(t)$.

Pues de lo contrario existiría $t_1 \in (-t_0, t_0)$ tal que $\varphi(t_1) = 0$, que como ya vimos no puede ocurrir.

Si $\mathcal{F}_n = \sigma(S_1, \ldots, S_n)$ entonces $\{Z_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala pues

$$E(Z_n|\mathcal{F}_{n-1}) = E\left(\frac{e^{itS_n}}{\prod_{j=1}^{n} \varphi_j(t)}\middle|S_1, \ldots, S_{n-1}\right) = \frac{1}{\prod_{j=1}^{n} \varphi_j(t)} E\left(e^{itS_{n-1}}e^{itX_n}\middle|S_1, \ldots, S_{n-1}\right) =$$

$$\frac{e^{itS_{n-1}}}{\prod_{j=1}^{n} \varphi_j(t)} E\left(e^{itX_n}|S_1, \ldots, S_{n-1}\right) = \frac{e^{itS_{n-1}}}{\prod_{j=1}^{n} \varphi_j(t)} E\left(e^{itX_n}\right) \text{ por independencia.}$$

Así $E(Z_n|\mathcal{F}_{n-1}) = \dfrac{e^{itS_{n-1}}\varphi_n(t)}{\varphi_1(t)\cdots\varphi_{n-1}(t)\varphi_n(t)} = \dfrac{e^{itS_{n-1}}}{\prod_{j=1}^{n-1} \varphi_j(t)} = Z_{n-1}$.

Notemos que $Re(Z_n)$ e $Im(Z_n)$ son martingalas puesto que

$E(Z_n|\mathcal{F}_{n-1}) = E[Re(Z_n)+iIm(Z_n)|\mathcal{F}_{n-1}] = E[Re(Z_n)|\mathcal{F}_{n-1}]+iE[Im(Z_n)|\mathcal{F}_{n-1}] = Z_{n-1} = Re(Z_{n-1}) + iIm(Z_{n-1})$.

Entonces $E[Re(Z_n)|\mathcal{F}_{n-1}] = Re(Z_{n-1}),\ E[Im(Z_n)|\mathcal{F}_{n-1}] = Im(Z_{n-1})$.

Por el Teorema 2.3.14 (Convergencia de Submartingalas) existe una v.a. Z

con $E(|Z|) < \infty$ tal que $Z_n \xrightarrow[n\to\infty]{\text{c.s.}} Z$, es decir, $\dfrac{e^{itS_n}}{\prod_{j=1}^{n} \varphi_j(t)} \xrightarrow[n\to\infty]{\text{c.s.}} Z$.

Pero $\lim\limits_{n\to\infty} \dfrac{e^{itS_n}}{\prod_{j=1}^{n} \varphi_j(t)} = \dfrac{\lim\limits_{n\to\infty} e^{itS_n}}{\lim\limits_{n\to\infty} \prod_{j=1}^{n} \varphi_j(t)} = \dfrac{\lim\limits_{n\to\infty} e^{itS_n}}{\varphi(t)}$.

Por lo tanto $e^{itS_n} \xrightarrow[n\to\infty]{\text{c.s.}} \varphi(t)Z$.

Consideremos el mapeo $e^{itS_n(\omega)} : (-t_0, t_0) \times \Omega \longrightarrow \mathbb{C}$ con la medida producto $\lambda \times P$, λ la medida de Lesbegue en $(-t_0, t_0)$.

Este mapeo es medible para todo $n \geqslant 1$ por ser continuo. Definimos $A = \{(t, \omega) : \lim\limits_{n\to\infty} e^{itS_n(\omega)} \text{ existe}\}$, A es medible porque para todo $t \in (-t_0, t_0)$ el conjunto $A(t) = \{\omega : \lim\limits_{n\to\infty} e^{itS_n(\omega)} \text{ existe }\}$ tiene probabilidad 1 y el conjunto $A(\omega) = \{t \in (-t_0, t_0) : \lim\limits_{n\to\infty} e^{itS_n(\omega)} \text{ existe }\}$ tiene medida $2t_0$.

Luego hay un $\Omega^* \subseteq \Omega$ con $P(\Omega^*) = 1$ tal que para todo $\omega \in \Omega^*$ existe $E_\omega \subseteq (-t_0, t_0)$ con $\lambda(E_\omega) = 2t_0 > 0$.

Entonces $\lim\limits_{n\to\infty} E\left(e^{itS_n(\omega)}\right)$ existe para $t \in E_\omega$.

Por lo tanto $S_n(\omega)$ converge para todo $\omega \in \Omega^*$.

De este modo S_n converge c.s. $\square$

2.5. Tiempos de Paro

Ahora se establecerá un nuevo concepto que tiene que ver con el momento en que un jugador que apuesta sucesivamente sobre los valores de una sucesión de variables aleatorias que son los resultados del juego y cuya fortuna depende de estas, decide parar, esto es, decide dejar de jugar.

Definición 2.5.1. Sean $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}$ una sucesión creciente de sub $\sigma-$álgebras. Un tiempo de paro con respecto a $\mathcal{F}_n$ es una v.a. N tal que $P(N < \infty) = 1$ y el evento $[N = n] \in \mathcal{F}_n$ para todo $n \geqslant 1$. Si N es una v.a. extendida, esto es, $P(N = \infty) > 0$ y $[N = n] \in \mathcal{F}_n$ para todo $n \geqslant 1$ entonces N se llama regla de paro.

Notemos que la familia de tiempos paro depende de la familia de $\sigma-$álgebras.

Ejemplo 2.5.2. $N = c$ con $c \in \mathbb{N}$ es tiempo de paro porque $[N = n] = \begin{cases} \varnothing & \text{si } n \neq c \\ \Omega & \text{si } n = c \end{cases}$ y tanto $\varnothing$ como Ω están en $\mathcal{F}_n$ para todo $n \in \mathbb{N}$. Aquí N es un tiempo de paro para toda familia de $\sigma-$álgebras ya que N es determinista. ■

Ejemplo 2.5.3. Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de v.a. definidas en $(\Omega, \mathcal{F}, P)$, $\{\mathcal{F}_n : n \in \mathbb{N}\}$ una sucesión creciente de sub $\sigma-$álgebras de $\mathcal{F}$ tal que X_n es $\mathcal{F}_n-$medible para todo $n \geqslant 1$.

Si $B \in \mathcal{B}(\mathbb{R})$ fijo definimos

$$N_B(\omega) = \begin{cases} \inf\{n : X_n(\omega) \in B\} & \text{si existe } n \in \mathbb{N} \text{ tal que } X_n(\omega) \in B \\ \infty & \text{si } X_n(\omega) \notin B \text{ para todo } n \in \mathbb{N}. \end{cases}$$

es decir, N_B es el tiempo en que el proceso está por primera vez en B, que se define como ∞ si el proceso nunca está en B.

N_B es una regla de paro.

$[N_B = n] = [X_1 \notin B, X_2 \notin B, \ldots, X_{n-1} \notin B, X_n \in B] \in \mathcal{F}_n$ y para $n = \infty$: $[N_B = \infty] = \bigcap_{i=1}^{\infty}[X_i \notin B] = \Omega - [N_B < \infty] \in \mathcal{F}$. ■

Ejemplo 2.5.4. En el ejemplo 2.1.2 se considero la estrategia de doblar la apuesta en cada paso.

El jugador gana o pierde con probabilidad $\frac{1}{2}$ en cada juego.

Inicia con S1 y dobla la apuesta hasta ganar.

Sea S_n la ganancia neta después de n juegos.

$P[S_1 = 1] = \frac{1}{2} = P[S_1 = -1]$ y para toda $n \geqslant 1$ $P[S_{n+1} = 1|S_n = 1] = 1$, $P[S_{n+1} = 1|S_n = 1 - 2^n] = \frac{1}{2} = P[S_{n+1} = 1 - 2^{n+1}|S_n = 1 - 2^n]$ porque si pierde los primeros n juegos tiene una pérdida de $1+2+2^2+\cdots+2^{n-1} = 2^n-1$, esto es, una ganancia de $1 - 2^n$.

Pero al juego $(n+1)$ apuesta 2^n y de ganar tiene una ganancia $2^n+(1-2^n) = 1$, entonces $E[S_{n+1}|\mathcal{F}_n] = E(S_{n+1}|X_1, \ldots, X_n) = E(S_{n+1}|S_n)$.

Luego $E(S_{n+1}|S_n = 1) = 1 \cdot P[S_{n+1}|S_n = 1] = 1$ y $E(S_{n+1}|S_n = 1 - 2^n) = \frac{1}{2}[1 - 2^{n+1}] + \frac{1}{2} \cdot 1 = \frac{1}{2}(2 - 2^{n+1}) = 1 - 2^n$.

Por lo tanto $E[S_{n+1}|S_n] = S_n$.

Sea $N = \inf\{n \geqslant 1 : S_n = 1\}$ = Tiempo en que el jugador gana.

Como $P[N = n] = \frac{1}{2}(1 - \frac{1}{2})^{n-1}$ $n = 1, 2, \ldots$ entonces $P[N < \infty] = 1$.

De este modo N es un tiempo de paro.

Es claro que $[N = n] \in \mathcal{F}_n$.

Además $P[S_N = 1] = 1$, $E(S_N) = 1 \geqslant E(S_1) = 0$.

Bajo esta estrategia el jugador no puede perder.

$\{S_n : n \in \mathbb{N}\}$ es uniformemente acotada por arriba por 1 pero no lo es por abajo.

$\{S_n : n \in \mathbb{N}\}$ no es u.i.

Para $n \geqslant 1$ sea S_n una v.a. $\mathcal{F}_n$−medible definida en Ω y N un tiempo de paro.

Se define S_N como $S_N = \sum_{n=1}^{\infty} S_n I_{[N=n]}$ para cada $B \in \mathcal{B}(\mathbb{R})$

$[S_N \in B] =$

$\bigcup_{n=1}^{\infty} [S_n \in B, N = n] \in \mathcal{F}$ por lo que S_N es variable aleatoria.

Con el tiempo de paro N se tiene un σ−álgebra $\mathcal{F}_N$ asociado a N que es $\mathcal{F}_N = \{A \in \mathcal{F} : A \cap [N = n] \in \mathcal{F}_n$ para todo $n \in \mathbb{N}\}$.

i) $\varnothing \in \mathcal{F}_N$ porque $\varnothing \in \mathcal{F}$ y además $\varnothing \cap [N = n] = \varnothing \in \mathcal{F}_n$ para todo n.
$\Omega \in \mathcal{F}_N$ pues $\Omega \in \mathcal{F}$ y más aún $\Omega \cap [N = n] = [N = n] \in \mathcal{F}_n$ para todo n.

ii) Sea $A \in \mathcal{F}_N$ entonces $A \in \mathcal{F}$ y $A \cap [N = n] \in \mathcal{F}_n$ para todo n.
Pero $[N = n] \in \mathcal{F}_n$ y $[N = n] = \{[N = n] \cap A\} \cup \{[N = n] \cap A^c\}$.
Luego $[N = n] \cap A^c = [N = n] - A \cap [N = n]$ asi $A^c \cap [N = n] \in \mathcal{F}_n$ para todo n.

iii) Sean $A_1, A_2, \ldots \in \mathcal{F}_N$ entonces $A_k \cap [N = n] \in \mathcal{F}_n$ para todo n y para todo k.
Luego $\left(\bigcup_{k=1}^{\infty} A_k\right) \cap [N = n] \in \mathcal{F}_n$ para todo n asi $\bigcup_{k=1}^{\infty} A_k \in \mathcal{F}_N$.
Si existe $n_0 \in \mathbb{N}$ tal que $[N = n_0] = \Omega$ tenemos que $\mathcal{F}_N = \mathcal{F}_{n_0}$, es decir, N es constante.
Si $B \in \mathcal{F}_{n_0}$ sucede que $B \cap [N = n_0] = B \cap \Omega = B \in \mathcal{F}_N$.

Si $n \neq n_0$ tenemos que $B \cap [N = n] = B \cap \varnothing = \varnothing \in \mathcal{F}_N$.

Si $B \in \mathcal{F}_N$ entonces para todo $n \in \mathbb{N}$ $B \cap [N = n] \in \mathcal{F}_n$ pero

$$B \cap [N = n] = \begin{cases} \varnothing & \text{si } n \neq n_0 \\ B & \text{si } n = n_0 \end{cases} \in \mathcal{F}_{n_0}. \quad \blacksquare$$

Proposición 2.5.5. *Las variables aleatorias N y S_N son $\mathcal{F}_N$ medibles.*

Demostración:

1. Debemos mostrar que $[N = m] \in \mathcal{F}_N$ para todo $m \geqslant 1$ pero,
   $[N = m] \cap [N = n] = \begin{cases} [N = n] & \text{si } m = n \\ \varnothing & \text{si } m \neq n. \end{cases}$
   Como $\varnothing \in \mathcal{F}_n$ para todo n y $[N = n] \in \mathcal{F}_n$ para todo n se tiene que $[N = m] \in \mathcal{F}_N$.
   Por lo tanto N es $\mathcal{F}_N$—medible.

2. Tomemos $B \in \mathcal{B}(\mathbb{R})$.
   Debemos mostrar que $[S_N \in B] \in \mathcal{F}_N$, pero $[S_N \in B] = \bigcup_{m=1}^{\infty} [S_m \in B. N = m]$.
   Ahora $[S_N \in B] \cap [N = n] = \left\{ \bigcup_{m=1}^{\infty} [S_m \in B, N = m] \right\} \cap [N = n] = \bigcup_{m=1}^{\infty} [S_m \in B, N = m, N = n]$.
   El conjunto $[S_m \in B, N = m, N = n] = \begin{cases} \varnothing & \text{si } m \neq n, \\ [S_n \in B, N = n] & \text{si } m = n \end{cases}$
   y como $\varnothing \in \mathcal{F}_n$ para todo n y $[S_n \in B, N = m] \in \mathcal{F}_n$ para todo n en-

tonces $\bigcup_{m=1}^{\infty}[S_m \in B, N = m, N = n] \in \mathcal{F}_n$ para todo n.
Por lo tanto $[S_N \in B] \in \mathcal{F}_N$. $\square$

Observaciones.

a) Sea $\mathcal{G} = \sigma\left(\bigcup_{n=1}^{\infty}\mathcal{F}_n\right) \subseteq \mathcal{F}$.
Si N es tiempo de paro entonces $\mathcal{F}_N \subseteq \mathcal{G}$ porque si $A \in \mathcal{F}_N$ tenemos que $A = A \cap \Omega = A \cap \left(\bigcup_{n=1}^{\infty}[N = n]\right) = \bigcup_{n=1}^{\infty}(A \cap [N = n]) \in \mathcal{G}$ debido que $A \cap [N = n] \in \mathcal{F}_n$.
Si N es regla de paro entonces $[N = \infty] = \Omega - [N < \infty] = \Omega - \bigcup_{n=1}^{\infty}[N = n] \in \mathcal{G}$ pues $\bigcup_{n=1}^{\infty}[N = n] \in \mathcal{G}$.
De este modo N es $\mathcal{G}-$medible.

b) $[N \leqslant n] = \bigcup_{k=1}^{n}[N = k] \in \mathcal{F}_n$ porque $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \cdots \subseteq \mathcal{F}_n$.
Luego $[N = n] = [N \leqslant n] - [N < n] = [N \leqslant n] \cap [N \leqslant n-1]^c$.
Pero $[N \leqslant n] \in \mathcal{F}_n$ y $[N \leqslant n-1]^c \in \mathcal{F}_{n-1} \subseteq \mathcal{F}_n$.
Por lo tanto $[N = n] \in \mathcal{F}_n$ para todo n si y sólo si $[N \leqslant n] \in \mathcal{F}_n$ para todo n. Notemos que $[N > n] = [N \leqslant n]^c \in \mathcal{F}_n$ para todo n.

c) Si N_1, N_2 son tiempos de paro tales que $N_1 \leqslant N_2$ c.s. entonces $\mathcal{F}_{N_1} \subseteq \mathcal{F}_{N_2}$.
Como $N_1 \leqslant N_2$ sucede que $[N_2 \leqslant n] \subseteq [N_1 \leqslant n]$ para todo $n \geqslant 1$ c.s.
Si $A \in \mathcal{F}$ tenemos $A \cap [N_2 = n] = A \cap ([N_2 = n] \cap [N_1 \leqslant n])$.
Si $A \in \mathcal{F}_{N_1}$ entonces $A \cap [N_1 \leqslant n] \in \mathcal{F}_n$ para todo n.
Y como $[N_2 = n] \in \mathcal{F}_n$ para todo n sucede que $A \cap [N_2 = n] \in \mathcal{F}_n$ para todo n.

De este modo $A \in \mathcal{F}_{N_2}$.

Por lo tanto $\mathcal{F}_{N_1} \subseteq \mathcal{F}_{N_2}$.

d) Si N_1, N_2 son tiempos de paro con respecto a $\mathcal{F}_n$ entonces $N_1 + N_2$, $\min\{N_1, N_2\}$ y $\max\{N_1, N_2\}$ son tiempos de paro.

Porque $[N_1 + N_2 \leqslant n] =$

$\bigcup_{\substack{k,l \\ k+l \leqslant n}} [N_1 \leqslant k, N_2 \leqslant l]$ pero $[N_1 \leqslant k] \in \mathcal{F}_k$ y $[N_2 \leqslant l] \in \mathcal{F}_l$.

Así $[N_1 \leqslant k, N_2 \leqslant l] \in \mathcal{F}_{\max\{k,l\}} \subseteq \mathcal{F}_n$.

$[\min\{N_1, N_2\} \leqslant n] = [\min\{N_1, N_2\} > n]^c = [N_1 > n, N_2 > n]^c = [N_1 \leqslant n] \cup [N_2 \leqslant n] \in \mathcal{F}_n$.

$[\max\{N_1, N_2\} \leqslant n] = [N_1 \leqslant n, N_2 \leqslant n] \in \mathcal{F}_n$.

En particular $\min\{N, n_0\}$ es tiempo de paro si $n_0 \in \mathbb{N}$.

Ahora sea $\{S_n, \mathcal{F}_n : n \in \mathbb{N}\}$ una martingala, si S_n es el capital del jugador después de n juegos, un tiempo de paro es una estrategia para parar el juego y S_N es el capital final.

La condición $[N = n]$ o $[N \leqslant n] \in \mathcal{F}_n$ para todo $n \in \mathbb{N}$ dice que la decisión de parar el juego depende de las primeras n variables aleatorias y no del futuro.

Si el juego es justo debería ocurrir que $E(S_N) = E(S_1)$.

Teorema 2.5.6 (Paro Opcional de Doob). *Sean $\{S_n, \mathcal{F}_n : n \in \mathbb{N}\}$ una martingala (submartingala), N un tiempo de paro.*

Si

i) $E(|S_N|) < \infty$, *y*

ii) $\displaystyle\liminf_{n\to\infty} \int_{[N>n]} |S_n| dP = 0$ *entonces* $E(S_N) \overset{=}{(\geqslant)} E(S_1)$.

Demostración:

$S_N = \displaystyle\sum_{j=1}^{\infty} S_j I_{[N=j]}$ por definición.

Por *i)* $E(S_N) = \displaystyle\sum_{j=1}^{\infty} \int_{[N=j]} S_j dP =$

$$\lim_{n\to\infty} \sum_{j=1}^{n} \int_{[N=j]} S_j dP = \lim_{n\to\infty} \sum_{j=1}^{n} \left[\int_{[N\geqslant j]} S_j dP - \int_{[N\geqslant j+1]} S_j dP \right].$$

Sea $A_{j-1} = \{\omega : N(\omega) \geqslant j\} = \{\omega : N(\omega) > j-1\} =$
$\{\omega : N(\omega) \leqslant j-1\}^c \in \mathcal{F}_{j-1}$.

Como $E(S_j|\mathcal{F}_{j-1}) = S_{j-1}$ c.s. por ser martingala entonces $\displaystyle\int_{A_{j-1}} E(S_j|\mathcal{F}_{j-1}) dP =$

$\displaystyle\int_{A_{j-1}} S_j dP$ por definición de esperanza condicional.

Asi $\displaystyle\int_{A_{j-1}} S_j dP = \int_{A_{j-1}} S_{j-1} dP$.

$$\text{Luego } \sum_{j=1}^{n} \left[\int_{[N\geqslant j]} S_j dP - \int_{[N\geqslant j+1]} S_j dP \right] = \left(\int_{[N\geqslant 1]} S_1 dP - \int_{[N\geqslant 2]} S_1 dP \right) +$$

$$\left(\int_{[N\geqslant 2]} S_2 dP - \int_{[N\geqslant 3]} S_2 dP \right) + \cdots + \left(\int_{[N\geqslant n]} S_n dP - \int_{[N\geqslant n+1]} S_n dP \right) =$$

$$\int_{[N\geqslant 1]} S_1 dP - \int_{[N\geqslant n+1]} S_n dP = \int_{\Omega} S_1 dP - \int_{[N\geqslant n+1]} S_n dP = E(S_1) - \int_{[N\geqslant n+1]} S_n dP.$$

Por lo tanto $E(S_N) = \lim\limits_{n\to\infty}\left[E(S_1) - \int\limits_{[N\geqslant n+1]} S_n dP\right] =$

$E(S_1) - \liminf\limits_{n\to\infty} \int\limits_{[N\geqslant n+1]} S_n dP = E(S_1)$. $\square$

Observaciones.

1. Las condiciones *i)* $E(|S_N|) < \infty$ y *ii)* $\liminf\limits_{n\to\infty} \int\limits_{[N\geqslant n]} |S_n| dP = 0$ se cumplen si existe un $n_0 \in \mathbb{N}$ tal que $[N \leqslant n_0] = \Omega$.

   Demostración:

   $$\text{i)}\quad E(|S_N|) = \sum_{j=1}^{\infty} \int\limits_{[N=j]} |S_j| dP = \sum_{j=1}^{n_0} \int\limits_{[N=j]} |S_j| dP < \infty..$$

   $$\text{ii)}\quad [N > n] = \begin{cases} \varnothing & \text{si } n \geqslant n_0 \\ \bigcup\limits_{k=n+1}^{n_0} [N = k] & \text{si } n = 1, 2, \ldots, n_0 - 1 \end{cases} \quad \text{entonces}$$

   $$\int\limits_{[N>n]} |S_n| dP = \begin{cases} 0 & \text{si } n \geqslant n_0 \\ \sum\limits_{k=n+1}^{n_0} \int\limits_{[N=k]} |S_n| dP & \text{si } n = 1, 2, \ldots, n_0 - 1 \end{cases}$$

   Entonces $\liminf\limits_{n\to\infty} \int\limits_{[N>n]} |S_n| dP = 0$. $\square$

2. *i)* y *ii)* se cumplen si $\{S_n : n \in \mathbb{N}\}$ es uniformemente acotada c.s. En efecto, si $|S_n| \leqslant c$ c.s. para todo $n = 1, 2, \ldots$ entonces

   $$|S_N| = \left|\sum_{n=1}^{\infty} S_n I_{[N=n]}\right| \leqslant \sum_{n=1}^{\infty} |S_n| I_{[N=n]} \leqslant c \sum_{n=1}^{\infty} I_{[N=n]} = c.$$

Por lo tanto $E(|S_N|) \leqslant c < \infty$.

Además $\int\limits_{[N>n]} |S_n| dP \leqslant \int\limits_{[N>n]} c dP = cP[N > n] \xrightarrow[n\to\infty]{} 0$ porque $[N > n] \downarrow \varnothing$ y así $P[N > n] \downarrow 0$.

3. La condición *ii)* se cumple si existe una constante $c > 0$ tal que $E(S_n^2) \leqslant nc$ para toda $n \geqslant 1$ y $E(N) < \infty$.

Sabemos que la desigualdad de Cauchy-Schwartz establece que $|E(UV)| \leqslant [E(U^2)E(V^2)]^{\frac{1}{2}}$.

Haciendo $U = |S_n|$ y $V = I_{[N>n]}$ tenemos

$\int\limits_{[N>n]} |S_n| dP = \int\limits_{\Omega} |S_n| I_{[N>n]} dP = \left| \int\limits_{\Omega} |S_n| I_{[N>n]} dP \right| \leqslant \left[E(S_n^2) E(I_{[N>n]}) \right]^{\frac{1}{2}} = \left[E(S_n^2) P[N > n] \right]^{\frac{1}{2}} \leqslant (cnP[N > n])^{\frac{1}{2}}$.

Pero $nP[N > n] = n \sum\limits_{k=n+1}^{\infty} P[N = k] = \sum\limits_{k=n+1}^{\infty} nP[N = k] < \sum\limits_{k=n+1}^{\infty} kP[N = k]$ y $\sum\limits_{k=n+1}^{\infty} kP[N = k] \xrightarrow[n\to\infty]{} 0$ por ser cola de una serie convergente porque $E(N) = \sum\limits_{k=1}^{\infty} kP[N = k] < \infty$.

4. *i)* y *ii)* se cumplen si $\{S_n : n \in \mathbb{N}\}$ es uniformemente integrable, esto es, si

$$\sup_{n\in\mathbb{N}} \int\limits_{[|S_n|>a]} |S_n| dP \xrightarrow[a\to\infty]{} 0.$$

$$\int\limits_{[N>n]} |S_n| dP = \int\limits_{[N>n]\cap[|S_n|>a]} |S_n| dP + \int\limits_{[N>n]\cap[|S_n|\leqslant a]} |S_n| dP.$$

Pero $\displaystyle\int_{[N>n]\cap[|S_n|\leqslant a]} |S_n|dP \leqslant \int_{[N>n]\cap[|S_n|\leqslant a]} adP \leqslant aP[N>n, |S_n| \leqslant a] \leqslant$

$aP[N>n] \xrightarrow[a\to\infty]{} 0$ porque $[N>n, |S_n| \leqslant a] \subseteq [N>n]$.

Mientras que por otro lado

$$\int_{[N>n]\cap[|S_n|>a]} |S_n|dP = \int_\Omega |S_n| I_{[N>n]\cap[|S_n|>a]}dP \leqslant \int_\Omega |S_n| I_{[|S_n|>a]}dP$$

porque $I_{A\cap B} \leqslant I_A$,

$\displaystyle\int_\Omega |S_n| I_{[|S_n|>a]} = \int_{[|S_n|>a]} |S_n|dP \xrightarrow[a\to\infty]{} 0.$

Por lo tanto *ii*) se cumple.

Ahora como $\{S_n : n \in \mathbb{N}\}$ es u.i. entonces por la proposición 2.4.5 existe una constante $c > 0$ tal que $\sup\limits_{n\in\mathbb{N}} E(|S_n|) \leqslant c$.

Por lo tanto $E(|S_n|) \leqslant c$ para todo $n \geqslant 1$.

Como $\Omega = \bigcup\limits_{j=1}^{\infty} [N=j]$ tenemos:

$c \geqslant E(|S_n|) = \displaystyle\int_\Omega |S_n|dP = \sum_{j=1}^{\infty} \int_{[N=j]} |S_n|dP = \sum_{j=1}^{n} \int_{[N=j]} |S_n|dP +$

$\displaystyle\sum_{j=n+1}^{\infty} \int_{[N=j]} |S_n|dP = \sum_{j=1}^{n} \int_{[N=j]} |S_n|dP + \int_{[N>n]} |S_n|dP.$

Como $\{S_n : n \in \mathbb{N}\}$ es martingala y $[N=j] \in \mathcal{F}_j$ tenemos por la proposición 2.1.10:

$$\int_{[N=j]} |S_n|dP \geqslant \int_{[N=j]} |S_{n-1}|dP \geqslant \int_{[N=j]} |S_{n-2}|dP \geqslant \cdots \geqslant \int_{[N=j]} |S_j|dP$$

pues $\{|S_n| : n \in \mathbb{N}\}$ es submartingala.

Por tanto $\displaystyle\sum_{j=1}^{n} \int_{[N=j]} |S_n|dP \geqslant \sum_{j=1}^{n} \int_{[N=j]} |S_j|dP$, asi

$c \geqslant \sum_{j=1}^{n} \int\limits_{[N=j]} |S_n| dP + \int\limits_{[N>n]} |S_n| dP \geqslant \sum_{j=1}^{n} \int\limits_{[N=j]} |S_j| dP + \int\limits_{[N>n]} |S_n| dP$ entonces $\sum_{j=1}^{n} \int\limits_{[N=j]} |S_j| dP \leqslant c - \int\limits_{[N>n]} |S_n| dP$ luego $E\left[\sum_{j=1}^{n} |S_j| I_{[N=j]}\right] \leqslant c - \int\limits_{[N>n]} |S_n| dP.$

Pero $E(|S_n|) = \lim\limits_{n\to\infty} E\left[\sum_{j=1}^{n} |S_j| I_{[N=j]}\right] \leqslant c - \lim\limits_{n\to\infty} \int\limits_{[N>n]} |S_n| dP = c$ porque S_n es integrable.

Proposición 2.5.7. *Sean $\{S_n : n \in \mathbb{N}\}$ martingala (submartingala) y N un tiempo de paro. Supongamos que $E(N) < \infty$ y que $E[|S_{n+1} - S_n| \mid S_1, \ldots, S_n] \leqslant c < \infty$. Si $N \geqslant n$ entonces $E[|S_N|] < \infty$ y $E(S_N) \underset{(\geqslant)}{=} E(S_1)$.*

Demostración:

Veremos que se cumplen *i)* y *ii)* del Teorema 2.5.6.

S_N se puede escribir como

$S_N = S_1 + (S_2 - S_1) + (S_3 - S_2) + \cdots + (S_N - S_{N-1})$ luego

$|S_N| \leqslant |S_1| + |S_2 - S_1| + |S_3 - S_2| + \cdots + |S_N - S_{N-1}| =$

$|S_1| + \sum_{j=2}^{N} |S_j - S_{j-1}|.$

Si $S_0 = 0$ entonces

$E(|S_N|) \leqslant E\left[\sum_{j=1}^{N} |S_j - S_{j-1}|\right] = \sum_{n=1}^{\infty} \sum_{j=1}^{n} \int\limits_{[N=n]} |S_j - S_{j-1}| dP =$

$\sum_{j=1}^{\infty} \sum_{n=j}^{\infty} \int\limits_{[N=n]} |S_j - S_{j-1}| dP = \sum_{j=1}^{\infty} \int\limits_{[N\geqslant j]} |S_j - S_{j-1}| dP$ cambiando el orden de la suma doble y aplicando la primer suma.

Pero el evento $[N \geqslant j] = [N < j]^c = [N \leqslant j-1]^c \in \mathcal{F}_{j-1} = \sigma(S_1, \ldots, S_{j-1}).$

Entonces $\int\limits_{[N\geqslant j]} |S_j - S_{j-1}|dP = \int\limits_{[N\geqslant j]} E[|S_j - S_{j-1}||\mathcal{F}_{j-1}]dP$ por definición de esperanza condicional.

Luego, $\int\limits_{[N\geqslant j]} E[|S_j - S_{j-1}||S_1, \ldots, S_{j-1}]dP \leqslant \int\limits_{[N\geqslant j]} cdP = cP[N \geqslant j]$.

Por lo tanto $E[|S_n|] \leqslant \sum_{j=1}^{\infty} cP[N \geqslant j] = cE(N) < \infty$.

De este modo, se cumple $i)$ del Teorema 2.5.6.

Ahora si $N > n$, $|S_n| \leqslant \sum_{j=1}^{n} |S_j - S_{j-1}| \leqslant \sum_{j=1}^{N} |S_j - S_{j-1}|$ con $S_0 = 0$.

Entonces $\int\limits_{[N>n]} |S_n|dP \leqslant \int\limits_{[N>n]} \sum_{j=1}^{N} |S_j - S_{j-1}|dP$ y como $E\left[\sum_{j=1}^{N} |S_j - S_{j-1}|\right] < \infty$ y $[N > n] \downarrow \varnothing$ tomando $g_n = I_{[N>n]} \sum_{j=1}^{N} |S_j - S_{j-1}|$ por el Teorema de Convergencia Dominada $E(g_n) \longrightarrow E(0) = 0$.

Por lo tanto $ii)$ del Teorema 2.5.6 se cumple.

Asi $E(S_N) = E(S_1)$. □

Los siguientes resultados son consecuencia del Teorema de Paro Opcional y son importantes en la teoría de martingalas.

Proposición 2.5.8 (Ecuación de Wald). *Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con $E(X_1) = \mu$. $S_n = \sum_{i=1}^{n}(X_i - \mu)$ $n \in \mathbb{N}$. Si N es tiempo de paro con $E(N) < \infty$ entonces $E(S_N)$ existe y $E\left(\sum_{k=1}^{N} X_k\right) = \mu E(N)$.*

Demostración:

Primero veremos que $\{S_n, \sigma(S_1, \ldots, S_n)\}$ es martingala.

$E(S_n|S_1,\ldots,S_{n-1}) = E(S_{n-1} + (X_n-\mu)|S_1,\ldots,S_{n-1}) = S_{n-1} + E[(X_n-\mu)|S_1,\ldots,S_{n-1}] = S_{n-1} + E[(X_n-\mu)|X_1,\ldots,X_{n-1}] = S_{n-1} + E(X_n-\mu) = S_{n-1}$ por independencia y porque $E(X_n)=\mu$.

Por lo tanto $\{S_n : n \in \mathbb{N}\}$ es martingala.

Además $E[|S_{n+1}-S_n||S_1,\ldots,S_n] = E[|X_{n+1}-\mu||X_1,\ldots,X_n] = E[|X_{n+1}-\mu|] = E(|X_1-\mu|) \leqslant E(|X_1|)+\mu = c < \infty$ porque $|X_1-\mu| \leqslant |X_1|+\mu$.

Por la proposición 2.5.7 tenemos que $E(S_N) = E(S_1)$.

Pero $E(S_1) = E(X_1-\mu) = 0$.Por lo tanto $E(S_N)=0$.

Asi que $E\left[\sum_{i=1}^{N}(X_i-\mu)\right] = E(X_1-\mu) = 0$.

Entonces $E\left[\sum_{i=1}^{N}(X_i-\mu)\right] = E\left[\sum_{i=1}^{N}X_i - N\mu\right] = E\left[\sum_{i=1}^{N}X_i\right] - \mu E(N) = 0$.

Por lo tanto $E\left[\sum_{i=1}^{N}X_i\right] = \mu E(N)$. □

Proposición 2.5.9 (Identidad Fundamental de Wald). *Sea $\{X_n : n \in \mathbb{N}\}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas. $S_n = \sum_{k=1}^{n} X_k$ $n = 1,2,\ldots$ supongamos que para algún número real $t \neq 0$, $M(t) = E(e^{tX_1})$ existe y $M(t) \geqslant 1$. Sea N tiempo de paro para las sumas $\{S_n : n \in \mathbb{N}\}$ tal que $|S_n| < c$ c.s. para $n \leqslant N$ con $E(N) < \infty$. Entonces $E\left(\frac{e^{tS_N}}{(M(t))^N}\right) = 1$*

Demostración:

Definimos para todo $n \geqslant 1$ las variables aleatorias

$Z_n = \frac{e^{tS_n}}{(M(t))^n}$.

$\{Z_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala con $\mathcal{F}_n = \sigma(S_1, S_2, \ldots, S_n)$. porque $E(Z_n|\mathcal{F}_{n-1}) =$

$E[Z_n|S_1,\ldots,S_{n-1}] = E\left[\frac{e^{tS_n}}{(M(t))^n}\Big|S_1,\ldots,S_{n-1}\right] =$
$E\left[\frac{e^{tS_{n-1}}e^{tX_n}}{(M(t))^n}\Big|S_1,\ldots,S_{n-1}\right] = \frac{e^{tS_{n-1}}}{(M(t))^n}E(e^{tX_n}|X_1,\ldots,X_{n-})$ pues $\sigma(S_1,\ldots,S_{n-1}) = \sigma(X_1,\ldots,X_n)$ y debido a qie $e^{tS_{n-1}}$ es $\sigma(S_1,\ldots,S_{n-1})-$medible.
Así $\frac{e^{tS_{n-1}}}{(M(t))^n}E(e^{tX_n}|X_1,\ldots,X_{n-1}) = \frac{e^{tS_{n-1}}E(e^{tX_n})}{(M(t))^n} = e^{tS_{n-1}}\frac{M(t)}{(M(t))^n} =$
$\frac{e^{tS_{n-1}}}{(M(t))^{n-1}} = Z_{n-1}.$

Por lo tanto $\{Z_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala.

Por otra parte $E(Z_n) = \frac{1}{(M(t))^n}E(e^{tS_n}) = \frac{1}{(M(t))^n}E(e^{tX_1}\cdots e^{tX_n}) =$
$\frac{E(e^{tX_1})\cdots E(e^{tX_n})}{(M(t))^n} = \frac{(M(t))^n}{(M(t))^n} = 1.$
Ahora $E[|Z_{n+1}-Z_n||S_1,\ldots,S_n] = E\left[\left|\frac{e^{tS_{n+1}}}{(M(t))^{n+1}} - \frac{e^{tS_n}}{(M(t))^n}\right|\Big|S_1,\ldots,S_n\right] =$

$$E\left[\left|\frac{e^{tX_{n+1}}e^{tS_n}}{M(t)(M(t))^n} - \frac{e^{tS_n}}{(M(t))^n}\right|\Big|S_1,\ldots,S_n\right]$$
$$= E\left[\frac{e^{tS_n}}{(M(t))^n}\left|\frac{e^{tX_{n+1}}}{M(t)} - 1\right|\Big|S_1,\ldots,S_n\right] = \frac{e^{tS_n}}{(M(t))^n}E\left[\left|\frac{e^{tX_{n+1}}}{M(t)} - 1\right|\Big|S_1,\ldots,S_n\right]$$
$$= \frac{e^{tS_n}}{(M(t))^n}E\left[\frac{|e^{tX_{n+1}} - M(t)|}{M(t)}\right]\text{(por independencia)}$$
$$= Z_nE\left[\frac{|e^{tX_{n+1}} - M(t)|}{M(t)}\right].$$
Como $S_n \leqslant |S_n| \leqslant c$ c.s. y la función exponencial es creciente, tenemos que $e^{tS_n} \leqslant e^{tc}$.
Asi $Z_n = \frac{e^{tS_n}}{(M(t))^n} \leqslant \frac{e^{tc}}{(M(t))^n} \leqslant e^{tc}$ puesto que $M(t) \geqslant 1$.
De este modo $Z_n \leqslant e^{tc}$.
Por otra parte $E\left[\frac{|e^{tX_1} - M(t)|}{M(t)}\right] = \frac{1}{M(t)}E\left[|e^{tX_1} - M(t)|\right] \leqslant \frac{1}{M(t)}\Big\{E(e^{tx_1})+$

$M(t)\Big\} = \frac{1}{M(t)}(M(t) + M(t)) = 2.$

Por lo tanto $Z_n E\left[\left|\frac{e^{tX_1} - M(t)}{M(t)}\right|\right] \leqslant 2e^{tc} < \infty.$

Así por la proposición 2.5.7 $E(Z_n) = E(Z_1)$ pero $E(Z_N) = E\left[\frac{e^{tS_N}}{M(t)^N}\right]$ y $E(Z_1) = 1$.

Concluyendo que $E\left[\frac{e^{tS_N}}{M(t)^N}\right] = 1.$ $\square$

Un caso especial del resultado anterior aparece en el problema siguiente.

Ejemplo 2.5.10. Supongamos que se tiene una sucesión de variables aleatorias i.i.d. no constantes X_i, $S_n = X_1 + \ldots + X_n$.

Se define N como:

$N = \inf_{n \in \mathbb{N}}\{n : S_n \leqslant -a \text{ ó } S_n \geqslant b\}$ donde $a > 0$ y $b > 0$ son fijos.

En este caso necesitamos verificar solamente que $E(N) < \infty$.

Existen $n_0 \in \mathbb{N}$ y $\delta > 0$ tal que $P[|S_{n_0}| > a+b] > \delta$ puesto que de lo contrario se cumpliría que nunca es superado el número $a + b$.

n_0 es el número de sumandos que se necesitan para que la suma $X_1 + \cdots + X_{n_0}$ exceda en valor absoluto al número $a + b$ con probabilidad mayor que δ.

$P\left[|S_{n_0}| > a + b\right] = 1 - P\left[|S_{n_0}| \leqslant a + b\right] > \delta$ si y sólo si $P\left[|S_{n_0}| \leqslant a + b\right] < 1 - \delta$, es decir, la función de distribución de $|S_{n_0}|$ evaluada en $a + b$ no es 1 por tanto existe tal $\delta > 0$.

Ahora tomemos $k \geqslant 0$ cualquier entero.

El evento $[N \geqslant kn_0]$ significa que la primera vez que la sucesión S_n cayó fuera del intervalo $(-a, b)$ ocurre después o en el tiempo kn_0.

Es decir, $[N \geqslant kn_0] = [-a < S_1 < b, -a < S_2 < b,$

$\ldots, -a < S_{kn_0-1} < b]$ la expresión del lado derecho es porque lo único que se sabe es que en los primeros $(kn_0 - 1)$ pasos se mantiene la sucesión en el intervalo $(-a, b)$.

$[-a < S_1 < b, -a < S_2 < b, \ldots, -a < S_{kn_0-1} < b] \subseteq [-a < S_1 < b,$
$-a < S_2 < b, \ldots, -a < S_{kn_0-1} < b, -a < S_{kn_0} < b] \subseteq [|S_{n_0}| \leqslant a + b,$
$|S_{2n_0} - S_{n_0}| \leqslant a + b, \ldots, |S_{kn_0} - S_{(k-1)n_0}| \leqslant a + b]$.

Esta última contención se explica porque la longitud del intervalo $(-a, b)$ es $a + b$.

Se tienen kn_0 puntos en tal intervalo entonces cualquiera dos de ellos distan a lo más $a + b$.

De todo lo anterior

$[N \geqslant kn_0] \subseteq [|S_{n_0}| \leqslant a+b, |S_{2n_0} - S_{n_0}| \leqslant a+b, \ldots, |S_{kn_0} - S_{(k-1)n_0}| \leqslant a+b]$.

Entonces $P[N \geqslant kn_0] \leqslant P[|S_{n_0}| \leqslant a + b, |S_{2n_0} - S_{n_0}| \leqslant a + b, \ldots, |S_{kn_0} - S_{(k-1)n_0}| \leqslant a + b] = P[|S_{n_0}| \leqslant a + b]P[|S_{2n_0} - S_{n_0}| \leqslant a + b] \cdots P[|S_{kn_0} - S_{(k-1)n_0}| \leqslant a + b] \leqslant (1 - \delta)^k$.

La última igualdad es por independencia.

Luego $E(N) = \sum_{n=1}^{\infty} P[N \geqslant n] \leqslant P[N \geqslant 0] + P[N \geqslant 1] + \cdots + P[N \geqslant n_0 - 1] + P[N \geqslant n_0] + P[N \geqslant n_0 + 1] + \cdots + P[N \geqslant 2n_0 - 1] + P[N \geqslant 2n_0] + P[N \geqslant 2n_0 + 1] + \cdots + P[N \geqslant 3n_0 - 1] + \cdots \leqslant n_0 P[N \geqslant 0] + n_0 P[N \geqslant n_0] + n_0 P[N \geqslant 2n_0] + \cdots = n_0 \sum_{k=0}^{\infty} P[N \geqslant kn_0] \leqslant n_0 \sum_{k=0}^{\infty} (1-\delta)^k = n_0 \dfrac{1}{1 - (1 - \delta)} = \dfrac{n_0}{\delta} < \infty$.

Por lo tanto $E(N) < \infty$.

Así $E\left[\dfrac{e^{tS_N}}{(M(t))^N}\right] = 1$ para $t \neq 0$ tal que $M(t) \geqslant 1$. ■

Ejemplo 2.5.11. (Ruina del jugador). Sean $\{X_n : n \in \mathbb{N}\}$ una sucesión de v.a.i.i.d. con $P[X_i = 1] = \frac{1}{2} = P[X_i = -1]$, $S_n = \sum_{k=1}^{n} X_k$ para

$n = 1, 2, \ldots$

$\{S_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala con $\mathcal{F}_n = \sigma(X_1, \ldots, X_n) = \sigma(S_1, \ldots, S_n)$ porque $E(S_n|\mathcal{F}_{n-1}) = E(S_{n-1}+X_n\big|S_1, \ldots, S_{n-1}) = S_{n-1}+E(X_n|X_1, \ldots, X_{n-1}) = S_{n-1} + E(X_n) = S_{n-1} + 0 = S_{n-1}$.

Si hacemos $N = \inf\limits_{n\in\mathbb{N}}\{n : S_n = 1\}$, N es el tiempo en que por primera vez se tiene una ganancia de un peso.

Entonces $P[N < \infty] = 1$ y por definición $P[S_N = 1] = 1$.

Asi $E(S_N) = 1$ pero $E(S_1) = E(X_1) = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$ por lo que $E(S_N) \neq E(S_1)$.

De este modo se debe cumplir que $E(N) = \infty$ por la proposición 2.5.8.

Sean a, b dos enteros positivos y N definido por $N = \inf\limits_{n\in\mathbb{N}}\{n : S_n = -a \text{ ó } S_n = b\}$.

X_i es la ganancia del $i-$ésimo juego y S_n es la ganancia durante n juegos, a y b los capitales iniciales de los jugadores. Si juegan hasta la ruina, N es la duración del juego.

Por el Ejemplo 2.5.10 se tiene que $E(N) < \infty$ y por la ecuación de Wald $E(S_N) = E(X_1)E(N) = 0 \cdot E(N) = 0$.

Sean p_{-a} la probabilidad de que el juego termine en el estado $-a$ y p_b la probabilidad de que el juego termine en el estado b.

Tenemos que $E(S_N) = (-a)p_{-a} + bp_b$

Entonces $0 = (-a)p_{-a} + b(1 - p_{-a})$.

Por lo tanto $b = (a + b)p_{-a}$.

De este modo $p_{-a} = \dfrac{b}{a+b}$ y $p_b = \dfrac{a}{a+b}$.

La probabilidad de ruina es proporcional al dinero que se tiene.

$\{S_n^2 - n : n \in \mathbb{N}\}$ es martingala.

Porque $E[S_n^2 - n|S_1, \ldots, S_{n-1}] = E[S_n^2|S_1, \ldots, S_{n-1}] - n$
$= E[(S_{n-1} + X_n)^2|S_1, \ldots, S_{n-1}] - n =$
$E[S_{n-1}^2 + 2X_nS_{n-1} + X_n^2|S_1, \ldots, S_{n-1}] - n = S_{n-1}^2 + 2S_{n-1}E(X_n|X_1, \ldots, X_{n-1}) +$
$E(X_n^2|S_1, \ldots, S_{n-1}) - n = S_{n-1}^2 + 2S_{n-1}E(X_n) + E(X_n^2) - n = S_{n-1}^2 + E(X_n^2) - n$
$= S_{n-1}^2 + 1 - n = S_{n-1}^2 - (n-1)$.

Entonces $\{S_n^2 - n : n \in \mathbb{N}\}$ es martingala.

Si N es como antes veremos que se satisfacen las condiciones de la proposición 2.5.8.

$E[|S_{n+1}^2 - (n+1) - S_n^2 + n||S_1, \ldots, S_n] = E[|S_{n+1}^2 - S_n^2 - 1||S_1, \ldots, S_n]$ pero $S_{n+1}^2 = (S_n + X_{n+1})^2 = S_n^2 + 2X_{n+1}S_n + X_{n+1}^2$ entonces $S_{n+1}^2 - S_n^2 - 1 = 2X_{n+1}S_n + X_{n+1}^2 - 1$.

Luego $|S_{n+1}^2 - S_n^2 - 1| = |2X_{n+1}S_n + X_{n+1}^2 - 1| \leqslant 2|S_n||X_{n+1}| + 1 + 1 = 2|S_n||X_{n+1}| + 2$ porque $X_{n+1}^2 = 1$ c.s.

Asi $E[|S_{n+1}^2 S_n^2 - 1||S_1, \ldots, S_n] \leqslant E[2|S_n||X_{n+1}||S_1, \ldots, S_n] + 2 = 2|S_n|E(|X_{n+1}|) + 2 = 2|S_n| + 2$ porque $|X_{n+1}| = 1$ c.s. y por independencia.

Ahora si $n \leqslant N$ entonces $|S_n| \leqslant a + b$.

Por lo tanto $E[|S_{n+1}^2 - S_n^2 - 1||S_1, \ldots, S_n] \leqslant 2(a+b) + 2$.

Por lo tanto por la proposición 2.5.8 $E(S_N^2 - N) = E(S_1^2 - 1) = E(X_1^2 - 1) = E(1-1) = 0$.

De este modo $E(N) = E(S_N^2) = (-a)^2 p_{-a} + b^2 p_b = a^2 p_{-a} + b^2(1 - p_{-a}) = (a^2 - b^2)p_{-a} + b^2$.

Luego $E(N) = (a^2 - b^2)\dfrac{b}{a+b} + b^2 = \dfrac{(a-b)(a+b)b}{a+b} + b^2 = (a-b)b + b^2 = ab$.

Esto significa que la duración media del juego es el producto de los capitales de los jugadores.

Si un jugador tiene 10 pesos y otro 15 pesos en promedio jugarán 150 veces hasta la ruina de alguno de ellos.

En el caso general donde $P[X_i = 1] = p$ y $P[X_i = -1] = q$ con $q = 1 - p$, suponiendo que $p > q$ entonces $E(X_1) = p - q = \mu > 0$.

Tenemos que $\{S_n - n\mu, \mathcal{F}_n : n \in \mathbb{N}\}$ y $\left\{ \left(\frac{q}{p}\right)^{S_n}, \mathcal{F}_n : n \in \mathbb{N}\right\}$ con $\mathcal{F}_n = \sigma(X_1, \ldots, X_n) = \sigma(S_1, \ldots, S_n)$ son martingalas.

De la ecuación de Wald $E(S_N) = \mu E(N)$ para cualquier tiempo de paro con $E(N) < \infty$.

Se podría aplicar la proposición 2.5.8 si la martingala $\left\{ \left(\frac{q}{p}\right)^{S_n} : n \in \mathbb{N}\right\}$ cumple con $E\left[\left|\left(\frac{q}{p}\right)^{S_{n+1}} - \left(\frac{q}{p}\right)^{S_n}\right| \Big| S_1, \ldots, S_n\right] \leqslant c$.

$E\left[\left|\left(\frac{q}{p}\right)^{S_{n+1}} - \left(\frac{q}{p}\right)^{S_n}\right| \Big| S_1, \ldots, S_n\right] = E\left[\left(\frac{q}{p}\right)^{S_n}\left|\left(\frac{q}{p}\right)^{X_{n+1}} - 1\right| \Big| S_1, \ldots, S_n\right] = \left(\frac{q}{p}\right)^{S_n} E\left[\left|\left(\frac{q}{p}\right)^{X_{n+1}} - 1\right|\right]$ porque S_n es $\sigma(S_1, \ldots S_n)$−medible y por la independencia de X_{n+1} con $S_1, \ldots, S_n$.

Pero $\left(\frac{q}{p}\right)^{S_n} E\left[\left|\left(\frac{q}{p}\right)^{X_{n+1}} - 1\right|\right] \leqslant \left(\frac{q}{p}\right)^{S_n} \left\{E\left[\left(\frac{q}{p}\right)^{X_{n+1}}\right] + 1\right\} =$
$\left(\frac{q}{p}\right)^{S_n} \left[\frac{q}{p}p + \left(\frac{q}{p}\right)^{-1} q + 1\right] = \left(\frac{q}{p}\right)^{S_n} [q + p + 1] = 2\left(\frac{q}{p}\right)^{S_n} \leqslant c$ porque $p > q$.

Además $S_n < a + b$, entonces $E\left[\left(\frac{q}{p}\right)^{S_N}\right] = E\left[\left(\frac{q}{p}\right)^{S_1}\right] = E\left[\left(\frac{q}{p}\right)^{X_1}\right] = \frac{q}{p}p + \left(\frac{q}{p}\right)^{-1} q = 1$.

De esta manera $E\left[\left(\frac{q}{p}\right)^{S_N}\right] = 1$ pero $E\left[\left(\frac{q}{p}\right)^{S_N}\right] = \left(\frac{q}{p}\right)^{-a} p_{-a} + \left(\frac{q}{p}\right)^{b} p_b = \left(\frac{q}{p}\right)^{-a} p_{-a} + \left(\frac{q}{p}\right)^{b}(1 - p_{-a})$.

Asi

$1 = p_{-a}\left[\left(\frac{q}{p}\right)^{-a} - \left(\frac{q}{p}\right)^{b}\right] + \left(\frac{q}{p}\right)^{b}.$

Por lo tanto $p_{-a} = \dfrac{1-\left(\frac{q}{p}\right)^{b}}{\left(\frac{q}{p}\right)^{-a} - \left(\frac{q}{p}\right)^{b}}.$

Si $b \longrightarrow \infty$ la probabilidad de ser arruinado ante un oponente infinitamente rico es: $\dfrac{1}{\left(\frac{q}{p}\right)^{-a}} = \left(\dfrac{q}{p}\right)^{a}.$

Un argumento similar se aplica si $p < q$ y en tal caso $\lim\limits_{b\to\infty} p_{-a} = 1$, es decir, $P[S_n = -a$ para algún $n \in \mathbb{N}] = 1$ cuando $p < q$ lo que significa la ruina es segura en este caso. ■

Proposición 2.5.12. *Sean $\{S_n, \mathcal{F}_n : n \in \mathbb{N}\}$ una submartingala no negativa y $\{\varphi_n : n \in \mathbb{N}\}$ una sucesión de funciones decrecientes no negativas c.s. tal que φ_n es $\mathcal{F}_{n-1}-$medible para todo $n \in \mathbb{N}$, $\mathcal{F}_0 = \{\varphi, \Omega\}$. Supongamos que $E[\varphi_n|S_n - S_{n-1}|] < \infty$ para todo $n \in \mathbb{N}$ y sea N cualquier regla de paro entonces*

$$\int_{[N<\infty]} \varphi_N S_N dP \leqslant \sum_{k=1}^{\infty} E[\varphi_k E(S_k - S_{k-1}|\mathcal{F}_{k-1})] \textit{ con } S_0 = 0.$$

Demostración:

Sean $X_k = S_k - S_{k-1}$ $k = 1, 2, \ldots$ y $Z_n = \varphi_n S_n - \sum\limits_{k=1}^{n} \varphi_k E(X_k|\mathcal{F}_{k-1}) + \sum\limits_{k=1}^{n}(\varphi_{k-1} - \varphi_k)S_{k-1}$ entonces Z_n es $\mathcal{F}_n-$medible.

Mientras que $Z_n - Z_{n-1} = \varphi_n S_n - \sum\limits_{k=1}^{n} \varphi_k E(X_k|\mathcal{F}_{k-1}) + \sum\limits_{k=1}^{n}(\varphi_{k-1} - \varphi_k)S_{k-1} - \varphi_{n-1}S_{n-1} + \sum\limits_{k=1}^{n-1} \varphi_k E(X_k|\mathcal{F}_{k-1}) - \sum\limits_{k=1}^{n-1}(\varphi_{k-1} - \varphi_k)S_{k-1} = \varphi_n S_n - \varphi_{n-1}S_{n-1} - \varphi_n E(X_n|\mathcal{F}_{n-1}) + (\varphi_{n-1} - \varphi_n)S_{n-1} = \varphi_n(S_n - S_{n-1}) - \varphi_n E(X_n|\mathcal{F}_{n-1}) = \varphi_n X_n - \varphi_n E(X_n|\mathcal{F}_{n-1}) = \varphi_n[X_n - E(X_n|\mathcal{F}_{n-1})].$

Por lo tanto $Z_n = Z_{n-1} + \varphi_n[X_n - E(X_n|\mathcal{F}_{n-1})]$.

Asi $E[Z_n|\mathcal{F}_{n-1}] = E[Z_{n-1}|\mathcal{F}_{n-1}] + E[\varphi_n(X_n - E(X_n|\mathcal{F}_{n-1}))\big|\mathcal{F}_{n-1}] = Z_{n-1} + \varphi_n[E(X_n|\mathcal{F}_{n-1}) - E[X_n|\mathcal{F}_{n-1}]] = Z_{n-1} + 0 = Z_{n-1}$.

Pues φ_n y $E(X_n|\mathcal{F}_{n-1})$ son $\mathcal{F}_{n-1}-$medibles.

Esto significa que $\{Z_n, \mathcal{F}_n : n \in \mathbb{N}\}$ es martingala.

Si existe un n_0 tal que $P[N \leqslant n_0] = 1$ entonces por el Teorema 2.5.6 tenemos que $E(Z_N) = E(Z_1) = E[\varphi_1 S_1 - \varphi_1 E(X_1|\mathcal{F}_0)] = E[\varphi_1 X_1] - E[E(\varphi_1 X_1|\mathcal{F}_0)] = E(\varphi_1 X_1) - E(\varphi_1 X_1) = 0$. Y $Z_N = \varphi_N S_N - \sum_{k=1}^{N} \varphi_k E(X_k|\mathcal{F}_{k-1}) + \sum_{k=1}^{N} (\varphi_{k-1} - \varphi_k) S_{k-1}$.

De este modo $0 = E(Z_N) = E(\varphi_N S_N) - E\left[\sum_{k=1}^{N} \varphi_k E(X_k|\mathcal{F}_{k-1}) - \sum_{k=1}^{N} (\varphi_{k-1} - \varphi_k) S_{k-1}\right]$.

Por lo tanto $E(\varphi_N S_N) = E\left[\sum_{k=1}^{N} \varphi_k E(X_k|\mathcal{F}_{k-1}) - \sum_{k=1}^{N} (\varphi_{k-1} - \varphi_k) S_{k-1}\right] \leqslant E\left[\sum_{k=1}^{N} \varphi_k E(X_k|\mathcal{F}_{k-1})\right]$ porque $\varphi_{k-1} - \varphi_k \geqslant 0$ puesto que la sucesión $\{\varphi_n : n \in \mathbb{N}\}$ es decreciente y también $S_{k-1} \geqslant 0$.

Así $E(\varphi_N S_N) \leqslant E\left[\sum_{k=1}^{\infty} \varphi_k E(X_k|\mathcal{F}_{k-1}) I_{[N \geqslant k]}\right]$.

En el caso general, sea $N_0 = \min\{N, n_0\}$ entonces $E[\varphi_{N_0} S_{N_0}] \leqslant E\left[\sum_{k=1}^{\infty} \varphi_k E(X_k|\mathcal{F}_{k-1}) I_{[N_0 \geqslant k]}\right]$ y si $n_0 \longrightarrow \infty$ obtenemos que

$$\int_{[N<\infty]} \varphi_N S_N dP = E[\varphi_N S_N] \leqslant E\left[\sum_{k=1}^{\infty} \varphi_k E(S_k - S_{k-1}|\mathcal{F}_{k-1})\right]. \quad \square$$

Corolario 2.5.13 (Chow). *Sean* $\{S_n, \mathcal{F}_n : n \in \mathbb{N}\}$ *una submartingala y* $\{c_n : n \in \mathbb{N}\}$ *una sucesión decreciente de números reales positivos. Para cada*

$\epsilon > 0$:

$$P[\max_{1\leqslant k\leqslant n} c_k S_k \geqslant \epsilon] \leqslant \frac{1}{\epsilon}\left[c_1 E(S_1^+) + \sum_{k=2}^{n} c_k E(S_k^+ - S_{k-1}^+)\right]$$

Demostración:

Sea $N = \inf\{n : c_n S_n^+ \geqslant \epsilon\}$.

Esto significa que $[N = k] = [c_1 S_1^+ < \epsilon, c_2 S_2^+ < \epsilon, \ldots, c_{k-1} S_{k-1}^+ < \epsilon, c_k S_k^+ \geqslant \epsilon]$.

Hacemos $E = [\max_{1\leqslant k\leqslant n} c_k S_k^+ \geqslant \epsilon]$.

Mostraremos que $\epsilon I_E(\omega) \leqslant c_N S_N^+(\omega) I_E(\omega)$.

Hay dos casos:

i) $\omega \notin E$ entonces se obtiene que $0 \leqslant 0$.

ii) $\omega \in E$ entonces $\epsilon I_E(\omega) = \epsilon$

y por otro lado sabemos que $c_1 S_1^+(\omega) < \epsilon, c_2 S_2^+(\omega) < \epsilon, \ldots, c_{N-1} S_{N-1}^+(\omega) < \epsilon$, $c_N S_N^+(\omega) \geqslant \epsilon$.

Asi $c_N S_N^+(\omega) I_E(\omega) \geqslant \epsilon I_E(\omega) = \epsilon$.

Por lo tanto $\epsilon I_E \leqslant c_N S_N^+ I_E$.

Tomando esperanzas

$\epsilon P(E) \leqslant \int_E c_N S_N^+ dP$.

Entonces $\epsilon P[\max_{1\leqslant k\leqslant n} c_k S_k^+ \geqslant \epsilon] \leqslant \int_E c_N S_N^+ dP = \int_{[N\leqslant n]} c_N S_N^+ dP$

porque $E[\max_{1\leqslant k\leqslant n} c_k S_k^+ \geqslant \epsilon] = \bigcup_{k=1}^{n} [c_k S_k^+ \geqslant \epsilon]$. Además

$$\left(\bigcup_{k=1}^{n} [c_k S_k^+ \geqslant \epsilon]\right)^c = \bigcap_{k=1}^{n} [c_k S_k^+ < \epsilon] = [c_1 S_1^+ < \epsilon, \ldots, c_N S_1^+ < \epsilon] = [N > n].$$

Por lo tanto $E^c = [N > n]$ implica que $E = [N \leqslant n]$.

Hacemos $X_k = S_k^+ - S_{k-1}^+$ para $k \geqslant 1$, $S_0^+ = 0$.

De este modo $\{S_n^+, \mathcal{F}_n : n \in \mathbb{N}\}$ es submartingala $\geqslant 0$ y por el Teorema 2.5.6

$$\int_{[N\leqslant n]} c_N S_N^+ dP \leqslant E\left[\sum_{k=1}^{n} c_k E(X_k|\mathcal{F}_{k-1})\right] = c_1 E(S_1^+) + \sum_{k=2}^{n} c_k E(S_k^+ - S_{k-1}^+).$$

Luego $\epsilon P[\max_{1\leqslant k\leqslant n} c_k S_k^+ \geqslant \epsilon] \leqslant c_1 E(S_1^+) + \sum_{k=2}^{n} c_k E(S_k^+ - S_{k-1}^+)$.

Por lo tanto $P[\max_{1\leqslant k\leqslant n} c_k S_k^+ \geqslant \epsilon] \leqslant \frac{1}{\epsilon}\left[c_1 E(S_1^+) + \sum_{k=2}^{n} c_k E(S_k^+ - S_{k-1}^+)\right]$. □

Corolario 2.5.14 (Hájek-Rényi). *Sean $\{X_n : n \in \mathbb{N}\}$ una sucesión de v.a.i. con $E(X_n) = 0$ y $E(X_n^2) < \infty$ para todo $n \in \mathbb{N}$, $\{c_n : n \in \mathbb{N}\}$ una sucesión de números reales decreciente y positiva. Para cada $\epsilon > 0$*

$$P[\max_{1\leqslant k\leqslant n} c_k |S_k| \geqslant \epsilon] \leqslant \frac{1}{\epsilon^2}\sum_{k=1}^{n} c_k^2 E(X_k^2) \textit{ donde } S_n = \sum_{k=1}^{n} X_k \; n = 1, 2, \ldots$$

Demostración:

Notemos que $[\max_{1\leqslant k\leqslant n} c_k|S_k| \geqslant \epsilon] = [\max_{1\leqslant k\leqslant n} c_k^2 S_k^2 \geqslant \epsilon^2]$ y por el corolario 2.5.13

$$P[\max_{1\leqslant k\leqslant n} c_k|S_k| \geqslant \epsilon] = P[\max_{1\leqslant k\leqslant n} c_k^2 S_k^2 \geqslant \epsilon^2] \leqslant$$

$$\frac{1}{\epsilon^2}\left[c_1^2 E(S_1^2) + \sum_{k=2}^{n} c_k^2 E(S_k^2 - S_{k-1}^2)\right] \text{ (porque } S_k^{2+} = S_k^2)$$

$$= \frac{1}{\epsilon^2}\left[c_1^2 E(S_1^2) + \sum_{k=2}^{n} c_k^2 E[(S_{k-1} + X_k)^2 - S_{k-1}^2]\right] =$$

$$\frac{1}{\epsilon^2}\left[c_1^2 E(S_1^2) + \sum_{k=2}^{n} c_k^2 E[X_k^2 + 2X_k S_{k-1}^2]\right] =$$

$$\frac{1}{\epsilon^2}\left[c_1^2 E(S_1^2) + \sum_{k=2}^{n} c_k^2 E(X_k^2)\right] \text{ porque } E(X_k S_{k-1}^2) = E(X_k)E(S_{k-1}^2) = 0$$

$$= \frac{1}{\epsilon^2}\sum_{k=1}^{n} c_k^2 E(X_k^2).$$

Notemos que si $c_n = 1$ para todo $n \in \mathbb{N}$ se obtiene la desigualdad de Kolmogorov. $\square$

Capítulo 3

Caracterización de Distribuciones Mediante Esperanzas Condicionales

3.1. Introducción

Las distribuciones suelen pertenecer a familias de acuerdo al tipo de función que las representa, tal como la familia exponencial en donde $f(x) = a(\theta)b(x)e^{c(\theta)d(x)}$, $x \in I$ en donde θ es el parámetro e I es un intervalo en $\mathbb{R}$. En esta parte estudiaremos familias de distribuciones de acuerdo a la forma que tienen las esperanzas condicionales de sus estadísticas de orden o algunas otras esperanzas condicionales.

Veremos algunos resultados de los últimos diez años que fueron publicados por Ouyang [1995], Balasubramanian y Dey [1997], Wu y Ouyang [1996], Zoroa, Ruiz y Marin [1990].

El extremo izquierdo de una distribución $F(x)$ absolutamente continua se define como $\alpha = \sup\{x : F(x) = 0\}$, es decir, $\alpha \geqslant -\infty$ y el extremo derecho de $F(x)$ es $\beta = \inf\{x : F(x) = 1\}$ por lo que $\beta \leqslant \infty$.

Sea $F(x)$ una distribución continua.

Recordemos que si $X_1, \ldots, X_n$ es una muestra aleatoria (v.a.i.i.d.) las estadísticas de orden de la muestra son las variables aleatorias

$X_{(1)} = \text{mín}\{X_1, \ldots, X_n\}$ = Estadística de orden 1.

$\vdots$

$X_{(j)}$ =La $j-$ésima más chica de las X_i= Estadística de orden j.

$\vdots$

$X_{(n)} = \text{máx}\{X_1, \ldots, X_n\}$ = Estadística de orden n.

Calcularemos $f_{X_{(k)}, X_{(k+1)}}(x, y)$ la densidad conjunta de la estadística de orden $X_{(k)}$ y $X_{(k+1)}$. También $f_{X_{(k)}}(x)$ la densidad de $X_{(k)}$.

Para calcular $f_{X_{(k)}}(x)$ consideremos lo siguiente:

$P[x < X_{(k)} \leqslant x + h] = \int_x^{x+h} f_{X_{(k)}}(t)dt \approx h f_{X_{(k)}}(x)$ con h pequeña pues $f_{X_{(k)}}$ debe ser continua, de no ser así la igualdad es válida para casi toda x.

Entonces $f_{X_{(k)}}(x) = \lim_{h \to 0} \dfrac{P[x < X_{(k)} \leqslant x + h]}{h}$.

El cálculo de $P[x < X_{(k)} \leqslant x + h]$ es como sigue:

El evento $[x < X_{(k)} \leqslant x + h]$ es igual al evento que consiste en que haya $(k-1)$ observaciones de la muestra menores o iguales que x, una entre x y $x+h$, $(n-k)$ mayores que $x+h$.

Luego

$$P[x < X_{(k)} \leqslant x + h] = \binom{n}{k-1}\binom{n-k+1}{1}F^{k-1}(x)[F(x+h) - F(x)][1 - F(x+h)]^{n-k}$$

De este modo:

$$f_{X_{(k)}}(x) = \frac{n!}{(k-1)!(n-k)!}F^{k-1}(x)[1 - F(x)]^{n-k}f(x) \quad \alpha < x < \beta.$$

Ahora para calcular $f_{X_{(k)},X_{(k-1)}}(x,y)$ consideramos de la misma forma:

$P[x < X_{(k)} \leqslant x + h_1, y < X_{(k+1)} \leqslant y + h_2] = \int\limits_{y}^{y+h_2}\int\limits_{x}^{x+h_1} f_{X_{(k)},X_{(k-1)}}(t,v)dtdv \approx$ $h_1h_2f_{X_{(k)},X_{(k-1)}}(x,y)$ con h_1, h_2 pequeñas.

Por lo que $f_{X_{(k)},X_{(k+1)}}(x,y) = \lim\limits_{\substack{h_1\to 0\\ h_2\to 0}} \dfrac{P[x < X_{(k)} \leqslant x + h_1, y < X_{(k+1)} \leqslant y + h_2]}{h_1h_2}$.

El evento $[x < X_{(k)} \leqslant x + h_1, y < X_{(k+1)} \leqslant y + h_2]$ es igual al evento que consiste en que haya $(k-1)$ observaciones menores o iguales que x, una entre x y $x + h_1$, una entre y y $y + h_2$, $(n-k-1)$ mayores que $y + h_2$, por lo que:

$P[x < X_{(k)} \leqslant x + h_1, y < X_{(k+1)} \leqslant y + h_2]$

$=\binom{n}{k-1}\binom{n-k+1}{1}\binom{n-k}{1}F^{k-1}(x)[F(x+h_1)-F(x)][F(h+y_2)-F(y)][1-F(y+h_2)]^{n-k-1}.$

De donde

$$f_{X_{(k)},X_{(k-1)}}(x,y) = \frac{n!}{(k-1)!(n-k-1)!}F^{k-1}(x)[1 - F(y)]^{n-k-1}f(x)f(y),$$

para $\alpha < x < y < \beta$.

Ahora es posible calcular $f_{X_{(k+1)}|X_{(k)}}(y|x)$ puesto que

$$f_{X_{(k+1)}|X_{(k)}}(y|x) = \frac{f_{X_{(k)},X_{(k+1)}}(x,y)}{f_{X_{(k)}}(x)} \quad \text{si } \alpha < x < y < \beta.$$

$$f_{X_{(k+1)}|X_{(k)}}(y|x) = \frac{\frac{n!}{(k-1)!(n-k-1)!}F^{k-1}(x)[1-F(y)]^{n-k-1}f(x)f(y)}{\frac{n!}{(k-1)!(n-k)!}F^{k-1}(x)[1-F(y)]^{n-k}f(x)}$$

$$= \frac{(n-k)[1-F(y)]^{n-k-1}f(y)}{[1-F(x)]^{n-k}} \quad \text{si } \alpha < x < y < \beta.$$

Luego

$$F_{X_{(k+1)}|X_{(k)}}(y|x) = P[X_{k+1} \leqslant y|X_{(k)} = x] = \int_x^y f_{X_{(k+1)}|X_{(k)}}(t|x)dt$$

$$= (n-k)\int_x^y \frac{[1-F(t)]^{n-k-1}}{[1-F(x)]^{n-k}}f(t)dt = \frac{1}{[1-F(x)]^{n-k}}[1-F(t)]^{n-k}\Big|_y^x$$

$$= 1 - \frac{[1-F(x)]^{n-k}}{[1-F(y)]^{n-k}} \quad \text{para } x < y < \beta.$$

Con técnicas similares a las usadas anteriormente podemos mostrar que:

a) $f_{X_{(r)},X_{(i)},X_{(s)}}(x,t,y) = \frac{n!}{(r-1)!(i-r-1)!(s-i-1)!(n-s)!}F^{r-1}(x)[F(t) - F(x)]^{i-r-1} \cdot [F(y)-F(t)]^{s-i-1}[1-F(y)]^{n-s}f(x)f(t)f(y).$

para $1 \leqslant r < i < s \leqslant n$, $x < t < y$.

b) $f_{X_{(r)},X_{(s)}}(x,y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}F^{r-1}(x)[F(y)-F(x)]^{s-r-1} \cdot [1-F(y)]^{n-s}f(x)f(y).$

para $1 \leqslant r < s \leqslant n$, $x < y$.

Entonces de a) y b) obtenemos que:

$$f_{X_{(i)}|X_{(r)},X_{(s)}}(t|x,y) = \frac{(s-r-1)!}{(i-r-1)!(s-i-1)!}\left[\frac{F(t)-F(x)}{F(y)-F(x)}\right]^{i-r-1} \cdot \left[\frac{F(y)-F(t)}{F(y)-F(x)}\right]^{s-i-1}\left[\frac{f(t)}{F(y)-F(x)}\right] \text{ para } x < t < y.$$

Necesitamos obtener $f_{X_{(k)},X_{(r)}}(t,x)$ y $f_{X_{(k)}}(t)$ para calcular $f_{X_{(r)}|X_{(k)}}(x|t)$, sabemos que

$$f_{X_{(k)}}(t) = \frac{n!}{(k-1)!(n-k)!}F^{k-1}(t)[1-F(t)]^{n-k}f(t) \quad \alpha < t < \beta, \text{ y que}$$

$$f_{X_{(k)},X_{(r)}}(t,x) = \frac{n!}{(k-1)!(r-k-1)!(n-r)!}F^{k-1}(t)[F(x)-F(t)]^{r-k-1}\cdot[1-F(x)]^{n-r}f(t)f(x)$$

con $1 \leqslant k < r \leqslant n, \quad \alpha < t < x < \beta$

entonces:

$$f_{X_{(r)}|X_{(k)}}(x|t) = \frac{\frac{n!}{(k-1)!(r-k-1)!(n-r)!}F^{k-1}(t)[F(x)-F(t)]^{r-k-1}[1-F(x)]^{n-r}f(t)f(x)}{\frac{n!}{(k-1)!(n-k)!}F^{k-1}(t)[1-F(t)]^{n-k}f(t)} =$$

$$\frac{(n-k)!}{(r-k-1)!(n-r)!}\frac{[F(x)-F(t)]^{r-k-1}[1-F(x)]^{n-r}}{[1-F(t)]^{n-k}}f(x) \quad \alpha < t < x < \beta.$$

3.2. Caracterizaciones de Ouyang

Teorema 3.2.1. *Sea $F(x)$ absolutamente continua con extremos izquierdo y derecho α y β respectivamente. Consideremos $u : [\alpha, \beta) \longrightarrow \mathbb{R}$ continua, tal que u' existe y es continua en (α, β) además supongamos que $0 < F(x) < 1$ para $\alpha < x < \beta$ y que $E(|u(X)|) < \infty$.*

$E[u(X_{(k+1)})|X_{(k)}] = u(X_{(k)}) + c$ *c.s. para toda $k < n$ y alguna $c \neq 0$ si y sólo si*

$F(x) = 1 - e^{\frac{(u(\alpha) - u(x))}{c^*}}$ *para $x \in (\alpha, \beta)$ con $c^* = c(n-k)$ en donde*

$$\lim_{x \to \beta} u(x) = \begin{cases} \infty & si\ c > 0 \\ -\infty & si\ c < 0. \end{cases}$$

Demostración:

Denotaremos $\overline{F}(x) = 1 - F(x)$

y de este modo $F_{X_{(k+1)}|X_{(k)}}(y|x) = 1 - \dfrac{(\overline{F}(y))^{n-k}}{(\overline{F}(x))^{n-k}} \quad x < y < \beta.$

Entonces

$$E[u(X_{(k+1)})|X_{(k)} = x] = \int_x^\beta u(y) dF_{X_{(k+1)}|X_{(k)}}(y|x) = -\int_x^\beta \frac{u(y)d[(\overline{F}(y))^{n-k}]}{(\overline{F}(x))^{n-k}} =$$

$$-\frac{1}{(\overline{F}(x))^{n-k}} \int_x^\beta u(y) d[(\overline{F}(y))^{n-k}].$$

Ahora integraremos por partes usando la fórmula conocida $\int l dm = lm - \int m dl$.

Sean $l = u(y)$, $dl = u'(y)dy$, $dm = d[(\overline{F}(y))^{n-k}]$, $m = (\overline{F}(y))^{n-k}$

Así,

$$E[u(X_{(k+1)})|X_(k) = x] = -\frac{1}{(\overline{F}(x))^{n-k}} \left[u(y)(\overline{F}(y))^{n-k}\Big|_x^\beta - \int_x^\beta u'(y)(\overline{F}(y))^{n-k} dy \right].$$

Como

$0 < \overline{F}(y) < 1$ tenemos que $0 < (\overline{F}(y))^{n-k} < \overline{F}(y) < 1$.

De este modo $|u(y)|(\overline{F}(y))^{n-k} \leqslant |u(y)|(\overline{F}(y))$.

Luego $\lim\limits_{x \to \beta} |u(y)|\overline{F}(y) = 0$ como consecuencia de que $E[|u(X)|]$ existe.

Así,

$$E[u(X_{(k+1)})|X_{(k)} = x] = -\frac{1}{(\overline{F}(x))^{n-k}}\left[-u(x)(\overline{F}(x))^{n-k} - \int_x^\beta u'(y)(\overline{F}(y))^{n-k}dy\right] =$$

$$u(x) + \frac{1}{(\overline{F}(x))^{n-k}}\int_x^\beta u'(y)(\overline{F}(y))^{n-k}dy \text{ si } \alpha < x < \beta.$$

Supongamos ahora que $E[u(X_{(k+1)})|X_{(k)} = x] = u(x) + c$ entonces

$u(x) + c = u(x) + \frac{1}{(\overline{F}(x))^{n-k}}\int_x^\beta u'(y)(\overline{F}(y))^{n-k}dy,$ es decir

$\int_x^\beta u'(y)(\overline{F}(y))^{n-k}dy = c(\overline{F}(x))^{n-k}$.

Derivando ambas partes con respecto a x obtenemos

$-u'(x)(\overline{F}(x))^{n-k} = c(n-k)(\overline{F}(x))'(\overline{F}(x))^{n-k-1}$

o

$u'(x)(\overline{F}(x)) = c^*(\overline{F}(x))'$ por lo que $\frac{(\overline{F}(x))'}{\overline{F}(x)} = -\frac{u'(x)}{c^*}$ para $\alpha < x < \beta$.

Entonces $\frac{d}{dx}\ln(\overline{F}(x)) = -\frac{1}{c^*}\frac{d}{dx}u(x)$, integrando: $\int_\alpha^x \left(\ln\left(\overline{F}(t)\right)\right)' dt = -\frac{1}{c^*}\int_\alpha^x (u(t))'dt$.

Así $\ln(\overline{F}(x)) - \ln(\overline{F}(\alpha)) = -\frac{1}{c^*}(u(x) - u(\alpha))$, es decir $\ln\left(\frac{\overline{F}(x)}{\overline{F}(\alpha)}\right) = \frac{(u(\alpha) - u(x))}{c^*}$.

Como $F(\alpha) = 0$ entonces $\overline{F}(\alpha) = 1 - F(\alpha) = 1$.

Por tanto $\ln\left(\frac{\overline{F}(x)}{\overline{F}(\alpha)}\right) = \ln(\overline{F}(x))$.

Concluyendo que $\overline{F}(x) = e^{\frac{(u(\alpha)-u(x))}{c^*}}$ es decir $F(x) = 1 - e^{\frac{(u(\alpha)-u(x))}{c^*}}$ $\alpha < x < \beta$.

Como $\lim\limits_{x\to\beta} F(x) = 1$ entonces $\lim\limits_{x\to\beta} e^{\frac{(u(\alpha)-u(x))}{c^*}} = 0$, o equivalentemente $\lim\limits_{x\to\beta} e^{\frac{-u(x)}{c(n-k)}} = 0$

que significa que $\lim\limits_{x\to\beta} \dfrac{-u(x)}{c(n-k)} = -\infty$.
Esto es $\lim\limits_{x\to\beta} \dfrac{u(x)}{c} = \infty$.

Concluimos que $\lim\limits_{x\to\beta} u(x) = \begin{cases} \infty & \text{si } c > 0 \\ -\infty & \text{si } c < 0. \end{cases}$

Supongamos que $F(x) = 1 - e^{\frac{(u(\alpha)-u(x))}{c^*}}$ y $\overline{F}(x) = e^{\frac{(u(\alpha)-u(x))}{c^*}}$ entonces $(\overline{F}(x))^{n-k} = (e^{\frac{(u(\alpha)-u(x))}{c^*}})^{n-k} = e^{\frac{u(\alpha)-u(x)}{c}}$.

Entonces $E[u(X_{(k+1)})|X_{(k)} = x] = -\dfrac{1}{(\overline{F}(x))^{n-k}} \displaystyle\int_x^\beta u(y)d[(\overline{F}(y))^{n-k}]$

$$= -\frac{1}{e^{\frac{u(\alpha)-u(x)}{c}}} \int_x^\beta u(y)d(e^{\frac{u(\alpha)-u(y)}{c}}) = -e^{\frac{u(x)-u(\alpha)}{c}} e^{\frac{u(\alpha)}{c}} \int_x^\beta u(y)d(e^{\frac{-u(y)}{c}})$$

$= e^{\frac{u(x)}{c}} \displaystyle\int_x^\beta u(y)d(e^{\frac{-u(y)}{c}})$. Sean $l = u(y)$, $dl = u'(y)dy$, $dm = d(e^{-\frac{u(y)}{c}})$, $m = e^{-\frac{u(y)}{c}}$

$$E[u(X_{(k+1)})|X_{(k)} = x] = -e^{\frac{u(x)}{c}} \left[u(y)e^{-\frac{u(y)}{c}} \Big|_x^\beta - \int_x^\beta u'(y)e^{-\frac{u(y)}{c}}dy \right]$$

$$= -e^{\frac{u(x)}{c}} \left[-u(x)e^{-\frac{u(x)}{c}} - \int_x^\beta u'(y)e^{-\frac{u(y)}{c}}dy \right] = u(x) + e^{\frac{u(x)}{c}} \int_x^\beta u'(y)e^{-\frac{u(y)}{c}}dy.$$

Sea $v = -\frac{u(y)}{c}$, $u(y) = -cv$ entonces $u'(y)dy = -cdv$

$$E[u(X_{(k+1)})|X_{(k)} = x] = u(x) - ce^{\frac{u(x)}{c}} \int_{-\frac{u(x)}{c}}^{-\frac{u(\beta)}{c}} e^v dv = u(x) - ce^{\frac{u(x)}{c}} \left(e^{-\frac{u(\beta)}{c}} - e^{-\frac{u(x)}{c}} \right)$$

$= u(x) + c$ por tanto $E[u(X_{(k+1)})|X_{(k)}] = u(X_{(k)}) + c$. □

Teorema 3.2.2. *Sea $F(x)$ absolutamente continua con extremos izquierdo y derecho α y β respectivamente. Consideremos $u : (\alpha, \beta] \longrightarrow \mathbb{R}$ continua tal que u' existe y es continua en (α, β), además supongamos que $E[|u(X)|] < \infty$. $E[u(X_{(k)})|X_{(k+1)}] = u(X_{(k+1)}) + c$ c.s. para toda $k < n$ y alguna $c \neq 0$ si y sólo si*

$F(x) = e^{\frac{u(\beta)-u(x)}{c^*}}$ con $\alpha < x < \beta$ y $c^* = ck$ en particular

$$\lim_{x \to \alpha} u(x) = \begin{cases} \infty & si\ c > 0 \\ -\infty & si\ c < 0. \end{cases}$$

Demostración:

Por el Teorema 3.2.1 se tiene que

$$f_{X_{(k+1)}}(y) \frac{n!}{k!(n-k-1)!} F^k(y)[1 - F(y)]^{n-k-1} f(y) \quad \alpha < y < \beta,$$

y

$$f_{X_{(k)}, X_{(k+1)}}(x, y) \frac{n!}{(k-1)!(n-k-1)!} F^{k-1}(x)[1 - F(y)]^{n-k-1} f(x) f(y)$$

$\alpha < x < y < \beta$.

Entonces:

$$f_{X_{(k)}|X_{(k+1)}}(x|y) = \frac{f_{X_{(k)}, X_{(k+1)}}(x, y)}{f_{X_{(k+1)}}(y)} = \frac{\frac{n!}{(k-1)!(n-k-1)!} F^{k-1}(x)[1 - F(y)]^{n-k-1} f(x) f(y)}{\frac{n!}{k!(n-k-1)!} F^k(y)[1 - F(y)]^{n-k-1} f(y)}$$

si $\alpha < x < y$, es decir

$$f_{X_{(k)}|X_{(k+1)}}(x|y) = \frac{k F^{k-1}(x) f(x)}{F^k(y)} \quad \alpha < x < y.$$

Luego $F_{X_{(k)}|X_{(k+1)}}(x|y) = P[X_{(k)} \leqslant x | X_{(k+1)} = y] = \int\limits_{\alpha}^{x} f_{X_{(k)}|X_{(k+1)}}(t|y)dt =$
$\dfrac{k}{F^k(y)} \int\limits_{\alpha}^{x} F^{k-1}(t)f(t)dt = \dfrac{k}{F^k(y)} \dfrac{F^k(t)}{k}\Big|_{\alpha}^{x} = \dfrac{F^k(x)}{F^k(y)}$ para $\alpha < x < y$,
entonces

$$E[u(X_{(k)})|X_{(k+1)} = y] = \int\limits_{\alpha}^{y} u(x)dF_{X_{(k)}|X_{(k+1)}}(x|y) = \frac{1}{F^k(y)} \int\limits_{\alpha}^{y} u(x)d[F^k(x)],$$

tomando $l = u(x)$, $dl = u'(x)dx$, $dm = d[F^k(x)]$, $m = F^k(x)$

$$E[u(X_{(k)})|X_{(k+1)} = y] = \frac{1}{F^k(y)} \left[u(x)F^k(x)\Big|_{\alpha}^{y} - \int\limits_{\alpha}^{y} u'(x)F^k(x)dx \right].$$

Como $0 < F(x) < 1$, también $0 < F^k(x) < F(x) < 1$, de aquí que $|u(x)|F^k(x) \leqslant |u(x)F(x)|$.
De este modo $\lim\limits_{x \to \alpha} |u(x)F(x)| = 0$ como consecuencia de que $E[|u(X)|]$ existe.
Así $E[u(X_{(k)})|X_{(k+1)} = y] = \dfrac{1}{F^k(y)} \left[u(y)F^k(y) - \int\limits_{\alpha}^{y} u'(x)F^k(x)dx \right]$

$$= u(y) - \frac{1}{F^k(y)} \int\limits_{\alpha}^{y} u'(x)F^k(x)dx \quad \text{con} \quad \alpha < y < \beta.$$

Supongamos ahora que: $E[u(X_{(k)})|X_{(k+1)} = y] = u(y) + c$,
entonces $u(y) + c = u(y) - \dfrac{1}{F^k(y)} \int\limits_{\alpha}^{y} u'(x)F^k(x)dx$,
es decir $cF^k(y) = -\int\limits_{\alpha}^{y} u'(x)F^k(x)dx$.

Derivando ambas partes con respecto a y obtenemos

$ckF^{k-1}(y)f(y) = -u'(y)F^k(y)$ o $ckF'(y) = -u'(y)F(y)$,

por lo que $\dfrac{F'(y)}{F(y)} = -\dfrac{u'(y)}{c^*}$ para $\alpha < y < \beta$.

Ahora $\dfrac{d}{dy}\ln F(y) = -\dfrac{1}{c^*}\dfrac{d}{dy}u(y)$, integrando $\int_y^\beta (\ln F(t))'dt = -\dfrac{1}{c^*}\int_y^\beta (u(t))'dt$.

Asi $\ln(F(\beta) - \ln(F(y))) = -\dfrac{1}{c^*}(u(\beta) - u(y))$

es decir, $\ln\left(\dfrac{F(\beta)}{F(y)}\right) = -\dfrac{1}{c^*}(u(\beta) - u(y))$.

Como $F(\beta) = 1$ entonces $\ln\left(\dfrac{1}{F(y)}\right) = -\dfrac{1}{c^*}(u(\beta) - u(y))$.

Luego $\dfrac{1}{F(y)} = e^{-\frac{(u(\beta)-u(y))}{c^*}}$ $\quad \alpha < y < \beta$

es decir $F(y) = e^{\frac{(u(\beta)-u(y))}{c^*}}$ $\quad$ con $\quad \alpha < y < \beta$.

Como $\lim\limits_{y\to\alpha} F(y) = 0$ entonces $\lim\limits_{y\to\alpha} e^{\frac{(u(\beta)-u(y))}{c^*}} = 0$ de donde $\lim\limits_{y\to\alpha} e^{\frac{-u(y)}{c^*}} = 0$, que es equivalente a

$$\lim_{y\to\alpha} -\frac{u(y)}{c^*} = -\infty \text{ entonces } \lim_{y\to\alpha} \frac{u(y)}{c} = \infty.$$

$$\text{concluyendo que } \lim_{y\to\infty} u(y) = \begin{cases} \infty & \text{si } c > 0 \\ -\infty & \text{si } c < 0. \end{cases}$$

Supongamos que $F(y) = e^{\frac{(u(\beta)-u(y))}{c^*}}$ con $c^* = ck$ entonces $F^k(y) = e^{\frac{(u(\beta)-u(y))}{c}}$.

Ahora $E[u(X_{(k)})|X_{(k+1)} = y] = \int_\alpha^y u(x)d[F_{X_{(k)}|X_{(k+1)}}(x|y)] =$

$\dfrac{1}{F^k(y)}\int_\alpha^y u(x)d[F^k(x)] = \dfrac{1}{e^{\frac{u(\beta)-u(y)}{c}}}\int_\alpha^y u(x)d[e^{\frac{u(\beta)-u(x)}{c}}] = e^{\frac{u(y)}{c}}\int_\alpha^y u(x)d[e^{-\frac{u(x)}{c}}]$

tomando $l = u(x)$, $dl = u'(x)dx$, $dm = d(e^{-\frac{u(x)}{c}})$, $m = e^{-\frac{u(x)}{c}}$,

$$E[u(X_{(k)})|X_{(k+1)} = y] = e^{\frac{u(y)}{c}} \left[u(x)e^{-\frac{u(x)}{c}} \Big|_{\alpha}^{y} - \int_{\alpha}^{y} u'(x)e^{-\frac{u(x)}{c}} dx \right]$$

$$= e^{\frac{u(y)}{c}} \left[u(y)e^{-\frac{u(y)}{c}} - \int_{\alpha}^{y} u'(x)e^{-\frac{u(x)}{c}} dx \right] = u(y) - e^{\frac{u(y)}{c}} \int_{\alpha}^{y} u'(x)e^{-\frac{u(x)}{c}} dx.$$

Sea $v = -\frac{u(x)}{c}$, $u(x) = -cv$, entonces $u'(x)dx = -cdv$

$$E[u(X_{(k)})|X_{(k+1)} = y] = u(y) + ce^{\frac{u(y)}{c}} \left(\int_{-\frac{u(\alpha)}{c}}^{-\frac{u(y)}{c}} e^{v} dv \right)$$

$$= u(y) + ce^{\frac{u(y)}{c}} (e^{-\frac{u(y)}{c}} - e^{-\frac{u(\alpha)}{c}}) = u(y) + c.$$

Por lo tanto $E[u(X_{(k)})|X_{(k+1)}] = u(X_{(k+1)}) + c.$ $\square$

Ejemplo 3.2.3. La pareja formada por la función $u(x) = -\ln(1 - x^{\theta})$ con $0 \leqslant x < 1$, $\theta > 0$ y la función de distribución $F(x) = x^{\theta}$ $0 \leqslant x \leqslant 1$ satisfacen el Teorema 3.2.1.

$F(x)$ es distribución claramente, $\alpha = 0$ y $\beta = 1$.

Luego $u : [0, 1) \longrightarrow \mathbb{R}$ es continua, $u(x) = -\ln(1 - x^{\theta})$, $u(0) = 0$ y si $x_0 \in [0, 1)$ $\lim\limits_{x \to x_0} u(x) = \lim\limits_{x \to x_0} [-\ln(1 - x^{\theta})] = -\lim\limits_{x \to x_0} \ln(1 - x^{\theta})$.

Pero como $0 \leqslant x < 1$ entonces $0 \leqslant x^{\theta} < 1$ pues $\theta > 0$.

Además $0 < 1 - x^{\theta} \leqslant 1$ pero ln es continua en $(0, 1]$ por lo que $\lim\limits_{x \to x_0} u(x) = -\lim\limits_{x \to x_0} \ln(1 - x^{\theta}) = -\ln(1 - x_0^{\theta})$.

Por lo tanto $u(x)$ es continua en $[0, 1)$.

Ahora veremos que $u'(x)$ existe y es continua en $(0, 1)$

$u'(x) = [-\ln(1 - x^{\theta})]' = \dfrac{\theta x^{\theta-1}}{1 - x^{\theta}}$.

Si $0 < x_0 < 1$ $\quad \lim\limits_{x \to x_0} \dfrac{\theta x^{\theta-1}}{1 - x^{\theta}} = \dfrac{\theta x_0^{\theta-1}}{1 - x_0^{\theta}}$.

Falta comprobar que $E[|u(X)|] < \infty$, $u(x) = -\ln(1 - x^{\theta})$ con $0 \leqslant x < 1$,

entonces $0 < 1 - x^\theta \leqslant 1$ porque $\theta > 0$, por lo que $-\infty < \ln(1 - x^\theta) \leqslant 0$.
De este modo $0 \leqslant -\ln(1-x^\theta) < \infty$, por lo tanto $u(x) \geqslant 0$ asi $|u(x)| = u(x)$.
Luego $E[|u(X)|] = E[u(X)] = -\int_0^1 \ln(1 - x^\theta)dF(x) = -\int_0^1 \ln(1 - x^\theta)d(x^\theta)$.
$t = 1 - x^\theta,\ dt = -d(x^\theta)$.
Por lo tanto $E[|u(X)|] = \int_0^1 \ln t(-dt) = -\int_0^1 \ln t dt = -[t \ln t - t]\Big|_0^1 = 1$ porque
$\lim\limits_{t \to 0} t \ln t = \lim\limits_{t \to 0} \frac{\ln t}{1/t} = \lim\limits_{t \to 0} \frac{1/t}{1/t^2} = -\lim\limits_{t \to 0} t = 0$.
De este modo se cumplen las condiciones del Teorema 3.2.1 por lo que

$$E[u(X_{(k+1)})|X_{(k)}] = u(X_k) + c \quad \text{c.s.}$$

Esto es,

$$E[-\ln(1 - X^\theta_{(k+1)})|X_{(k)}] = -\ln(1 - X^\theta_{(k)}) + c \text{ para } k = 1, \ldots, n-1.\ c \neq 0.$$

Además $F(x) = x^\theta$ $0 < x < 1$, $\alpha = 0$, $\beta = 1$ cumple con la igualdad $F(x) = 1 - e^{\frac{(u(\alpha)-u(x))}{c^*}}$ para $0 < x < 1$ con $c^* = c(n-k)$.
Luego $u(\alpha) = u(0) = 0$, por lo tanto $F(x) = 1 - e^{-\frac{u(x)}{c^*}}$ sustituyendo $u(x) = -ln(1 - x^\theta)$ tenemos que $F(x) = 1 - e^{\frac{\ln(1-x^\theta)}{c^*}} = 1 - e^{\ln[(1-x^\theta)^{\frac{1}{c^*}}]} = 1 - (1 - x^\theta)^{\frac{1}{c^*}}$ pero $F(x) = x^\theta$ entonces $c^* = 1$ por lo que $c = \frac{1}{n-k}$.
De este modo $E[-ln(1 - X^\theta_{(k+1)})|X_{(k)}] = -\ln(1 - X^\theta_{(k)}) + \frac{1}{n-k}$.
También se debe cumplir que $\lim\limits_{x \to \beta} u(x) = \begin{cases} \infty & \text{si } c > 0 \\ -\infty & \text{si } c < 0. \end{cases}$
Pero sabemos que $c = \frac{1}{n-k} > 0$ entonces $\lim\limits_{x \to \beta} u(x) = \lim\limits_{x \to 1}(-\ln(1 - x^\theta)) = -\lim\limits_{x \to 1} \ln(1 - x^\theta) = \infty$.
Otras parejas de funciones $u(x)$ y $F(x)$ que satisfacen el Teorema 3.2.1 son:

a) $u(x) = -\theta \ln(\beta - x) \quad 0 \leqslant x < \beta,\ \beta > 0,\ \theta > 0.$
$F(x) = 1 - (1 - \frac{x}{\beta})^b \quad 0 < x < \beta,\ b > 0.$

b) $u(x) = \ln(1 + x^\theta) \quad 0 \leqslant x < \infty,\ \theta > 0.$
$F(x) = 1 - (1 + x^\theta)^{-b} \quad x \geqslant 0, b > 0.$

c) $u(x) = \ln x \quad \alpha \leqslant x < \infty.$
$F(x) = 1 - \left(\frac{\alpha}{x}\right)^b \quad \alpha \leqslant x < \infty,\ b > 0.$

d) $u(x) = \ln[1 - \frac{1}{1 + e^{-\theta x}}] \quad x \in \mathbb{R},\ \theta > 0.$
$F(x) = \frac{1}{1 + e^{-\theta x}} \quad x \in \mathbb{R}.$

e) $u(x) = (x - \alpha)^\theta \quad \alpha \leqslant x < \infty,\ \alpha > 0,\ \theta > 0.$
$F(x) = 1 - e^{-\frac{(x-\alpha)^\theta}{b}} \quad \alpha \leqslant x < \infty,\ b > 0.$ ■

Ejemplo 3.2.4. La pareja formada por la función $u(x) = x$, $-\infty < x \leqslant \beta,\ \beta > 0$ y la función de distribución $F(x) = e^{b(x-\beta)} \quad -\infty < x \leqslant \beta,\ b > 0$ satisfacen el Teorema 3.2.2. $F(x)$ es distribución porque $F(x) = e^{-b\beta}e^{bx}$ que claramente es creciente ($b > 0$) y continua, $\lim\limits_{x\to-\infty} F(x) = \lim\limits_{x\to-\infty} e^{-b\beta}e^{bx} = e^{-b\beta} \lim\limits_{x\to-\infty} e^{bx} = 0$,
$\lim\limits_{x\to\beta} F(x) = \lim\limits_{x\to\beta} e^{-b\beta}e^{bx} = e^{-b\beta}e^{b\beta} = 1.$
$\alpha = -\infty$, luego $u : (-\infty, \beta] \longrightarrow \mathbb{R}$ definida por $u(x) = x$ es continua, $u'(x) = 1$ que es continua.

Falta verificar que $E[|u(X)|] < \infty$ pero

$$E[|u(X)|] = E[|X|] = \int_{-\infty}^{\beta} |x| dF(x) = -\int_{-\infty}^{0} x dF(x) + \int_{0}^{\beta} x dF(x) = \int_{0}^{\beta} x d(e^{-b\beta}e^{bx}) -$$

$$\int_{-\infty}^{0} xd(e^{-b\beta}e^{bx}) = b\int_{0}^{\beta} e^{-b\beta}xe^{bx}dx - b\int_{-\infty}^{0} e^{-b\beta}xe^{bx}dx =$$

$$be^{-b\beta}\int_{0}^{\beta} xe^{bx}dx - be^{-b\beta}\int_{-\infty}^{0} xe^{bx}dx.$$

Tomando $l = x$, $dl = dx$, $dm = e^{bx}dx$, $m = \frac{1}{b}e^{bx}$ tenemos

$$E[|u(X)|] = be^{-b\beta}\left[\frac{x}{b}e^{bx}|_0^\beta - \frac{1}{b}\int_0^\beta e^{bx}dx\right] - be^{-b\beta}\left[\frac{x}{b}e^{bx}|_{-\infty}^0 - \frac{1}{b}\int_{-\infty}^0 e^{bx}dx\right] =$$

$$be^{-b\beta}\left[\frac{\beta}{b}e^{b\beta} - \frac{1}{b^2}(e^{b\beta}-1)\right] - be^{-b\beta}[-\frac{1}{b^2}] = \beta - \frac{1}{b} + \frac{e^{-b\beta}}{b} + \frac{e^{-b\beta}}{b} < \infty.$$

De este modo se cumplen las condiciones del Teorema 3.2.2 por lo que $E[u(X_{(k)})|X_{(k+1)}] = u(X_{(k+1)}) + c$ c.s., esto es, $E[X_{(k)}|X_{(k+1)}] = X_{(k+1)} + c$ c.s. para $k = 1, 2, \ldots, n-1$ y $c \neq 0$.

Además $F(x) = e^{-b\beta}e^{bx}$ para $-\infty < x \leqslant \beta$, $b > 0$ cumple con la igualdad $F(x) = e^{\frac{u(\beta)-u(x)}{c^*}}$

con $-\infty < x \leqslant \beta$ y $c^* = ck$.

Pero $u(\beta) = \beta$, por lo tanto $F(x) = e^{\frac{\beta - x}{c^*}} = e^{\frac{\beta}{c^*}}e^{-\frac{x}{c^*}}$.

Sabemos que $F(x) = e^{-b\beta}e^{bx}$ por lo que $c^* = -\frac{1}{b} = ck$ entonces $c = -\frac{1}{bk} < 0$.

Así se tiene que $E[X_{(k)}|X_{(k+1)}] = X_{(k+1)} - \frac{1}{bk}$ c.s.

También del Teorema 3.2.2 se debe cumplir que $\lim\limits_{x\to\alpha} u(x) = \begin{cases} \infty & \text{si } c > 0 \\ -\infty & \text{si } c < 0, \end{cases}$

es este caso $c < 0$ por lo que $\lim\limits_{x\to-\infty} u(x) = \lim\limits_{x\to\infty} x = -\infty$

Otras parejas de funciones $u(x)$ y $F(x)$ que satisfacen el Teorema 3.2.2 son:

a) $u(x) = \theta\ln(x-\alpha)$ $\quad \alpha < x \leqslant \beta$, $\alpha < \beta$.

$$F(x) = \left(\frac{x-\alpha}{\beta-\alpha}\right)^b \quad \alpha \leqslant x \leqslant \beta,\ b > 0.$$

b) $u(x) = \theta \ln(1 + e^{-x}),\ x \in \mathbb{R},\ \theta > 0.$
$$F(x) = \left(\frac{1}{1+e^{-x}}\right)^b,\ x \in \mathbb{R}, b > 0.$$

c) $u(x) = \ln(1 - \frac{1}{x^\theta}),\ 1 < x < \infty,\ \theta > 0.$
$$F(x) = 1 - \frac{1}{x^\theta} \quad 1 \leqslant x < \infty.$$

d) $u(x) = \ln\left(1 + e^{-\theta x}\right) \quad x \in \mathbb{R},\ \theta > 0.$
$$F(x) = \frac{1}{1+e^{-\theta x}} \quad x \in \mathbb{R}.$$

e) $u(x) = \ln\left[1 - e^{-\lambda(x-\alpha)^\theta}\right] \quad \alpha < x < \infty,\ \theta > 0,\ \alpha > 0.$
$$F(x) = 1 - e^{-\lambda(x-\alpha)^\theta} \quad \alpha \leqslant x < \infty. \blacksquare$$

3.3. Caracterizaciones de Balasubramanian y Dey

Después retomaremos la investigación de propiedades de esperanzas condicionales de estadísticas de orden, el siguiente resultado da una caracterización de la distribución de una variable aleatoria continua.

Teorema 3.3.1. *Sea X una variable aleatoria absolutamente continua con función de distribución $F(x)$ y densidad $f(x)$ continua en el soporte.*
$E[g(X)|x \leqslant X \leqslant y] = \dfrac{2g(x)g(y)}{g(x)+g(y)}$ $x \leqslant y$ *si y sólo si* $g(x) = \sqrt{\dfrac{b}{a+F(x)}}$ *donde g es no negativa con derivada continua, a y b constantes.*

Demostración:

Primero supongamos que

$$E[g(X)|x \leqslant X \leqslant y] = \int_x^y \frac{g(t)f(t)}{F(y)-F(x)}dt = \frac{2g(x)g(y)}{g(x)+g(y)}.$$

Entonces $\displaystyle\int_x^y g(t)f(t)dt = (F(y)-F(x))\frac{2g(x)g(y)}{g(x)+g(y)}$.

Derivando ambas partes con respecto a y tenemos

$$g(y)f(y) = \frac{2\partial}{\partial y}\left[(F(y)-F(x))\frac{g(x)g(y)}{g(x)+g(y)}\right] =$$

$$2g(x)\frac{(g(x)+g(y))[g'(y)(F(y)-F(x))+f(y)g(y)]-[F(y)-F(x)]g(y)g'(y)}{(g(x)+g(y))^2} =$$

$$\frac{2[F(y)-F(x)]}{(g(x)+g(y))^2}[g^2(x)+g(x)g(y)-g(x)g(y)]g'(y)+\frac{2g(x)f(y)g(y)}{(g(x)+g(y))} =$$

$$\frac{2[F(y)-F(x)]g'(y)}{(\frac{g(x)+g(y)}{g(x)})^2}+\frac{2f(y)}{\frac{g(x)+g(y)}{g(x)g(y)}}=\frac{2[F(y)-F(x)]g'(y)}{[1+\frac{g(y)}{g(x)}]^2}+\frac{2f(y)}{\frac{1}{g(x)}+\frac{1}{g(y)}}.$$

De este modo

$$g(y)f(y)=\frac{2[F(y)-F(x)]g'(y)}{[1+\frac{g(y)}{g(x)}]^2}+\frac{2f(y)}{\frac{1}{g(x)}+\frac{1}{g(y)}},$$

esto es,

$$g(y)f(y)-\frac{2f(y)}{\frac{1}{g(x)}+\frac{1}{g(y)}}=\frac{2[F(y)-F(x)]g'(y)}{[1+\frac{g(y)}{g(x)}]^2}.$$

Asi $f(y)g(y)-\dfrac{2g(x)g(y)f(y)}{g(x)+g(y)}$

$$=\frac{2[F(y)-F(x)]g'(y)g^2(x)}{(g(x)+g(y))^2}.$$

Pero $\dfrac{f(y)g(y)g(x)+f(y)g^2(y)-2g(x)g(y)f(y)}{g(x)+g(y)}$

$$=\frac{2[F(y)-F(x)]g'(y)g^2(x)}{(g(x)+g(y))^2}.$$

De este modo $\dfrac{f(y)g^2(y)-g(x)g(y)f(y)}{F(y)-F(x)}=\dfrac{2g'(y)g^2(x)}{g(x)+g(y)}.$

Entonces $\dfrac{f(y)[g^2(y)-g(x)g(y)]}{F(y)-F(x)}=\dfrac{2g'(y)g^2(x)}{g(x)+g(y)}.$

Por lo que $\dfrac{f(y)g(y)(g(y)-g(x))}{F(y)-F(x)}$

$$=\frac{2g'(y)g^2(x)}{g(x)+g(y)}.$$

Por tanto $\dfrac{f(y)}{F(y)-F(x)}=\dfrac{2g^2(x)g'(y)}{g(y)[g^2(y)-g^2(x)]}.$

Notemos que $\dfrac{\partial}{\partial y}\ln\left[1-\dfrac{g^2(x)}{g^2(y)}\right]=\dfrac{2g^2(x)g^{-3}(y)g'(y)}{1-\frac{g^2(x)}{g^2(y)}}=\dfrac{2g^2(x)g^{-1}(y)g'(y)}{g^2(y)-g^2(x)}=$

$$\frac{2g^2(x)g'(y)}{g(y)[g^2(y) - g^2(x)]}.$$

Luego

$$\frac{\partial}{\partial y}\ln(F(y) - F(x)) = \frac{\partial}{\partial y}\ln\left[1 - \frac{g^2(x)}{g^2(y)}\right].$$

Integrando tenemos

$$\ln(F(y) - F(x)) = \ln\left[1 - \frac{g^2(x)}{g^2(y)}\right] + c$$

que se puede escribir como

$$\ln(F(y) - F(x)) = \ln\left[1 - \frac{g^2(x)}{g^2(y)}\right] + \ln k, \quad k > 0,$$

esto es

$$F(y) - F(x) = k\left[1 - \frac{g^2(x)}{g^2(y)}\right].$$

Entonces $F(y) - F(x) = k - \dfrac{kg^2(x)}{g^2(y)}$ implica que

$\dfrac{kg^2(x)}{g^2(y)} = k + F(x) - F(y)$.

Por lo que $\dfrac{g^2(y)}{kg^2(x)} = \dfrac{1}{k + F(x) - F(y)}$.

Asi $g^2(y) = \dfrac{kg^2(x)}{k + F(x) - F(y)}$.

Es decir $g^2(y) = \dfrac{-kg^2(x)}{F(y) - k - F(x)}$ entonces $g(y) = \sqrt{\dfrac{b}{F(y) + a}}$ donde a, b son constantes tales que el cociente $\dfrac{b}{F(y) + a} > 0$.

Supongamos ahora que $g(y) = \sqrt{\dfrac{b}{F(y) + a}}$ de este modo tenemos que

$$\int_x^y \frac{g(t)f(t)}{F(y) - F(x)}dt = \frac{1}{F(y) - F(x)}\int_x^y \sqrt{\frac{b}{F(t) + a}}f(t)dt$$

$$= \frac{1}{F(y)-F(x)}\sqrt{b}\int_x^y (a+F(t))^{-1/2} f(t)dt = \frac{1}{F(y)-F(x)}\sqrt{b}\left[2(a+F(t))^{1/2}\Big|_x^y\right] =$$

$$\frac{2\sqrt{b}}{F(y)-F(x)}\left[\sqrt{a+F(y)}-\sqrt{a+F(x)}\right] =$$

$$\frac{2\sqrt{b}}{(F(y)+a)-(F(x)+a)}\left[\sqrt{a+F(y)}-\sqrt{a+F(x)}\right]$$

$$= \frac{2\sqrt{b}}{\sqrt{a+F(y)}-\sqrt{a+F(x)}} \frac{\left[\sqrt{a+F(y)}-\sqrt{a+F(x)}\right]}{\sqrt{a+F(y)}+\sqrt{a+F(x)}}$$

$$= \frac{2\sqrt{b}}{\sqrt{a+F(y)}+\sqrt{a+F(x)}}, \text{ pero } \sqrt{a+F(y)} = \frac{\sqrt{b}}{g(y)}.$$

Entonces

$$E[g(X)|x \leqslant X \leqslant y] = \frac{2\sqrt{b}}{\frac{\sqrt{b}}{g(y)}+\frac{\sqrt{b}}{g(x)}} = \frac{2}{\frac{1}{g(y)}+\frac{1}{g(x)}} = \frac{2g(x)g(y)}{g(x)+g(y)}. \qquad \square$$

Corolario 3.3.2. *Sean X y g como en el Teorema 3.3.1, $X_1,\ldots,X_n$ una muestra aleatoria de F_X con $X_{(1)} < \cdots < X_{(n)}$ sus estadísticas de orden. Entonces*

$$E\left[\frac{1}{s-r-1}\sum_{i=r+1}^{s-1} g(X_{(i)})\Big| X_{(r)}, X_{(s)}\right] = \frac{2g(X_{(r)})g(X_{(s)})}{g(X_{(r)})+g(X_{(s)})}$$

con $1 \leqslant r < s \leqslant n$, $(s-r) \geqslant 2$ *si y sólo si* $g(x) = \sqrt{\dfrac{b}{F(x)+a}}$.

Demostración:

$$E\left[\frac{1}{s-r-1}\sum_{i=r+1}^{s-1} g(X_{(i)})\Big| X_{(r)} = x, X_{(s)} = y\right] =$$

$$\frac{1}{s-r-1}\sum_{i=r+1}^{s-1} E\left[g(X_{(i)})\Big| X_{(r)} = x, X_{(s)} = y\right] \text{ asi } r < i < s$$

$$E[g(X_{(i)})|X_{(r)} = x, X_{(s)} = y] = \int_x^y g(t) f_{X_{(i)}|X_{(r)},X_{(s)}}(t|x,y)dt =$$

$\int_x^y \frac{(s-r-1)!}{(i-r-1)!(s-i-1)!} \left[\frac{F(t)-F(x)}{F(y)-F(x)}\right]^{i-r-1} \left[\frac{F(y)-F(t)}{F(y)-F(x)}\right]^{s-i-1} \left[\frac{f(t)}{F(y)-F(x)}\right] g(t)dt.$

Sea $l = \frac{F(t)-F(x)}{F(y)-F(x)}$ entonces $F(t) = F(x) + (F(y) - F(x))l$ luego $dF(t) = f(t)dt = (F(y) - F(x))dl$ y también $t = F^{-1}[c + dl]$ con $c = F(x)$ y $d = F(y) - F(x)$.

Por lo tanto $E[g(X_{(i)})|X_{(r)} = x, X_{(s)} = y] =$

$$\frac{(s-r-1)!}{(i-r-1)!(s-i-1)!} \int_0^1 l^{i-r-1}(1-l)^{s-i-1} g[F^{-1}(c+dl)]dl$$

porque $1 - l = \frac{F(y) - F(t)}{F(y) - F(x)}$.

Pero $E\left[\frac{1}{s-r-1} \sum_{i=r+1}^{s-1} g(X_{(i)})|X_{(r)} = x, X_{(s)} = y\right] =$

$$\frac{1}{s-r-1} \sum_{i=r+1}^{s-1} \frac{(s-r-1)!}{(i-r-1)!(s-i-1)!} \int_0^1 l^{i-r-1}(1-l)^{s-i-1} g[F^{-1}(c+dl)]dl =$$

$$\int_0^1 g[F^{-1}(c+dl)] \sum_{i=r+1}^{s-1} \frac{(s-r-2)!}{(i-r-1)!(s-i-1)!} l^{i-r-1}(1-l)^{s-i-1}dl.$$

Sea $j = i - r - 1$, entonces

$$E\left[\frac{1}{s-r-1} \sum_{i=r+1}^{s-1} g(X_{(i)})|X_{(r)} = x, X(s) = y\right] =$$

$$\int_0^1 g[F^{-1}(c+dl)] \sum_{j=0}^{s-r-2} \binom{s-r-2}{j} l^j (1-l)^{s-r-2-j}dl =$$

$$\int_0^1 g[F^{-1}(c+dl)][l+(1-l)]^{s-r-2}dl = \int_0^1 g[F^{-1}(c+dl)]dl = \int_x^y \frac{g(t)f(t)}{F(y)-F(x)}dt =$$

$$\frac{2g(x)g(y)}{g(x)+g(y)}$$

si y sólo si $g(x) = \sqrt{\frac{b}{a + F(x)}}$.

Luego $E\left[\frac{1}{s-r-1}\sum_{i=r+1}^{s-1} g(X_{(i)})|X_{(r)},X_{(s)}\right] = \frac{2g(X_{(r)})g(X_{(s)})}{g(X_{(r)})+g(X_{(s)})}$

si y sólo si $g(x) = \sqrt{\frac{b}{a+F(x)}}$. □

Un resultado que nos da otra caracterización de $F(x)$ es el siguiente

Teorema 3.3.3. *Sean X y g como en el Teorema 3.3.1. Entonces $E[g(X)|x \leqslant X \leqslant y] = \sqrt{g(x)g(y)}$ si y sólo si $g(x) = \frac{b^2}{(F(x)-a)^2}$ donde a y b son constantes.*

Demostración:

Primero supongamos que

$E[g(X)|x \leqslant X \leqslant y] = \int_x^y \frac{g(t)f(t)}{F(y)-F(x)}dt = \sqrt{g(x)}\sqrt{g(y)}$ por hipótesis,

entonces $\int_x^y g(t)f(t)dt = (F(y)-F(x))\sqrt{g(x)}\sqrt{g(y)}$.

Derivando ambas partes con respecto a y obtenemos

$g(y)f(y) = \sqrt{g(x)}[(F(y)-F(x))\frac{g'(y)}{2\sqrt{g(y)}} + f(y)\sqrt{g(y)}]$.

Asi $2\sqrt{g(y)}g(y)f(y) = \sqrt{g(x)}\,[(F(y)-F(x))g'(y) + 2f(y)g(y)]$ por lo que

$2f(y)g^{3/2}(y) - 2f(y)g(y)\sqrt{g(x)} = \sqrt{g(x)}(F(y)-F(x))g'(y)$.

Entonces $2f(y)g(y)[\sqrt{g(y)} - \sqrt{g(x)}] = \sqrt{g(x)}(F(y)-F(x))g'(y)$.

Es decir $\frac{f(y)}{F(y)-F(x)} = \frac{\sqrt{g(x)}g'(y)}{2g(y)[\sqrt{g(y)}-\sqrt{g(x)}]}$.

Notemos que

$$\frac{\partial}{\partial y}\ln\left(1-\frac{\sqrt{g(x)}}{\sqrt{g(y)}}\right) = \frac{\frac{1}{2}\sqrt{g(x)}g^{-3/2}(y)g'(y)}{1-\frac{\sqrt{g(x)}}{\sqrt{g(y)}}} = \frac{\sqrt{g(x)}g'(y)}{2g^{3/2}(y)\left[\frac{\sqrt{g(y)}-\sqrt{g(x)}}{\sqrt{g(y)}}\right]} =$$

$$\frac{\sqrt{g(x)}g'(y)}{2g(y)\left[\sqrt{g(y)}-\sqrt{g(x)}\right]}.$$

Por tanto
$$\frac{f(y)}{F(y)-F(x)} = \frac{\sqrt{g(x)}g'(y)}{2g(y)[\sqrt{g(y)}-\sqrt{g(x)}]}.$$
Luego $\dfrac{\partial}{\partial y}\ln(F(y)-F(x)) =$

$$\frac{\partial}{\partial y}\ln\left[1-\frac{\sqrt{g(x)}}{\sqrt{g(y)}}\right].$$
Entonces $\ln(F(y)-F(x)) = \ln\left[k\left(1-\dfrac{\sqrt{g(x)}}{\sqrt{g(y)}}\right)\right]$,

para alguna $k > 0$.

Es decir,
$$F(y)-F(x) = k\left(1-\frac{\sqrt{g(x)}}{\sqrt{g(y)}}\right).$$
De este modo $\dfrac{F(y)-F(x)}{k} = 1-\dfrac{\sqrt{g(x)}}{\sqrt{g(y)}}$.

Luego
$$\frac{\sqrt{g(x)}}{\sqrt{g(y)}} = 1-\frac{F(y)-F(x)}{k}.$$

Por lo tanto $\dfrac{\sqrt{g(x)}}{\sqrt{g(y)}} = \dfrac{k+F(x)-F(y)}{k}$.

Asi $\sqrt{g(y)} = \dfrac{k\sqrt{g(x)}}{k+F(x)-F(y)}$.

Esto es $\sqrt{g(y)} = \dfrac{-k\sqrt{g(x)}}{F(y)-k-F(x)}$.

Por lo que $g(y) = \dfrac{b^2}{(F(y)-a)^2}$ donde a y b son constantes.

Supongamos que $g(y) = \dfrac{b^2}{(F(y)-a)^2}$ entonces
$$\int_x^y \frac{g(t)f(t)}{F(y)-F(x)}dt =$$

$$\frac{1}{F(y)-F(x)}\int_x^y \frac{b^2}{(F(t)-a)^2}f(t)dt = \frac{b^2}{F(y)-F(x)}\int_x^y \frac{f(t)dt}{(F(t)-a)^2} =$$
$$\frac{b^2}{F(y)-F(x)}\left(\frac{1}{F(t)-a}\Big|_y^x\right) = \frac{b^2}{F(y)-F(x)}\left[\frac{1}{F(x)-a}-\frac{1}{F(y)-a}\right] =$$
$$\frac{b^2}{F(y)-F(x)}\left[\frac{F(y)-a-F(x)+a}{(F(x)-a)(F(y)-a)}\right] = \frac{b^2}{(F(x)-a)(F(y)-a)}.$$

Pero $(F(y)-a)^2 = \dfrac{b^2}{g(y)}$ entonces $E[g(X)|x \leqslant X \leqslant y] = \dfrac{b^2}{\frac{|b|}{\sqrt{g(x)}}\frac{|b|}{\sqrt{g(y)}}} =$ $\sqrt{g(x)g(y)}$. □

Corolario 3.3.4. *Sean X y g como en el Teorema 3.3.1. Entonces* $E\left[\dfrac{1}{s-r-1}\displaystyle\sum_{i=r+1}^{s-1} g(X_{(i)})\Big|X_{(r)}, X_{(s)}\right] = \sqrt{g(X_{(r)})g(X_{(s)})}$.

Demostración:

Es exactamente igual a la demostración del corolario 3.3.2. □

Ejemplo 3.3.5. Las funciones de distribución continuas que sean de la forma $F(x) = Ax^k + B$ son caracterizadas por una de las siguientes dos condiciones:

1) $E\left[\dfrac{1}{s-r-1}\displaystyle\sum_{i=r+1}^{s-1} \frac{1}{X_{(i)}^{k/2}}\Big|X_{(r)}, X_{(s)}\right] = \dfrac{2}{X_{(r)}^{k/2}+X_{(s)}^{k/2}}$.

2) $E\left[\dfrac{1}{s-r-1}\displaystyle\sum_{i=r+1}^{s-1} \frac{1}{X_{(i)}^{2k}}\Big|X_{(r)}, X_{(s)}\right] = \dfrac{1}{X_{(r)}^{k}X_{(s)}^{k}}$.

Estas igualdades son consecuencia de los corolarios 3.3.2 y 3.3.4 respectivamente.

Si $g(x) = \frac{1}{x^{k/2}}$, $F(x) = Ax^k + B$ y se cumple el corolario 3.3.2 entonces $g(x) = \sqrt{\dfrac{b}{a+F(x)}}$.

Luego $g^2(x) = \dfrac{b}{a+F(x)}$, pero $g^2(x) = \dfrac{1}{x^k}$ asi $\dfrac{1}{x^k} = \dfrac{b}{a+Ax^k+B}$ que se cumple si $a=-B$ y $b=A$.

Por lo que

$$E\left[\frac{1}{s-r-1}\sum_{i=r+1}^{s-1}\frac{1}{X_{(i)}^{k/2}}\Big|X_{(r)},X_{(s)}\right] = \frac{2-g(X_{(r)})g(X_{(s)})}{g(X_{(r)})+g(X_{(s)})} = \frac{2-\frac{1}{X_{(r)}^{k/2}}\frac{1}{X_{(s)}^{k/2}}}{\frac{1}{X_{(r)}^{k/2}}+\frac{1}{X_{(s)}^{k/2}}} = \frac{2}{X_{(r)}^{k/2}+X_{(s)}^{k/2}}.$$

Ahora si $g(x) = \dfrac{1}{x^{2k}}$, $F(x) = Ax^k + B$ y se cumple el corolario 3.3.4 entonces $g(x) = \dfrac{b^2}{(F(x)-a)^2}$.

De este modo $\dfrac{1}{x^{2k}} = \dfrac{b^2}{(Ax^k+B-a)^2}$ que se cumple si $a=B$ y $b^2=A^2$ por lo que

$$E\left[\frac{1}{s-r-1}\sum_{i=r+1}^{s-1}\frac{1}{X_{(i)}^{k/2}}\Big|X_{(r)},X_{(s)}\right] = \sqrt{g(X_{(r)})g(X_{(s)})} = \sqrt{\frac{1}{X_{(r)}^{2k}}\frac{1}{X_{(s)}^{2k}}} = \frac{1}{X^{k_{(r)}}}\frac{1}{X^{k_{(s)}}}. \quad \blacksquare$$

Ejemplo 3.3.6. Las funciones las distribución continuas que sean de la forma $F(x) = Ae^{-\theta x^N} + B$ son caracterizadas por una de las siguientes dos condiciones:

1) $E\left[\dfrac{1}{s-r-1}\displaystyle\sum_{i=r+1}^{s-1} e^{\frac{\theta}{2}X_{(i)}^N}\Big|X_{(r)},X_{(s)}\right] = \dfrac{2e^{\frac{\theta}{2}(X_{(r)}^N+X_{(s)}^N)}}{e^{\frac{\theta}{2}X_{(r)}^N}+e^{\frac{\theta}{2}X_{(s)}^N}}.$

2) $E\left[\dfrac{1}{s-r-1}\displaystyle\sum_{i=r+1}^{s-1} e^{2\theta X_{(i)}^N}\Big|X_{(r)},X_{(s)}\right] = e^{\theta(X_{(r)}^N+X_{(s)}^N)}.$

De nuevo la primer igualdad es consecuencia del corolario 3.3.2 porque si $g(x) = e^{\frac{\theta}{2}X^N}$, $F(x) = Ae^{-\theta x^N} + B$ entonces $g(x) = \sqrt{\dfrac{b}{a+F(x)}}$.

De este modo
$e^{\frac{\theta}{2}X^N} = \sqrt{\dfrac{b}{a + Ae^{-\theta x^N} + B}}.$
Por lo que $e^{\theta x^N} = \dfrac{b}{a + Ae^{-\theta x^N} + B}$ que se cumple

si $a = -b$ y $b = A$.

Luego

$$E\left[\frac{1}{s-r-1}\sum_{i=r+1}^{s-1} e^{\frac{\theta}{2}X_{(i)}^N}\Big|X_{(r)}, X_{(s)}\right] = \frac{2g(X_{(r)})g(X_{(s)})}{g(X_{(r)}) + g(X_{(s)})} = \frac{2e^{\frac{\theta}{2}X_{(r)}^N}e^{\frac{theta}{2}X_{(s)}^N}}{e^{\frac{\theta}{2}X_{(r)}^N} + e^{\frac{\theta}{2}X_{(s)}^N}} =$$

$$\frac{2e^{\frac{\theta}{2}(X_{(r)}^N + X_{(s)}^N)}}{e^{\frac{\theta}{2}X_{(r)}^N} + e^{\frac{\theta}{2}X_{(s)}^N}}.$$

Por otro lado si $g(x) = e^{2\theta x^N}$, $F(x) = Ae^{-\theta x^N} + B$ y se cumple el corolario 3.3.4 entonces $g(x) = \dfrac{b^2}{(F(x) - a)^2}$.
Luego $e^{2\theta x^N} = \dfrac{b^2}{(Ae^{-\theta x^N} + B - a)^2}$ que se cumple si $a = B$ y $b^2 = A^2$.
Por lo que

$$E\left[\frac{1}{s-r-1}\sum_{i=r+1}^{s-1} e^{2\theta X_{(i)}^N}\Big|X_{(r)}, X_{(s)}\right] = \sqrt{g(X_{(r)})g(X_{(s)})} = \sqrt{e^{2\theta X_{(r)}^N}e^{2\theta X_{(s)}^N}}$$

$$= e^{\theta(X_{(r)}^N + X_{(s)}^N)}. \quad \blacksquare$$

3.4. Caracterizaciones de Wu y Ouyang

Hay más resultados que involucran esperanzas condicionales de estadísticas de orden, el siguiente Teorema nos da una caracterización en donde aparecen las estadísticas de orden $X_{(k)}$ y $X_{(r)}$ con $1 \leqslant k < r \leqslant n$.

Teorema 3.4.1. *Sean $\alpha \geqslant -\infty$ y $\beta \leqslant \infty$ los extremos izquierdo y derecho respectivamente de una distribución absolutamente continua F, g una función real continua en $[\alpha, \beta)$ con derivada g' definida y continua en (α, β).*
$0 < F(x) < 1$ para $\alpha < x < \beta$, $E[g(x)] < \infty$.
$E[g(X_{(r)})|X_{(k)} = t] = g(t) + c\sum_{j=k}^{r-1} \frac{1}{n-j} \quad 1 \leqslant k < r \leqslant n$ *si y sólo si*

$$F(x) = 1 - e^{-\frac{1}{c}(g(x)-g(\alpha))} \quad \alpha < x < \beta \text{ con } \lim_{x\to\beta^-} g(x) = \begin{cases} \infty & si\ c > 0 \\ -\infty & si\ c < 0. \end{cases}$$

Demostración:

$$E[g(X_{(r)})|X_{(k)} = t] = \tfrac{(n-k)!}{(r-k-1)!(n-k)!} \int_t^\beta \frac{[F(x) - F(t)]^{r-k-1}[1 - F(x)]^{n-r} g(x)}{[1 - F(t)]^{n-k}} dF(x).$$

Supongamos que $F(x) = 1 - e^{-\frac{1}{c}(g(x)-g(\alpha))}$.
Entonces $1-F(x) = e^{-\frac{1}{c}(g(x)-g(\alpha))}$ y $F(x)-F(t) = e^{-\frac{1}{c}(g(t)-g(\alpha))} - e^{-\frac{1}{c}(g(x)-g(\alpha))}$.
Por tanto

$$E[g(X_{(r)})|X_{(k)} = t] = \frac{(n-k)!}{(r-k-1)!(n-r)!} \frac{1}{[1 - F(t)]^{n-k}} \int_t^\beta g(x) e^{-\frac{(n-r)}{c}(g(x)-g(\alpha))} \cdot$$

$$\cdot \left[e^{-\frac{1}{c}(g(t)-g(\alpha))} - e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-1} e^{-\frac{1}{c}(g(x)-g(\alpha))} \frac{g'(x)}{c} dx =$$

$$\frac{(n-k)!}{(r-k-1)!(n-r)!}\frac{1}{[1-F(t)]^{n-k}}\int_t^\beta g(x)e^{-\frac{(n-r+1)}{c}(g(x)-g(\alpha))}\cdot$$
$$\cdot\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-1}\cdot\frac{g'(x)}{c}dx.$$

Sean

$u=g(x)\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-1}$. $dv=\frac{1}{c}g'(x)e^{-\frac{(n-r+1)}{c}(g(x)-g(\alpha))}dx.$

Asi $du=g'(x)\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-1}dx+$

$(r-k-1)g(x)\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-2}e^{-\frac{1}{c}(g(x)-g(\alpha))}\frac{g'(x)}{c}dx.$

$v=-\frac{1}{n-r-1}e^{-\frac{1}{c}(n-r+1)(g(x)-g(\alpha))}.$

Por lo que

$$E[g(X_{(r)})|X_{(k)}=t]=\frac{(n-k)!}{(r-k-1)!(n-r)!}\frac{1}{[1-F(t)]^{n-k}}\cdot\left\{uv\Big|_t^\beta-\int_t^\beta vdu\right\}=$$

$$D\left\{-\frac{g(x)}{n-r+1}\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-1}e^{-\frac{1}{c}(n-r+1)(g(x)-g(\alpha))}\Big|_t^\beta+\right.$$

$$\int_t^\beta\frac{1}{n-r+1}g'(x)\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-1}e^{-\frac{1}{c}(n-r+1)(g(x)-g(\alpha))}dx+$$

$$\left.\int_t^\beta\frac{r-k-1}{n-r+1}\frac{g(x)g'(x)}{c}\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-2}e^{-\frac{1}{c}(n-r+2)(g(x)-g(\alpha))}dx\right\}$$

donde $D=\frac{(n-k)!}{(r-k-1)!(n-r)!}\frac{1}{[1-F(t)]^{n-k}}$.

Sin pérdida de generalidad podemos suponer que $c>0$ y que $\lim_{x\to\beta^-}g(x)=\infty$, entonces

$E[g(X_{(r)})|X_{(k)}=t]=$

$$D\int_t^\beta \frac{1}{n-r+1}g'(x)\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-1}e^{-\frac{1}{c}(n-r+1)(g(x)-g(\alpha))}dx+$$

$$D\int_t^\beta \frac{r-k-1}{n-r+1}\frac{g(x)g'(x)}{c}\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-2}e^{-\frac{1}{c}(n-r+2)(g(x)-g(\alpha))}dx.$$

Notemos que

$$D\int_t^\beta \tfrac{r-k-1}{n-r+1}\tfrac{g(x)g'(x)}{c}\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-2}e^{-\frac{1}{c}(n-r+2)(g(x)-g(\alpha))}dx=$$

$$\tfrac{1}{[1-F(t)]^{n-k}}\tfrac{(n-k)!}{(r-k-2)!(n-r+1)!}\int_t^\beta \tfrac{g(x)g'(x)}{c}\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-2}\cdot$$

$$\cdot\, e^{-\frac{1}{c}(n-r+2)(g(x)-g(\alpha))}dx.$$

Pero por definición sabemos que

$E[g(X_{(r-1)})|X_{(k)}=t]=$

$$\tfrac{1}{[1-F(t)]^{n-k}}\tfrac{(n-k)!}{(r-k-2)!(n-r+1)!}\int_t^\beta \tfrac{g(x)g'(x)}{c}\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-2}\cdot$$

$$\cdot\, e^{-\frac{1}{c}(n-r+2)(g(x)-g(\alpha))}dx.$$

Luego

$$E[g(X_{(r)})|X_{(k)}=t]=\frac{D}{n-r+1}\int_t^\beta g'(x)\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-1}\cdot$$

$$\cdot\, e^{-\frac{1}{c}(n-r+1)(g(x)-g(\alpha))}dx+E[g(X_{(r-1)})|X_{(k)}=t].$$

Sean ahora

$u=\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-1}$, $dv=g'(x)e^{-\frac{1}{c}(n-r+1)(g(x)-g(\alpha))}dx.$

Asi $du=(r-k-1)\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-2}\dfrac{g'(x)}{c}e^{-\frac{1}{c}(g(x)-g(\alpha))}dx,$

$v=-\dfrac{c}{n-r+1}e^{-\frac{1}{c}(n-r+1)(g(x)-g(\alpha))}.$

Por lo que

$$E[g(X_{(r)})|X_{(k)}=t]=E[g(X_{(r-1)})|X_{(k)}=t]+$$

$$\frac{D}{n-r+1}\Bigg\{-\frac{c}{n-r+1}\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-1}e^{-\frac{1}{c}(n-r+1)(g(x)-g(\alpha))}\Big|_t^{\beta}+$$

$$\frac{r-k-1}{n-r+1}\int_t^{\beta}\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-2}g'(x)e^{-\frac{1}{c}(n-r+2)(g(x)-g(\alpha))}dx\Bigg\}=$$

$$E[g(X_{(r-1)})|X_{(k)}=t]+$$

$$\frac{D(r-k-1)}{(n-r+1)(n-r+1)}\int_t^{\beta}\left[e^{-\frac{1}{c}(g(t)-g(\alpha))}-e^{-\frac{1}{c}(g(x)-g(\alpha))}\right]^{r-k-2}g'(x)e^{-\frac{1}{c}(n-r+2)(g(x)-g(\alpha))}dx.$$

Si seguimos integrando por partes obtenemos:

$$E[g(X_{(r)})|X_{(k)}=t]=E[g(X_{(r-1)})|X_{(k)}=t]+$$

$$\frac{D(r-k-1)(r-k-2)\cdots 2\cdot 1}{(n-r+1)(n-r+1)(n-r+2)\cdots(n-k-1)}\int_t^{\beta}g'(x)e^{-\frac{1}{c}(n-k)(g(x)-g(\alpha))}dx=E[g(X_{(r-1)})|X_{(k)}=t]+$$

$$\frac{(n-k)!(r-k-1)(r-k-2)\cdots 2\cdot 1}{(r-k-1)!(n-r)!(n-r-1)(n-r-1)(n-r-2)\cdots(n-k-1)}\cdot\frac{1}{[1-F(t)]^{n-k}}\int_t^{\beta}g'(x)e^{-\frac{1}{c}(n-k)(g(x)-g(\alpha))}dx=$$

$$E[g(X_{(r-1)})|X_{(k)}=t]+\frac{(n-k)!}{(n-r+1)(n-k-1)!}\frac{1}{[1-F(t)]^{n-k}}\cdot\int_t^{\beta}g'(x)e^{-\frac{1}{c}(n-k)(g(x)-g(\alpha))}dx=$$

$$E[g(X_{(r-1)})|X_{(k)}=t]+\frac{(n-k)}{(n-r+1)}\frac{1}{[1-F(t)]^{n-k}}\left[-\frac{c}{(n-k)}e^{-\frac{1}{c}(n-k)(g(x)-g(\alpha))}\Big|_t^{\beta}\right]=$$

$$E[g(X_{(r-1)})|X_{(k)}=t]+\frac{c}{n-r+1}\frac{e^{-\frac{1}{c}(n-k)(g(t)-g(\alpha))}}{[1-F(t)]^{n-k}}.$$

Pero por hipótesis $F(t)=1-e^{-\frac{1}{c}(g(t)-g(\alpha))}$ entonces $[1-F(t)]^{n-k}=e^{-\frac{1}{c}(n-k)(g(t)-g(\alpha))}$.

De este modo:

$$E[g(X_{(r)})|X_{(k)}=t]=E[g(X_{(r-1)})|X_{(k)}=t]+\frac{c}{n-r+1}.$$

Y usando esta expresión de manera recurrente tenemos:

$E[g(X_{(r)})|X_{(k)} = t] = \dfrac{c}{n-r+1} + E[g(X_{(r-2)})|X_{(k)} = t] + \dfrac{c}{n-r+2} = c\left(\dfrac{1}{n-r+1}+\dfrac{1}{n-r+2}\right) + E[g(X_{(r-2)})|X_{(k)} = t] = \cdots = c\Big(\dfrac{1}{n-r+1} + \dfrac{1}{n-r+2}+\dfrac{1}{n-r+3}+\cdots+\dfrac{1}{n-k}\Big)+E[g(X_{(k)})|X_{(k)} = t] = g(t)+c\displaystyle\sum_{j=k}^{r-1}\frac{1}{n-j}$

con $1 \leqslant k < r \leqslant n$.

Ahora supongamos que $E[g(X_{(r)})|X_{(k)} = t] = g(t) + c\displaystyle\sum_{j=k}^{r-1}\frac{1}{n-j}$ para $1 \leqslant k < r \leqslant n, \quad \alpha < t < \beta$.

Pero se tiene también que

$E[g(X_{(r)})|X_{(k)} = t] = \dfrac{(n-k)!}{(r-k-1)!(n-r)!}\dfrac{1}{[1-F(t)]^{n-k}} \cdot \displaystyle\int_t^\beta g(x)[F(x) - F(t)]^{r-k-1}[1-F(x)]^{n-r}dF(x)$.

Entonces

$\left[g(t) + c\displaystyle\sum_{j=k}^{r-1}\frac{1}{n-j}\right][1 - F(t)]^{n-k} = \dfrac{(n-k)!}{(r-k-1)!(n-r)!} \cdot \displaystyle\int_t^\beta g(x)[F(x) - F(t)]^{r-k-1}[1-F(x)]^{n-r}dF(x)$.

Sean $u = g(x)[F(x) - F(t)]^{r-k-1} \quad dv = [1-F(x)]^{n-r}dF(x)$.

Entonces $du = (r-k-1)g(x)[F(x)-F(t)]^{r-k-2}dF(x)+g'(x)[F(x)-F(t)]^{r-k-1}dx$

$v = -\dfrac{1}{n-r+1}[1-F(x)]^{n-r+1}$

Por lo que

$\left[g(t) + c\displaystyle\sum_{j=k}^{r-1}\frac{1}{n-j}\right][1 - F(t)]^{n-k} =$

$$\frac{(n-k)!}{(r-k-1)!(n-r)!}\left\{-\frac{g(x)}{n-r+1}[1-F(x)]^{n-r+1}[F(x)-F(t)]^{r-k-1}\Big|_t^{\beta}+\right.$$

$$\frac{1}{n-r+1}\int_t^{\beta} g'(x)[1-F(x)]^{n-r+1}[F(x)-F(t)]^{r-k-1}dx+$$

$$\left.\frac{r-k-1}{n-r+1}\int_t^{\beta} g(x)[1-F(x)]^{n-r+1}[F(x)-F(t)]^{r-k-2}dF(x)\right\}.$$

Como $\lim\limits_{x\to\beta^-} F(x)=1$ tenemos

$$\left[g(t)+c\sum_{j=k}^{r-1}\frac{1}{n-j}\right][1-F(t)]^{n-k}=$$

$$\frac{(n-k)!}{(r-k-2)!(n-r+1)!}\int_t^{\beta} g(x)[1-F(x)]^{n-r+1}[F(x)-F(t)]^{r-k-2}dF(x)+$$

$$\frac{(n-k)!}{(r-k-1)!(n-r+1)!}\int_t^{\beta} g'(x)[1-F(x)]^{n-r+1}[F(x)-F(t)]^{r-k-1}dx.$$

Por otro lado sabemos por hipótesis que $\left[g(t)+c\sum_{j=k}^{r-2}\frac{1}{n-j}\right][1-F(t)]^{n-k}=$

$$\frac{(n-k)!}{(r-k-2)!(n-r+1)!}\int_t^{\beta} g(x)[1-F(x)]^{n-r+1}[F(x)-F(t)]^{r-k-2}dF(x).$$

De este modo

$$\left[g(t)+c\sum_{j=k}^{r-1}\frac{1}{n-j}\right][1-F(t)]^{n-k}=\left[g(t)+c\sum_{j=k}^{r-2}\frac{1}{n-j}\right][1-F(t)]^{n-k}+$$

$$\frac{(n-k)!}{(r-k-1)!(n-r+1)!}\int_t^{\beta} g'(x)[1-F(x)]^{n-r+1}[F(x)-F(t)]^{r-k-1}dx.$$

Entonces

$$\frac{c}{n-r+1}[1-F(t)]^{n-k}=$$

$$\frac{(n-k)!}{(r-k-1)!(n-r+1)!}\int_t^\beta g'(x)[1-F(x)]^{n-r+1}[F(x)-F(t)]^{r-k-1}dx.$$

Recordemos que como consecuencia de la regla de la cadena, se tiene que:

Si $H(y)=\int_{a(y)}^{b(y)} h(x,y)dx$ entonces $H'(y)=\int_{a(y)}^{b(y)} \frac{\partial}{\partial y}h(x,y)dx+$

$h(b(y),y)b'(y)-h(a(y),y)a'(y)$, la demostración de esta fórmula aparece en Sagan (1974).

De este modo si

$$H(t)=\int_t^\beta g'(x)[F(x)-F(t)]^{r-k-1}[1-F(x)]^{n-r+1}dx$$

entonces $H'(t)=-\int_t^\beta (r-k-1)g'(x)[F(x)-F(t)]^{r-k-2}[1-F(x)]^{n-r+1}f(t)dx.$

Es decir $\frac{d}{dt}\Big(\int_t^\beta g'(x)[F(x)-F(t)]^{r-k-1}[1-F(x)]^{n-r+1}dx\Big)=$

$-\int_t^\beta (r-k-1)g'(x)[F(x)-F(t)]^{r-k-2}[1-F(x)]^{n-r+1}f(t)dx\Big).$

Por lo que $\frac{d}{dt}\left[\frac{c}{n-r+1}(1-F(t))^{n-k}\right]=$

$\frac{(n-k)!}{(r-k-1)!(n-r+1)!}\frac{d}{dt}\int_t^\beta g'(x)[F(x)-F(t)]^{r-k-1}[1-F(x)]^{n-r+1}dx.$

Luego $-\frac{c(n-k)}{n-r+1}[1-F(t)]^{n-k-1}f(t)=$

$-\frac{(n-k)!(r-k-1)}{(r-k-1)!(n-r+1)!}\int_t^\beta g'(x)[F(x)-F(t)]^{r-k-2}[1-F(x)]^{n-r+1}f(t)dx.$

Entonces $\frac{c(n-k)}{n-r+1}[1-F(t)]^{n-k-1} =$
$\frac{(n-k)!(r-k-1)}{(r-k-1)!(n-r+1)!}\int_t^\beta g'(x)[F(x)-F(t)]^{r-k-2}[1-F(x)]^{n-r+1}dx.$
Derivando de nuevo ambos miembros con respecto a t se tiene

$$-\frac{c(n-k)(n-k-1)}{n-r+1}[1-F(t)]^{n-k-2}f(t) = -\frac{(n-k)!(r-k-1)(r-k-2)}{(r-k-1)!(n-r+1)!}\cdot \int_t^\beta g'(x)[F(x)-F(t)]^{r-k-3}[1-F(x)]^{n-r+1}f(t)dx.$$

Asi $\frac{c(n-k)(n-k-1)}{n-r+1}[1-F(t)]^{n-k-2} = \frac{(n-k)!(r-k-1)(r-k-2)}{(r-k-1)!(n-r+1)!}\cdot$
$\int_t^\beta g'(x)[F(x)-F(t)]^{r-k-3}[1-F(x)]^{n-r+1}dx.$

Si seguimos derivando

$$\frac{c(n-k)(n-k-1)\cdots(n-r+2)}{n-r+1}[1-F(t)]^{n-r+1} = \frac{(n-k)!(r-k-1)(r-k-2)\cdots 2\cdot 1}{(r-k-1)!(n-r+1)!}\cdot\int_t^\beta g'(x)[1-F(x)]^{n-r+1}dx.$$

De este modo $\frac{c(n-k)(n-k-1)\cdots(n-r+2)}{n-r+1}[1-F(t)]^{n-r+1} =$
$\frac{(n-k)!}{(n-r+1)!}\int_t^\beta g'(x)[1-F(x)]^{n-r+1}dx.$
Y tomando una última derivada

$$-\frac{c(n-k)(n-k-1)\cdots(n-r+2)(n-r+1)}{n-r+1}[1-F(t)]^{n-r}f(t) = -\frac{(n-k)!}{(n-r+1)!}g'(t)[1-F(x)]^{n-r+1}.$$

Luego $\frac{c(n-k)!}{(n-r+1)!}[1-F(t)]^{n-r}f(t) = \frac{(n-k)!}{(n-r+1)!}g'(t)[1-F(t)]^{n-r+1},$

es decir, $cf(t) = g'(t)[1 - F(t)]$ implica que $\dfrac{f(t)}{1-F(t)} = \dfrac{g'(t)}{c}$.

Entonces $\displaystyle\int_{\alpha}^{x} \frac{f(t)}{1-F(t)}dt = \int_{\alpha}^{x} \frac{g'(t)}{c}dt$.

Así tenemos que $\ln[1-F(t)]\Big|_{x}^{\alpha} = \dfrac{1}{c}(g(x) - g(\alpha))$.

Por lo tanto $-\ln(1-F(x)) = \dfrac{1}{c}(g(x) - g(\alpha))$.

Entonces $\ln(1-F(x)) = -\dfrac{1}{c}(g(x) - g(\alpha))$,

y luego $1 - F(x) = e^{-\frac{1}{c}(g(x)-g(\alpha))}$.

Concluyendo que $F(x) = 1 - e^{-\frac{1}{c}(g(x)-g(\alpha))}$.

Ahora tomando límite cuando $x \longrightarrow \beta^{-}$ tenemos

$1 = 1 - e^{-\frac{1}{c}\lim\limits_{x\to\beta^-}(g(x)-g(\alpha))}$ implica que $-\dfrac{1}{c}\lim\limits_{x\to\beta^-}(g(x) - g(\alpha)) = -\infty$.

Asi $\lim\limits_{x\to\beta^-}(g(x) - g(\alpha)) = \infty$.

Por lo que $\dfrac{1}{C}\lim\limits_{x\to\beta^-} g(x) = \infty$.

Entonces $\lim\limits_{x\to\beta^-} g(x) = \begin{cases} \infty & \text{si } c > 0 \\ -\infty & \text{si } c < 0. \end{cases}$ $\square$

Ejemplo 3.4.2. La pareja formada por $g(x) = \ln(x)$ $\alpha \leqslant x < \infty$ con $\alpha > 0$ y la función de distribución $F(x) = 1 - \left(\dfrac{\alpha}{x}\right)^b$, $\alpha < x < \infty$, $b > 1$ satisfacen el Teorema 3.4.1, porque $\beta = \infty$, $g(x) = \ln x$ es continua en $[\alpha, \infty)$, $g'(x) = \dfrac{1}{x}$ es también continua en (α, ∞).

$F(x)$ es de distribución absolutamente continua en (α, ∞) porque $\lim\limits_{x\to\alpha^+} F(x) = 1 - \left(\dfrac{\alpha}{\alpha}\right)^b = 1 - 1 = 0$.

$\lim\limits_{x\to\infty} F(x) = \lim\limits_{x\to\infty}\left[1 - \left(\dfrac{\alpha}{x}\right)\right]^b = 1$, $\lim\limits_{x\to x_0} F(x) = \lim\limits_{x\to x_0}\left[1 - \left(\dfrac{\alpha}{x}\right)^b\right] = 1 - \left(\dfrac{\alpha}{\alpha}\right)^b = F(x_0)$ si $\alpha < x_0 < \infty$.

Y si tomamos $\alpha < x_1 < x_2$ entonces $\dfrac{1}{x_2} < \dfrac{1}{x_1} < \dfrac{1}{\alpha}$.
Asi $\dfrac{\alpha}{x_2} < \dfrac{\alpha}{x_1} < 1$
entonces $\left(\dfrac{\alpha}{x_2}\right)^b < \left(\dfrac{\alpha}{x_1}\right)^b < 1$.
Por lo que $-\left(\dfrac{\alpha}{x_1}\right)^b < -\left(\dfrac{\alpha}{x_2}\right)^b$.
De este modo $1-\left(\dfrac{\alpha}{x_1}\right)^b < 1-\left(\dfrac{\alpha}{x_2}\right)^b$.

Por lo tanto $F(x_1) < F(x_2)$ y concluimos que F es función de distribución continua.
Solo tenemos que verificar que $E[|g(X)|] < \infty$, esto es, $\int_\alpha^\infty |g(x)|dF(x) < \infty$, lo que significa que $\int_\alpha^\infty |\ln x| d\left[1-\left(\frac{\alpha}{x}\right)^b\right] < \infty$ hay dos casos:

1) Si $\alpha \geqslant 1$ entonces $\int_\alpha^\infty |\ln x| d\left[1-\left(\frac{\alpha}{x}\right)^b\right] = \int_\alpha^\infty \ln x \left(\frac{b\alpha^b}{x^{b+1}}\right)^b dx$
$= b\alpha^b \int_\alpha^\infty \frac{\ln x}{x^{b+1}} dx.$
Sean $u = \ln x$, $du = dx$, $dv = x^{-b-1}$, $b = -\frac{1}{b}x^{-b}$
por lo que $\int_\alpha^\infty |\ln x| d\left[1-\left(\frac{\alpha}{x}\right)^b\right] = b\alpha^b \left[-\frac{1}{b}\frac{\ln x}{x^b}\Big|_\alpha^\infty + \frac{1}{b}\int_\alpha^\infty x^{-b}dx\right] =$
$\alpha^b \left[\frac{\ln\alpha}{\alpha^b} + \frac{x^{-b+1}}{(-b+1)}\Big|_\alpha^\infty\right] = \ln\alpha - \frac{\alpha^{-b+1}}{1-b} < \infty.$

De igual manera se hace el caso 2), así se tiene que:
$E[\ln(X_{(r)})|X_{(k)} = t] = \ln t + c\sum_{j=k}^{r-1}\frac{1}{n-j}$ para $1 \leqslant k < r \leqslant n$ si y sólo si $F(x) = 1 - e^{-\frac{1}{c}(g(x)-g(\alpha))}$.

Pero $F(x) = 1 - \left(\frac{\alpha}{x}\right)^b = 1 - e^{b\ln(\alpha/x)} = 1 - e^{b(\ln\alpha - \ln x)} = 1 - e^{-b(\ln x - \ln\alpha)}$ implica que $b = \frac{1}{c}$ ó $c = \frac{1}{b}$. $\blacksquare$

3.5. Caracterizaciones de Zoroa, Ruiz y Marin

Veremos ahora un problema de caracterización diferente: si h es una función real, continua y estrictamente monótona, denotamos por $\mathcal{L}_h$ el conjunto de las funciones de distribución reales y continuas que cumplen con que la integral de Riemann-Stieltjes $\int_a^\infty h(x)dF_X(x)$ es finita para cualquier $a \in \mathbb{R}$. Para cada $F_X \in \mathcal{L}_h$ definimos la función ψ como:

$$\psi(x) = E[h(X)|X \geqslant x] = \frac{1}{1 - F_X(x)} \int_x^\infty h(t)dF_X(t).$$

Notamos que $\psi(x)$ está definida en $D = \mathbb{R}$ o en $D = (-\infty, \beta)$ con $\beta \in \mathbb{R}$ en donde $\beta = \inf\{x : F_X(x) = 1\}$.

Implícitamente existe un mapeo T que va de $\mathcal{L}_h$ en el conjunto de las funciones reales, definido por $T(F) = \psi$ con $F_X \in \mathcal{L}_h$, la pregunta que contestaremos a continuación, es: ¿Cuáles son las condiciones necesarias y suficientes para que una función real ψ pertenezca a la imagen de T? o en otras palabras, ¿qué condiciones tiene que cumplir ψ función real para que exista una función de distribución real F_X tal que $\psi(x) = \frac{1}{1 - F_X(x)} \int_x^\infty h(t)dF_X(t)$ donde h está dada?

Estableceremos las condiciones necesarias y suficientes que debe cumplir ψ para la existencia de una función de distribución F tal que $\psi = T(F)$. Primero daremos un Lema y luego el teorema principal de esta sección.

Lema 3.5.1. *Sea h una función real, continua, estrictamente monótona y ψ cualquier función real que satisface las siguientes condiciones:*

i) Su dominio D es $\mathbb{R}$ o en su caso $(-\infty, \beta)$ con $\beta \in \mathbb{R}$.

ii) Es continua.

iii) a) Si h es estrictamente creciente entonces $\psi(x) > h(x)$ para toda $x \in D$ y si $D = (-\infty, \beta)$ entonces $h(\beta) > \psi(x)$ para todo $x \in D$.

b) Si h es estrictamente decreciente entonces $\psi(x) < h(x)$ para todo $x \in D$ y si $D = (-\infty, \beta)$ entonces $h(\beta) < \psi(x)$ para todo $x \in D$.

iv) ψ es creciente si h es estrictamente creciente, ψ es decreciente si h es estrictamente decreciente.

v) La integral de Riemann-Stieltjes $\int\limits_{-\infty}^{x} \frac{d\psi(t)}{\psi(t) - h(t)}$ converge para todo $x \in D$.

Entonces la función $G(x) = e^{-\int_{-\infty}^{x} \frac{d\psi(t)}{\psi(t)-h(t)}}$ satisface la igualdad $G(c)\psi(c) - G(b)\psi(b) = \int\limits_{b}^{c} h(t)dG(t)$ para todo $b < c$ con $b.c \in D$.

Demostración:

Supongamos h estrictamente creciente, el caso estrictamente decreciente se haría de la misma manera.

Por $v)$ $\int\limits_{-\infty}^{x} \frac{d\psi(t)}{\psi(t) - h(t)} \in \mathbb{R}$ para todo $x \in \mathbb{R}$.

Por lo que $G(x) = e^{-\int_{-\infty}^{x} \frac{d\psi(t)}{\psi(t)-h(t)}}$ existe.

Si tomamos $x < y$ veremos que $G(x) \geqslant G(y)$, es decir,

$e^{-\int_{-\infty}^{x} \frac{d\psi(t)}{\psi(t)-h(t)}} \geqslant e^{-\int_{-\infty}^{y} \frac{d\psi(t)}{\psi(t)-h(t)}}$ si y sólo si

$\int\limits_{-\infty}^{x} \frac{d\psi(t)}{\psi(t) - h(t)} \leqslant \int\limits_{-\infty}^{y} \frac{d\psi(t)}{\psi(t) - h(t)}$ si y sólo si $\int\limits_{x}^{y} \frac{d\psi(t)}{\psi(t) - h(t)} \geqslant 0.$

Como ψ es creciente, continua y además $\psi(t) > h(t)$, se satisface la desigualdad anterior.

$G(x+\triangle x) - G(x) = e^{-\int_{-\infty}^{x+\triangle x} \frac{d\psi(t)}{\psi(t)-h(t)}} - e^{-\int_{-\infty}^{x} \frac{d\psi(t)}{\psi(t)-h(t)}} =$
$e^{-\int_{-\infty}^{x} \frac{d\psi(t)}{\psi(t)-h(t)}} \left[e^{-\int_{x}^{x+\triangle x} \frac{d\psi(t)}{\psi(t)-h(t)}} - 1 \right].$

Pero $\int_{x}^{x+\triangle x} \frac{d\psi(t)}{\psi(t)-h(t)} = [\psi(x+\triangle x) - \psi(x)] \frac{1}{\psi(\eta) - h(\eta)} \xrightarrow[\triangle x \to 0]{} 0$ por la continuidad de ψ donde $x \leqslant \eta \leqslant x + \triangle x$ entonces $G(x)$ es continua.

Como $G(x)$ es decreciente y continua tenemos que $\int_{b}^{c} h(t) dG(t)$ existe.

Sean $\pi = \{t_0, t_1, \ldots t_n : b = t_0 < t_1 < \cdots, < t_n = c\}$ una partición de $[b, c]$, $N(\pi) = \text{máx}\{t_{i+1} - t_i : 0 \leqslant i \leqslant n-1\}$ la norma de la partición, $\triangle G(t_i) = G(t_{i+1}) - G(t_i)$, $\triangle \psi(t_i) = \psi(t_{i+1}) - \psi(t_i)$, $I = \int_{b}^{c} h(t) dG(t)$,

$I(\pi) = \sum_{i=0}^{n-1} h(t_{i+1}) \triangle G(t_i)$ la aproximación a I usando la partición π, (ver Rudin, 1980).

$Z_i = \int_{t_i}^{t_{i+1}} \frac{d\psi(u)}{\psi(u) - h(u)}$ entonces

$\sum_{i=0}^{n-1} [\psi(t_{i+1}) - h(t_{i+1})] \triangle G(t_i) + \sum_{i=0}^{n-1} G(t_i) \triangle \psi(t_i) = \sum_{i=0}^{n-1} [\psi(t_{i+1}) - h(t_{i+1})][G(t_{i+1}) -$
$G(t_i)] + \sum_{i=0}^{n-1} G(t_i)[\psi(t_i) - \psi(t_i)] = -I(\pi) + \sum_{i=0}^{n-1} \psi(t_{i+1}) G(t_{i+1}) - \sum_{i=0}^{n-1} \psi(t_i) G(t_i) =$
$-I(\pi) + \psi(c)G(c) - \psi(b)G(b),$

por otro lado

$\triangle G(t_i) = G(t_{i+1}) - G(t_i) = e^{-\int_{-\infty}^{t_{i+1}} \frac{d\psi(u)}{\psi(u)-h(u)}} - e^{-\int_{-\infty}^{t_i} \frac{d\psi(u)}{\psi(u)-h(u)}} =$

$$e^{-\int_{-\infty}^{t_i} \frac{d\psi(u)}{\psi(u)-h(u)}} e^{-\int_{t_i}^{t_{i+1}} \frac{d\psi(u)}{\psi(u)-h(u)}} - e^{-\int_{-\infty}^{t_i} \frac{d\psi(u)}{\psi(u)-h(u)}} = G(t_i)\left[e^{-\int_{t_i}^{t_{i+1}} \frac{d\psi(u)}{\psi(u)-h(u)}}\right] =$$

$G(t_i)[e^{-z_i} - 1]$.

Por lo tanto $\triangle G(t_i) = G(t_i)[e^{-z_i} - 1]$.

Como $e^{-z} = 1 - z + \rho z^2$ con $|\rho| \leqslant 1$ para $|z| < 1/2$ entonces

$$G(c)\psi(c) - G(b)\psi(b) - I(\pi) = \sum_{i=0}^{n-1}[\psi(t_{i+1}) - h(t_{i+1})]\triangle G(t_i) + \sum_{i=0}^{n-1} G(t_i)\triangle\psi(t_i) =$$
$$\sum_{i=0}^{n-1}[\psi(t_{i+1}) - h(t_{i+1})]G(t_i)(e^{-z_i} - 1) + \sum_{i=0}^{n-1} G(t_i)\triangle\psi(t_i) = \sum_{i=0}^{n-1}[\psi(t_{i+1}) - h(t_{i+1})]G(t_i)\cdot$$
$$(-z_i + \rho z_i^2) + \sum_{i=0}^{n-1} G(t_i)\triangle\psi(t_i) = \sum_{i=0}^{n-1}[\psi(t_{i+1}) - h(t_{i+1})]G(t_i)(-z_i + \rho z_i^2) +$$
$$\sum_{i=0}^{n-1} G(t_i)\triangle\psi(t_i)\left[\frac{\psi(t_{i+1}) - h(t_{i+1})}{\psi(t_{i+1}) - h(t_{i+1})}\right] = \sum_{i=0}^{n-1} G(t_i)[\psi(t_{i+1}) - h(t_{i+1})]\Big[- z_i +$$
$$\frac{\triangle\psi(t_i)}{\psi(t_{i+1}) - h(t_{i+1})}\Big] + \sum_{i=0}^{n-1} G(t_i)[\psi(t_{i+1}) - h(t_{i+1})]\rho z_i^2 = R + S$$

donde $R = \sum_{i=0}^{n-1} G(t_i)[\psi(t_{i+1}) - h(t_{i+1})]\left[- z_i + \frac{\triangle\psi(t_i)}{\psi(t_{i+1}) - h(t_{i+1})}\right]$

y $S = \sum_{i=0}^{n-1} G(t_i)[\psi(t_{i+1}) - h(t_{i+1})]\rho z_i^2$.

Así $G(c)\psi(c) - G(b)\psi(b) - I(\pi) = R + S$, ahora sea $T = I(\pi) - I$, es decir, $I(\pi) = T + I$.

Por lo tanto $G(c)\psi(c) - G(b)\psi(b) - T - I = R + S$,

es decir $G(c)\psi(c) - G(b)\psi(b) - I = R + S + T$

no olvidar que hemos supuesto que $|z_i| < 1/2$ para todo i.

Tomemos $\epsilon > 0$ con $0 < \epsilon < \frac{1}{2}$, sabemos que se cumple:

a) Por definición de integral de Riemann-Stieltjes existe $\delta_1 > 0$ tal que si

$N(\pi) < \delta_1$ entonces $|T| = |I - I(\pi)| < \epsilon$.

b) Como la función $\dfrac{1}{\psi(x) - h(x)}$ es uniformemente continua en $[b, c]$ (ψ y h son continuas, $\psi - h$ es continua, $\psi(x) \neq h(x)$ para todo x.

Por lo tanto $\frac{1}{\psi - h}$es continua y cualquier función continua en un intervalo cerrado es uniformemente continua) entonces existe $\delta_2 > 0$ tal que si $N(\pi) < \delta_2$ se tiene

$$\left| -z_i + \frac{\triangle\psi(t_i)}{\psi(t_{i+1}) - h(t_{i+1})} \right| = \left| \frac{\triangle\psi(t_i)}{\psi(t_{i+1}) - h(t_{i+1})} - \int_{t_i}^{t_{i+1}} \frac{d\psi(u)}{\psi(u) - h(u)} \right| =$$

$$\left| \frac{\triangle\psi(t_i)}{\psi(t_{i+1}) - h(t_{i+1})} - \frac{\psi(t_{i+1}) - \psi(t_i)}{\psi(\eta) - h(\eta)} \right| = \left| \frac{\triangle\psi(t_i)}{\psi(t_{i+1}) - h(t_{i+1})} - \frac{\triangle\psi(t_i)}{\psi(\eta) - h(\eta)} \right| =$$

$$\triangle\psi(t_i) \left| \frac{1}{\psi(t_{i+1}) - h(t_{i+1})} - \frac{1}{\psi(\eta) - h(\eta)} \right| \leqslant \triangle\psi(t_i)\epsilon,$$

donde $t_i \leqslant \eta \leqslant t_{i+1}$ se obtiene del teorema del valor intermedio.

Por lo tanto $|R| = \left| \sum_{i=0}^{n-1} G(t_i)[\psi(t_{i+1}) - h(t_{i+1})] \left[-z_i + \frac{\triangle\psi(t_i)}{\psi(t_{i+1}) - h(t_{i+1})} \right] \right| \leqslant$

$\sum_{i=0}^{n-1} G(t_i)[\psi(t_{i+1}) - h(t_{i+1})]\triangle\psi(t_i)\epsilon$ por la desigualdad del triángulo y porque $\psi(x) > h(x)$ para todo $x \in D$.

Por lo tanto $|R| \leqslant \sum_{i=0}^{n-1} G(t_i)[\psi(t_{i+1}) - h(t_{i+1})]\triangle\psi(t_i)\epsilon = \epsilon \sum_{i=0}^{n-1} G(t_i)[\psi(t_{i+1}) - h(t_{i+1})]\triangle\psi(t_i)$.

Como ψ es creciente se tiene que

$\psi(b) \leqslant \psi(t_1) \leqslant \cdots \leqslant \psi(t_{n-1}) \leqslant \psi(c)$.

Y como h es estrictamente creciente

$h(b) < h(t_1) < \cdots < h(t_{n-1}) < h(c)$

implica que $-h(c) < -h(t_{n-1}) < \cdots < -h(t_1) < -h(b)$ entonces

$|R| \leqslant \epsilon \sum_{i=0}^{n-1} G(t_i)[\psi(c) - h(b)]\triangle\psi(t_i).$

Ahora como $G(x)$ es decreciente

$G(c) \leqslant G(t_{n-1}) \leqslant \cdots \leqslant G(t_1) \leqslant G(b)$ obtenemos

$|R| \leqslant \epsilon \sum_{i=0}^{n-1} G(b)[\psi(c) - h(b)]\triangle\psi(t_i) = \epsilon G(b)[\psi(c) - h(b)] \sum_{i=0}^{n-1}[\psi(t_{i+1}) - h(t_i)] = \epsilon G(b)[\psi(c) - h(b)][\psi(c) - \psi(b)].$

Por lo que

$|R| \leqslant \epsilon G(b)[\psi(c) - h(b)][\psi(c) - \psi(b)] = \epsilon P.$

c) Por la continuidad uniforme de $f(t) = \int_b^t \frac{d\psi(u)}{\psi(u) - h(u)}$ sobre $[b, c]$ existe un $\delta_3 > 0$ tal que si $N(\pi) < \delta_3$ tenemos que $0 < z_i = \int_{t_i}^{t_{i+1}} \frac{d\psi(u)}{\psi(u) - h(u)} = f(t_{i+1}) - f(t_i) < \epsilon < \frac{1}{2}.$

Pero por hipótesis $|z_i| < \frac{1}{2}$ y entonces

$G(c)\psi(c) - G(b)\psi(b) - I = R + S + T,$

mientras que $S = \sum_{i=0}^{n-1} G(t_i)[\psi(t_{i+1}) - h(t_{i+1})]\rho z_i^2 < \sum_{i=0}^{n-1} G(b)[\psi(c) - h(b)]\rho z_i^2$ por la argumentación de b),

es decir, $S < G(b)[\psi(c) - h(b)] \sum_{i=0}^{n-1} \rho z_i^2$ pero $|\rho| \leqslant 1$.

Por lo tanto $S < G(b)[\psi(c) - h(b)] \sum_{i=0}^{n-1} z_i^2$ también $0 < z_i < \epsilon$ entonces

$S < G(b)[\psi(c) - h(b)]\epsilon \sum_{i=0}^{n-1} z_i.$

Y como $z_i = \int_{t_i}^{t_{i+1}} \frac{d\psi(u)}{\psi(u) - h(u)}$ se cumple que $\sum_{i=0}^{n-1} z_i = \int_b^c \frac{d\psi(u)}{\psi(u) - h(u)}.$

Por lo tanto $S < G(b)[\psi(c) - h(b)]\epsilon \int\limits_b^c \frac{d\psi(u)}{\psi(u) - h(u)} = Q\epsilon.$

Ahora si consideramos una partición π del intervalo $[b, c]$ tal que $N(\pi) < \text{mín}\{\delta_1, \delta_2, \delta_3\}$, como consecuencia de a), b) y c) tenemos que $|T| < \epsilon$, $|R| < \epsilon P$, $S = |S| < \epsilon Q$.

Entonces $|G(c)\psi(c) - G(b)\psi(b) - I| = |R + S + T| < (P + Q + 1)\epsilon$

donde $P = G(b)[\psi(c) - \psi(b)][\psi(c) - h(b)]$.

$Q = G(b)[\psi(c) - h(b)] \int\limits_b^c \frac{d\psi(u)}{\psi(u) - h(u)}$

$|T| = I - I(\pi)$.

Por lo tanto $\int\limits_b^c h(t)dE(t) = G(c)\psi(c) - G(b)\psi(b)$. $\square$

Teorema 3.5.2. *Sean h una función real, continua y estrictamente monótona y ψ una función real.*

$\psi \in I_m(T)$ si y sólo si se cumple $i), \ldots . v)$ del Lema 3.5.1 y además

vi) La integral de Riemann-Stieltjes $\int\limits_D \frac{d\psi(t)}{\psi(t) - h(t)}$ *diverge.*

vii) Si $D = \mathbb{R}$ entonces $\lim\limits_{x\to\infty} \psi(x)e^{-\int_{-\infty}^{x} \frac{d\psi(t)}{\psi(t)-h(t)}} = 0.$

Demostración:

Supongamos que se cumplen $i), \ldots, vii)$ para una función ψ^* cuyo dominio es D^*.

Definimos la función $F^*(x)$ por $F^*(x) = 1 - e^{-\int_{-\infty}^{x} \frac{d\psi^*(t)}{\psi^*(t)-h(t)}}$ si $x \in D^*$ y además $F^*(x) = 1$ si $x \geqslant \beta$ cuando $D^* = (-\infty, \beta)$.

La defición de F^* tiene sentido porque por $v)$ $\int\limits_{-\infty}^{x} \frac{d\psi^*(t)}{\psi^*(t) - h(t)}$ converge para todo $x \in D^*$.

Notemos que $F^*(x) = 1 - G(x)$ y como $G(x)$ es continua y decreciente entonces $F^*(x) = 1 - G(x)$ es continua y creciente.

Como $\psi^*(x) = \frac{1}{1-F^*(x)} \int_x^{\infty} h(t)dF^*(t) = \frac{1}{1-F^*(x)} \int_x^{\infty} h(t)d(1-G(t))$ entonces $D^* = \{x \in \mathbb{R} : F^*(x) < 1\}$, por otro lado $\lim\limits_{x \to \infty} F^*(x) = 1 - \lim\limits_{x \to \infty} e^{-\int_{-\infty}^{x} \frac{d\psi^*(t)}{\psi^*(t)-h(t)}} = 1 - e^{-\int_D \frac{d\psi^*(t)}{\psi^*(t)-h(t)}} = 1$ porque $\int_D \frac{d\psi^*(t)}{\psi^*(t)-h(t)} = \infty$ (condición vi)).

$\lim\limits_{x \to -\infty} F^*(x) = 1 - \lim\limits_{x \to -\infty} e^{-\int_{-\infty}^{x} \frac{d\psi^*(t)}{\psi^*(t)-h(t)}} = 1 - e^0 = 0.$

Por lo tanto $F^*(x)$ es función de distribución real absolutamente continua.

Por otro lado las hipótesis $i), \ldots, v)$ del Lema 3.5.1 se cumplen para ψ^*.

En este caso sabemos que $G(x) = 1 - F^*(x)$ y la conclusión del Lema 3.5.1 es $G(c)\psi(c) - G(b)\psi(b) = \int_b^c h(t)dG(t)$ entonces

$[1 - F^*(c)]\psi^*(c) - [1 - F^*(b)]\psi^*(b) = -\int_b^c h(t)dF^*(t)$ con $b < c$, $b, c, \in D^*$.

En el caso $D^* = \mathbb{R}$, tomamos el límite cuando $c \longrightarrow \infty$:

$\lim\limits_{c \to \infty} [1-F^*(c)]\psi^*(c) - [1-F^*(b)]\psi^*(b) = -\int_b^{\infty} h(t)dF^*(t)$ entonces $\lim\limits_{c \to \infty} \psi^*(c) \cdot e^{-\int_{-\infty}^{c} \frac{d\psi^*(t)}{\psi^*(t)-h(t)}} - [1 - F^*(b)]\psi^*(b) = -\int_b^{\infty} h(t)dF^*(t).$

pero por vii)

$\lim\limits_{c \to \infty} \psi^*(c) e^{-\int_{-\infty}^{c} \frac{d\psi^*(t)}{\psi^*(t)-h(t)}} = 0.$

Asi $[1 - F^*(b)]\psi^*(b) = \int_b^{\infty} h(t)dF^*(t).$

Luego $\psi^*(b) = \dfrac{1}{1 - F^*(b)} \displaystyle\int_b^\infty h(t)dF^*(t)$.

De este modo $\psi^* = T(F^*)$ entonces $\psi^* \in Im(T)$.

En el caso $D = (-\infty, \beta)$ de la condición $iii)$ del Lema 3.5.1 obtenemos que $\lim\limits_{x \to \beta^-} \psi^*(x) = h(\beta)$ porque $h(x) < \psi^*(x)$ para todo $x \in (-\infty, \beta)$.

Pero $h(\beta) > \psi^*(x)$ para todo $x \in (-\infty, \beta)$.

Asi $h(x) < \psi^*(x) < h(\beta)$

y $h(\beta) = \lim\limits_{x \to \beta^-} h(x) \leqslant \lim\limits_{x \to \beta^-} \psi^*(x) \leqslant \lim\limits_{x \to \beta^-} h(\beta) = h(\beta)$.

Por tanto $\lim\limits_{x \to \beta^-} \psi^*(x) = h(\beta)$.

Luego $\lim\limits_{c \to \beta^-} [1 - F^*(c)]\psi^*(c) = 0 \quad h(\beta) = 0$.

Por lo que $\psi^*(b) = \dfrac{1}{1 - F^*(b)} = \displaystyle\int_b^\infty h(t)dF^*(t)$.

De este modo $\psi^* = \omega(F^*)$, entonces $\psi^* \in Im(T)$. $\square$

Conclusiones

Las variables aleatorias, en general, tienen alguna estructura de dependencia, esto es evidente por ejemplo en el caso de las estadísticas de orden de una muestra aleatoria continua, la relación de dependencia se puede expresar en términos de las esperanzas condicionales. Diversos autores, citados en este trabajo demostraron que sobre el conjunto de distribuciones se puede hacer una caracterización basada en esperanzas condicionales.

Más específicamente, se demuestra en esta tesis que la función de distribución común de un conjunto finito de variables aleatorias continuas e independientes queda determinada de manera única por medio de las esperanzas de funciones de estadísticas de orden condicionadas a otras estadísticas de orden.

Las condiciones que deben satisfacer las funciones de estas estadísticas de orden son mínimas y el condicionamiento se puede hacer sobre una o dos estadísticas de orden, consecutivas o no.

Si se conoce la forma de las esperanzas condicionales mencionadas antes, se conoce la distribución, esto da una caracterización de las distribuciones tal como lo hace, por ejemplo, la función característica.

También se exponen las condiciones que debe satisfacer una función para ser la esperanza condicional de alguna función dada.
En los dos primeros capítulos se exponen de manera detallada los principales conceptos y resultados sobre Esperanzas Condicionales, Probabilidades Condicionales y Martingalas con el objeto de que se cuente con un material más accesible que los textos tradicionales.

Bibliografía

[1] Ash, R.B. (1972), *Real Analysis and Probability*. Ed. Academic Press, New York.

[2] Balasubramanian, K. Dey, A. (1997), Distributions characterized through conditional expectations. *Metrika*, **45**, pp. 189-196.

[3] Billinsgley, P. (1986), *Probability and Measure*. Ed. John Wiley, New York.

[4] Bojdecki, T. (1985), ¿Qué es y para qué sirve una martingala? *Ciencia*, **36**, pp. 59-65.

[5] Chow, Y. Teicher, H. (1978), *Probability Theory*. Ed. Springer-Verlag, New York.

[6] Dudley, R.M. (1989), *Real Analysis and Probability*. Ed. Wadsworth and Brooks/Cole, Pacific Grove.

[7] Ibarrola, P. Pardo, L. Quesada,V. (1997), *Teoría de la Probabilidad*. Ed. Sintesis, Madrid.

[8] Laha, R.G. Rohatgi, V.K. (1979), *Probability Theory*. Ed. John Wiley, New York.

[9] Neuts, M.F. (1973), *Probability*. Ed. Allyn and Bacon, Boston.

[10] Ouyang, L.Y. (1995), Characterization through the conditional expectation of a function of one order statistics relative to an adjacent one. *Sankhya*, **57**, pp. 500-503.

[11] Rudin, W. (1980), *Principios de Análisis Matemático*. Ed. McGrawHill, México.

[12] Sagan, H. (1974), *Advanced Calculus*. Ed. Houghton Mifflin, Boston.

[13] Wu, J.W. Ouyang, L.Y. (1996), On characterizing distributions by conditional expectations of functions of order statistics. *Metrika*, **43**, pp. 135-147.

[14] Zoroa, P. Ruiz, J.M. Marin, J. (1990), A characterization based on conditional expectations. *Communication in Statisitcs, Theory and Methods*, **19**, pp. 3127-3135