UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

MARTINGALAS DISCRETAS

T E S I S

QUE PARA OBTENER EL TITULO DE:

M A T E M A T I C O

P R E S E N T A:

MARTHA RICO DIENER

AGOSTO 1990

Contenido

INTRODUCCION

Veamos antes que nada algunos resultados de teoría de la medida.

Definición: 1.1

Sea $\Omega \neq \emptyset$ fijo. Denotamos por $P(\Omega)$ al conjunto potencia de Ω, i.e. $P(\Omega) = \{A: A \subset \Omega\}$.

Definición. 1.2

Una σ-álgebra, $\mathcal{F} \subset P(\Omega)$ es una familia que satisface:

i) $\Omega \in \mathcal{F}$.

ii) $A, B \in \mathcal{F} \Rightarrow A - B \in \mathcal{F}$.

iii) Si $(A_n)_1^\infty \subset \mathcal{F} \Rightarrow \bigcup_{n=1}^\infty A_n \in \mathcal{F}$, donde $(A_n)_1^\infty$ es una sucesión de elementos de $\mathcal{F}$.

Observación. 1.3

i) $\emptyset \in \mathcal{F}$, pues $\emptyset = \Omega - \Omega \in \mathcal{F}$.

ii) $A^c \in \mathcal{F}$, pues $A^c = \Omega - A \in \mathcal{F}$.

iii) Si $(A_n)_1^\infty \subset \mathcal{F} \Rightarrow \bigcap_{n=1}^\infty A_n \in \mathcal{F}$, esto se deduce directamente de la definición.

Definición. 1.4

Sea $\varepsilon \subset P(\Omega)$. La σ-álgebra generada por ε, denotada por $\mathcal{F}(\varepsilon)$, es una σ-álgebra tal que:

i) $\varepsilon \in \mathcal{F}(\varepsilon)$.

ii) Si $\mathcal{F}'$ es una σ-álgebra ($\subset P(\Omega)$) que contenga a ε, entonces también contiene a $\mathcal{F}(\varepsilon)$.

La condición (ii) nos dice en realidad que de existir $\mathcal{F}(\varepsilon)$, $\mathcal{F}(\varepsilon)$ es única, y de hecho que es la mínima σ-álgebra que contiene a ε.

Teorema de Existencia. 1.5

Sea $\Omega \neq \emptyset$, $\varepsilon \subset P(\Omega)$ entonces existe una única σ-álgebra $\subset P(\Omega)$ denotada $\mathcal{F}(\varepsilon)$ tal que:

i) $\varepsilon \subset \mathcal{F}(\varepsilon)$.

ii) Si $\mathcal{F}' \subset P(\Omega)$ es una σ-álgebra y $\varepsilon \subset \mathcal{F}' \Rightarrow \mathcal{F}(\varepsilon) \subset \mathcal{F}'$.

La demostración de este teorema no se incluirá en este trabajo pero puede consultarse en [ ].

Un ejemplo importantísimo de una σ-álgebra generada por un conjunto es la σ-álgebra de Borel.

Sea $\Omega = \mathbb{R}$ y sea $\varepsilon = \{(a,b): a < b \quad (a, b \in \mathbb{R})\}$. Se define la σ-álgebra de Borel de R como $\mathcal{F}(\varepsilon)$, y se denota por $\mathcal{B}(\mathbb{R})$.

Algunos conjuntos contenidos en $\mathcal{B}(\mathbb{R})$ son:

i) $(a,b) \in \mathcal{B}(\mathbb{R}) \qquad \forall\, a < b$.

ii) Todo conjunto abierto de R pertenece a $\mathcal{B}(\mathbb{R})$.

iii) Si F es un conjunto cerrado entonces $F \in \mathcal{B}(\mathbb{R})$.

iv) $\{x\} = \bigcap_{n=1}^{\infty}(x - 1/n, x + 1/n) \in \mathcal{B}(\mathbb{R})$.

v) Todo conjunto finito o numerable pertenece a $\mathcal{B}(\mathbb{R})$.

Definición. 1.6

Un espacio medible es una pareja $(\Omega, \mathcal{F})$ con $\Omega \neq \emptyset$ y $\mathcal{F} \in \mathcal{P}(\Omega)$ una σ-álgebra. A los elementos de $\mathcal{F}$ se les llama conjuntos $\mathcal{F}$-medibles.

Así por ejemplo $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ es un espacio medible.

Definición. 1.7

Sea $(\Omega, \mathcal{F})$ un espacio medible. Una medida μ en $\mathcal{F}$ es una función $\mu: \mathcal{F} \to \mathbb{R}^+$ tal que:

i) $\mu(A) \geq 0 \qquad \forall A \in \mathcal{F}$.

ii) $\mu(\emptyset) = 0$.

iii) Si $(A_n)_1^{\infty}$, es una sucesión disjunta de elementos de $\mathcal{F}$ entonces $\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n)$. A esta última propiedad se le conoce con el nombre de σ-aditividad.

Proposicion. 1.8

Toda medida μ es aditiva, monótona y sustractiva.

i) Aditiva. $(\mu(A \cup B)) = \mu(A) + \mu(B)$.

ii) Monótona. $(A \subset B \Rightarrow \mu(A) \leq \mu(B)) \quad A, B \in \mathcal{F}$.

iii) Sustractiva. Si $A \subset B$ y $\mu(A) < +\infty \Rightarrow \mu(B - A) = \mu(B) - \mu(A)$.

Demostración:

i) $A \cup B \cup \emptyset \cup \emptyset \cup \emptyset \cup \ldots = A \cup B \Rightarrow$ por la σ-aditividad, $\mu(A \cup B) = \mu(A) + \mu(B) + 0 + 0 + \ldots = \mu(A) + \mu(B)$.

ii) $B = A \cup (B - A)$ entonces $\mu(B) = \mu(A) + \mu(B - A) \geq \mu(A)$.

iii) Como en (ii) $\mu(B) = \mu(A) + \mu(B - A) \Rightarrow \mu(B) - \mu(A) = \mu(B - A)$.

Podemos ahora definir el concepto de probabilidad.

Definición 1.9
Sea $(\Omega, \mathcal{F})$ un espacio medible. Una probabilidad en este espacio es una medida P definida en $\mathcal{F}$, y tal que $P(\Omega) = 1$.

Nota: A los elementos de $\mathcal{F}$ se les llama eventos.

Algunas de las propiedades de esta medida P son las siguientes:
i) $P(\emptyset) = 0$.
ii) Si $(A_n)_{n=1}^{\infty}$ es una sucesión disjunta de eventos de Ω entonces

$$P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n).$$

iii) $P(A^c) = 1 - P(A)$.
iv) $P(A) \leq 1 \qquad \forall A \subset \Omega$.
v) Si $A \subset B \Rightarrow P(A) \leq P(B)$.
vi) $P(\cup_n A_n) \leq \sum P(A_n)$, con (A_n) no necesariamente ajenos.

Las propiedades (i) y (ii) se deducen directamente de la definición. La propiedad (iii) se deduce así: $1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c)$ por lo tanto $P(A^c) = 1 - P(A)$. Las propiedades (iv) y (v) son consecuencia directa de que toda medida es monótona, i.e. $A \subset B \Rightarrow P(A) \leq P(B)$ y por lo tanto si $A \subset \Omega \Rightarrow P(A) \leq P(\Omega) = 1$. La propiedad (vi) se debe a la σ-sub-aditividad de una medida.

Llegamos entonces a la siguiente definición.

Definición: 1.10
Un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ está definido por la especificación de un conjunto Ω, una σ-álgebra $\mathcal{F}$ de subconjuntos de Ω y una probabilidad P definida en $\mathcal{F}$.

Veamos ahora lo que es una variable aleatoria, para ello veamos una definición preliminar.

Definición: 1.11

Sea $(\Omega, \mathcal{F})$ un espacio medible fijo.
Una función $f: \Omega \to \mathbb{R}$ se dice que es $\mathcal{F}$-medible, si $\forall \quad \alpha \in \mathbb{R}$,
$A\alpha = \{x \in \Omega: f(x) > \alpha\} \in \mathcal{F}$. i.e. f es $\mathcal{F}$-medible si $\forall \alpha \in \mathbb{R}$, $f^{-1}\{(\alpha, \infty)\} \in \mathcal{F}$.

Definición: 1.12

Una variable aleatoria X en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ es una función $\mathcal{F}$-medible de Ω a $\mathbb{R}$; i.e. $X: (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.
Una variable aleatoria extendida es una función $X: (\Omega, \mathcal{F}) \to (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ tal que $X^{-1}(\alpha, \infty) \cup \{\infty\} \in \mathcal{F}$ para toda $\alpha \in \mathbb{R}$.

Es lo mismo que considerar a X una función real en Ω que tenga rango finito y que cumpla $[\omega: X(\omega) > x] \in \mathcal{F} \quad \forall x \in \mathbb{R}$. Podemos notar entonces que el que X sea una variable aleatoria depende unicamente de $\mathcal{F}$ y no de P.

Si X es una variable aleatoria en $(\Omega, \mathcal{F}, P)$, la medida de probabilidad inducida por X es P_X en $\mathcal{B}(\mathbb{R})$, dada por

$$P_X(B) = P\{\omega: X(\omega) \in B\}, \quad B \in \mathcal{B}(\mathbb{R}).$$

Los números $P_X(B)$ y $B \in \mathcal{B}(\mathbb{R})$ caracterizan por completo a la variable aleatoria X en el sentido de que estos proveen de probabilidades a todos los eventos que incluyen a X.

Veamos porque afirmamos que $P_X(B)$ es una medida de probabilidad, para ello basta ver que P_X es una medida y que $P_X(\mathbb{R}) = 1$.

i) $P_X(B) = P\{\omega: X(\omega) \in B\} \geq 0 \qquad \forall B \in \mathcal{B}(\mathbb{R})$ por ser P un probabilidad.
ii) $P_X(\emptyset) = P\{\omega: X(\omega) \in \emptyset\} = P(\emptyset) = 0$.
iii) Si $(A_n)_1^\infty$ es una sucesión disjunta de elementos de $\mathcal{B}(\mathbb{R})$ entonces
$P_X(\bigcup_{n=1}^\infty A_n) = P\{(\omega: X(\omega) \in \bigcup_{n=1}^\infty A_n)\} = \sum_{n=1}^\infty P\{\omega: X(\omega) \in A_n\} = \sum_{n=1}^\infty P_X(A_n)$.
iv) $P_X(\mathbb{R}) = P(\omega: X(\omega) \in \mathbb{R}) = P(\Omega) = 1$.
por lo tanto P_X es una probabilidad.

Si X es cualquier variable aleatoria discreta, las propiedades de X se pueden determinar por completo por la función de probabilidad P_X, definida por

$$P_X(x) = P\{X = x\}, \quad x \in \mathbb{R}.$$

Explícitamente, $P_X(B) = \sum_{x \in B} P_X(x)$; y se trata de una suma numerable, pues P_X es 0 salvo en las x_n, donde por x_n se entienden todos los valores de X.

Por tanto podemos afirmar que una variable aleatoria discreta se puede especificar dando un conjunto numerable $\{x_n, n = 1, 2, \ldots\}$ de números reales y un conjunto de probabilidades $\{P_n, n = 1, 2, \ldots\}$ $(P_n \geq 0,\ \sum_n P_n = 1)$, donde P_n funciona como $P\{X = x_n\}$. La probabilidad de que X pertenezca a B se encuentra sumando los P_n para las cuales X_n pertenece a B.

Algunas definiciones y comentarios 1.13

Sea X una variable aleatoria en $(\Omega, \mathcal{F}, P)$. Decimos que X es simple si y sólo si X toma únicamente un número finito de valores, discreta si y sólo si el conjunto de valores de X es finito o infinito numerable. Cualquier variable aleatoria en un espacio de probabilidad discreto es discreta, pues Ω es numerable.

Estamos ahora en condiciones de definir lo que es una esperanza.

Definición 1.14

Si X es una variable aleatoria en $(\Omega, \mathcal{F}, P)$, la esperanza de X se define por $E(X) = \int X\,dP = \int_\Omega X(\omega)P(d\omega)$ siempre que ésta integral exista.

Como $E(X)$ es la integral de una función X $\mathcal{F}$-medible con respecto a la medida de probabilidad P, tenemos entonces que se cumplen todos los resultados de la teoría de integración. Ahora bien, sabemos que $\mathcal{L}_1(\Omega, \mathcal{F}, P)$ es la clase de funciones integrables que toman valores en R y que $\int_E f\,dP < +\infty \quad \forall E \in \mathcal{F}$.

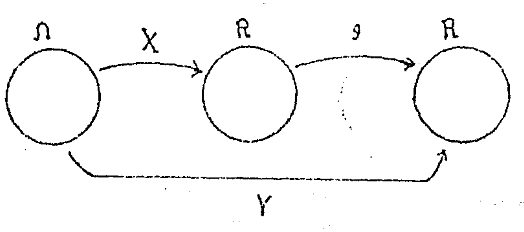
Podemos expresar la esperanza de otra manera:

Teorema: 1.15

Sea X una variable aleatoria en $(\Omega, \mathcal{F}, P)$. Sea $g\colon \mathrm{R} \to \mathrm{R}$ una función Borel-medible. Si $Y = g \circ X$, entonces

$$E(Y) = \int_{\mathrm{R}} g\,dP_X.$$

en el sentido de que si algún lado existe, entonces existe el otro, y ambos son iguales.



Demostración:

La demostración se hara por pasos.

Sea g un indicador I_B, $B \in \mathcal{B}(\mathbb{R})$. Entonces

$$E(Y) = E(I_B \circ X) = E(I_{\{X \in B\}}) = P_X(B) = \int_{\mathbb{R}} g \, dP_X.$$

probando entonces que $E(Y)$ y $\int_{\mathbb{R}} g \, dP_X$ existen y son iguales.

Consideremos ahora a g como una función simple no-negativa.

Sea $g(x) = \sum_{j=1}^{n} X_j I_{B_j}(X)$, los B_j conjuntos disjuntos en $\mathcal{B}(\mathbb{R})$. Entonces

$$\begin{aligned}
E(Y) &= \sum_{j=1}^{n} X_j E(I_{B_j} \circ X) \\
&= \sum_{j=1}^{n} X_j \int_{\mathbb{R}} I_{B_j} \, dP_X \quad \text{por lo que acabamos de probar} \\
&= \int_{\mathbb{R}} \Big(\sum_{j=1}^{n} X_j I_{B_j}\Big) \, dP_X \quad \text{pues} \quad g \geq 0 \\
&= \int_{\mathbb{R}} g \, dP_X
\end{aligned}$$

Nuevamente ambas integrables existen y son iguales.

Ahora consideremos a g como una función Borel-medible y no-negativa. Sean $g_1, g_2, \ldots$ funciones simples no-negativas tales que $g_n \uparrow g$.

Acabamos de probar que $E(g_n \circ X) = \int_{\mathbb{R}} g_n \, dP_X$ y por el teorema de la convergencia monótona $E(g \circ X) = \int_{\mathbb{R}} g \, dP_X$ y otra vez, ambas integrales existen y son iguales.

Finalmente, si $g = g^+ - g^-$ es una función Borel-medible cualquiera y $Y = g \circ X$, tenemos

$$\begin{aligned}
E(Y) &= E(Y^+) - E(Y^-) \\
&= E(g^+ \circ X) - E(g^- \circ X) \\
&= \int_{\mathbb{R}} g^+ \, dP_X - \int_{\mathbb{R}} g^- \, dP_X \\
&= \int_{\mathbb{R}} g \, dP_X
\end{aligned}$$

Si $E(Y)$ existe y $E(Y^-)$ es finito, entonces $\int_{\mathbb{R}} g^- \, dP_X$ es finita y por lo tanto $\int_{\mathbb{R}} g \, dP_X$ existe. Por un razonamiento similar, si $\int_{\mathbb{R}} g \, dP_X$ existe implica la existencia de $E(Y)$. $\square$

Definición 1.16
Sea X una variable aleatoria en $(\Omega, \mathcal{F}, P)$. Si $k > 0$, el número $E(X^k)$ es llamado el k-ésimo momento de X; $E(|X|^k)$ es llamado el k-ésimo momento absoluto de X. $E\big[(X - E(X))^k\big]$ es llamado el k-ésimo momento central, $E\big[|X - E(X)|^k\big]$ el k-ésimo momento central absoluto; los momentos centrales están definidos únicamente si $E(X)$ es finita.
Se definen la media := si $k = 1$ i.e. $E(X)$; la varianza $:= \sigma^2 = E\big[(X - E(X))^2\big]$.

Podemos ahora hacer una comparación, en esta ocasión con los espacios $\mathcal{L}_p$, a saber:

Definición 1.17
Sea $p \in [1, +\infty)$ fijo y sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad.
Definimos $\mathcal{L}_p(\Omega, \mathcal{F}, P) = \{X\ \mathcal{F}\text{−medible}: \int |X|^p\, dP < \infty\}$.
Si $p = 1 \Rightarrow \mathcal{L}_1$ coincide con las funciones p-integrables.

Lema 1.18
Si $k > 0$ y $E(X^k)$ es finita, entonces $E(X^j)$ es finita para $0 < j < k$.

Demostración:
Tenemos que $\|X\|_j \leq \|X\|_k$ para $0 < j < k$. □

El siguiente resultado es casi inmediato

Lema 1.19
Si n es un entero positivo mayor que 1, $E(X^{n-1})$ es finita y $E(X^n)$ existe, entonces

$$E\big[(X - E(X))^n\big] = \sum_{k=0}^{n} \binom{n}{k} \big[-E(X)\big]^{n-k} E(X^k).$$

En particular, si $E(X)$ es finita, ($E(X^2)$ siempre existe si $X^2 \geq 0$), entonces
$\mathrm{var}X = E(X^2) - \big[E(X)\big]^2$.

Desigualdad de Chebyshev 1.20
Sea f una función en $(\Omega, \mathcal{F}, \mu)$, no-negativa y Borel-medible. Si $0 < p < \infty$ y $0 < \varepsilon < \infty$,

$$\mu\{\omega: f(\omega) \geq \varepsilon\} \leq \frac{1}{\varepsilon^p} \int_\Omega f^p\, d\mu.$$

Pero la siguiente versión se usa mucho en probabilidad.
Si g es una función en $(\Omega, \mathcal{F})$ Borel-medible y P es una medida de probabilidad en $\mathcal{F}$, se definen:

$$m = \int_\Omega g\, dP \quad \text{se supone finita la integral.}$$

$$\sigma^2 = \int_\Omega (g-m)^2\, dP.$$

Si $0 < k < \infty$,

$$P\{\omega : |g(\omega) - m| \geq k\sigma\} \leq \frac{1}{k^2}.$$

Esto se deduce de la primera versión, tomando $f = |g - m|$, $\varepsilon = k\sigma$, $p = 2$.

Demostración:

$$\int_\Omega f^p\, d\mu \geq \int_{\{\omega : f(\omega) \geq \varepsilon\}} f^p\, d\mu \geq \varepsilon^p \mu\{\omega : f(\omega) \geq \varepsilon\}.$$

□

Estudiemos ahora un nuevo concepto.

Definición (Variables aleatorias independientes) 1.21
Sean $X_1, \ldots, X_n$ variables aleatorias en $(\Omega, \mathcal{F}, P)$. Se dice que $X_1, \ldots, X_n$ son independientes si y sólo si para todos los conjuntos $B_1, \ldots, B_n \in \mathcal{B}(\mathbb{R})$, tenemos:

$$P\{X_1 \in B_1, \ldots, X_n \in B_n\} = P\{X_1 \in B_1\} \ldots P\{X_n \in B_n\}.$$

Lo anterior significa que una afirmación sobre X_i corresponde a un evento de la forma $A_i = \{X_i \in B_i\}$, por lo tanto los eventos $A_1, \ldots, A_n$ van a ser independientes. En otras palabras, si $X = (X_1, \ldots, X_n)$ entonces P_X es el producto de P_{X_i}, $i = 1, \ldots, n$.

Notación: v.a.i abrevia variables aleatorias independientes.

Observaciones: 1.22
a) Si $X_1, \ldots, X_n$ son independientes, entonces $X_1, \ldots, X_k$ $\forall k < n$ también lo son.

Demostración:

$$P\{X_1 \in B_1, \ldots, X_k \in B_k\} = P\{X_1 \in B_1, \ldots, X_k \in B_k, X_{k+1} \in \mathbb{R}, \ldots, X_n \in \mathbb{R}\}$$
$$= P\{X_1 \in B_1\} \ldots P\{X_k \in B_k\}$$

b) Sea $\{X_i, i \in I\}$ (un conjunto arbitrario de índices) una familia arbitraria de variables aleatorias. Se dice que $\{X_i, i \in I\}$ son independientes si y sólo si $X_{i_1}, \ldots, X_{i_n}$ son independientes para cada conjunto finito de subindices distintos $\{i_1, \ldots, i_n\}$ en I.

Teorema: 1.23
Sean $X_1, \ldots, X_n$ v.a.i en $(\Omega, \mathcal{F}, P)$. Si todas las X_i son no- negativas o si $E(X_i) < +\infty \quad \forall i$, entonces $E(X_1, \ldots, X_n) = E(X_1)E(X_2) \ldots E(X_n)$.

Demostración:
Si $X_i \geq 0 \quad \forall i \Rightarrow E(X_1, \ldots, X_n) = \int_{\mathbb{R}} X_1 \ldots X_n \, dP_X(X_1, \ldots, X_n)$ donde $X = (X_1, \ldots, X_n)$
Como se mencionó al dar la definición de v.a.i., P_X es el producto de $P_{X_i}, i = 1, 2, \ldots,$ y entonces el teorema de Fubini implica que:

$$E(X_1 \ldots, X_n) = \int_{\mathbb{R}} X_1 \, dP_{X_1} \ldots \int_{\mathbb{R}} X_n \, dP_{X_n} = E(X_1)E(X_2) \ldots E(X_n)$$

Si $E(X_i) < +\infty \quad \forall i \Rightarrow E(|X_1, \ldots, X_n|) = \Pi_{i=1}^{n} E(|X_i|) < +\infty$ y podemos aplicar nuevamente el teorema de Fubini. □

Definición 1.24
La función de distribución de una v.a. X es la función $F = F_X : \mathbb{R} \to [0, 1]$, dada por

$$F(x) = P\{\omega : X(\omega) \leq x\}, x \in \mathbb{R}.$$

Observación:
Como para $a < b$, $F(b) - F(a) = P\{\omega : a < X(\omega) \leq b\} = P_x(a, b]$, F es una función de distribución que corresponde a la medida P_x de Lebesgue- -Stieltjes.
En particular, F es monótona creciente y contínua por la derecha.

Teorema 1.25 Ley débil de los grandes números.
Sean $X_1, X_2, \ldots$ v.a.i., (no necesariamente con la misma distribución), cada una con media y varianza finitas. Supongamos que las varianzas están uniformemente acotadas por $M < +\infty$. Sea $S_n = X_1 + \ldots X_n$. Entonces $[S_n - E(S_n)]/n$ converge en probabilidad a 0, esto es, dada $\varepsilon > 0$,

$$P\{|\frac{S_n - E(S_n)}{n}| \geq \varepsilon\} \to 0 \qquad \text{cuando} \qquad n \to \infty.$$

Demostración:
Antes de hacer la prueba probaremos un resultado que resulta útil a saber:
$var(X_1 + \ldots + X_n) = \sum_{i=1}^{n} varX_i$

$$\begin{aligned} var(X_1 + \ldots + X_n) &= E[(X_1 + \ldots + X_n - E(X_1) - \ldots - E(X_n))^2] \\ &= \sum_{i=1}^{n} E[(X_i - E(X_i))^2] + 2 \sum_{\substack{i,j=1 \\ i<j}}^{n} E[(X_i - E(X_i))(X_j - E(X_j))] \\ &= \sum_{i=1}^{n} E[(X_i - E(X_i))^2] + 2 \sum_{\substack{i,j=1 \\ i<j}} E(X_i X_j) - E(X_i)E(X_j) \\ &= \sum_{i=1}^{n} varX_i \end{aligned}$$

Ahora sí, la demostración del teorema:
Por la desigualdad de Chebyshev:

$$\begin{aligned} P\{|\frac{S_n - E(S_n)}{n}| \geq \varepsilon\} &\leq \frac{1}{\varepsilon^2} E[(\frac{S_n - E(S_n)}{n})^2] \\ &= \frac{1}{\varepsilon^2 n^2} varS_n \\ &= \frac{1}{\varepsilon^2 n^2} \sum_{k=1}^{n} varX_k \qquad \text{por lo que acabamos de probar.} \end{aligned}$$

Como las varianzas están uniformemente acotadas tenemos que

$$varX_k \leq M \qquad \forall k \Rightarrow \sum_{k=1}^{n} varX_k \leq nM$$

$$\text{por lo tanto } \frac{1}{\varepsilon^2 n^2} \sum_{k=1}^{n} varX_k \leq \frac{M}{\varepsilon^2 n} \to 0$$

□

Probabilidad y Esperanza Condicional

La noción de probabilidad condicional de un evento A "dado un evento B" corresponde a la frecuencia de A en pruebas sucesivas donde B ocurre. Para cada evento A, la relación

$$P(B)P(A|B) = P(A \cap B)$$

es decir,

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

siempre y cuando $P(B) \neq 0$, define la probabilidad condicional de A "dado un evento B". Denotaremos a $P(A|B)$ como $P_B(A)$.

En el caso en que $P(B) = 0$, tendremos que $P(A \cap B) = 0$, por lo que $P(A|B)$ queda indeterminada. Por el momento vamos a analizar únicamente el caso en que $P(B) \neq 0$.

La función P_B en la σ-álgebra de eventos $\mathcal{F}$, cuyos valores son $P_B(A)$, $A \in \mathcal{F}$, se llama probabilidad condicional dado B.

La idea que subyace a la definición es la siguiente:

Si consideramos a la probabilidad condicional en términos de un observador en posesión de información parcial, un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ describe el trabajo de un proceso, gobernado por el azar, que produce un resultado ω distribuido de acuerdo a P. $P(A)$ es para el observador la probabilidad de que el punto ω producido esté en A.

Supongamos ahora que la única información de qúe dispone el observador consiste en saber que ω está en B.

Desde el punto de vista del observador, ahora en posesión de esta información parcial sobre ω, la probabilidad de que ω también esté en A es $P(A|B)$ en vez de $P(A)$.

Ahora bien, de la definición se deriva que si P es una probabilidad sobre $\mathcal{F}$, también lo es P_B pues:

i) $P_B(\Omega) = P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$

ii) $P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)} \geq 0 \quad \forall A \in \mathcal{F}$ por ser P una probabilidad.

iii) $P_B(\cup_{i=1}^{n} A_i) = P(\cup_{j=1}^{\infty} A_j|B) = \sum_{j=1}^{\infty} P(A_j|B)$ pues P es σ-aditiva.

Así, la expresión condicionante "dado B" significa que el espacio inicial $(\Omega, \mathcal{F}, P)$ se puede sustituir por $(\Omega, \mathcal{F}, P_B)$.

Tenemos entonces que si X es una v.a., en este nuevo espacio de probabilidad, la esperanza, en caso de existir, se denomina esperanza condicional dada B y se denota por

$$E_B(X) = E(X|B) = \int X\, dP_B.$$

Como $P_B = 0$ en $(A \cap B^c)$ $\forall A \in \mathcal{F}$, entonces

$$E(X|B) = \int_B X\, dP_B$$

y como $P_B = \frac{1}{P(B)} P$ sobre $(A \cap B)$ $\forall A \in \mathcal{F}$, entonces

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

Podemos entonces definir la esperanza condicional de X dado B como:

$$P(B)E(X|B) = \int_B X\, dP$$

siempre y cuando $P(B) \neq 0$, i.e., que B no sea nulo.
Si consideramos ahora el caso en que X sea un indicador I_A, tendremos

$$P(B)E(I_A|B) = \int_B I_A\, dP = P(A \cap B)$$

de tal forma que la probabilidad condicional se puede definir por

$$P(A|B) = E(I_A|B).$$

Esto quiere decir que si $E(\cdot|B) = E_B$ es la esperanza condicional dada B, con valores $E(X|B)$ en la familia $\mathcal{G}_B$, donde $\mathcal{G}_B$ consiste en la familia de todas las v.a. X cuya integral sobre B existe, la probabilidad condicional P_B viene a ser una restricción de E_B sobre la familia $I_{\mathcal{F}}$ de indicadores de eventos.

Analicemos ahora el concepto de esperanza condicional, pero concebido como valores de una función, i.e., vamos a considerar a la esperanza condicional como una función con dos valores reales, de la manera que sigue:
El número $E(X|B)$ ya no se va a asignar, como antes, a B, sino que se va a asignar a todo $\omega \in B$, e igual procederemos con $E(X|B^c)$. De esta forma hemos obtenido una función definida de Ω a $\mathbb{R}$, cuyos dos posibles valores son:

$$f_X(\omega) = \begin{cases} E(X|B), & \text{si } \omega \in B; \\ E(X|B^c), & \text{si } \omega \in B^c. \end{cases}$$

o lo que es lo mismo

$$(E(X|B)I_B + E(X|B^c)I_B^c)(\omega) \qquad ...(1)$$

Para el caso más general, consideremos a $\{B_j\}$ como una partición numerable de Ω y sea $\mathcal{B} = \sigma(\{B_j\})$. Sea $\mathcal{G}$ la familia de todas las v.a. X cuya esperanza $E(X)$ existe, de tal manera que su integral indefinida existe y, por lo tanto, su esperanza condicional también existe.

Al igual que en (1), obtenemos la siguiente función:

$$E(X|\mathcal{B}) = \sum E(X|B_j) I_{B_j}, \quad X \in \mathcal{G}$$

que no es otra cosa que una función discreta.

Nuevamente, tendríamos problemas si algún B_j es nulo, pues su valor correspondiente $E(X|B_j)$ estaría indeterminado; sin embargo, vamos a dejar el caso de algún B_j nulo, junto con su definición $E(X|B_j)$, para después.

De esta manera hemos llegado a una definición:

La variable aleatoria discreta $E(X|\mathcal{B})$ definida salvo equivalencias por

$$E(X|\mathcal{B}) = \sum \left(\frac{1}{P(B_j)} \int_{B_j} X\, dP\right) I_{B_j}, \quad \forall X \in \mathcal{G}$$

es la esperanza condicional de X dada $\mathcal{B}$.

Particularizando nuevamente sobre los indicadores, la función $\mathcal{B}$-medible $P(A|\mathcal{B})$, definida salvo equivalencias por

$$P(A|\mathcal{B}) = E(I_A|\mathcal{B}), A \in \mathcal{F}$$

va a ser la probabilidad condicional de A dado $\mathcal{B}$: la contracción de $E(\bullet|\mathcal{B}) = E_{\mathcal{B}}$ sobre $I_{\mathcal{F}}$, denotada por $P(\bullet|\mathcal{B}) = P_{\mathcal{B}}$, va a ser la probabilidad condicional dada $\mathcal{B}$, y sus valores van a ser las funciones $\mathcal{B}$-medibles $P(A|\mathcal{B})$ con $A \in \mathcal{F}$, definidas salvo equivalencias.

Hablamos ya de la σ-álgebra $\mathcal{B}$ y no de la partición $\{B_j\}$ porque $E(X|\mathcal{B})$ determina la esperanza condicional de X dado un evento cualquiera no nulo $B \in \mathcal{B}$.

De hecho si Σ' denota la suma sobre alguna subclase de $\{B_j\}$, entonces cualquier evento $B \in \mathcal{B}$ es de la forma $\Sigma' B_j$ y tendremos

$$P(B)E(X|B) = \int_{\Sigma' B_j} X\, dP = \Sigma' \int_{B_j} X\, dP = \Sigma' P(B_j) E(X|B_j)$$

pero si consideramos a $P_{\mathcal{B}}$ como la restricción de P a $\mathcal{B}$ definida por

$$P_{\mathcal{B}}(B) = P(B), B \in \mathcal{B}$$

tendermos que $\int_B X\, dP = \int_B E(X|\mathcal{B})\, dP_{\mathcal{B}}$.

Lo anterior nos lleva a la siguiente definición:

La esperanza condicional $E(X|\mathcal{B})$ de $X \in \mathcal{G}$ dada $\mathcal{B}$ es cualquier función $\mathcal{B}$-medible cuya integral indefinida con respecto a $P_{\mathcal{B}}$ es la restricción a $\mathcal{B}$ de la integral indefinida de X con respecto a P. Como la integral indefinida con respecto a $P_{\mathcal{B}}$ de cualquier función $\mathcal{B}$-medible determina a esta función salvo equivalencias, esta definición nos dice precisamente que $\forall X \in \mathcal{G}$, $E(X|\mathcal{B})$ esta definida por

$$\int_B E(X|\mathcal{B})\, dP_{\mathcal{B}} = \int_B X\, dP, \qquad B \in \mathcal{B}.$$

Con esto la definición de probabilidad condicional se convierte en

$$\int_B P(A|B)\, dP_{\mathcal{B}} = P(A \cap B), \quad A \in \mathcal{F}, \quad B \in \mathcal{B}$$

o equivalentemente en

$$P(A|B) = \sum \frac{P(A \cap B_j)}{P(B_j)} I_{B_j}, \quad \forall A \in \mathcal{F},$$

salvo equivalencias.

Veamos ahora que sucede con los eventos nulos.
Sea B_0 la suma de todos los B_j nulos, y para cada A, seleccionemos $P(A|\mathcal{B})$ con su clase de equivalencia con $\omega \in B_j$, tomando los valores:

$$P^{\mathcal{B}}_{\omega}(A) = \begin{cases} \frac{P(A \cap B_j)}{P(B_j)} & \text{si } B_j \text{ es no nulo} \\ P(A) & \text{si } B_j \text{ es nulo } (\subset B_0). \end{cases}$$

Entonces, para cada $\omega \in \Omega$, la función $P^{\mathcal{B}}_{\omega}$ en $\mathcal{F}$, con valores $P^{\mathcal{B}}_{\omega}(A)$, es una probabilidad y podemos considerar integrales con respecto a ella .
Sea $X \in \mathcal{G}$ y escribamos

$$E^{\mathcal{B}}_{\omega}(X) = \int X\, dP^{\mathcal{B}}_{\omega}, \qquad \omega \in \Omega.$$

Como $\forall \quad \omega \in B_j$ no contenido en el evento B_0, tenemos

$$\int X\, dP^{\mathcal{B}}_{\omega} = \frac{1}{P(B_j)} \int_{B_j} X\, dP,$$

se sigue que la función en Ω con valores $E^{\mathcal{B}}_{\omega}(X)$ pertenece a la clase de equivalencia de $E^{\mathcal{B}}(X)$. Así, podemos definir a $E^{\mathcal{B}}(X)$ como $P_{\mathcal{B}}$-equivalente a la integral $\int X\, dP^{\mathcal{B}}$, donde $P^{\mathcal{B}}$, y de aqui la integral, son funciones de $\omega \in \Omega$; en símbolos

$$E^{\mathcal{B}}(X) = \int X\, dP^{\mathcal{B}} \quad \text{c.s.}$$

Estamos ahora en condiciones de analizar el caso general:
Hasta ahora lo que hemos hecho tiene el defecto de que hemos considerado σ-álgebras generadas por particiones numerables y este no es el caso de cualquier σ-álgebra dada. Sin embargo todo lo dicho anteriormente es posible gracias al teorema de Radon-Nikodym, que se enuncia a continuación:

Teorema: 1.26
Sean $(\Omega, \mathcal{F})$ un espacio medible, $\mu: \mathcal{F} \to \overline{\mathbb{R}}$ una medida σ- finita y $\nu: \mathcal{F} \to \overline{\mathbb{R}}$ una medida con signo σ-finita tales que $\nu \ll \mu$. Entonces existe f medible en $(\Omega, \mathcal{F})$ tal que $\nu(E) = \int_E f\, d\mu \quad \forall E \in \mathcal{F}$. Si existe f' medible en $(\Omega, \mathcal{F})$ tal que $\nu(E) = \int_E f'\, d\mu \quad \forall E \in \mathcal{F}$ entonces $f = f'$ (c.d. relativo a μ).
La prueba de este teorema no se incluirá en este trabajo, pero puede consultarse en [ ].

Definición: 1.27
Supongamos que X es una v.a. integrable en $(\Omega, \mathcal{F}, P)$ y que $\mathcal{G}$ es una σ-álgebra contenida en $\mathcal{F}$. Entonces existe una v.a. $E(X|\mathcal{G})$, llamada el valor de la esperanza condicional de X dada $\mathcal{G}$, que cumple con estas dos propiedades:

i) $E(X|\mathcal{G})$ es $\mathcal{G}$-medible e integrable.
ii) $E(X|\mathcal{G})$ satisface la ecuación funcional

$$\int_G E(X|\mathcal{G})\, dP = \int_G X\, dP, \qquad G \in \mathcal{G}$$

Nota:
$\mathcal{G}$ es cualquier σ-álgebra contenida en $\mathcal{F}$, mientras que $\mathcal{B}$ era una σ-álgebra generada por una partición numerable. Es por ello, que esta definición es más general.

Para probar la existencia de tal variable aleatoria, consideremos primero el caso de X no-negativa. Definamos una medida ν en $\mathcal{G}$ por

$$\nu(G) = \int_G X\, dP.$$

Esta medida es finita porque X es integrable, y es absolutamente continua con respecto a P. Por el teorema de Radon-Nikodym existe una función f, $\mathcal{G}$-medible, tal que $\nu(G) = \int_G f\, dP$.

Observación: 1.28
La integral es la misma en $(\Omega, \mathcal{F}, P)$ que en $(\Omega, \mathcal{G})$ con P restringida a $\mathcal{G}$.
Esta f cumple con las propiedades (i) y (ii).
Si X no es necesariamente no-negativa, podemos considerar $E(X^+|\mathcal{G}) - E(X^-|\mathcal{G})$ que claramente tiene las propiedades requeridas.

Observación: 1.29

En general, vamos a tener varias v.a. $E(X|\mathcal{G})$, a cada una de ellas la llamaremos una **versión** del valor de la esperanza condicional. Esto se da por el propio teorema de Radon-Nikodym donde puede haber mas de una función que lo satisfaga, pero siempre dos cualesquiera difieren por un conjunto de medida cero. Por lo tanto cualesquiera dos versiones son iguales con probabilidad 1, con P restringida a $\mathcal{G}$.

Dos resutados inmediatos:

1) $E(X|\{\emptyset,\Omega\}) = E(X)$ con probabilidad 1
2) $E(X|\mathcal{F}) = X$ con probabilidad 1

Demostración:

1) $\int_G X\,dP = \int_G E(X)\,dP$ si $G = \emptyset$ o Ω.
2) $\int_G X\,dP = \int_G X\,dP$, $\quad G \in \mathcal{F}$ y X es $\mathcal{F}$-medible. □

El valor $E(X|\mathcal{G})_\omega$ en ω se puede interpretar como el valor esperado de X para alguien que sabe para cada G en $\mathcal{G}$ si esta contiene o no al punto ω, mismo que en general permanece desconocido.

La primera condición (i) asegura que $E(X|\mathcal{G})$, puede ser en principio calculada a partir de esta información parcial.

La condición (ii) puede expresarse como sigue

$$\int_G [X - E(X|\mathcal{G})]\,dP = 0$$

Si al observador con la información parcial contenida en $\mathcal{G}$ se le ofrece la oportunidad de apostar, pagando una entrada de $E(X|\mathcal{G})$ y se le regresa la cantidad X y si adopta la estrategia de apostar dado que G ocurre, esta ecuación implica que el juego es justo.

Esto es por la interpretación de la esperanza.

$X - E(X|\mathcal{G})$ = ganancia si ocurre G.

$1_G[X - E(X|\mathcal{G})]$ = ganancia

y $\int_\Omega 1_G[X - E(X|\mathcal{G})]\,dP = E(\text{ganancia})$

En general se dice que un juego es justo si X es la ganancia y cumple con $E(X) = 0$. Esto es, si se juega un número muy grande de veces un juego, la ganancia promedio es igual a cero.

Vamos a retomar esta discusión cuando hablemos de martingalas.

También hay un concepto general de probabilidad condicional dada una σ-álgebra.

Teorema: 1.30
Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, $\mathcal{G}$ una sub-σ-álgebra de $\mathcal{F}$; fijemos $B \in \mathcal{F}$. Existe entonces una función $P(B|\mathcal{G}): (\Omega, \mathcal{G}) \to (\mathbb{R}, \mathcal{B})$, llamada la probabilidad condicional de B dada $\mathcal{G}$, tal que

$$P(G \cap B) = \int_G P(B|\mathcal{G})\, dP \quad \forall G \in \mathcal{G}.$$

Cualesquiera dos funciones de estas deben coincidir casi dondequiera con respecto a P, y en realidad $P(B|\mathcal{G}) = E(I_B|\mathcal{G})$ c.d. rel P.

Demostración:

Sea $\lambda(G) = P(G \cap B)$, $G \in \mathcal{G}$. Entonces λ es una función en $\mathcal{G}$ aditiva, finito valuada y numerable (contable), absolutamente continua con respecto a P; la existencia y la unicidad c.d. rel P de $P(B|\mathcal{G})$ se sigue directamente del teorema de Radon-Nikodym.
La conección entre esperanza condicional y probabilidad condicional, se sigue directamente de la definición anterior poniendo I_B en vez de Y.

Propiedades de la Esperanza Condicional

Definición: 1.31
Sea $X: (\Omega, \mathcal{F}) \to (\Omega', \mathcal{F})$ un objeto aleatorio. La σ-álgebra inducida por X esta dada por $\mathcal{F}(X) = X^{-1}(\mathcal{F}')$.

Observación: 1.32
Los teoremas concernientes a Probabilidad Condicional se obtienen sustituyendo Y por I_B, y los resultados que se refieren a $E(Y|X)$ se obtienen haciendo $\mathcal{G} = \mathcal{F}(X)$, donde $\mathcal{F}(X)$ es la σ-álgebra inducida por X.
Para obtener los teoremas para el caso discreto podemos reemplazar a $\mathcal{G}$ por $(X = x)$, en general la redacción es casi la misma.

Teorema: 1.33
Supongamos que X, Y, X_n son integrables.

i) Si $X = a$ con probabilidad 1, entonces $E(X|\mathcal{G}) = a$.
ii) Si a, b son constantes, entonces $E(aX + bY|\mathcal{G}) = aE(X|\mathcal{G}) + bE(Y|\mathcal{G})$
iii) Si $X \leq Y$ con probabilidad 1, entonces $E(X|\mathcal{G}) \leq E(Y|\mathcal{G})$
iv) $|E(X|\mathcal{G})| \leq E(|X||\mathcal{G})$

Demostración:

Si $X = a$ con probabilidad 1, la función idénticamente igual a a satisface las condiciones (i) y (ii) de la definición de $E(X|\mathcal{G})$, y por tanto (i) se sigue por unicidad.
Para (ii) tenemos que $aE(X|\mathcal{G}) + bE(Y|\mathcal{G})$ es integrable y $\mathcal{G}$-medible, y

$$\int_G [aE(X|\mathcal{G}) + bE(Y|\mathcal{G})]\, dP = a\int_G E(X|\mathcal{G})\, dP + b\int_G E(Y|\mathcal{G})\, dP$$

$$= a\int_G X\, dP + b\int_G Y\, dP = \int_G (aX + bY)\, dP$$

y esto para toda G en $\mathcal{G}$, por tanto esta función satisface la ecuación funcional de la definición de $E(X|\mathcal{G})$.
(iii) Si $X \leq Y$ con probabilidad 1, entonces por (ii)

$$\int_G [E(Y|\mathcal{G}) - E(X|\mathcal{G})]\, dP = \int_G (Y - X)\, dP \geq 0 \quad \forall G \in \mathcal{G}.$$

Como $E(Y|\mathcal{G}) - E(X|\mathcal{G})$ es $\mathcal{G}$-medible, este tiene que ser no-negativo con probabilidad 1; por lo tanto $E(X|\mathcal{G}) \leq E(Y|\mathcal{G})$. Esto prueba (iii) así como el hecho de que $E(X|\mathcal{G}) = E(Y|\mathcal{G})$ si $X = Y$ con probabilidad 1.
(iv) Se sigue directamente de (iii) considerando que $-|X| \leq X \leq |X|$. □

De aqui en adelante, $Y, Y_1, Y_2, \ldots$ son v.a. en $(\Omega, \mathcal{F}, P)$ y supondremos que sus esperanzas existen.

Teorema: 1.34
Si $Y_n \geq 0 \quad \forall n$ y $Y_n \uparrow Y$ c.d. P, entonces
a) $E(Y_n|\mathcal{G}) \uparrow E(Y|\mathcal{G})$ c. d. P
Si todas $Y_n \geq 0$, entonces
b) $E(\sum_{n=1}^{\infty} Y_n|\mathcal{G}) = \sum_{n=1}^{\infty} E(Y_n|\mathcal{G})$ c.d. P
En particular, si $B_1, B_2, \ldots$ son conjuntos disjuntos en $\mathcal{F}$
c) $P(\cup_{n=1}^{\infty} B_n|\mathcal{G}) = \sum_{n=1}^{\infty} P(B_n|\mathcal{G})$ c.d. P

Demostración:

a) $\int_G Y_n\, dP = \int_G E(Y_n|\mathcal{G})\, dP$, G en $\mathcal{G}$; por (iii) del teorema (1.33), los $E(Y_n|\mathcal{G})$ crecen a una función h, $\mathcal{G}$-medible.
Por el teorema de la convergencia monótona $\int_G Y\, dP = \int_G h\, dP$ con $h = E(Y|\mathcal{G})$ c.d. P.
b) Por (ii) del teorema 1.33, $E(\sum_{k=1}^{n} Y_k|\mathcal{G}) = \sum_{k=1}^{n} E(Y_k|\mathcal{G})$ c.d. P
Consideremos $n \to \infty$ y aplicando (a) obtenemos el resultado deseado.
c) Es un caso especial de (b). □

Teorema: 1.35
$E[E(Y|\mathcal{G})] = E(Y)$; por tanto si Y es integrable, también lo es $E(Y|\mathcal{G})$.

Demostración:
$E(Y) = \int_\Omega Y\,dP = \int_\Omega E(Y|\mathcal{G})\,dP = E(E(Y|\mathcal{G}))$ □

Teorema: 1.36
Si $|Y_n| \leq Z \quad \forall n$, con $E(Z)$ finita, Z $\mathcal{G}$-medible, y $Y_n \to Y$ c.d. P entonces
$E(Y_n|\mathcal{G}) \to E(Y|\mathcal{G})$ c.d. P.
Este teorema es análogo al teorema de la convergencia dominada para esperanzas condicionales.

Demostración:
Sea $Z_n = \sup_{k\geq n}|Y_k - Y|$, entonces Z_n es variable aleatoria y $Z_n \downarrow 0$ c.d. P. Ahora por el teorema anterior $E(Y_n|\mathcal{G})$ y $E(Y|\mathcal{G})$ son finitas c.d. P, y $|E(Y_n|\mathcal{G}) - E(Y|\mathcal{G})| \leq E[|Y_n - Y||\mathcal{G}]$ por (i) y (ii) del teorema 1.33. Esto es menor o igual que $E(Z_n|\mathcal{G})$ por (iii) del teorema 1.33. Por tanto es suficiente con probar que $E(Z_n|\mathcal{G}) \to 0$ c.d. P.
Por (iii) del teorema 1.33, $E(Z_n|\mathcal{G}) \downarrow h$ c.d. P para alguna función h $\mathcal{G}$-medible.
Como $0 \leq Z_n \leq 2Z$, que es integrable, tenemos que

$$0 \leq \int_\Omega h\,dP \leq \int_\Omega E(Z_n|\mathcal{G})\,dP = \int_\Omega Z_n\,dP \to 0$$

por el teorema de la convergencia dominada, por lo tanto $h = 0$ c.d. P. □

Teorema: 1.37
Consideremos $Y_n \geq Z \quad \forall n$, donde $E(Z) > -\infty$, Z $\mathcal{G}$- medible.
a) Si $Y_n \uparrow Y$ c.d. P, entonces $E(Y_n|\mathcal{G}) \uparrow E(Y|\mathcal{G})$ c.d. P
b) $\underline{\lim}_{n\to\infty} E(Y_n|\mathcal{G}) \geq E(\underline{\lim}_{n\to\infty} (Y_n|\mathcal{G})$ c.d. P
Si consideramos ahora $Y_n \leq Z \quad \forall n$, donde $E(Z) < +\infty$
c) Si $Y_n \downarrow Y$ c.d. P, entonces $E(Y_n|\mathcal{G}) \downarrow E(Y|\mathcal{G})$ c.d. P
d) $\overline{\lim}_{n\to\infty} E(Y_n|\mathcal{G}) \leq E(\overline{\lim}_{n\to\infty} (Y_n|\mathcal{G})$ c.d. P

Demostración:

a) Tenemos $0 \leq Y_n^+ \uparrow Y^+$, por lo tanto $E(Y_n^+|\mathcal{G}) \to E(Y^+|\mathcal{G})$ c.d. P por el teorema 1.34 inciso (a).

Como $0 \leq Y_n^- \leq Z^-$ que es integrable, por lo tanto $E(Y_n^-|\mathcal{G}) \to E(Y^-|\mathcal{G})$ c.d. P por el teorema anterior. Ahora

$$E(Y_n|\mathcal{G}) = E(Y_n^+|\mathcal{G}) - E(Y_n^-|\mathcal{G}) \quad \text{por (ii) del teorema 1.33}$$

$$\Rightarrow E(Y^+|\mathcal{G}) - E(Y^-|\mathcal{G}) = E(Y|\mathcal{G}) \quad \text{por (ii) del teorema 1.33}$$

(Notemos que Y_n^+ y Y_n^- son integrables).

b) Sea $Y_n' = \inf_{k \geq n} Y_k; \Rightarrow Y_n' \uparrow Y' = \lim_{n \to \infty} Y_n$

Por (a) $E(Y_n'|\mathcal{G}) \uparrow E(Y'|\mathcal{G})$ c.d. P, pero $Y_n' \leq Y_n$ por tanto

$$E(Y'|\mathcal{G}) = \lim_{n \to \infty} E(Y_n'|\mathcal{G}) = \underline{\lim}_{n \to \infty} E(Y_n'|\mathcal{G}) \leq \underline{\lim}_{n \to \infty} E(Y_n|\mathcal{G}) \quad \text{por (iii) del teorema1.33}$$

c) Se sigue de (a) substituyendo todas las variables aleatorias por sus inversas.

d) $E(\overline{\lim}_{n \to \infty} Y_n|\mathcal{G}) = -E(\underline{\lim}_{n \to \infty} (-Y_n)|\mathcal{G}) \geq -\underline{\lim}_{n \to \infty} E(-Y_n|\mathcal{G})$ por (b)

$$= \overline{\lim}_{n \to \infty} E(Y_n|\mathcal{G})$$

Nótese la similitud de este teorema, con el teorema de la convergencia monótona extendido, y con el lema de Fatou.

Teorema: 1.38

Si Z es $\mathcal{G}$-medible, y si Y y ZY son integrables, entonces $E(ZY|\mathcal{G}) = ZE(Y|\mathcal{G})$ con probabilidad 1.

En particular $E(Z|\mathcal{G}) = Z$.

Observación: 1.39

Si X es $\mathcal{G}$-medible, entonces claramente $E(X|\mathcal{G}) = X$ con probabilidad 1. Este teorema es una generalización de este hecho. Para un observador con la información en $\mathcal{G}$, X es efectivamente una constante si es $\mathcal{G}$-medible.

Demostración.

Si Z es un indicador I_B, $B \in \mathcal{G}$, y $C \in \mathcal{G}$, tenemos que

$$\int_C YZ\,dP = \int_{C \cap B} Y\,dP = \int_{C \cap B} E(Y|\mathcal{G})\,dP = \int_C I_B E(Y|\mathcal{G})\,dP = \int_C Z\,E(Y|\mathcal{G})\,dP,$$

y $ZE(Y|\mathcal{G})$ es $\mathcal{G}$-medible.

Por lo tanto el resultado es válido para indicadores.
Sea ahora Z una función simple, es decir $Z = \sum_{j=1}^{n} Z_j I_{B_j}$ con $B_j \in \mathcal{G}$.
Por (ii) del teorema 1.33

$$E(ZY|\mathcal{G}) = \sum_{j=1}^{n} Z_j E(Y I_{B_j}|\mathcal{G}) = \sum_{j=1}^{n} Z_j I_{B_j} E(Y|\mathcal{G}) = Z E(Y|\mathcal{G}).$$

Si Z es una función arbitraria $\mathcal{G}$-medible, consideremos $Z_1, Z_2, \ldots$ funciones simples y $\mathcal{G}$-medibles tales que $|Z_n| \leq |Z|$ y $Z_n \to Z$.
Ahora $E(YZ_n|\mathcal{G}) = Z_n E(Y|\mathcal{G})$ por lo que acabamos de probar, y $E(YZ_n|\mathcal{G}) \to E(YZ|\mathcal{G})$ por el teorema 1.36 y porque YZ es integrable.
Como Y es integrable, también lo es $E(Y|\mathcal{G})$; por tanto $E(Y|\mathcal{G})$ es finito c.d. y en consecuencia $Z_n E(Y|\mathcal{G}) \to Z E(Y|\mathcal{G})$ c.d.
Por lo tanto $E(ZY|\mathcal{G}) = Z E(Y|\mathcal{G})$. □

Nota:
Es importante que $E(Y|\mathcal{G})$ sea finito pues, por ejemplo, $\frac{1}{n} \to 0$ pero $\frac{1}{n}(\infty)$ no converge a 0.

El tomar un valor de esperanza condicional, puede pensarse como una operación promedio o suavizadora (smoothing).
Esto nos hace pensar que promediando X con respecto a $\mathcal{G}_2$, y luego el resultado promediarlo con respecto a $\mathcal{G}_1$, una σ-álgebra más pequeña, nos debe resultar lo mismo que promediar a X desde un principio con $\mathcal{G}_1$.

Teorema: 1.40 De la proyección
Si X es integrable y las σ-álgebras $\mathcal{G}_1$ y $\mathcal{G}_2$ satisfacen $\mathcal{G}_1 \subset \mathcal{G}_2$, entonces
a) $E[E(X|\mathcal{G}_2)|\mathcal{G}_1] = E(X|\mathcal{G}_1)$ con probabilidad 1.
b) $E[E(X|\mathcal{G}_1)|\mathcal{G}_2] = E(X|\mathcal{G}_1)$ c.d. P

Demostración:

a) Si $G \in \mathcal{G}_1$, entonces $\int_G E(X|\mathcal{G}_1)\, dP = \int_G X\, dP = \int_G E(X|\mathcal{G}_2)\, dP$ pues $G \in \mathcal{G}_2$
$\Rightarrow E[E(X|\mathcal{G}_2)|\mathcal{G}_1] = E(X|\mathcal{G}_1)$ c.d.
Alternativamente, si $G \in \mathcal{G}_1$, entonces $\int_G E[E(X|\mathcal{G}_2)|\mathcal{G}_1]\, dP = \int_G E(X|\mathcal{G}_2)\, dP = \int_G X\, dP$ pues $G \in \mathcal{G}_2$; $\Rightarrow E(X|\mathcal{G}_1) = E[E(X|\mathcal{G}_2)|\mathcal{G}_1]$ c.d.
b) $E(X|\mathcal{G}_1)$ es $\mathcal{G}_1$-medible, y por tanto $\mathcal{G}_2$-medible, y

$$\int_G E(X|\mathcal{G}_1)\, dP = \int_G E(X|\mathcal{G}_1)\, dP, \quad G \in \mathcal{G}_2.$$

□

Esperanza Condicional como Operadores de Proyección.

Hasta aqui hemos estudiado a la esperanza condicional como una función medible, ahora vamos a estudiarla desde otro punto de vista. La idea en este punto es probar que la misma esperanza condicional con la que hemos venido trabajando se puede estudiar también como un operador funcional en los espacios $\mathcal{L}_p$, $1 \leq p \leq \infty$. Buscamos probar esta equivalencia porque en algunos casos puede ser más cómodo trabajar a la esperanza condicional como un operador funcional.

Sin embargo, antes de entrar en materia enunciaremos y probaremos la desigualdad de Jensen pues va a ser una herramienta fundamental en lo que resta de este trabajo:

Desigualdad de Jensen: 2.1
Sea $g: I \to \mathbb{R}$ convexa. Sea $X \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$ v. a., con $X(\omega) \in I \ \forall\ \omega$. Si $\mathcal{H}$ es una sub-σ-álgebra de $\mathcal{F}$, entonces

$$E\big(g(X) \mid \mathcal{H}\big) \geq g\big(E(X \mid \mathcal{H})\big) \quad \text{c.d.}$$

En particular $E\big(g(X)\big) \geq g\big(E(X)\big)$.

Demostración:

Primero notemos que $E\big(X \mid \mathcal{H}\big) \in I$ c.d. pues, supongamos que $a \in \mathbb{R}$ y $X > a$, entonces $E\big(X \mid \mathcal{H}\big) > a$ c.d. pues

$$0 \geq \int_{\{E(X|\mathcal{H}) \leq a\}} E\big(X - a \mid \mathcal{H}\big)\, dP = \int_{\{E(X|\mathcal{H}) \leq a\}} (X - a)\, dP \geq 0$$

por lo tanto $X = a$ c.d. en $\{E(X|\mathcal{H}) \leq a\}$, lo que implica que $P\{E(X|\mathcal{H}) \leq a\} = 0$.
por lo tanto $g\big(E(X \mid \mathcal{H})\big)$ está bien definida.

En el teorema anterior podemos escribir $g(y) = \sup\limits_n (a_n y + b_n)$, $y \in I$, así

$g(x) \geq a_n x + b_n \quad \forall\ n$.
por lo tanto $E\big(g(x) \mid \mathcal{H}\big) \geq a_n E\big(X \mid \mathcal{H}\big) + b_n$ c.d.
Si consideramos el supremo sobre las n obtenemos el resultado. $\square$

Teorema: 2.2
Sea $\mathcal{B} \subset \mathcal{F}$ una σ-álgebra y $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Entonces el operador esperanza condicional $E(\circ|\mathcal{B})$: $\mathcal{L}_p(\Omega, \mathcal{F}, P) \to \mathcal{L}_p(\Omega, \mathcal{F}, P)$, $1 \leq p \leq \infty$, es una proyección lineal positiva y una contracción (*i.e.* $\|E(f|\mathcal{B})\|_p \leq \|f\|_p$). Su rango es $\mathcal{L}_p(\Omega, \mathcal{B}, P_{\mathcal{B}})$.

Observación:
Como se había mencionado anteriormente, $P_{\mathcal{B}}$ es la probabilidad restringida a la σ–álgebra $\mathcal{B}$.

Demostración:
Se probará primero la existencia de dicho operador.
Como $P(\Omega) = 1$ tenemos que $\mathcal{L}_p(\Omega, \mathcal{F}, P) \subset \mathcal{L}_1(\Omega, \mathcal{F}, P)$ así si $f \in \mathcal{L}_p(\Omega, \mathcal{F}, P) \Rightarrow f \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$, *i.e.* f es integrable y por lo tanto $E(\bullet|\mathcal{B}): f \longrightarrow E(f|\mathcal{B}) \in \mathcal{L}_1(\Omega, \mathcal{B}, P_{\mathcal{B}})$ existe. Esto de hecho nos dice que la función que a cada $f \in \mathcal{L}_p$ le asocia $E(f|\mathcal{B})$ pertenece a $\mathcal{L}_1$.

Para ver que $E(\bullet|\mathcal{B})$ es un operador lineal positivo:

i) consideremos $\nu(A) = \int\limits_A f\,dP$, entonces $\nu: \mathcal{B} \to \mathbb{R}$ es una función σ–aditiva que es absolutamente contínua relativa a $P_{\mathcal{B}}$, y con mayor razón a P.

Sea $f \geq 0$ c.d. $\Rightarrow$ $\nu(A) \geq 0$ $\Rightarrow \frac{d\nu}{dP} \geq 0 \Rightarrow E(f|\mathcal{B}) \geq 0$ c.d. pues por el teorema de Radon–Nikodym $E(f|\mathcal{B}) = f$ c.d. siempre que f sea $\mathcal{B}$–medible.

ii) $E(\bullet|\mathcal{B}): f \longrightarrow E(f|\mathcal{B})$ es un operador lineal *i.e.* si f_1, f_2 son integrables, entonces $E(af_1 + bf_2|\mathcal{B}) = aE(f_1|\mathcal{B}) + bE(f_2|\mathcal{B})$ c.d. y esto es gracias a las propiedades de la integral.

Probaremos a continuación que $E(\circ|\mathcal{B})$ es una proyección.

Ya hemos demostrado que si $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \mathcal{F}$ son σ– álgebras y X es integrable, entonces

$$E(E(X|\mathcal{B}_2)|\mathcal{B}_1) = E(E(X|\mathcal{B}_1)|\mathcal{B}_2) = E(X|\mathcal{B}_1) \quad c.d. \qquad ...(\Diamond)$$

$$\Rightarrow E(E(\bullet|\mathcal{B})|\mathcal{B}) = E(\bullet|\mathcal{B}) \quad i.e. \quad E(\circ|\mathcal{B}) \text{ es un operador idempotente.}$$

Vamos a probar ahora que se trata de una contracción;
i.e. $\|E(f|\mathcal{B})\|_p \leq \|f\|_p$.

i) Si $p = +\infty$, entonces $|f| \leq \|f\|_\infty$ c.d., de tal forma que

$$|E(f|\mathcal{B})| \leq E(|f| \mid \mathcal{B}) \leq E(\|f\|_\infty \mid \mathcal{B}) = \|f\|_\infty \quad \text{c.d.}$$

por lo tanto $\|E(f \mid \mathcal{B})\|_\infty \leq \|f\|_\infty$ y $E(\bullet|\mathcal{B})$ es una contracción en este caso.

Por lo tanto, por tratarse de un operador lineal idempotente y contractivo, tenemos que $E(\bullet|\mathcal{B})$ es una proyección.

Observación:
Esto también muestra que $E(f|\mathcal{B}) \in \mathcal{L}_\infty(\Omega, \mathcal{B}, P)$.

ii) Si $1 \leq p < \infty$, sea $\varphi(X) = |X|^p$. Entonces $\varphi(\bullet)$ es una función contínua convexa en R, y $f, \varphi(f) \in \mathcal{L}_1(\Omega, \mathcal{F}, P)$.
Por la desigualdad de Jensen esto implica que

$$E(\varphi(f)|\mathcal{B}) \geq \varphi(E(f|\mathcal{B})) \quad \text{c.d.} \qquad ...(*)$$

tenemos entonces;

$$\begin{aligned} ||f||_p^p &= \int_\Omega \varphi(f)\, dP \quad \text{por la definición de } \varphi \text{ y de } ||f||_p \\ &= \int_\Omega E(\varphi(f)|\mathcal{B})\, dP_{\mathcal{B}} \quad \text{por definición de esperanza condicional} \\ &\geq \int_\Omega \varphi(E(f|\mathcal{B}))\, dP_{\mathcal{B}} \quad \text{por } (*) \\ &= ||E(f|\mathcal{B})||_p^p \end{aligned}$$

Se sigue entonces que $E(f|\mathcal{B}) \in \mathcal{L}_p(\Omega, \mathcal{B}, P_{\mathcal{B}})$ y que $||E(f|\mathcal{B})||_p \leq ||f||_p$
Por lo tanto $E(\bullet|\mathcal{B})$ es efectivamente una contracción y en este caso con rango contenido en $\mathcal{L}_p(\Omega, \mathcal{B}, P_{\mathcal{B}})$.
Como $E(\bullet|\mathcal{B})$ es la identidad en $\mathcal{L}_p(\Omega, \mathcal{B}, P_{\mathcal{B}})$, su rango es precisamente $\mathcal{L}_p(\Omega, \mathcal{B}, P_{\mathcal{B}})$. □

Teorema: 2.3
Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $\mathcal{B} \subset \mathcal{F}$ una $\sigma-$ álgebra.
Si $\mathcal{L}_p(\mathcal{F})$ y $\mathcal{L}_p(\mathcal{B})$, $1 \leq p < \infty$ son los espacios de Lebesgue en $(\Omega, \mathcal{F}, P)$ y $(\Omega, \mathcal{B}, P_{\mathcal{B}})$ respectivamente, entonces existe una proyección $Q: \mathcal{L}_p(\mathcal{F}) \to \mathcal{L}_p(\mathcal{F})$, $Q(\mathcal{L}_p(\mathcal{F})) = \mathcal{L}_p(\mathcal{B})$, llamada $Q = E(\bullet|\mathcal{B})$, que es positiva, y es una contracción, y $Q1 = 1$ c.d. Si $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \mathcal{F}$ son $\sigma-$ álgebras y $Q_i = E(\bullet|\mathcal{B}_i)$, $i = 1, 2$ son las proyecciones correspondientes, entonces $Q_1Q_2 = Q_2Q_1 = Q_1$.

Más aún, para cada $1 < p \leq \infty$, $Q(\mathcal{L}_p(\mathcal{F})) = \mathcal{L}_p(\mathcal{B})$, $Q = E(\cdot|\mathcal{B})$ tiene una única extensión a $\hat{Q}$, $\hat{Q}(\mathcal{L}_1(\mathcal{F})) = \mathcal{L}_1(\mathcal{B})$, donde $\hat{Q}$ es una esperanza condicional $\mathcal{L}_1$- contínua.

Observación:
Este teorema es "casi" el teorema anterior.

Demostración:

Por el teorema anterior y por ($\Diamond$) se cumple que $Q_1Q_2 = Q_2Q_1 = Q_1$. Lo demás excepto lo de la extensión ya se probó en el teorema anterior.

Para probar la extensión:
Sabemos que $\mathcal{L}_p(\mathcal{F}) \subset \mathcal{L}_1(\mathcal{F})$ así que Q no sólo está definido en un subespacio de $\mathcal{L}_1(\mathcal{F})$, sino que por el teorema anterior tenemos que es contínuo en norma $\mathcal{L}_1$.
Como $\mathcal{L}_p(\mathcal{F})$, es denso en $\mathcal{L}_1(\mathcal{F})$, la función lineal Q **uniformemente** contínua en un subconjunto denso en $\mathcal{L}_1(\mathcal{F})$, tiene una única extensión a todo $\mathcal{L}_1(\mathcal{F})$ por el "principio de extensión por continuidad" en espacios métricos. □

Nota:
En la prueba de la última parte del teorema, se usó el siguiente teorema de topología:

Si $\mathcal{X}$ y $\mathcal{Y}$ son dos espacios métricos y $\mathcal{Y}$ es completo, si $\mathcal{D} \subset \mathcal{X}$ es un subconjunto denso, entonces una función $f : \mathcal{D} \longrightarrow \mathcal{Y}$ uniformemente contínua tiene una única extensión $\hat{f}: \mathcal{X} \to \mathcal{Y}$ que es uniformemente contínua en $\mathcal{X}$.

Leyendo este teorema, surgen de momento dos preguntas:
1.- Existirán algunos otros operadores en $\mathcal{L}_p(\mathcal{F})$ sobre $\mathcal{L}_p(\mathcal{B})$, $p \geq 1$, que sean una proyección y una contracción que no sean necesariamente una esperanza condicional?

2.- Si $Q: \mathcal{L}_p(\mathcal{F}) \to \mathcal{L}_p(\mathcal{F})$ es una proyección y una contracción, bajo qué condiciones es una esperanza condicional?

Estas dos preguntas vienen a ser en cierto sentido un inverso a este último teorema; para analizarlas estudiemos el siguiente resultado:

Teorema: 2.4
Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $\mathcal{B} \subset \mathcal{F}$ una σ-álgebra. Si $\mathcal{L}_p(\mathcal{F})$ y $\mathcal{L}_p(\mathcal{B})$ $(\subset \mathcal{L}_p(\mathcal{F}))$, $1 \leq p < \infty$ son los espacios de Lebesgue en $(\Omega, \mathcal{F}, P)$ y $(\Omega, \mathcal{B}, P_{\mathcal{B}})$ respectiva-

mente, entonces toda proyección Q en $\mathcal{L}_p(\mathcal{F})$ lineal y que es una contracción, con rango $\mathcal{L}_p(\mathcal{B})$, coincide con la esperanza condicional $E(\cdot|\mathcal{B})$, *i.e.* $Q(f) = E(f|\mathcal{B}) \quad \forall\, f \in \mathcal{L}_p(\mathcal{F})$.

Demostración:

La demostración se va a dividir en dos casos $1 < p < \infty$ y $p = 1$.

Caso 1) $1 < p < \infty$.
Sea $q = \frac{p}{p-1}$. Si $f_0 \in \mathcal{L}_p(\mathcal{B})$ es un elemento acotado, entonces es claro que

$$f_0 \in \mathcal{L}_q(\mathcal{B}) \subset \mathcal{L}_p(\mathcal{F}) \qquad \ldots(*)$$

La idea en ambos casos es probar que

$$\int_A Qf\, dP = \int_A f\, dP \quad \forall\, f \in \mathcal{L}_p(\mathcal{F}) \quad \text{y} \quad A \in \mathcal{B},$$

de tal suerte que el resultado se concluya.

Recordemos que $(\mathcal{L}_p(\mathcal{F}))^*$, el espacio dual de $\mathcal{L}_p(\mathcal{F})$, se identifica también con $\mathcal{L}_q(\mathcal{F})$ para $1 \leq p < \infty$.
El operador $Q^* : \mathcal{L}_q \longrightarrow \mathcal{L}_q$, definido por:

$$\langle Qf, g\rangle = \int_\Omega (Qf)\, g\, dP = \int_\Omega f(Q^*g)\, dP = \langle f, Q^*g\rangle \quad \forall\, f \in \mathcal{L}_p(\mathcal{F}) \quad \text{y} \quad g \in \mathcal{L}_q(\mathcal{F})$$

es llamado el adjunto de Q y es un operador lineal acotado.

Más aún $||Q|| = ||Q^*||$, y $(Q^*)^2 = Q^*$ si $Q^2 = Q$, por lo tanto Q^* es una proyección y una contracción si Q también lo es.

De hecho, si $Q^2 = Q$ y $g \in \mathcal{L}_q(\mathcal{F})$, entonces

$$\langle f, Q^*g\rangle = \langle Qf, g\rangle = \langle Q^2 f, g\rangle = \langle Qf, Q^*g\rangle = \langle f, Q^{*^2} g\rangle \qquad \ldots(1)$$

Como $f \in \mathcal{L}_p(\mathcal{F})$ es arbitraria, $Q^*g = Q^{*^2}g$ y por lo tanto Q^* es una proyección.

Respecto a la ecuación de normas, por la definición de norma de un operador, tenemos:

$$\begin{aligned}||Q^*|| &= \sup\{||Q^*g||_q : ||g||_q \leq 1\} \quad \text{por definición}\\ &= \sup\{\sup\{|\langle f, Q^*g\rangle| : ||f||_p \leq 1\} : ||g||_q \leq 1\}\end{aligned}$$

como $||h||_q = \sup\{|\int_\Omega hf\,dP| : ||f||_p \leq 1\}$,

$$\begin{aligned}&= \sup\{\sup\{|\langle Qf, g\rangle| : ||f||_p \leq 1\} : ||g||_q \leq 1\}\\ &= \sup\{\sup\{|\langle Qf, g\rangle| : ||g||_q \leq 1\} : ||f||_p \leq 1\} \quad \text{por definición de } Q \text{ y } Q^*\\ &= \sup\{||Qf||_p : ||f||_p \leq 1\} = ||Q||\end{aligned}$$

por lo tanto $||Q^*|| \leq 1$ si $||Q|| \leq 1$, lo que se cumple, pues Q es una contracción.

Estos cálculos son ciertos para todos los operadores acotados en un espacio de Banach.

Si $\mathcal{L}_q$ es considerado como $(\mathcal{L}_p)^*$ y si $h_0 = Q^*f_0 \in \mathcal{L}_q(\mathcal{F})$ con $f_0 \in \mathcal{L}_q(\mathcal{B})$ entonces h_0 no sólo está en $\mathcal{L}_q(\mathcal{F})$ sino que está en $\mathcal{L}_q(\mathcal{B})$. En efecto:

$$||h_0||_q = ||Q^*f_0||_q \leq ||Q^*||\,||f_0||_q \leq ||f_0||_q \qquad ...(2)$$

Por otro lado, si $A \in \mathcal{B}$ es tal que $\chi_A \in \mathcal{L}_p(\mathcal{B}) = Q(\mathcal{L}_p(\mathcal{F}))$, y como $f_0 \in \mathcal{L}_p(\mathcal{F}) \Rightarrow f_0$ es integrable, tenemos:

$$\int_A f_0\,dP_{\mathcal{B}} = \int_\Omega \chi_A f_0\,dP_{\mathcal{B}} = \int_\Omega Q(\chi_A)f_0\,dP_{\mathcal{B}}, \quad \text{pues} \quad Q\chi_A = \chi_A,$$

$$= \int_\Omega \chi_A(Q^*f_0)\,dP = \int_\Omega \chi_A h_0\,dP = \int_A h_0\,dP = \int_A E(h_0|\mathcal{B})\,dP_{\mathcal{B}} \qquad ...(3)$$

Pero f_0 y $E(h_0|\mathcal{B})$ son $\mathcal{B}$-medibles y $A \in \mathcal{B}$ es arbitrario, se sigue entonces de (3) que $f_0 = E(h_0|\mathcal{B})$ c.d.

Ahora, por un teorema anterior $E(\bullet|\mathcal{B})$ es una contración en $\mathcal{L}_q$ así que:

$$||f_0||_q = ||E(h_0|\mathcal{B})||_q \leq ||h_0||_q \qquad ...(4)$$

De (2) y (4) podemos concluír que $||f_0||_q = ||h_0||q$.

Como $1 < q < \infty$, $\varphi(X) = |X|^q$ es convexa, $\Rightarrow$ por la desigualdad de Jensen

$$\varphi(f_0) = \varphi(E(h_0|\mathcal{B})) \leq E(\varphi(h_0)|\mathcal{B}) \quad \text{c.d.} \qquad ...(5)$$

y la desigualdad se cumple en un conjunto de medida positiva siempre que h_0 no sea $\mathcal{B}$-medible.

Si la primera posibilidad se cumple

$$||f_0||_q^q = \int_\Omega \varphi(f_0)\, dP < \int_\Omega E(\varphi(h_0)|\mathcal{B})\, dP = \int_\Omega \varphi(h_0)\, dP = ||h_0||_q^q \qquad ...(6)$$

pero esto contradice el hecho que habíamos probado de que

$$||f_0||_q = ||h_0||_q \quad \Rightarrow \quad h_0 \quad \text{es} \quad \mathcal{B}-\text{medible}.$$

Así $f_0 = E(h_0|\mathcal{B}) = h_0$ c.d. y f_0 es un punto fijo en Q^*.

$$\Rightarrow \forall \quad A \in \mathcal{B},\ \chi_A \in \mathcal{L}_p(\mathcal{B}) \cap \mathcal{L}_q(\mathcal{B}), \quad \text{y} \quad \forall f \leq \mathcal{L}_p(\mathcal{F}),$$

$$\int_A Qf\, dP_{\mathcal{B}} = \int_\Omega \chi_A Qf\, dP_{\mathcal{B}} = \int_\Omega Q^*(\chi_A) f\, dP = \int_\Omega \chi_A f\, dP, \quad \text{ya que} \quad Q^*(\chi_A) = \chi_A,$$

$$= \int_A f\, dP = \int_A E(f|\mathcal{B})\, dP_{\mathcal{B}} \qquad ...(7)$$

Aqui usamos nuevamente que $\mathcal{L}_p(\mathcal{F}) \subset \mathcal{L}_1(\mathcal{F})$.
De (7) se sigue que, como $Qf \in \mathcal{L}_p(\mathcal{B}) \Rightarrow Qf$ es $\mathcal{B}$–medible, $Qf = E(f|\mathcal{B}) \quad \forall f \in \mathcal{L}_p(\mathcal{F})$.
Con esto concluímos el primer caso.

Caso 2) $p = 1$
Como Q es la identidad en $\mathcal{L}_1(\mathcal{B})$ y las constantes están en $\mathcal{L}_1(\mathcal{B}) \Rightarrow Q1 = 1$ c.d. de hecho $Q1 \cdot 1_\Omega = 1 \cdot Q1_\Omega = 1 \cdot 1_\Omega = 1$
por lo tanto si $0 \leq f \leq 1$ c.d. es un elemento arbitrario de $\mathcal{L}_1(\mathcal{F})$, entonces por la linealidad y contractividad de Q,

$$\begin{aligned} 1 - \int_\Omega f\, dP &= \int_\Omega (1-f)\, dP_{\mathcal{B}} \geq \int_\Omega |Q(1-f)|\, dP_{\mathcal{B}} \quad \text{pues} \quad ||Q|| \leq 1 \\ &= \int_\Omega |(1 - Qf)|\, dP \quad \text{pues} \quad Q1 = 1 \\ &\geq 1 - \int_\Omega Qf\, dP \end{aligned}$$

Consecuentemente

$$\int_\Omega f\,dP \le \int_\Omega Qf\,dP \le \int_\Omega |Qf|\,dP \le \int_\Omega f\,dP \qquad ...(8)$$

Se sigue entonces que existe una igualdad por todas partes y que $Qf \geq 0$ c.d.

p.d. $Q(g) \geq 0$ para toda $g \geq 0$ en $\mathcal{L}_1(\mathcal{F})$

i) Si g es acotada $\Rightarrow ||g||_\infty < \infty$
$0 \leq \frac{g}{||g||_\infty} \leq 1$ si $g \neq 0$ c.d. (pues $g \geq 0$).
$\Rightarrow Q(\frac{g}{||g||_\infty}) \geq 0$, por la linealidad de Q y como $\frac{1}{||g||_\infty}$ es constante
$\Rightarrow \frac{1}{||g||_\infty}Q(g) \geq 0$ por lo tanto $Q(g) \geq 0$ pues $||g||_\infty \geq 0$.

ii) Si $g \geq 0$ y $g \in \mathcal{L}_1(\mathcal{F})$
considero la sucesión
$g_n = g \wedge n$

Observación:
$g_n \xrightarrow{\mathcal{L}_1} g$ entonces como Q es contínua $\Rightarrow 0 \leq Q(g_n) \to Q(g) \quad \forall\, n$
por lo tanto $Q(g) \geq 0$
demostración de la observación:
$0 \leq g - g_n \leq g \in \mathcal{L}_1(\mathcal{F})$
$g - g_n \longrightarrow 0$ puntualmente cuando $n \to \infty$
$\Rightarrow$ por el T.C.D.L. $\int g - g_n \longrightarrow \int 0 = 0 \quad g_n \xrightarrow{\mathcal{L}_1} g$
por lo tanto $Q_g \geq 0 \quad \forall\, g \geq 0$ en $\mathcal{L}_1(\mathcal{F})$
Mas aún (8) también implica

$$\int_\Omega f\,dP = \int_\Omega Qf\,dP_{\mathcal{B}} \qquad ...(9)$$

para $0 \leq f \leq 1$. Donde Qf es $\mathcal{B}$-medible.

p.d. esta igualdad se cumple $\forall\, g \in \mathcal{L}_1(\mathcal{F})$

i) para g acotada
$\Rightarrow 0 \leq \frac{g}{||g||_\infty} \leq 1$

$$\Rightarrow \int_\Omega \frac{g}{\|g\|_\infty}\, dP = \int_\Omega Q\left(\frac{g}{\|g\|_\infty}\right) dP_{\mathcal{B}}$$

como Q es lineal $\quad \frac{1}{\|g\|_\infty}\int_\Omega g\, dP = \frac{1}{\|g\|_\infty}\int_\Omega Q(g)\, dP_{\mathcal{B}}$

multiplico ambos lados por $\|g\|_\infty$ y $\int_\Omega g\, dP = \int_\Omega Q(g)\, dP_{\mathcal{B}} \quad \forall\, g$ acotada.

ii) $\forall g \in \mathcal{L}_1(\mathcal{F})$

Sea $g_n = g \wedge n \Rightarrow g_n \xrightarrow{\mathcal{L}_1} g$

$$\Rightarrow \int g\, dP = \lim_{n\to\infty} \int g_n\, dP = \lim_{n\to\infty}\int Q(g_n)\, dP_{\mathcal{B}} = \int Q(g)\, dP_{\mathcal{B}}$$

pues $\|Q(g_n) - Q(g)\|_1 = \|Q(g_n - g)\|_1$ por la linealidad

$$\le \|g_n - g\|_1 \longrightarrow 0$$

$$\Rightarrow \int |Q(g_n) - Q(g)|\, dP \longrightarrow 0 \quad \text{cuando } n \to \infty$$

Como $\lim_{n\to\infty}\left|\int (Q(g) - Q(g_n))\, dP\right| \le \lim_{n\to\infty}\int |Q(g) - Q(g_n)|\, dP \to 0$.

$$\Rightarrow \left|\lim_{n\to\infty}\int Q(g)\, dP - \lim_{n\to\infty}\int Q(g_n)\, dP\right| \to 0$$

$$\Rightarrow \lim_{n\to\infty}\int Q(g)\, dP = \int Q(g)\, dP = \lim_{n\to\infty}\int Q(g_n)\, dP$$

por lo tanto (9) se cumple $\forall\, f\ \mathcal{L}_1(\mathcal{F})$.

Afirmación:

(9) se cumple si Ω es sustituída por $A \in \mathcal{B}$.

De hecho, consideremos para $A \in \mathcal{B}$, $\quad 0 \le f \le \chi_A$ y $f \in \mathcal{L}_1(\mathcal{F})$

$$h = Q(\chi_A f) - Q(\chi_A Q f)$$

Pedimos que $h = 0$ c.d. Pues si Q es positivo se tiene que $0 \le Qf \le Q\chi_A$ y

$$\chi_A \pm h = \chi_A \pm \left[Q(\chi_A f) - Q(\chi_A Q f)\right] = \chi_A \pm \left[Q(\chi_A f - \chi_A Q f)\right]$$

$$= \chi_A \pm \left[Q(\chi_A(f - Qf))\right] = Q\left[\chi_A(\chi_A \pm f \mp Qf)\right] \quad \text{pues} \quad Q(\chi_A\chi_A) = Q(\chi_A) = \chi_A$$

$$\le Q(\chi_A \pm f \mp Qf) = Q\chi_A \pm Qf \mp Q^2 f = \chi_A \pm Qf \mp Qf = \chi_A \qquad ...(10)$$

Entonces (10) implica que el par de desigualdades $\chi_A + h \le \chi_A$ y $\chi_A - h = \chi_A$ son ambas ciertas.

De igual manera, $h \leq 0$ y $-h \leq 0$ simultáneamente.
Pero $h \in \mathcal{L}_1(\mathcal{F})$, por lo tanto $h = 0$ c.d. y $Q(\chi_A f) = Q(\chi_A Q f) = \chi_A Q f$
pues $\chi_A Q f \in \mathcal{L}_1(\mathcal{B})$ y Q es la identidad en $\mathcal{L}_1(\mathcal{B})$.

Vamos a probar que esta última ecuación es cierta $\forall\, 0 \leq f \leq 1$ c.d. en $\mathcal{L}_1(\mathcal{F})$ (y por lo tanto por linealidad para toda acotada, y por continudad de Q $\forall f \in \mathcal{L}_1(\mathcal{F})$).

Si $A \in \mathcal{B}$ es arbitrario, $0 \leq f\chi_A \leq \chi_A$ $(f \leq 1)$ y $0 \leq f\chi_A^c \leq \chi_A^c$ de tal forma que el caso especial implica (con $f\chi_A$ y $f\chi_A^c$)

$$Q(f) = Q(f\chi_A) + Q(f\chi_A^c) = \chi_A Q(f\chi_A) + \chi_A^c Q(f\chi_A^c) \qquad ...(P)$$

$$\Rightarrow \chi_A Q(f) = \chi_A Q(f\chi_A) + 0 = \chi_A Q(f\chi_A) = Q(f\chi_A) \qquad ...(11)$$

por tanto

$$\int_A Q(f)\, dP_{\mathcal{B}} = \int_A Q(f\chi_A)\, dP_{\mathcal{B}} + \int_A Q(f\chi_A^c)\, dP_{\mathcal{B}} \quad \text{por } (P)$$

$$= \int_A Q(f\chi_A)\, dP \quad \text{pues } P(A) = P_{\mathcal{B}}(A) \quad \text{si } A \in \mathcal{B}$$

$$= \int_\Omega \chi_A Q(f\chi_A)\, dP = \int_\Omega Q(\chi_A Q(f\chi_A))\, dP$$

pues $\chi_A Q(f\chi_A) \in \mathcal{L}_1(\mathcal{B})$ y Q es la identidad en $\mathcal{L}_1(\mathcal{B})$

$$= \int_\Omega Q(\chi_A f\chi_A)\, dP \quad \text{pues si consideramos } g = f\chi_A$$

$$\Rightarrow Q(\chi_A g) = Q(\chi_A Q g),\ g \in \mathcal{L}_1(\mathcal{F})$$

$$= \int_\Omega Q(\chi_A f)\, dP \quad \text{pues } \chi_A\chi_A = \chi_A$$

$$= \int_\Omega \chi_A f\, dP \quad \text{por } (9)$$

$$= \int_A f\, dP_{\mathcal{B}} = \int_A E(f|\mathcal{B})\, dP_{\mathcal{B}}$$

...(12)

Como los integrandos extremos en (12) *i.e.*
$\int_A Qf\, dP_{\mathcal{B}} = \int_A E(f|\mathcal{B})\, dP_{\mathcal{B}}$ son $\mathcal{B}$-medibles y $A \subset \mathcal{B}$ es arbitraria, podemos concluir que

$E(f|\mathcal{B}) = Qf$ c.d., para acotadas y por lo tanto $\forall f \in \mathcal{L}_1(\mathcal{F})$ $\square$

MARTINGALAS

Hasta el momento hemos definido el concepto de esperanza condicional y hemos analizado algunas de sus propiedades, esto nos va a servir ahora para estudiar un nuevo concepto que es el de martingalas.

Daremos primero la definición de este nuevo concepto, para después poder analizarlo.

Definición: 3.1

Sea $X_1, X_2, \cdots$ una sucesión de v.a. en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$, y sean $\mathcal{F}_1, \mathcal{F}_2, \ldots$ una sucesión de σ–álgebras en $\mathcal{F}$. La sucesión $\left\{ (X_n, \mathcal{F}_n);\ n = 1, 2 \cdots \right\}$ es una martingala si:

i) $\mathcal{F}_n \subset \mathcal{F}_{n+1}$;
ii) X_n es $\mathcal{F}_n$ – medible;
iii) $E(|X_n|) < \infty$;
iv) con probabilidad 1, tenemos :

$$E\left(X_{n+1} \mid \mathcal{F}_n\right) = X_n \quad \text{c.d.} \qquad ...(*)$$

Se dice entonces que la sucesión $(X_n)_n$ es una martingala con respecto a las σ–álgebras $(\mathcal{F}_n)_n$.

Observación:

La condición (ii) se expresa diciendo qu X_n es adaptada a $\mathcal{F}_n$.

Nota:

Si X_n representa la ganancia de un jugador después de la n–ésima jugada y $\mathcal{F}_n$ representa la información que tiene del juego hasta ese momento, entonces $(*)$ nos dice que la ganancia esperada después de la siguiente jugada es la misma que la que tiene ahora. Esto es, la martingala representa un juego justo.

Ejemplo: Sumas de v. a. i. 3.2

Sean $\Psi_0 = 0$ y $\Psi_1, \Psi_2, \cdots$ v. a. i. con $E(|\Psi_n|) < \infty$ y $E(\Psi_n) = 0 \quad \forall\, n$. Si definimos $X_0 = 0$ y $X_n = \Psi_1 + \ldots + \Psi_n$ para $n \geq 1$, entonces $\{X_n\}$ es una martingala con respecto a Ψ_n.

Demostración:

$i) E(|X_n|) \leq E(|\Psi_1|) + \cdots + E(|\Psi_n|) < \infty,$

$$\begin{aligned} ii) E(X_{n+1} \mid \Psi_0, \cdots, \Psi_n) &= E(X_n + \Psi_{n+1} \mid \Psi_0, \cdots, \Psi_n) \\ &= E(X_n \mid \Psi_0, \cdots, \Psi_n) + E(\Psi_{n+1} \mid \Psi_0, \cdots, \Psi_n) \\ &= X_n + E(\Psi_{n+1}) \text{ ya que las } \{\Psi_i\} \text{ son independientes} \\ &= X_n \quad \text{debido a que } E(\Psi_m) = 0 \text{ por def.} \end{aligned}$$

por lo tanto (X_n) es una martingala. □

Observación: 3.3

La sucesión $X_1, X_2, \ldots$ se define como una martingala si es una martingala relativa a una sucesión $\mathcal{F}_1, \mathcal{F}_2, \ldots$. Podemos notar que las σ-álgebras $\mathcal{G}_n = \sigma(X_1, \ldots, X_n)$ siempre nos convierten a una suc. $(X_n)_n$ en una martingala.

Demostración:

$i)$ $\mathcal{G}_n \subset \mathcal{G}_{n+1}$

$ii)$ X_n es $\mathcal{G}_n$-medible

$iii)$ Si $(*)$ es cierto, entonces

$$E(X_{n+1} \mid \mathcal{G}_n) = E\Big(E(X_{n+1} \mid \mathcal{F}_n)\mathcal{G}_n\Big) = E(X_n \mid \mathcal{G}_n) = X_n$$

esto último por propiedades de la esperanza que ya hemos analizado.
por lo tanto $(*)$ se reduce en este caso a:

$$E(X_{n+1} \mid X_1, \ldots, X_n) = X_n.$$

□

Como: $\sigma(X_1, \ldots, X_n) \in \mathcal{F}_n \Leftrightarrow X_n$ es $\mathcal{F}_n$ medible $\forall\, n$, entonces $\sigma(X_1, \ldots, X_n)$ son las mínimas σ- álgebras con respecto a las cuales X_n es una martingala.

Analicemos con más detalle la definición de martingala que hemos enunciado.
La condición (iii) nos asegura que $E(X_{n+1} \mid \mathcal{F}_n)$ existe. La condición (iv) nos dice que que X_n es una versión de $E(X_{n+1} \mid \mathcal{F}_n)$;

Como X_n es $\mathcal{F}_n$-medible, entonces $(*)$ se reduce a:

$$\int_A X_{n+1}\, dP = \int_A X_n\, dP\,, \quad A \in \mathcal{F}_n. \qquad \ldots(1)$$

Como los $\mathcal{F}_n$ están anidados, $A \in \mathcal{F}_n$ implica que

$$\int_A X_n\, dP = \int_A X_{n+1}\, dP = \ldots = \int_A X_{n+k}\, dP\,.$$

Por lo tanto, X_n, al ser $\mathcal{F}_n$–medible, es una versión de $E(X_{n+k} \mid \mathcal{F}_n)$:

$$E(X_{n+k} \mid \mathcal{F}_n) = X_n$$

con probabilidad 1 para $k \geq 1$. Si $A = \Omega \Rightarrow E(X_1) = E(X_2) = \ldots$.

Podemos entonces enunciar las condiciones de definición de una martingala en términos de sus diferencias.

$$\Delta_n = X_n - X_{n-1} \quad \Delta_1 = X_1 \, .$$

Por linealidad (*) es lo mismo que $E(\Delta_{n+1} \mid \mathcal{F}_n) = 0$.

Notemos que como $X_k = \Delta_1 + \cdots + \Delta_k$ y $\Delta_k = X_k - X_{k-1}$, los conjuntos $X_1, \ldots, X_n$ y $\Delta_1, \ldots, \Delta_n$ generan la misma σ–álgebra:

$$\sigma(X_1, \ldots, X_n) = \sigma(\Delta_1, \ldots, \Delta_n).$$

Observación: 3.4
La teoría de martingalas se cumple para todo intervalo de la recta, *i.e.* si consideramos a $\{(X_n, \mathcal{F}_n)\}$ una martingala, el subíndice n puede ser continuo en el caso de que n corra sobre un intervalo de recta, o discreto en el caso en que corra sobre los naturales, o sobre conjuntos numerables. En este trabajo vamos a estudiar únicamente el caso discreto.

Definición: 3.5
Las v.a. X_n son una submartingala relativa a las σ–álgebras $\mathcal{F}_n$ si $(i), (ii)$ y (iii) de la definición de martingala se quedan iguales y si (iv) se convierte en:
(iv') con probabilidad 1;

$$E(X_{n+1} \mid \mathcal{F}_n) \geq X_n. \qquad (**)$$

Igual que antes, $(X_n)_n$ es una submartingala si lo es con respecto a alguna sucesión $(\mathcal{F}_n)_n$, y la sucesión $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ es la mínima sucesión que funciona para tal efecto.
Pensar en (**) es lo mismo que

$$\int_A X_{n+1} \, dP \geq \int_A X_n \, dP \, , \quad A \in \mathcal{F}_n. \qquad \ldots(2)$$

Inductivamente, al igual que antes, podemos ver que

$$E(X_{n+k} \mid \mathcal{F}_n) \geq X_n. \quad \forall \, k \geq 1. \qquad \ldots(3)$$

Tomando las esperanzas en la ecuación anterior obtenemos que

$$E(X_1) \leq E(X_2) \leq \ldots \qquad \ldots(4)$$

Ejemplo: 3.6

Consideremos $(\Psi_n)_n$ igual que en el ejemplo 3.2, pero en este caso supongamos que $E(\Psi_n) \geq 0 \quad \forall\, n \geq 2$. Si definimos $X_0 = 0$ y $X_n = \Psi_1 + \ldots + \Psi_n$ para $n \geq 1$, entonces $\{X_n\}$ es una submartingala con respecto a Ψ_n.

Demostración:

$i) E(|X_n|) \leq E(|\Psi_1|) + \cdots + E(|\Psi_n|) < \infty,$

$ii) E(X_{n+1} \mid \Psi_0, \ldots, \Psi_n) = X_n + E(\Psi_{n+1}) \geq X_n$

por lo tanto (X_n) es una submartingala. □

Definición: 3.7

Invirtiendo la desigualdad $(**)$ en la definición anterior, obtenemos la definición de supermartingala. Las desigualdades (2), (3) y (4) se invierten también.

Así pues, la teoría de supermartingalas es simétrica a la de submartingalas.

Nota:

Volvamos nuevamente al jugador cuya ganancia es X_n después del n–ésimo juego y cuya información sobre el juego en este momento se representa por la σ–álgebra $\mathcal{F}_n$. Si $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$, él conoce la sucesión de sus ganancias y nada más, pero $\mathcal{F}_n$ pudiera ser mayor. La condición $(*)$ de martingala estipula que su ganancia esperada después del siguiente juego iguala su ganancia actual, de tal manera que la martingala se puede considerar como un modelo de juego justo.

En cambio la condición $(**)$ de submartingala nos dice que gana o al menos no pierde en el promedio, la submartingala nos representa un juego favorable al jugador, de igual manera la supermartingala nos representa un juego desfavorable para el jugador.

Veamos ahora algunos resultados importantes que usaremos para probar algunos teoremas de convergencia de martingalas.

Teorema de recta de soporte: 3.8

Sea $g: I \to \mathbb{R}$, donde I es un intervalo abierto en $\mathbb{R}$, acotado o no acotado. Supongamos que g es convexa, esto es:

$$g(ax + (1-a)y) \leq ag(x) + (1-a)g(y) \quad \forall \quad x, y \in I \text{ y } \forall\, a \in [0,1].$$

Entonces existen sucesiones $\{a_n\}$ y $\{b_n\}$ de números reales tales que $\forall\, y \in I$, tenemos

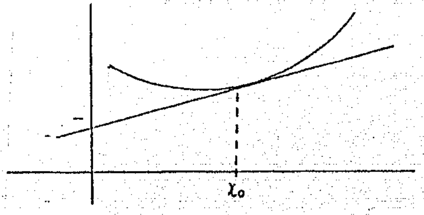
$$g(y) = \sup_n (a_n y + b_n).$$

para entender mejor el teorema veamos una definición:

Recta de soporte:

Sea φ una función convexa en (a,b) y $x_0 \in (a,b)$. La recta $y = m(x-x_0)+\varphi(x_0)$ a través de $(x_0, \varphi(x_0))$ es llamada una recta de soporte en x_0 si siempre está debajo de la gráfica de φ, esto es, si $\varphi(x) \geq m(x-x_0)+\varphi(x_0)$.

Tal linea es una linea de soporte si y sólo si la pendiente m está entre las derivadas izquierda y derecha de x_0



La idea del teorema es probar que g es el supremo de un número numerable de lineas de soporte.

Demostración:

Sea $g\colon I \to \mathbb{R}$ una función convexa, I un intervalo abierto en $\mathbb{R}$. Si $0 < h_1 < h_2$, entonces por la convexidad,

$$g(x \pm h_1) \leq \frac{(h_2-h_1)}{h_2}\, g(x) + \frac{h_1}{h_2}\, g(x \pm h_2);$$

$$\text{por tanto} \quad \frac{1}{h_1}\big[g(x+h_1)-g(x)\big] \leq \frac{1}{h_2}\big[g(x+h_2)-g(x)\big] \qquad \text{...(1)}$$

$$\text{y} \quad -\frac{1}{h_2}\big[g(x-h_2)-g(x)\big] \leq -\frac{1}{h_1}\big[g(x-h_1)-g(x)\big] \qquad \text{...(2)}$$

Además si $h, h' > 0$,

$$g(x) \leq \frac{h'}{h+h'}\, g(x-h) + \frac{h}{h+h'}\, g(x+h'),$$

$$\text{entonces} \quad \frac{g(x-h)-g(x)}{-h} \leq \frac{g(x+h')-g(x)}{h'} \qquad \text{...(3)}$$

Con (1) y (2) probamos que g'_+ y g'_- existen en I; y con (3) que estas derivadas son finitas. Más aún, de (1) y (2)

$$g'_+(x) = \inf_{y>x} \frac{g(y)-g(x)}{y-x} \quad , \quad g'_- = \sup_{y<x} \frac{g(x)-g(y)}{x-y} \qquad \text{...(4)}$$

$$\text{y por (3)} \quad g'_-(x) \leq g'_+(x) \qquad \text{...(5)}$$

Si $x_1 < x_2$, entonces

$$g'_+ \leq \frac{g(x_2) - g(x_1)}{x_2 - x_1} = \frac{g(x_1) - g(x_2)}{x_1 - x_2} \leq g'_-(x_2) \qquad ...(6)$$

por lo tanto g'_+ y g'_- son crecientes en I.

La existencia de las derivadas izquierda y derecha nos dice que g tiene límites izquierdo y derecho en cada punto, por tanto g es contínua, pues si no lo fuera g'_+ o g'_- sería infinita para algún punto.
Si $y \geq x$, entonces $g(y) \geq g(x) + (y - x)g'_+(x)$ y si $y < x$ entonces por (4)

$$g(y) \geq g(x) + (y - x)\, g'_-(x)$$

Podemos concluir entonces que si $g'_-(x) \leq a_x \leq g'_+(x)$, entonces

$$g(y) \geq g(x) + (y - x)a_x \quad \forall\, y \in I. \qquad ...(7)$$

Ahora bien, la función L_x dada por $L_x(y) = g(x) + (y - x)a_x, g \in I$, se llama linea de soporte de g en el punto x.

$$\Rightarrow \quad g(x) = sup\{L_s(x) \colon s \in I\}.$$

Se puede probar también que g es el supremo de un número numerable de lineas de soporte.
Si $x \in I$, sea r que se aproxima a x através de una sucesión de números racionales. Entonces $L_r(x) = g(x) + (x - r)a_r$.
Pero $g'_-(r) \leq a_r \leq g'_+(r)$, y como las g' son funciones crecientes y acotadas en un intervalo finito y cerrado dentro de I, tenemos que las a_r forman una sucesión acotada.
por lo tanto $L_r(x) \to g(x)$ por (7), $L_r(y) \leq g(y) \quad \forall\, y \in I$ y $\forall\, r$;
por lo tanto $g(x) = sup\{Ls(x) \colon s \in I, s$ racional$\}$, esto es,
$g(x) = sup\{g(s) + (x - s)as \colon s \in I,\ s$ racional$\}$. □

Veamos ahora un teorema que nos muestra una manera de construir submartingalas desde martingalas o desde otras submartingalas.

Teorema: 3.9
a) Sea $\{X_n, \mathcal{F}_n\}$ una submartingala, $g \colon \mathbb{R} \longrightarrow \mathbb{R}$ una función convexa creciente.
Si $g(X_n)$ es integrable $\forall\, n$, entonces $\{g(X_n), \mathcal{F}_n\}$ es una submartingala.

b) Sea $\{X_n, \mathcal{F}_n\}$ una martingala, $g \colon \mathbb{R} \longrightarrow \mathbb{R}$ una función convexa. Si $g(X_n)$ es integrable $\forall\, n$, entonces $\{g(X_n)|\mathcal{F}_n\}$ es una submartingala. En particular si $r \geq 1$, $\{X_n\}$ es una martingala y $|X_n|^r$ es integrable $\forall\, n$, entonces $\{|X_n|^r\}$ es una submartingala.

Demostración:

Por la desigualdad de Jensen tenemos que

$$E(g(X_{n+1} \mid \mathcal{F}_n)) \geq g(E(X_{n+1} \mid \mathcal{F}_n))$$

a) Por tratarse de una submartingala tenemos que $E(X_{n+1} \mid \mathcal{F}_n) \geq X_n$; por tanto $g(E(X_{n+1} \mid \mathcal{F}_n)) \geq g(X_n)$ pues g es creciente.

b) $E(X_{n+1}|\mathcal{F}_n) = X_n$ por las propiedades de martingala, entonces $g(E(X_{n+1} \mid \mathcal{F}_n)) = g(X_n)$. □

Estudiemos ahora algunas desigualdades fundamentales para martingalas, que generalizan la clásica desigualdad de Chebyshev y de Kolmogorov.

Teorema: (Desigualdad Maximal) **3.10**
Sea $\{X_k, \mathcal{F}_k, 1 \leq k \leq n\}$ una submartingala. Entonces para cada $\lambda \in \mathbb{R}$,

$$\lambda P\{\omega\colon \max_{1\leq k\leq n} X_k(\omega) \geq \lambda\} \leq \int_{[\max_{1\leq k\leq n} X_k \geq \lambda]} X_n \, dP \leq E(|X_n|), \qquad \text{...(1)}$$

y

$$\lambda P\left\{\omega\colon \min_{1\leq k\leq n} X_k(\omega) \leq \lambda\right\} \geq \int_{[\min_{1\leq k\leq n} X_k \leq \lambda]} X_n \, dP - E(X_n - X_1) \geq E(X_1) - E(|X_n|) \quad \text{...(2)}$$

Demostración:

Sea $M = \left\{\omega\colon \max_{1\leq k\leq n} X_k(\omega) \geq \lambda\right\}$.
Primero vamos a expresar el evento M como una unión finita y disjunta de conjuntos.
Sea $M_1 = \left\{\omega\colon X_1(\omega) \geq \lambda\right\}$ y para $k > 1$ definamos M_k como

$$M_k = \left\{\omega\colon X_k(\omega) \geq \lambda, \quad X_j(\omega) < \lambda, \quad 1 \leq j \leq k-1\right\},$$

i.e. M_k es el conjunto tal que X_k es la primera v.a. en ésta sucesión que excede a λ.
Claramente $M_k \in \mathcal{F}_k$ y $M = \cup_{k=1}^{n} M_k$, es una unión disjunta.
Por lo tanto

$$E(|X_n|) \geq \int_M |X_n| \, dP \geq \int_M X_n \, dP = \sum_{k=1}^{n} \int_{M_k} X_n \, dP \geq \sum_{k=1}^{n} \int_{M_k} X_k \, dP,$$

pues $\{X_k\}_1^n$ es una submartingala,

$$\geq \lambda \sum_{k=1}^{n} P(M_k) = \lambda P\left(\cup_{k=1}^{n} M_k\right) = \lambda P(M) \text{ y esto prueba (1).}$$

Para probar (2) vamos a hacer una descomposición parecida a la de (1).

$$\text{Sea} \quad N = \{\omega: \min_{1\leq k\leq n} X_k(\omega) \leq \lambda\}$$

Igual que antes sea $N_1 = \{\omega: X_1(\omega) \leq \lambda\}$ y N_k es el conjunto donde X_k es la primera v.a. que no excede a λ, así para $k > 1$

$$N_k = \{\omega: X_k(\omega) \leq \lambda, X_j(\omega) > \lambda, 1 \leq j \leq k-1\}.$$

Claramente $N = \cup_{k=1}^{n} N_k$, $N_k \in \mathcal{F}_k$ disjuntos y $N_k \subset \left(\cup_{i=1}^{k-1} N_i\right)^c$

Por tanto

$$E(X_1) = \int_{N_1} X_1\, dP + \int_{N_1^c} X_1\, dP \leq \lambda \int_{N_1} dP + \int_{N_1^c} X_2\, dP\,,$$

pues $N_1^c \in \mathcal{F}_1$ y $\{X_1, X_2\}$ es una submartingala,

$$= \lambda P(N_1) + \int_{N_2} X_2\, dP + \int_{N_1^c \cap N_2^c} X_2\, dP$$

$$\leq \lambda\left[P(N_1) + P(N_2)\right] + \int_{N_1^c \cap N_2^c} X_3\, dP, \text{ pues } N_1^c \cap N_2^c \in \mathcal{F}_2,$$

$$\vdots$$

$$\leq \lambda \sum_{i=1}^{n} P(N_i) + \int_{\bigcap_{i=1}^{n} N_i^c} X_n\, dP$$

$$= \lambda P(N) + \int_{\Omega} X_n\, dP - \int_{N} X_n\, dP$$

$$\text{por lo tanto } \lambda P(N) \geq \int_{N} X_n\, dP - \int_{\Omega} (X_n - X_1)\, dP\,. \text{ y esto prueba (2)}$$

$\square$

Teorema: (Desigualdad de Kolmogorov) 3.11

Sean $X_1, \ldots, X_n$ v.a.i. tales que $\alpha_i = E(X_i)$ y $\sigma_i^2 = Var\, X_i = E(X_i - \alpha_i)^2$ existen. Entonces para cualquier $\varepsilon > 0$,

$$P\left[\max_{1\le k\le n} \left| \sum_{i=1}^{k} (x_i - \alpha_i) \right| \ge \varepsilon\right] \le \frac{1}{\varepsilon^2}\sum_{i=1}^{n} \sigma_i^2 \qquad \ldots(3)$$

Demostración:

Sean $\Psi_i = X_i - \alpha_i$; $S_k = \sum\limits_{i=1}^{k} \Psi_i$ y $\sigma_k = \sigma\{X_i, 1 \le i \le k\}$.

Observemos que : $\{S_k, \mathcal{F}_k, 1 \le k \le n\}$ es una martingala.

Demostración:

$E(S_{k+1} \mid \mathcal{F}_k) = E(S_k + \Psi_{k+1} \mid \mathcal{F}_k) = S_k + E(\Psi_{k+1} | \mathcal{F}_k) = S_k$ c.s.

pues como Ψ_{k+1} es independiente de $\mathcal{F}_k$, $E(\Psi_{k+1}|\mathcal{F}_k) = E(\Psi_{k+1}) = 0$, y S_k es $\mathcal{F}_k$-medible.

Por el teorema 3.9(a) tenemos que : $\{S_k^2, \mathcal{F}_k, 1 \le k \le n\}$ es una submartingala pues $E(S_k^2) = \sum_{i=1}^{k} \sigma_i^2 < \infty$.

Entonces por (1) $\varepsilon^2 P\left[\max\limits_{1\le k\le n} S_k^2 \ge \varepsilon^2\right] \le E(S_n^2) = \sum_{i=1}^{n} \sigma_i^2$, que es (3)

Ahora que si consideramos $n = 1$, entonces (3) se convierte en:

$P[|X_1 - \alpha_1| \ge \varepsilon] \le \frac{\sigma_1^2}{\varepsilon^2}$ que es la desigualdad de Chebysev. □

Definición: 3.12

Si $\sup\limits_{n\ge 1} E|X_n| < \infty$ decimos que la sucesión $(X_n, n \ge 1)$ es $\mathcal{L}_1$-acotada, mientras que si $E\left[\sup\limits_{n\ge 1}|X_n|\right] < \infty$ decimos que la sucesión $(X_n, n \ge 1)$ es $\mathcal{L}_1$- dominada.

Observación: 3.13

Sea $X = (X_n, \mathcal{F}_n, n \ge 1)$ una martingala. La relación $\sup\limits_{n\ge 1} E|X_n| \le E\left[\sup\limits_{n\ge 1}|X_n|\right]$ implica que toda martingala $\mathcal{L}_1$-dominada es también $\mathcal{L}_1$-acotada. Sin embargo la relación inversa no necesariamente es cierta, para probarlo daremos unos ejemplos.

Ejemplo:

Consideremos el espacio discreto $\Omega = \{1, 2, \ldots\}$ y definámosle la siguiente probabilidad P como $P(\{n\}) = \frac{1}{n} - \frac{1}{n+1}$, $n \in \mathbb{N}$. Sea $(\mathcal{F}_n, n \ge 1)$ la sucesión creciente de σ-álgebras donde $\mathcal{F}_n$ es generada por las particiones $\{\{1\},\{2\},\ldots,\{n\},[n+1,\infty)\}$. Definamos la sucesión $(X_n, n \geq 1)$ de v.a. por

$$X_n = X_n(\omega) = n 1_{[n,\infty)}(\omega), n \in \mathbb{N}.$$

Entonces $X = (X_n, \mathcal{F}_n, n \geq 1)$ es una martingala positiva tal que $E(X_n) = 1 \quad \forall\, n \in \mathbb{N}$ y por lo tanto X es $\mathcal{L}_1$-acotada. Sin embargo, $\sup_{n\in\mathbb{N}} X_n(\omega) = \omega$ que claramente no es integrable.
Por lo tanto la martingala X no es $\mathcal{L}_1$-dominada.

Ejemplo:
Sea $\Omega = [0,1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$ y P la medida de Lebesgue.
Definimos

$$X_n = X_n(\omega) = \begin{cases} 0 & \text{si} \quad \frac{1}{n} \leq \omega \leq 1 \\ -n^2\omega + n & \text{si} \quad 0 \leq \omega < \frac{1}{n} \end{cases}$$

y $\mathcal{F}_n = \sigma\{X_1, \ldots, X_n\}$. Entonces $(X_n, \mathcal{F}_n, n \geq 1)$ es una martingala. Como $E(X_n) = \frac{1}{2}$ para cada $n \in \mathbb{N}$, ésta martingala es $\mathcal{L}_1$-acotada; sin embargo su supremo, $\sup_{n\in\mathbb{N}} |X_n|$ no es integrable y por lo tanto no es $\mathcal{L}_1$-dominada.

Estudiemos ahora algunos teoremas de convergencia para las martingalas.

Teorema del salto óptimo (Halmos) 3.14
Sea $\{X_n, \mathcal{F}_n\}$ una submartingala.

Sean $\varepsilon_1, \varepsilon_2, \ldots$ v.a. definidas por

$$\varepsilon_k = \begin{cases} 1 & \text{si} \quad (X_1, \ldots, X_k) \in B_k \\ 0 & \text{si} \quad (X_1, \ldots, X_k) \notin B_k \end{cases}$$

donde los B_k son conjuntos arbitrarios en $\mathcal{B}(\mathbb{R}^n)$.
Consideremos

$$\begin{aligned} \Psi_1 &= X_1, \\ \Psi_2 &= X_1 + \varepsilon(X_2 - X_1) \\ &\vdots \\ \Psi_n &= X_1 + \varepsilon(X_2 - X_1) + \ldots + \varepsilon_{n-1}(X_n - X_{n-1}). \end{aligned}$$

Entonces $\{\Psi_n, \mathcal{F}_n\}$ es también una submartingala y $E(\Psi_n) \leq E(X_n)\ \forall n$. Si $\{X_n, \mathcal{F}_n\}$ es una martingala, entonces $\{\Psi_n, \mathcal{F}_n\}$ también lo es y $E(\Psi_n) = E(X_n)\ \forall n$.

Una manera de interpretar este teorema es volviendo a la idea del jugador:
Sea X_n la ganancia del jugador después de n juegos; entonces Ψ_n es nuestra ganancia si seguimos una estrategia de salto opcional. Después de observar $X_1,\ldots,X_k$, podemos escoger si apostamos con el jugador en el juego $k+1$ (en este caso $\varepsilon_k=\varepsilon_k(X_1,\ldots,X_k)=1$) o pasamos ($\varepsilon_k=0$). Nuestra ganancia en el juego $k+1$ es $\varepsilon_k(X_{k+1}-X_k)$. Este teorema lo que nos dice es que no importa que estrategia sigamos si el juego es justo (una martingala) o favorable (una submartingala), así se queda, es decir justo o favorable según el caso; y ninguna estrategia de este tipo puede incrementar la ganancia esperada.

Demostración:

Tenemos que:

$$\begin{aligned} E(\Psi_{n+1}|\mathcal{F}_n) &= E(\Psi_n+\varepsilon(X_{n+1}-X_n)|\mathcal{F}_n) \\ &= \Psi_n+\varepsilon_n E\big((x_{n+1}-X_n)\mid\mathcal{F}_n\big) \end{aligned}$$

por ser ε_n una función Borel-medible de $X_1,\ldots,X_n$, y por lo tanto es

$$\mathcal{F}(X_1,\ldots,X_n)\subset\mathcal{F}_n\text{ –medible.}$$

$\Rightarrow$

$$\begin{aligned} E(\Psi_{n+1}|\mathcal{F}_n) &= \Psi_n+\varepsilon_n(X_n-X_n)=\Psi_n \text{ en el caso de las martingalas} \\ &\geq \Psi_n+\varepsilon_n(X_n-X_n)=\Psi_n \text{ en el caso de las submartingalas.} \end{aligned}$$

Como $\Psi_1=X_1$, tenemos $E(X_1)=E(\Psi_1)$, y $E(X_k-\Psi_k)\geq 0$ ($=0$ en el caso de martingalas),

$$X_{k+1}-\Psi_{k+1}=X_{k+1}-\Psi_k-\varepsilon_k(X_{k+1}-X_k)=(1-\varepsilon_k)(X_{k+1}-X_k)+X_k-\Psi_k$$

$$\Rightarrow\ E(X_{k+1}-\Psi_{k+1}|\mathcal{F}_k)=(1-\varepsilon_k)E(X_{k+1}-X_k|\mathcal{F}_k)+E(X_k-\Psi_k|\mathcal{F}_k)$$

$$\geq E(X_k-\Psi_k)|\mathcal{F}_k)=X_k-\Psi_k\ ,$$

con igualdad en el caso de martingalas.
Tomando la esperanza en ambos lados y usando la propiedad $E\big(E(X|\mathcal{G})\big)=E(X)$ obtenemos:

$$E(X_{k+1}-\Psi_{k+1})\geq E(X_k-\Psi_k)\geq 0\ ,$$

nuevamente con igualdad en el caso de las martingalas. □

Llegamos ahora a una nueva desigualdad que va a ser una herramienta importante en la prueba de algunos teoremas de convergencia. Se trata de un teorema debido a Doob, que consiste de la siguiente idea: Dada una sucesión de v.a. $\{X_n\}$, para cada punto $\omega \in \Omega$, las propiedades de convergencia de una sucesión numérica $\{X_n(\omega)\}$ dependen de la oscilación de segmentos finitos $\{X_n(\omega), n \in N_m\}$ cuando $m \to \infty$. En particular la sucesión tendrá un límite, finito o infinito, si y sólo si el número de sus oscilaciones entre dos números (reales) $a < b$ es finito (dependiendo de a, b y ω). Lo importante aquí, es que es que para una submartingala puede ser obtenida una medida precisa del número esperado de oscilaciones.

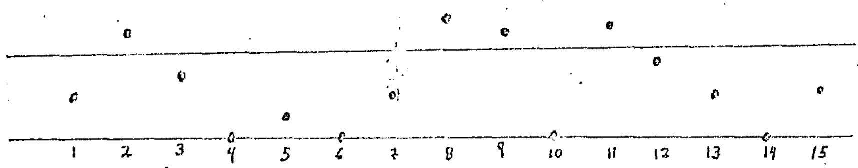
Teorema "de los cruzamientos hacia arriba" 3.15
Sea $\{X_k, \mathcal{F}_k, k = 1, 2, \ldots, n\}$ una submartingala. Si $a, b \in \mathbb{R}$, con $a < b$, sea U_{ab} el número de cruzamientos de (a, b) por $X_1, \ldots, X_n$, definido de la siguiente manera:

Sea $T_1 = T_1(\omega)$ el primer entero en $\{1, 2, \ldots, n\}$ tal que $X_{T_1} \leq a$, T_2 el primer entero mayor que T_1 tal que $X_{T_2} \geq b$, T_3 el primer entero después de T_2 tal que $X_{T_3} \leq a$, etc.
(Decimos que $T_i = \infty$ si la condición no se satisface).
Si N es el número de los T_i finitos, difinimos $\mathsf{U}_{ab} = \frac{N}{2}$ si n es par y $\frac{N-1}{2}$ si n es impar.
Entonces

$$E(\mathsf{U}_{ab}) \leq \frac{1}{b-a} E\big((X_n - a)^+\big).$$

Demostración:

Supongamos que $a = 0$, y que toda $X_j \geq 0$. Definamos las T_i como dice el teorema. Sea $\varepsilon_j = 0$ para $j < T_1$; $\varepsilon_j = 1$ para $T_1 \leq j < T_2$, $\varepsilon_j = 0$ para $T_2 \leq j < T_3$; $\varepsilon_j = 1$ para $T_3 \leq j < T_4$ etc.
Pongamos un ejemplo:



Aquí tendremos que $n = 15$, $T_1 = 4$, $T_2 = 8$, $T_3 = 10$, $T_4 = 11$, $T_5 = 14$, $T_n = \infty$, $n > 5$, $\mathsf{U}_{ab} = 2$;
$\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 0$; $\varepsilon_4, \varepsilon_5 = \varepsilon_6 = \varepsilon_7 = 1$.
$E_8 = \varepsilon_9 = 0$; $\varepsilon_{10} = 1$, $\varepsilon_{11} = \varepsilon_{12} = \varepsilon_{13} = 0$; $\varepsilon_{14} = 1$.
$X_1 + \varepsilon_1(X_2 - X_1) + \ldots + \varepsilon_{14}(X_{15} - X_{14}) = X_1 + X_8 - X_4 + X_{11} - X_{10} + X_{15} - X_{14}$.

Notemos que Ψ_n definida como en el teorema del "salto óptimo" es el incremento total durante los cruzamientos, mas un posible cruzamiento parcial al final, mas una contribución debida a X_1.

Entonces $\Psi_n \geq b\cup_{ab}$. Pero las ε_j pueden expresarse en términos de $X_1, \ldots, X_j$, por tanto podemos aplicar el teorema anterior;

por lo tanto $\{\Psi_k, \mathcal{F}_k, k = 1, 2, \ldots, n\}$ es una submartingala, y $E(\Psi_n) \leq E(X_n)$ y por tanto $E(\cup_{ab}) \leq \frac{1}{b}E(\Psi_n) \leq \frac{1}{b}E(X_n)$ se cumple.

Para el caso general, consideremos $\{(X_k - a)^+, \mathcal{F}_k, k = 1, 2, \ldots, n\}$ que es una submartingala y el número de cruzamientos de (a, b) por $\{X_j\}$ es el mismo que el número de cruzamientos de $(0, b - a)$ por $\{(X_j - a)^+\}$. Como $X_j \leq a$, $X_j - a \leq 0$ y $(X_j - a)^+ \leq 0$ son equivalentes a $X_j \geq b$, $X_j - a \geq b - a$ y $(X_j - a)^+ \geq b - a$, el resultado se concluye del resultado anterior. □

Teorema de convergencia para submartingalas. 3.16

Sea $\{X_n, \mathcal{F}_n, n = 1, 2, \ldots\}$ una submartingala. Si $\sup_n E(X_n^+) < \infty$, existe entonces una v.a. X_∞ tal que $X_n \longrightarrow X_\infty$ c.s., y $E(|X_\infty|) \leq \varliminf_{n\to\infty} E(|X_n|) < \infty$.

Demostración:

Consideremos

$P\{\omega: X_n(\omega)$ no converge a un límite, ya sea finito o infinito$\}$

$= P\left(\bigcup_{\substack{a<b \\ a,b \text{ racional}}} \left\{\omega: \varliminf_{n\to\infty} X_n(\omega) < a < b < \varlimsup_{n\to\infty} X_n(\omega)\right\}\right)$

Si para algunas $a < b$, $P\left\{\varliminf_{n\to\infty} X_n < a < b < \varlimsup_{n\to\infty} X_n\right\} > 0$,

entonces $\{X_n\}$ tiene un número infinito de cruzamientos de (a, b) en un conjunto con probabilidad positiva, y por tanto $E(\cup_{ab}) = \infty$. Pero $\cup_{ab}$ es el límite de la sucesión monótona $\cup_{ab;n}$ = número de cruzamientos de (a, b) por $X_1, X_2, \ldots, X_n$, asi que $E(\cup_{ab;n}) \longrightarrow E(\cup_{ab})$. Pero por el teorema anterior

$$E(\cup_{ab;n}) \leq \frac{1}{b-a} E\left[(X_n - a)^+\right] \leq \frac{1}{b-a}\left(\sup_n E(X_n^+) + a^-\right) < \infty$$

lo cual es una contradicción.

Por tanto X_n converge a un límite X_∞ c.s.

Ahora $|X_n| = X_n^+ + X_n^- = 2X_n^+ - X_n$, y $E(X_n) \geq E(X_1)$ por ser submartingala. Por lo tanto $E(|X_n|) \leq 2\sup_n E(X_n^+) - E(X_1) < \infty$. Y por el lema de Fatou

$$E(|X_\infty|) \leq \varliminf_{n\to\infty} E(|X_n|) < \infty$$

y por ello X_∞ es integrable y por tanto finita c.d.
Podemos considerar a X_∞ como una v.a. aunque probablemente haya que "cambiarla" en un conjunto de medida cero. $\square$

Corolario 3.17

Sea $\{X_n, \mathcal{F}_n, n = 1, 2, \ldots\}$ una submartingala inversa, *i.e.* $\mathcal{F}_n$ decrece cuando n crece y $E(X_n|\mathcal{F}_{n+1}) \geq X_{n+1}$ c.d. Si $\inf_n E(X_n) > -\infty$, entonces existe una v.a. X_∞ tal que $X_n \longrightarrow X_\infty$ c.d.

Demostración:

Esta prueba se lleva a cabo casi como la anterior, sólo que en éste caso hay que considerar a $\mathsf{U}_{ab,n}$ como el número de cruzamientos de (a, b) por $\{X_n, X_{n-1}, \ldots, X_1\}$ la cual es una submartingala pues:

$$E(X_k|X_{k+1}, \ldots, X_n) = E\big(E(X_k|\mathcal{F}_{k+1}) \mid X_{k+1}, \ldots, X_n\big) \geq X_{k+1}$$

Podemos entonces concluir $E(\mathsf{U}_{ab,n}) \leq \frac{1}{b-a} E\big((X_1 - a)^+\big) < \infty$,
y por lo tanto $X_n \longrightarrow X_\infty$ c.d.

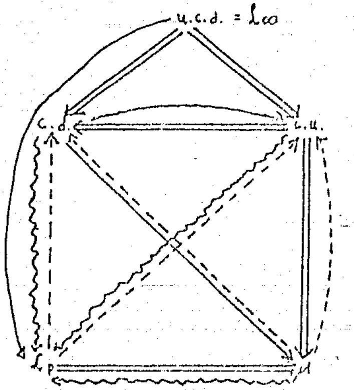
Ahora $|X_n| = 2X_n^+ - X_n$ y $E(X_n) \geq \inf_n E(X_n) > -\infty$.

Por tanto $\{X_n^+, \ldots, X_1^+\}$ es una submartingala y así $E(X_n^+) < E(X_1^+)$
$\Rightarrow E(|X_n|) \leq 2E(X_1^+) - \inf_n E(X_n) < \infty$, y nuevamente usando el lema de Fatou concluimos que X_∞ es integrable. $\square$

Observación:

Los dos resultados anteriores también pueden probarse para supermartingalas.

Teoremas de Convergencia



$\Rightarrow$ implicación sin restricciones
$\longrightarrow$ $\mu(X) < +\infty$
$- - - \to \exists$ subsucesión (f_{n_k}) de (f_n) que converge en el sentido que marca la flecha
$\sim\sim\sim\to \exists\, g \in \mathcal{L}_p(\mu)$ tal que $|f_n| < g \ \forall n \in N$.

Definición. Una sucesión (f_n) de funciones en $D \subseteq R^p$ a R^q converge a una función f en $D_0 \subseteq I$
$\Leftrightarrow \forall \varepsilon > 0$ y $\forall x \in D_0 \ \exists\, k(\varepsilon, x)$ tal que $\forall n \geq k(\varepsilon, x) \quad \|f_n(x) - f(x)\| < \varepsilon$.

1) $f_n \xrightarrow{c.d.} f \Leftrightarrow \exists\, N \in S$ tal que $\mu(N) = 0$ tal que $f_n(x) \longrightarrow f(x) \ \forall x \notin N$.

Definición. Una sucesión (f_n) de funciones en $D \subseteq R^p$ a R^q converge uniformemente en un subconjunto D_0 de D a una función f cuando $\forall\ \varepsilon > 0 \ \exists\, k(\varepsilon)$ tal que $\forall\ n \geq k(\varepsilon)$ y $x \in D_0$ $\|f_n(x) - f\| < \varepsilon$.

2) $f_n \underset{u.c.d.}{\xrightarrow{\mathcal{L}_\infty}} f \Leftrightarrow \exists\, N \in S$ tal que $\big(f_n(x)\big)$ converge uniformemente $\forall\, x \in X - N$, $\mu(N) = 0$.

3) $f_n \xrightarrow{\mathcal{L}_p} f (f_n, f \in \mathcal{L}_p(\mu)) \Leftrightarrow \|f_n - f\|_p \longrightarrow 0 \quad (n \longrightarrow \infty)$

4) $f_n \xrightarrow{\mu} f \Leftrightarrow \forall \varepsilon > 0 \ \exists\, N$ tal que $\forall\, n \geq N$ tenemos que:
$\mu\underbrace{\{x \in x : |f_n(x) - f(x)| \geq \varepsilon\}}_{A_n(\varepsilon)} < \varepsilon. \qquad i.e.\ \mu(A_n(\varepsilon)) \longrightarrow 0 \quad (n \to \infty) \quad \forall\, \varepsilon > 0$ fija.

5) $f_n \xrightarrow{c.u.} f \Leftrightarrow \forall\, \varepsilon > 0 \quad \exists\, \delta = \delta(\varepsilon) > 0$ y un conjunto medible E en S con $\mu(E) < \delta$ tal que $|f_n(x) - f(x)| < \varepsilon$ si $n \geq N(\varepsilon, E)$, $x \in X - E$

Integrabilidad Uniforme

Definición 3.18
Sean $f_1, f_2, \ldots$ funciones (reales o complejas) Borel medibles en el espacio de medida $(\Omega, \mathcal{F}, \mu)$, μ finita.
Se dice que f_n son uniformemente integrables si y sólo si

$$\int_{\{|f_n| \geq c\}} |f_n|\, d\mu \longrightarrow 0 \text{ cuando } c \to \infty$$

uniformemente en n.

Observación:
Esta es la definición más común que hay para el concepto de integrabilidad uniforme, sin embargo hay otra definición que deseamos discutir:

Definición 3.19
Sea $\mathcal{H}$ un subconjunto del espacio $\mathcal{L}_1(\Omega, \mathcal{F}, P)$. Se dice que $\mathcal{H}$ es un conjunto uniformemente integrable de v.a., si las integrales

$$\int_{\{|f| \geq c\}} |f(\omega)|\, dP(\omega) \quad f \in \mathcal{H}$$

tienden uniformemente a cero cuando $c \to +\infty$.

Observación:
Evidentemente se trata del mismo concepto, salvo que en la definición 2 hablamos de v.a. en $\mathcal{L}_1(\Omega, \mathcal{F}, P)$ en vez de funciones Borel medibles en $(\Omega, \mathcal{F}, \mu)$ como en la definición 1.
La razón por la que se mencionan ambas definiciones es para poder explicar mejor algunos resultados de integrabilidad uniforme.

Observaciones: 3.20
a) Es inmediato que si las f_n son uniformemente integrables, cada f_n es integrable.

b) Si $|f_n| \leq g \ \forall\ n$, donde g es integrable, en particular, si las f_n son uniformemente acotadas, entonces las f_n son uniformemente integrables.

c) Si las f_n son uniformemente integrables, entonces $\sup_n \int_\Omega |f_n|\, d\mu < \infty$

Sea $\varepsilon > 0$, $\int_\Omega |f_n|\, d\mu = \int_{\{|f_n| \geq c\}} |f_n|\, d\mu + \int_{\{|f_n| < c\}} |f_n|\, d\mu \leq \varepsilon + c\mu(\Omega)$ para n suficientemente grande.

Teorema: 3.21
Sean $f_1, f_2, \ldots$ real-valuadas y uniformemente integrables

a)

$$\int_\Omega \left(\underline{\lim}_n f_n\right) d\mu \leq \underline{\lim}_n \int_\Omega f_n\, d\mu \leq \overline{\lim}_n \int_\Omega f_n\, d\mu \leq \int (\overline{\lim}_n f_n)\, d\mu$$

[Esto es una extensión del lema de Fatou]

b) Si $f_n \longrightarrow f$ c.d., o en medida, entonces f es integrable y $\int_\Omega f_n\, d\mu \longrightarrow \int_\Omega f\, d\mu$

[Esto es una extensión del teorema de Lebesgue]

Demostración:

a) Sabemos que:

$$\int_\Omega f_n\, d\mu = \int_{\{f_n < -c\}} f_n\, d\mu + \int_{\{f_n \geq -c\}} f_n\, d\mu \quad c > 0$$

Por integrabilidad uniforme, podemos tomar c lo suficientemente grande de tal manera que:

$$\left|\int_{\{f_n < -c\}} f_n\, d\mu\right| < \varepsilon \quad \forall\, n, \text{ donde } \quad \varepsilon > 0 \text{ es dada.}$$

Como $f_n I_{\{f_n \geq -c\}} \geq -c$, que es integrable, pues μ es finita, podemos aplicar el lema de Fatou:

$$\underline{\lim}_n \int_{\{f_n \geq -c\}} f_n\, d\mu \geq \int_\Omega \underline{\lim}(f_n I_{\{f_n \geq -c\}})\, d\mu$$

Como $f_n I_{\{f_n \geq -c\}}) \geq f_n$, esta integral es mayor o igual a $\int_\Omega \left(\underline{\lim}_n f_n\right) d\mu$

$$\Rightarrow \underline{\lim}_n \int_\Omega f_n\, d\mu \geq \int_\Omega (\underline{\lim}\, f_n)\, d\mu - \varepsilon.$$

Esto prueba el teorema para la parte de $\underline{\lim}$, la parte de $\overline{\lim}$ se hace con un argumento análogo.

b) Si $f_n \longrightarrow f$ c.d. el resultado se sigue de (a).
Consideremos entonces el caso en que $f_n \longrightarrow f$ en medida, esiste entonces una subsucesión $\{f_{n_k}\}$ de $\{f_n\}$ tal que $f_{n_k} \longrightarrow f$ c.d. (ver cuadro de teoremas de convergencia); aplicamos entonces el inciso (a) a f_{n_k}, y entonces f es integrable; tenemos: $\int_\Omega f_{n_k}\, d\mu \longrightarrow \int_\Omega f\, d\mu$

Si $\int_\Omega f_n\, d\mu$ no converge a $\int_\Omega f\, d\mu$, entonces dada $\varepsilon > 0$ tenemos que

$\left|\int_\Omega f_n\, d\mu - \int_\Omega f\, d\mu\right| \geq \varepsilon$, que podemos pensar se cumple para toda n.
Pero entonces podemos encontrar una subsucesión $f_{m_j} \to f$ c.d. y como antes,

$$\int_\Omega f_{m_j}\, d\mu \longrightarrow \int_\Omega f\, d\mu \quad ! \quad \text{lo cual no es posible.}$$

$\square$

Teorema: 3.22
Las funciones (complejo-valuadas), medibles son uniformemente integrables si y sólo si las integrales

a) $\int_\Omega |f_n|\, d\mu$ son uniformemente acotadas y

b) uniformemente contínuas, esto es, $\int_A |f_n|\, d\mu \to 0$ cuando $\mu(A) \to 0$ uniformemente en n.

Observación:
Si nos concentramos en el espacio $\mathcal{L}_1(\Omega, \mathcal{F}, P)$, y consideramos las f_n como v.a. entonces el lado derecho del $\Leftrightarrow$ se puede reescribir de la
Siguiente manera:
Las esperanzas $E(|f_n|)$, $\forall\, n$, son uniformemente acotadas y uniformemente contínuas.

Demostración:

$\Leftarrow$) Como las integrales son uniformemente acotadas y uniformemente contínuas, entonces por la desigualdad de Chebyshev tenemos:

$$\mu\{|f_n| \geq c\} \leq \frac{1}{c}\int_\Omega |f_n|\, d\mu \to 0 \text{ cuando } c \longrightarrow \infty$$

uniformemente en n, esto es por el hecho de que son uniformemente acotadas.
Entonces $\int\limits_{\{|f_n|\geq c\}} |f_n|\, d\mu \to 0$ cuando $c \longrightarrow \infty$ uniformemente en n, esto por la continuidad uniforme.

$\Rightarrow$) Suponemos ahora la integrabilidad uniforme:

$$\int\limits_A |f_n|\, d\mu = \int\limits_{A\cap\{|f_n|\geq c\}} |f_n|\, d\mu + \int\limits_{A\cap\{|f_n|<c\}} |f_n|\, d\mu \leq \int\limits_{\{|f_n|\geq c\}} |f_n|\, d\mu + c\mu(A) \quad ...(*)$$

Sea c tal que:

$$\int\limits_{\{|f_n|\geq c\}} |f_n|\, d\mu < \frac{\varepsilon}{2} \quad \forall\, n; \quad \text{si } \mu(A) < \frac{\varepsilon}{2c} \Rightarrow \text{ por } (*)$$

$\int\limits_A |f_n|\, d\mu < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \;\forall\, n$, lo que prueba la continuidad uniforme.
Ya con anterioridad, en las observaciones, se probó uniformemente acotada. □

Ya hemos visto que convergencia en $\mathcal{L}_p \Rightarrow$ convergencia en medida (ver cuadro).
Se puede dar un teorema inverso considerando la hipótesis de integrabilidad uniforme.

Teorema: 3.23
Sea μ una medida finita en $(\Omega, \mathcal{F})$, y sea $0 < p < \infty$. Si $f_n \xrightarrow{\mu} f$ y $|f_n|^p$ son uniformemente integrables, entonces $f_n \xrightarrow{\mathcal{L}_p} f$.

Demostración:

Primero suponemos que las $|f_n - f|^p$ son uniformemente integrables. Por un resultado de análisis (ver cuadro) existe una subsucesión $\{f_{n_k}\}$ que converge a f c.d. y en medida. Por la extensión del teorema de convergencia dominada ya probado tenemos que:

$$\int\limits_\Omega |f_{n_k} - f|^p\, d\mu \to 0 \text{ cuando } k \longrightarrow \infty$$

El mismo argumento prueba que cualquier subsucesión de $\{f_n\}$ tiene una sub-sucesión que converge a f en $\mathcal{L}_p$. Por lo tanto $f_n \xrightarrow{\mathcal{L}_p} f$, si no, existiría $\varepsilon > 0$ y una subsucesión $\{f_{n_i}\}$ tal que $\int\limits_\Omega |f_{n_i} - f|^p\, d\mu \geq \varepsilon \;\forall\, i$

Ahora supongamos que las $|f_n|^p$ son uniformemente integrables. Entonces

$$|f_n - f|^p \le |f_n|^p + |f|^p \text{ si } p \le 1,$$
$$\text{y } |f_n - f|^p \le 2^{p-1}\left(|f_n|^p + |f|^p\right) \text{ si } p \ge 1$$

Como antes, tenemos una subsucesión $f_{n_k} \longrightarrow f$ c.d. por $(*)$, $|f|^p$ es integrable y se sigue entonces que las integrales $\int\limits_{\Omega} |f_n - f|^p \, d\mu$ son uniformemente acotadas y uniformemente contínuas. Por el teorema anterior tenemos entonces, que las $|f_n - f|^p$ son uniformemente integrables. □

Corolario: 3.24

En 3.21 inciso (b), $f_n \xrightarrow{\mathcal{L}_1} f$, esto es, $\int\limits_{\Omega} |f_n - f| \, d\mu \to 0$

Demostración:

Las $|f_n|$ son uniformemente integrables por hipótesis y $f_n \xrightarrow{\mu} f$ por hipótesis o por el teorema de Egoroff (ver cuadro). □

El siguiente resultado es de La Vallée Poussin.

Teorema: 3.25

Sea $\mathcal{H}$ un subconjunto de $\mathcal{L}_1(\Omega, \mathcal{F}, P)$. Las siguientes propiedades son equivalentes:

(1) $\mathcal{H}$ es uniformemente integrable.

(2) Existe una función $G(t)$ definida en $\mathbb{R}_+$ que es positiva creciente y convexa, tal que:

$$(i) \quad \lim_{t \to +\infty} \frac{G(t)}{t} = +\infty$$

y

$$(ii) \sup_{f \in \mathcal{H}} E(G \circ |f|) < \infty$$

Demostración:

$2) \Rightarrow 1)$

Sea $\varepsilon > 0$, por (ii) podemos considerar

$$M = \sup_{f \in \mathcal{H}} E(G \circ |f|) \text{ y sea } a \doteq \frac{M}{\varepsilon}.$$

Por (i) podemos considerar a c tan grande como sea necesario tal que $\frac{G(t)}{t} \ge a$ para $t \ge c$. De aquí tenemos entonces que

$$\frac{G(|f|)}{|f|} \geq a \quad ó \quad \frac{G(|f|)}{a} \geq |f| \text{ en el conjunto } \{|f| \geq c\}$$

y por lo tanto

$$\int_{\{|f|\geq c\}} |f|\, dP \leq \frac{1}{a} \int_{\{|f|\geq c\}} (G \circ |f|)\, dP \leq \frac{1}{a} M = \varepsilon \quad \forall f \in \mathcal{H}$$

Por lo tanto hemos probado que $\mathcal{H}$ es un conjunto de v.a. uniformemente integrable pues de hecho se cumple la definición 2.

1)$\Rightarrow$ 2)
Tenemos ahora a $\mathcal{H} \in \mathcal{L}_1$, donde $\mathcal{H}$ es uniformemente integrable.

Construyamos ahora una función $G(t)$ de la forma $\int_0^t g(t)\, dt$, donde g es cualquier función creciente fija, que tiende a $+\infty$ cuando t lo hace y que tiene valores constantes g_n en cada intervalo $[n, n+1) \quad (n \in \mathbb{N})$.

$$g(t) = \sum \chi_{[n,n+1)} g_n(t) \quad g(t) \uparrow \infty \text{ si } t \longrightarrow \infty; \text{ si } t \in [n, n+1)\ g(t) = g_n(t) \text{ y } g_n \leq g_{n+1}.$$

$$\lim_{x\to\infty} \frac{1}{x} \int_0^x g(t)\, dt = \infty \quad i.e.\ (i) \text{ se cumple.}$$

Consideremos para cada función $f \in \mathcal{H}$, $a_n(f) = P\{|f| > n\}$ y $g_0 = 0$; tenemos entonces que:

$$E[G \circ |f|] \leq g_1 P\{1 < |f| \leq 2\} + (g_1 + g_2) P\{2 < |f| \leq 3\} + \ldots = \sum_{n=1}^{\infty} g_n a_n(f)$$

Como $\mathcal{H}$ es uniformemente integrable podemos encontrar una sucesión de enteros c_n que crece a $+\infty$ tal que

$$\sup_{f\in\mathcal{H}} \int_{\{|f|\geq c_n\}} |f|\, dP \leq 2^{-n}$$

Entonces

$$\int_{\{|f|\geq c_n\}} |f|\, dP \geq \sum_{m=c_n}^{\infty} m P\{m < |f| \leq m+1\} \geq \sum_{m=c_n}^{\infty} P\{|f| > m\} = \sum_{m=c_n}^{\infty} a_m(f)$$

Se sigue entonces que la suma $\sum_n \sum_{m=c_n}^{\infty} a_m(f)$ es **uniformemente acotada** para $f \in \mathcal{H}$, pero esta suma es de la forma $\sum_m g_m a_m(f)$ donde g_m denota el número de enteros n tales que $c_n \leq m$.
Por lo tanto $\sup_{f \in \mathcal{H}} E(G \circ |f|) < \infty$ □

A continuación enunciamos y demostramos algunos resultados que nos relacionan al concepto de Integrabilidad uniforme con la teoría de martingalas.

Lema 3.26
Sea Ψ una v.a. en $\mathcal{L}_1(\Omega, \mathcal{F}, P)$ y sean $\mathcal{G}_i$, $i \in I$, sub-σ-álgebras arbitrarias de $\mathcal{F}$. Entonces las v.a. $X_i = E(\Psi|\mathcal{G}_i)$, $i \in I$, son uniformemente integrables, esto es, $\int_{\{|X_i| \geq c\}} |X_i| \, dP \longrightarrow 0 \quad (c \longrightarrow \infty)$ uniformemente en i.

Demostración:

Por propiedades de la esperanza condicional tenemos que $|E(\Psi|\mathcal{G}_i)| \leq E(|\Psi| \mid \mathcal{G}_i) \Rightarrow \int_{\{|X_i| \geq c\}} |X_i| \, dP \leq \int_{\{|X_i| \geq c\}} E(|\Psi| \mid \mathcal{G}_i) \, dP = \int_{\{|X_i| \geq c\}} |\Psi| \, dP$ pues $\{|X_i| \geq c\} \in \mathcal{G}_i$.
Pero por la desigualdad de Chebyshev

$$P\{|X_i| \geq c\} \leq c^{-1} E(|X_i|) \leq c^{-1} E\big(E(|\Psi| \mid \mathcal{G}_i)\big) = c^{-1} E(|\Psi|) \longrightarrow 0 \text{ si } \quad c \longrightarrow \infty$$

uniformemente en i, por lo tanto las v.a. X_i son uniformemente integrables. □

El siguiente es un teorema debido a Lévy.

Teorema 3.27
Sea $\{\mathcal{F}_n\}$ una sucesión creciente de sub-σ-álgebras de $\mathcal{F}$, y sea $\mathcal{F}_\infty = \sigma\left(\cup_{n=1}^{\infty} \mathcal{F}_n\right)$ si Ψ es integrable y $X_n = E(\Psi|\mathcal{F}_n)$, $\quad n = 1, 2, \ldots$, entonces $X_n \to E(\Psi|\mathcal{F}_\infty)$ c.d. y en $\mathcal{L}_1$.

Demostración:

Si consideramos $\{X_n, \mathcal{F}_n\}$ tenemos que:

$$\begin{aligned} E(X_{n+1}|\mathcal{F}_n) &= E\big(E(\Psi|\mathcal{F}_{n+1})|\mathcal{F}_n\big) \\ &= E(\Psi|\mathcal{F}_n) \text{ pues } \mathcal{F}_n \subset \mathcal{F}_{n+1} \\ &= X_n \end{aligned}$$

por lo tanto se trata de una martingala, y por el lema anterior podemos afirmar que $\{X_n, \mathcal{F}_n\}$ son uniformemente integrables.

Ahora, por el teorema de convergencia de submartingalas, y como $E(|X_n|) \leq E(|\Psi|) < \infty$, podemos afirmar que X_n converge c.d. a una v.a. X_∞ integrable.
La convergencia en $\mathcal{L}_1$ se sigue del corolario 3.24.
Resta probar únicamente que $X_\infty = E(\Psi|\mathcal{F}_\infty)$ c.d.

Consideremos $A \in \mathcal{F}_n \Rightarrow$

$$\int_A \Psi \, dP = \int_A E(\Psi|\mathcal{F}_n) \, dP = \int_A X_n \, dP \longrightarrow \int_A X_\infty \, dP$$

por la convergencia en $\mathcal{L}_1$ ya probada.
Por tanto

$$\int_A \Psi \, dP = \int_A X_\infty \, dP \quad A \in \cup_{n=1}^{\infty} \mathcal{F}_n$$ y como consecuencia del teorema de las clases monótonas para todo $A \in \mathcal{F}_\infty$.
Como X_n es $\mathcal{F}_n \subset \mathcal{F}_\infty$–medible, X_∞ es $\mathcal{F}_\infty$– medible y por lo tanto $X_\infty = E(\Psi|\mathcal{F}_\infty)$ c.d. que no es más que una consecuencia de la definición de esperanza condicional. □

En el caso de que $\{\mathcal{F}_n\}$ sea una sucesión decreciente, obtenemos el siguiente resultado.

Teorema 3.28
Sea $\{\mathcal{F}_n\}$ una sucesión decreciente de sub-σ-álgebras de $\mathcal{F}$, y sea $\mathcal{F}_0 = \bigcap_{n=1}^{\infty} \mathcal{F}_n$. Si Ψ es integrable y $X_n = E(\Psi|\mathcal{F}_n)$, $n = 1, 2, \ldots$, entonces $X_n \to E(\Psi|\mathcal{F}_0)$ c.d. y en $\mathcal{L}_1$.

Demostración:

La demostración en este caso es muy similar a la del teorema anterior. Sin embargo, en vez del teorema de convergencia de submartingalas utilizamos el corolario de éste. Así de manera análoga concluímos que $X_n \longrightarrow X_\infty$ c.d. y en $\mathcal{L}_1$.
Nuevamente resta únicamente probar que $X_\infty = E(\Psi|\mathcal{F}_0)$ c.d.

Si $A \in \mathcal{F}_0 \subset \mathcal{F}_n$, entonces

$$\int_A \Psi \, dP = \int_A E(\Psi|\mathcal{F}_n) \, dP = \int_A X_n \, dP \longrightarrow \int_A X_\infty \, dP$$

Como antes X_n es $\mathcal{F}_n \subset \mathcal{F}_k$-medible para $n \geq k$, X_∞ es $\mathcal{F}_k$ medible para todo k. Por lo tanto X_n es $\mathcal{F}_0$-medible. □

Teorema 3.29
Sea $\{X_n, \mathcal{F}_n, n = 1, 2, \ldots\}$ una submartingala uniformemente integrable.
Entonces $\sup\limits_n E(X_n^+) < \infty$, y X_n converge a un límite X_∞ c.d. y en $\mathcal{L}_1$.
Más aún, si $\mathcal{F}_\infty$ es la σ-álgebra, generada por $\cup_{n=1}^{\infty} \mathcal{F}_n$, $\Rightarrow$
$\{X_n, \mathcal{F}_n, n = 1, 2, \ldots, \infty\}$ es una submartingala.

Demostración:
Por (3.22) $\sup\limits_n E(|X_n|) < \infty$,
Por (3.16) $X_n \longrightarrow X_\infty$ c.d.
Por (3.24) $X_n \xrightarrow{\mathcal{L}_1} X_\infty$

Falta probar la segunda afirmación:
Si $A \in \mathcal{F}_n$, y $k \geq n$, entonces por propiedades de martingala pag. 33

$$\int_A X_n \, dP \leq \int_A X_k \, dP$$

Ahora bien, si hacemos tender k a ∞, la convergencia en $\mathcal{L}_1$ se tiene, y

$$\int_A X_n \, dP \leq \int_A X_\infty \, dP$$

entonces por propiedades de martingala $X_n \leq E(X_\infty | \mathcal{F}_n)$,
y por lo tanto $\{X_n, \mathcal{F}_n, n = 1, \ldots, \infty\}$ es una submartingala.
La prueba de la última afirmación se hace igual que ésta, únicamente sustituyendo "$\leq$" por "$=$". □

Observación: 3.30
Si $\{X_n, \mathcal{F}_n, n = 1, 2, \ldots, \infty\}$ es una (sub ó super) martingala, donde $\mathcal{F}_\infty$ es cualquier σ-álgebra que incluye todos los $\mathcal{F}_n$, X_∞ se dice que es el "último elemento".

Definición 3.31
Decimos que una martingala (supermartingala) $(X_t, \mathcal{F}_t)_{t \in T}$ es cerrada por la derecha si $\exists$ $\mathcal{F}_\infty$-una σ álgebra tal que $\forall\, t$, $\mathcal{F}_t \subset \mathcal{F}_\infty \subset \mathcal{F}$ y una v.a. X_∞, $\mathcal{F}_\infty$-medible tal que $(X_t, \mathcal{F}_t)_{T \cup \{+\infty\}}$ es martingala (supermartingala).

Teorema 3.32

$\{X_n, \mathcal{F}_n, n = 1, 2, \ldots\}$ es una martingala uniformemente integrable $\Leftrightarrow$ existe una v.a. integrable Ψ tal que $X_n = E(\Psi|\mathcal{F}_n)$, $n = 1, 2, \ldots$, en este caso, $X_n \to E(\Psi|\mathcal{F}_\infty)$ c.d. y en $\mathcal{L}_1$, donde $\mathcal{F}_\infty$ es el σ–álgebra generado por $\cup_{n=1}^{\infty} \mathcal{F}_n$.

Observación:

Podrá notarse que este resultado es una especie de resumen de los resultados anteriores.

Demostración:

$\Rightarrow$ 1 y 2

$\Leftarrow$ 4 con $\Psi = X_\infty$ □

Teorema 3.33

Sea $\{X_n, \mathcal{F}_n, n = 1, 2, \ldots, \infty\}$ una submartingala no–negativa con un último elemento. Entonces X_n es uniformemente integrable.

Demostración:

Por hipótesis

$$\int_{\{X_n \geq c\}} X_n \, dP \leq \int_{\{X_n \geq c\}} X_\infty \, dP$$

y por la desigualdad de Chebishev

$$P\{X_n \geq c\} \leq \frac{E(X_n)}{c} \leq \frac{E(X_\infty)}{c} \longrightarrow 0 \quad \text{si } c \longrightarrow \infty \text{ uniformemente en } n$$

por lo tanto X_n es uniformemente integrable. □

Estudiemos ahora el problema de la convergencia en $\mathcal{L}_p$.

Lema 3.34

Sean X_i, $i \in I$ v.a. $\in \mathcal{L}_1(\Omega, \mathcal{F}, P)$.

Si $h: [0, \infty) \longrightarrow [0, \infty)$ es Borel medible, $\frac{h(t)}{t} \longrightarrow \infty$ si $t \to \infty$ y $\sup_{i \in I} E(h(|X_i|)) < \infty$, entonces las X_i son uniformemente integrables.

Demostración:

Es el último teorema que vimos en integración uniforme

Teorema 3.35
Sea $\{X_n, \mathcal{F}, n = 1, 2, \ldots\}$ una martingala, o una submartingala no- negativa con

$$E(|X_n|^p) \leq M < \infty \quad \forall\, n, \quad \text{donde} \quad P > 1.$$

Entonces $X_n \longrightarrow X_\infty$ c.d. y en $\mathcal{L}_p$.

Demostración:

Por el lema anterior X_n es uniformemente integrable, y por (3.29) podemos afirmar que $X_n \xrightarrow{c.d.} X_\infty$ y que X_∞ es el último elemento.
Ahora $\{|X_n|^p, \mathcal{F}_n, n = 1, 2, \ldots, \infty\}$ es una submartingala no-negativa, pues hacemos $g(X) = |X|^p$ y aplicamos (3.9), por el teorema 3.33 las $|X_n|^p$ son uniformemente integrables y por (3.23) $X_n \longrightarrow X_\infty$ en $\mathcal{L}_p$. □

Teoremas de Convergencia

Antes que nada probaremos algunos resultados de análisis que serán de interés independiente para la obtención de algunos resultados importantes.
Así mismo probaremos algunos teoremas de convergencia de sucesiones de variables aleatorias.

Lema: 4.1

Sea $A = [a_{ij}]$ una matriz infinita de números reales; asumamos que $a_{nj} \longrightarrow 0$ cuando $n \longrightarrow \infty$ para cada j fija, y que para algún número c no-negativo $\sum_{j=1}^{\infty} |a_{n_j}| \leq c \quad \forall\, n$. Si $\{X_n\}$ es una sucesión acotada de números reales, definamos

$$\Psi_n = \sum_{j=1}^{\infty} a_{n_j} X_j \,, \quad n = 1, 2, \ldots$$

entonces:

a) Si $X_n \longrightarrow 0$, entonces $\Psi_n \longrightarrow 0$

b) Si $\sum_{j=1}^{\infty} a_{n_j} \longrightarrow 1$ y $X_n \longrightarrow X$ (X real), entonces $\Psi_n \longrightarrow X$

Demostración:

a) Podemos escribir

$$|\Psi_n| \leq \sum_{j=1}^{N} |a_{n_j}|\,|X_j| + \sum_{j=N+1}^{\infty} |a_{n_j}|\,|X_j| \qquad \ldots(*)$$

Dada $\varepsilon > 0 \ni N$ tal que para todo $|X_j| \leq \varepsilon/c \quad \forall\, j > N$, por tanto el segundo término de $(*)$ es a lo más $c(\varepsilon/c) = \varepsilon$.
Como el primer término $\longrightarrow 0$ cuando $n \longrightarrow \infty \quad \forall\, N$, se sigue que $\Psi_n \longrightarrow 0$.

$$\text{Por (a)}, \sum_{j=1}^{\infty} a_{n_j} (X_j - X) \longrightarrow 0 \text{ y por lo tanto } \Psi_n \longrightarrow X$$

□

El lema de Toeplitz 4.2
Sea $\{a_n\}$ una sucesión de números reales no-negativos y sea $b_n = \sum_{j=1}^{n} a_j$; supongamos que $b_n > 0 \ \forall n$, y $b_n \longrightarrow \infty$ cuando $n \longrightarrow \infty$. Si $\{X_n\}$ es una sucesión de números reales que converge al número real X, entonces

$$\frac{1}{b_n} \sum_{j=1}^{n} a_j X_j \longrightarrow X$$

Demostración:
Podemos formar una matriz infinita A cuyo n–ésimo renglón sea

$$\left(\frac{a_1}{b_n}, \frac{a_2}{b_n}, \frac{a_n}{b_n}, 0, 0, \ldots\right)$$

aplicamos entonces el inciso (b) del lema anterior y ya. □

Lema de Kronecker 4.3
Sea $\{b_n\}$ una sucesión creciente de números reales positivos con $b_n \longrightarrow \infty$, y sea $\{X_n\}$ una sucesión de números reales con $\sum_{n=1}^{\infty} X_n = X$ (finito). Entonces

$$\frac{1}{b_n} \sum_{j=1}^{\infty} b_j X_j \longrightarrow 0 \quad \text{cuando} \quad n \longrightarrow \infty$$

Demostración:
Si $S_n = \sum_{j=1}^{n} X_j$ con $S_0 = 0$,

$$\sum_{j=1}^{n} b_j X_j = \sum_{j=1}^{n} b_j (S_j - S_{j-1}) = b_n S_n - b_0 S_0 - \sum_{j=1}^{n} S_{j-1}(b_j - b_{j-1}), \quad \text{con} \quad b_0 = 0$$

$$\Rightarrow \frac{1}{b_n} \sum_{j=1}^{n} b_j X_j = S_n - \frac{1}{b_n} \sum_{j=1}^{n} a_j S_{j-1}, \quad \text{con} \quad a_j = b_j - b_{j-1} \geq 0$$

Como $S_n \longrightarrow X$ cuando $n \longrightarrow \infty$ y $S_{j-1} \longrightarrow X$ cuando $j \longrightarrow \infty$, entonces por el lema de Toeplitz $\frac{1}{b_n} \sum_{j=1}^{n} b_j X_j \longrightarrow 0$ □

El siguiente resultado es una especie de generalización de la desigualdad de Chebyshev.

Desigualdad de Kolmogorov 4.4
Sea $X_1, \ldots, X_n$ v.a.i. con esperanza finita, y sea $S_j = X_1 + \ldots + X_j$, $j = 1, \ldots, n$. Entonces para cualquier $\varepsilon > 0$,

$$P\left\{\max_{1\le j\le n} |\, S_j - E(S_j)\,| \ge \varepsilon\right\} \le \frac{var\, S_n}{\varepsilon^2}$$

Demostración:

Sin pérdida de generalidad podemos asumir que $E(S_j) \equiv 0$.

Definamos $A_k = \{|S_j| < \varepsilon,\ j = 1, \ldots, k-1, |S_k| \ge \varepsilon\}$, y $A = \left\{\max_{1\le j\le n} |S_j| \ge \varepsilon\right\}$

A es la unión disjunta de A_k, $k = 1, \ldots, n$.

$$\text{Ahora} \quad var\, S_n = \int_\Omega S_n^2\, dP \ge \int_A S_n^2\, dP = \sum_{k=1}^n \int_{A_k} S_n^2\, dP \qquad \ldots(*)$$

pero $S_n = S_k + \Psi_k$, donde $\Psi_k = X_{k+1} + \ldots + X_n$; por tanto

$$\int_{A_k} S_n^2\, dP = \int_{A_k} S_k^2\, dP + 2\int_{A_k} S_k \Psi_k\, dP + \int_{A_k} \Psi_k^2\, dP \qquad \ldots(**)$$

Analizando $(**)$.
El segundo término del lado derecho es $2E(I_{A_k} S_k \Psi_k)$, que es cero pues $I_{A_k}S_k$ y Ψ_k, son funciones de v.a.i., y por lo tanto son independientes.
Como el tercer término del lado derecho es no-negativo, tenemos

$$\int_{A_k} S_n^2\, dP \ge \int_{A_k} S_k^2\, dP \ge \varepsilon^2 P(A_k) \quad \text{por definición de} \quad A_k$$

por lo tanto por $(*)$

$$var\, S_n \ge \varepsilon^2 \sum_{k=1}^n P(A_k) = \varepsilon^2 P(A)$$

□

Segundo lema de Borel-Cantelli 4.5
Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, y sean $A_1, A_2, \ldots$ eventos independientes en $\mathcal{F}$.
Si $\sum_{n=1}^{\infty} P(A_n) = \infty$, entonces $P(\overline{\lim_n}\, A_n) = 1$

Observación:
Se trata de un casi-inverso del lema de Borel–Cantelli que reza:
Si $A_1, A_2, \ldots$ son eventos tales que $\sum_n P(A_n) < \infty$, entonces $\overline{\lim_n}\ A_n$ tiene probabilidad 0.

Demostración:

$$P\left(\overline{\lim_n}\ A_n\right) = P\left(\cap_n \cup_{k \geq n} A_k\right) = \lim_{n \to \infty} P\left(\cup_{k=n}^{\infty} A_k\right) = \lim_{n \to \infty} \lim_{m \to \infty} P\left(\cup_{k=n}^{m} A_k\right)$$

Ahora

$$\begin{aligned} P\left(\cup_{k=n}^{m} A_k\right)^c &= P\left(\cap_{k=n}^{m} A_k^c\right) \\ &= \prod_{k=n}^{m} P(A_k^c) \quad \text{por independencia} \\ &\leq \prod_{k=n}^{m} exp[-P(A_k)] \text{ pues } P(A_k^c) = 1 - P(A_k) \leq exp[-P(A_k)] \\ &\longrightarrow 0 \quad \text{cuando } m \to \infty \quad \text{pues } \sum P(A_k) = \infty \end{aligned}$$

□

Estamos ahora en posición de establecer algunos resultados básicos en lo que es la convergencia de sucesiones de v.a.

Teorema: 4.6
Sea $X_1, X_2, \ldots$, v.a.i. con esperanza finita. Si $\sum_{n=1}^{\infty} var\ X_n < \infty$, entonces

$$\sum_{n=1}^{\infty} [X_n - E(X_n)] \quad \text{converge c.d.}$$

Demostración:
Sin pérdida de generalidad podemos pensar que $E(X_n) \equiv 0$.
Sea $S_n = X_1 + \ldots + X_n$. Entonces S_n converge $\Leftrightarrow S_j - S_k \longrightarrow 0$ cuando $j, k \to \infty$, y esto ocurre c.d. $\Leftrightarrow \forall \varepsilon > 0 \quad P\left[\cup_{j,k \geq n}\{|S_j - S_k| \geq \varepsilon\}\right] \longrightarrow 0$ cuando $n \to \infty$.

De manera equivalente tenemos que probar que $\forall \varepsilon > 0$,

$$P\left[\cup_{k=1}^{\infty}\{|S_{m+k} - S_m| \geq \varepsilon\}\right] \longrightarrow 0 \quad \text{cuando} \quad m \longrightarrow \infty$$

Pero...

$$\begin{aligned}
P\left[\cup_{k=1}^{\infty}\{|S_{m+k}-S_m|\geq\varepsilon\}\right] &= \lim_{n\to\infty} P\left[\cup_{k=1}^{n}\{|S_{m+k}-S_m|\geq\varepsilon\}\right] \\
&= \lim_{n\to\infty} P\left\{\max_{1\leq k\leq n}|S_{m+k}-S_m|\geq\varepsilon\right\} \\
&\leq \frac{1}{\varepsilon^2}\, var\,(S_{m+n}-S_m) \quad \text{por la desigualdad de Kolmogorov} \\
&= \frac{1}{\varepsilon^2}\sum_{j=1}^{n} var\,(X_{m+j}) \\
&\longrightarrow 0 \quad \text{cuando} \quad n\longrightarrow\infty \quad \text{pues} \quad \sum_{n=1}^{\infty} var\, X_n < \infty
\end{aligned}$$

□

Ley fuerte de los grandes números de Kolmogorov. 4.7

Sean $X_1, X_2, \ldots$ v.a.i., cada una de ellas con esperanza y varianza finitas, y sea $\{b_n\}$ una sucesión creciente de números reales positivos con $b_n \longrightarrow \infty$.
Si $\sum_{n=1}^{\infty} \frac{var\, X_n}{b_n^2} < \infty$, entonces (con $S_n = X_1 + \ldots + X_n$) $\quad \frac{S_n - E(S_n)}{b_n} \longrightarrow 0$ c.d.

Demostración:

$$\sum_{n=1}^{\infty} var\left(\frac{X_n - E(X_n)}{b_n}\right) = \sum_{n=1}^{\infty} var\,\frac{X_n}{b_n^2} < \infty \quad \text{por hipótesis}$$

por el teorema anterior $\sum_{n=1}^{\infty}\left(\frac{X_n - E(X_n)}{b_n}\right)$ converge c.d.

Pero $\frac{S_n - E(S_n)}{b_n} = \frac{1}{b_n}\sum_{k=1}^{n} b_k\left(\frac{X_k - E(X_k)}{b_k}\right)$ y esto tiende a cero pues por el lema de Kronecker. □

Definición: 4.8

Si todas las v.a. X_n tienen la misma distribución, se dice que son idénticamente distribuídas. La frase "independientes e idénticamente distribuídas" se abreviará: iid.

Definición: 4.9

Sea $X_1, X_2, \ldots$ una sucesión de v.a., y sea $\mathcal{F}_n = \mathcal{F}(X_n, X_{n+1}, \ldots)$, $n = 1, 2, \ldots$, $\mathcal{F}_n$ se puede pensar como la σ-álgebra de eventos que involucra a $X_n, X_{n+1}, \ldots$. La σ-álgebra

$\mathcal{F}_0 = \cap_{n=1}^{\infty} \mathcal{F}_n$ se llama la σ-álgebra cola de X_n, y conjuntos en $\mathcal{F}_0$ se llaman eventos cola y funciones $\mathcal{F}_0$-medibles, esto es, funciones $f : (\Omega, \mathcal{F}_0) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$ se llaman funciones cola (relativo a X_n).

Ley cero–uno de Kolmogorov 4.10

Todos los eventos cola relativos a una sucesión de v.a.i. tienen probabilidad 0 ó 1, y todas las funciones cola son constantes c.d.

Demostración:

Sea $A \in \mathcal{F}_0$; la idea es probar que A es independiente de si mismo, de tal manera que $P(A \cap A) = P(A)P(A)$ y en consecuencia $P(A) = 0$ ó 1.

Como $\mathcal{F}_0 \subset \mathcal{F}_1$, A es de la forma $\{(X_1, X_2, \ldots) \in A'\}$ para alguna $A' \in \mathcal{B}(\mathbb{R})^\infty$.

Sea $\mathcal{C}$ la clase de conjuntos $C' \in \mathcal{B}(\mathbb{R})^\infty$ tales que A y C son independientes, donde $C = \{(X_1, X_2, \ldots) \in C'\}$.

Si C' es un cilindro medible, entonces C es de la forma $\{(X_1, \ldots, X_n) \in \mathcal{B}_n\}$; pero como $A \in \mathcal{F}_{n+1}$, A se puede escribir como $\{(X_{n+1}, X_{n+2}, \ldots) \in A_{n+1}\}$, y se puede concluír que A y C son independientes, por lo tanto $\mathcal{C}$ contiene todos los cilindros medibles.

Pero si $C_n' \in \mathcal{C}$, $C_n' \uparrow C'$ (ó $C_n' \downarrow C'$), y $P(A \cap C_n) = P(A)P(C_n)$, $n = 1, 2, \ldots$ entonces $C_n \uparrow C$ (ó $C_n \downarrow C$), entonces $P(A \cap C) = P(A)P(C)$ por lo tanto $\mathcal{C}$ es una clase monótona que contiene a todos los cilindros medibles; entonces $\mathcal{C}$ contiene todos los conjuntos en $\mathcal{B}(\mathbb{R})^\infty$, y en particular A'. Pero entonces A es independiente de si mismo.

Finalmente, si f es una función cola, entonces para cada $c \in \overline{\mathbb{R}}$, $\{\omega : f(\omega) < c\}$ es un evento cola, y por lo tanto tiene probabilidad 0 ó 1.

Si $k = \sup\{c \in \overline{\mathbb{R}} : P\{f < c\} = 0\}$, entonces $f = k$ c.d. □

Definición: 4.11

Sea $X_1, X_2, \ldots$ una sucesión de v.a. y definamos las σ– álgebras como en la definición (4.9).

Sea $A \in \mathcal{F}_1$, tal que $A = \{(X_1, X_2, \ldots) \in A'\}$, $A' \in \mathcal{B}(\mathbb{R})^\infty$. Se dice que el evento A es simétrico $\Leftrightarrow$ la ocurrencia de A o la no ocurrencia no se ve afectada por una permutación de X_i numerable.

Formalmente, si $T\{1, 2, \ldots\} \longrightarrow \{1, 2, \ldots\}$ es una permutación de algunas X_i, se requiere que $A = \{(X_{T(1)}, X_{T(2)}, \ldots) \in A'\}$.

La ley cero–uno de Hewitt–Savage. 4.12

Sea $X_1, X_2, \ldots$ v.a. iid.

Si A es un conjunto simétrico en $\mathcal{F}(X_1, X_2, \ldots)$, entonces $P(A) = 0$ ó 1.

Demostración:

Sea $A = \{(X_1, X_2, \ldots) \in A'\}$ así que $P(A) = P_X(A')$, $X = (X_1, X_2, \ldots)$. Podemos encontrar cilindros medibles C'_k tales que $P_X(A'\Delta C'_k) \longrightarrow 0$ cuando $k \longrightarrow \infty$, y sea $C_k = \{X \in C'_k\}$; digamos $C_k = \{(X_1, \ldots, X_{n_k}) \in B_k\}$. Sea T_k un intercambio de $(1, \ldots, n_k)$ y $(n_{k+1} \ldots 2_{n_k})$.
Como las X_n son iid, X y $X(T_k)$ tienen la misma distribución:

$$
\begin{aligned}
P(A\Delta C_k) &= P_X(A'\Delta C'_k) = P_{X(Tk)}(A'\Delta C'_k) \\
\text{por lo tanto} \qquad &= P[\{X(T_k) \in A'\}\Delta\{X(T_k) \in C'_k\}] \\
&= P[\{X \in A'\}\Delta\{X(T_k) \in C'_k\}] \quad \text{pues } A \text{ es simétrico} \\
&= P(A\Delta C_k(T_k))
\end{aligned}
$$

Entonces $P(A\Delta C_k)$ y $P(A\Delta C_k(T_k)) \longrightarrow 0$, y por lo tanto también
$P(A\Delta[C_k \cap C_k(T_k)]) \Rightarrow P(C_k)$, $P(C_k(T_k))$ y $P[C_k \cap C_k(T_k)] \longrightarrow 0$
Pero $P[C_k \cap C_k(T_k)] = P[\{(X_1, \ldots, X_{nk}) \in B_k,\ (X_{nk+1}, \ldots, X_{2nk}) \in B_k\}] = P(C_k)P(C_k(T_k))$
Si hacemos tender $k \longrightarrow \infty$ obtenemos $P(A) = P(A)P(A)$ $\square$

Tiempos de Paro

El término "tiempo de paro" es una expresión del juego. Un juego cuyos sucesos se desarrollan en el tiempo (por ejemplo, el lanzar un número infinito de veces una moneda) puede ser representado de manera adecuada por un espacio $(\Omega, \mathcal{F}, P)$ con $(\mathcal{F}_n, n \in \mathrm{N})$ una filtración.

Una regla de paro para el jugador consiste en dar una receta para abandonar el juego, basada en cada tiempo n, $(n \in \mathrm{N})$ con la información que éste tiene hasta ese momento. Esta receta puede ser por ejemplo, cuando alcance cierta ganancia, o cuando ya no tenga fondos para seguir apostando.

Denotamos por T el tiempo de parar el juego dada tal regla. En el caso en que el juego no se detenga tendremos que $T = +\infty$.

Definición: 5.1
Sea $\{\mathcal{F}_n\}$ una sucesión creciente de sub–σ–álgebras de $\mathcal{F}$. Un tiempo de paro para las $\{F_n\}$ es una v.a. $T : \Omega \longrightarrow \{0, 1, \ldots, \infty\}$ tal que $\{T \leq n\} \in \mathcal{F}_n \quad \forall\, n \in \mathrm{Z}^+$

Observación 5.2
La definición es equivalente al requerimiento de que $\{T = n\} \in \mathcal{F}_n \; \forall\, n \in \mathrm{Z}^+$

Demostración:

a) $\{\omega \in \Omega | T(\omega) = n\} = \{\omega \in \Omega | T(\omega) \leq n\} - \{\omega \in \Omega | T(\omega) \leq n - 1\} \in \mathcal{F}_n$
ya que $\mathcal{F}_n$ es σ–álgebra.

b) $\{\omega \in \Omega | T(\omega) \leq n\} = \cup_{k=0}^{n} \{\omega \in \Omega | T(\omega) = k\} \; \in \mathcal{F}_n$.

Observación 5.3
Si $\{X_n, n = 0, 1, 2, \ldots\}$ es una sucesión de v.a., un tiempo de paro para $\{X_n\}$ es, por definición un tiempo de paro relativo a las σ–álgebras $\mathcal{F}_n = \mathcal{F}(X_0, \ldots, X_n)$

Ejemplo:
Uno de los ejemplos más importantes es el conocido como ("el primer ") tiempo de ocurrencia (hitting time).
Sea $\{X_n\}$ una sucesión de v.a. y $B \in \mathcal{B}(\mathrm{R})$, sea

$$T(\omega) = \begin{cases} min\{n : X_n(\omega) \in B\} & \text{si} \quad X_n(\omega) \in B \text{ para alguna } n \\ +\infty & \text{si} \quad X_n(\omega) \text{ no está nunca en } B \end{cases}$$

i.e. T es el momento de la primera entrada de X a B.
T es un tiempo de paro pues

$$\{T \leq n\} = \cup_{k \leq n}\{X_k \in B\} \in \mathcal{F}(X_k, k \leq n)$$

Definición 5.4
Sea T un tiempo de paro para las σ–álgebras $\mathcal{F}_n, n = 0, 1, \ldots$, y sea $A \in \mathcal{F}$.
El conjunto A se dice que es previo a $T \Leftrightarrow A \cap \{T \leq n\} \in \mathcal{F}_n \; \forall \; n = 0, 1, \ldots$
La clase de todos los conjuntos previos a T se denota por $\mathcal{F}_T$.

Observación: 5.5
$\mathcal{F}_T$ es una σ–álgebra. $\quad \mathcal{F}_T = \{A \in \mathcal{F} : A \cap \{T \leq n\} \in \mathcal{F}_n \; \forall n \in \mathbb{N}\}$.

Demostración:
$i)$ p.d. $\Omega \in \mathcal{F}_T$.

$$\text{pero} \quad \Omega = \cup_{n=0}^{\infty}\{T \leq n\} \quad \text{por lo tanto} \quad \Omega \in \mathcal{F}_T$$

$ii)$ p.d. $A, B \in \mathcal{F}_T \Rightarrow A - B \in \mathcal{F}_T$.

$$A \in \mathcal{F}_T \Rightarrow A \cap \{T \leq n\} \in \mathcal{F}_n \; \forall n \in \mathbb{N}$$

$$B \in \mathcal{F}_T \Rightarrow B \cap \{T \leq n\} \in \mathcal{F}_n \; \forall n \in \mathbb{N}$$

$$\Rightarrow (A - B) \cap \{T \leq n\} = (A \cap \{T \leq n\}) - (B \cap \{T \leq n\}) \in \mathcal{F}_n \; \forall n$$

por ser $\mathcal{F}_n$ σ–álgebra

$$\text{por lo tanto } A - B \in \mathcal{F}_T$$

$iii)$ p.d. Si $(A_n)_{n=1}^{\infty} \in \mathcal{F}_T \Rightarrow \cup_{n=1}^{\infty} A_n \in \mathcal{F}_T$, donde $(A_n)_1^{\infty}$ es una sucesión de elementos de $\mathcal{F}_T$

$$(\cup_{n=1}^{\infty} A_n) \cap \{T \leq n\} = \cup_{n=1}^{\infty}(A_n \cap \{T \leq n\}) \in \mathcal{F}_n \; \forall n$$

pues $\mathcal{F}_n$ es σ–álgebra

$$\text{por lo tanto } \cup_{n=1}^{\infty} A_n \in \mathcal{F}_T$$

□

Observación: 5.6

La σ-álgebra $\mathcal{F}_T$ aparece con frecuencia de la siguiente manera. Sea T un tiempo de paro finito para $\{X_n\}$, y definamos X_T de la siguiente manera. Si $T(\omega) = n$, sea $X_T(\omega) = X_n(\omega)$. Si $B \in \mathcal{B}(\mathbb{R})$, entonces $\{X_T \in B\} \in \mathcal{F}_T$, en otras palabras, X_T es $\mathcal{F}_T$ medible. (Como por definición $\mathcal{F}_T \subset \mathcal{F}$, se sigue en particular que X_T es una v.a.).

Para verlo, consideremos

$$\{X_T \in B\} \cap \{T \leq n\} = \cup_{n=0}^{n} \big[\{X_k \in B\} \cap \{T = k\}\big]$$

$$\text{Como} \quad \{X_k \in B\} \cap \{T = k\} \in \mathcal{F}(X_0, \ldots, X_n) \quad \forall\, k \leq n,$$

$$\text{tenemos} \quad \{X_T \in B\} \cap \{T \leq n\} \in \mathcal{F}(X_0, \ldots, X_n).$$

Además, como T es finito, tenemos que $\cup_{n=0}^{\infty} \{T \leq n\} = \Omega$,

$$\Rightarrow \{X_T \in B\} \in \mathcal{F}$$

$\square$

Si consideramos la fortuna del jugador en los tiempos de paro $T_1, T_2, \ldots$, donde $T_1 \leq T_2 \leq \ldots$ forman una sucesión creciente de tiempos de paro. La pregunta que surge es: cuándo se conserva la condición de martingala?.

Si $X_{T_1}, X_{T_2}, \ldots$ forma una martingala, entonces $E(X_{T_n}) = E(X_{T_1}) \quad \forall\, n$, y si $T_1 \equiv 1$, entonces $E(X_{T_1}) = E(X_1)$.

El siguiente teorema nos da condiciones suficientes para que esto suceda.

Teorema de muestreo opcional (J. L. Doob) 5.7

Sea $\{X_1, X_2, \ldots\}$ una submartingala, y sean $T_1, T_2, \ldots$ una sucesión creciente de tiempos de paro finitos para $\{X_n\}$, con $\Psi_n = X_{T_n}$, $n = 1, 2, \ldots$. Si

$$\text{A)} \quad E(|\Psi_n|) < \infty\, \forall\, n \qquad \text{y} \qquad \text{B)} \quad \lim_{k \to \infty} \int_{\{T_n > k\}} |X_k|\, dP = 0\, \forall\, n$$

entonces $\{\Psi_n\}$ es una submartingala relativa a las σ-álgebras $\mathcal{F}_{T_n}$. Si $\{X_n\}$ es una martingala, también lo es $\{\Psi_n\}$.

Demostración:

La haremos en dos partes:

i) p.d. $\mathcal{F}_{T_n}$ crece con n, *i.e.* $\mathcal{F}_{T_n}$ es una filtración.

Sea $A \in \mathcal{F}_{T_n} \Rightarrow$

$$A \cap \{T_{n+1} \leq k\} = \cup_{i=1}^{k} \big[A \cap \{T_n = i\} \cap \{T_{n+1} \leq k\}\big]$$

pero $A \cap \{T_n = i\} \in \mathcal{F}(X_1, \ldots, X_i) \subset \mathcal{F}(X_1, \ldots, X_k)$ y $\{T_{n+1} \leq k\} \in \mathcal{F}(X_1, \ldots, X_k)$ por lo tanto $A \in \mathcal{F}_{n+1}$.

ii) Si $A \in \mathcal{F}_{T_n} \Rightarrow$ p.d. $\int\limits_A \Psi_{n+1}\, dP \geq \int\limits_A \Psi_n\, dP$. Con igualdad en el caso de martingalas.

Por definición tenemos que $A = \cup_j \big[A \cap \{T_n = j\}\big]$ podemos entonces sustituír A con $D_j = A \cap \{T_n = j\} \in \mathcal{F}(X_1, \ldots, X_j)$
Ahora si $k > j$, tenemos que: $T_n = j \Rightarrow T_{n+1} \geq j$ de tal suerte que

$$\int\limits_{D_j} \Psi_{n+1}\, dP = \sum_{i=j}^{k} \int\limits_{D_j \cap \{T_{n+1} = i\}} \Psi_{n+1}\, dP + \int\limits_{D_j \cap \{T_{n+1} > k\}} \Psi_{n+1}\, dP$$

Así

$$\int\limits_{D_j} \Psi_{n+1}\, dP = \sum_{i=j}^{k} \int\limits_{D_j \cap \{T_{n+1} = i\}} X_i\, dP + \int\limits_{D_j \cap \{T_{k+1} > k\}} X_k\, dP - \int\limits_{D_j \cap \{T_{n+1} > k\}} (X_k - \Psi_{n+1}\, dP \qquad \ldots(1)$$

Combinando ahora el término $i = k$ en (1) con el término $\int X_k\, dP$ obtenemos:

$$\int\limits_{D_j \cap \{T_{n+1} = k\}} X_k\, dP + \int\limits_{D_j \cap \{T_{n+1} > k\}} X_k\, dP = \int\limits_{D_j \cap \{T_{k+1} \geq k\}} X_k\, dP \geq \int\limits_{D_j \cap \{T_{n+1} \geq k\}} X_{k-1}\, dP$$

pues $\{X_n, \mathcal{F}_n\}$ es una submartingala y como $\{T_{n+1} \geq k\} = \{T_{n+1} \leq k-1\}^c \in \mathcal{F}(X_1, \ldots, X_{n-1})$ y $D_j \in \mathcal{F}(X_1, \ldots, X_j) \subset \mathcal{F}(X_1, \ldots, X_{k-1})$.
Pero

$$\int\limits_{D_j \cap \{T_{n+1} \geq k\}} X_{k-1}\, dP = \int\limits_{D_j \cap \{T_{n+1} > k-1\}} X_{k-1}\, dP$$

por tanto este término puede ser combinado con el término $i = k-1$ de (1) para obtener:

$$\int\limits_{D_j \cap \{T_{n+1} > k-2\}} X_{k-2}\, dP$$

Procediendo de manera inductiva tenemos que

$$\int\limits_{D_j} \Psi_{n+1}\, dP \geq \int\limits_{D_j \cap \{T_{n+1} \geq j\}} X_j\, dP - \int\limits_{D_j \cap \{T_{n+1} > k\}} (X_k - \Psi_{n+1})\, dP$$

Pero por hipótesis, $\int\limits_{D_j \cap \{T_{n+1} > k\}} X_k\, dP \to 0$ y $\int\limits_{D_j \cap \{T_{n+1} > k\}} \Psi_{n+1}\, dP \to 0$ cuando $k \longrightarrow \infty$ pues $\{T_{n+1} > k\} \longrightarrow \emptyset$.

Finalmente; $D_j \cap \{T_{n+1} \geq j\} = D_j$ pues $D_j \subset \{T_n = j\}$, y $X_j = \Psi_n$ en D_j. Por lo tanto

$$\int_{D_j} \Psi_{n+1}\, dP \geq \int_{D_j} \Psi_n\, dP$$

lo que concluye la prueba.
Para probar el caso de martingalas se hace igual, pero en vez de desigualdades ponemos igualdades y ya.

□

Hay algunos casos especiales de este teorema, que son tan importantes, que conviene enunciarlos como teoremas.

Teorema 5.8
Sea $\{X_n\}$ una submartingala, $\{T_n\}$ una sucesión creciente de tiempos de paro finitos para $\{X_n\}$. Las condiciones (A) y (B) del teorema anterior se sostienen en cualquiera de las siguientes situaciones.

a) Cada T_n es acotada, esto es, $\forall n \ni k_n$ constante tal que $T_n \leq k_n$ c.d.

Demostración:

$$\int_{\Omega} |X_{T_n}|\, dP = \sum_{i \leq k_n} \int_{\{T_n = i\}} |X_i|\, dP \leq \sum_{i \leq k_n} E(|X_i|) < \infty$$

Por lo tanto (A) es cierto.
y (B) se sigue de que $\{T_n > k\}$ es vacío para k suficientemente grande. □

b) $E\left(\sup_n |X_n|\right) < \infty$ (en particular si X_n son uniformemente acotadas)

Demostración:
Consideremos $Z = \sup_n |X_n|$, entonces $\int_{\Omega} |X_{T_n}|\, dP \leq E(z) < \infty$
y por lo tanto A se cumple.

B se prueba de la siguiente manera

$$\int_{\{T_n > k\}} |X_k|\, dP \leq \int_{\{T_n > k\}} Z\, dP \longrightarrow 0 \quad \text{cuando} \quad k \longrightarrow \infty$$

$$\text{pues} \quad \{T_n > k\} \downarrow \emptyset \quad \text{y} \quad Z \quad \text{es integrable}$$

□

El siguiente teorema generaliza el inciso (b) del teorema anterior.

Teorema 5.9
Si $\{T_n\}$ es una sucesión creciente de tiempos de paro finitos para una submartingala $\{X_n\}$ con último elemento X_∞, entonces $\{X_{T_n}\}$ es una submartingala relativa a las σ–álgebras $\mathcal{F}_{T_n}$; si $\{X_n\}$ es una martingala, también lo es $\{X_{T_n}\}$. En particular, esto es cierto si X_n son uniformemente integrables.

Demostración:
La demostración se hará por casos:

Caso 1:
Sea $X_n \leq 0 \,\forall n$, y $X_\infty = 0$.
Para cualquier n fija consideremos $S_k = T_n \wedge k = min\ (T_n, k)$, $\quad k = 1, 2, \ldots;$ $\quad S_k$ es entonces un tiempo de paro para X_n.
Ahora, como $X_{s_k} \longrightarrow \Psi_n = X_{T_n}$ cuando $k \longrightarrow \infty$; entonces por el lema de Fatou

$$\overline{\lim_{k\to\infty}} \int_\Omega X_{s_k}\, dP \leq \int_\Omega \Psi_n\, dP \leq 0$$

pero por el teorema 5.8 inciso a), $\{X_{s_k}\}$ es una submartingala, entonces

$$\int_\Omega X_{s_k}\, dP \geq \int_\Omega X_{s_1}\, dP = \int_\Omega X_1\, dP, \quad \text{que es finito}$$

y por lo tanto Ψ_n es integrable.

Otra vez por el inciso (a) del teorema 5.8, $\{X_{T_n\wedge k}, \mathcal{F}_{T_n\wedge k}, n = 1, 2, \ldots\}$ es una submartingala. Nota: k es fija.

$$\text{Entonces} \quad \int_A X_{T_n\wedge k}\, dP \leq \int_A X_{T_{n+1}\wedge k}\, dP, \quad \text{si} \quad A \in \mathcal{F}_{T_n\wedge k}$$

Pero si $A \in \mathcal{F}_{T_n}$, entonces $A \cap \{T_n \leq k\} \in \mathcal{F}_{T_n\wedge k}$, pues

$$A \cap \{T_n \leq k\} \cap \{T_n \wedge k \leq i\} = \begin{cases} A \cap \{T_n \leq i\} & \text{para} \quad i \leq k \\ A \cap \{T_n \leq k\} & \text{para } i > k \end{cases}$$

$$\text{Entonces} \int_{A\cap\{T_n\le k\}} X_{T_n\wedge k}\,dP \le \int_{A\cap\{T_n\le k\}} X_{T_{n+1}\wedge k}\,dP, \quad A\in \mathcal{F}_{T_n}$$

Pero en $\{T_n \le k\}$, $T_n\wedge k = T_n$; y $\{T_{n+1}\le k\}\subset\{T_n\le k\}$ y $X_{T_{n+1}\wedge k}\le 0$;

$$\text{Entonces} \int_{A\cap\{T_n\le k\}} X_{T_n}\,dP \le \int_{A\cap\{T_{n+1}\le k\}} X_{T_{n+1}}\,dP$$

haciendo tender $k\longrightarrow\infty$ obtenemos la propiedad de submartingala.

Caso 2:
$X_n = E(X_\infty|\mathcal{F}_n)$, $n=1,2,\ldots$; entonces tenemos una martingala uniformemente integrable y por resultados anteriores podemos concluír que $X_{T_n\wedge k}$, $k=1,2,\ldots$, son u.i. y que $X_{T_n\wedge k}\longrightarrow X_{T_n}$ cuando $k\longrightarrow\infty$. Entonces X_{T_n} es integrable y por lo tanto se satisface la condición A.
La condición (B) se sigue directamente de la integrabilidad uniforme pues

$$\int_A |X_k|\,dP\to 0 \quad \text{cuando} \quad P(A)\to 0 \quad \text{uniformemente en } k$$

y $P\{T_n>k\}\longrightarrow 0$ cuando $k\longrightarrow\infty$

Para el caso general:
Podemos escribir a $X_n = X_n' + X_n''$, donde $X_n' = X_n - E(X_\infty|\mathcal{F}_n)$, y $X_n'' = E(X_\infty|\mathcal{F}_n)$.
Por lo que el resultado para X_n se sigue de lo ya probado. □

Teorema: (para el caso de una sucesión finita) 5.10
Sea $\{X_1,\ldots,X_n\}$ una submartingala y sean $T_1,T_2,\ldots$, una sucesión creciente de tiempos de paro para $\{X_i, 1\le i\le n\}$.
Entonces las X_{T_i} forman una submartingala relativa a las $\sigma-$ álgebras $\mathcal{F}_{T_i}$, y una martingala si $\{X_i\}$ es una martingala.

Demostración:

Podemos extender la sucesión a una sucesión infinita $X_1,X_2,\ldots,X_n,X_n,\ldots$, y aplicar el inciso (a) del teorema 5.8. □

Aquí concluyen los teoremas referentes al teorema de muestreo opcional, veamos ahora algunos resultados que se pueden obtener en lo que es convergencia de martingalas a partir de estos resultados.

Los siguientes dos lemas los vamos a usar para probar un resultado muy importante que es la ley fuerte de los grandes números para el caso iid.; donde iid. abrevia independientes, idénticamente distribuídas.

Lema 5.11
Si $X_1, X_2, \ldots, X_n$ son v.a. independientes, idénticamente distribuídas con esperanza finita, y

$$S_n = \sum_{k=1}^{n} X_k, \quad \text{entonces} \quad E(X_k|S_n) = \frac{S_n}{n} \; c.d. \quad k = 1, \ldots, n$$

Demostración:
Si $B = \mathcal{B}(\mathbb{R})$, entonces

$$\int_{\{S_n \in B\}} X_k \, dP = E[X_k I_{\{S_n \in B\}}]$$

$$= \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} X_k I_B(X_1 + \ldots + X_n) \, dF(X_1) \ldots dF(X_n)$$

donde F es la función de distribución de X_i. Por el teorema de Fubini, esto es independiente de la k; y así

$$\int_{\{S_n \in B\}} X_n \, dP = \frac{1}{n} \int_{\{S_n \in B\}} \sum_{k=1}^{n} X_k \, dP = \int_{\{S_n \in B\}} \frac{S_n}{n} \, dP$$

□

Lema 5.12
Si $X_1, X_2, \ldots$ son v.a. y $S_n = \sum_{n=1}^{n} X_k$, entonces

$$\mathcal{F}(S_n, S_{n+1}, S_{n+2}, \ldots) = \mathcal{F}(S_n, X_{n+1}, X_{n+2}, \ldots)$$

Demostración:

Lo haremos probando ambas contenciones.
Como $X_{n+k} = S_{n+k} - S_{n+k-1}$, $S_n, X_{n+1}, X_{n+2}, \ldots$ son $\mathcal{F}(S_n, S_{n+1}, S_{n+2}, \ldots)$-medibles; entonces $\mathcal{F}(S_n, X_{n+1}, X_{n+2}, \ldots) \subset \mathcal{F}(S_n, S_{n+1}, S_{n+2}, \ldots)$
De igual manera $S_n, S_{n+1}, S_{n+2}, \ldots$ son $\mathcal{F}(S_n, X_{n+1}, X_{n+2}, \ldots)$-medibles entonces $\mathcal{F}(S_n, S_{n+1}, S_{n+2}, \ldots) \subset \mathcal{F}(S_n, X_{n+1}, X_{n+2}, \ldots)$ □

Teorema: La ley fuerte de los grandes números caso iid. 5.13

Si $X_1, X_2, \ldots$ son v.a. iid. con esperanza finita m, y $S_n = X_1 + \ldots + X_n$, entonces $\frac{S_n}{n} \longrightarrow m$ *c.d.* y en $\mathcal{L}_1$.

Demostración:

Vamos a hacer la demostración en dos partes, primero probaremos que $\frac{S_n}{n}$ converge a un límite finito y luego que éste es m.

Si consideramos $(X_1, \ldots, X_n)$ y $(X_{n+1}, X_{n+2}, \ldots)$ que son independientes y como sabemos que "Funciones de objetos aleatorios independientes son independientes" entonces (X_1, S_n) y $(X_{n+1}, X_{n+2}, \ldots)$ son funciones de objetos aleatorios independientes, y por lo mismo son independientes.

Recordemos ahora el siguiente resultado:
Si Ψ es una v.a. integrable y X y Z son objetos aleatorios, y si (X, Ψ) y Z son independientes, entonces $E(\Psi|X, Z) = E(\Psi|X)$

tenemos por lo tanto
$$\begin{aligned} E(X_1|S_n) &= E(X_1|S_n, X_{n+1}, \ldots) \\ &= E(X_1|S_n, S_{n+1}, \ldots) \quad \text{por el lema 5.12} \end{aligned}$$

Pero entonces por el lema 5.11

$$E(X_1|S_n, S_{n+1}, \ldots) = \frac{S_n}{n} \quad c.d.$$

Pero por el teorema 3.28 , $E(X_1|S_n, S_{n+1}, \ldots) \to E(X_1|\mathcal{G}_\infty)$ *c.d.* y en $\mathcal{L}_1$, donde $\mathcal{G}_\infty$ es la σ-álgebra cola de S_n.
Por lo tanto $\frac{S_n}{n}$ converge c.d. y en $\mathcal{L}_1$ a un límite finito.

Ahora hay que probar que este límite finito es m, para lo cual podemos proceder de dos formas.

primera forma
Notemos que $\lim\limits_{n\to\infty} (\frac{S_n}{n})$ es una * función cola de X_n, y por tanto es constante c.d. por la ley cero–uno de Kolmogorov.

* La σ-álgebra $\mathcal{F}_0 = \cap_{n=1}^{\infty} \mathcal{F}_n$ es la σ-álgebra cola de X_n. Funciones $\mathcal{F}_0$-medibles son funciones cola
$f : (\Omega, \mathcal{F}_0) \longrightarrow (\overline{\mathbb{R}}, \mathcal{B}(\overline{\mathbb{R}}))$.

Como $\frac{S_n}{n}$ es convergente en $\mathcal{L}_1$ y $E(\frac{S_n}{n}) \equiv m$, la constante tiene que ser m.

segunda forma

Podemos utilizar también la ley cero–uno de Hewitt–Savage para probar que cada conjunto en $\mathcal{G}_\infty$ tiene probabilidad cero ó uno. [$\mathcal{G}_\infty$ es la σ–álgebra cola de S_n].

Si $A \in \mathcal{G}_\infty$ y T permuta n coordenadas, entonces $A \in \mathcal{F}(S_n, S_{n+1}, \ldots)$; por lo tanto A es de la forma $\{(S_n, S_{n+1}, \ldots) \in A'\}$ para alguna $A \in [\mathcal{B}(\mathbb{R})]^\infty$.

Como $S_k = X_1 + \ldots + X + k = X_{T(1)} + \ldots + X_{T(k)} = S_{T(k)}$, $k \geq n$, A es simétrico, y por lo tanto la ley cero–uno de Hewitt–Savage es aplicable.

Por lo tanto $\mathcal{G}_\infty$ es trivial, y se sigue que $E(X_1|\mathcal{G}_\infty) = E(X_1)$ c.d. $\square$

Teorema 5.14

Sea $\{X_n, \mathcal{F}_n, n = 1, 2, \ldots\}$ una submartingala, y sea $Z = \sup_n (X_n - X_{n-1})$, definimos $X_0 = 0$.

Si $E(Z) < \infty$, entonces X_n converge a un límite finito c.d. en el conjunto $\left\{\sup_n X_n < \infty\right\}$.

Si $\{X_n, \mathcal{F}_n\}$ es una supermartingala y $E\left[\inf_n (X_n - X_{n-1})\right] > -\infty$, entonces X_n converge a un límite finito c.d. en $\left\{\inf_n X_n > -\infty\right\}$.

Demostración:

Fijemos $M > 0$, y sea $T = \inf\{n : X_n > M\}$; si no hay tal n sea $T = \infty$.

Definamos $T_n = T \wedge n$; si $\Psi_n = X_{T_n}$, $n = 1, 2, \ldots$, entonces $\{\Psi_n\}$ es una submartingala por (5.8) inciso (a).

Ahora si $n < T$, entonces $\Psi_n = X_n = X_{n-1} + (X_n - X_{n-1}) \leq M + Z$, y si $n \geq T$, entonces $\Psi_n = X_{T-1} + (X_T - X_{T-1}) \leq M + Z$.

Por lo tanto en cualquier caso $\Psi_n \leq M + Z$, así $\sup_n E(\Psi_n^+) \leq M + E(Z^+) < \infty$ por hipótesis.

Pero por el teorema de convergencia de submartingalas tenemos que Ψ_n converge c.d. a un límite finito.

Pero si $T = \infty$, entonces $X_n \equiv \Psi_n$, por lo tanto X_n converge c.d. en $\left\{\sup_n X_n \leq M\right\}$.

Como M es arbitraria, X_n converge c.d. en $\left\{\sup_n X_n < \infty\right\}$.

Para probar el resultado para supermartingalas, ponemos $-X_n$ en vez de X_n. $\square$

Teorema: 5.15

Si $\{X_1, X_2, \ldots\}$ es una martingala y $E(X_n^2) < \infty \, \forall \, n$, entonces la martingala de diferencias

$X_1, X_2 - X_1, \ldots, X_n - X_{n-1}, \ldots,$ es ortogonal.

Demostración:
Si $j < k$ y $\mathcal{F}_j = \mathcal{F}(X_1, \ldots, X_j)$,

$$\begin{aligned} E[(X_j - X_{j-1})(X_k - X_{k-1})] &= E[E((X_j - X_{j-1})(X_k - X_{k-1})|\mathcal{F}_j)] \\ &= E[(X_j - X_{j-1})E(X_k - X_{k-1}|\mathcal{F}_j)] \end{aligned}$$

pues $X_j - X_{j-1}$ es $\mathcal{F}_j$–medible. Pero $E(X_k - X_{k-1}|\mathcal{F}_j) = X_j - X_j = 0$ por la propiedad de martingala, por lo tanto el teorema es cierto. □

Teorema: 5.16
Sean $\Psi_1, \Psi_2, \ldots,$ v.a. independientes con media 0, y asumamos $E(\sup_k \Psi_k^2) < \infty$.
Si $\sum_{k=1}^{\infty} \Psi_k$ converge c.d., entonces $\sum_{k=1}^{\infty} E(\Psi_k^2) < \infty$.

Demostración:
Las $X_n = \sum_{k=1}^{n} \Psi_k$ forman una martingala, con $\mathcal{F}_n = \mathcal{F}(\Psi_1, \Psi_2, \ldots, \Psi_n)$ *i.e.* $\{X_n, \mathcal{F}_n\}$ es martingala.
Sea M tal que $P\{\sup_n |X_n| \leq M\} > 0$; esto es posible pues la serie converge c.d.

$$\text{Sea} \quad T = \begin{cases} inf & \{n : |X_n| > M\} \\ \infty & \text{si no hay tal } n. \end{cases}$$

p.d. que T es un tiempo de paro
i) $\mathcal{F}_n$ es una filtración por ser submartingala.
ii) p.d. $\{T \leq n\} \in \mathcal{F}_n \ \forall n \in Z^+$

$$\{T \leq n\} = \{\omega \in \Omega | T(\omega) \leq n\} = \cup_{k \leq n} \{\omega \in \Omega | |X_k(\omega)| > M\}$$

a) Si no hay la n de la definición y $T = \infty$, entonces $A = \emptyset \in \mathcal{F}_n$
b) Si hay tal n, $\cup_{k \leq n}\{\omega \in \Omega \mid |X_k(\omega)| > M\} \in \mathcal{F}(\Psi_k, k \leq n)$.
Por lo tanto T es tiempo de paro.

Si $T_n = T \wedge n$, entonces X_{T_n} es una martingala; esto se debe a que $T \wedge n$ es un tiempo de paro pues T lo era, aplicamos entonces el teorema de muestreo opcional, pues cada T_n es acotada.
Observemos ahora que

$$X_n^2 = \sum_{i \leq j}^{n} \Psi_i \Psi_j \qquad X_{n-1}^2 = \sum_{i \leq j}^{n-1} \Psi_i \Psi_j$$

$$\Rightarrow X_n^2 - X_{n-1}^2 = (X_n - X_{n-1})(X_n + X_{n-1}) = \Psi_n(\Psi_n + 2X_{n-1}) = \Psi_n^2 + 2\sum_{i=1}^{n-1} \Psi_n\Psi_i$$

$$\text{pero} \quad E(X_n^2 - X_{n-1}^2) = E(\Psi_n^2) + 2\sum_{i=1}^{n-1} E(\Psi_n)E(\Psi_i) = E(\Psi_n^2) < \infty \quad \text{por hipótesis.}$$

Entonces:

$$|X_{T_n}^2| = (X_{T_n}^2 - X_{T_{n-1}}^2) + X_{T_{n-1}}^2 \leq \sup_j |X_j^2 - X_{j-1}^2| + X_{T_{n-1}}^2 = Z + M$$

Se sigue entonces que los números $E(X_{T_n}^2)$ están uniformemente acotados y por lo tanto X_{T_n} converge c.d. y en $\mathcal{L}_2$.

$$\text{Pero} \quad E(X_{T_n}^2) = \sum_{j=1}^{n} E(X_{T_j}^2 - X_{T_{j-1}}^2) \quad \text{con} \quad X_{T_0} = 0$$

$$\text{pero} \quad E(X_{T_j} - X_{T_{j-1}})^2 = E(X_{T_j}^2 - 2X_{T_j}X_{T_{j-1}} + X_{T_{j-1}}^2) = E(X_{T_j}^2 - X_{T_{j-1}}^2)$$

$$\text{pues} \quad E(X_{T_j}X_{T_{j-1}}) = E\big(E(X_{T_j}X_{T_{j-1}}|\mathcal{F}_{T_{j-1}})\big) =$$

$$= E\big(X_{T_{j-1}}E(X_{T_j}|\mathcal{F}_{T_{j-1}})\big) = E(X_{T_{j-1}}^2) \quad \text{ya que por ser martingala :}$$

$$E(X_{T_j}|\mathcal{F}_{T_{j-1}}) = X_{T_{j-1}}$$

por lo tanto $E(X_{T_n}^2) = \sum_{j=1}^{\infty} E\big((X_{T_j} - X_{T_{j-1}})^2\big)$

Como X_{T_n} es $\mathcal{L}_2$–convergente, $E(X_{T_n}^2)$ se aproxima a un límite finito, por lo tanto $\sum_{n=1}^{\infty} E\big[(X_{T_n} - X_{T_{n-1}})^2\big] < \infty$. Pero

$$X_{T_n} - X_{T_{n-1}} = \Psi_n I_{\{T\geq n\}} \begin{cases} i) & \text{Si } n \leq T \Rightarrow X_{T_n} - X_{T_{n-1}} = \sum_{k=1}^{n} \Psi_k - \sum_{k=1}^{n-1} \Psi_k = \Psi_n \\ ii) & \text{Si } T < n \Rightarrow X_{T_n} - X_{T_{n-1}} = \sum_{k=1}^{T} \Psi_k - \sum_{k=1}^{T} \Psi_k = 0 \\ iii) & \text{Si } n-1 \leq T < n \Rightarrow X_{T_n} - X_{T_{n-1}} = \sum_{k=1}^{T} \Psi_k - \sum_{k=1}^{n-1} \Psi_k = 0 \end{cases}$$

Así $\sum_{n=1}^{\infty} E(\Psi_n^2 I_{\{T\geq n\}}) < \infty$

Notemos que si $Z_n \geq 0$ y $\sum_n E(Z_n) < \infty$, $\Rightarrow E(\sum_n Z_n) < \infty$, así $\sum_n Z_n < \infty$ c.d.
Sea $Z_n = E(\Psi_n^2 I_{\{T\geq n\}}|\Psi_1, \ldots, \Psi_{n-1})$ y como $E(\Psi_n^2 I_{\{T\geq n\}}) = E\big[E(\Psi_n^2 I_{\{T\geq n\}}|\mathcal{F}_{n-1})\big]$

$$\Rightarrow \sum_{n=1}^{\infty} E(\Psi_n^2 I_{\{T\geq n\}}|\Psi_1, \ldots, \Psi_{n-1}) < \infty \quad c.d.$$

Ahora $I_{\{T\geq n\}} = I_{\{T\leq n-1\}^c}$, que es $\mathcal{F}(\Psi_1,\ldots,\Psi_{n-1})$-medible

$$\Rightarrow \sum_{n=1}^{\infty} I_{\{T\geq n\}} E(\Psi_n^2|\Psi_1,\ldots,\Psi_{n-1}) = \sum_{n=1}^{\infty} I_{\{T\geq n\}} E(\Psi_n^2) < \infty \quad c.d.$$

Tomemos entonces un ω donde la serie converja y donde $\sup_n |X_n(\omega)| \leq M$.

Entonces $T(\omega) = \infty \geq n \quad \forall\, n$; se sigue que $\sum_{n=1}^{\infty} E(\Psi_n^2) < \infty$ □

Teorema: 5.17

Sean $\Psi_1, \Psi_2, \ldots$ v.a. independientes con media 0; supongamos $E\left(\sup_k \Psi_k^2\right) < \infty$.

Entonces $\sum_{k=1}^{\infty} \Psi_k$ converge c.d. $\Leftrightarrow \sum_{k=1}^{\infty} E(\Psi_k^2) < \infty$

Demostración:

$\Longrightarrow$ es el teorema anterior.

$\Longleftarrow$ $var\, \Psi_n = E\left[(\Psi_n - E(\Psi_n))^2\right] = E\left([\Psi_n]^2\right) < \infty$ por hipótesis.

$\Rightarrow$ por el teorema 4.6

$$\sum_{n=1}^{\infty} [\Psi_n - E(\Psi_n)] = \sum_{n=1}^{\infty} \Psi_n \quad \text{converge c.d.}$$

□

Teorema: 5.18

Sean Ψ_1, Ψ_2, independientes, v.a. uniformemente acotadas. Entonces $\sum_{k=1}^{\infty} \Psi_k$ converge c.d.

$$\Leftrightarrow \sum_{k=1}^{\infty} var\, \Psi_k < \infty \quad \text{y} \quad \sum_{k=1}^{\infty} E(\Psi_k) \quad \text{converge.}$$

Observación:

En este teorema quitamos la hipótesis de que la media sea cero.

Demostración:

$\Longleftarrow$) Por el teorema anterior, $\sum_{k=1}^{\infty} (\Psi_k - E(\Psi_k))$ converge c.d.; como $\sum_{k=1}^{\infty} E(\Psi_k)$ converge, tenemos que $\sum_{k=1}^{\infty} \Psi_k$ converge c.d.

$\Longrightarrow$) Sean $\Psi_1, Z_1, \Psi_2, Z_2, \ldots$ independientes, donde para cada j, Z_j tiene la misma distribución que Ψ_j.
Entonces $E(\Psi_j - Z_j) = E(\Psi_j) - E(Z_j) = 0$.

$$E\left[(\Psi_j - Z_j)^2\right] = var\,(\Psi_j - Z_j) = var\,\Psi_j + var\,Z_j = 2\,var\,\Psi_j$$

Ahora como Ψ_j tiene la misma distribución que Z_j podemos observar que

$$P\{(\Psi_1, \ldots, \Psi_n) \in B\} = P\{(Z_1, \ldots, Z_n) \in B\}\ \forall n \quad \text{y} \quad \forall B \in \mathcal{B}(\mathrm{R}^n)$$

por lo tanto $P_\Psi = P_Z$.
Podemos concluír entonces que cuando $\sum_j \Psi_j$ converge c.d. también converge $\sum_j Z_j$ c.d.
Se sigue entonces que $\sum_j (\Psi_j - Z_j)$ converge c.d., entonces por el teorema anterior $\sum_j var\,\Psi_j < \infty$ y nuevamente por el teorema anterior

$$\sum_j (\Psi_j - E(\Psi_j)) \quad \text{converge c.d. por lo tanto} \quad \sum_j E(\Psi_j) \quad \text{converge}$$

$\square$

De esta manera llegamos al momento en que podemos enunciar el teorema de las 3 series de Kolmogorov que de alguna manera generaliza lo anterior.

Teorema de las tres series de Kolmogorov. 5.19
Sean $\Psi_1, \Psi_2, \ldots$ v.a.i. Si $M > 0$, definamos

$$\Psi'_j = \begin{cases} \Psi_j & \text{si } |\Psi_j| \leq M \\ 0 & \text{si } |\Psi_j| > M \end{cases}$$

a) Si $\sum_j \Psi_j$ converge c.d., $\Rightarrow$ para cualquier M, las tre series

$$\sum_j P\{\Psi_j \neq \Psi'_j\}, \quad \sum_j E(\Psi'_j), \quad \sum_j var\,\Psi'_j \quad \text{convergen.}$$

Demostración:

Por hipótesis, $\Psi_j \longrightarrow 0$ c.d., así generalmente $\Psi_j = \Psi'_j$, entonces $\Psi_j \neq \Psi'_j$ para un número contable de j, esto es, $P\left(\overline{\lim_j}\{\Psi_j \neq \Psi'_j\}\right) = 0$.

Por el segundo lema de Borel-Cantelli $\sum_j P\{\Psi_j \neq \Psi'_j\} < \infty$.

La convergencia de las otras dos series se prueba aplicando el teorema anterior.

b) Si para alguna $M > 0$, las tres series convergen, entonces $\sum_j \Psi_j$ converge c.d.

Por el teorema anterior $\sum_j \Psi'_j$ converge c.d.

Como $\sum_j P\{\Psi_j \neq \Psi'_j\} < \infty$, tenemos que, casi seguramente, $\Psi_j \neq \Psi'_j$ generalmente, por lo tanto

$\sum_j \Psi_j$ converge c.d. □

La descomposición de submartingalas de Doob

Es un teorema útil en el desarrollo de algunos resultados importantes en la teoría de martingalas, en particular en el estudio de la convergencia de las martingalas cuadrado integrables; y estamos ahora en condiciones de analizarlo.

Nota:

Al igual que en todo este trabajo, vamos a considerar un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ y una sucesión creciente $(\mathcal{B}_n, n \in \overline{\mathbb{N}})$ de sub σ-álgebras de $\mathcal{F}$, donde $\overline{\mathbb{N}} = \mathbb{N} \cup \{0\}$.

Definiciones: 6.1

Una sucesión de v.a. $(X_n, n \in \overline{\mathbb{N}})$ definida en $(\Omega, \mathcal{F}, P; (\mathcal{B}_n, n \in \overline{\mathbb{N}}))$ se dice que es **adaptada** si $\forall n \in \overline{\mathbb{N}}$, las v.a. X_n son $\mathcal{B}_n$- medibles.

Si la sucesión $(X_n, n \in \overline{\mathbb{N}})$ de v.a. es adaptada, lo mismo se cumple para toda sucesión $(X'_n = f_n(X_0, X_1, \ldots, X_n), n \in \overline{\mathbb{N}})$ construida usando funciones $f_n (n \in \overline{\mathbb{N}})$ medibles.

Por otro lado, una sucesión $(X_n, n \in \overline{\overline{\mathbb{N}}})$ definida para $n = +\infty$ va a ser llamada adaptada si las v.a. X_n son $\mathcal{B}_n$-medibles y también para $n = +\infty$.

Definición 6.2

Una sucesión $(U_n, n \in \overline{\mathbb{N}})$ de v.a. se dice que es predecible si la v.a. U_0 es $\mathcal{B}_0$-medible y si $\forall \quad n \in \overline{\mathbb{N}}$ las v.a. U_{n+1} son $\mathcal{B}_n$-medibles.

Un proceso creciente se define como una sucesión predecible $(A_n, n \in \overline{\mathbb{N}})$ de v.a. finitas tales que

$$0 = A_0 \leq A_1 \leq A_2 \leq \ldots \quad \text{c.s. en } \Omega.$$

Notemos que $(U_n, n \in \overline{\mathbb{N}})$ es adaptada en cualquier caso (predecible o proceso creciente).

Es importante hacer notar que para una sucesión predecible y en particular para un proceso creciente, no solo la sucesión $(U_n)_{n \in \overline{\mathbb{N}}}$ sino la sucesión $(U_{n+1})_{n \in \overline{\mathbb{N}}}$ es adaptada a la sucesión $(\mathcal{B}_n, n \in \overline{\mathbb{N}})$ de σ-álgebras.

El contenido de la siguiente proposición ya fue demostrado anteriormente, pero la incluyo aqui y redactada de esta forma para usarla en lo que sigue.

Proposición: 6.3

Para cualquier martingala integrable $(X_n, n \in \overline{\mathbb{N}})$, las siguientes condiciones son equivalentes:

a) La sucesión $(X_n, n \in \overline{\mathbb{N}})$ converge en $\mathcal{L}_1$.
b) $\sup\limits_n E(|X_n|) < \infty$ y el c.s. límite $X_\infty = \lim\limits_n X_n$ de la martingala satisface $X_n = E(X_\infty|\mathcal{B}_n)$ $\forall n \in \overline{\mathbb{N}}$.
b') Existe una v.a. integrable X tal que $X_n = E(X|\mathcal{B}_n)$ $\forall n \in \overline{\mathbb{N}}$.
c) La sucesión $(X_n, n \in \overline{\mathbb{N}})$ satisface la condición de integrabilidad uniforme.

Definición: 6.4
La martingala integrable $(X_n, n \in \overline{\mathbb{N}})$ se dice que es regular si satisface alguna de las condiciones anteriores.

Definición: 6.5
Se dice que el tiempo de paro ν es regular para la martingala $(X_n, n \in \overline{\mathbb{N}})$ si la martingala $(X_{\nu\wedge n}, n \in \overline{\mathbb{N}})$ es regular.

Teorema: 6.6

(1) Toda submartingala integrable puede ser escrita, de manera única, como la suma de una martingala integrable $(M_n, n \in \overline{\mathbb{N}})$ y un proceso creciente $(A_n, n \in \overline{\mathbb{N}})$, i.e.

$$X_n = M_n + A_n \quad (n \in \overline{\mathbb{N}}).$$

(2) Mas aún, la condición $\sup\limits_{\overline{\mathbb{N}}} E(X_n^+) < \infty$ (que es suficiente para asegurar la c.s. convergencia de la submartingala) es equivalente a la conjunción de las dos condiciones:

$$\sup_{\overline{\mathbb{N}}} E(|M_n|) < \infty \quad \text{y} \quad A_\infty \in \mathcal{L}_1.$$

(3) Entre tanto la convergencia en $\mathcal{L}_1$ de la submartingala $(X_n, n \in \overline{\mathbb{N}})$ es equivalente a la regularidad de la martingala $(M_n, n \in \overline{\mathbb{N}})$ junto con la condición $A_\infty \in \mathcal{L}_1$.
(4) Para todo tiempo de paro ν regular para la martingala $(M_n, n \in \overline{\mathbb{N}})$ la v.a. X_ν es integrable si y sólo si $E(A_\nu) < \infty$, y entonces $E(X_\nu) = E(M_0) + E(A_\nu)$.

Demostración:

La demostración se hará por pasos.
(1) Las v.a. M_n y A_n $(n \in \overline{\mathbb{N}})$ se pueden definir a través de sus diferencias por las fórmulas

$$(a) \quad M_0 = X_0 \qquad M_{n+1} - M_n = X_{n+1} - E(X_{n+1}|\mathcal{B}_n)$$

$$(b) \quad A_0 = 0 \qquad A_{n+1} - A_n = E(X_{n+1}|\mathcal{B}_n) - X_n$$

de donde podemos observar que

$$\begin{aligned}(a) + (b) \Rightarrow X_0 &= M_0 + A_0\\ (M_{n+1} - M_n) + (A_{n+1} - A_n) &= X_{n+1} - E(X_{n+1}|\mathcal{B}_n) + E(X_{n+1}|\mathcal{B}_n) - X_n\\ \Rightarrow (M_{n+1} + A_{n+1}) - (M_n + A_n) &= X_{n+1} - X_n\end{aligned}$$

p.d. $(M_n, n \in \overline{\mathbb{N}})$ es una martingala integrable
i.e. p.d. i)Se trata de una sucesión adaptada

Demostración:
Hay que probar que M_n son $\mathcal{B}_n$-medibles
Por como los definimos $M_{n+1} = X_{n+1} - E(X_{n+1}|\mathcal{B}_n) + M_n$
por definición $X_0 = M_0$ y por lo tanto M_0 es $\mathcal{B}_0$-medible.
Vamos a utilizar un proceso inductivo

$$M_1 = X_1 - E(X_1|\mathcal{B}_0) + M_0$$

pero X_1 es $\mathcal{B}_1$-medible y $E(X_1|\mathcal{B}_0)$ y M_0 son $\mathcal{B}_0$-medibles, y por tanto $\mathcal{B}_1$-medibles.
por tanto M_1 es $\mathcal{B}_1$-medible.
Supongamos que M_n es $\mathcal{B}_n$-medible

$$M_{n+1} = X_{n+1} - E(X_{n+1}|\mathcal{B}_n) + M_n$$

donde X_{n+1} es $\mathcal{B}_{n+1}$-medible y $E(X_{n+1}|\mathcal{B}_n)$ y M_n son $\mathcal{B}_n$-medibles
y por tanto M_{n+1} es $\mathcal{B}_{n+1}$-medible.
Por lo tanto M_n es $\mathcal{B}_n$-medible $\quad \forall n$.
p.d. ii) $M_n = E(M_{n+1}|\mathcal{B}_n) \quad \forall n \in \overline{\mathbb{N}}$

Demostración:
Tenemos que

$$M_n = -X_{n+1} + E(X_{n+1}|\mathcal{B}_n) + M_{n+1}$$

como M_n es $\mathcal{B}_n$-medible

$$\begin{aligned}M_n = E(M_n|\mathcal{B}_n) &= E[-X_{n+1} + E(X_{n+1}|\mathcal{B}_n) + M_{n+1}|\mathcal{B}_n]\\ &= E(-X_{n+1}|\mathcal{B}_n) + E(X_{n+1}|\mathcal{B}_n) + E(M_{n+1}|\mathcal{B}_n) = E(M_{n+1}|\mathcal{B}_n)\end{aligned}$$

por lo tanto $(M_n, n \in \overline{\mathbb{N}})$ es una martingala integrable.
p.d. $(A_n, n \in \overline{\mathbb{N}})$ es un proceso creciente

i.e. p.d. i) $(A_n, n \in \overline{N})$ es predecible.
Recurrimos nuevamente a un proceso inductivo
por definición $A_0 = 0$ y por lo que A_0 es $\mathcal{B}_0$-medible.

$$A_1 = E(X_1|\mathcal{B}_0) - X_0 + A_0$$

como todos los elementos del lado derecho de la igualdad son $\mathcal{B}_0$-medibles, podemos concluir que A_1 es $\mathcal{B}_0$-medible
Supongamos que A_n es $\mathcal{B}_{n-1}$-medible $\Rightarrow A_n$ es $\mathcal{B}_n$-medible y por lo tanto

$$A_{n+1} = E(X_{n+1}|\mathcal{B}_n) - X_n + A_n \quad \text{es} \mathcal{B}_n\text{-medible}$$

Por lo tanto $(A_n, n \in \overline{N})$ es predecible.
p.d. ii) $0 = A_0 \leq A_1 \leq \ldots$
Para ello basta probar que $A_{n+1} - A_n = E(X_{n+1}|\mathcal{B}_n) - X_n \geq 0$, pero esto es inmediato pues por ser $(X_n, n \in \overline{N})$ una submartingala, $E(X_{n+1}|\mathcal{B}_n) \geq X_n$ y por lo tanto $A_{n+1} - A_n \geq 0$.
Por lo tanto $(A_n, n \in \overline{N})$ es un proceso creciente.
Como $M_0 + A_0 = X_0$ y los incrementos de ambas sucesiones $(M_n + A_n, n \in \overline{N})$ y $(X_n, n \in \overline{N})$ coinciden, es claro que

$$X_n = M_n + A_n \quad \forall n \in \overline{N}.$$

Vamos ahora a establecer la unicidad de la descomposición de Doob.
Supongamos que $X_n = M'_n + A'_n (n \in \overline{N})$ es una descomposición de la submartingala $(X_n, n \in \overline{N})$ como la suma de una martingala $(M'_n, n \in \overline{N})$ y un proceso creciente $(A'_n, n \in \overline{N})$, entonces la igualdad de los incrementos

$$A'_{n+1} - A'_n = (X_{n+1} - X_n) - (M'_{n+1} - M'_n) \quad n \in \overline{N}$$

implica, tomando la esperanza condicional dodo $\mathcal{B}_n$ de ambos lados

$$E(A'_{n+1} - A'_n|\mathcal{B}_n) = E(A'_{n+1}|\mathcal{B}_n) - E(A'_n|\mathcal{B}_n)$$

y por ser $(A'_n, n \in \overline{N})$ un proceso creciente

$$= A'_{n+1} - A'_n = E(X_{n+1}|\mathcal{B}_n) - E(X_n|\mathcal{B}_n) - E(M'_{n+1}|\mathcal{B}_n) + E(M'_n|\mathcal{B}_n)$$

y por ser X_n, M'_n $\mathcal{B}_n$-medibles y $(M'_n, n \in \overline{N})$ ser una martingala

$$E(X_{n+1}|\mathcal{B}_n) - X_n - M'_n + M'_n$$

por tanto $A'_{n+1} - A'_n = E(X_{n+1}|\mathcal{B}_n) - X_n = A_{n+1} - A_n$
$\Rightarrow A'_n = A_n \quad \forall n \in \overline{N}$ pues $A'_0 = A_0$
y por tanto $M'_n = X_n - A'_n = X_n - A_n = M_n$

Por lo tanto la descomposición es única.
De esta manera queda probado el punto (1) del teorema.

(2) Probaremos la equivalencia en dos pasos.
a) p.d. $\sup_{\mathbb{N}} E(X_n^+) < \infty \Rightarrow \sup_{\mathbb{N}} E(|M_n|) < \infty$ y $A_\infty \in \mathcal{L}_1$
i) Como $X_n = M_n + A_n$ y $A_n \geq 0 \quad \forall n \Rightarrow M_n \leq X_n \quad \forall n$
pero $M_n^+ = \max\{M_n, 0\}$ y $X_n^+ = \max\{X_n, 0\} \Rightarrow X_n^+ \geq M_n^+ \quad \forall n$

$$\Rightarrow \sup_{\mathbb{N}} E(M_n^+) \leq \sup_{\mathbb{N}} E(X_n^+) < \infty$$

pero $|M_n| = M_n^+ + M_n^- = 2M_n^+ - M_n$

$$\Rightarrow \sup_{\mathbb{N}} E(|M_n|) \leq 2\sup_{\mathbb{N}} E(M_n^+) - \sup_{\mathbb{N}} E(M_n) = 2\sup_{\mathbb{N}} E(M_n^+) - E(M_0)$$

pero $E(M_0) < \infty$ pues M_n es martingala y por lo tanto

$$\sup_{\mathbb{N}} E(|M_n|) < \infty$$

ii) Como $A_n = X_n - M_n \leq X_n^+ - M_n (n \in \mathbb{N})$

$$\Rightarrow E(A_\infty) \leq \sup_{\mathbb{N}} E(X_n^+) - E(M_0)$$

ya que $E(M_n) = E(M_0) \quad \forall n \in \mathbb{N}$

$$\Rightarrow E(A_\infty) \quad \text{por lo tanto } A_\infty \in \mathcal{L}_1$$

b)p.d. $\sup_{\mathbb{N}} E(|M_n|) < \infty$ y $A_\infty \in \mathcal{L}_1 \Rightarrow \sup_{\mathbb{N}} E(X_n^+) < \infty$
como $X_n = M_n + A_n \Rightarrow X_n^+ = (M_n + A_n)^+ \leq M_n^+ + A_n \quad \forall n$ pues $A_n \geq 0$
y por lo tanto $\sup_{\mathbb{N}} E(X_n^+) \leq \sup_{\mathbb{N}} E(M_n^+) + E(A_\infty)$ ya que A_n es un proceso creciente
como $|M_n| = M_n^+ + M_n^- \Rightarrow$ si $\sup_{\mathbb{N}} E(|M_n|) < \infty \Rightarrow \sup_{\mathbb{N}} E(M_n^+) < \infty$
y como $A_\infty \in \mathcal{L}_1 \Rightarrow E(A_\infty) < \infty$
por lo tanto $\sup_{\mathbb{N}} E(X_n^+) \leq \sup_{\mathbb{N}} E(M_n^+) + E(A_\infty) < \infty$
De esta manera queda probado el punto (2) de la proposición.

(3) $\Leftarrow$) Si $A_\infty \in \mathcal{L}_1$, el teorema de la convergencia dominada de Lebesgue nos dice que $A_n \longrightarrow A_\infty$ en $\mathcal{L}_1$ cuando $n \longrightarrow \infty$
De igual manera la regularidad de la martingala $(M_n,\ n \in \overline{\mathbb{N}})$ nos dice que ésta converge en $\mathcal{L}_1$. Por lo tanto podemos concluír que la submartingala $(X_n, n \in \overline{\mathbb{N}})$ converge en $\mathcal{L}_1$ pues

$$X_n = M_n + A_n$$

$\Leftarrow$) Si la submartingala $(X_n, n \in \overline{\mathbb{N}})$ converge en $\mathcal{L}_1$ a una v.a. X_∞

$$\Rightarrow E(A_\infty) = \lim_{n\uparrow\infty} E(A_n) = \lim_{n\uparrow\infty}\big(E(X_n) - E(M_n)\big) = E(X_\infty) - E(M_0) < +\infty$$

esto prueba que $A_\infty \in \mathcal{L}_1$ y por ello que $A_n \longrightarrow A_\infty$ en $\mathcal{L}_1$.
pero entonces la martingala $(M_n = X_n - A_n, n \in \overline{\mathbb{N}})$ también converge en $\mathcal{L}_1$.
Así hemos probado el punto (3)

(4) Si ν es un tiempo de paro regular para la martingala $(M_n, n \in \overline{\mathbb{N}})$, la v.a. $M_\nu (= \lim_{n\to\infty} M_n$ *c.s.* en$\{\nu = \infty\})$ es integrable, y por un teorema probado anteriormente satisface $E(M_\nu) = E(M_0) = E(X_0)$.
Por tanto, la v.a. $X_\nu (= \lim_{n\to\infty} X_n$ *c.s.* en$\{\nu = \infty\})$ existe y es igual a $M_\nu + A_\nu$ *c.s.*;
por ello X_ν es integrable si y sólo si A_ν lo es, y cuando esto sucede, satisface $E(X_\nu) = E(M_\nu) + E(A_\nu)$.
Esto termina la prueba del teorema. $\square$

BIBLIOGRAFIA

1.- *ASH, Robert B.* Real Analysis and Probability *Academic Press, Inc., 1974.*

2.- *BILLINGSLEY, Patrick.* Probability and Measure, *Second Edition, John Wiley & Sons, 1986.*

3.- *BREIMAN, Leo.* Probability, *Addison-Wesley Publishing Company, 1968.*

4.- *BREIMAN, Leo.* Probability and Stochastic Processes. *Second Edition, The scientific Press, 1986.*

5.- *BOJDECKI, Tomasz.* Qué es y para qué sirve una martingala?, *Ciencia 36, 59-65 (1985).*

6.- *CHEN MOY, Shu-Teh.* Characterizations of Conditional expectation as a transformations on Functions Spaces, *Pacific J. Math, 4, 47-64 (1954).*

7.- *CHUNG, Kai Lai.* A course in Probability Theory, *Second Edition, Academic Press, Inc. 1974.*

8.- *DOOB, J. L.*, Stochastic Processes, *John Wiley & Sons, 1953.*

9.- *DOOB, J. L.*, What is a Martingale. *American Mathematical Review. May (1971), 451-463.*

10.- *KOLMOGOROV, A. N.* Foundations of the Theory of Probability, *Second English edition, Chelsea Publishing Co., 1956.*

11.- *LOÈVE, Michel.* Probability Theory, *Third Edition, Van Nostrand Reinhold Co., 1963.*

12.- *MEYER, Paul A.* Probability and Potentials, *Blaisdell Publishing Company, 1966.*

13.- *NEVEU, Jacques.* Discrete-Parameter Martingales, *North Holland Publishing Co., 1975.*

14.- *NEVEU, Jacques.* Mathematical Foundations of the Calculus of Probability, *Holden-Day Inc., 1965.*

15.- *RAO, M. M.*, Stochastic Processes and Integration, *Sijhtoff & Noordhoff, 1979.*

16.- *SHILOV, G. E. & GUREVICH, B. L.*, Integral, Measure & Derivative: A Unified Aproach., *Dover Publications, Inc. 1977.*

17.- *STOYANOV, Jordan M.* Counterexamples in Probability, *John Wiley & Sons, 1987.*