FACULTAD DE MATEMÁTICAS
PONTIFICIA UNIVERSIDAD
CATÓLICA DE CHILE

Pontificada Universidad Católica

MAT2805

Teoría de Probabilidad

Autor:
Sebastián Lepe V.

2 de julio de 2025

Índice

1. Fundamentos de la Teoría de la Probabilidad

1.1. Introducción

La teoría de probabilidad es el estudio de la probabilidad de forma formal. Por lo que incluye el estudio de la teoría de la medida aplicada a la probabilidad.

Por otro lado, la probabilidad es una parte importante en el estudio de distintas cosas, ya que nos dice que tan probable es que ocurra algo.

Veamos como la teoría de medida se aplica en la probabilidad.

Definición (Número normal): *Sea $x \in \mathbb{R}$, diremos que es normal en base 10 si la proporción de los números $\{0, 1 \ldots, 9\}$ es de $\frac{1}{10}$, dicho de otro forma, sea $i \in \{0, 1 \ldots, 9\}$,*

$$\frac{\#\{\textit{aparición de } i \textit{ en las primeras } n \textit{ decimales de } x\}}{n} \xrightarrow{n\to\infty} \frac{1}{10}$$

De forma análoga podemos definir un número normal en base q con $q \in \mathbb{N}$. Ahora un número normal es un número tal que es normal en todas las bases.

Es decir, un número es normal si lo es en toda base al mismo tiempo. Lo interesante es que la cantindad de números normales no es despreciable, al contrário, estudiemos el siguiente resultado:

Teorema (Teorema del número normal de Borel): *El conjunto de los números no normales, tiene medida Lebesgue 0.*

Antes de estudiar el teorema, necesitamos la siguiente convenencia:

Nota: Si $x \in \mathbb{R}$ tiene dos o más expansiones decimales, siempre tomaremos la infinita (menos la que termina en 0's).

Primero, demostremos el siguiente resultado:

Lema: *Se cumple lo siguiente:*

$$\lambda(\{x \in [0, 1] : x \textit{ no es normal en base } 2\}) = 0$$

Pensemos en la base 2, entonces $x \in [0, 1]$ se puede escribir de la siguiente manera:

$$x = 0, d_1 d_2 \ldots$$

donde $d_i \in \{0, 1\}$. Entonces podemos pensar en x como la siguiente serie:

$$x = \sum_{i=1}^{\infty} \frac{d_i}{2^i}$$

Ahora vamos a considerar de la siguiente forma d_i:

$$d_i : [0, 1] \to \{0, 1\}$$

donde el primer d_1 se expresa como una función por partes simple, en particular,

$$d_1(t) = \begin{cases} 1 & t \in (1/2, 1] \\ 0 & t \in [0, 1/2] \end{cases}$$

el segundo d_2 es tomar los intervalos $[0, 1/2], (1/2, 1]$ y dividir cada uno en dos partes, obteniendo cuatro intervalos, siguiendo una forma similar a d_1, obteniendo,

$$d_2(t) = \begin{cases} 1 & t \in (3/4, 1] \\ 0 & t \in (1/2, 3/4] \\ 1 & t \in (1/4, 1/2] \\ 0 & t \in [0, 1/4] \end{cases}$$

esta construcción se replica para cada d_i. Digamos que $\mathbb{P}(A)$ es una "probabilidad" que representa $\lambda(A)$, donde $A \subseteq [0, 1]$ es un boreleano por lo que tiene medida. Notemos que,

$$\begin{aligned} &\mathbb{P}(\{x \in [0, 1] : d_1(x) = 1, d_2(x) = 0, d_3(x) = 1, d_4(x) = 1\}) \\ &= \lambda(\{x \in [0, 1] : d_1(x) = 1, d_2(x) = 0, d_3(x) = 1, d_4(x) = 1\}) \end{aligned}$$

Para determinar el resultado pensemos en lo siguiente, d_1 toma valor 1 en el intervalo $(1/2, 1]$ que en medida de Lebesgue tiene medida $1/2$. Pero queremos el intervalo donde $d_2 = 0$ y eso es en el intervalo $(1/2, 3/4]$, siguiendo así, obtenemos un intervalo de la forma $(a, b]$ donde $d_1 = 1, d_2 = 0, d_3 = 1, d_4 = 1$ y si somos cuidadosos, vemos que $\lambda(a, b] = 1/16$, siendo este el valor de la "probabilidad".

Lo que acabamos de hacer es asociar un problema de probabilidad con medida de Lebesgue. Ahora antes de seguir veamos la ley débil de los grandes números.

Ley débil de los grandes números: *Sea $\{X_i\}_{i\in\mathbb{N}}$ una colección de variables aleatorias iid (independientes e identicamente distribuidas). Supongamos que $\mathbb{E}(X_i) = \mu$, entonces para todo $\varepsilon > 0$ se tiene que,*

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \overset{n\to\infty}{\longrightarrow} 0$$

En el contexto del teorema de Borel, la media es $\mu = 1/2$, las variables aleatorias son los d_i y por último, la suma parcial de los primeros n d_i se puede pensar como:

$$S_n(x) = \sum_{i=1}^{n} d_i(x) = \#\{\text{Cantidad de veces que aparece 1 en los primeros números dígitos binario}\}$$

Observación: La interpretación de los $\{d_i\}$ como resultados de tirar una moneda al azar, inmediatamente nos permite trabajar con un espacio muestral en "∞" tiradas independientes.

En virtud de la ley débil de los grandes números, se tiene que el sumado S_n/n se va pareciendo a $1/2 = \mu$ cuando $n \to \infty$, pero no están directo, aunque podemos sacar información. Para ver el resultado de forma más directo se debe estudiar la ley fuerte de los grandes números, la cual en nuestro contexto nos dice que,

$$\mathbb{P}\left(\frac{S_n}{n} \to \frac{1}{2}\right) = 1$$

que se puede pensar como:

$$\lambda\left(\left\{x \in [0,1] : \sum_{i=1}^{n} \frac{d_i(x)}{n} \to \frac{1}{2}\right\}\right) = 1$$

Por tando, asumiendo que se cumple la ley fuerte de los grandes números, podemos demostrar que la medida de los números no normales en base 2 en el intervalo $[0,1]$, tiene medida 0 (puesto que son aquellos números donde el promedio de la suma de los primeros dígitos no tiende a $1/2$). Ahora, sabemos que $\mathbb{R}$ se puede expresar como unión numerable de intervalos de la forma $[k, k+1]$ con $k \in \mathbb{Z}$, por tanto, la colección de los números no normales en base 2, tiene medida 0.

Observación: El resultado previo se puede extender para toda base q.

Ahora vamos a demostrar algunos resultados sin saber mucho de probabilidad, sino con la teoría de medida.

Teorema: *Se cumple que:*

$$\lambda\left(\left\{x \in [0,1] : \left|\sum_{i=1}^{n} \frac{d_i(x)}{n} - \frac{1}{2}\right| \geq \varepsilon\right\}\right) \overset{n\to\infty}{\longrightarrow} 0$$

y también que:

$$\lambda\left(\left\{x \in [0,1] : \sum_{i=1}^{n} \frac{d_i(x)}{n} \to \frac{1}{2}\right\}\right) = 1$$

Dem: Probemos la primera afirmación. Consideremos la función $r_i(x) := 2d_i(x) - 1$, el cual por construcción toma los valores $\{-1, 1\}$. Por lo que basta demostrar el siguiente límite:

$$\lambda\left(\left\{x \in [0,1] : \left|\sum_{i=1}^{n} \frac{r_i(x)}{n}\right| \geq \varepsilon\right\}\right) \overset{n\to\infty}{\longrightarrow} 0$$

Y esto es debido a que,

$$\sum_{i=1}^{n} \frac{r_i(x)}{n} = 2\left(\sum_{i=1}^{n} \frac{d_i(x)}{n} - \frac{1}{2}\right)$$

Calculemos la integral de Lebesgue de $r_i(x)$, notemos que es Riemann integrable, por lo que está bien definida como integral de Lebesgue. Con esto en mente, tenemos que,

$$\int_0^1 r_i(x)dx = 0$$

para todo $i \in \mathbb{N}$. Para ver esto notemos que r_i divide al intervalo $[0,1]$ en 2^i trozos, viendo que la integral se anula.

Estudiemos $r_i(x)r_j(x)$, supongamos que $i > j$, por lo que de alguna forma r_i está "contenido" en r_j, en el sentido que en un segmento de largo $1/2^j$ este está separado por cierta cantidad de intervalos de largo $1/2^i$, es más, hay 2^{i-j} intervalos de largo $1/2^i$. Supongamos que $x \in [0, 1/2^i]$, entonces $r_j(x) = -1$ y luego obtenemos que $r_i(x)r_j(x) = -r_i(x)$, por otro lado notemos que $r_i(x)$ sigue la siguiente distribución:

$$r_i(x) = \begin{cases} 1, & x \in \left(\frac{2k-1}{2^i}, \frac{2k}{2^i}\right] \\ -1, & x \in \left[0, \frac{1}{2^i}\right] ; \left[\frac{2k}{2^i}, \frac{2k+1}{2^i}\right] \end{cases}$$

donde $k = 1, \ldots, 2^{i-j-1}$. Lo importante es que en el primer intervalo (en $[0, 1/2^i]$) $-r_i(x)$ toma valor 1, mientras que en el "último intervalo", el cual sería cuando $x \in [(2^{i-j}-1)/2^i, 1/2^j]$, toma valor $r_i(x) = -1$, y sabiendo que los $1, -1$ van intercalando, se tiene claramente que,

$$0 = \int_0^{1/2^j} -r_i(x)dx = \int_0^{1/2^j} r_i(x)r_j(x)dx$$

Finalmente realizando este estudio para todo intervalo de largo $1/2^j$ se concluye que,

$$\int_0^1 r_i(x)r_j(x)dx = 0$$

cuando $j \neq i$. Si $i = j$ entonces basta observar que $r_i^2(x) = 1$ para todo $x \in [0,1]$, por tanto,

$$\int_0^1 r_i^2(x)dx = 1$$

Definimos la n-ésima suma parcial de los r_i por:

$$S_n(x) := \sum_{i=1}^{n}$$

Demostremos que,

$$\lambda\left(\left\{x \in [0,1] : \left|\frac{S_n(x)}{n}\right| \geq \varepsilon\right\}\right) \overset{n\to\infty}{\longrightarrow} 0$$

para todo $\varepsilon > 0$. Para ello vamos a necesitar el siguiente resultado:

Teorema: *Sea $f : [0,1] \to [0,\infty)$ una función medible, entonces se cumple:*

$$\lambda\{x : f(x) \geq a\} \leq \frac{1}{a}\int_0^1 f(x)dx$$

para $a > 0$.

Dem: Suponiendo que está bien definido la integral de Lebesgue, se tiene que,

$$\begin{aligned}\int_0^1 f(x)dx &= \int_{\{x:f(x)\geq a\}} f(x)dx + \int_{\{x:f(x)<a\}} f(x)dx \\ &\geq \int_{x:f(x)\geq a\}} f(x)dx \\ &= \int_{\{x:f(x)\geq a\}} adx = a\lambda\{x : f(x) \geq a\}\end{aligned}$$

dado que $a > 0$ podemos despejar y obtener el resultado:

$$\lambda\{x : f(x) \geq a\} \leq \frac{1}{a}\int_0^1 f(x)dx$$

■

Volviendo a la demostración principal. Observemos lo siguiente:

$$\lambda\left(\left\{x \in [0,1] : \left|\frac{S_n(x)}{n}\right| \geq \varepsilon\right\}\right) = \lambda\left(\left\{x \in [0,1] : S_n^2(x) \geq n^2\varepsilon^2\right\}\right) \leq \frac{1}{n^2\varepsilon^2}\int_0^1 S_n^2(x)dx$$

Por otro lado se tiene que,

$$S_n^2(x) = \sum_{i,j=1,\ldots,n} r_i(x)r_j(x)$$

De esta forma,

$$\lambda\left(\left\{x \in [0,1] : \left|\frac{S_n(x)}{n}\right| \geq \varepsilon\right\}\right) \leq \frac{1}{n^2\varepsilon^2}\int_0^1 \sum_{i,j=1,\ldots,n} r_i(x)r_j(x) = \frac{1}{n\varepsilon^2} \overset{n\to\infty}{\longrightarrow} 0$$

para todo $\varepsilon > 0$. Nos falta probar la otra parte. Notemos que,

$$\lambda\left(\left\{x \in [0,1] : \lim_{n\to\infty} \frac{S_n(x)}{n} = 0\right\}\right) = 1$$

es equivalente a demostrar que,

$$\lambda\left(\left\{x \in [0,1] : \sum_{i=1}^{n} \frac{d_i(x)}{n} \to \frac{1}{2}\right\}\right) = 1$$

Para ello usaremos el mismo truco de elevar al cuadrado, pero en vez, elevaremos a la cuarte, obteniendo la siguiente desigualdad:

$$\lambda\left(\left\{x \in [0,1] : \left|\frac{S_n(x)}{n}\right| \geq \varepsilon\right\}\right) \leq \frac{1}{n^4\varepsilon^4}\int_0^1 S_n^4(x)dx$$

donde,

$$S_n^4(x) = \sum_{i,j,s,l=1,\ldots,n} r_i(x) r_j(x) r_s(x) r_l(x)$$

falta

Por tanto,

$$\lambda\left(\left\{x \in [0,1] : \left|\frac{S_n(x)}{n}\right| \geq \varepsilon\right\}\right) \leq \frac{3}{n^2\varepsilon^4} \overset{n\to\infty}{\longrightarrow} 0$$

Para concluir el resultado usaremos otro teorema.

Teorema (Borel-Cantelli): *Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida y sea $\{A_n\}_{n\in\mathbb{N}}$ una secuencia de medibles, tales que,*

$$\sum_{n=1}^{\infty} \mu(A_n) < \infty$$

Entonces, $\mu(\limsup_{n\to\infty} A_n) = 0$.

Dem: por hacer

Sea,

$$A_n := \left\{x \in [0,1] : \left|\frac{S_n(x)}{n}\right| \geq \varepsilon_n\right\}$$

donde más adelante escogeremos ε_n. Se cumple para todo n que,

$$\lambda(A_n) \leq \frac{3}{n^2\varepsilon^4}$$

Entonces,

$$\sum_{n=1}^{\infty} \lambda(A_n) < \infty$$

y por el teorema anterior se tiene que,

$$\lambda\left(\limsup_{n\to\infty} A_n\right) = 0$$

Digamos que $\varepsilon_n = 1/n^\alpha$, entonces,

$$\lambda(A_n) \leq \frac{1}{n^{2-4\alpha}}$$

si tomamos $\alpha > 1/2$ se tiene que la serie de los $\lambda(A_n)$ converge y luego el límite superior de los A_n tiene medida 0. Esto lo podemos pensar de la siguente forma:

$$\lambda\left(\left\{x \in [0,1] : \left|\frac{S_n(x)}{n}\right| \geq \frac{1}{n^{1/8}} \text{ para infinitos valores de } n\right\}\right) = 0$$

Por lo tanto,

$$\lambda\left(\left\{x \in [0,1] : \left|\frac{S_n(x)}{n}\right| < \frac{1}{n^{1/8}} \text{ para todo } n \text{ suficientemente grande}\right\}\right) = 1$$

Que es equivalente a decir,

$$\lambda\left(\left\{x \in [0,1] : \lim_{n\to\infty} \frac{S_n(x)}{n} = 0\right\}\right) = 1$$

Probando el teorema. ■

1.2. Fundamentos

Empezaremos con la formalización de la teoría de probabilidad definiendo un σ-álgebra hasta conceptos generales de probabilidad.

Definición 1.1 (σ-álgebra y Medida): *Sea Ω un conjunto no vacío. Sea $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ una colección de subconjunto de Ω. Decimos que $\mathcal{F}$ es un σ-álgebra si cumple las siguientes propiedades:*

(a) $\emptyset \in \mathcal{F}$,

(b) Si $A \in \mathcal{F}$, *entonces* $A^c \in \mathcal{F}$,

(c) Si $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$, *entonces,*

$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$$

Al par $(\Omega, \mathcal{F})$ le decimos espacio medible. Definimos una medida sobre el espacio de medible $(\Omega, \mathcal{F})$ como una función $\mu : \mathcal{F} \to [0, \infty]$ tal que cumple las siguientes propiedades:

(a) $\mu(\emptyset) = 0$,

(b) Si $A \subseteq B$ *(con* $A, B \in \mathcal{F}$*), entonces* $\mu(A) \leq \mu(B)$,

(c) Si $\{A_n\}_{n\in\mathbb{N}}$ *es una secuencia de eventos (elementos del σ-álgebra) disjuntos, entonces,*

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$$

La tupla $(\Omega, \mathcal{F}, \mu)$ le decimos espacio de medida. Cuando $\mu(\Omega) = 1$, entonces diremos que es un espacio de probabilidad. En tal caso denotaremos $\mathbb{P}$ en vez de μ.

Ejemplo 1.2: Consideremos el experimento de lanzar una moneda, entonces $\Omega = \{C, S\}$, nuestro σ-álgebra puede ser $\mathcal{F} = \mathcal{P}(\Omega)$ con una medida tal que,

$$\mu(\{C\}) = \frac{1}{2}$$
$$\mu(\{S\}) = \frac{1}{2}$$

Siendo un caso de una probabilidad equiprobable (cuando la probabilidad de cada punto es igual).

Ejemplo 1.3: En general podemos tomar $\Omega = \{x_i : i \in \mathbb{N}\}$ un conjunto numerable (o finito) con σ-álgebra $\mathcal{F} = \mathcal{P}(\Omega)$, en este caso decimos que $(\Omega, \mathcal{F})$ es un espacio discreto. Para definir una probabilidad consideramos la sucesión $\{p_i\}_{i\in\mathbb{N}}$ de números no negativos tales que,

$$\sum_{n\in\mathbb{N}} p_i = 1$$

Podemos definir la siguiente medida de probabilidad:

$$\mathbb{P}(A) = \sum_{i:x_i \in A} p_i$$

para todo $A \in \mathcal{P}(\Omega)$. Probemos que es una medida. Para ello debemos probar los axiomas de medida de probabilidad.

(a) $\mathbb{P}(\emptyset) = 0$ por definición,

(b) Sea $A \subseteq B$, entonces,

$$\mathbb{P}(B) = \sum_{i:x_i \in B} p_i = \underbrace{\sum_{i:x_i \in A} p_i}_{\geq 0} + \underbrace{\sum_{i:x_i \in B \setminus A} p_i}_{\geq 0} \geq \sum_{i:x_i \in A} p_i = \mathbb{P}(A)$$

(c) Sea $\{A_n\}_{n\in\mathbb{N}}$ una secuencia de eventos disjuntos, entonces,

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{i:x_i \in \bigcup_n A_n} p_i = \sum_{n\in\mathbb{N}} \left(\sum_{i:x_i \in A_n} p_i\right) = \sum_{n\in\mathbb{N}} \mathbb{P}(A_n)$$

Finalmente $\mathbb{P}(\Omega) = 1$. Siendo, en efecto, $\mathbb{P}$ una medida de probabilidad.

Nota 1.4: Notemos que si Ω es un conjunto finito, entonces podemos definir una probabilidad equiprobable, sin embargo, este ejercicio no se puede replicar cuando Ω es numerable infinito por obvias razones. Por lo que el ejemplo 1.3 define bien la probabilidad sobre un conjunto numerable infinito.

Ejemplo 1.5: Sea $\Omega = [0,1]$, sea $\mathcal{F} = \sigma(\{O : O \subseteq [0,1], \text{ abierto}\}) =: \mathcal{B}[0,1]$ y sea $\mathbb{P} = \lambda$ (la medida de Lebesgue), entonces $([0,1], \mathcal{B}[0,1], \lambda)$ es un espacio de probabilidad. Verifiquemos este hecho.

Recordemos que en general la medida de Lebesgue se define como la función $\lambda : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ donde $\mathcal{B}(\mathbb{R})$ es subconjunto de los λ-medibles, dada por:

$$\lambda(A) := \text{ínf}\left\{\sum_{n\in\mathbb{N}} |I_n| : A \subseteq \bigcup_{n\in\mathbb{N}} I_n; \{I_n\}_{n\in\mathbb{N}} \subseteq \mathcal{B}(\mathbb{R})\right\}$$

Debemos probar los axiomas de espacio de probabilidad.

(a) $\lambda(\emptyset) = 0$ por definición,

(b) Sean $A, B \in \mathcal{B}[0,1]$ tal que $A \subseteq B$. Consideremos la colección $\{I_n\}_{n\in\mathbb{N}} \subseteq \mathcal{B}[0,1]$ un cubrimiento de B, entonces también es un cubrimiento de A y luego,

$$\lambda(A) \leq \sum_{n\in\mathbb{N}} |I_n|$$

Esto se puede hacer para todo cubrimiento de B, por lo que $\lambda(A)$ es cota inferior de la colección,

$$\left\{\sum_{n\in\mathbb{N}} |I_n| : A \subseteq \bigcup_{n\in\mathbb{N}} I_n; \{I_n\}_{n\in\mathbb{N}} \subseteq \mathcal{B}[0,1]\right\}$$

Por tanto,

$$\lambda(A) \leq \lambda(B)$$

(c) Sea $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{B}[0,1]$ una colección disjunta. Hay un resultado que nos dice que: *Si tenemos una colección disjuntas de conjuntos medibles sobre una medida, entonces la medida de unión numerable es igual a la suma numerable de cada medida del elemento de la colección.* Por tanto,

$$\lambda\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \lambda(A_n)$$

Por último observemos que,

$$\lambda([0,1]) = 1$$

Probando que λ es una medida de probabilidad.

Nota 1.6: En teoría de medida podemos definir el producto de dos espacios de probabilidad mediante el teorema de extensión de Caratheodory, debido a que dados los espacios de medida $(\Omega_1, \mathcal{F}_1, \mu_1), (\Omega_2, \mathcal{F}_2, \mu_2)$ podemos considerar el conjunto $\Omega = \Omega_1 \times \Omega_2$ con σ-álgebra,

$$\mathcal{F} = \sigma(\{A \times B : A \in \mathcal{F}_1, B \in \mathcal{F}_2\})$$

con una medida con la condición:

$$\mu(A \times B) = \mu(A)\mu(B)$$

al extender por el teorema de Caratheodory obtenemos una medida sobre $(\Omega, \mathcal{F})$ obteniendo un espacio de medida sobre el producto y por tanto $(\Omega, \mathcal{F}, \mu)$ es un espacio de medida. Además, si $(\Omega_1, \mathcal{F}_1, \mu_1), (\Omega_2, \mathcal{F}_2, \mu_2)$ son espacios de probabilidad, entonces el espacio producto también es de probabilidad. Estando todo bien definido.

Definición 1.7: *Sea E un espacio métrico o topológico. Definimos los boreleanos como el σ-álgebra generado por los abiertos de E y lo denotaremos por $\mathcal{B}(E)$. Los boreleanos de $\mathbb{R}^n$ lo denotaremos por $\mathcal{R}^n$.*

Definición 1.8: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Una variable aleatoria es una función $\mathcal{F}|\mathcal{R}$-medible,*

$$X : \Omega \to \mathbb{R}$$

donde $\mathbb{R}$ *está dotado por el espacio medible* $(\mathbb{R}, \mathcal{R})$.

Ejemplo 1.9: Consideremos un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Sea $A \in \mathcal{F}$, entonces definimos la indicatriz de A por:

$$X(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

Entonces X es una variable aleatoria, para ver esto notemos que por definición, para $B \in \mathcal{R}$ se tiene que,

$$\begin{aligned} X^{-1}(B) &= \{\omega \in \Omega : X(\omega) \in B\} \\ &= \begin{cases} \emptyset, & 0, 1 \notin B \\ A^c, & 0 \in B, 1 \notin B \\ A, & 0 \notin B, 1 \in B \\ \Omega, & 0, 1 \in B \end{cases} \end{aligned}$$

Por lo que $X^{-1}(B) \in \mathcal{F}$ para todo $B \in \mathcal{R}$. Como variante se escribe la indicatriz de A como:

$$\mathbb{1}_A = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$$

De forma que $\mathbb{1}_A$ es una variable aleatoria si y sólo si $A \in \mathcal{F}$.

Definición 1.10: *Sea* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espacio de probabilidad y sea* $X : \Omega \to \mathbb{R}$ *una variable aleatoria. Definimos la distribución de* X *como la medida de probabilidad* μ_X *del espacio medible* $(\mathbb{R}, \mathcal{R})$ *definida por:*

$$\mu_X(A) = \mathbb{P}(\{\omega : X(w) \in A\})$$

Nota 1.11: En probabilidad tenemos la siguiente conveniencia:

$$\{X \in A\} := \{\omega : X(\omega) \in A\}$$

Es más, escribimos la probabilidad de $\{X \in A\}$ de la siguiente forma:

$$\mathbb{P}(X \in A)$$

para reducir notación.

Observación 1.12: La medida μ_X está bien definida y es una medida de probabilidad. Probemos que es una medida de probabilidad. Para ello debemos probar los axiomas de medida.

(a) Se tiene que,

$$\mu_X(\emptyset) = \mathbb{P}(\{\omega : X(\omega) \in \emptyset\}) = \mathbb{P}(\emptyset) = 0$$

(b) Sea $A \subseteq B$ $(A, B \in \mathcal{R})$, entonces,

$$\{X \in A\} \subseteq \{X \in B\}$$

Y esto implica,

$$\mu_X(A) = \mathbb{P}(X \in A) \leq \mathbb{P}(X \in B) = \mu_X(B)$$

(c) Sea $\{A_n\}_{n\in\mathbb{N}}$ una secuencia de eventos disjuntos, por lo que,

$$\mu_X\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mathbb{P}\left(X \in \bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mathbb{P}(X \in A_n) = \sum_{n\in\mathbb{N}} \mu_X(A_n)$$

Finalmente $\mu_X(\mathbb{R}) = \mathbb{P}(\Omega) = 1$.

Definición 1.13 (Función Distribución): *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sea $X : \Omega \to \mathbb{R}$ una variable aleatoria. Definimos la función:*

$$\begin{aligned} &F_X : \mathbb{R} \to [0, 1] \\ &F_X(x) := \mathbb{P}(\{\omega : X(\omega) \leq x\}) = \mathbb{P}(X \leq x) \end{aligned}$$

como la función distribución acumulada (fda).

Proposición 1.14: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sea $X : \Omega \to \mathbb{R}$ una variable aleatoria. Sea F_X la fda de X, entonces se cumple las siguientes afirmaciones:*

(a) F_X es decreciente,

(b) Se cumplen los límites:

$$\lim_{x\to+\infty} F_X(x) = 1, \quad \lim_{x\to-\infty} F_X(x) = 0$$

(c) F_X es continua por la derecha, es decir,

$$\lim_{x\to x_0^+} F_X(x) = F(x_0)$$

para todo $x_0 \in \mathbb{R}$.

(d) Se cumple que:

$$F_X(x_0^-) := \lim_{x\to x_0^-} F_X(x) = \mathbb{P}(X < x_0)$$

(e) Para $x_0 \in \mathbb{R}$ se cumple que,

$$\mathbb{P}(X = x_0) = F_X(x_0) - F_X(x_0^-)$$

Dem: Consideremos F_X una fda de una variable aleatoria X. Entonces,

(a) Sean $x \leq y$ con $x, y \in \mathbb{R}$, luego $\{X \leq x\} \subseteq \{X \leq y\}$, luego por definición de la medida $\mathbb{P}$, se cumple que,

$$F_X(x) = \mathbb{P}(X \leq x) \leq \mathbb{P}(X \leq y) = F_X(y)$$

(b) Para x arbitrario se tiene que,

$$\begin{aligned}\lim_{x\to+\infty} F_X(x) &= \lim_{x\to+\infty} \mathbb{P}(\{\omega : X(\omega) \leq x\}) \\ &= \lim_{x\to+\infty} \mathbb{P}(\{\omega : X(\omega) \in (-\infty, x]\}) \\ &= \lim_{n\to+\infty} \mathbb{P}\left(\{\omega : X(\omega) \in (-\infty, n]\}\right)\end{aligned}$$

Definimos,

$$A_n := \{\omega : X(\omega) \in (-\infty, n]\}$$

Notemos que la siguiente cadena creciente:

$$A_1 \subseteq A_2 \subseteq \cdots \subseteq A_n \subseteq \ldots$$

Como $\mathbb{P}(\Omega) = 1$, entonces por propiedades de teoría de medida se tiene que,

$$\lim_{x\to+\infty} F_X(x) = \lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \mathbb{P}(\Omega) = 1$$

Para el otro límite debemos notar que podemos definir un conjunto análogo A_n solo que se trabaja en $(-\infty, -n]$ en vez de $(-\infty, n]$, y además, se cumple que,

$$\lim_{x\to-\infty} F_X(x) = \lim_{n\to\infty} \mathbb{P}(\{\omega : X(\omega) \in (-\infty, -n])$$

Como se tiene la cadena decreciente:

$$A_1 \supseteq A_2 \supseteq \cdots \supseteq A_n \supseteq \ldots$$

Entonces,

$$\lim_{x\to-\infty} F_X(x) = \lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \mathbb{P}(\emptyset) = 0$$

Probando los límites.

(c) Sea $x_0 \in \mathbb{R}$ cualquiera. Notemos que,

$$(-\infty, x_0] = \bigcap_{n\in\mathbb{N}} (-\infty, x_n]$$

donde $x_n \downarrow x_0$. Luego,

$$\begin{aligned}\lim_{x\to x_0^+} F_X(x) &= \lim_{x_n\downarrow x_0} \mathbb{P}\left(\{\omega : X(\omega) \in (-\infty, x_n]\}\right)\\ &= \mathbb{P}\left(\bigcap_{n\in\mathbb{N}} \{\omega : X(\omega) \in (-\infty, x_n]\}\right)\\ &= \mathbb{P}(\{\omega : X(\omega) \in (-\infty, x_0]\})\\ &= F_X(x_0)\end{aligned}$$

(d) El procedimiento es análogo al punto anterior, solamente debemos notar que,

$$(-\infty, x_0) = \bigcup_{n\in\mathbb{N}} (-\infty, x_n)$$

donde $x_n \uparrow x$, luego,

$$\begin{aligned}\lim_{x\to x_0^-} F_X(x) &= \lim_{x_n\uparrow x_0} \mathbb{P}(\{\omega : X(\omega) \in (-\infty, x_n))\\ &= \mathbb{P}\left(\bigcup_{n\in\mathbb{N}} \{\omega : X(\omega) \in (-\infty, x_n)\right)\\ &= \mathbb{P}(\{\omega : X(\omega) \in (-\infty, x))\\ &= \mathbb{P}(X < x_0)\end{aligned}$$

(e) Usaremos los puntos anteriores. Queremos probar que,

$$\mathbb{P}(X = x_0) = \mathbb{P}(X \leq x_0) - \mathbb{P}(X < x_0)$$

Notemos que,

$$\mathbb{P}(X = x_0) = 1 - \mathbb{P}(X \neq x_0)$$

usando propiedades de medida y que $\mathbb{P}(\Omega) = 1$. Luego vemos que,

$$\{X \neq x_0\} = \{X < x_0\} \cup \{X > x_0\}$$

Que son disjuntos, entonces,

$$\begin{aligned}\mathbb{P}(X = x_0) &= 1 - (\mathbb{P}(X < x_0) + \mathbb{P}(X > x_0))\\ &= 1 - \mathbb{P}(X > x_0) - \mathbb{P}(X < x_0)\\ &= \mathbb{P}(X \leq x_0) - \mathbb{P}(X < x_0)\end{aligned}$$

Como queriamos probar.

Probando ls proposición. ■

Teorema 1.15: *Sea $F : \mathbb{R} \to \mathbb{R}$ una función que satisface las condiciones (a),(b),(c) de fda. Entonces F es una fda de alguna variable aleatoria.*

Dem: Consideremos el espacio de medida $([0,1], \mathcal{B}[0,1], \lambda)$. Definimos la función $X : [0,1] \to \mathbb{R}$ dada por,

$$X(\omega) := \sup\{y \in \mathbb{R} : F(y) < \omega\}$$

donde $\omega \in [0,1]$. Necesitamos probar la siguiente igualdad:

$$\{\omega : X(\omega) \leq x\} = \{\omega : \omega \leq F(x)\}$$

Supongamos que $\omega \in [0,1]$ es tal que $\omega \leq F(x)$, entonces $x \notin \{y \in \mathbb{R} : F(y) < \omega\}$, esto implica que $y < x$ ya que si $x \leq y$ entonces por F creciente se tieen que $F(x) \leq F(y) < \omega$. Ahora x es cota superior del conjunto $\{y \in \mathbb{R} : F(y) < \omega\}$ y entonces,

$$X(\omega) \leq x$$

Probando la inclusión $\supseteq$).

Ahora por contraposición supongamos que $\omega > F(x)$. Como F es continua por la derecha, se cumple que,

$$\lim_{\varepsilon \to 0^+} F(x + \varepsilon) = F(x)$$

Luego existe un $\delta > 0$ suficientemente pequeño de forma que,

$$F(x + \delta) < \omega$$

Esto implica que $x + \delta \leq X(\omega)$, es decir, $x \leq X(\omega)$. Por tanto si $\omega \notin \{\omega : \omega \leq F(x)\}$ entonces $\omega \notin \{\omega : X(\omega) \leq x\}$ que es básicamente decir que,

$$\{\omega : X(\omega) \leq x\} \subseteq \{\omega : \omega \leq F(x)\}$$

Probando la igualdad. Ahora esto significa que X es una variable aleatoria, puesto que dado $(-\infty, x] \in \mathcal{R}$, se tiene que,

$$\begin{aligned} X^{-1}((-\infty, x]) &= \{\omega : X(\omega) \leq x(\omega) \leq x\} \\ &= \{\omega : \omega \leq F(x) \leq 1\} \\ &= [0, F(x)] \in \mathcal{R} \end{aligned}$$

Para concluir que X es variable aleatoria basta saber que la colección $\{(-\infty, x] : x \in R\}$ tiene por generado σ-álgebra los boreleanos $\mathcal{R}$. Por tanto $X : [0,1] \to \mathbb{R}$ es una variable aleatoria que induce una función distribución acumulada F_X.

Para concluir el resultado notemos que,

$$F_X(x) = \mathbb{P}(\{\omega : X(\omega) \leq x\}) = \mathbb{P}(\{\omega : \omega \leq F(x)\}) = \lambda(\{\omega : \omega \leq F(x)\})$$

Y observemos que,

$$\begin{aligned}\lambda(\{\omega : \omega \leq F(x))\} &= \int_{[0,1]} \mathbb{1}_{\{\omega:\omega\leq F(x)\}}(\omega)d\lambda(\omega) \\ &= \int_0^{F(x)} 1ds \\ &= F(x)\end{aligned}$$

Usando que la constante 1 es Riemann-integrable. Por tanto $F_X(x) = F(x)$ para todo $x \in \mathbb{R}$. Como queriamos probar. ■

Corolario 1.16: *Sea $F : \mathbb{R} \to \mathbb{R}$ una función que satisface las condiciones (a),(b),(c) de la **proposición 1.14**, entonces existe una única medida de probabilidad μ_F en $(\mathbb{R}, \mathcal{R})$ tal que,*

$$\mu_F(a, b] = F(b) - F(a)$$

Dem: Vamos a probar la existencia y luego la unicidad.

- **Existencia:** Por el teorema anterior, existe una variable aleatoria X sobre un espacio de medida. Entonces podemos definir la distribución de la variable X por μ_X, que es una medida en $(\mathbb{R}, \mathcal{R})$, para $a < b$ notemos que:

$$\begin{aligned}\mu_X(a, b] &= \mathbb{P}(\{\omega : a < X(\omega) \leq b\}) \\ &= 1 - \mathbb{P}(\{\omega : X(\omega) \leq a\} \cup \{\omega : X(\omega) > b\}) \\ &= 1 - (\mathbb{P}(\{\omega : X(\omega) \leq a\}) + \mathbb{P}(\{\omega : X(\omega) > b\}) \\ &= 1 - \mathbb{P}(\{\omega : X(\omega) > b\}) - \mathbb{P}(\{\omega : X(\omega) \leq a\}) \\ &= \mathbb{P}(\{\omega : X(\omega) \leq b\}) - \mathbb{P}(\{\omega : X(\omega) \leq a\}) \\ &= F(b) - F(a)\end{aligned}$$

  Probando que existe tal medida.

- Tenemos el espacio de probabilidad $(\mathbb{R}, \mathcal{R}, \mu_X)$, si $\mathcal{R}$ es un álgebra y si $\mathbb{R}$ es fuertemente σ-infinito, se tiene que $\overline{\mu_X} : \sigma(\mathcal{R}) \to [0, 1]$ (extensión de μ_X dada por restringir la medida exterior μ_X^* asociada a μ_X sobre los σ-álgebras de los μ_X^*-medibles) es única extensión de μ_X, sin embargo, por construcción,

$$\overline{\mu_X} = \mu_X$$

  para todo evento de $\sigma(\mathcal{R}) = \mathcal{R}$. Por tanto la medida μ_X es única.

Probando el corolario. ■

Definición 1.17: *Sean $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1), (\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ dos espacios de probabilidad. Sean X e Y variables aleatorias de cada espacio de probabilidad respectivamente (X con la primera e Y con la segunda), decimos que son iguales en distribución si sus distribuciones son iguales, es decir, $\mu_X = \mu_Y$ (medida en $(\mathbb{R}, \mathcal{R})$), en tal caso denotamos $X \stackrel{d}{=} Y$.*

En general podemos decir que X, Y están en el mismo espacio. Otra cosa que podemos definir es una colección de variables aleatorias que sean identicamente distribuidos.

Definición 1.18: *Sea $(\Omega, \mathcal{F}, \mu, \nu)$ un espacio de medida con dos medidas. Decimos que μ es absolutamente continua con respecto a la medida ν si para todo $A \in \mathcal{F}$ tal que $\nu(A) = 0$, se cumple que $\mu(A) = 0$.*

Teorema 1.19: *Sea $(\Omega, \mathcal{F}, \mu, \nu)$ un espacio de medida con dos medidas. Si μ es absolutamente continua con respecto a ν, entonces existe una función $f \in L^1(\nu)$ tal que,*

$$\mu(A) = \int_A f(x) d(\nu(x))$$

En general, si la distribución de X puede ser expresado de la siguiente forma:

$$\mu_X(A) = \int_A f(x) d\nu(x)$$

donde ν es una medida $(\mathbb{R}, \mathcal{R})$, decimos que $f(x)$ es la función densidad de μ_X.

La unica medida en $(\mathbb{R}, \mathcal{R})$ tal que $|[a, b]| = b - a$

Definición 1.20: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sea X una variable aleatoria cuya distribución μ_X es absolutamente continua con respecto a la medida de Lebesgue, es decir, existe una función $f_X : \mathbb{R} \to [0, \infty]$ medible tal que,*

$$\mu_X(A) = \int_A f_X(x) d(\lambda(x))$$

para todo $A \in \mathcal{R}$. Entonces decimos que tal f_X es la función densidad de μ_X.

Ejemplo 1.21: Sea X una variable aletoria uniformemente distribuida en $[0, 1]$, es decir, tiene función distribución acumulada dada por:

$$F_X(x) = \begin{cases} 0, & x < 0 \\ x, & x \in [0, 1] \\ 1, & x > 1 \end{cases}$$

Observemos que,

$$\mu_X(A) = \mathbb{P}(X)$$

Esta distribución tiene una función densidad dada por:

$$f_X(x) = \begin{cases} 0, & x < 0 \\ 1, & x \in [0, 1] \\ 0, & x > 1 \end{cases}$$

Observación 1.22: Sabemos que una variable aleatoria X induce una medida y de ser posible, induce una función densidad f_X. Pero puede pasar que $f : \mathbb{R} \to [0, \infty]$ medible tal que,

$$\int_{\mathbb{R}} f(x)d(\lambda(x)) = 1$$

Entonces podemos definir una medida de probabilidad en $\mathbb{R}$ de la siguiente forma:

$$\mu(A) = \int_A f(x)d(\lambda(x)) \in [0, 1]$$

Esto induce una función distribución acumulada y por tanto por el **teorema 1.15** una variable aleatoria.

Ejemplo 1.22: Sea la función,

$$f(x) = ke^{-kx}\mathbb{1}_{[0,\infty)}(x)$$

Veamos si podemos deducir una variable aleatoria. Determinemos el valor de,

$$\int_{\mathbb{R}} f(x)d(\lambda(x))$$

Sea $f_n := f(x)\mathbb{1}_{[0,n]}$, claramentre $f_n \uparrow f$, entonces por el teorema de convergencia monótona se tiene que,

$$\begin{aligned}
\int_{\mathbb{R}} f d\lambda &= \lim_{n\to\infty} \int_{[0,n]} ke^{-kx}d\lambda \\
&= \lim_{n\to\infty} \int_0^n ke^{-kx}dx \\
&= 1
\end{aligned}$$

Entonces, podemos definir una medida y por tanto una función de distribución acumulada, en particular, la fda se determina de la siguinete forma:

$$\begin{aligned}
F(x) &= \int [0, x]ke^{-kx}d\lambda \\
&= \int_0^x ke^{-kx}dx \\
&= (1 - e^{-kx})\mathbb{1}_{[0,\infty)}(x)
\end{aligned}$$

Las variables aleatorias que tienen densidad f se denominan **exponenciales o exponencialmente distribuido con parámetro k.**

Ejemplo 1.23: La distribución normal es aquella que tiene función densidad:

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

Nota 1.24: No todas las distribuciones tienen densidad. Sea C el conjunto de 1/3-Cantor. Podemos definir la escalera del diablo por:

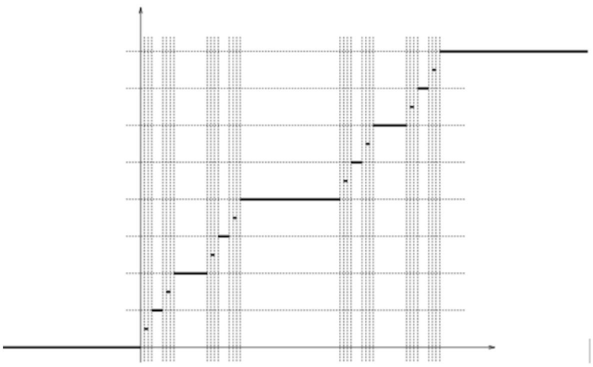

Figura 1: Escalera del Diablo

Esta función distribución acumulada tiene asociada a una variable aleatoria al cumplir las condiciones de fda, sin embargo, no poseé una función densidad.

Definición 1.25: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ una medida de probabilidad y sea $X : \Omega \to \mathbb{R}$ una variable aleatoria. Se define la esperanza dde X de la siguiente forma:*

$$\mathbb{E}(X) := \int_\Omega X(\omega)\mathbb{P}(d\omega)$$

siempre y cuando esté bien definido la integral.

Teorema 1.26: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabildiad y sea $X : \Omega \to \mathbb{R}$ una variable aleatoria. Si $f : \mathbb{R} \to \mathbb{R}$ es una función medible, entonces se satisface que,*

$$\mathbb{E}(f(X)) = \int_\mathbb{R} f(x)\mu_X(dx)$$

siempre y cuando la integral esté bien definida. Si además la distribución de X tiene una densidad, se cumple que,

$$\mathbb{E}(f(X)) = \int_\mathbb{R} f(x)g_X(x)\lambda(dx)$$

donde g_X es la densidad de X.

Dem: Vamos a probar por construcción de la función medible la primera igualdad. Supongamos que $f : \mathbb{R} \to \mathbb{R}$ es medible indicatriz, de la forma $f = \mathbb{1}_A$ con $A \in \mathcal{F}$, entonces se cumple por definición,

$$\begin{aligned}\mathbb{E}(f(X)) &= \int_\Omega f \circ X \mathbb{P} \\ &= \int_\Omega \mathbb{1}_{X^{-1}(A)} \mathbb{P} \\ &= \mathbb{P}(X^{-1}(A)) \\ &= \mu_X(A) \\ &= \int_\mathbb{R} \mathbb{1}_A d\mu_X\end{aligned}$$

Suppongamos que f es una función simple positiva, es decir,

$$f = \sum_{i=1}^n a_i \mathbb{1}_{A_i}$$

con $A_i \in \mathcal{F}$ y $a_i \geq 0$. Luego por linealidad obtenemos,

$$\begin{aligned}\mathbb{E}(f(X)) &= \int_\Omega \sum_{i=1}^n a_i \mathbb{1}_{X^{-1}(A_i)} d\mathbb{P} \\ &= \int_\mathbb{R} \sum_{i=1}^n a_i \mathbb{1}_{A_i} d\mu_X\end{aligned}$$

Ahora consideramos f una función no negativa, de forma que existe una sucesión de funciones simples tales que $f_n \uparrow f$, luego por el teorema de la convergencia monótona se concluye,

$$\mathbb{E}(f(X)) = \int_\mathbb{R} f d\mu_X$$

Para concluir el resultado debemos ver que para toda función medible f se puede descomponer de la siguiente forma: $f = f^+ - f^-$ donde f^+, f^- son funciones medibles no negativas, luego por linealidad se concluye el primer resultado.

Para la segunda igualdad notemos que si $f = \mathbb{1}_A$ con $A \in \mathcal{F}$ se tiene que,

$$\mathbb{E}(f(X)) = \mu_X(A) = \int_\mathbb{R} \mathbb{1}_A g_X d\lambda$$

Luego haciendo el mismo argumento anterior, se concluye que si la distribución de X tiene una densida g_X, entnces

$$\mathbb{E}(f(X)) = \int_\mathbb{R} f g_X d\lambda$$

■

Definición 1.27: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de medida. Sea X una variable aleatoria con esperanza finita $\mathbb{E}(X) \in \mathbb{R}$). Se define la variana de X por:*

$$Var(X) := \mathbb{E}((X - \mathbb{E}(X))^2)$$

Pedimos que la esperanza sea finita puesto que si no lo fuera, entonces la varianza puede tener problemas, ya que por el teorema anterior se tiene que,

$$\mathrm{Var}(X) = \int_{\mathbb{R}} (x-\mu)^2 d\mu_X$$

donde $\mu = \mathbb{E}(X)$. Y entonces,

$$\mathrm{Var}(X) = \int_{\mathbb{R}} x^2 d\mu_X - \mu^2$$

Por lo que por conveniencia pedimos que μ sea finito. Además, con esto concluimos que,

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

Ejemplo 1.28: Sea X una variable aleatoria exponencial de tasa 1, entonces,

$$\mathbb{E}(X^k) = \int_{\mathbb{R}} x^k \mu_X(dx) = \int_0^\infty x^k e^{-x} dx = (k-1)!$$

donde $k \in \mathbb{N}$. Si X es una variable aleatoria con densidad normal, entonces,

$$\mathbb{E}(X) = \int_{\mathbb{R}} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 0$$

Determinemos la varianza de X, por definición:

$$\begin{aligned}\mathrm{Var}(X) = \mathbb{E}(X^2) &= \int_{\mathbb{R}} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} d\lambda \\ &= 2\int_{[0,\infty)} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} d\lambda\end{aligned}$$

Consideremos la función $f_n : \mathbb{R} \to [0,\infty)$ medible:

$$f_n(x) := x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \mathbb{1}_{[0,n]}$$

Entonces f_n es creciente que además,

$$\lim_{n\to\infty} f_n(x) = x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Por lo tanto, por el teorema de la convergencia monótona se tiene que,

$$\int_{[0,\infty)} x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} d\lambda = \int_0^\infty x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

Tomando el cambio de variable $t = x^2/2$ obtenemos que,

$$\int_0^\infty x^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{\pi}} \int_0^\infty t^{1/2} e^{-t} dt = \frac{1}{2}$$

Por tanto,

$$\operatorname{Var}(X) = 2 \cdot \frac{1}{2} = 1$$

Proposición 1.29: *Sean X, Y variables aleatorias y sea $a \in \mathbb{R}$ una constante. Entonces se cumple que,*

- ***Esperanza lineal:*** $\mathbb{E}(aX + Y) = a\mathbb{E}(X) + \mathbb{E}(Y)$,
- ***Esperanza creciente:*** *Si $X \leq Y$ entonces,*

$$\mathbb{E}(X) \leq \mathbb{E}(Y)$$

- ***Propiedad de Varianza:*** $Var(aX) = a^2\, Var(X)$.

Dem:

2. Ley de los Grandes Números

En esta sección vamos a estudiar la ley de los grandes número, y como introducción, necesitamos hablat de independiencia.

2.1. Independencia

Definición 2.1: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de medida. Sean $A, B \in \mathcal{F}$ eventos. Decimos que son independientes si,*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

Si X, Y son variables aleatorias del mismo espacio de probabilidad, decimos que son independientes si,

$$\mathbb{P}(X \leq a, Y \leq b) = \mathbb{P}(X \leq a)\mathbb{P}(Y \leq b)$$

También se puede definir de forma más débil que para todo $A, B \in \mathcal{R}$ se tiene que,

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

Si $\Sigma_1, \Sigma_2 \subseteq \mathcal{F}$ son σ-álgebras, entonces diremos que son independientes si para todo $A \in \Sigma_1, B \in \Sigma_2$ se tiene que,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

Nota 2.2: La definición de eventos independiente no parece tener mucho sentido, pero si usamos la definición de la probabilidad condicional dada por:

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

que es básicamente estudiar la probabilidad de A dado que ocurrio el evento B, aplicando que A y B son independientes se llega a que,

$$\mathbb{P}(A|B) = \mathbb{P}(A)$$

Es decir, no importa si ocurre B o no, la probabilidad de A no se ve afectada, es decir, A es independiente de B, pero no necesariamente B es independiente de A, pero si estudiamos la probabilidad de B dado A se obtiene,

$$\mathbb{P}(B|A) = \mathbb{P}(B)$$

Es decir, B es independiente de A, por tanto tiene sentido que la definición de independencia menciona que A, B son independientes entre si.

Observación 2.3: Sea X una variable aleatoria sobre un espacio de probabilidad. Podemos definir el σ-álgebra de la variable aleatoria como:

$$\sigma(X) := \{X^{-1}(A) : A \in \mathcal{A}\}$$

que es en efecto, un σ-álgebra. Para poder verlo hay que verificar los tres axiomas de σ-álgebra.

(a) $X^{-1}(\emptyset) = \emptyset \in \sigma(X)$,

(b) Si $A \in \sigma(X)$, entonces $A = X^{-1}(B)$ donde $B \in \mathcal{R}$, luego,

$$A^c = X^{-1}(B^c) \in \mathcal{R}$$

(c) Sea $\{A_n\}_{n\in\mathbb{N}} \subseteq \sigma(X)$ una colección, entonces,

$$\begin{aligned}\bigcup_{n\in\mathbb{N}} A_n &= \bigcup_{n\in\mathbb{N}} X^{-1}(B_n)\\ &= X^{-1}\left(\bigcup_{n\in\mathbb{N}} B_n\right) \in \mathcal{R}\end{aligned}$$

Verificando que $\sigma(X)$ es un σ-álgebra.

Teorema 2.4: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean $X, Y : \Omega \to \mathbb{R}$ variables aleatorias, entonces X, Y son independientes si y sólo si $\sigma(X), \sigma(Y)$ lo son.*

Dem: Supongamos que $X, Y : \Omega \to \mathbb{R}$ son variables aleatorias. Debemos probar que,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

para todo $A \in \sigma(X), B \in \sigma(Y)$, por definición se tiene que $A = X^{-1}(A_x), B = Y^{-1}(B_y)$ donde $A_x, B_y \in \mathcal{R}$, luego

$$\begin{aligned}\mathbb{P}(A) &= \mathbb{P}(X^{-1}(A_x))\\ &= \mathbb{P}(X \in A_x)\end{aligned}$$

Con esto podemos concluir el resultado, puesto que,

$$\begin{aligned}\mathbb{P}(A \cap B) &= \mathbb{P}(X^{-1}(A_x) \cap Y^{-1}(B_y))\\ &= \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A_x, Y(\omega) \in B_y\})\\ &= \mathbb{P}(X \in A_x, Y \in B_y)\\ &= \mathbb{P}(X \in A_x)\mathbb{P}(Y \in B_y)\\ &= \mathbb{P}(A)\mathbb{P}(B)\end{aligned}$$

usando que X, Y son independientes.

Supongamos que $\sigma(X), \sigma(Y)$ son independientes. Sean $A, B \in \mathcal{R}$, entonces $X^{-1}(A) \in \sigma(X), Y^{-1}(B) \in \sigma(Y)$ y por tanto,

$$\begin{aligned}\mathbb{P}(X \in A, Y \in B) &= \mathbb{P}(X^{-1}(A) \cap Y^{-1}(B))\\ &= \mathbb{P}(X^{-1}(A))\mathbb{P}(Y^{-1}(B))\\ &= \mathbb{P}(X \in A)\mathbb{P}(Y \in B)\end{aligned}$$

Por lo tanto, X, Y son variables aleatorias independientes. ■

Observación 2.5: En un espacio de probabilidad, sean dos σ-álgebras independientes Σ_1, Σ_2. Si X es Σ_1-medible e Y es Σ_2-medible, entonces X e Y son independientes. Verifiquemos este hecho. Debemos probar que para todo $A, B \in \mathcal{R}$ se tiene que,

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

Por definición de preimagen, se tiene que,

$$\begin{aligned}
\{X \in A, Y \in B\} &= \{\omega \in \Omega : X(\omega) \in A, Y(\omega) \in B\} \\
&= \{\omega \in \Omega : X(\omega) \in A\} \cap \{\omega \in \Omega : Y(\omega) \in B\} \\
&= X^{-1}(A) \cap Y^{-1}(B)
\end{aligned}$$

Por tanto, como X, Y son variables aleatorias, se tiene que $X^{-1}(A) \in \Sigma_1, Y^{-1}(B) \in \Sigma_2$, por lo que,

$$\begin{aligned}
\mathbb{P}(X \in A, Y \in B) &= \mathbb{P}(X^{-1}(A)) \cap Y^{-1}(B)) \\
&= \mathbb{P}(X^{-1}(A))\mathbb{P}(Y^{-1}(B)) \\
&= \mathbb{P}(X \in A)\mathbb{P}(Y \in B)
\end{aligned}$$

Como queriamos verificar.

Teorema 2.6: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Entonces $A, B \in \mathcal{F}$ son independientes si y sólo si $\mathbb{1}_A, \mathbb{1}_B : \Omega \to \mathbb{R}$ son variables aleatorias independientes.*

Antes de demostrar el teorema, requerimos de un pequeño lema.

Lema 2.7: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean A, B eventos independientes, entonces se cumple la siguientes propiedades:*

(a) A^c, B son independientes,

(b) A, B^c son independientes,

(c) A^c, B^c son independientes.

Dem: Para demostrar el lema antes necesitamos observar algunas cosas, como estamos trabajando en un conjunto σ-finito tenemos muchas propiedades útiles. Notemos que para todo $A \in \mathcal{F}$ se cumple que,

$$\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$$

puesto que $A \cap A^c = \emptyset$ luego por propiedades de medida se llega a lo anterior.

(a) Notemos lo siguiente,

$$\begin{aligned}
\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) &\iff \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(B) - \mathbb{P}(A)\mathbb{P}(B) \\
&\iff \mathbb{P}(B \cap A^c) = \mathbb{P}(B)(1 - \mathbb{P}(A)) \\
&\iff \mathbb{P}(B \cap A^c) = \mathbb{P}(B)\mathbb{P}(A^c)
\end{aligned}$$

Debemos notar que $B \cap A^c \cap B \cap A = \emptyset$ y que $(B \cap A^c) \cup (B \cap A) = B$, obteniendo el resultado.

(b) En este caso basta aplicar el punto anterior, obteniendo que,

$$\mathbb{P}(A \cap B^c) = \mathbb{P}(A)\mathbb{P}(B^c)$$

siendo A, B^c independientes.

(c) Notemos que,

$$\begin{aligned}
\mathbb{P}(A^c \cap B) = \mathbb{P}(A^c)\mathbb{P}(B) &\iff \mathbb{P}(A^c) - \mathbb{P}(A^c \cap B) = \mathbb{P}(A^c) - \mathbb{P}(A^c)\mathbb{P}(B) \\
&\iff \mathbb{P}(A^c \cap B^c) = \mathbb{P}(A^c)(1 - \mathbb{P}(B)) \\
&\iff \mathbb{P}(A^c \cap B^c) = \mathbb{P}(A^c)\mathbb{P}(B^c)
\end{aligned}$$

Por lo que A^c, B^c son independientes.

Probando el lema.■

Dem (Teorema 2.6): Sea $A, B \in \mathcal{F}$ por lo que las indicatrices $\mathbb{1}_A, \mathbb{1}_B$ son variables aleatorias. Supongamos que las indicatrices son independientes. Consideremos $\{1\} \in \mathcal{R}$ (se puede expresar como $\{1\} = (-\infty, 1) \cup (1, \infty)$, es decir, la unión de dos conjuntos de $\mathcal{R}$), luego,

$$\begin{aligned}
\mathbb{P}(A \cap B) &= \mathbb{P}(\mathbb{1}_A \in \{1\}, \mathbb{1}_B \in \{1\}) \\
&= \mathbb{P}(\mathbb{1}_A \in \{1\})\mathbb{P}(\mathbb{1}_B \in \{1\}) \\
&= \mathbb{P}(A)\mathbb{P}(B)
\end{aligned}$$

Por tanto, A, B son eventos independientes.

Supongamos ahora que A, B son eventos independientes. Vamos a usar el teormea 1.32, es decir, demostraremos que los σ-álgebras $\sigma(\mathbb{1}_A), \sigma(\mathbb{1}_B)$ son independientes. Notemos que,

$$\begin{aligned}
\sigma(\mathbb{1}_A) &= \{\emptyset, A, A^c, \Omega\} \\
\sigma(\mathbb{1}_B) &= \{\emptyset, B, B^c, \Omega\}
\end{aligned}$$

Por lo que debemos probar que $\mathbb{P}(U \cap V)$ son independientes para todo $U \in \sigma(\mathbb{1}_B), V \in \sigma(\mathbb{1}_B)$. Pero observemos que si U, V no es $\emptyset$ ni Ω, entonces tenemos que probar los casos de independencia del lema 1.35, cosa que ya fue probada por tanto por el teorema 1.32 se tiene que las indicatrices son independientes. ■

Definición 2.7: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean $\Sigma_1, \ldots, \Sigma_n$ σ-álgebras subconjunto de $\mathcal{F}$. decimos que son independientes si,*

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) = \prod_{i=1}^{n} \mathbb{P}(A_i), \quad \textit{para todo } A_i \in \Sigma_i$$

Sean $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ variables aleatorias. Decimos que son independientes si,

$$\mathbb{P}(X_1 \in A_1, \ldots, X_n \in A_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \in A_i), \quad \textit{para todo } A_i \in \mathcal{R}$$

Sean $A_1, \ldots, A_n$ eventos. Decimos que son independientes si para toda subcolección $I \subseteq \{1, \ldots, n\}$ se tiene que,

$$\mathbb{P}\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} \mathbb{P}(A_i)$$

Observación 2.8: La independencia sobre n eventos se define de esa forma debido a que queremos una relaciación entre la otras independencia, además, uno podría intuitivamente pensar en la independencia de n eventos como que sean independientes a pares, pero esta condición no es tan buena, el siguiente ejemplo verifica este hecho.

Ejemplo 2.9: Consideremos un espacio de probabilidad y sean $X_1, X_2, X_3 : \Omega \to \{0,1\} \subseteq \mathbb{R}$ variables aleatorias independientes tales que,

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = 0) = \frac{1}{2}$$

para $i = 1, 2, 3$. Sean los eventos:

$$A = \{X_1 = X_2\},\ B = \{X_1 = X_3\},\ C = \{X_2 = X_3\}$$

Entonces se cumple que,

$$\begin{aligned}
\mathbb{P}(A) &= \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = X_2(\omega\}\}) \\
&= \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = X_2(\omega\} = 1\} \cup \{\omega \in \Omega : X_1(\omega) = X_2(\omega\} = 0\}) \\
&= \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = X_2(\omega\} = 1\}) + \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = X_2(\omega\} = 0\}) \\
&= \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = 1, X_2(\omega\} = 1\}) + \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = 0, X_2(\omega\} = 0\}) \\
&= \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = 1\})\mathbb{P}(\{\omega \in \Omega : X_0(\omega) = 1\}) \\
&+ \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = 0\})\mathbb{P}(\{\omega \in \Omega : X_2(\omega) = 0\}) \\
&= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}
\end{aligned}$$

Usando la independencia de las variables aleatorias. De forma análoga podemos concluir que,

$$\mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{2}$$

También podemos ver que,

$$\mathbb{P}(A \cap B) = \frac{1}{2}^3 + \frac{1}{2}^3 = \frac{1}{4}$$

Por tanto,

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

Por lo que A, B son independientes, es más, A, B, C son independientes a pares, sin embargo,

$$\mathbb{P}(A \cap B \cap C) = \frac{1}{4} \neq \frac{1}{8} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$$

Es decir, A, B, C no son independientes.

Ahora, la independencia se puede generalizar sobre un índice infinito no numerable, lo cual haremos en la siguiente definición:

Definición 2.10: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sea $\mathcal{C}$ una colección infinita (numerable o no numerable) de σ-álgebras, eventos o de variables aleatorias. Decimos que la colección es independiente si para toda subcolección finita de $\mathcal{C}$ es independiente.*

Nota 2.11: Sabemos que una forma de definir la independencia de dos variables aleatorias X, Y se necesita probar que:

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

Pero esta definición implica la otra definición:

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

para todo $x, y \in (-\infty, \infty)$, y esto se debe a la forma de $\mathcal{R}$, se puede comprobar que $(-\infty, x] \in \mathcal{R}$ con la siguiente interpretación:

$$(-\infty, x] = \bigcap_{n \in \mathbb{N}} \left(-\infty, x + \frac{1}{n}\right)$$

donde claramente $(-\infty, x + 1/n) \in \mathcal{R}$. Lo más útil, es que ambas definiciones son equivalentes, para probarlo notemos la siguiente identificación:

$$\sigma(I_s) = \sigma(\{O : O \subseteq \mathbb{R}, \text{abierto}\}) = \mathcal{R}$$

donde $I_s = \{(a, b] : -\infty \leq a < b < \infty\}$, es decir, $A \in \mathcal{R}$. Para concluir la equivalencia se requiere de un teorema muy importante.

Teorema 2.12 (Teorema $\pi - \lambda$): *Sea Ω un conjunto no vacío y sea $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ una colección. Decimos que $\mathcal{A}$ es un π-sistema si es cerrado bajo la intersección. Sea $\mathcal{B} \subseteq \mathcal{P}(\Omega)$ otra colección. Decimos que es un λ-sistema si $\Omega \in \mathcal{B}$, si para todo $A \subset B$ donde $A, B \in \mathcal{B}$ son tales que $B \setminus A \in \mathcal{B}$ y si $A_n \in \mathcal{B}$ para todo $n \geq 1$ es una secuencia de conjuntos tales que $A_n \uparrow A$ (A_n es creciente y converge a A), entonces $A \in \mathcal{B}$. El teorema dice que si $\mathcal{A}$ es un π-sistema y $\mathcal{L}$ es un λ-sistema que contiene a $\mathcal{A}$, entonces $\sigma(\mathcal{A}) \subseteq \mathcal{L}$.*

Dem Durret.

Definición 2.13: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean $\mathcal{B}_1, \ldots, \mathcal{B}_n$ finitas colecciones de conjuntos contenidos en $\mathcal{F}$. Decimos que son independientes si para todo subcolección $I \subseteq \{1, \ldots, n\}$ y para todo $\{B_i\}_{i \in I}$ tal que $B_i \in \mathcal{B}_i$, se tiene que,*

$$\mathbb{P}\left(\bigcap_{i \in I} B_i\right) = \prod_{i \in I} \mathbb{P}(B_i)$$

Observación 2.14: No confundir la definición anterior con independencia de σ-álgebras, en la definición anterior estamos tomando colecciones arbitrarias.

Teorema 2.15: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un conjunto no vacío y sean $\mathcal{A}_1, \ldots, \mathcal{A}_n$ π-sistemas independientes. Entonces $\sigma(\mathcal{A}_1), \ldots, \sigma(\mathcal{A}_n)$ son independientes.*

Dem: Vamos a demostrar que,

$$\mathbb{P}\left(A_1 \cap \cdots \cap A_n\right) = \prod_{i=1}^{n} \mathbb{P}(A_i)$$

para todo $A_i \in \sigma(\mathcal{A}_i)$ con $i = 1, \ldots, n$. Para ello demostraremos que se cumple el caso $A_1 \in \sigma(\mathcal{A}_1)$ y con $A_i \in \mathcal{A}_i$ para $i = 2, \ldots, n$, es decir, que,

$$\mathbb{P}\left(A_1 \cap \bigcap_{i=2}^{n} A_i\right) = \mathbb{P}(A_1)\mathbb{P}\left(\bigcap_{i=2}^{n} A_i\right)$$

En particular, como $\mathcal{A}_2, \ldots \mathcal{A}_n$ son independientes, se tiene que,

$$\mathbb{P}\left(A_1 \cap \bigcap_{i=2}^{n} A_i\right) = \mathbb{P}(A_1)\mathbb{P}\left(\bigcap_{i=2}^{n} A_i\right) = \prod_{i=1}^{n} \mathbb{P}(A_i)$$

considerando $A_1 \in \sigma(\mathcal{A}_1)$ y $A_i \in \mathcal{A}_i$ para $i = 2, \ldots, n$. Fijemos $A_i \in \mathcal{A}_i$ para $i = 2, \ldots, n$ y definamos el siguiente conjunto:

$$\mathcal{L}_1 := \left\{ A_1 \in \sigma(\mathcal{A}_1) : \mathbb{P}\left(A_1 \cap \bigcap_{i=2}^{n} A_i\right) = \mathbb{P}(A_1)\mathbb{P}\left(\bigcap_{i=2}^{n} A_i\right) \right\}$$

Notemos que $\mathcal{A}_1 \subseteq \mathcal{L} \subseteq \sigma(\mathcal{A}_1)$. Vamos a probar que $\mathcal{L}$ es un λ-sistema, para ello debemos probar los axiomas de λ-sistema.

(a) Claramente $\Omega \in \mathcal{L}$ puesto que,

$$\mathbb{P}\left(\Omega \cap \bigcap_{i=2}^{n} A_i\right) = \mathbb{P}\left(\bigcap_{i=2}^{n} A_i\right) = \mathbb{P}(\Omega)\mathbb{P}\left(\bigcap_{i=2}^{n} A_i\right)$$

(b) Sean $A, B \in \mathcal{L}$ tales que $A \subset B$, debemos probar que,

$$\mathbb{P}\left((B \setminus A) \cap \bigcap_{i=1}^{n} A_i\right) = \mathbb{P}(B \setminus A)\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right)$$

Por comodidad diremos que:

$$C := \bigcap_{i=1}^{n} A_i$$

Notemos que,

$$\begin{aligned} \mathbb{P}((B \cap C) \setminus (A \cap C) &= \mathbb{P}(B \cap C) - \mathbb{P}(A \cap C) \\ &= \mathbb{P}(B)\mathbb{P}(C) - \mathbb{P}(A)\mathbb{P}(C) \\ &= \mathbb{P}(B \setminus A)\mathbb{P}(C) \end{aligned}$$

Luego $B \setminus A \in \lambda$.

(c) Consideremos la secuencia:

$$A_1 \subseteq \cdots \subseteq A_n \subseteq \ldots$$

donde $A_i \in \mathcal{L}$ para todo $i \in \mathbb{N}$. Entonces se tiene la siguiente secuencia creciente:

$$A_1 \cap C \subseteq \cdots \subseteq A_n \cap C \subseteq \ldots$$

Por tanto, tenemos que,

$$\begin{aligned}
\mathbb{P}\left(\left(\bigcup_{i=1}^{\infty} A_i\right) \cap C\right) &= \mathbb{P}\left(\bigcup_{i=1}^{\infty} (A_i \cap C)\right) \\
&= \lim_{n\to\infty} \mathbb{P}(A_n \cap C) \\
&= \lim_{n\to\infty} \mathbb{P}(A_n)\mathbb{P}(C) \\
&= \mathbb{P}(A)\mathbb{P}(C)
\end{aligned}$$

por tanto $A \in \mathcal{L}$ (donde $A_n \uparrow A$).

Por tanto, por el teorema de $\pi - \lambda$ se tiene que $\sigma(\mathcal{A}_1) \subseteq \mathcal{L}$ y por tanto,

$$\mathcal{L} = \sigma(\mathcal{A}_1)$$

Es decir, la condición:

$$\mathbb{P}\left(A_1 \cap \bigcap_{i=2}^{n} A_i\right) = \mathbb{P}(A_1)\mathbb{P}\left(\bigcap_{i=2}^{n} A_i\right) = \prod_{i=1}^{n} \mathbb{P}(A_i)$$

se cumple para todo para todo $A_1 \in \sigma(\mathcal{A}_1)$ y para todo $A_i \in \sigma(\mathcal{A}_i)$ con $i = 2, \ldots, n$. Ahora para concluir el resultado debemos definir el conjunto:

$$\mathcal{L}_2 := \left\{ A_2 \in \sigma(\mathcal{A}_2) : \mathbb{P}\left(A_2 \cap \bigcap_{i=1, i\neq 2}^{n} A_i\right) = \mathbb{P}(A_1)\mathbb{P}\left(\bigcap_{i=2}^{n} A_i\right)\right\}$$

donde $A_1 \in \sigma(\mathcal{A}_1), A_i \in \mathcal{A}_i$ con $i = 3, \ldots, n$. Luego inductivamente se puede probar que,

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} \mathbb{P}(A_i)$$

para todo $A_i \in \sigma(\mathcal{A}_i)$, es decir, $\sigma(\mathcal{A}_1), \ldots, \sigma(\mathcal{A}_n)$ son independientes. ■

Corolario 2.16: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ espacio de probabilidad y sean $X_1, \ldots, X_n$ variables aleatorias. Entonces para que $X_1, \ldots, X_n$ sean independientes, es suficiente que se satisfaga:*

$$\mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq x_i)$$

para todo $(x_1, \ldots, x_n) \in \mathbb{R}^n$.

Dem: Definimos $\mathcal{A}_i \subseteq \mathcal{P}(\Omega)$ como los conjuntos de la forma $\{\omega \in \Omega : X_i(\omega) \leq x\}$ para algún $x \in \mathbb{R}$ para todo $i = 1, \ldots, n$. Entonces $\mathcal{A}_i$ es un π-sistema puesto que.

$$\{\omega \in \Omega : X_i(\omega) \leq x\} \cap \{\omega \in \Omega : X_i(\omega) \leq y\} = \{\omega \in \Omega : X_i(\omega) \leq \text{mín}(x, y)\}$$

Además, por hipótesis, los $\mathcal{A}_i$ son independientes. Por tanto, por el **teorema 2.15** se tiene que $\sigma(\mathcal{A}_1), \ldots, \sigma(\mathcal{A}_n)$ son independientes. Probemos que para todo $i = 1, \ldots, n$ se tiene que $\sigma(\mathcal{A}_i) = \sigma(X_i)$. Recordemos que $\mathcal{R} = \sigma(\{(-\infty, x] : x \in \mathbb{R}\})$, entonces,

$$\sigma(X_i) = \{X_i^{-1}(A) : A \in \sigma(\{(-\infty, x] : x \in \mathbb{R}\})\}$$

Y esto implica la igualdad. Probando el corolario. ∎

Ahora vamos a estudiar una función aleatoria. En un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ se tiene las variables aleatorias independientes $X_1, \ldots, X_n$, entonces se tiene que las variables aleatorias:

$$X_1 + \cdots + X_{n-1}, \ \ X_n$$

Para verificarlo, definimos la siguiente función

$$\begin{aligned} f : \mathbb{R}^{n-1} &\to \mathbb{R} \\ (x_1, \ldots, x_{n-1}) &\mapsto x_1 + \cdots + x_{n-1} \end{aligned}$$

Entonces podemos pensar en que,

$$f(X_1, \ldots, X_{n-1}) = X_1(\omega) + \cdots + X_n(\omega)$$

es una variable aleatoria independiente de X_n. **verficar**.

Veamos que son independientes. Por definición, para $A, B \in \mathbb{R}$ tenemos,

$$\begin{aligned} \mathbb{P}(f(X_1, \ldots, X_{n-1}) \in A, X_n \in B) &= \mathbb{P}((X_1, \ldots, X_{n-1}) \in f^{-1}(A), X_n \in B)\} \\ &= ... \\ &= \mathbb{P}(f(X_1, \ldots, X_{n-1}) \in A)\mathbb{P}(X_n \in B) \end{aligned}$$

Teorema 2.17: *Sea* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espacio de medida. Sean* $\mathcal{F}_{i,j} \subseteq \mathcal{F}$*, donde* $1 \leq i \leq n, 1 \leq j \leq m(i)$ *σ-álgebras independientes. Definimos:*

$$\mathcal{G}_i := \sigma\left(\bigcup_{j=1}^{m(i)} \mathcal{F}_{i,j}\right)$$

Entonces, se tiene que,

$$\mathcal{G}_1, \ldots, \mathcal{G}_n$$

son independientes.

Dem: Sea $\mathcal{A}_i$ la colección de los conjuntos de la forma $\bigcap_{j=1}^{m(i)} A_{i,j}$ donde $A_{i,j} \in \mathcal{F}_{i,j}$. Entonces $\mathcal{A}_i$ es un π-sistema que contiene a Ω y al conjunto,

$$\bigcup_{j=1}^{m(i)} \mathcal{F}_{i,j}$$

También observemos que $\mathcal{A}_i$ son independientes, entonces por el **teorema 2.15** se tiene que, $\sigma(\mathcal{A}_i) = \mathcal{G}_i$ son independientes. ■

Corolario 2.18: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean $X_{i,j}$ variablea aleatorias independientes donde $i = 1, \ldots, n$ y $j = 1, \ldots, m(i)$ y sean $f_i : \mathbb{R}^{m(i)} \to \mathbb{R}$ funciones medibles para todo $i = 1, \ldots, n$. Entonces,*

$$f_i(X_{i,1}, \ldots, X_{i,m(i)})$$

son funciones independientes para $i = 1, \ldots, n$.

Dem: Probar que f_i**es medible**. Vamos a probar que $\sigma(f_i(X_{i,1}, \ldots, X_{i,m(i)}))$ son independientes. Si los $X_{i,j}$ con $1 \leq i \leq n, 1 \leq j \leq m(i)$ son indpendientes, entonces la colección $\sigma(X_{i,j})$ con $1 \leq i \leq n, 1 \leq j \leq m(i)$ son independientes, luego por el teorema anterior se tiene que,

$$\mathcal{G}_i = \sigma\left(\bigcup_{j=1}^{m(i)} \sigma(X_{i,j})\right)$$

son independientes. Veamos que $\mathcal{G}_i = \sigma(f_i(X_{i,1}, \ldots, X_{i,m(i)}))$. Observemos que, **blabla** luego por el teorema anterior, se tiene que los $f_i(X_{i,1}, \ldots, X_{i,m(i)})$ son independientes. ■

Teorema 2.19: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean $X_1, \ldots, X_n$ variables aleatorias independientes tales que X_i tiene distribución μ_i. Entonces el vector aleatorio:*

$$(X_1, \ldots, X_n)$$

tiene distribución $\mu_1 \otimes \cdots \otimes \mu_n$.

Dem: Sea $A := A_1 \times \cdots \times A_n$ donde $A_i \in \mathcal{R}$ para todo $i = 1, \ldots, n$ un réctangulo. Entonces, por definición de medida producto, se cumple que:

$$\begin{aligned}
\mathbb{P}((X_1, \ldots, X_n) \in A) &= \mathbb{P}(X_1 \in A_1, \ldots, x_n \in A_n) \\
&= \mu_1(A_1) \ldots \mu_n(A_n) \\
&= \mu_1 \otimes \cdots \otimes \mu_n(A)
\end{aligned}$$

Por otro lado sabemos que los conjuntos cilíndricos generan $\mathcal{R}^n$. Por tanto la distribución del vector aleatorio $(X_1, \ldots, X_n)$ coincide con la medida $\mu_1 \otimes, \cdots \otimes \mu$ para todo $\mathcal{R}^n$. ■

Teorema 2.19: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean X, Y variables aleatorias independientes y sea $h : \mathbb{R}^2 \to \mathbb{R}$ una función medible. Se tiene que,*

$$\mathbb{E}(h(X, Y)) = \int_{\mathbb{R}^2} h(x, y) \mu \otimes v$$

donde μ es la distribución de X y ν es la distribución de Y y $h(x,y)$ es una función tal que:

$$\begin{aligned}&\Omega \to \mathbb{R}^2\\ &\omega \mapsto h(X(\omega), Y(\omega))\end{aligned}$$

En particular, si $h(x,y) = f(x)g(y)$ y si $f, g \geq 0$ o $\mathbb{E}(|f(x)|) < \infty$ y $\mathbb{E}(|g(y)|) < \infty$. Entonces,

$$\mathbb{E}(h(X,Y)) = \mathbb{E}(f(X)) \cdot \mathbb{E}(g(Y)) = \mathbb{E}(f(X)g(Y))$$

Dem: Supongamos que $f, g \geq 0$, luego por Fubini tenemos que,

$$\begin{aligned}\mathbb{E}(f(X), g(Y)) &= \int_{\mathbb{R}^2} f(x)g(y)\mu \otimes \nu\\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x)g(y)d\mu(x) \right) \nu(dy)\\ &= \int_{\mathbb{R}} g(y)\nu(dy) \int_{\mathbb{R}} f(x)d\mu(x)\\ &= \mathbb{E}(f(X))\mathbb{E}(g(Y))\end{aligned}$$

Para el otro caso debemos probar que $h \in L^1(\Omega)$ para aplicar Fubini. Por definición,

$$\begin{aligned}\int_{\mathbb{R}^2} |h(x,y)| d\mu \otimes d\nu &= \int_{\mathbb{R}^2} |f(x)||g(x)| d\mu \otimes d\nu\\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} |f(x)||g(y)| d\mu \right) d\nu\\ &= \mathbb{E}(|f(X)|)\mathbb{E}(g(Y)) < \infty\end{aligned}$$

Luego aplicando lo anterior se concluye que,

$$\mathbb{E}(f(X), g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y))$$

■

Teorema 2.20: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabildiad. Sean $X_1, \ldots, X_n$ variables aleatorias independientes tales que $X_i \geq 0$ para $i = 1, \ldots, n$, o $\mathbb{E}(|X_i|) < \infty$ para todo i. Entonces,*

$$\mathbb{E}\left(\prod_{i=1}^{n} X_i \right) = \prod_{i=1}^{n} \mathbb{E}(X_i)$$

Dem: Sean $X_1, \ldots, X_n$ variables aleatorias independientes. Consideremos $h : \mathbb{R}^n \to \mathbb{R}$ dado por $h(x_1, \ldots, x_n) = x_1 \ldots x_n$ que es medible puesto que es el producto de funciones medibles. Entonces si $X_i \geq 0$ tenemos por Fubini que:

$$\begin{aligned}\mathbb{E}\left(\prod_{i=1}^{n} X_i \right) &= \int_{\mathbb{R}^n} x_1 \ldots x_n \mu_{X_1} \otimes \cdots \otimes \mu_{X_n}\\ &= \left(\int_{\mathbb{R}} x_1 \mu_{X_1}(dx) \right) \ldots \left(\int_{\mathbb{R}} x_n \mu_{X_n}(dx_n) \right)\\ &= \mathbb{E}(X_1) \ldots \mathbb{E}(X_n)\end{aligned}$$

Como querimos probar. ■

Nota 2.21. Puede pasar que $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ pero con X, Y variables aleatorias no independientes, es tal caso decimos que X e Y están **no corralacionados**.

Teorema 2.22: *Sea* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espacio de probabilidad. Sean* X, Y *variables aleatorias independientes. Entonces,*

$$\mathbb{P}(X + Y \leq z) = \int_{\mathbb{R}} F(z - y)\nu(dy)$$

donde F *es la función distribución de* X *y* ν *es la distribución de* Y*. Si* G *es la función distribución de* Y*, entonces a veces se escribe como:*

$$\int_{\mathbb{R}} F(z - y)dG(y)$$

Dem terminar: Tenemos que,

$$\mathbb{P}(X + Y \leq z) = \mathbb{E}(\mathbb{1}_{\{X+Y\leq z\}}(x, y))$$

y para ver esto basta estudiar por definición, por lo que,

$$\begin{aligned} \{X + Y \leq z\} &= \{\omega \in \Omega : X(\omega) + Y(\omega) \leq z\} \\ &= \{\omega \in \Omega : h(X(\omega), Y(\omega)) \leq z\}) \end{aligned}$$

donde $h(x, y) = x + y$. Ahora,

$$\mathbb{P}(X + Y \leq z) = \int_{\mathbb{R}^2} \mathbb{1}_{\{X+Y\leq z\}}(x, y)d\mu(x) \otimes d\nu(y)$$

donde μ es la distribución de X y ν es la distribución de Y. Luego por Fubini tenemos que,

$$\begin{aligned} \int_{\mathbb{R}^2} \mathbb{1}_{\{X+Y\leq z\}}(x, y)d\mu(x) \otimes d\nu(y) &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{1}_{\{X\leq z-Y\}}(x, y)d\mu(x) \right) d\nu(y) \\ &= \int_{\mathbb{R}} \mathbb{P}(X \leq z - y)d\nu(y) \\ &= \int_{\mathbb{R}} F(z - y)\nu(dy) \\ &= \int_{\mathbb{R}} F(z - y)dG(y) \end{aligned}$$

revisar ■

Teorema 2.23: *Sea* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espacio de probabilidad. Sean* X, Y *variables aleatorias independientes. Si* X *tiene una denisdad* f*, entonces* $X + Y$ *tiene densidad,*

$$\int f(z - y)dG(y)$$

Si además, Y tiene densidad g, entonces $X+Y$ tiene densidad,

$$\int f(z-y)g(y)d\lambda$$

Dem: Se tiene que,

$$\begin{aligned}
\mathbb{P}(X+Y\leq z) &= \int F(z-y)dG(y)\\
&= \int \left(\int_{-\infty}^{z-y} f(x)dx\right) dG(y)\\
&= \int \left(\int_{-\infty}^{z} f(x-y)dx\right) dG(y)\\
&= \int_{-\infty}^{z} \left(\underbrace{\int f(x-y)dG(y)}_{\text{depende solamente de } x}\right) dx
\end{aligned}$$

■

Ejemplo 2.24: Decimos que una variable aleatoria tiene distribución normla $N(\mu,\sigma)$ si tiene función densidad:

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

Esta variable aleatoria tiene media/esperanza μ y varianza σ^2. Verifiquemos este hecho. Para la esperanza se tiene que,

$$\mathbb{E}(X)=\int_{\mathbb{R}}\frac{x}{\sqrt{2\pi\sigma}}\exp\left(-\frac{(x-\mu)^2}{2\sigma}\right)d\lambda$$

Queremos usar integral de Riemann, apra ello observemos que la función,

$$\frac{x}{\sqrt{2\pi\sigma}}\exp\left(-\frac{(x-\mu)^2}{2\sigma}\right)$$

Es Riemann integrable en $\mathbb{R}$, probemos que es $L^1(\mathbb{R})$. Por definición se tiene que,

$$\begin{aligned}
\int_{\mathbb{R}}\frac{|x|}{\sqrt{2\pi\sigma}}\exp\left(-\frac{(x-\mu)^2}{2\sigma}\right)d\lambda &= 2\int_{[0,\infty)}\frac{x}{\sqrt{2\pi\sigma}}\exp\left(-\frac{(x-\mu)^2}{2\sigma}\right)d\lambda\\
&= 2\int_{0}^{\infty}\frac{x}{\sqrt{2\pi\sigma}}\exp\left(-\frac{(x-\mu)^2}{2\sigma}\right)dx
\end{aligned}$$

terminar Luego se tiene que,

$$\mathbb{E}(X)=\int_{-\infty}^{\infty}\frac{x}{\sqrt{2\pi\sigma}}\exp\left(-\frac{(x-\mu)^2}{2\sigma}\right)dx=...=\mu$$

Para la varianza se usa un argumento similar, de forma que,

$$\operatorname{Var}(X) = \mathbb{E}((X-\mu)^2) = ... = \sigma^2$$

Ejemplo 2.25: Sea $X \sim N(\mu_1, \sigma_1)$ y $Y \sim N(\mu_2, \sigma_2)$. Si X e Y son variables aleatorias independientes, entonces $X+Y \sim N(\mu_1+\mu_2, \sigma_1+\sigma_2)$. Vamos a demostrar el caso cuando $\mu_1 = \mu_2 = 0$ y con $\sigma_1 = a, \sigma_2 = b > 0$ ya que sigue el mismo procedimiento.

Sabeos que la densidad de $X+Y$ es,

$$\int_{\mathbb{R}} f_X(z-x) f_Y(x) d\lambda$$

donde f_X, f_Y son las densidades de X, Y respectivamente. En particular, $f_X \cdot f_Y$ es $L^1(\mathbb{R})$ además de estar bien definida como integral de Riemann, por lo que simplemente debemos estudiar,

$$\int_{-\infty}^{\infty} f_Y(z-x) f_X(x) dx$$

con z fijo. Entonces tenemos,

$$\begin{aligned}\int_{-\infty}^{\infty} f_X(z-x) f_Y(x) dx &= \int_{-\infty}^{\infty} \frac{1}{2\pi\sqrt{ab}} e^{-\frac{x^2}{2a}} e^{-\frac{(z-x)^2}{2b}} dx \\ &= \frac{1}{2\pi\sqrt{ab}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x^2}{a}+\frac{(z-x)^2}{b}\right)} dx\end{aligned}$$

Estudiemos la potenica de e.

$$-\frac{1}{2}\left(\frac{x^2}{a} + \frac{(z-x)^2}{b}\right) = -\frac{(a+b)}{2ab}\left(x - \frac{a}{a+b}z\right)^2 - \frac{z^2}{2(a+b)}$$

Luego,

$$\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x^2}{a}+\frac{(z-x)^2}{b}\right)} dx = e^{-\frac{z^2}{(a+b)}} \int_{-\infty}^{\infty} e^{-\frac{(a+b)}{2ab}\left(x-\frac{a}{a+b}z\right)^2} dx$$

Ahora necesitamos determinar la integral.

$$\int_{-\infty}^{\infty} e^{-\frac{(a+b)}{2ab}\left(x-\frac{a}{a+b}z\right)^2} dx = 2\int_{az/(a+b)}^{\infty} e^{-\frac{(a+b)}{2ab}\left(x-\frac{a}{a+b}z\right)^2} dx$$

Tomando el siguiente cambio de variable,

$$s = \frac{(a+b)}{2ab}\left(z - \frac{a}{a+b}z\right)^2$$

obteenmos,

$$2\int_{az/(a+b)}^{\infty} e^{-\frac{(a+b)}{2ab}\left(x-\frac{a}{a+b}z\right)^2}dx = 2\sqrt{\frac{(a+b)\pi}{2ab}}$$

Finalmente,

$$\int_{\mathbb{R}} f_Y(z-x)f_X(x)d\lambda = \frac{1}{\sqrt{2\pi(a+b)}}e^{-\frac{z^2}{2(a+b)}}$$

Es decir, $X+Y \sim N(0, a+b)$, como queriamos probar.

Teorema 2.26 (Extensión de Kolmogorov): *Sea $(\Omega, \mathcal{F}, P)$ espacio de probabilidad.* ***reviar*** *Consideremos $\{\mu_n\}_{n\in\mathbb{N}}$ una colección de distribuciones en $(\mathbb{R}^n, \mathcal{R}^n)$ consistentes, es decir, para todo $n \in \mathbb{N}$ se cumple que,*

$$\mu_{n+1}((a_1,b_1]\times\cdots\times(a_n,b_n]\times\mathbb{R}) = \mu_n((a_1,b_1]\times\ldots(a_n,b_n])$$

Entonces, existe una única medida μ en $(\mathbb{R}^{\mathbb{N}}, \mathcal{R}^{\mathbb{N}})$ tal que,

$$\mu(\omega = (\omega_1,\omega_2,\ldots)\in\mathbb{R}^{\mathbb{N}}, \omega_1\in(a_1,b_1],\ldots,\omega_n\in(a_n,b_n]) = \mu_n((a_1,b_1]\times\ldots(a_n,b_n])$$

para todo $n \in \mathbb{N}$. Donde $\mathcal{R}^{\mathbb{N}}$ es el σ-álgebra generado por conjuntos de la forma:

$$\{\omega = (\omega_1,\omega_2,\ldots)\in\mathbb{R}^{\mathbb{N}} : \omega_1\in A_1,\ldots,\omega_n\in A_n\}$$

donde $A_i \in \mathcal{R}$.

La extensión de Kolmogorov, nos dice que podemos extender de forma única medidas consistentes a un Boreleano numerable como $\mathcal{R}^{\mathbb{N}}$. Así por ejemplo, una moneda puede ser lanzado infinitamente, generando distribuciones, el problema es que en el infinito no sabemos como se comporta. Con este teorema sabemos que sigue la recurrencia de la distribución en casos finitos y además que es única la probabilidad.

Proposición 2.27: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean $X_1,\ldots,X_n$ variables aleatorias no correlacionadas (o independientes). Se tiene que,*

$$Var\left(\sum_{i=1}^{n} X_n\right) = \sum_{i=1}^{n} Var(X_i)$$

(si la varianza $Var(X_i)$ existe). También se cumple que para todo $c\in\mathbb{R}$ que,

$$Var(cX) = c^2\,Var(X)$$

Dem: Consideremos $X_1,\ldots,X_n$ variables aleatorias no correlacionadas con esperanza $\mu_i :=$

$\mathbb{E}(X)$. Por definición se tiene que,

$$\begin{aligned}\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) &= \mathbb{E}\left(\left(\sum_{i=1}^{n} X_i - \mathbb{E}\left(\sum_{i=1}^{n} X_i\right)\right)^2\right)\\ &= \mathbb{E}\left(\left(\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \mu_i\right)^2\right)\\ &= \mathbb{E}\left(\left(\sum_{i=1}^{n} X_i - \mu_i\right)^2\right)\end{aligned}$$

Observemos que,

$$\left(\sum_{i=1}^{n} X_i - \mu_i\right)^2 = \sum_{i,j=1,\ldots,n} (X_i - \mu_i)(X_j - \mu_j)$$

Para llegar al resulado, queremos probar que $\mathbb{E}((X_i - \mu_i)(X_j - \mu_j)) = 0$ y en efecto, ya que por propiedades de esperanza, se tiene que,

$$\begin{aligned}\mathbb{E}((X_i - \mu_i)(X_j - \mu_j)) &= \mathbb{E}(X_iX_j - X_i\mu_j - X_j\mu_i + \mu_i\mu_j)\\ &= \mathbb{E}(X_i)\mathbb{E}(X_j) - \mu_j\mathbb{E}(X_i) - \mu_i\mathbb{E}(X_j) + \mu_i\mu_j\\ &= 0\end{aligned}$$

Por tanto,

$$\begin{aligned}\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) &= \mathbb{E}\left(\sum_{i,j=1,\ldots,n} (X_i - \mu_i)(X_j - \mu_j)\right)\\ &= \sum_{i=1}^{n} \mathbb{E}((X_i - \mu_i)^2)\\ &= \sum_{i=1}^{n} \operatorname{Var}(X_i)\end{aligned}$$

Ahora consideremos $c \in \mathbb{R}$, luego por definición de varianza sobre una variable aleatoria X se tiene que,

$$\begin{aligned}\operatorname{Var}(cX) &= \mathbb{E}((cX - \mathbb{E}(cX))^2)\\ &= \mathbb{E}((cX - c\mathbb{E}(X))^2)\\ &= \mathbb{E}(c^2(X - \mathbb{E}(X))^2)\\ &= c^2\mathbb{E}((X - \mathbb{E}(X))^2)\\ &= c^2\operatorname{Var}(X)\end{aligned}$$

■

2.2. Ley Débil de los Grandes Números

Teorema 2.28 (Desigualdad de Chebysher): *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sea X un avariable aleatoria que es $X \geq 0$ casi siempre (es $\mathbb{P}$-ctp), $a > 0$. Entonces,*

$$\mathbb{P}(X \geq a) \leq \frac{1}{a}\mathbb{E}(X)$$

Dem: Por definición se tiene que,

$$\begin{aligned}
\mathbb{E}(X) &= \mathbb{E}(X\mathbb{1}_{\{X\geq a\}}) + \mathbb{E}(X\mathbb{1}_{\{X>a\}}) \\
&\geq \mathbb{E}(X\mathbb{1}_{\{X\geq a\}}) \\
&= \int_\Omega X\mathbb{1}_{\{X\geq a\}} d\mathbb{P} \\
&= a\int_\Omega \frac{X}{a}\mathbb{1}_{\{X\geq a\}} d\mathbb{P} \\
&\geq a\int_\Omega \mathbb{E}(\mathbb{1}_{\{X\geq a\}}) d\mathbb{P} \\
&= a\mathbb{P}(X \geq a)
\end{aligned}$$

Por tanto,

$$\mathbb{P}(X \geq a) \leq \frac{1}{a}\mathbb{E}(X)$$

■

Teorema 2.29 (Ley de los grandes números con varianza finita): *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de medida. Sean $\{X_i\}_{i\in\mathbb{N}}$ variables aleatorias i.i.d (independientes e idénticamente distribuidas) tal que $\mathbb{E}(X_i^2) < \infty$. Digamos que $S_n = X_1 + \cdots + X_n$ y que $\mu := \mathbb{E}(X_i)$ para todo $i = 1, \ldots, n$. Entonces para todo $\varepsilon > 0$ se tiene que,*

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \overset{n\to\infty}{\longrightarrow} 0$$

También se cumple que $S_n/n \to \mu$ en L_2. Dicho de otra forma, S_n/n converge a μ en probabilidad (análogo a medida) y en L^2.

Dem: Sea $\varepsilon > 0$ fijo. Notemos que,

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) = \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right|^2 \geq \varepsilon^2\right)$$

Notemos que,

$$\left|\frac{S_n}{n} - \mu\right|^2$$

es una variable aleatoria positiva. Por lo que podemos aplicar la desigualdad anterior, obteniendo,

$$\mathbb{P}\left(\left|\frac{S_n}{n}-\mu\right|^2 \geq \varepsilon^2\right) \leq \frac{1}{\varepsilon^2}\mathbb{E}\left(\left(\frac{S_n}{n}-\mu\right)^2\right)$$

Ovservemos que,

$$\begin{aligned}\mathbb{E}\left(\left(\frac{S_n}{n}-\mu\right)^2\right) &= \frac{1}{n^2}\mathrm{Var}(S_n)\\ &= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i)\\ &= \frac{1}{n}\mathrm{Var}(X_i) \overset{n\to\infty}{\longrightarrow} 0\end{aligned}$$

Por tanto,

$$\mathbb{P}\left(\left|\frac{S_n}{n}-\mu\right| \geq \varepsilon\right) \overset{n\to\infty}{\longrightarrow} 0$$

para todo $\varepsilon > 0$. Esto implica que converge en medida y en medida $L2$. ■

El siguiente teorema es más una aplicación de la ley de los grande números.

Teorema 2.3 (Teorema de Bernstein): *Sea $f:[0,1]\to\mathbb{R}$ una función continua. Definimos las funciones $f_n:[0,1]\to\mathbb{R}$ dada por:*

$$f_n(x)=\sum_{m=0}^{n}\binom{n}{m}x^m(1-x)^{n-m}f\left(\frac{m}{n}\right)$$

Entonces f_n converge uniformemente a f.

no será $f:[0,1]\to[0,1]$

Dem: Consideremos $(\{0,1\},\mathcal{P}(\{0,1\}),\mathbb{P})$ un espacio de probabilidad y sean $\{X_i^{(x)}\}_{i\in\mathbb{N}}$ variables aleatorias independientes e identicamente distribuidas tales que,

$$\mathbb{P}(X_i=1)=x,\ \mathbb{P}(X_i=0)=1-x$$

Entonces $X_i\,$sinBernoulli(x) y $S_n=X_1+\cdots+X_n$ tiene distribución Binomial(n,x). Notemos que, **revisar**

$$\mathbb{E}\left(f\left(\frac{S_n}{n}\right)\right)=\int_{\mathbb{R}} x..$$

terminar

Teorema 2.31 (Ley de los Grandes Números para arreglos Triangulares): *Sea* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espacio de probabilidad. Consideremos* $X_{n,k}$ *para* $n \in \mathbb{N}$ *y* $k = 1, \ldots, n$*; variables aleatorias. Sean,*

$$\begin{aligned} S_n &:= \sum_{k=1}^{n} X_{n,k} \\ \mu_n &:= \mathbb{E}(S_n) \\ \sigma_n^2 &:= Var(S_n) \end{aligned}$$

Entonces, si,

$$\lim_{n\to\infty} \frac{\sigma_n^2}{b_n^2} = 0$$

donde b_n *es una sucesión tal que el límite anterior esté bien definido. Entonces se tiene que,*

$$\frac{S_n - \mu_n}{b_n} \longrightarrow 0$$

en probabilidad.

Dem: Notemos que.

$$\begin{aligned} \mathbb{P}\left(\left|\frac{S_n - \mu_n}{b_n}\right| > \varepsilon\right) &= \mathbb{P}\left(\left|\frac{S_n - \mu_n}{b_n}\right|^2 > \varepsilon^2\right) \\ &\leq \frac{1}{\varepsilon^2}\mathbb{E}\left(\left|\frac{S_n - \mu_n}{b_n}\right|^2\right) \\ &= \frac{1}{\varepsilon^2}\frac{1}{b_n^2}\mathrm{Var}(S_n) \overset{n\to\infty}{\longrightarrow} 0 \end{aligned}$$

Probando que,

$$\frac{S_n - \mu_n}{b_n} \longrightarrow 0$$

en probabilidad. ■

Ejemplo 2.32: Supongamos que coleccionamos un álbum con n láminas. Sea $\mathcal{T}_k^n$ = cantidad de láminas que comparamos hasta tener k láminas no repetidas. Entonces $\mathcal{T}_n^n$ es la cantindad de lámina que comparamos hasta completar el album. Se tiene que,

$$\mathcal{T}_n^n = \underbrace{\mathcal{T}_1^n}_{X_1^n} + \underbrace{(\mathcal{T}_2^n - \mathcal{T}_1^n)}_{X_2^n} + \cdots + \underbrace{(\mathcal{T}_n^n - \mathcal{T}_{n-1}^n)}_{X_n^n}$$

Donde X_i^n son variables aleatorias para todo $i = 1, \ldots, n$. Observemos que,

$$X_i^n \sim \mathrm{Geom}\left(1 - \frac{i-1}{n}\right)$$

En general,

$$\mathbb{E}(\text{Geom}(p)) = \frac{1}{p}$$
$$\text{Var}(\text{Geom}(p)) = \frac{1}{p^2}$$

Luego la esperanza cumple,

$$\mathbb{E}(\mathcal{T}_n^n) = \sum_{i=1} n\left(1 - \frac{i-1}{n}\right)^{-1}$$
$$= n\sum_{i=1}^{n}\frac{1}{i} \sim n\log(n)$$

Por tanto,

$$\sum_{i=1}^{n}\frac{1}{i} \sim \log(n)$$

Y la varianza cumple,

$$\text{Var}(\mathcal{T}_n^n) = \sum_{i=1}^{n}\left(1 - \frac{i-1}{n}\right)^2 = n^2\sum_{i=1} ni^{-2}$$

Tomando $b_n := n\log(n)$ se obtiene que,

$$\text{Var}(S_n) = \frac{\sigma_n^2}{b_n^2} \overset{n\to\infty}{\longrightarrow} 0$$

Por lo tanto,

$$\frac{\mathcal{T}_n^n - n\log(n)}{n\log(n)} \longrightarrow 0$$

en probabilidad.

Teorema 2.33 (Ley débil de los grandes números Versió Estándar): *Sea $\{X_i\}_{i\in\mathbb{N}}$ variables aleatorias independientes e identicamente distribuidos con $\mathbb{E}(|X_1|) < \infty$. Entonces se tiene que,*

$$\frac{S_n}{n} \longrightarrow \mu$$

en probabilidad, donde $\mu = \mathbb{E}(X_1)$.

Dem Durret

Ejemplo 2.34: Estudiemos la paradoja de Sam Petersburgo. **blabla** consideremos variables aleatorias tales que,

$$\mathbb{P}(X_1 = 2^j) = \frac{1}{2^j}$$

para todo $j \in \mathbb{N}$. Si $\mathbb{E}(X_1) = \infty$, entonces,

$$\frac{S_n}{n\log(n)} \longrightarrow 1$$

en probabilidad.

Dem (2.7): Vamos a usar el teorema 2.6, se tiene que,

$$\begin{aligned}\mathbb{P}(|X_1| > x) &= x\mathbb{E}\left(\mathbb{1}_{\{|X_1|>x\}}\right)\\ &= \mathbb{E}\left(x\mathbb{1}_{\{|X_1|>x\}}\right)\\ &\leq \mathbb{E}\left(|X_1|\mathbb{1}_{\{|X_1|>x\}}\right) \longrightarrow \mathbb{E}(0) = 0\end{aligned}$$

por el teorema de converne **blabla**. $|X_1| \in L_1$.

La ley débil de los grandes números nos dice que, para una sucesión de variables aleatorias $\{X_i\}_{i\in\mathbb{N}}$ independientes e identicamente distribuidas con $\mathbb{E}(|X_i|) < \infty$, y considerando $\mu = \mathbb{E}(X_i)$. Entonces para todo $\varepsilon > 0$ se tiene que,

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \overset{n\to\infty}{\longrightarrow} 0$$

donde $S_n = X_1 + \cdots + X_n$. Ahora veremos una versión pero con condiciones más débiles.

Teorema 2.35 (Ley muy débil de los grandes números): *Sea $\{X_i\}_{i\in\mathbb{N}}$ una sucesión de variables aleatorias independientes e indeticamente distribuidas tales que,*

$$x\mathbb{P}(|X_i| \geq x) \overset{n\to\infty}{\longleftarrow} 0$$

para todo $i \geq 1$. Sea $\mu_n := \mathbb{E}(|X_1|\mathbb{1}_{\{|X_1|\leq n\}})$. Entonces,

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu_n\right| \geq \varepsilon\right) \overset{n\to\infty}{\longrightarrow} 0$$

Dem: Observemos que,

$$\mu_n = \mathbb{E}\left(|X_1|\mathbb{1}_{\{|X_i|\leq n\}}\right)$$

para todo $i \in \mathbb{N}$, debido a que X_i tienen igual distribución. Sea $\varepsilon > 0$, definimos,

$$\overline{S}_n := \sum_{i=1}^{n} X_i\mathbb{1}_{\{|X_i|\leq n\}}$$

Entonces se cumple la siguiente desigualdad:

$$\mathbb{P}\left(\left|\frac{S_n}{n}-\mu_n\right|\geq\varepsilon\right)\leq\underbrace{\mathbb{P}(S_n\neq\overline{S}_n)}_{=:A}+\underbrace{\mathbb{P}\left(\left|\frac{\overline{S}_n}{n}-\mu_n\right|\geq\varepsilon\right)}_{=:B}$$

Vamos a encontrar cotas superiores para A, B.

- **Cota para** A**:** Por definición,

$$\begin{aligned}\mathbb{P}(S_n\neq\overline{S}_n)&\leq\mathbb{P}\left(\bigcup_{i=1}^{n}\left\{|x_i|\mathbb{1}_{\{|X_i|\leq n\}}\neq x_i\right\}\right)\\&=\mathbb{P}\left(\bigcup_{i=1}^{n}\{|X_i|>n\right)\\&\leq\sum_{i=1}^{n}\mathbb{P}(\{|X_i|>n\})=n\mathbb{P}(\{|X_i|>n\})\overset{n\to\infty}{\longrightarrow}0\end{aligned}$$

  Cota para B : Notemos que,

$$\mathbb{P}\left(\left|\frac{\overline{S}_n}{n}-\mu_n\right|\geq\varepsilon\right)=\mathbb{P}\left(\left|\frac{\overline{S}_n}{n}-\mu_n\right|^2\geq\varepsilon^2\right)$$

  Luego por la desigualdad **cebicher** se cump,e que,

$$\begin{aligned}\mathbb{P}\left(\left|\frac{\overline{S}_n}{n}-\mu_n\right|\geq\varepsilon\right)&\leq\frac{1}{\varepsilon^2}\mathbb{E}\left(\left(\frac{\overline{S}_n}{n}-\mu_n\right)^2\right)\\&=...\\&=\frac{1}{\varepsilon^2 n}\mathrm{Var}\left(X_1\mathbb{1}_{\{|X_1|\leq n\}}\right)\end{aligned}$$

  Para continudar necesitamos el siguiente lema:

  Lema 2.26: *Sea* $(\Omega,\mathcal{F},\mathbb{P})$ *un espacio de probabilidad. Sea* $Y:\Omega\to\mathbb{R}$ *una variable aleatorias tal que* $Y\geq 0$ *casi seguramente. Entonces,*

$$\mathbb{E}(Y^p)=\int_0^\infty py^{p-1}\mathbb{P}(Y>y)dy$$

Dem: Tenemos que,

$$\begin{aligned}\int_0^\infty py^{p-1}\mathbb{P}(Y>y)dy &= \int_0^\infty py^{p-1}\mathbb{E}\left(\mathbb{1}_{\{Y>y\}}\right)dy\\ &= \int_0^\infty \mathbb{E}\left(py^{p-1}\mathbb{1}_{\{Y>y\}}\right)dy\\ &= \textbf{fubini}\\ &= \mathbb{E}\left(\int_0^\infty py^{p-1}\mathbb{1}_{\{Y>y\}}dy\right)\\ &= \mathbb{E}\left(\int_0^Y py^{p-1}dy\right)\\ &= \mathbb{E}(Y^p)\end{aligned}$$

■

Volvamos a la cota de B. Con el lema anterior tenemos que,

$$\frac{1}{\varepsilon^2 n}\mathrm{Var}\left(X_1\mathbb{1}_{\{|X_1|\le n\}}\right) \le \frac{1}{\varepsilon^2 n}\mathbb{E}\left(\left(X_1\mathbb{1}_{\{|X_1|\le n\}}\right)^2\right)$$

puesto que $\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$. Finalmente,

$$\begin{aligned}\frac{1}{\varepsilon^2 n}\mathbb{E}\left(\left(X_1\mathbb{1}_{\{|X_1|\le n\}}\right)^2\right) &\le \frac{1}{\varepsilon^2 n}\int_0^\infty 2y\mathbb{P}\left(|X_1|\mathbb{1}_{\{|X_1|\le n\}} > y\right)dy\\ &\le \frac{2}{\varepsilon^2 n}\int_0^n y\mathbb{P}(|X_1|>y)dy\\ &= \frac{2}{\varepsilon^2}\underbrace{\left(\int_0^n y\mathbb{P}(|X_1|>y)dy\right)}_{\text{promedio}}\end{aligned}$$

Por hipótesis sabemos que $x\mathbb{P}(|X_1|x) \to 0$ cuando $y \to \infty$. Entonces el promedio integral también converge a cero.

Probando que,

$$\mathbb{P}\left(\left|\frac{\overline{S}_n}{n} - \mu_n\right| \ge \varepsilon\right) \overset{n\to\infty}{\longrightarrow} 0$$

■

2.3. Ley Fuerte de los Grandes Números

Definición 2.27: *Sea $(\Omega, \mathcal{F})$ un espacio medible. Sea $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ una colección. Definimos el límite superior de la colección $\{A_n\}_{n\in\mathbb{N}}$ por:*

$$\limsup_{n\to\infty} A_n := \bigcap_{n\in\mathbb{N}}\bigcup_{k\ge n} A_n$$

Análogamente se define el límite inferior por:

$$\liminf_{n\to\infty} A_n := \bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} A_n$$

Nota 2.28: Se puede definir el límite superior e inferior sobre cualquier colección de conjuntos.

Proposición 2.29: *Sea* $(\Omega, \mathcal{F})$ *un espacio medible. Entonces,*

$$\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n$$

También se cumple que,

$$\limsup_{n\to\infty} \mathbb{1}_{A_n} = \mathbb{1}_{\limsup_{n\to\infty} A_n}$$
$$\liminf_{n\to\infty} \mathbb{1}_{A_n} = \mathbb{1}_{\liminf_{n\to\infty} A_n}$$

Lema (Borel-Cantelli) 2.30: *Sea* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espacio de probabilidad. Sea* $\{A_n\}_{n\in\mathbb{N}}$ *una secuencia de eventos. Si,*

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < \infty$$

Entonces,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 0$$

Dem 1: Si la serie,

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) < \infty$$

Entonces se tiene que la serie,

$$\sum_{k=n}^{\infty} \mathbb{P}(A_k) \overset{n\to\infty}{\longrightarrow} 0$$

Ahora por definición de límite superir se cumple que,

$$\begin{aligned}\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) &= \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k\geq n} \mathbb{P}(A_k)\right)\\ &\leq \lim_{n\to\infty} \sum_{k=n}^{\infty} \mathbb{P}(A_k) = 0\end{aligned}$$

Probando el lema. ■

Dem 2: Consideremos,

$$N = \sum_{i=1}^{\infty} \mathbb{1}_{A_i}$$

Luego como A_i es medible, se tiene que N es una variable aleatoria. Luego,

$$\begin{aligned}\mathbb{E}(N) &= \mathbb{E}\left(\sum_{i=1}^{\infty} \mathbb{1}_{A_i}\right) \\ &= \sum_{i=1}^{\infty} \mathbb{E}(\mathbb{1}_{A_i}) \\ &= \sum_{i=1}^{\infty} \mathbb{P}(A_i) < \infty\end{aligned}$$

Por lo que $\mathbb{E}(N) < \infty$. Entonces,

$$\mathbb{P}(N = \infty) = 0$$

En particular,

$$\{\omega \in \Omega : N = \infty\} = \limsup_{n\to\infty} A_n$$

De esta forma,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 0$$

■

Teorema 2.31 (Ley fuerte de los grandes númetos con cuarto momento finito): *Sea $\{X_i\}_{i\in\mathbb{N}}$ independientes e indeticamente distribuidad con $\mathbb{E}(X_i^4) < \infty$. Entonces,*

$$\mathbb{P}\left(\lim_{n\to\infty} \frac{S_n}{n} = \mu\right) = 1$$

donde $\mu = \mathbb{E}(X_i)$

Dem: Sea $\varepsilon > 0$, entonces se tiene que,

$$\begin{aligned}\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) &= \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right|^4 > \varepsilon^4\right) \\ &\leq \frac{1}{\varepsilon^4}\mathbb{E}\left(\left(\frac{S_n}{n} - \mu\right)^4\right)\end{aligned}$$

Estudiemos el cuarto momento no centralizado. Observemos que,

$$\begin{aligned}\frac{S_n}{n}-\mu &= \frac{S_n-n\mu}{n}\\ &= \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)\end{aligned}$$

Consideremos $\overline{X}_i = X_i - \mu$, de esta forma,

$$\begin{aligned}\mathbb{E}\left(\left(\frac{S_n}{n}-\mu\right)^4\right) &= \mathbb{E}\left(\left(\sum_{i=1}^{n}\overline{X}_i\right)^4\right)\frac{1}{n^4}\\ &= \mathbb{E}\left(\sum_{\substack{i_1,i_2,i_3,i_4\\ \in\{1,\dots,n\}}}\overline{X}_{i_1}\overline{X}_{i_2}\overline{X}_{i_3}\overline{X}_{i_4}\right)\frac{1}{n^4}\\ &= \sum_{\substack{i_1,i_2,i_3,i_4\\ \in\{1,\dots,n\}}}\mathbb{E}\left(\overline{X}_{i_1}\overline{X}_{i_2}\overline{X}_{i_3}\overline{X}_{i_4}\right)\frac{1}{n^4}\end{aligned}$$

Como $\{X_n\}_{n\in\mathbb{N}}$ son variables aleatorias independientes, entonces podemos separar los $\overline{X}_{i_1},\overline{X}_{i_2}$ siempre y cuando $i_1 \neq i_2$, ya que si son iguales, no necesariamente son no correlacionados. Notemos lo siguiente,

$$\mathbb{E}\left(\overline{X}_{i_1}\overline{X}_{i_2}\overline{X}_{i_3}\overline{X}_{i_4}\right) = \begin{cases}\text{Todos distintos:} & \mathbb{E}\left(\overline{X}_{i_1})\mathbb{E}(\overline{X}_{i_2})\mathbb{E}(\overline{X}_{i_3})\mathbb{E}(\overline{X}_{i_4}\right)\\ \text{Un par iguales:} & \mathbb{E}(\overline{X}_{i_1}^2)\mathbb{E}(\overline{X}_{i_2})\mathbb{E}(\overline{X}_{i_3})\\ \text{Un trío iguales:} & \mathbb{E}(\overline{X}_{i_1}^3)\mathbb{E}(\overline{X}_{i_2})\\ \text{Todos iguales:} & \mathbb{E}(\overline{X}_{i_1}^4)\\ \text{Dos pares iguales:} & \mathbb{E}(\overline{X}_{i_1}^2)\mathbb{E}(\overline{X}_{i_2}^2)\end{cases}$$

Por otro lado se cumple que $\mathbb{E}(\overline{X}_i) = 0$ por linealidad de la esperanza, por tanto los únicos casos que sobreviven son cuando las cuatro variables aleatorias son iguales o cuando dos pares son iguales, por lo que todo se reduce a contar la cantidad de estas, de forma que,

$$\begin{aligned}\mathbb{E}\left(\left(\frac{S_n}{n}-\mu\right)^4\right) &= \sum_{i=1}^{n}\mathbb{E}(\overline{X}_i^4) + \sum_{i\neq j}\mathbb{E}(\overline{X}_i^2)\mathbb{E}(\overline{X}_j^2)\\ &= n\mathbb{E}(\overline{X}_1^4) + \frac{(n-1)n}{2}\binom{4}{2}\mathbb{E}(\overline{X}_1^2)^2\end{aligned}$$

Por lo tanto,

$$\begin{aligned}\mathbb{P}\left(\left|\frac{S_n}{n}-\mu\right| > \varepsilon\right) &\leq \frac{1}{\varepsilon^4 n^4}\left(n\mathbb{E}(\overline{X}_1^4) + \frac{(n-1)}{2}\binom{4}{2}\mathbb{E}(\overline{X}_1^2)^2\right)\\ &= \frac{1}{\varepsilon^4}\left(\frac{1}{n^3}\mathbb{E}(\overline{X}_1^4) + \frac{3(n-1)}{n^3}\mathbb{E}(\overline{X}_1^2)^2\right) \xrightarrow{n\to\infty} 0\end{aligned}$$

Esto es para todo $\varepsilon > 0$, ahora para concluir debemos notar que lo anterior implica que,

$$\mathbb{P}\left(\limsup_{n\to\infty}\left|\frac{S_n}{n}-\mu\right|>\varepsilon\right)=0$$

Tomando $\varepsilon = 1/i$ se obtiene que,

$$\mathbb{P}\left(\bigcup_{i\in\mathbb{N}}\left|\frac{S_n}{n}-\mu\right|>\frac{1}{i}\right)\leq\sum_{i=1}^{\infty}\mathbb{P}\left(\limsup_{n\to\infty}\left|\frac{S_n}{n}-\mu\right|>\frac{1}{i}\right)=0$$

■

Teorema 2.32: *Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias. Entonces $X_n \to X$ en probabilidad para alguna variable aleatoria X si y sólo si, toda subsucesión $\{X_{n_k}\}_{k\in\mathbb{N}}$ tiene una subsucesión $\{X_{n_{k_j}}\}$ tal que converge a X casi seguro.*

Dem: Supongamos que $X_n \to X$ en probabilidad. Entonces de teoría de la medida, se cumple que $X_n \to X$ casi seguramente, esto implica que toda subsucesión tiene una subsucesión que converge a X casi seguramente.

Supongamos que toda subsucesión **terminar**

Del lema de Borel-Cantelli sabemos que si una serie de eventos es finita,

$$\sum_{n\in\mathbb{N}}\mathbb{P}(A_n)<\infty$$

Entonces la probabilidad del límite superior de los eventos es 0. Por lo que naturalmente uno se puede preguntar que pasa si la serie de los eventos es infinita,

$$\sum_{n\in\mathbb{N}}\mathbb{P}(A_n)=\infty$$

Entonces la probabilidad del límite superior de los eventos debe ser 1 o bien, positivo. Sin embargo esto no es necesariamente cierto. Consideremos el siguiente espacio de probabilidad $(\Omega,\mathcal{F},\mathbb{P})=([0,1],\mathcal{B}([0,1]),\lambda)$, sean los evento $A_i=[0,1/i]$, luego,

$$\mathbb{P}(A_i)=\frac{1}{i}$$

Y entonces,

$$\sum_{i\in\mathbb{N}}\frac{1}{i}=\infty$$

Sin embargo,

$$\limsup_{n\to\infty}A_n=\bigcap_{n\in\mathbb{N}}\left[0,\frac{1}{n}\right]=\{0\}$$

Por lo que,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 0$$

Pero, podemos arreglar todo esto pidiendo que los eventos sean independientes, obteniendo el segundo lema de Borel-Cantelli.

Lema 2.33 (Borel-Cantelli II): *Sean $\{A_n\}_{n\in\mathbb{N}}$ una secuencia de eventos independientes tales que,*

$$\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \infty$$

Entonces,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 1$$

Dem: Si $\{A_n\}_{n\in\mathbb{N}}$ son eventos independientes, entonces $\{A_n^c\}$ también son independientes, luego,

$$\mathbb{P}\left(\bigcap_{i=M}^{N} A_i^c\right) = \prod_{i=M}^{N}(1-\mathbb{P}(A_i))$$

Ahora consideremos la siguiente desigualdad $1-x \leq e^{-x}$ para todo $x \in \mathbb{R}$, entonces,

$$\begin{aligned}\prod_{i=M}^{N}(1-\mathbb{P}(A_i)) &\leq \prod_{i=M}^{N}\exp(-\mathbb{P}(A_i)) \\ &= \exp\left(-\sum_{i=M}^{N}\mathbb{P}(A_i)\right) \overset{N\to\infty}{\longrightarrow} 0\end{aligned}$$

Entonces,

$$\mathbb{P}\left(\bigcup_{i\geq M} A_i\right) = 1$$

para todo $M \in \mathbb{N}$. Para concluir observamos que,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{i=n}^{\infty} A_i\right) = 1$$

■

Teorema 2.34: *Sean $\{X_n\}_n$ variables aleatorias iid con $\mathbb{E}(|X_1|) < \infty$. Entonces,*

$$\mathbb{P}\left(\lim_{n\to\infty}\frac{S_n}{n} \in (-\infty,\infty)\right) = 0$$

Dem 14/09

Teorema 2.35: *Sean A_i eventos independientes a pares tales que,*

$$\sum_{i=1}^{\infty} \mathbb{P}(A_i) = \infty$$

Entonces,

$$\frac{\sum_{i=1}^{n} \mathbb{1}_{A_i}}{\sum_{i=1}^{n} \mathbb{P}(A_i)} \longrightarrow 1$$

en probabulidad. Es decir,

$$\sum_{i=1}^{n} \mathbb{1}_{A_i} \sim \sum_{i=1}^{n} \mathbb{P}(A_i)$$

Dem: Sea $X_n := \mathbb{1}_{A_1} + \cdots + \mathbb{1}_{A_n}$. Sea $\varepsilon > 0$, luego se cumple que,

$$\begin{aligned}
\mathbb{P}\left(\left|\frac{X_n}{\mathbb{E}(X_n)} - 1\right| > \varepsilon\right) &= \mathbb{P}\left(\left|\frac{X_n}{\mathbb{E}(X_n)} - 1\right|^2 > \varepsilon^2\right) \\
&= \mathbb{P}\left(|X_n - \mathbb{E}(X_n)|^2 > \varepsilon^2\mathbb{E}^2(X_n)\right) \\
&\leq \frac{1}{\varepsilon^2\mathbb{E}^2(X_n)}\mathrm{Var}(X_n)
\end{aligned}$$

Notemos lo siguiente,

$$\begin{aligned}
\mathrm{Var}(X_n) &= \mathbb{E}\left(\left(\sum_{i=1}^{n}(\mathbb{1}_{A_i} - \mathbb{P}(A_i))\right)^2\right) \\
&= \mathbb{E}\left(\sum_{i=1}^{n}(\mathbb{1}_{A_i} - \mathbb{P}(A_i))^2 + \sum_{i \neq j}(\mathbb{1}_{A_i} - \mathbb{P}(A_i))(\mathbb{1}_{A_j} - \mathbb{P}(A_j))\right)
\end{aligned}$$

Si $\{A_n\}$ es una secuencia de eventos independientes a partes, entonces $\mathbb{1}_{A_n}$ son variables aleatorias independientes, en particular,

$$\mathbb{E}((\mathbb{1}_{A_i} - \mathbb{P}(A_i))(\mathbb{1}_{A_i} - \mathbb{P}(A_i))) = \mathbb{E}(\mathbb{1}_{A_i} - \mathbb{P}(A_i))\mathbb{E}(\mathbb{1}_{A_i} - \mathbb{P}(A_i)) = 0$$

Por tanto,

$$\mathrm{Var}(X_n) = \sum_{i=1}^{n} \mathbb{E}((\mathbb{1}_{A_i} - \mathbb{P}(A_i))^2) = \sum_{i=1}^{n} \mathrm{Var}(\mathbb{1}_{A_i})$$

Por tanto,

$$\frac{1}{\varepsilon^2\mathbb{E}^2(X_n)}\mathrm{Var}(X_n) = \frac{1}{\varepsilon^2\mathbb{E}^2(X_n)}\mathrm{Var}(\mathbb{1}_{A_i}) \overset{n\to\infty}{\longrightarrow} 0$$

Probando el teorema. ■

Teorema 2.36 (Ley Fuerte de los Grandes números): *Sea $\{X_i\}_{i\in\mathbb{N}}$ variables aleatorias independientes e identciamente distribuidas con esperanza finita. Sea $\mu = \mathbb{E}(X_i)$ y sea $S_n = X_1 + \cdots + X_n$, entonces,*

$$\frac{S_n}{n} \overset{n\to\infty}{\longrightarrow} \mu$$

casi seguramente.

2.4. Ley 0-1 Kolmogorov

Estudiemos un pequeño resultado pero importante que nos sirve para determinar la probabilidad de un evento si cumple cierta condición.

Necesitamos construir la σ-álgebra cola. Sea $\{X_i\}_{i\in\mathbb{N}}$ variables aleatorias. Definimos la σ-álgebra cola (o de la cola) como:

$$\mathcal{T} := \bigcap_{k\in\mathbb{N}} \sigma(X_k, X_{k+1}, \dots)$$

donde $\sigma(X_k, X_{k+1}, \dots)$ es la menor σ-álgebra donde $X_k, X_{k+1}, \dots$ son medibles.

Algunos eventos que pertenecen a la σ-álgebra cola son:

$$\left\{\sup_{n\in\mathbb{N}} X_n < \infty\right\}, \quad \left\{\inf_{n\in\mathbb{N}} X_n < \infty\right\}, \quad \left\{\lim_{n\to\infty} X_n < \infty\right\}, \quad \left\{\limsup_{n\to\infty} X_n < \infty\right\}$$

Entre otros.

Teorema 2.37 (0-1 Kolmogorov): *Sea $\{X_i\}_{i\in\mathbb{N}}$ variables aleatorias independientes. Entonces, si $A \in \mathcal{T}$, entonces $\mathbb{P}(A) = 0$ o $\mathbb{P}(A) = 1$.*

Dem: La ideal central de la demostración, es ver que si $A \in \mathcal{T}$ entonces A es independiente de si mismo, por lo que se cumple que,

$$\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}^2(A)$$

Entonces $\mathbb{P}(A)(1 - \mathbb{P}(A)) = 0$, concluyendo el resultado.

- **Primer Paso:** Sean $A \in \sigma(X_k, \dots)$ y $B \in \sigma(X_1, \dots, X_{k-1})$, entonces A es independiente de B. Si $A \in \sigma(X_k, \dots X_{k+i})$ entonces A es indpendiente de B.

  Consideremos,

$$\bigcup_{i=1}^{\infty} \sigma(X_k, \dots, X_{k+i})$$

Sean C, D conjunto de esta unión, luego existen i_1, i_2 tales que $C \in \sigma(X_k, \ldots, X_{k+i_1}), D \in \sigma(X_k, \ldots, X_{k+i_2})$. Pero C, D pertenecen a $\sigma(X_k, \ldots, X_{k+i})$ donde i es el mínimo de i_1, i_2. Entonces es un π-sistema, luego,

$$\bigcup_{i=1}^{\infty} \sigma(X_k, \ldots, X_{k+i}), \quad \sigma(X_1, \ldots, X_{k-1})$$

son π-sistemas independientes. Por otro lado, se cumple,

$$\sigma(\bigcup_{i=1}^{\infty} \sigma(X_k, \ldots, X_{k+i})) = \sigma(X_k, X_{k+1}, \ldots)$$

Concluyendo lo que queriamos probar.

- **Segundo Paso:** Vamos a probar que si $A \in \mathcal{T}$ y que $B \in \sigma(X_1, \ldots)$ entonces A, B son independientes.

  Si $B \in \sigma(X_1, \ldots, X_k)$ estamos listo puesto que $A \in \sigma(X_{k+1}, \ldots)$ y por el paso anterior son independiente.

  Ahora,

$$\bigcup_{i=1}^{\infty} \sigma(X_1, \ldots, X_i)$$

  es un π-sistema independiente de $\mathcal{T}$ y que,

$$\sigma\left(\bigcup_{i=1}^{\infty} \sigma(X_1, \ldots, X_i)\right) = \sigma(X_1, \ldots)$$

Con esto probamo el teorema. ■

3. Teorema del Límite Central

Un resultado fundamental en la teoría de probabilidad es el teorema del límite central. Consideremos $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e identicamente distribuidas tal que,

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = 1/2$$

Consideremos las sumas n-ésima suma parcial de las variables aleatorias,

$$S_n = \sum_{i=1}^{n} X_i$$

Observemos que X_n tiene distribución similar a Bernoulli, entonces,

$$\begin{aligned}\mathbb{E}(X) &= \int_\Omega X(\omega) d\mathbb{P}(\omega)\\ &= \mathbb{P}(\{\omega : X(\omega) = 1\}) - \mathbb{P}(\{\omega : X(\omega) = -1)\\ &= 0\end{aligned}$$

Y luego $\mathbb{E}(S_n) = 0$.

El teorema del límite central mostrará $S_n/\sqrt{n}$ tiene un comportamiento no trivial, pero antes observemos que la raiz cuadrada de n no es mero capricho, veamos que pasa si estudiamos S_n/n. Como $\mathbb{E}(X_n) = 0$, entonces por la ley fuerte de los grande números se cumple que,

$$\mathbb{P}\left(\lim_{n\to\infty} \frac{S_n}{n} = 0\right) = 1$$

Ahora se tiene que,

$$a \leq \frac{S_n}{n} \leq b \xrightarrow{\text{c.s}} a \leq 0 \leq b$$

Entonces se obtiene que,

$$\mathbb{P}\left(a \leq \frac{S_n}{n} \leq b\right) \overset{n\to\infty}{\longrightarrow} \begin{cases} 1, & a < 0, b > 0 \\ 0, & a > 0 \\ 0, & b < 0 \end{cases}$$

Es decir, que S_n/n tiene un comportamiento trivial al compararlo con $a < b$. Ahora veremos que en escala $\sqrt{n}$ se cumple que,

$$\mathbb{P}\left(a \leq \frac{S_n}{\sqrt{n}} \leq b\right) \overset{n\to\infty}{\longrightarrow} \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

Estudiemos un poco más profundamente S_n. Como $X_n \in \{-1, 1\}$, entonces,

$$S_n \in \{-n, -(n-1), \ldots, n-1, n\}$$

Ahora, como $-1, 1$ son impares, entonces si n es par, se tiene que S_n es un valor par, si n es impar, entonces S_n es impar. Por lo que el valor que tomen k está directamente condicionado por si n es par o impar. Determinemos la probabilidad de que $S_{2n} = 2k$ con $2k \in \{-2n, \dots, 2n\}$, entonces es simplemente contar las combinaciones de X_n de tal forma que la suma sea $2k$. De esta forma,

$$\mathbb{P}(S_{2n} = 2k) = \binom{2n}{n+k} \frac{1}{2^{2n}}$$

Para estudiar la probabilidad de que $S_{2n+1} = 2k+1$ se debe observar que, $S_{2n+1} = S_{2n} \pm 1$. Aquí debemos ver que $k = -(n+1), -n, \dots, n$, luego,

$$\mathbb{P}(S_{2n+1} = 2k+1) = \mathbb{P}(S_{2n} = 2k) + \mathbb{P}(S_{2n} = 2(k+1))$$

Si $k = -n, \dots, n-1$, entonces,

$$\begin{aligned}\mathbb{P}(S_{2n+1} = 2k+1) &= \left(\binom{2n}{n+k} + \binom{2n}{n+k+1} \right) \frac{1}{2^{2n}} \\ &= \binom{2n+1}{n+k+1} \frac{1}{2^{2n}}\end{aligned}$$

Teorema 3.1 (De Moivre-Laplace): *Sea $\{X_n\}_{n\in\mathbb{N}}$ variables aleatoria independientes e identicamente distribuidas tales que,*

$$\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$$

Entonces si $a < b$ se cumple que,

$$\mathbb{P}\left(a \leq \frac{S_n}{\sqrt{n}} \leq b \right) \overset{n\to\infty}{\longrightarrow} \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

Para la demostración del teorema es importante conocer la **fórmula de Stirling** que dice que,

$$n! \sim n^k e^{-n} \sqrt{2\pi n}$$

para $n \to \infty$, donde $\sim$ es la relación dada por,

$$a_n \sim b_n \iff \frac{a_n}{b_n} \overset{n\to\infty}{\longrightarrow} 1$$

Dem: terminar

Lema 3.2: *Si $a_n \to 0, c_n \to \infty$ y $a_n c_n \to \lambda$, entonces,*

$$(1 + a_n)^{c_n} \to e^{\lambda}$$

3.1. Convergecia Débil

Nuestro objetivo es demostrar el teorema del límite central, y para ello necesitamos una noción de convergencia entre variables aleatorias.

Definición 3.3: *Sean F_n, F funciones de distribución acumulada. Decimos que F_n* ***converge débilmente*** *a F si,*

$$\lim_{n\to\infty} F_n(x) = F(x)$$

para todo x donde F es continua. En tal caso denotamos $F_n \Rightarrow F$.

Observemos que no pedimos variables aleatorias ni espacio de probabilidad. Solo nos interesa sean funciones de distribución acumuladas, aunque se puede definir sobre cualquier tipo de función.

Definición 3.4: *Sean X_n, X variables aleatorias. Decimos que X_n* ***converge a X en distribución*** *si sus funciones de distribución convergen débilmente, en tal caso denotamos $X_n \Rightarrow X$.*

Ejemplo 3.5: Sea $\{X_n\}$ una secuencia de variables aleatorias independientes e indenticamente distribuidas tales que,

$$\mathbb{P}(X_i = -1) = \mathbb{P}(X_i = 1) = \frac{1}{2}$$

Consideremos $Y_n = \frac{S_n}{\sqrt{n}}$ donde $S_n = \sum_{i=1}^{n} X_i$. Entonces por el teorema de De Moivre-Laplace se cumple que,

$$F_n(z) = \mathbb{P}\left(\frac{S_n}{\sqrt{n}} \leq z\right) \overset{n\to\infty}{\longrightarrow} \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

Observamos que,

$$F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

es continua para todo $z \in \mathbb{R}$. Definimos,

$$F(z) := \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

que es la función distribución de una variable aleatoria X con distribución normal con media 0 y varianza 1. Claramente F es continua en todo z y por tanto, $F_n \Rightarrow F$. Por otro lado $X_n \Rightarrow X$.

Ejemplo 3.6: Sea X una variable aleatoria tal que $\mathbb{P}(X = 0) = 1$ con función distribución,

$$F(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

Sea $X_n = X + 1/n$ un variable aleatoria. Entonces $X_n \to X$ casi seguramente, por otro lado

$$\begin{aligned} F_n(x) &= \mathbb{P}(X_n \leq x) \\ &= \mathbb{P}\left(X \leq x - \frac{1}{n}\right) \\ &= F\left(x - \frac{1}{n}\right) \\ &= \begin{cases} 0, & x < \frac{1}{n} \\ 1, & x \geq \frac{1}{n} \end{cases} \end{aligned}$$

Por tanto, es claro que si F es continua en $x \in \mathbb{R}$, entonces,

$$\lim_{n\to\infty} F_n(x) = F(x)$$

Por tanto, $F_n \Rightarrow F$.

Ejemplo 3.7: Sea X^p una variable aleatoria con distribución,

$$\mathbb{P}(X^p = k) = (1-p)^k p$$

Luego,

$$\begin{aligned} \mathbb{P}(X^p \geq k) &= \sum_{i=k}^{\infty} (1-p)^i p \\ &= (1-p)^k p \sum_{i=k}^{\infty} (1-p)^{i-k} \\ &= (1-p)^k \end{aligned}$$

Sea $P_n = 1/n$ por lo que $P_n \to 0$ para $n \to \infty$ y sea $X_n = P_n X^{P_n}$. Sea F_n la distribución de X_n, entonces,

$$\begin{aligned} F_n(x) &= \mathbb{P}\left(\frac{1}{n} X^{1/n} \leq x\right) \\ &= 1 - \mathbb{P}\left(\frac{1}{n} X^{1/n} > x\right) \\ &= \mathbb{P}(X^{1/n} \leq nx) \\ &= 1 - \left(1 - \frac{1}{n}\right)^{nx+1} \xrightarrow{n\to\infty} (1 - e^{-x})\mathbb{1}_{[0,\infty)} \end{aligned}$$

Por tanto $F_n \Rightarrow F$ donde F es la exponencia de parámetro 1 y si X es la variable aleatoria con esta distribución, se tiene que X_n converge a X en distribución.

Teorema 3.8 (terminar): *Sean X_n, X variables aleatorias tales que $X_n \Rightarrow X$, entonces existe un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ con variables aleatorias Y_n, Y tales que,* *(a)* $X_n \stackrel{d}{=} Y_n$,

(b) $X \stackrel{d}{=} Y$,

(c) $Y_n \to Y$ *casi seguramente.*

Por otro lado, si $X_n \to X$ *en probabilidad, entonces,* $X_n \Rightarrow X$.

Dem: Consideremos el espacio de probabilidad $([0,1], \mathcal{B}([0,1]), \lambda)$. Definimos,

$$F_n^{-1}(x) := \sup\{y \in \mathbb{R} : F_n(y) < x\}$$
$$F^{-1}(x) := \sup\{y \in \mathbb{R} : F(y) < x\}$$

Sean,

$$a_x := \sup\{y \in \mathbb{R} : F(y) < x\}$$
$$b_x := \inf\{y \in \mathbb{R} : F(y) > x\}$$

Siempre se tiene que $a_x \leq b_x$. Sea,

$$\Omega_0 := \{x \in [0,1] : a_x < b_x\}$$

Observemos que Ω_0 es, a lo más, numerable, dado que si $x_1, , x_2 \in \Omega_0$, entonces (a_{x_1}, b_{x_1}) es disjunto de (a_{x_2}, b_{x_2}) **revisar** y, para cada uno de estos intervalos se puede escoger $q_x \in \mathbb{Q} \cap (a_x, b_x)$ lo que produce una inyección entre Ω_0 y $\mathbb{Q}$, en otras palabras,

$$\lambda(\Omega_0) = 0$$

Demostremos que $F_n^{-1}(x) \longrightarrow F^{-1}(x)$ para todo $x \in [0,1] \setminus \Omega_0$. Si $x \in [0,1] \setminus \Omega_0$ entonces, si $F^{-1}(x) < y$ implica que $x < F(y)$ y si $F^{-1}(x) > y$ implica que $x > F(y)$.

Por demostrar

- $\liminf_n F_n^{-1}(x) \geq F^{-1}(x)$
- $\limsup_n F_n^{-1}(x) \leq F^{-1}(x)$

Teorema 3.9: *Sea* $\{X_n\}_{n\in\mathbb{N}}$ *variables aleatorias. Entonces* $X_n \Rightarrow X$ *si y sólo si,*

$$\lim_{n\to\infty} \mathbb{E}(g(X_n)) = \mathbb{E}(g(X))$$

para toda función $g \in C_b(\mathbb{R})$ *(funciones continuas acotadas).*

Dem: Supongamos que $X_n \Rightarrow X$, entonces por el teorema anterior existe un espacio de probabilidad $(\Omega, \mathcal{B}([0,1]), \mathbb{P})$ donde existen Y_n, Y variables aleatorias tales que,

$$Y_n \stackrel{d}{=} X_n, \;\; Y \stackrel{d}{=} X, \;\; Y_n \xrightarrow{\text{c.s}} Y$$

Sea g una función continua, entonces $g(Y_n) \xrightarrow{\text{c.s}} g(Y)$. Observemos que,

$$\mathbb{E}(|g(Y_n)|) \leq C$$

donde C es una constante y luego $C \in L^1(\Omega)$, por lo que podemos aplicar el teorema de la convergencia dominada/acotada, por lo que se cumple que,

$$\lim_{n\to\infty} \mathbb{E}(g(Y_n)) = \mathbb{E}(g(Y))$$

Por otro lado,

$$\mathbb{E}(g(Y_n)) = \int_{\mathbb{R}} g(x)\mu_n(x)$$

donde μ_n es la distribución de Y_n y como X_n es igual a Y_n en distribución, se tiene que X_n también tiene distribución μ_n y por tanto se concluye que,

$$\lim_{n\to\infty} \mathbb{E}(g(X_n)) = \mathbb{E}(g(X))$$

Como queriamos probar.

Supongamos ahora que para todo $g \in C_b(\mathbb{R})$ se tiene que,

$$\lim_{n\to\infty} \mathbb{E}(g(X_n)) = \mathbb{E}(g(X))$$

Queremos demostrar que si F_n, F son las distribuciones acumuladas de X_n, X respectivamente, entonces $F_n(x) \to F(x)$ puntualmente donde F es continua en x.

Consideremos la siguiente función,

$$(g_\varepsilon(x))(y) := \begin{cases} 1, & y \leq x \\ \text{lineal} & x < y < x+\varepsilon \\ 0, & y > x+\varepsilon \end{cases}$$

donde $x \in \mathbb{R}$, de forma que $g_\varepsilon(x)$ sea continua y acotada. Luego se cumple la siguiente desigualdad:

$$F_n(x) = \mathbb{P}(X_n \leq x) = \mathbb{E}(\mathbb{1}_{(-\infty,x]}(X_n)) \leq \mathbb{E}(g(X_n))$$

Aplicando el límite superior obtenemos,

$$\limsup_{n\to\infty} F_n(x) \leq \limsup_{n\to\infty} \mathbb{E}(g(X_n)) = \mathbb{E}(g(X)) \leq F(x+\varepsilon)$$

(verificar última desigualdad). Tomando $\varepsilon \to 0$ se concluye que,

$$\limsup_{n\to\infty} F_n(x) \leq F(x)$$

Este truco se puede hacer para la siguiente función **falta**.

Por lo que se concluye que,

$$\liminf_{n\to\infty} F_n(x) \geq F(x)$$

Por tanto, para x donde F es continua se cumple que,

$$\lim_{n\to\infty} F_n(x) = F(x)$$

Como queriamos probar. ■

Teorema 3.10 (Teorema del Mapa Continuo): *Sea g una función medible y sea $D(g) = \{x \in \mathbb{R} : \ g \text{ es discontinua en } x\}$. Si $X_n \Rightarrow X$ y $\mathbb{P}(X \in D(g)) = 0$ entonces $g(X_n) \Rightarrow g(X)$. Si además, g es acotada se tiene que,*

$$\lim_{n\to\infty} \mathbb{E}(g(X_n)) = \mathbb{E}(g(X))$$

Dem ¿por que la condición Discontinua?: Antes que nada debemos probar que $D(g)$ es un boreleano. **probar**.

Como $X_n \Rightarrow X$, entonces existe un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ con variables aleatorias Y_n, Y tales que,

$$Y_n \overset{d}{=} X_n, \ \ Y \overset{d}{=} X, \ \ Y_n \xrightarrow{\text{c.s}} Y$$

Sea $f : \mathbb{R} \to \mathbb{R}$ una función continua, entonces se cumple que $D(f \circ g) \subseteq D(g)$ y por tanto se cumple que,

$$f(g(Y_n)) \longrightarrow f(g(Y))$$

casi seguramente, además, si f es acotada, entonces podemos aplicar el teorema de convergencia acotada, de forma que,

$$\lim_{n\to\infty} \mathbb{E}(f(g(Y_n))) = \mathbb{E}(f(g(Y)))$$

Y usando que X_n, X tiene igual distribución que Y_n, Y respectivamente, se cumple que,

$$\lim_{n\to\infty} \mathbb{E}(f(g(X_n))) = \mathbb{E}(f(g(X)))$$

para todo f continua acotada. Luego por el teorema anterior se cumple que $g(X_n) \Rightarrow g(X)$.

Ahora, para la segunda conclusión observemos que si $\mathbb{P}(Y \in D(g)) = 0$ entonces $g(Y_n) \to g(Y)$ casi seguramente y como es acotada se puede aplicar el teorema de la convegencia acotada y usando que X_n, X tiene igual distribución a Y_n, Y respectivamente se concluye que,

$$\lim_{n\to\infty} \mathbb{E}(g(X_n)) = \mathbb{E}(g(X))$$

Como queriamos probar. ■

Teorema 3.11: *Sean X_n, X variables aleatorias. Entonces las siguientes afirmaciones son equivalentes:*

i) $X_n \Rightarrow X$.

ii) $\liminf \mathbb{P}(X_n \in A) \geq \mathbb{P}(X \in A)$ *para todo* A *abierto real.*

iii) $\limsup \mathbb{P}(X_n \in C) \leq \mathbb{P}(X \in C)$ *para todo* C *cerrado real.*

iv) $\lim \mathbb{P}(X_n \in B) = \mathbb{P}(X \in B)$ *para todo* B *Boreleando tal que* $\mathbb{P}(X \in \partial B) = 0$.

Dem:

- **i) implica ii):** Supongamos que $X_n \Rightarrow X$. Entonces existe un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ con variables aleatorias Y_n, Y tales que X_n, Y_n tienen igual distribución, lo mismo para X, Y y que Y_n converge a Y casi seguramente. Observemos que,

$$\mathbb{1}_{\{Y \in A\}}(\omega) \leq \liminf_{n\to\infty} \mathbb{1}_{\{Y_n \in A\}}(\omega)$$

Donde A es abierto. Por el lema de Fatou se tiene que,

$$\begin{aligned}\mathbb{E}(\mathbb{1}_{Y \in A}) &\leq \mathbb{E}\left(\liminf_{n\to\infty} \mathbb{1}_{\{Y_n \in A\}}\right)\\ &\leq \liminf_{n\to\infty} \mathbb{E}(\mathbb{1}_{\{Y_n \in A\}})\end{aligned}$$

Es decir,

$$\mathbb{P}(Y \in A) \leq \liminf_{n\to\infty} \mathbb{P}(Y_n \in A)$$

Finalmente aplicando que X_n, X tiene igual distribución a Y_n, Y respectivamente se concluye ii).

- **ii) implica iii):** Sea C un cerrado, entonces C^c es abierto, de forma que se cumple que,

$$\mathbb{P}(X \in C^c) \leq \liminf_{n\to\infty} \mathbb{P}(X_n \in C^c)$$

Si además $\mathbb{P}(X_n \in C^c) = 1 - \mathbb{P}(X_n \in C)$ (lo mismo para X), se tiene que,

$$\begin{aligned}\limsup_{n\to\infty} \mathbb{P}(X_n \in C) &= \limsup_{n\to\infty}(1 - \mathbb{P}(X_n \in C^c))\\ &= 1 - \liminf_{n\to\infty} \mathbb{P}(X_n \in C^c)\\ &\leq 1 - \mathbb{P}(X \in C^c)\\ &= \mathbb{P}(X \in C)\end{aligned}$$

Conluyendo iii). Observemos que podemos concluir que ii) es equivalente a iii) usando un argumento similar.

- **ii) y iii) implican iv):** Sea B un boreleano con borde despreciable. Entonces se cumple que,

$$\mathbb{P}(X \in B) = \mathbb{P}(X \in \overline{B}) = \mathbb{P}(X \in B^o)$$

Luego aplicando ii) y iii) se concluye que,

$$\begin{aligned}\mathbb{P}(X \in B) &= \mathbb{P}(X \in B^o)\\ &\leq \liminf_{n\to\infty} \mathbb{P}(X_n \in B^o)\\ &\leq \liminf_{n\to\infty} \mathbb{P}(X_n \in B)\end{aligned}$$

Y que,

$$\begin{aligned}\mathbb{P}(X \in B) &= \mathbb{P}(X \in \overline{B}) \\ &\geq \limsup_{n\to\infty} \mathbb{P}(X_n \in \overline{B}) \\ &\leq \liminf_{n\to\infty} \mathbb{P}(X_n \in B)\end{aligned}$$

Probando que,

$$\lim_{n\to\infty} \mathbb{P}(X_n \in B) = \mathbb{P}(X \in B)$$

- **iv) implica i):** Sean $F_n(x) = \mathbb{P}(X_n \in (-\infty, x])$ y $F(x) = \mathbb{P}(X \in (-\infty, x])$ las funciones distribución de X_n, X respectivamente. Sea x un punto de continuidad de F. Sabemos que,

  $$F(x^-) = F(x^+) = F(x)$$

  por la continudad de F, pero esto es equivalente a decir que,

  $$\mathbb{P}(X \in (-\infty, x)) = \mathbb{P}(X \in (-\infty, x])$$

  Esto implica que $\mathbb{P}(X \in \{x\}) = 0$, es decir, el borde $(-\infty, x]$ es despreciable y por tanto se cumple que,

  $$\lim_{n\to\infty} F_n(x) = F(x)$$

  por la hipótesis de iv). Por lo tanto $X_n \Rightarrow X$.

Probando el teorema. ■

3.2. Teorema de Selección de Helly

Teorema 3.12 (Helly): *Sea F_n una sucesión de fnciones de distribución. Entonces existe una función F creciente (no necesariamente estrica), continua por la derecha y $0 \leq F(x) \leq 1$ para todo $x \in \mathbb{R}$, tal que existe una subsucesión $F_{n(k)}$ tal que,*

$$\lim_{k\to\infty} F_{n(k)}(x) = F(x)$$

para todo $x \in \mathbb{R}$ donde F es continua.

Nota 3.13: La convergencia descrita en el teorema de Holly se le denomina **convergencia vaga** y se denota por $\overset{v}{\Rightarrow}$.

Teorema 3.14: *Supongamos que para todo $\varepsilon > 0$ existe $M(\varepsilon)$ tal que,*

$$\liminf_{n\to\infty} (F_n(-M(\varepsilon)) - F_n(M(\varepsilon)) + 1) \leq \varepsilon$$

Entonces la F asegurada por el teorema de Helly es una función de distribución acumulada.

Nota 3.15: La condición de que para todo existe $\varepsilon > 0$ existe un $M(\varepsilon)$ tal que,

$$\liminf_{n\to\infty}(F_n(-M(\varepsilon)) - F_n(M(\varepsilon)) + 1) \leq \varepsilon$$

Se le llama **tensión**. Entonces la secuencia F_n le decimos que es tensa.

Observación 3.16: Sea F_n funciones distribuciones de variables aleatorias X_n. Supongamos que la $\{F_n\}_{n\in\mathbb{N}}$ es tensa, entonces además se cumple que para todo $\varepsilon > 0$ existe $M(\varepsilon)$ tal que,

$$\mathbb{P}(X_n \in (-\infty, -M(\varepsilon)]) + 1 - \mathbb{P}(X_n \in (-\infty, -M(\varepsilon)]) \leq \varepsilon$$

O dicho de otra forma,

$$\mathbb{P}(X_n \notin (-M(\varepsilon), M(\varepsilon)) \leq \varepsilon$$

Dem Idea (Teorema de Helly): Sea $\{q_i\}_{i\in\mathbb{N}}$ una enumeración de los racionales de forma que $F_n(q_1) \in [0,1]$ para todo$n \in \mathbb{N}$. Existe una subsucesión $n_1(k)$ tal que,

$$\lim_{k\to\infty} F_{n_1(k)}(q_1) = G(q_1)$$

Ahora también se cumple que $F_{n_1(k)}(q_2) \in [0,1]$, luego existe una subsucesión $n_2(k)$ tal que,

$$\lim_{k\to\infty} F_{n_1(k)}(q_2) = G(q_2)$$

De forma que podemos construir una subsucesión diagonal $F_{n_k(k)}$ de forma que,

$$\lim_{k\to\infty} F_{n_k(k)}(q_i) = G(q_i)$$

para todo $i \in \mathbb{N}$. Luego se define $F(x) := \inf\{G(q) : q > x\}$. Se puede demostrar que F es monóntona (no estrica), continua por la derecha y que $F_n(x)$ converge a $F(x)$ para todo $x \in \mathbb{R}$ donde F es continua. ■

Dem (Tensión): Supongamos que para todo $\varepsilon > 0$ existe $M(\varepsilon)$ tal que,

$$\liminf_{n\to\infty}(F_n(-M(\varepsilon)) + 1 - F_n(M(\varepsilon))) \leq \varepsilon$$

Podemos escoger $M(\varepsilon)$ conociendo un punto de continuidad de F. Porlo tanto,

$$F(-M(\varepsilon)) + 1 - F(M(\varepsilon)) < \varepsilon$$

Pero como $0 \leq F(-M(\varepsilon)) \leq 1$ y $0 \leq F(M(\varepsilon)) \leq 1$, se tiene que,

$$F(-M(\varepsilon)) < \varepsilon, \quad F(M(\varepsilon)) > 1 - \varepsilon$$

Por lo tanto,

$$\lim_{x\to-\infty} F(x) = 0$$
$$\lim_{x\to\infty} F(x) = 1$$

Como queriamos demostrar. ■

3.3. Función Característica

Definición 3.17: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sea $X : \Omega \to \mathbb{R}$ una variable aleatoria. Definimos la función característica de X por:*

$$\varphi_X(t) := \mathbb{E}(e^{itx})$$

donde $\varphi_X : \mathbb{R} \to \mathbb{C}$.

Observación 3.18: La función característica está bien definida, puesto que para todo $t \in \mathbb{R}$ se cumple que,

$$\int_{\mathbb{R}} |e^{itx}| d\mathbb{P}_X(x) < \infty$$

Además, también se puede escribir de la siguiente forma:

$$\varphi_X(t) = \mathbb{E}(\cos(tX)) + i\mathbb{E}(\operatorname{sen}(tX))$$

Nota 3.19: Podemos definir la función característica mediante una medida real, ya que esta define una función de distribución acumulada y por tanto, una variable aleatoria.

Ejemplo 3.20: Sea X una variable aleatoria con distribución de Bernoulli(1/2) que toma valores en $\{-1, 1\}$, entonces,

$$\begin{aligned}
\varphi_X(t) &= \mathbb{E}(e^{itx}) \\
&= \int_{\Omega} e^{itX(\omega)} d\mathbb{P}(\omega) \\
&= \int_{\{X=1, X=-1\}} e^{itX(\omega)} d\mathbb{P}(\omega) \\
&= e^{it}\mathbb{P}(X = 1) + e^{-it}\mathbb{P}(X = -1) \\
&= \frac{e^{it} + e^{-it}}{2} = \cos(t)
\end{aligned}$$

Ejemplo 3.21: Sea $X \sim N(0, \sigma^2)$. Entonces la distribución de X tiene una densidad y entonces,

$$\begin{aligned}
\varphi_X(t) &= \int_{\mathbb{R}} e^{itx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}} d\lambda \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{t^2\sigma^2}{2}} \int_{\mathbb{R}} e^{-\frac{(x-it\sigma^2)^2}{2\sigma^2}} d\lambda \\
&= e^{-\frac{t^2\sigma^2}{2}}
\end{aligned}$$

Proposición 3.22: *Se cumple la siguientes propiedades de la función característica:*

(a) $\varphi_X(0) = 1$.

(b) Para todo $t \in \mathbb{R}$ se tiene que,

$$|\varphi_X(t)| \leq 1$$

(c) Sean $a, b \in \mathbb{R}$, entonces,

$$\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$$

(d) Sean X, Y variables aleatorias independientes, entonces,

$$\varphi_{X+Y}(t) = \varphi_{X_1}(t)\varphi_{X_2}(t)$$

(e) Sean $t, h \in \mathbb{R}$, entonces,

$$|\varphi_X(t+h) - \varphi(t)| \leq \mathbb{E}(|e^{ihX} - 1|)$$

Dem:

(a) Por definición se tiene que,

$$\varphi_X(0) = \mathbb{E}(1) = 1$$

(b) Sea $t \in \mathbb{R}$ fijo, luego,

$$\begin{aligned} |\varphi_X(t)| &= |\mathbb{E}(e^{itX})| \\ &\leq \left| \int_{\mathbb{R}} e^{itx} d\mathbb{P}_X \right| \\ &\leq \int_{\mathbb{R}} |e^{itx}| d\mathbb{P}_X \\ &= \mathbb{P}_X(\mathbb{R}) = 1 \end{aligned}$$

Luego $|\varphi_X(t)| \leq 1$ para todo $t \in \mathbb{R}$.

(c) Sean $a, b \in \mathbb{R}$ cosntantes, luego por definición,

$$\begin{aligned} \varphi_{aX+b}(t) &= \mathbb{E}(e^{it(aX+b)}) \\ &= \mathbb{E}(e^{i(at)Xe^{itb}}) \\ &= e^{itb}\varphi_X(at) \end{aligned}$$

(d) Sean X, Y independientes, entonces e^{itX}, e^{itY} son independientes, de forma que,

$$\begin{aligned} \varphi_{X+Y}(t) &= \mathbb{E}(e^{itX}e^{itY}) \\ &= \mathbb{E}(e^{itX})\mathbb{E}(e^{itY}) \\ &= \varphi_X(t)\varphi_Y(t) \end{aligned}$$

(e) Sean $t, h \in \mathbb{R}$, entonces se cumple que,

$$\begin{aligned}
|\varphi_X(t+h) - \varphi(t)| &= |\mathbb{E}(e^{i(t+h)X}) - e^{itX})| \\
&= \left| \int_{\mathbb{R}} e^{itx}(e^{ihx} - 1) d\mathbb{P}_X \right| \\
&\leq \int_{\mathbb{R}} |e^{ihx}| d\mathbb{P}_X \\
&= \mathbb{E}(|e^{ihX}|)
\end{aligned}$$

Como se quería probar.

Probando la proposición. ■

Corolario 3.23: *Sea X una variable aleatoria. Entonces la función característica φ_X es absolutamente continua.*

Que la función φ sea absolutamente continua es equivalente a decir que, existe una función f que es la derivada de φ casi seguramente tal que,

$$\varphi(t) = \varphi(a) + \int_a^t f(x) d\lambda(x)$$

a tal f lo denotamos por $d\varphi/dt$.

Teorema 3.24 (De inversión): *Sea μ una medida de probabilidad sobre $\mathbb{R}$ y sea,*

$$\varphi(t) = \int_{\mathbb{R}} e^{itx} d\mu(x)$$

definido para todo $t \in \mathbb{R}$. Sean $a < b$, entonces,

$$\mu(a,b) + \frac{1}{2}\mu(\{a,b\}) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt$$

Dem: Para T fijo consideramos,

$$I_T := \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt$$

Observamos es una integral impropia ya que no está bien definido en $t = 0$, sin embargo veremos más adelante que no hay problema con esto ya que veremos que,

$$\frac{e^{-ita} - e^{-itb}}{it}$$

es controlable, en el sentido que es acotado en una vecindad pinchada de $t = 0$, con única cota para todo $t \in \mathbb{R}$. Observemos que,

$$I_T := \int_{-T}^{T} \int_{\mathbb{R}} \frac{e^{-ita} - e^{-itb}}{it} e^{itx} d\mu(x) dt$$

Como estamos trabajando en una medida de probabilidad y en $\mathbb{R}$, entonces podriamos ocupar Fubini, pero nos falta comprobar que la función de adentro sea L^1, y en efecto, notemos que,

$$\begin{aligned}\left|\frac{e^{-ita}-e^{-itb}}{it}e^{itx}\right| &= \left|\frac{e^{-ita}-e^{-itb}}{it}\right| \\ &= \left|\int_a^b e^{itx}dx\right| \leq b-a\end{aligned}$$

Por lo que podemos ignorar $t=0$ y por Fubini se tiene que,

$$\begin{aligned}I_T &= \int_{\mathbb{R}}\left(\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}e^{itx}dt\right)d\mu(x) \\ &= \int_{\mathbb{R}}\left(\int_{-T}^{T}\frac{e^{it(x-a)}}{it}dt-\int_{-T}^{T}\frac{e^{it(x-b)}}{it}dt\right)d\mu(x)\end{aligned}$$

Definimos,

$$R(\theta,T) := \int_{-T}^{T}\frac{e^{it\theta}}{it}dt$$

Entonces se cumple que,

$$R(\theta,T) = \int_{-T}^{T}\frac{\cos(t\theta)}{it}dt+\int_{-T}^{T}\frac{\operatorname{sen}(t\theta)}{t}dt$$

Vemos que $\cos(t\theta)/it$ es impar, de forma que,

$$R(\theta,T) = \int_{-T}^{T}\frac{\operatorname{sen}(t\theta)}{t}dt$$

Recordemos que,

$$S(T) := \int_0^T\frac{\operatorname{sen}(y)}{y}dy \overset{T\to\infty}{\longrightarrow} \frac{\pi}{2}$$

Además que,

$$\left|\frac{\operatorname{sen}(t\theta)}{t}\right| \leq \frac{|t\theta|}{|t|} = |\theta|$$

Vemos que si $\theta>0$ entonces,

$$\begin{aligned}R(\theta,T) &= 2\int_0^T\frac{\operatorname{sen}(t\theta)}{t}dt \\ &= 2\int_0^{T\theta}\frac{\operatorname{sen}(y)}{y}dt \\ &= 2S(T\theta)\end{aligned}$$

Por otro lado si $\theta < 0$ se puede concluir que $R(\theta, T) = -2S(|T\theta|)$, así para todo $\theta \in \mathbb{R}$ se cumple que,

$$R(\theta, T) = 2\text{sgn}(\theta)S(|\theta|T)$$

Tomando $T \to \infty$ se observa que,

$$\lim_{T\to\infty} R(\theta, T) = \text{sgn}(\theta)\pi$$

Ahora notemos lo siguiente,

$$\lim_{T\to\infty} R(x-a, T) - R(x-b, T) = \begin{cases} 2\pi, & a < x < b \\ \pi, & a = x \text{ o } x = b \\ 0, & x < a \text{ o } x > b \end{cases}$$

Por otro lado se tiene que,

$$|R(\theta, T)| \leq 2 \sup_u |S(u)| < \infty$$

Por lo tanto por convergencia dominada obtenemos que,

$$\begin{aligned}
\lim_{T\to\infty} \frac{1}{2\pi} I_T &= \frac{1}{2\pi} \lim_{T\to\infty} I_T \\
&= \frac{1}{2\pi} \lim_{T\to\infty} \int_{\mathbb{R}} R(x-a, T) - R(x-b, T) d\mu(x) \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \lim_{T\to\infty} R(x-a, T) - R(x-b, T) d\mu(x) \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} (2\pi \mathbb{1}_{(a,b)} + \pi \mathbb{1}_{\{a,b\}}) d\mu(x) \\
&= \mu(a, b) + \frac{1}{2}\mu(\{a, b\})
\end{aligned}$$

Probando el teorema. ■

Nota 3.25: El teorema de la inversión nos sirve para caracterisar las distribuciones, es decir, si X, Y son variables aleatorias donde sus funciones características coinciden, entonces X, Y son igual en distribución. Esto se debe a que se puede verificar que,

$$\mathbb{P}_X(a, b) + \frac{1}{2}(\mathbb{P}_X(\{a, b\})) \overset{\substack{a\to-\infty \\ b\to y^+}}{\longrightarrow} \mathbb{P}_X((-\infty, y])$$

Y si,

$$\mathbb{P}_X(a, b) + \frac{1}{2}(\mathbb{P}_X(\{a, b\})) = \mathbb{P}_Y(a, b) + \frac{1}{2}(\mathbb{P}_Y(\{a, b\}))$$

Entonces finalmente se concluye que $F_X = F_Y$ que es equivalente a decir que ambas variables aleatorias tienen la misma distribución.

Corolario 3.26: *Si $\varphi_X(t) \in \mathbb{R}$, entonces $-X, X$ tienen la misma distribución.*

Dem: Por definición,

$$\begin{aligned}\varphi_{-X}(t) &= \mathbb{E}(e^{-itX}) \\ &= \mathbb{E}(\overline{e^{itX}}) \\ &= \overline{\mathbb{E}(e^{itX})} \\ &= \overline{\varphi_X(t)} = \varphi_X(t)\end{aligned}$$

De esta forma $X, -X$ tienen la misma distribución. ■

Corolario 3.27: *Sean X_1, X_2 variables aleatorias independientes. Si $X_1 \sim N(0, \sigma_1^2), X_2 \sim N(0, \sigma_2^2)$, entonces $X_1 + X_2 \sim N(0, \sigma_1^2 + \sigma_2^2)$.*

Dem: Se cumple que,

$$\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\varphi_{X_2}(t) = e^{-\frac{\sigma_1^2 t^2}{2}} e^{-\frac{\sigma_1^2 2t^2}{2}} = e^{-\frac{(\sigma_1^2+\sigma_2^2)t^2}{2}}$$

Por tanto, $X_1 + X_2$ tiene distribución $N(0, \sigma_1^2 + \sigma_2^2)$. ■

Toerema 3.28: *Sea μ una medida de probabilidad sobre $\mathbb{R}$ y sea $\varphi(t)$ su caracterítica. Supongamos que,*

$$\int_{\mathbb{R}} |\varphi(t)| dt < \infty$$

Entonces μ es una medida continua con densidad,

$$f(y) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ity} \varphi(t) dt$$

Dem: Por el teorema de inversión, se cumple que,

$$\mu(a,b) + \frac{1}{2}\mu(\{a,b\}) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt$$

Como μ es no negativa, se tiene que es invariante sobre el valor absoluto, además de que el valor absoluto escontinua, por lo que,

$$\begin{aligned}\mu(a,b) + \frac{1}{2}\mu(\{a,b\}) &= \lim_{T\to\infty} \frac{1}{2\pi} \left| \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt \right| \\ &\le \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \left| \frac{e^{-ita} - e^{itb}}{it} \right| dt \\ &\le \frac{(b-a)}{2\pi} \lim_{-T}^{T} |\varphi(t)| dt \\ &= \frac{b-a}{2\pi} \int_{\mathbb{R}} |\varphi(t)| dt < \infty\end{aligned}$$

Esto implica que,

$$\lim_{a\to b} \mu(a,b) + \frac{1}{2}\mu(\{a,b\}) = 0$$

Luego la medida de los átomos es cero, de forma que μ es una medida continua. Ahora, del cálculo anterior se observa que por convergencia dominada se tiene obtiene que,

$$\begin{aligned}
\mu(a,b) &= \frac{1}{2\pi}\int_{\mathbb{R}} \frac{e^{-ita} - e^{-tb}}{it}\varphi(t)dt \\
&= \frac{1}{2\pi}\int_{\mathbb{R}} \left(\int_a^b e^{-itx}dx\right)\varphi(t)dt \\
&= \int_a^b \left(\frac{1}{2\pi}\int_{\mathbb{R}} e^{-ity}\varphi(t)dt\right)dx
\end{aligned}$$

Por tanto, la medida μ tiene densidad

$$\frac{1}{2\pi}\int_{\mathbb{R}} e^{-ity}\varphi(t)dt$$

Es más, la densidad de μ es continua. ■

Teorema 3.29 (Teorema de continuidad): *Sean μ, μ_n medidas de probabilidad sobre $\mathbb{R}$ con funciones características φ, φ_n. Luego,*

i) Si $\mu_n \Rightarrow \mu$, entonces $\varphi_n(t) \to \varphi(t)$ para todo $t \in \mathbb{R}$.

ii) Si $\varphi_n(t) \to \varphi(t)$ para todo $t \in \mathbb{R}$ y φ es continua en 0, entonces $\{\mu_n\}_{n\in\mathbb{N}}$ es tensa y $\mu_n \Rightarrow \mu$.

Dem:

i) Supongamos que $\mu_n \Rightarrow \mu$, debemos probar que,

$$\int_{\mathbb{R}} e^{itx}d\mu_n(x) \overset{n\to\infty}{\longrightarrow} \int_{\mathbb{R}} e^{itx}d\mu(x)$$

Observamos que e^{itx} es acotada como sabemos, luego por el teorema de la convergencia dominada se cumple que,

$$\lim_{n\to\infty}\int_{\mathbb{R}} e^{itx}d\mu_n(x) = \int_{\Omega} e^{itX}d\mathbb{P}$$

ii) Consideremos la siguiente afirmación.

Afirmación: *si ν es una medida de probabilidad con función característica ψ, entonces,*

$$\nu\left(\left\{x : |x| > \frac{2}{u}\right\}\right) \le \frac{1}{u}\int_{-u}^{u}(1-\psi(t))dt$$

para todo $u > 0$.

Obseremos que si tomamos $\varepsilon > 0$, entonces como $\varphi(t) \to \varphi(0) = 1$ para $t \to 0$, entonces existe un $u_0 > 0$ tal que,

$$\frac{1}{u}\int_{-u}^{u}(1-\varphi(t))dt < \varepsilon$$

para todo $0 < u < u_0$. Definimos,

$$g_\nu(u) := \frac{1}{u}\int_{-u}^{u}(1-\psi(t))dt$$

Sea $0 < u < u_0$. Como $\varphi_n \to \varphi$, entonces existe $n_0 \geq 1$ tal que,

$$|g_{\mu_n}(u) - g_\mu(u)| < \varepsilon$$

para todo $n \geq n_0$.

Combinando estas dos cosas obtenemos que, si $n \geq n_0$, entonces,

$$\begin{aligned}\mu_n\left(\left\{x : |x| > \frac{2}{u}\right\}\right) &\leq g_{\mu_n}(u) \\ &\leq g_\mu(u) + |g_\mu(u) - g_\mu(u)| \\ &< 2\varepsilon\end{aligned}$$

Por lo tanto $\{\mu_n\}_{n\in\mathbb{N}}$ es tensa.

Sea $\{\mu_{n'}\}_{n'}$ una subsucesión de la sucesión de medidas tensa, luego es tensa, de esta forma existe una subsucesión $\{\mu_{n''}\}_{n''}$ y una medida de probabilidad $\widetilde{\mu}$ tal que $\mu_{n''} \Rightarrow \widetilde{\mu}$. Por el punto i) se tiene que,

$$\varphi_{n''}(t) \overset{n\to\infty}{\longrightarrow} \widetilde{\varphi}(t) = \int_{\mathbb{R}} e^{itx}\widetilde{\mu}(dx)$$

Por lo tanto $\widetilde{\varphi} = \varphi$ y, por lo tanto, $\widetilde{\mu} = \mu$ y por tanto $\mu_{n''} \Rightarrow \mu$. Por el argumento de subsucesiones se tiene que $\mu_n \Rightarrow \mu$.

Demostrando el teorema. Nos falta probar la afirmación.

Dem Afirmación: Debemos probar que,

$$\nu(\{x : |x| > \frac{2}{u}\}) \leq \frac{1}{u}\int_{-u}^{u}(1-\psi(t))dt$$

Notemos lo siguiente,

$$\int_{-u}^{u}(1-e^{itx})dt = 2u - \int_{-u}^{u}(\cos(tx) + i\operatorname{sen}(tx))dx$$

Pero el seno no participa al ser impar, entonces,

$$\begin{aligned}\int_{-u}^{u}(1-e^{itx})dt &= 2u - 2\int_0^u \cos(tx)dt \\ &= 2u - 2\frac{\operatorname{sen}(ux)}{x}\end{aligned}$$

Luego,

$$\begin{aligned}\frac{1}{u}\int_{-u}^{u}(1-\psi(t))dt &= \frac{1}{u}\int_{-u}^{u}\int_{\mathbb{R}}(1-e^{itx})dtd\nu(x) \\ &= \frac{1}{u}\int_{\mathbb{R}}\int_{-u}^{u}(1-e^{itx})dtd\nu(x) \\ &= \frac{2}{u}\int_{\mathbb{R}}(u-\frac{\operatorname{sen}(ux)}{x})d\nu(x) \\ &= 2\int_{\mathbb{R}}(1-\frac{\operatorname{sen}(ux)}{ux})d\mu(x) \\ &\geq 2\int_{|x|\geq 2/u}(1-\frac{1}{|ux|})d\nu(x)\end{aligned}$$

aplicando Fubini y que $|\operatorname{sen}(ux)|\leq 1, |\operatorname{sen}(ux)|\leq |ux|$. Si,

$$1-\frac{1}{|ux|}\geq\frac{1}{2}$$

Finalmente se obtiene que,

$$2\int_{|x|\geq 2/u}(1-\frac{1}{|ux|})d\nu(x)\geq\int_{|x|\geq 2/u}1d\nu(x)=\nu(\{x:|x|\geq 2/u\})$$

■

Teorema 3.30: *Sea μ una medida de probabilidad sobre $\mathbb{R}$ con función característica φ. Sea $n\geq 1$, entonces si,*

$$\int_{\mathbb{R}}|x|^n d\mu(x)<\infty$$

Entonces la función característica tiene n-ésima derivada continua de la forma,

$$\frac{d^n}{dt^n}\varphi(t)=\int_{\mathbb{R}}(ix)^n e^{itx}d\mu(x)$$

En particular, si X es una variable con distribución μ, entonces,

$$\frac{d^n}{dt^n}\varphi(0)=i^n\mathbb{E}(X^n)$$

Dem: Estudiemos la derivada e^{itx} en función de t.

$$|(e^{itx})'| = |ixe^{itx}| = |x| \in L^1(\mu)$$

Luego por el teorema de la convergencia dominada nos dice que podemos meter la derivada adentro de la integral, de esta forma,

$$\begin{aligned}\frac{d}{dt}\varphi(t) &= \frac{d}{dt}\int e^{itx}d\mu(x) \\ &= \int_{\mathbb{R}} \frac{d}{dt}e^{itx}d\mu(x) \\ &= \int_{\mathbb{R}} ixe^{itx}dx\end{aligned}$$

Probando para el caso base, el caso general sigue de forma análoga. Para concluir la última igualdad basta tomar $t = 0$, luego,

$$\frac{d^n}{dt^n}\varphi(0) = i^n\mathbb{E}(X^n)$$

■

Observación 3.31: Determinemos las dos primeras derivadas, supongamos que $\mathbb{E}(|X|) < \infty$ y $\mathbb{E}(|X|^2) < \infty$. Entonces por el teorema anterior,

$$\varphi'(0) = i\mathbb{E}(X), \;\; \varphi''(0) = i^2\mathbb{E}(X^2) = -\mathbb{E}(X^2)$$

Veamos que pasa si aplicamos Taylor de centrado en 0 a una funcion cacracterística . Para t pequeño se tiene que,

$$\begin{aligned}\varphi(t) &\approx \varphi(0) + \varphi'(0)t + \frac{\varphi''(0)}{2}t^2 \\ &= 1 + it\mathbb{E}(X) - \frac{t^2}{2}\mathbb{E}(X^2)\end{aligned}$$

Supongamos que $\mathbb{E}(X) = 0, \mathbb{E}(X^2) = \sigma^2 \in (0, \infty)$, entonces,

$$\varphi(t) \approx 1 - \frac{\sigma^2 t^2}{2}$$

Consideremos $\{X_n\}_{n\in\mathbb{N}}$ variables aleatorias i.i.d con distribución igual a la de X. Se tiene que,

$$\begin{aligned}\mathbb{E}\left(\exp\left(it\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right)\right) &= \mathbb{E}\left(\prod_{i=1}^{n}\exp\left(it\frac{X_i}{\sqrt{n}}\right)\right) \\ &= \prod_{i=1}^{n}\mathbb{E}(e^{itX_i/\sqrt{n}}) \\ &= \left(\varphi\left(\frac{t}{\sqrt{n}}\right)\right)^n\end{aligned}$$

Usando la aproximación de Taylor se tiene que,

$$\left(\varphi\left(t/\sqrt{n}\right)\right)^n \approx \left(1-\frac{\sigma^2 t^2}{2n}\right)^n \overset{n\to\infty}{\longrightarrow} e^{-\frac{\sigma^2 t^2}{2}}$$

Necesitamos justificar algunos pasos, cosa que haremos más adelante, pero si lo hacemos obtenemos que,

$$\frac{X_1+\cdots+X_n}{\sqrt{n}} \Rightarrow N(0,\sigma^2)$$

Estudiemos la aproximación de Taylor. Se cumple que,

$$\left|e^{iy}-1-iy+\frac{y^2}{2}\right| \leq C|y|^3$$

para $|y|$ pequeño con C una constante. Si $|y|$ no es tan pequeño se tiene que,

$$\left|e^{iy}-1-iy+\frac{y^2}{2}\right| \leq C|y|^3 \leq C|y^2|$$

Por lo tanto,

$$\left|e^{iy}-1-iy+\frac{y^2}{2}\right| \leq \text{mín}\{|y|^3, |y|2\}$$

Luego considerando la esperanza se obtiene,

$$\left|\mathbb{E}\left(e^{iX}-1-iX+\frac{X^2}{2}\right)\right| \leq C|t|^2\mathbb{E}(|X|^2)$$

Si $\mathbb{E}(|X|^2)<\infty$, entonces,

$$\varphi(t) = 1+it\mathbb{E}(X)-\frac{t^2}{2}\mathbb{E}(X^2)+O(t^2)$$

Esto lo podemos generalizar con el siguiente resultado.

Lema 3.32:

$$\left|e^{ix}-\sum_{m=0}^{n}\frac{(ix)^m}{m!}\right| \leq \text{mín}\left\{\frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!}\right\}$$

Demostración durret.

El caso $n=2$ va a ser clave en la demostración del teorema del límite central.

Teorema 3.33: *Si,*

$$\limsup_{h\downarrow 0} \frac{\varphi(h)+\varphi(-h)-2\varphi(0)}{h^2} > -\infty$$

Entonces,

$$\mathbb{E}(|X|^2) < \infty$$

Dem: Observemos que,

$$\frac{e^{ihx}+e^{-ihx}-2}{h^2} = -2\frac{1-\cos(hx)}{h^2} \leq 0$$

Además,

$$2\frac{1-\cos(hx)}{h^2} \xrightarrow{h\downarrow 0} x^2$$

Luego,

$$\begin{aligned}
\mathbb{E}(|X|^2) &= \int_{\mathbb{R}} |x^2| d\mu(x) \\
&\leq \liminf_{h\downarrow\infty} \int_{\mathbb{R}} 2\frac{1-\cos(hx)}{h^2} d\mu(x) \\
&= \liminf_{h\downarrow 0} - \limsup_{h\downarrow 0} \frac{\varphi(h)+\varphi(-h)-2\varphi(0)}{h^2} < \infty
\end{aligned}$$

Probando el teorema. ■

3.4. Teorema Central del Límite

Teorema 3.34 (Teorema Central del Límite): *Sea $\{X_n\}_{n\in\mathbb{N}}$ variables aleatorias independientes e identicamente distribuidas tal que $\mathbb{E}(X_1)=\mu$ y $Var(X_1)=\sigma^2 \in (0,\infty)$. Entonces,*

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \Rightarrow \mathcal{X}$$

Donde $\mathcal{X}$ es una variable aleatoria con distribución normal estándar.

La demostración del teorema del central del límite es sencillo de demostrar, basta usar la función característica. Por otro lado, sin pérdida de generalidad podemos suponer que $\mathbb{E}(X_1)=0$ puesto que en caso contrario podemos trabajar con $\widetilde{X}_i = X_i - \mu$ y luego $\mathbb{E}(\widetilde{X}_i)=0$. Ahora, es necesario el siguiente lema.

Lema 3.35: *Sea $\{c_n\} \subseteq \mathbb{C}$ una sucesión de complejos tal que,*

$$c_n \xrightarrow{n\to\infty} c$$

Entonces,

$$\left(1 + \frac{c_n}{n}\right)^n \stackrel{n\to\infty}{\longrightarrow} e^c$$

Para demostrar el **lema 3.35** se necesitan los siguientes dos lemas.

Lema 3.36: *Sean* $\{z_1, \dots, z_n, w_1, \dots, w_n\} \subseteq \mathbb{C}$ *una colección finitos de complejos tales que,* $|z_i|, |w_i| \leq \theta$ *para todo* $i = 1, \dots, n$*. Entonces se cumple que,*

$$\left|\prod_{i=1}^{n} z_i - \prod_{i=1}^{n} w_i\right| \leq \theta^{n-1} \sum_{i=1} n|z_i - w_i|$$

Dem: Probaremos por inducción. Si $n = 2$, entonces,

$$\begin{aligned} |z_1 z_2 - w_1 w_2| &= |z_1 z_2 - z_1 w_2 + z_1 w_2 - w_1 w_2| \\ &\leq |z_1||z_2 - w_2| + |w_2||z_1 - w_2| \\ &\leq \theta(|z_1 - w_1| + |z_2 - w_2|) \end{aligned}$$

Supongamos que se cumple para $n - 1$, entonces,

$$\begin{aligned} \left|\prod_{i=1}^{n} z_i - \prod_{i=1}^{n} w_i\right| &= \left|\prod_{i=1}^{n} z_i - \prod_{i=1}^{n-1} z_i w_n + \prod_{i=1}^{n-1} z_i w_n \prod_{i=1}^{n} w_i\right| \\ &\leq \left|\prod_{i=1}^{n-1} z_i\right| |z_n - w_n| + |w_n| \left|\prod_{i=1}^{n-1} z_i - \prod_{i=1}^{n-1} w_i\right| \\ &\leq \theta^{n-1} \sum_{i=1}^{n} |z_i - w_o| \end{aligned}$$

Probando el lema. ■

Lema 3.37: *Sea* $z \in \mathbb{C}$ *tal que* $|z| < 1$*, entonces,*

$$|e^z - (1 + z)| \leq |z|^2$$

Dem: Para todo $z \in \mathbb{C}$ se cumple que,

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \sum_{n=2}^{\infty} \frac{z^n}{n!}$$

Entonces,

$$\begin{aligned}
|e^z - (1+z)| &= \left|\sum_{n=2}^{\infty} \frac{z^n}{n!}\right| \\
&\leq \sum_{n=2}^{\infty} \frac{|z|^n}{n!} \\
&= |z|^2 \left(\frac{1}{2} + \frac{|z|}{2\cdot 3\cdot 4} + \dots\right) \\
&\leq \frac{|z|^2}{2}\left(1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots\right) \\
&= |z|^2
\end{aligned}$$

■

Dem Lema 3.35: Sea $\gamma > |c|$, con respecto al **lema 3.36** consideramos,

$$\begin{aligned}
z_i &= \left(1 + \frac{c_n}{n}\right) \\
w_i &= e^{\frac{c_n}{n}}
\end{aligned}$$

para $i = 1, \dots, n$.

Afirmación: *Se cumple la siguiente desigualdad:*

$$\left|\left(1 + \frac{c_n}{n}\right)^n - e^{c_n}\right| \leq \left(e^{\frac{\gamma}{n}}\right)^{n-1} \sum_{i=1}^{n} \left|1 + \frac{c_n}{n} - e^{\frac{c_n}{n}}\right|$$

Dem: Observemos lo siguiente,

$$\begin{aligned}
|z_i| &\leq e^{\frac{\gamma}{n}} \\
|w_i| &\leq e^{\frac{\gamma}{n}}
\end{aligned}$$

La desigualdad con respecto a z_i basta considerar que $1 + \gamma/n \leq \exp(\gamma/n)$, entonces,

$$|z_i| \leq 1 + \frac{\gamma}{n} \leq e^{\frac{\gamma}{n}}$$

Y con respecto a w_i, hay que notar lo siguiente,

$$\begin{aligned}
|w_i| &= \left|e^{\frac{c_n}{n}}\right| \\
&= \left|\sum_{m=0}^{\infty} \frac{(c_n)^m}{m!}\right| \\
&\leq \sum_{m=0}^{\infty} \frac{|c_n|^m}{m!} \\
&\leq \sum_{m=0}^{\infty} \frac{\gamma^m}{m!} \\
&= e^{\frac{\gamma}{n}}
\end{aligned}$$

Ahora por el **lema 3.36** se tiene que,

$$\left|\prod_{i=1}^{n} z_i - \prod_{i=1}^{n} w_i\right| = \left|\left(1+\frac{c_n}{n}\right)^n - e^{c_n}\right| \leq \left(e^{\frac{\gamma}{n}}\right)^{n-1} \sum_{i=1}^{n} \left|1+\frac{c_n}{n} - e^{\frac{c_n}{n}}\right|$$

■

Si $c_n \to c$, entonces $c_n/n \to 0$, por lo que para n suficientemente grande, se tiene que $|c_n/n| < 1$ y por tanto, por el **lema 3.37** se tiene que,

$$\left|1+\frac{c_n}{n} - e^{\frac{c_n}{n}}\right| \leq \left|\frac{c_n}{n}\right|^2$$

Para n suficientemente grande, además, también para n suficientemente grande se cumple que $|c_n| < \gamma$, por lo que

$$\left|\left(1+\frac{c_n}{n}\right)^n - e^{c_n}\right| \leq e^{\frac{\gamma(n-1)}{n}} \frac{|c_n|^2}{n} \leq e^{\gamma} \frac{\gamma^2}{n} \overset{n\to\infty}{\longrightarrow} 0$$

Como queriamos probar. ■

Dem. Teorema 3.34: Hemos visto que se tiene que si X_i es una variable aleatoria, entonces,

$$\varphi_{X_i}(t) = 1 - \frac{\sigma^2 t^2}{2} + \mathcal{O}(t^2)$$

donde $\mathcal{O}(t^2) \to 0$ cuando $t \to 0$. Por lo tanto,

$$\begin{aligned}
\varphi_{S_n/\sigma\sqrt{n}} &= \mathbb{E}(e^{itS_n/\sigma\sqrt{n}}) \\
&= \prod_{i=1}^{n} \mathbb{E}(e^{itX_i/\sigma\sqrt{n}}) \\
&= \prod_{i=1}^{n} \varphi_{X_1}\left(\frac{t}{\sigma\sqrt{n}}\right)^n \\
&= \left(1 - \frac{t^2}{2n} + \mathcal{O}\left(\frac{t^2}{\sigma^2 n}\right)\right)^n \overset{n\to\infty}{\longrightarrow} e^{-\frac{t^2}{2}}
\end{aligned}$$

Para concluir el límite observemos que

$$\begin{aligned}
&-\frac{t^2}{2n} + \mathcal{O}\left(\frac{t^2}{\sigma^2 n}\right) \in \mathbb{C} \\
&-\frac{t^2}{2n} + \mathcal{O}\left(\frac{t^2}{\sigma^2 n}\right) \overset{n\to\infty}{\longrightarrow} -\frac{t^2}{2}
\end{aligned}$$

Aplicando el **lema 3.35** se concluye el límite mencionado. Y si además, la función característica de una distribución normal estandar es $e^{-t^2/2}$, esto demuestre que,

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \Rightarrow \mathcal{X}$$

Demostrando el teorema central del límite. ■

4. Esperanza Condicional

Sea X una variable aleaotria integrable, es decir, $\mathbb{E}(|X|) < \infty$, definida en un espacio de probabilidad $(\Omega, \mathcal{F}_0, \mathbb{P})$. Sea $\mathcal{F} \subseteq \mathcal{F}_0$ un σ-álgebra. Definimos la esperanza condicional de X dado $\mathcal{F}$, como la única función de la forma,

$$\mathbb{E}(X|\mathcal{F}) : \Omega \to \mathbb{R}$$

Que satisface lo siguiente:

(i) $\mathbb{E}(X|\mathcal{F})$ es $\mathcal{F}$-medible.

(ii) Para todo $A \in \mathcal{F}$ se cumple que,

$$\int_A \mathbb{E}(X|\mathcal{F}) d\mathbb{P} = \int_A X d\mathbb{P}$$

Notación: Sea X variable aleatoria. Diremos que $X \in \mathcal{F}$ si X es $\mathcal{F}$-medible.

Intuitivamente $\mathcal{F}$ se puede pensar como la información de algo y $A \in \mathcal{F}$ es una ocurrencia específica de la información. $\mathbb{E}(X|\mathcal{F})$ es el mejor valor que puede obtener X con respecto a la información $\mathcal{F}$.

Ejemplo 4.1: Si $X \in \mathcal{F}$, entonces la mejor información de X con respecto a $\mathcal{F}$ es literalmente el mismo X, es decir, $\mathbb{E}(X|\mathcal{F}) = X$. Luego más adelante probaremos la unicidad de la esperanza condcional, pero esto nos verifica la igualdad.

Ejemplo 4.2: Si X es una variable aleatoria independiente de $\mathcal{F}$, es decir,

$$\mathbb{P}(\{X \in B\} \cap A) = \mathbb{P}(X \in B)\mathbb{P}(A)$$

Para todo $B \in \mathcal{B}(\mathbb{R}), A \in \mathcal{F}$. Por lo que $\sigma(X)$ es indepeneitne de $\mathcal{F}$ y se cumple que $\mathbb{E}(X|\mathcal{F}) = \mathbb{E}(X)$ pensando en la esperanza como una constante. Y esto se cumple debido a que claramente $\mathbb{E}(X)$ es $\mathcal{F}$-medible al ser una constante y además,

$$\int_A \mathbb{E}(X) d\mathbb{P} = \mathbb{E}(X)\mathbb{P}(A)$$

para todo $A \in \mathcal{F}$, y por otro lado por la independencia de X y $\mathcal{F}$, se tiene que,

$$\int_A X d\mathbb{P} = \mathbb{E}(X\mathbb{1}_A) = \mathbb{E}(X)\mathbb{E}(\mathbb{1}_A) = \mathbb{E}(X)\mathbb{P}(A)$$

Luego por la unicidad se verifica la igualdad.

Lema 4.3: *Sea X una variable aleatoria tal que $\mathbb{E}(|X|) < \infty$, entonces $\mathbb{E}(X|\mathcal{F})$ es integrable*

Dem: Consideremos el siguiente conjunto $A := \{\omega \in \Omega : \mathbb{E}(X|\mathcal{F}) > 0\}$. Observemos que $A \in \mathcal{F}$ puesto que la esperanza de X dado $\mathcal{F}$ es $\mathcal{F}$-medible, entonces A es la preimagen de un Borelano de $\mathbb{R}$. Luego se tiene que,

$$\begin{aligned}\int_A \mathbb{E}(X|\mathcal{F})d\mathbb{P} &= \int_A Xd\mathbb{P} \\ &\leq \int_A Xd\mathbb{P} < \infty\end{aligned}$$

Por la integrabilidad de X. Por otro lado,

$$\begin{aligned}-\int_{A^c} \mathbb{E}(X|\mathcal{F})d\mathbb{P} &= -\int_{A^c} Xd\mathbb{P} \\ &\leq \int_{A^c} Xd\mathbb{P} < \infty\end{aligned}$$

También por la integrablilidad de X. Finalmente,

$$\int_\Omega |\mathbb{E}(X|\mathcal{F})|d\mathbb{P} = \int_A \mathbb{E}(X|\mathcal{F})d\mathbb{P} - \int_{A^c} \mathbb{E}(X|\mathcal{F})d\mathbb{P} < \infty$$

Por tanto, $\mathbb{E}(X|\mathcal{F})$ es integrable. ■

Lema 4.4: *La esperanza es única casi seguramente.*

Dem: Supongamos que la esperanza condicional existe (el siguiente resultado demustra la existencia). Sea $E : \Omega \to \mathbb{R}$ otra esperanza condicional asociada a la variable aleatoria X con respecto a $\mathcal{F}$. Entonces por definición se cumple que E es $\mathcal{F}$-medible y

$$\int_A \mathbb{E}(X|\mathcal{F})d\mathbb{P} = \int_A Ed\mathbb{P}$$

para todo $A \in \mathcal{F}$.

Sea $A_\varepsilon = \{\omega \in \Omega : \mathbb{E}(X|\mathcal{F}) - E \geq \varepsilon > 0\}$. Observemos que $A_\varepsilon \in \mathcal{F}$ puesto que $\mathbb{E}(X|\mathcal{F}), E$ son $\mathcal{F}$-medibles por definición. Luego se cumple que,

$$0 = \int_{A_\varepsilon} Xd\mathbb{P} - \int_{A_\varepsilon} Xd\mathbb{P} = \int_{A_\varepsilon} (\mathbb{E}(X|\mathcal{F}) - E)d\mathbb{P} \geq \int_{A_\varepsilon} \varepsilon d\mathbb{P} = \varepsilon\mathbb{P}(A_\varepsilon)$$

Por lo tanto $\mathbb{P}(A_\varepsilon) = 0$ para todo $\varepsilon > 0$. Dicho de otra forma, el evento,

$$A = \{\omega \in \Omega : \mathbb{E}(X|\mathcal{F}) \geq E\}$$

tiene probabilidad 0. De forma análoga se puede concluir el conjunto de los $\omega \in \Omega$ tales que $E \geq \mathbb{E}(X|\mathcal{F})$. Por tanto $\mathbb{E}(X|\mathcal{F}) = E$ casi seguramente. Probando la unicidad. ■

Lema 4.5: *Sea X integrable. Entonces siempre existe una función de X dado $\mathcal{F}$ que satisface las condiciones de esperanza condicional.*

Dem: Sea $(\Omega, \mathcal{F}_0, \mathbb{P})$ un espacio de probabilidad y sea $\mathcal{F} \subseteq \mathcal{F}_0$ una σ-álgebra. Supongamos que $X \geq 0$ es una variable aleatoria y sea la medida μ sobre $(\Omega, \mathcal{F})$ definido de la siguiente forma:

$$\mu(A) := \int_A Xd\mathbb{P}$$

para todo $A \in \mathcal{F}$. Claramente μ es una medida finita, por lo que por **RN** existe una función $Y : \Omega \to [0, \infty)$ que es $\mathcal{F}$-medible tal que,

$$\nu(A) = \int_A Y d\mathbb{P}$$

De aquí basta ver que $Y = \mathbb{E}(X|\mathcal{F})$.

Supongamos ahora que solamente se cumple $\mathbb{E}(|X|) < \infty$. Entonces $X = X^+ - X^-$ donde X^+, X^- son funciones no negativas, por el caso anterior existen $\mathbb{E}(X^+|\mathcal{F}), \mathbb{E}(X^-|\mathcal{F})$ las esperanzas condicionales de X^+, X^- respectivamente. Definimos,

$$E := \mathbb{E}(X^+|\mathcal{F}) - \mathbb{E}(X^-|\mathcal{F})$$

Veamos que E satisface las condiciones de esperanza condicional con respecto a la variable aleatoria X y $\mathcal{F}$. Claramente es $\mathcal{F}$-medible y,

$$\begin{aligned} \int_A E d\mathbb{P} &= \int_A \mathbb{E}(X^+|\mathcal{F}) d\mathbb{P} - \int_A \mathbb{E}(X^-|\mathcal{F}) d\mathbb{P} \\ &= \int_A X^+ d\mathbb{P} - \int_A X^- d\mathbb{P} \\ &= \int_A X d\mathbb{P} \end{aligned}$$

Por lo tanto $\mathbb{E}(X|\mathcal{F}) = \mathbb{E}(X^+|\mathcal{F}) - \mathbb{E}(X^-|\mathcal{F})$, probando la existencia. ■

Ejemplo 4.6: Sea $\{\Omega_n\}_{n\in\Gamma}$ una partición disjunta finita o infinita de Ω con probabilidad positiva y sea $\mathcal{F} = \sigma(\Omega_1, \Omega_2, \dots)$. Entonces,

$$\mathbb{E}(X|\mathcal{F}) = \frac{\mathbb{E}(X|_{\Omega_n})}{\mathbb{P}(\Omega_n)}$$

sobre Ω_n. Para verificar esto debemos las condiciones de esperanza condicional. Por definición se tiene que,

$$\frac{\mathbb{E}(X|_{\Omega_n})}{\mathbb{P}(\Omega_n)}$$

Es una constante en Ω_n. Por lo que en Ω se tiene una composición de constantes, de forma que es $\mathcal{F}$-medible. Por otro lado se cumple que,

$$\int_{\Omega_i} \frac{\mathbb{E}(X|_{\Omega_n})}{\mathbb{P}(\Omega_n)} d\mathbb{P} = \frac{\mathbb{E}(X|_{\Omega_n})}{\mathbb{P}(\Omega_n)} \mathbb{P}(\Omega_i) = \int_{\Omega_i} X d\mathbb{P}$$

Como queriamos probar.

4.1. Probabilidad Condicional

Como cabría esperar, la esperanza condicional esta relacionado con la condicionalidad en probabilidad. Sea $X = \mathbb{1}_A$ con $A \in \mathcal{F}_0$, definimos la probabilidad condicional de A con respecto a $\mathcal{F}$ por,

$$\mathbb{P}(A|\mathcal{F}) := \mathbb{E}(\mathbb{1}_A|\mathcal{F})$$

Consideremos ahora $B \in \mathcal{F}_0$, definimos la probabilidad condicional de A dado B por,

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

Por último, sean X, Y variables aleatorias, entonces definimos la esperanza condicional de X dado Y por,

$$\mathbb{E}(X|Y) := \mathbb{E}(X|\sigma(Y))$$

Ejemplo 4.7: Sean X e Y variables aleatorias no necesariamente independientes, con densidad conjunta $f(x, y)$, es decir,

$$\mathbb{P}(X \in A, Y \in B) = \iint\limits_{A\times B} f(x, y) d\mathbb{P} \otimes d\mathbb{P}$$

Supongamos que,

$$\int_{\mathbb{R}} f(x, y)dx > 0$$

para todo y. Vamos a probar que dado $g : \mathbb{R} \to \mathbb{R}$ medible tal que $\mathbb{E}(|g(X)|) < \infty$, entonces $\mathbb{E}(g(X)|Y) = h(Y)$ donde,

$$h(y) = \int_{\mathbb{R}} g(x)f(x, y)dy \Bigg/ \int_{\mathbb{R}} f(x, y)dx$$

Para deducir esta fórmula necesitamos estudiar valores particulaes, en particular,

$$\mathbb{P}(X = x|Y = y) = \frac{f(x, y)}{\int_{\mathbb{R}} f(x, y)dx}$$

De esta forma,

$$\mathbb{E}(g(X)|Y = y) = \int_{\mathbb{R}} g(x)\mathbb{P}(X = x|Y = y)dx$$

De aquí se obtiene la fórmula. Ahora solo debemos verificar que h satisface las propiedades de esperanza condicional. Claramente se cumple la primera puesto que $h(Y) \in \sigma(Y)$ y para la segunda se toma $A = \{\omega \in \Omega : Y(\omega) \in B\} \in \sigma(Y)$ donde $B \in \mathcal{R}$. Luego,

$$\begin{aligned}\mathbb{E}(h(Y)|_A) &= \int_B \int_{\mathbb{R}} h(y) f(x,y) dx dy \\ &= \int_B \int_{\mathbb{R}} g(x) f(x,y) dx dy \\ \mathbb{E}(g(X)\mathbb{1}_B(Y)) &= \mathbb{E}(g(X)|_A)\end{aligned}$$

En general podemos pedir que $\int_{\mathbb{R}} f(x,y)dx$ no sea positiva.

Proposición 4.8: *Sean X, Y variables aleatorias y sean a, b constantes, entonces,*

(a) Se satisface la linealidad con respecto a constantes, es decir,

$$\mathbb{E}(aX + b|\mathcal{F}) = a\mathbb{E}(X|\mathcal{F}) + b$$

(b) Si $X \leq Y$ entonces se preserva el orden casi seguramente, es decir,

$$\mathbb{E}(X|\mathcal{F}) \leq \mathbb{E}(Y|\mathcal{F})$$

casi seguramente.

Dem:

(a) Es simplemente por definición. Notemos que $a\mathbb{E}(X|\mathcal{F}) + b$ es $\mathcal{F}$-medible al ser producto y suma de constantes que son medibles. Y además, para todo $A \in \mathcal{F}$ se tiene que,

$$\begin{aligned}\int_A \mathbb{E}(aX + b|\mathcal{F}) d\mathbb{P} &= \int_A (aX + b) d\mathbb{P} \\ &= a \int_A X d\mathbb{P} + \int_A b d\mathbb{P} \\ &= a \int_A \mathbb{E}(X|\mathcal{F}) d\mathbb{P} + \int_A b d\mathbb{P} \\ &= \int_A (a\mathbb{E}(X|\mathcal{F}) + b) d\mathbb{P}\end{aligned}$$

(b) Probemos que $\mathbb{E}(X - Y|\mathcal{F}) \leq 0$. Consideremos,

$$A_\varepsilon := \{\omega \in \Omega : \mathbb{E}(X|\mathcal{F}) - \mathbb{E}(Y|\mathcal{F}) \leq 0\} \in \mathcal{F}$$

Entonces,

$$\begin{aligned}0 &\geq \int_{A_\varepsilon} (X - Y) d\mathbb{P} \\ &= \int_{A_\varepsilon} \mathbb{E}(X|\mathcal{F}) d\mathbb{P} - \int_{A_\varepsilon} \mathbb{E}(Y|\mathcal{F}) d\mathbb{P} \\ &= \int_{A_\varepsilon} (\mathbb{E}(X|\mathcal{F}) - \mathbb{E}(Y|\mathcal{F})) d\mathbb{P} \\ &\geq \varepsilon \mathbb{P}(A_\varepsilon) \geq 0\end{aligned}$$

Siendo una contradicción, por lo tanto necesariamente $\mathbb{P}(A_\varepsilon) = 0$ para todo ε. De forma que se cumple la desigualdad casi seguramente.

Teorema 4.9: *Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias no negativas tal que $X_n \uparrow X$ (X_n son crecientes y convergen puntualmente a una variable aleatoria X). Entonces $\mathbb{E}(X_n|\mathcal{F}) \uparrow \mathbb{E}(X|\mathcal{F})$ con $\mathbb{E}(X) < \infty$.*

Dem: Sea $Y_n := X - X_n$, claramente $Y_n \geq 0$ y que $Y_n \downarrow 0$. Basta demostrar que $\mathbb{E}(Y_n|\mathcal{F}) \downarrow 0$. Observemos que $\mathbb{E}(Y_n|\mathcal{F})$ es una secuencia decreciente casi seguramente, por lo que,

$$\mathbb{E}(Y_n|\mathcal{F}) \geq \mathbb{E}(0|\mathcal{F}) = 0$$

La última igualdad se cumple debido a que toda constante es una función medible. Luego existe un Z_∞ yal que $\mathbb{E}(Y_n|\mathcal{F}) \downarrow Z_\infty$ donde Z_∞ es $\mathcal{F}$-medible. Finalmente,

$$\int_A \mathbb{E}(Y_n|\mathcal{F})d\mathbb{P} = \int_A Y_n d\mathbb{P} \overset{n\to\infty}{\longrightarrow} 0$$

Para todo $A \in \mathcal{F}$, por el teorema de la convergencia monótona, pero por otro lado

$$\int_A \mathbb{E}(Y_n|\mathcal{F})d\mathbb{P} \overset{n\to\infty}{\longrightarrow} \int_A Z_\infty d\mathbb{P}$$

para todo $A \in \mathcal{F}$. Si la esperanza condicional de una función integrable hace que la misma esperanza sea integrable, se obtiene que,

$$\int_A Z_\infty d\mathbb{P} = 0$$

para tood $A \in \mathcal{F}$, es decir, $Z_\infty \in \mathcal{F}$ y finalmente $Z_\infty = 0$ casi seguramente. ■

Teorema 4.10: *Si $\mathcal{F} \subseteq \mathcal{G}$, entonces se tiene que,*

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

(b) $\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{F}) = \mathbb{E}(X|\mathcal{F})$.

Dem:

(a) Observemos que $\mathbb{E}(X|\mathcal{F})$ es $\mathcal{G}$-medible, puesto que es $\mathcal{F}$-medible, luego para todo $A \in \mathcal{R}$ se tiene que $X^{-1}(A) \in \mathcal{F}$, o mejor dicho $X^{-1}(A) \in \mathcal{G}$. Ahora, para todo $A \in \mathcal{G}$ se tiene que,

$$\int_A \mathbb{E}(\mathbb{E}(X|\mathcal{F})|\mathcal{G})d\mathbb{P} = \int_A \mathbb{E}(X|\mathcal{F})d\mathbb{P}$$

(b) Queremos mostrar que $\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{F})$ satisface las propiedades que definen a $\mathbb{E}(X|\mathcal{F})$. Claramente $\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{F})$ es $\mathcal{F}$-medible y para todo $A \in \mathcal{F}$ se cumple que,

$$\begin{aligned}\int_A \mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{F})d\mathbb{P} &= \int_A \mathbb{E}(X|\mathcal{G})d\mathbb{P} \\ &= \int_A X d\mathbb{P}\end{aligned}$$

Por tanto,

$$\mathbb{E}(\mathbb{E}(X|\mathcal{G})|\mathcal{F}) = \mathbb{E}(X|\mathcal{F})$$

Probando el teorema. ■

Teorema 4.11: *Sea $X \in \mathcal{F}$ tal que, $\mathbb{E}(|XY|) < \infty$ y $\mathbb{E}(|Y|) < \infty$. Entonces,*

$$\mathbb{E}(XY|\mathcal{F}) = X\mathbb{E}(Y|\mathcal{F})$$

Dem:

- **Paso 1:** Demostramos por construcción de una función medible. Consideremos $X = \mathbb{1}_A$ donde $A \in \mathcal{F}$. Probemos que $\mathbb{1}_A\mathbb{E}(Y|\mathcal{F})$ satisface las propiedades de $\mathbb{E}(\mathbb{1}_AY|\mathcal{F})$. Claramente $\mathbb{1}_A\mathbb{E}(Y|\mathcal{F})$ es $\mathcal{F}$-medible al ser producto de funciones $\mathcal{F}$-medibles. Ahora consideremos $B \in \mathcal{F}$, entonces,

$$\begin{aligned}\int_B \mathbb{1}_A\mathbb{E}(Y|\mathcal{F})d\mathbb{P} &= \int_{B\cap A} \mathbb{E}(Y|\mathcal{F})d\mathbb{P}\\ &= \int_{B\cap A} Y d\mathbb{P}\\ &= \int_B \mathbb{1}_A Y d\mathbb{P}\end{aligned}$$

  Por tanto $\nVdash_A\mathbb{E}(Y|\mathcal{F}) = \mathbb{E}(\mathbb{1}_AY|\mathcal{F})$.

- **Paso 2:** Sea X una función simple de la forma no negativa,

$$X = \sum_{i=1}^{n} a_i\mathbb{1}_{A_i}$$

  con $A_i \in \mathcal{F}$ para todo $i = 1, \ldots, n$. Entonces,

$$\begin{aligned}\mathbb{E}(XY|\mathcal{F}) &= \mathbb{E}\left(\sum_{i=1}^{n} a_i\mathbb{1}_{A_i}Y|\mathcal{F}\right)\\ &= \sum_{i=1}^{n} a_i\mathbb{E}(\mathbb{1}_{A_i}Y|\mathcal{F})\\ &= \sum_{i=1}^{n} a_i\mathbb{1}_{A_i}\mathbb{E}(Y|\mathcal{F})\\ &= X\mathbb{E}(Y|\mathcal{F})\end{aligned}$$

  Paso 3: Sean $X \geq 0, Y \geq 0$ variables aleatorias no negativos. Como $X \geq 0$ entonces existe una secuencia de funciones simples no negativas $\{X_n\}_{n\in\mathbb{N}}$ tal que $X_n \uparrow X$, por lo tanto $X_nY \uparrow XY$. Como $\mathbb{E}(XY) < \infty$ se tiene que $\mathbb{E}(X_nY|\mathcal{F}) \uparrow \mathbb{E}(XY|\mathcal{F})$, por otro lado, como X_n es una función simple se tiene que,

$$\mathbb{E}(X_nY|\mathcal{F}) = X_n\mathbb{E}(Y|\mathcal{F}) \uparrow X\mathbb{E}(Y|\mathcal{F})$$

Por lo tanto,

$$\mathbb{E}(XY|\mathcal{F}) = X\mathbb{E}(Y|\mathcal{F})$$

- **Paso 4:** Para demostrar el caso general basta ver que,

$$XY = XY\mathbb{1}_{\{X\geq 0, Y\geq 0\}} + XY\mathbb{1}_{\{X<0, Y\geq 0\}} + XY\mathbb{1}_{\{X\geq 0, Y<0\}} + XY\mathbb{1}_{\{X<0, Y<0\}}$$

  Luego es aplicar lo visto en el paso anterior.

Probando el teorema. ■

4.2. Martingala

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias y sea $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ una filtración, es decir, $\mathcal{F}_n \subseteq \mathcal{F}$ son σ-álgebra tales que $\mathcal{F}_n \subseteq \mathcal{F}_{n+1}$ para todo $n \in \mathbb{N}$. Diremos que $\{X_n\}_{n\in\mathbb{N}}$ es un **martingala** con respecto a la filtración $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ si cumple las siguientes condiciones,

(a) $X_n \in \mathcal{F}_n$ para todo $n \in \mathbb{N}$.

(b) $\mathbb{E}(|X_n|) < \infty$ para todo $n \in \mathbb{N}$.

(c) $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$.

Nota 4.12: En muchos ejemplos de martingalas se trabaja en $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$.

Consideremos todas las hipótesis de una martingala. Supongamos que se cumple (a) y (b). Diremos que $\{X_n\}_{n\in\mathbb{N}}$ es una **supermartingala** con respecto a la filtración $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$, si $\mathbb{E}(X_{n+1}|\mathcal{F}_n) \leq X_n$ y diremos que es una **submartingala** si $\mathbb{E}(X_{n+1}|\mathcal{F}_n) \geq X_n$.

Ejemplo 4.13: Sea $\{\xi_i\}_{i\in\mathbb{N}}$ una sucesión de variables aleatorias independientes e identicamente distribuidas tales que $\mathbb{E}(|\xi_i|) < \infty$ y $\mathbb{E}(\xi_i) = 0$ para todo $i \in \mathbb{N}$. Consideremos $X_n = \xi_1 + \cdots + \xi_n$ una variable aleatoria. Consideremos la filtración $\mathcal{F}_n = \sigma(X_1, \ldots, X_n) = \sigma(\xi_1, \ldots, \xi_n)$. Entonces $\{X_n\}_{n\in\mathbb{N}}$ es un martingala con respecto a $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$. Para ello comprobemos que se cumplen los axioma de martingala.

(a) Claramente $X_n \in \mathcal{F}_n$, puesto que la suma de variables aleatorias $\mathcal{F}_n$-medibles.

(b) Por propiedades de esperanza, se cumple que,

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

Para todo $n \in \mathbb{N}$.

(c) Por propiedades de esperanza condicional, se cumple que,

$$\begin{aligned}
\mathbb{E}(X_{n+1}|\mathcal{F}_n) &= \mathbb{E}(\xi_{n+1} + X_n|\mathcal{F}_n) \\
&= \mathbb{E}(\xi_{n+1}|\mathcal{F}_n) + \mathbb{E}(X_n|\mathcal{F}_n) \\
&= \mathbb{E}(\xi_{n+1}) + X_n \\
&= X_n
\end{aligned}$$

donde ξ_{n+1} es independiente de $\sigma(\xi_1, \ldots, \xi_n) = \mathcal{F}_n$ y X_n es $\mathcal{F}_n$-medible. Esto para todo $n \in \mathbb{N}$

Verificando que $\{X_n\}_{n\in\mathbb{N}}$ es un martingala con respecto a $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$.

Ejemplo 4.14: Sea $\{\xi_i\}_{i\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e identicamente distribuidas tal que $\mathbb{E}(|\xi_i|) < \infty$, con $\mathbb{E}(\xi_i) = 0$ y con varianza $\text{Var}(\xi_i) = \sigma^2 < \infty$ para todo $i \in \mathbb{N}$. Sea $S_n = \xi_1 + \cdots + \xi_n$ y sea $X_n = S_n^2 - n\sigma^2$. Entonces $\{X_n\}_{n\in\mathbb{N}}$ es un martingala con respecto a la filtración dada por $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$. Probemos los tres axiomas de martingala.

(a) Claramente X_n es $\mathcal{F}_n$-medible, puesto que S_n^2 es variable aleatoria y al resta por $n\sigma^2$, una constante, se preserva la condición de ser medible.

(b) Observemos que,

$$\begin{aligned}\mathbb{E}(|X_n|) &= \mathbb{E}(|S_n^2 - n\sigma^2|)\\ &\leq \mathbb{E}(|S_n^2|) + n\sigma^2\\ &= n\text{Var}(\xi_i) + n\sigma^2\\ &= 2n\sigma^2 < \infty\end{aligned}$$

Para todo $n \in \mathbb{N}$.

(c) Por propiedades de esperanza condiciona, se cumple que,

$$\begin{aligned}\mathbb{E}(X_{n+1}|\mathcal{F}_n) &= \mathbb{E}(S_{n+1}^2 - (n+1)\sigma^2|\mathcal{F}_n)\\ &= \mathbb{E}(S_{n+1}^2|\mathcal{F}_n) - (n+1)\sigma^2\\ &= \mathbb{E}((S_n + \xi_{n+1})^2|\mathcal{F}_n) - (n+1)\sigma^2\\ &= \mathbb{E}(S_n^2|\mathcal{F}_n) + \mathbb{E}(2S_n\xi_{n+1}|\mathcal{F}_n) + \mathbb{E}(\xi_n^2) - (n+1)\sigma^2\\ &= S_n^2 + 2S_n\mathbb{E}(\xi_{n+1}|\mathcal{F}_n) - n\sigma^2\\ &= S_n^2 - n\sigma^2\\ &= X_n\end{aligned}$$

Para todo $n \in \mathbb{N}$.

Verificando que $\{X_n\}_{n\in\mathbb{N}}$ es un martingala con respecto a $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$.

Ejemplo 4.15: Sea $\{Y_n\}_{n\in\mathbb{N}}$ una variable aleatoria independientes e identicamente distribuidas no negativas con esperanza $\mathbb{E}(Y_i) = 1$ para todo $i \in \mathbb{N}$. Sea $X_n = Y_1Y_2\ldots Y_n$. Entonces $\{X_n\}_{n\in\mathbb{N}}$ es un martingala con respecto a la filtración dada por $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n)$. Probemos los tres axiomas de martingala.

(a) Claramente $X_n \in \mathcal{F}_n$, puesto que es producto de funciones $\mathcal{F}_n$-meidbles.

(b) Observemos que X_n son variables aleatorias no negativas independientes, de forma que,

$$\mathbb{E}(|X_n|) = \mathbb{E}(X_n) = \prod_{i=1}^{n} \mathbb{E}(Y_i) = 1 < \infty$$

Para todo $n \in \mathbb{N}$.

(c) Observemos que para todo $n \in \mathbb{N}$ se cumple que,

$$\begin{aligned}\mathbb{E}(X_{n+1}|\mathcal{F}_n) &= \mathbb{E}(X_n Y_{n+1}|\mathcal{F}) \\ &= X_n \mathbb{E}(Y_{n+1}|\mathcal{F}_n) \\ &= X_n \mathbb{E}(Y_{n+1}) = X_n\end{aligned}$$

Verificando que $\{X_n\}_{n\in\mathbb{N}}$ Nes un martingala con respecto a $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$.

Ejemplo 4.16: Sea $\{\xi_i\}_{i\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e identicamente distribuidas tales que $\mathbb{E}(e^{\theta\xi_i}) < \infty$ para todo $i \in \mathbb{N}$. Se define,

$$X_n = \frac{e^{\theta \sum_{i=1}^{n} \xi_i}}{\mathbb{E}(e^{\theta\xi_i})^n}$$

Observemos que al escoger,

$$Y_n = \frac{e^{\theta\xi_i}}{\mathbb{E}(e^{\theta\xi_i})}$$

se forma una colección de variables aleatorias indendientes, luego $X_n = \prod_{i=1}^{n} Y_i$. Tomando la filtración $\mathcal{F}_n = \sigma(\xi_1, \ldots, \xi_n)$ obtenemos por ejemplo anterior que $\{X_n\}_{n\in\mathbb{N}}$ es una martinala con respecto a $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$.

Teorema 4.17: *Sea $\{X_n\}_{n\in\mathbb{N}}$ una supermartingala con respecto a filtración $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$. Entonces se cumple que,*

$$\mathbb{E}(X_{n+k}|\mathcal{F}_n) \leq X_n$$

para todo $n \in \mathbb{N}$ y para todo $k \geq 0$.

Dem: Demostraremos por inducción sobre k. Si $k = 1$ claramente se cumple. Supongamos que se cumple para k, luego,

$$\begin{aligned}\mathbb{E}(X_{n+k+1}|\mathcal{F}_n) &= \mathbb{E}(\mathbb{E}(X_{n+k+1}|\mathcal{F}_{n+1})|\mathcal{F}_n) \\ &= \mathbb{E}(\mathbb{E}(X_{n+1+k}|\mathcal{F}_{n+1})|\mathcal{F}_n) \\ &\leq \mathbb{E}(X_{n+1}|\mathcal{F}_n) \leq X_n\end{aligned}$$

Probando el teorema. ■

Teorema 4.18: *Sea $\varphi : \mathbb{R} \to \mathbb{R}$ una función convexa y sea $\{X_n\}_{n\in\mathbb{N}}$ un martingala con respecto a la filtración $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ tal que $\mathbb{E}(|\varphi(X_n)|) < \infty$ para todo $n \in \mathbb{N}$. Entonces $\{\varphi(X_n)\}_{n\in\mathbb{N}}$ es una submartingala con respecto a la filtración $\{\mathcal{F}_n\}$.*

Antes de demostrar el teorema, necesitamos la desigualdad de Jensen para la esperanza condicional.

Lema 4.19: *Sea $\varphi : \mathbb{R} \to \mathbb{R}$ una función convexa. Entonces,*

$$\varphi(\mathbb{E}(X|\mathcal{F})) \leq \mathbb{E}(\varphi(X)|\mathcal{F})$$

Dem: Sea $D := \{(a, b) : ax + b \leq \varphi(x), \forall x \in \mathbb{R}\}$, es decir, D es la colección de puntos, tales que la función afín generada está debajo de φ con respecto a todos los puntos $x \in \mathbb{R}$.

Afirmación: *Se cumple que,*

$$\varphi(x) = \sup_{(a,b)\in D} \{ax + b\}$$

Dem: Por definición de D es claro que,

$$\varphi(x) \geq \sup_{(a,b)\in D} \{ax + b\}$$

Para concluir la igualdad debemos notar que si φ es convexa entonces se puede definir derivadas, en particular, para un puntos $x \in \mathbb{R}$ existe un punto (a, b) tal que $\varphi(x) = ax + b$. Esto para todo $x \in \mathbb{R}$. ■

Luego se tiene que, dado X variable aleatorio y para todo $(a, b) \in D$ se cumple que

$$\varphi(X) \geq aX + b$$

Por la afirmación anterior. Luego por la linealidad de la esperanza condicional se cumple que,

$$\mathbb{E}(\varphi(X)|\mathcal{F}) \geq a\mathbb{E}(X|\mathcal{F}) + b$$

Por lo tanto,

$$\mathbb{E}(\varphi(X)|\mathcal{F}) \geq \sup_{(a,b)\in D} \{a\mathbb{E}(X|\mathcal{F}) + b\} = \varphi(\mathbb{E}(X|\mathcal{F}))$$

Probando la desigualdad de Jensen para la esperanza condicional. ■

Dem (Teorema 4.18): Observemos algunas cosas. $\varphi(X_n)$ es una variable aleatoria $\mathcal{F}_n$-medible, puesto que φ es continua y luego por composición de medibles se obtiene que $\varphi(X_n) \in \mathcal{F}_n$. Y por la hipótesis del enunciado se obtiene la segunda condición de martingala. Para concluir que es una submartingala basta usar el lema anterior, de forma que,

$$\begin{aligned}\mathbb{E}(\varphi(X_{n+1})|\mathcal{F}_n) &\geq \varphi(\mathbb{E}(X_n|\mathcal{F})) \\ &= \varphi(X_n)\end{aligned}$$

Probando el teorema. ■

Nota 4.20: La desigualdad de Jensen se puede verificar sobre una función cóncava, con la única diferencia que cambia la desigualdad y el **teorema 4.18** se puede probar que si φ es cóncava, entonces $\{\varphi(X_n)\}$ es una supermartingala.

Corolario 4.21: *Sea $\{X_n\}_{n\in\mathbb{N}}$ una submartingala y sea φ una función creciente y convexa tal que $\mathbb{E}(\varphi(X_n)) < \infty$ para todo $n \in \mathbb{N}$. Entonces $\{\varphi(X_n)\}_{n\in\mathbb{N}}$ es una submartingala.*

Ejemplo 4.22: Sea $\{X_n\}_{n\in\mathbb{N}}$ una martingala, entonces $\{|X_n + a|_+\}_{n\in\mathbb{N}}$ es una submartingala para todo $a \in \mathbb{R}$, donde $|\cdot|_+$ toma valor 0 cuando el argumento es negativo y cuando el argumento es no negativo, se comporta igual al valor absoluto. Por lo que es creciente y convexa.

Definición 4.23: *Sea $\{H_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias definidas en el espacio $(\Omega, \mathcal{F}, \mathbb{P})$ y sea $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ una filtración. Decimos que $\{H_n\}_{n\in\mathbb{N}}$ es predecible con respecto a la filtración $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ si $H_{n+1} \in \mathcal{F}_n$.*

Teorema 4.24: *Sea $\{X_n\}_{n\in\mathbb{N}}$ una submartingala con filtración $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$. Definimos,*

$$(H, X)_n := \sum_{i=1}^{n} H_i(X_i - X_{i-1})$$

Entonces si $\{H_n\}_{n\in\mathbb{N}}$ es una secuencia acotada, no negativa y predecible. Se tiene que $\{(H, X)_n\}_{n\in\mathbb{N}}$ es una submartingala con la misma filtración.

Dem: Observemos que $(H, X)_n$ es $\mathcal{F}_n$ medible al ser suma y producto de funciones $\mathcal{F}_n$ medibles. Y $\mathbb{E}(|(H, X)_n|) < \infty$ por hipótesis. Nos queda probar la última condición de submartingala. Observemos que,

$$\begin{aligned}\mathbb{E}((H, X)_{n+1}|\mathcal{F}_n) &= \mathbb{E}((H, X)_n|\mathcal{F}_n) + \mathbb{E}(H_{n+1}(X_{n+1} - X_n)|\mathcal{F}_n)\\ &= (H, X)_n + H_{n+1}\mathbb{E}(X_{n+1} - X_n|\mathcal{F}_n)\\ &\geq (H, X)_n\end{aligned}$$

La última desigualdad se cumple dado que $\mathbb{E}(X_{n+1} - X_n|\mathcal{F}) = \mathbb{E}(X_{n+1}|\mathcal{F}) - \mathbb{E}(X_n|\mathcal{F}_n) \geq X_n - X_n = 0$. Probando el teorema. ■

Corolario 4.25: *Sea $\{X_n\}_{n\in\mathbb{N}}$ una supermartingala con filtración $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$. Sea $\{H_n\}_{n\in\mathbb{N}}$ acotada, no negativa y predecible. Entonces $\{(H, X)_n\}_{n\in\mathbb{N}}$ es una supermartingala con la filtración $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$.*

Nota 4.26: Con respecto al corolario anterior se puede probar si $\{X_n\}_{n\in\mathbb{N}}$ es martingala, entonces $\{(H, X)_n\}$ es un martingala.

Definición 4.27: *Sea T un variable aleatoria con valores en $\mathbb{N}_0$. Decimos que T es un tiempo de parada con respecto a una filtración $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ si $\{T = n\} \in \mathcal{F}_n$ para todo $n \in \mathbb{N}$.*

Ejemplo 4.28: Sea T un tiempo de parada y consideremos,

$$H_n = \begin{cases} 1, & n \leq T \\ 0, & n > T \end{cases}$$

Entonces H_n es predecible. Para ver esto notemos que $H_n = \mathbb{1}_{\{T \geq n\}}$, por lo que basta de demostrar que $\{T \geq n\} \in \mathcal{F}_{n-1}$ y para esto observemos lo siguiente,

$$\begin{aligned}\{T \geq n\} &= \{T < n\}^c \\ &= \left(\bigcup_{m=0}^{n-1} \{T = m\}\right)^n\end{aligned}$$

Claramente $\{T = m\} \in \mathcal{F}_{n-1}$, de forma que $\{T \geq n\} \in \mathcal{F}_{n-1}$.

Teorema 4.29: *Sea $\{X_n\}_{n\in\mathbb{N}}$ una supermartingala y sea T un tiempo de parada, entonces,*

$$X_{T\wedge n} - X_0$$

es un supermartingala. ($T \wedge n = \min\{T, n\}$).

Dem: Basta demostrar que $H_n := \mathbb{1}_{T\geq n}$ es una predicción cosa que hemos anteriormente. Luego al tomar $(H, X)_n$ se observa lo siguiente,

$$\begin{aligned}(H, X)_n &= \sum_{i=1}^{n} H_i(X_i - X_{i-1}) \\ &= \begin{cases} X_n - X_0, & n \leq T \\ X_T - X_0, & n > T \end{cases}\end{aligned}$$

De forma que $X_{T\wedge n} - X_0$ es supermartingala. ■

Sea $\{X_n\}_{n\in\mathbb{N}}$ una submartingala. Sea $a < b$. Definimos $N_i(\omega)$ siguiendo el siguente comportamiento en referente a la figura 2, $N_1(\omega) = 1$ es la cantidad de puntos antes del punto de afuera incluyendo el mismo punto afuera del intervalo $[a, b]$, $N_2(\omega) = 4$ ya que hay tres puntos antes del segundo y al incluir este segundo punto se tendría cuatro. Por lo tanto, $N_3(\omega) = 7$ y $N_4(\omega) = 9$

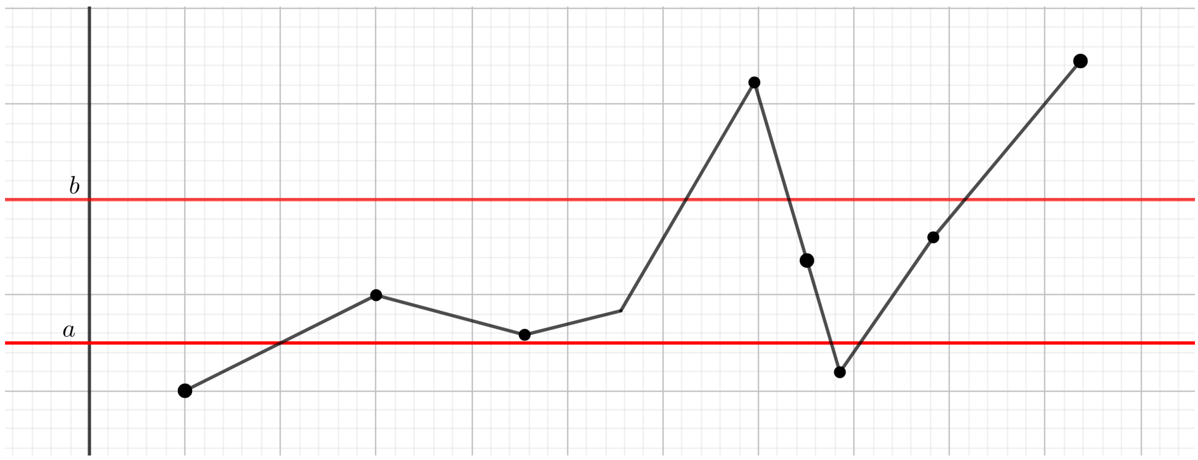

Figura 2: Marcha Aleatoria

De forma muy general definimos la secuencia $\{N_i\}_{n\in\mathbb{N}}$ por $N_0(\omega)=1$,

$$\begin{aligned} N_{2k-1}(\omega) &:= \min\{n > N_{2k-2}(\omega) : X_n(\omega) \leq a\} \\ N_{2k}(\omega) &:= \min\{n > N_{2k-1}(\omega) : X_n(\omega) \geq b\} \end{aligned}$$

Para todo $k \in \mathbb{N}$. De esta forma, que $N_{2k} = m$ significa que en el tiempo m se cumplió el k-ésimo cruce hacía arriba.

Definimos la siguiente secuencia predecible,

$$H_m = \begin{cases} 1, & N_{2k-1} < m \leq N_{2k} \\ 0, & \text{en otro caso} \end{cases}$$

Luego,

$$(H,X)_n = \sum_{i=1}^{n} H_i(X_i - X_{i-1})$$

Verifiquemos que $\{H_m\}_{m\in\mathbb{N}}$ es una secuencia predecible con respecto a la filtración $\mathcal{F}_{m-1}$ dad por,

$$\mathcal{F}_m = \begin{cases} 1, & m \leq T \\ 0, & m > T \end{cases}$$

donde T es un tiempo de pasada, es predecible. Notemos que para todo $k \in \mathbb{N}_0$ N_k es un tiempo de parada, entonces,

$$\{H_m = 1\} \in \mathcal{F}_{m-1}$$

de hecho,

$$\begin{aligned} \{H_m = 1\} &= \bigcup_{k\in\mathbb{N}} \{m > N_{2k}\} \cap \{m \leq N_{2k-1}\} \\ &= \bigcup_{k\in\mathbb{N}} \{m \leq N_{2k}\}^c \cap \{m \leq N_{2k-1}\} \end{aligned}$$

Teorema 4.30: *Si $\{X_n\}_{n\in\mathbb{N}}$ es una submartingala, entonces $\{(H,X)_n\}_{n\in\mathbb{N}}$ es una submartingala. Si $K_n := 1 - H_m \geq 0$, entonces $\{(K,X)_n\}_{n\in\mathbb{N}}$ es una submartingala.*

Teorema 4.31: *Sea $\{X_n\}_{n\in\mathbb{N}}$ una submartingala con respecto a una filtración $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$, entonces para todo $b > a$ se tiene que,*

$$(b-a)\mathbb{E}(U_n) \leq \mathbb{E}((X_n - a)^+) - \mathbb{E}((X_0 - a)^+)$$

donde U_n es el número de cruces hacia arriba ante el tiempo n, es decir, $U_n = \max\{k : N_{2k} \leq n\}$.

Dem: Definimos $Y_n := (X_n - a)^+ + a$ donde $(X_n - a)^+$ es una submartingala. Usando que $(\cdot)^+$ es convexa y creciente, se obtiene que,

$$U_n(b - a \leq (H, Y))_n$$

Por otro lado,

$$\begin{aligned} Y_n - Y_0 = (1, Y)_n &= (H + (1 - H), Y)_n \\ &= (H, Y)_n + (1 - H, Y)_n \end{aligned}$$

Luego,

$$\begin{aligned} \mathbb{E}(Y_n - Y_0) &= \mathbb{E}((H, Y)_n) + \mathbb{E}((1 - H, Y)_n) \\ &\geq \mathbb{E}(U_n(b - a)) + \mathbb{E}((1 - H, Y)_n) \\ &\geq \mathbb{E}(U_n(b - a)) + \mathbb{E}((1 - H, Y)_0) \end{aligned}$$

De forma que,

$$\mathbb{E}(Y_n - Y_0) \geq (b - a)\mathbb{E}(U_n)$$

Reemplazando se obtiene que,

$$\mathbb{E}((X_n - a)^+ - \mathbb{E}((X_0 - a)^+) \geq (b - a)\mathbb{E}(U_n)$$

Teorema 4.32: *Sea $\{X_n\}$ una submartingala con $\sup_{n\in\mathbb{N}} \mathbb{E}(X_n^+) < \infty$. Entonces X_n converge casi seguramente a un límite X cuando $n \to \infty$ y $\mathbb{E}(|X|) < \infty$.*

Dem: Sea $a < b$ reales. Por la desigualdad de cruces, se tiene que,

$$(b - a)\mathbb{E}(U_n) \leq \mathbb{E}((X_n - a)^+) - \mathbb{E}((X_0 - a)^+)$$

En particular,

$$\begin{aligned} \mathbb{E}(U_n) &\leq \frac{1}{b - a}\mathbb{E}((X_n - a)^+) \\ &\leq \frac{1}{b - a}(\mathbb{E}(X_n^+) - |a|) < M \end{aligned}$$

para todo $n \in \mathbb{N}$ y usando que para todo x, y que $(x + y)^+ \leq x^+ + y^+$. Por lo tanto $\mathbb{E}(U_n) < M$ para todo $n \in \mathbb{N}$. Sea,

$$U_\infty = \lim_{n\to\infty} U_n$$

entonces,

$$\mathbb{E}(U_\infty) = \lim_{n\to\infty} \mathbb{E}(U_n) < M$$

Por lo tanto $\mathbb{E}(U_\infty) < \infty$. Esto implica que $U_\infty < \infty$ casi seguramente.

Imporante ($\star$): Tomando $a, b \in \mathbb{Q}$, se tiene que la probabilidad que $\{X_n\}_{n\in\mathbb{N}}$ crece hacia arriba algún intervalo con extremos racionales, es 0. Por lo tanto, la probabilidad de cruzar algún intervalo infinitas veces es 0 y eso implica que,

$$\liminf_{n\to\infty} X_n = \limsup_{n\to\infty} X_n$$

Sea,

$$X = \lim_{n\to\infty} X_n$$

Luego por Fatou,

$$\infty > \liminf_{n\to\infty} \mathbb{E}(\mathbb{X}_\ltimes^+) \geq \left(\liminf_{n\to\infty} X_n^+\right) = \mathbb{E}(X^+)$$

Y por tanto $\mathbb{E}(X^+) < \infty$. Por otro lado, $X_n^- = X_n^+ - X_n$ y por lo tanto,

$$\begin{aligned}\mathbb{E}(X_n^-) &= \mathbb{E}(X_n^+) - \mathbb{E}(X_n)\\ &\leq \mathbb{E}(X_n^+) - \mathbb{E}(X_0)\end{aligned}$$

Aplicando nuevamente Fatou se obtiene que,

$$\begin{aligned}\mathbb{E}(X^-) &= \mathbb{E}\left(\liminf_{n\to\infty} X_n^-\right)\\ &\leq \liminf_{n\to\infty} \mathbb{E}(X_n^-)\\ &\leq \liminf_{n\to\infty} \mathbb{E}(X_n^+) - \mathbb{E}(X_0) < \infty\end{aligned}$$

Finalmente, $\mathbb{E}(X^-) < \infty$ y eso demuestra que,

$$\mathbb{E}(|X|) = \mathbb{E}()X^+ + X^-) < \infty$$

■

Corolario 4.33: *Sea $\{X_n\}_{n\in\mathbb{N}}$ una supermartingala tal que $X_n \geq 0$, entonces X_n es convergente.*

Dem: Se tiene que $\{-X_n\}_{n\in\mathbb{N}}$ es una submartingala, entonces $(-X_n)^+ = 0$, por lo que satisface las hipótesis del teorema anterior. ■

Ejemplo 4.34 (Marcha Aleatoria): Sea $\{\xi_n\}_{n\in\mathbb{N}}$ independientes e identicamente distribuidas tal que,

$$\mathbb{P}(\xi_i = 1) = \mathbb{P}(\xi_i = -1) = \frac{1}{2}$$

y sea,

$$X_n = \sum_{i=1}^{n} \xi_i + 1$$

con la consideración $X_0 = 1$. Sea $T = \text{mín}\{n \geq 0 : X_n = 0\}$, T es un tiempo de parada, entonces $\{X_{n\wedge T}\}_{n\neq 0}$ es una martingala. Entonces, por el corolario anterior $\{X_{n\wedge T}\}_{n\in\mathbb{N}}$ es convergente. En particular, $X_{n\wedge T} = 0$ para n suficientemente grande, o dicho de otra forma $T < \infty$ casi seguramente, (la marcha pasa por 0 infinitamente).

En este caso $X = \text{lím}_{n\to\infty} X_{n\wedge T}$ es identicamente 0, pero,

$$\mathbb{E}(X_{n\wedge T}) = 1$$

para todo $n \in \mathbb{N}$ y $\mathbb{E}(X) = 0$. Dicho de otra forma $X_{n\wedge T}$ converge casi seguramente pero no lo hace L^1 a X.

Ejemplo 4.35: Consideremos un proceso de ramifiación. Sea $\{\xi_i^n\}_{i\in\mathbb{N},n\in\mathbb{N}_0}$ variables aleaotrias que toma valores en $\mathbb{N}_0$ que son independientes e identicamente distribuidas. Definimos $\{Z_n\}_{n\in\mathbb{N}_0}$ de la siguiente forma,

$$\begin{aligned} Z_0 &= 1 \\ Z_n &= \sum_{i=1}^{Z_{n-1}} \xi_i^{n-1} \end{aligned}$$

donde ξ_i^n representa la cantidad de hijos de el i-ésimo individuo en la generación n. Luego $\{Z_n\}_{n\in\mathbb{N}_0}$ es una secuencia de variables aleatorias independientes e identicamente distribuidas que mide la cantidad de individuos en la generación n. Sea $\mu = \mathbb{E}(\xi_i^n)$, entonces se cumplen resultados muy interesantes.

Teorema 4.36: *Si $\mu < \infty$. Se tiene que $\{Z_n/\mu^n\}_{n\in\mathbb{N}}$ es una martingala.*

Dem: Observemos que,

$$\begin{aligned} \mathbb{E}\left(\frac{Z_{n+1}}{\mu^{n+1}}\middle| Z_n = k\right) &= \mathbb{E}\left(\frac{1}{\mu^{n+1}}\sum_{i=1}^{Z_n} \xi_i^n \middle| Z_n = k\right) \\ &= \mathbb{E}\left(\frac{1}{\mu^{n+1}}\sum_{i=1}^{k} \xi_i^n \middle| Z_n = k\right) \\ &= \mathbb{E}\left(\frac{1}{\mu^{n+1}}\sum_{i=1}^{k} \xi_i^n\right) \\ &= \sum_{i=1}^{k} \frac{1}{\mu^{n+1}}\mathbb{E}(\xi_i^n) \\ &= \sum_{i=1}^{k} \frac{1}{\mu^{n+1}} \cdot \mu \\ &= \sum_{i=1}^{k} \frac{1}{\mu^n} = \frac{k}{\mu^n} \end{aligned}$$

En particular, por condición impuesta, de que $Z_n = k$, por lo que,

$$\mathbb{E}\left(\frac{Z_{n+1}}{\mu^{n+1}}\middle| Z_n = k\right) = \frac{k}{\mu^n}$$

para todo $k \in \mathbb{N}$, es decir,

$$\mathbb{E}\left(\frac{Z_{n+1}}{\mu^{n+1}}\middle| Z_n\right) = \frac{Z_n}{\mu^n}$$

Por lo tanto $\{Z_n/\mu^n\}_{n\in\mathbb{N}}$ es una martingala con respecto a la filtración $\mathcal{F}_n = \sigma(\xi_k^i, i \in \mathbb{N}, k = 1, \dots, n)$. ■

Teorema 4.37: *Si $\mu < 1$, entonces,*

$$\mathbb{P}(Z_n = 0) \overset{n\to\infty}{\longrightarrow} 1$$

Dem: Observemos que,

$$\mathbb{P}(Z_n > 0) = \sum_{k=1}^{\infty} \mathbb{P}(Z_k = k) \le \sum_{k=0}^{\infty} k\mathbb{P}(Z_n = k) = \mathbb{E}(Z_n)$$

Ahora, por el teorema anterior se tiene que,

$$\mathbb{E}\left(\frac{Z_n}{\mu^n}\right) = 1$$

Entonces $\mathbb{E}(Z_n) = \mu^n$. Finalmente,

$$\mathbb{P}(Z_n > 0) \le \mu^n \overset{n\to\infty}{\longrightarrow} 0$$

■

Este teorema nos dice que si la esperanza de un individuo sea menor a 1, entonces en un futuro muy lejano no habrán más individuo y por tanto hay extinsión.

Teorema 4.38: *Si $\mu = 1$ y $p_1 < 1$, entonces $\mathbb{P}(Z_n = 0) \longrightarrow 1$ o $Z_n \to 0$ casi seguramente.*

Dem: Usaremos el teorema de convergencia de martingala. En este caso $Z_n/\mu^n = Z_n$, por lo que $\{Z_n\}_{n\in\mathbb{N}}$ es una martingala.

Además $\mathbb{E}(Z_n^+) = \mathbb{E}(Z_n) = 1$ para todo $n \in \mathbb{N}$. Luego $\mathbb{E}(Z_n^+) < 1$ y por el teorema de convergencia de martingala,

$$Z_n \overset{n\to\infty}{\longrightarrow} Z$$

Mostremos que $\mathbb{P}(Z=k)>0$ para todo $k\in\mathbb{N}$, de hecho,

$$\begin{aligned}\mathbb{P}(Z=k)&=\mathbb{P}\left(\lim_{n\to\infty} Z_n=k\right)\\&=\mathbb{P}\left(\bigcup_{i=1}^{\infty}\bigcap_{n=i}^{\infty}(Z_n=k)\right)\\&\leq\sum_{i=1}^{\infty}\mathbb{P}(Z_n=k\,\text{para todo}\,n\geq i)\end{aligned}$$

Pero $\mathbb{P}(Z_n=k\,\text{para todo}\,n\geq i)=0$ dado que si $Z_n=k$ hay una probabilidad positiva que $Z_n<k$, en efecto,

$$\mathbb{P}(Z_{n+1}=0|Z_n=k)=(p_0)^k$$

Pero $p_0>0$ porque,

$$\sum_{i=0}^{\infty}p_i=1,\;\;\mu=\sum_{i=0}^{\infty}ip_i=1$$

como $p_1<1$ tiene que existir $l\geq 2$ tal que $p_l>0$ y en este caso si $p_0=0$ se tiene que,

$$\mu=\sum_{i=0}^{\infty}ip_i\geq p_1+2\sum_{i=2}^{\infty}p_i=p_1+2(1-p_1)=2-p_1>1$$

Demostrando el teorema. ■

Teorema 4.39: *Si $\mu>1$, entonces,*

$$\mathbb{P}(Z_n>0)>0$$

para todo $n\in\mathbb{N}$.

Dem: Definimos,

$$f(x)=\sum_{i=0}^{\infty}x^ip_i$$

claramente está bien definido en $x\in(-1,1)$. Definimos $\theta_m:=\mathbb{P}(Z_m=0)$. Se cumple que,

$$\theta_{m+1}=f(\theta_m)$$

Y esto es debido que,

$$\begin{aligned}\mathbb{P}(Z_{m+1}=0)&=\sum_{i=1}^{\infty}\mathbb{P}(Z_{m+1}=0\cap Z_1=i)\\&=\sum_{i=1}^{\infty}\mathbb{P}(Z_{m+1}==|Z_1=i)\mathbb{P}(Z_1=i)\\&=\sum_{i=1}^{\infty}\mathbb{P}(Z_{m+1}=0|Z_1=i)p_i\end{aligned}$$

Observemos que $\mathbb{P}(Z_m = 0)^i$. Finalmente,

$$\theta_{m+1} = \sum_{i=1}^{\infty} \theta_m^i p_i = f(\theta_m)$$

Afirmación: *Existe un único $x \in [0,1)$ tal que $f(x) = x$.*

Dem: Demostraremos la existencia y luego la unicidad.

- **Existencia:** Observemos que f es una función creciente en $[0,1]$. Si $f(0) = p_0 > 0$ (si $p_0 = 0$ se tiene que $f(0) = 0$), por otro lado,

  $$f'(x) = \sum_{i=1}^{\infty} ix^{i-1}p_i$$

  es una serie bien definida en $(-1,1)$. Se puede ver que $f'(1) = \mu > 1$, esto significa que $f'(x) > 1$ para x muy cerca de 1, en particular,

  $$f(1) - f(1-h) = \int_{1-h}^{1} f'(s)ds \geq h > 0$$

  Luego $f(1) > f(1-h)$ para h muy pequeño. Por otro lado $f(1) = 1$ y lueho $1-f(1-h) > h$ lo que implica que $f(1-h) < 1-h$, como $x \mapsto x, x \mapsto f(x)$ son funciones continuas, esto implica que existe un $x_0 \in [0, 1-h]$ tal qeu $f(x_0) = x_0$, como queriamos probar.

- **Unicidad:** Sea y_0 otro punto fijo, entonces,

  $$f(y_0) = y_0$$

  Ahora observemos que,

  $$\theta$$

  termiar

Probando el teorema. ■

Ejemplo 4.40 (Urna de Polya): Consideremos una urna que contiene a bolas azules y m boas rojas y sea $c \in \mathbb{N}$. Digamos que cada vez que sacamos una bola al azar, se agregan c bolas del mismo color del que fue sacado.

Teorema 4.41: *La propoción de bolas azules (o rojas) es una martingala.*

Dem: Sea i la cantidad de bolas azules y sea j la cantindad de bolas rojas en el paso n-ésimo, de esta forma definimos $X_n = i/i + j$, entonces,

$$\begin{aligned}
\mathbb{E}(X_{n+1}|\mathcal{F}_n) &= \frac{i+c}{i+j+c}\mathbb{P}(\text{sacar bola azul}) + \frac{i}{i+j+c}\mathbb{P}(\text{sacar bola roja}) \\
&= \frac{i+c}{i+j+c}\frac{i}{i+j} + \frac{i}{i+j+c}\frac{j}{i+j} \\
&= \frac{i}{i+j}\left(\frac{i+c}{i+j+c} + \frac{j}{i+j+c}\right) \\
&= \frac{i}{i+j} = X_n
\end{aligned}$$

■

Con respecto a la urna, la martingala converge y será un límite aleatorio por el teorema de converge de martingalas.

Ejemplo 4.42: Observeos la siguiente probabilidad,

$$\begin{aligned}
&\mathbb{P}(\text{sacar } m \text{ bola azulues seguidas y luego sacar } n \text{ rojas seguidas}) \\
&= \frac{a}{a+r}\frac{a++c}{a+r+2c}\cdots\frac{a+(m-1)c}{a+r+c(m-1)}\frac{M}{a+M+cm}\frac{M+c}{a+r+c(m+1)}\cdots\frac{M+(n-1)c}{a+n+c(m+n-1)}
\end{aligned}$$

Observación 4.43: EL orden en que se saco las m bolas azules y las n bolas rojas no afecta a la probabilidad, finalmente,

$$\mathbb{P}(\text{sacar exactamente } m \text{ bolas azules dentro de las primeras } n+m \text{ tiradas}) = \binom{m+n}{m}$$

Casos particulares: si $a = r = c = 1$ vemos que ahy convergencia uniforme, puesto que,

$$\mathbb{P}(\text{sacar } k \text{ bolas azules en las primeras } n \text{ tiradas}) = \binom{n}{k}\frac{k!(n-k)!}{(n+1)!} = \frac{1}{n+1}$$

Por tanto $X_\infty = \lim_{n\to\infty} X_n$ es uniformemente distribuida en $[0,1]$. Si los parámetros son $a = 2, M = 1, c = 1$, entonces,

$$\mathbb{P}(\text{sacar } k \text{ bolas azules en las primeras } n \text{ tiradas}) = \binom{n}{k}\frac{(k+1)!(n-k)!}{3\cdot 4\ldots(n+2)} = \frac{2(k+1)}{(n+1)(n+2)}$$

Sea k tal que $k/n \to x$ cuando $n \to \infty$ donde $x \in [0,1]$. Entonces,

$$\frac{2(k+1)}{(n+1)(n+2)} \overset{n\to\infty}{\longrightarrow} 2x\frac{1}{n+1}$$

Esto nos dice que X_∞ tiene densidad $2x$.

5. Preliminares y Ayudantías

En esta sección aparte, vamos a introducir algunas definiciones importantes para comprender bien el estudio de la teoría de probabilidad.

Espacio de Medida

Vamos a definir un espacio de medida y algunas otras cosas, también repasaremos algunos resultados importantes de medida.

Definición 1 (Evento y σ-Álgebra): *Sea Ω un conjunto no vacío. Sea $\mathcal{F} \subseteq \mathcal{P}(\Omega)$. Decimos que $\mathcal{F}$ es un σ-álgebra de Ω, si cumple las siguientes propiedades:*

(a) $\emptyset \in \mathcal{F}$,

(b) Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$,

(c) Sea $\{A_n\}_{n\in\mathbb{N}}$ colección numerable de elementos de $\mathcal{F}$, entonces,

$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$$

En este contexto diremos que los elementos de $\mathcal{F}$ son los eventos de Ω y al par $(\Omega, \mathcal{F})$ le decimos pesacio medible.

Definición 2 (Medida y Medida de Probabilidad): *Sea $(\Omega, \mathcal{F})$ un espacio medible. Definimos una medida sobre el espacio medible $(\Omega, \mathcal{F})$ por una función $\mu : \mathcal{F} \to [0, \infty]$ que cumple las siguientes propiedades:*

(a) $\mu(\emptyset)$,

(b) Si $A, B \in \mathcal{F}$ tales que $A \subseteq B$, entonces $\mu(A) \leq \mu(B)$,

(c) Si $\{A_n\}_{n\in\mathbb{N}}$ es una colección de eventos disjuntos, entonces,

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$$

Decimos que es una medida de probabilidad si $\mu(\Omega) = 1$. En tal caso denotaremos $\mathbb{P}$ en vez de μ y diremos que $(\Omega, \mathcal{F}, \mathbb{P})$ es un espacio de probabilidad.

Veamos una propiedad fundamental de espacio de probabilidad.

Lema 3: *Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sea $\{A_n\}_{n\in\mathbb{N}}$ una colección de eventos. Entonces,*

(a) si $E_n \subseteq E_{n+1}$, entonces,

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n)$$

(b) si $E_n \supseteq E_{n+1}$*, entonces,*

$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n)$$

Nota 4: El lema anterior se puede enunciar para espacios de medidad, la única diferencia es que la condición (b) se pide que $\mu(E_1) < \infty$, el resto es igual.

El σ-álgebra es una colección bastante fuerte, pero existen nociones más débiles, donde además podemos definir una medidad. Las cuales se puede relacionar con la medida sobre un σ-álgebra, permitiendo una variedad de resultados útiles.

Definición 4 (Álgebra y Medida de un Álgebra): *Sea* Ω *un conjunto no vacío. Sea* $\mathcal{A} \subseteq \mathcal{P}(\Omega)$*. Diremos que* $\mathcal{A}$ *es un álgebra de* Ω*, si cumple las siguientes propiedades:*

(a) $\emptyset \in \mathcal{A}$*,*

(b) Si $A \in \mathcal{A}$*, entonces* $A^c \in \mathcal{A}$*,*

(c) Si $A_1, \ldots, A_n \in \mathcal{A}$*, entonces,*

$$\bigcup_{i=1}^{n} A_i \in \mathcal{A}$$

Sea $\mu : \mathcal{A} \to [0, \infty]$ *una función. Diremos que es una medida sobre* $\mathcal{A}$ *un álgebra si cumple las siguientes propiedades:*

(a) $\mu(\emptyset) = 0$*,*

(b) Si $A, B \in \mathcal{A}$ *tales que* $A \subseteq B$*, entocnes,*

$$\mu(A) \leq \mu(B)$$

(c) Si $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ *es una colección disjunta tal que,*

$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$$

Entonces,

$$\mu\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} \mu(A_n)$$

La noción de álgebra es más débil, pero con esto podemos deducir propiedades muy interesantes, pero antes de ver algunos de estos veamos otras nociones similares, donde una es más debil que un álgebra.

Definición 5 (Clase Monótona): *Sea* Ω *un conjunto no vacío, y sea* $\mathcal{M} \subseteq \mathcal{P}(\Omega)$*. Decimos que* $\mathcal{M}$ *es una clase monótona si,* *(a) dada una colección* $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}$ *tal que* $A_n \subseteq A_{n+1}$ *se cumple que:*

$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}$$

(b) dada una colección $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}$ *tal que* $A_n \supseteq A_{n+1}$ *se cumple que:*

$$\bigcap_{n\in\mathbb{N}} A_n \in \mathcal{M}$$

Definición 6 (Semi-Álgebra): *Sea* Ω *un conjunto no vacío. Sea* $\mathcal{S} \subseteq \mathcal{P}(\Omega)$. *Diremos que* $\mathcal{S}$ *es un semi-álgebra si cumple las siguientes propiedades:*

(a) $\emptyset \in \mathcal{S}$,

(b) Sea $\{S_n\}_{n=1}^{l}$ *una colección de conjuntos de la colección* $\mathcal{S}$ *tal que,*

$$\bigcap_{n=1}^{l} S_n \in \mathcal{S}$$

(c) Si $S \in \mathcal{S}$, *entonces* S^c *es unión finita disjunta de conjuntos de la colección* $\mathcal{S}$.

Un resultado no tan importante para nuesto estudio pero que se necesita conocer, es que un semi-álgebra puede construir un álgebra. Sea $(\Omega, \mathcal{S})$ un semi-álgebra. Entonces el conjunto:

$$\overline{S} := \{\text{unión disjunta de elementos de la colección } \mathcal{S}\}$$

es un álgebra. A este álgebra se le llama álgebra generado por $\mathcal{S}$.

Ahora lo realmente importante es que un espacio de medida sobre un álgebra $(\Omega, \mathcal{A}, \mu)$ puede ser extendida, permitiendo construir medidad más fuertes como una medida sobre un σ-álgebra. Este teorema conocido como **extensión de Caratheodory** relaciona la medida exterior μ^* inducida por μ, y la restringida de μ^* sobre los μ^*-medibles, de forma que estas tres componentes se conectan y construyen una medida sobre un σ-álgebra. Pero antes de ver el teorema recordemos que es la medida exterior inducidad sobre una medida y el conjunto de los μ^*-medibles.

Definición 7 (Medida Exterior Inducida por una Medida): *Sea* $(\Omega, \mathcal{A}, \mu)$ *un espacio de medida sobre un álgebra* $\mathcal{A}$. *Definimos la medida exterior inducida por* μ *como la función* $\mu^* : \mathcal{A} \to [0, \infty]$ *dada por:*

$$\mu^*(A) := \text{ínf}\left\{\sum_{n\in\mathbb{N}} \mu(A_n) : A \subseteq \bigcup_{n\in\mathbb{N}} A_n, \{A_n\} \subseteq \mathcal{A}\right\}$$

Nota 1.8: La medida exterior inducida usualmente se define sobre un pavimento y una premedida, pero no entraremos en detalles ya que no son relevantes en el repaso. También podemos definir la medida exterior de forma independiente (no inducida).

Definición 9 (Medida Exterior): *Sea* Ω *un conjunto no vacío. Sea* $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ *es una medida exterior si cumple las siguientes propiedades:*

(a) $\mu^*(\emptyset) = 0$,

(b) Si $A \subseteq B$, *entonces* $\mu^*(A) \leq \mu^*(B)$,

(c) Sea $\{A_n\}_{n\in\mathbb{N}}$ *una colección, entnces se cumple,*

$$\mu^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$$

Observación 10: Es evidente que la medida exterior μ^* inducida por una medida μ sobre un $\mathcal{A}$ está bien definida como medida exterior.

Definición 11 (μ^*-Medibles): *Sea* Ω *un conjunto no vacío y sea* μ^* *una medida exterior sobre* $\mathcal{P}(\Omega)$. *Decimos que* $E \subseteq \Omega$ *es* μ^**-medible si para todo* $A \subseteq \Omega$ *se cumple que:*

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

Observación 12: Notemos que no definimos μ-medible sobre medida sobre un σ-álgebra o un álgebra por que esta propiedad se cumple por definición (pensamos en E como elementos de la σ-álgebra o de la álgebra), por lo que no tiene mucho sentido mencionarlo. Lo que es importante es que el conjunto de los μ^*-medible es un σ-álgebra.

Observación 12: Sea μ^* una medida exterior sobre $\mathcal{P}(\Omega)$ y sea $\mathcal{F}$ la colección de los μ^*-medibles, entonces μ^* restringido a $\mathcal{F}$ es una medida.

Nota 13: Sea A un conjunto cualquier, definimos el menor σ-álgebra de la siguiente forma:

$$\sigma(A) = \bigcap_{\substack{\mathcal{F}\ \sigma\text{-álgebra} \\ A\subseteq\mathcal{F}}} \mathcal{F}$$

De forma análoga podemos definir la menor clase monotona que contiene a A por:

$$\mathcal{M}(A) = \bigcap_{\substack{\mathcal{M}'\ \text{clase monótona} \\ A\subseteq\mathcal{M}'}} \mathcal{M}'$$

Antes de pasar al teorema de extensión de Caratheodory, necesitamos saber sobre el teorema de las clases monótonas.

Teorema 14 (de las Clases Monótona): *Sea* Ω *un conjunto no vacío. Sea* $\mathcal{A}$ *un álgebra sobre* Ω, *entonces,*

$$\mathcal{M}(\mathcal{A}) = \sigma(\mathcal{A})$$

Teorema 12 (Caratheodory): *Sea* $(\Omega, \mathcal{A})$ *un álgebra. Sea* μ *una medida sobre el álgebra* $\mathcal{A}$, *sea* μ^* *la medida exterior asociada a* μ *y sea* $\overline{\mu}$ *la medida* μ^* *restringida a* $\mathcal{F}$ *donde* $\mathcal{F}$ *es el conjunto de todos los* μ^**-medibles. Entonces se cumple:*

(a) $\mathcal{F}$ *es un* σ*-álgebra que contiene a* $\mathcal{A}$*,*

(b) $\overline{\mu}$ *es una medida que extiende* μ *a todo* $\sigma(\mathcal{A})$*,*

(c) Si $\mu^*(E) = 0$*, entonces* $E \in \mathcal{F}$*,*

(d) Y si existe una colección $\{\Omega_n\}_{n\in\mathbb{N}}$ *de elementos de* $\mathcal{A}$ *tales que,*

$$\Omega = \bigcup_{n\in\mathbb{N}} \Omega_n$$

y $\mu(\Omega_n) < \infty$ *para todo* $n \geq 1$*. Entonces* $\overline{\mu}$ *es una única extensión de* μ *sobre* $\sigma(\mathcal{A})$*.*

A veces nos encontraremos con espacios de medida donde Ω tiene un comportamiento convenient.

Definición 13: *Sea* Ω *un conjunto no vacío y consideremos* $\mathcal{F}$ *un* σ*-álgebra y* $\mathcal{A}$ *un álgebra de* Ω*. Si* μ *es una medida sobre* $(\Omega, \mathcal{F})$*,*

(a) Decimos que μ *es finito si* $\mu(\Omega) < \infty$*,*

(b) Decimos que μ *es* σ*-finita si existe* $\{\Omega_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F}$ *tal que,*

$$\bigcup_{n=1}^{\infty} \Omega_n = \Omega, \;\; \mu(\Omega_n) < \infty$$

para todo $n \in \mathbb{N}$*.*

Si ν *es una medida sobre* $(\Omega, \mathcal{A})$*, decimos que es fuertemente* σ*-finito si existe* $\{\Omega_n\}_{n\in\mathbb{N}} \subseteq \mathcal{A}$ *tal que,*

$$\bigcup_{n=1}^{\infty} \Omega_n = \Omega, \;\; \nu(\Omega_n) < \infty$$

para todo $n \in \mathbb{N}$

Ejemplo 14: Sea $\Omega = \mathbb{R}, \mathcal{F} = \mathcal{B}(\mathbb{R})$ y consideremos la medida λ de Lebesgue. Sea $\Omega_n := [-n, n]$ para $n \geq 1$. Entonces,

$$\bigcup_{n=1}^{\infty} \Omega_n = \mathbb{R}$$

Claramente $\lambda(\Omega_n) = 2n < \infty$ para todo $n \in \mathbb{N}$, por tanto λ es σ-finita.

arreglar

Definimos la medida de Lebesgue por $\lambda : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ mediante la extesión de Caratheodory, y estas nos dice que es única. Ahora podemos definir la medida de Lebesgue para mayores deimsnones. Sea $d \geq 1$, decimos que $I \subseteq \mathbb{R}^d$ es un intervalo si es de la forma:

$$I = I_1 \times I_2 \times \cdots \times I_d$$

Luego podemos definir una premedida sobre los intervalos dado por:

$$\tau(I) = \prod_{i=1}^{d} |I_i|$$

Finalmente podemos concluir que $\mathcal{B}(\mathbb{R}^d) \subseteq \mathcal{L}_d$, finalmene obtenemos la medida de Lebesgue dada por:

$$\lambda : \mathcal{B}(\mathbb{R}^d) \to [0, \infty]$$
$$\lambda(A) = \text{ínf}\left\{ \sum_{i=1}^{\infty} \tau(I_i) : A \bigcup_{i=1}^{\infty} I_i \right\}$$

Funciones Medibles

Definiremos las funciones medibles y estudiaremos propiedades importantes de esta, además veremos convergencias asociado con funciones medibles.

Definición 1: *Sean $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$ espacios medibles y sea $f : \Omega_1 \to \Omega_2$ una función. Decimos que f es $(\mathcal{F}_1, \mathcal{F}_2)$-medible si para todo $B \in \mathcal{F}_2$ se tiene que $f^{-1}(B) \in \mathcal{F}_1$. Sea $(\Omega, \mathcal{F})$ espacio medible. Decimos que $f : \Omega \to \mathbb{R}^d$ es medible si para todo $A \in \mathcal{B}(\mathbb{R}^d)$ se tiene que $f^{-1}(A) \in \mathcal{F}$.*

Aun así ver que una función es medible puede llegar a ser complicado, por lo que necesitamos un resultado que sirva ver si una función es medible.

Lema 2: *Sea $(\Omega, \mathcal{F})$ un espacio medible y sea $f : \Omega \to \mathbb{R}$ una función. Entonces las siguientes afirmaciones son equivalentes:*

(a) f *es medible,*

(b) $f^{-1}(O) \in \mathcal{F}$ *para todo* $O \subseteq \mathbb{R}$ *abierto,*

(c) $f^{-1}(a, b) \in \mathcal{F}$ *para todo* $a < b$*,*

(d) $f^{-1}(a, \infty) \in \mathcal{F}$ *para todo* $a \in \mathbb{R}$*,*

(e) $f^{-1}(a, b] \in \mathcal{F}$ *para todo* $a < b$*.*

Lema 3: *Sea $(\Omega, \mathcal{F})$ espacio medible. Sean $f, g : \Omega \to \mathbb{R}$ medibles y $c \in \mathbb{R}$, entonces,*

(a) $cf, f + g, fg$ *son medibles,*

(b) $|f|$ *es medible,*

(c) $\max\{f, g\}, \min\{f, g\}$ *son medibles.*

Lema 4: *Sea $(\Omega, \mathcal{F})$ espacio medible y sean $f_n : \Omega \to \mathbb{R}$ medible para todo $n \geq 1$. Entonces,*

$$\sup_{n\geq 1} f_n, \quad \inf_{n\geq 1} f_n, \quad \limsup_{n\to\infty} f_n, \quad \limsup_{n\to\infty} f_n$$

son medible.

Corolario 5: *Sea $(\Omega, \mathcal{F})$ espacio medible. Sean $f_n, f : \Omega \to \overline{\mathbb{R}}$ funciones donde f_n es medible para todo $n \geq 1$. Si,*

$$\lim_{n\to\infty} f_n(\omega) = f(\omega)$$

para todo $\omega \in \Omega$. Entonces f es medible.

Definición 6: *Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida. Sean $f_n, f : \Omega \to \overline{\mathbb{R}}$ funciones. Decimos que f_n converge a f μ-casi en todas parte si existe $\Omega_0 \subseteq \Omega$ tal que,*

$$\lim_{n\to\infty} f_n(\omega) = f(\omega)$$

para todo $\omega \in \Omega_0$ y Ω_0^c es despreciable.

La condición μ-ctp nos ayuda bastante para determinar el comportamiento de cosas, ya que basta estuidar una parte reducida donde el complemento es despreciable para deducir propiedades importantes.

Lema 7: *Sea* $(\Omega, \mathcal{F}, \mu)$ *un espacio de medida completo. Sean,*

$$f_n, b : \Omega \to \overline{\mathbb{R}}$$

medibles tales que $f_n \to f$ *de forma* μ*-ctp. Entonces* f *es medible.*

Existe otro tipo de convergencia, que es en medida.

Definición 8: *Sea* $(\Omega, \mathcal{F}, \mu)$ *un espacio de medida. Sean* $f_n, f : \Omega \to \overline{\mathbb{R}}$ *medibles. Decimos que* f_n *converge a* f *en medida si para todo* $\delta > 0$*, se tiene que,*

$$\lim_{n \to infty} \mu(\{\omega \in \Omega : |f_n(\omega) - f(\omega)| > \delta\}) = 0$$

Lema 9: *Sea* $(\Omega, \mathcal{F}, \mu)$ *un espacio de medida. Supongamos que* $\mu(\Omega) < \infty$*. Luego si* $f_n \longrightarrow f$ *de forma* μ*-ctp. Entonces lo hace en* μ*-medida.*

Proposición 10: *Sea* $(\Omega, \mathcal{F}, \mu)$ *un espacio de medida. Sean* $f_n, f : \Omega \to \overline{\mathbb{R}}$ *funciones medibles. Si* $f_n \longrightarrow f$ *en medida. Entonces existe una subsucesión* f_{n_k} *que converge a* f *en* μ*-ctp.*

Lema 11: *Sea* $(\Omega, \mathcal{F}, \mu)$ *un espacio de medida. Sea* $\{E_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F}$*. Si,*

$$\sum_{n \in \mathbb{N}} \mu(E_n) < \infty$$

Entonces,

$$\mu\left(\limsup_{n \to \infty} E_n\right) = 0$$

Los Tres Princiíos de Littlewood

Definición 1: *Sea (X, τ) un espacio topológico. Definimos G_δ por los conjuntos que son intersección numerable de abiertos. De forma análoga definimos F_δ por los conjuntos que son intersección numerable de cerrados.*

Primer Principio: *Sea $E \subseteq \mathbb{R}$ un*

La integral de Lebesgue

Vamos a repasar el concepto de la integral de Lebesgue sabiendo propiedades de espacio de medida. También repasaremos algunos resultado sin demostrarlos.

Consideremos un espacio de medida $(\Omega, \mathcal{F}, \mu)$. Sea $s : \Omega \to [0, \infty)$ una función simple, es decir,

$$s(x) = \sum_{i=1}^{n} a_i \mathbb{1}_{A_i}(x)$$

donde $a_i \geq 0$ y $A_i \in \mathcal{F}$ para todo $i = 1, \ldots, n$. Definimos la integral de Lebesgue sobre la función simple s por:

$$\int_{\Omega} s(x) d(\mu(x)) = \int s d\mu := \sum_{i=1}^{n} a_i \mu(A_i)$$

Definimos la integral de s sobre un elemento $E \in \mathcal{F}$ por:

$$\int_E s d\mu = \int s \mathbb{1}_E d\mu$$

Observación 1: Claramente s es finita por cosntrucción y por otro lado, la integral,

$$\int s d\mu \in [0, \infty]$$

al poder expresarse como suma de medidas.

Nota 2: En general calculamos la integral sobre funciones medibles, que más adeltante repasaremos. Por otro lado podemos ver que s es medible al ser combinación lineal de funciones medibles, (que son de la forma $\mathbb{1}_A$).

Ejemplo 3: Sea $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ un espacio de medida, entonces,

$$\begin{aligned}
\int_{\mathbb{R}} \mathbb{1}_{\mathbb{Q}} d\lambda &= \lambda(\mathbb{Q}) \\
&= \lambda\left(\bigcup_{q \in \mathbb{Q}} \{q\}\right) \\
&= \sum_{q \in \mathbb{Q}} \lambda(\{q\}) \\
&= 0
\end{aligned}$$

Proposición 4: *Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida y sean $s, t : \Omega \to [0, \infty)$ funciones simples. Entonces se cumple:*

(a) ***Linealidad:*** *Si $\alpha, \beta \geq 0$ son constantes, entonces,*

$$\int (\alpha s + \beta t) d\mu = \alpha \int s d\mu + \beta \int t d\mu$$

(b) ***Monotonia:*** *Si $s(x) \leq t(x)$ para todo $x \in \Omega$, entonces,*

$$\int s d\mu \leq \int t d\mu$$

Proposición 5: *Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida y sea $s : \Omega \to [0, \infty)$* ***falta ¿se puede como f medible en general?***

Definido la integral sobre funciones simples positivas, podemos pasar a generalizar la integral de Lebesgue sobre funciones positivas medibles.

Sea $f : \Omega \to [0, \infty]$ una función medible y sea $\{s_n\}_{n \in \mathbb{N}}$ funciones simples no negativas tales que $s_n \uparrow f$ (s_n son crecientes y convergen a f puntualmente), entonces se define la integral de Lebesgue de f por:

$$\int df\mu = \lim_{n \to \infty} \int s_n d\mu$$

Observación 6: Notar que la integral de f está bien definida, dado que es independiente de la elección de $\{s_n\}_{n \in \mathbb{N}}$ (es decir, si consideramos dos sucesiones s_n, s'_n, se tiene que el límite de las integrales de las funciones positivas, son iguales).

Con la definición de la integral sobre una función f medible no negativa (o positiva) se puede caracterizar de la siguiente forma:

$$\int f d\mu = \sup \left\{ \int s d\mu : 0 \leq s \leq f, \ s \text{ simple no negativa} \right\}$$

Proposición 7: *Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida. Sean $f, g : \Omega \to [0, \infty]$ funciones medibles, entonces:*

(a) ***Linealidad:*** *Para $\alpha, \beta \geq 0$ se cumple que:*

$$\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu$$

(b) ***Monotonía de Funciones:*** *Si $f(x) \leq g(x)$ para todo $x \in \Omega$, entonces,*

$$\int f d\mu \leq \int g d\mu$$

(c) ***Monotonía de Medibles:*** *Sean $A, B \in \mathcal{F}$ tales que $A \subseteq B$, entonces,*

$$\int_A f d\mu \leq \int_B f d\mu$$

(d) ***Función Nula μ-ctp:*** *Se cumple que $f = 0$ de forma μ-ctp si y sólo si,*

$$\int f d\mu = 0$$

(e) ***Nulanidad de Medible:*** *Si $\mu(A) = 0$, entonces,*

$$\int_A f d\mu = 0$$

Calcular la integral de Lebesgue puede resultar muy complicado, ya que la noción de medida es muy abstracto, por lo que existen teoremas de convergencia que permite calcular mediante convergencia algunas integrales, sin embargo, solo estudiaremos un teorema de convergencia, más adelante veremos otro, haciendo un total de dos.

Teorema 8 (de Convergencia Monótona): *Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida. Sean $f_n : \Omega \to [0, \infty]$ funciones medibles tales que $f_n(\omega) \leq f_{n+1}(\omega)$, entonces,*

$$\lim_{n\to\infty} \int f_n d\mu = \int \lim_{n\to\infty} f_n d\mu$$

Teorema 9 (Fatou): *Sea $(\Omega, \mathcal{F}, \mu)$ espacio de medida. Sean $f_n : \Omega \to [0, \infty]$ funciones medibles, entonces,*

$$\int \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int f_n d\mu$$

Lema 10: *Sea $(\Omega, \mathcal{F}, \mu)$ espacio de medida. Sean $f_n : \Omega \to [0, \infty]$ funciones medibles, entonces,*

$$\int \sum_{n=1}^{\infty} f_n d\mu = \sum_{n=1}^{\infty} f_n d\mu$$

Finalmente vamos a definir la integral de Lebesgue sobre una función medible de Ω a $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\}$. El tema es que queremos trabajar sobre funciones no negativas, para ello necesitamos separar f en dos partes.

Observación 11: Si definimos $f^+ := \max\{f, 0\}$, $f^- := \max\{-f, 0\}$, entonces se cumple:

(a) $f = f^+ - f^-$,

(b) $|f| = f^+ f^-$,

(c) Si f es medible, entonces f^+, f^- son medibles.

Por tanto, consideremos un espacio de medida $(\Omega, \mathcal{F}, \mu)$ y una función $f : \Omega \to \overline{R}$ medible. Decimos que f es Lebesgue-integrable (L-integrable) si,

$$\int |f| d\mu < \infty$$

Aquí podemos definir el conjunto de las funciones L-integrables sobre Ω por:

$$L^1(\Omega, \mathcal{F}, \mu) = L^1(\Omega) := \{g : \Omega \to \overline{\mathbb{R}} : g \text{ es L-integrable}\}$$

Finalmente, si $f \in L^1(\Omega)$, calculamos su integral de la siguiente forma:

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu$$

Lema 12: *Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida, sean $f, g \in L^1(\Omega)$ y $\alpha, \beta \in \mathbb{R}$, entonces se cumple que:*

$$\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu$$

Ahora veremos el segundo teorema de convergencia, el cual dicta lo siguiente:

Teorema 13 (de Convergencia Dominada): *Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida y sean $f_n, f : \Omega \to \overline{\mathbb{R}}$ funciones medibles tales que,*

(a) $\lim\limits_{n\to\infty} f_n = f$ *de forma μ-ctp,*

(b) Existe un $g \in L^1(\Omega)$ tal que $|f_n| \leq g$ para todo $n \geq 1$ de forma μ-ctp.

Entonces se cumple que:

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu$$

Ahora estudiaremos una relación importange entre las funciones R-integrables con funciones L-integrable, viendo que son muy parecidos bajo la medida de Lebesgue.

Teorema 14: *Sea $f : [a, b] \to \mathbb{R}$ una función R-integrable, entonces en el espacio de medida $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$, se tiene que,*

(a) f es Lebesgue medible,

(b) $f \in L^1(\mathbb{R})$,

(c) Por último,

$$\int_{[a,b]} f d\lambda = \int_a^b f(x)dx$$

Veamos un ejemplo que aplica los teoremas visto anteriormente.

Ejemplo 15: Sea $f(x) = 1/1+x^2$ definido en $[0, \infty)$. Vamos a demostrar que $f \in L^1([0, \infty), \mathcal{B}(\mathbb{R}), \lambda)$. Queremos demostrar por el teormea de la convergencia monótona, por lo que debemos encontrar funciones f_n medibles crecientes. Consideremos,

$$f_n(x) = f(x)\mathbb{1}_{[0,n]}(x)$$

Claramente f_n es medible puesto que f es medible y la indicatriz es siempre medible, también observemos que $f_n \uparrow f$ y que $f \geq 0$, por tanto,

$$\begin{aligned}\int_{[0,\infty)} f d\lambda &= \int_{[0,\infty)} \lim_{n\to\infty} f_n d\lambda \\ &= \lim_{n\to\infty} \int_{[0,\infty)} f_n d\lambda \\ &= \lim_{n\to\infty} \int_{[0,n]} \frac{1}{1+x^2} d\lambda\end{aligned}$$

Notemos que $f_n : [0,n] \to \mathbb{R}$ es una función R-integrable al ser continua por tanto podemos pensarlo como una integral de Riemann, de forma que,

$$\begin{aligned}\int_{[0,\infty)} f d\lambda &= \lim_{n\to\infty} \int_0^n \frac{dx}{1+x^2} \\ &= \lim_{n\to\infty} \arctan(n) \\ &= \frac{\pi}{2} < \infty\end{aligned}$$

Como queriamos probar.

Existe otro resultado que relaciona de otra froma la integral de Lebesgue con la de Riemann pero con integrales impropias.

Teorema 16: *Consideremos el espacio medible $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$. Sea $f \in L^1(\mathbb{R})$ tal que f es R-integrable en el intervalo $[a,b]$ donde $a \leq 0, b \geq 0$ de forma que el límite,*

$$\lim_{\substack{a\to-\infty \\ b\to\infty}} \int_a^b f(x)dx$$

existe, entonces,

$$\int_{\mathbb{R}} f d\lambda = \lim_{\substack{a\to-\infty \\ b\to\infty}} \int_a^b f(x)dx = \int_{-\infty}^{\infty} f(x)dx$$

Por otro lado, si $f \in L^1(c,d)$ tal que f es R-integrable sobre todo interbalo $[c', d'] \subset (c,d)$ y tal que el límite,

$$\lim_{\substack{c'\to c \\ d'\to d}} \int_{c'}^{d'} f(x)dx$$

existe, entonces,

$$\int_{(c,d)} f d\lambda = \lim_{\substack{c'\to c \\ d'\to d}} \int_{c'}^{d'} f(x)dx = \int_c^d f(x)dx$$

Para cerrar las sección introductoria de la integral de Lebesgue, vamos a dar ejemplos cinco fundamentales para aplicar todo lo visto en esta sección.

Ejemplo 17: Determinemos la integral:

$$\lim_{n\to\infty} \int_{\mathbb{R}} \frac{n\operatorname{sen}(x/n)}{x(1+x^2)} d\lambda$$

Notemos que la función adentro de la integral no es no negativa, a tales funciones la denotaremos por f_n, por lo que queremos usar el teorema de la convergencia dominada, es decir, debemos probar que existe una función $g \in L^1(\mathbb{R})$ que acote laas funciones f_n para todo $n \geq 1$ de forma λ-ctp y que exista una función f donde f_n converge puntualemnte.

Recordemos que $|\operatorname{sen}(a)| \leq |a|$ para todo $a \in \mathbb{R}$, entonces,

$$|f_n| = \left|\frac{n\operatorname{sen}(x/n)}{x(1+x^2)}\right| \leq \frac{1}{1+x^2} \in L^1([0,\infty))$$

Nuestro problema es que queremos algo más fuerte, y lo hay, se puede comprobar que $1/(1+x^2) \in L^1(\mathbb{R})$ y para ello basta considerar $g_n(x) = g(x)\mathbb{1}_{[-n,n]}$, por tanto $f_n \in L^1(\mathbb{R})$ para todo $n \geq 1$ de forma λ-ctp (notar que f_n no está definido en $x \neq 0$, pero es claramente λ-despreciable). Por otro lado,

$$\begin{aligned}\lim_{n\to\infty} f_n(x) &= \lim_{n\to\infty} \frac{\operatorname{sen}(x/n)}{x/n}\frac{1}{1+x^2} \\ &= \frac{1}{1+x^2}\end{aligned}$$

Por tanto por el teorema de la convergencia dominada se tiene que,

$$\begin{aligned}\lim_{n\to\infty} \int_{\mathbb{R}} \frac{n\operatorname{sen}(x/n)}{x(1+x^2)} d\lambda &= \int_{\mathbb{R}} \frac{1}{1+x^2} d\lambda \\ &= \int_{-\infty}^{\infty} \frac{1}{1+x^2} = \pi\end{aligned}$$

Ejemplo 16: Determinemos el valor de,

$$\lim_{n\to\infty} \int_0^1 \frac{1+nx}{(1+x^2)^n} dx$$

Notemos que queremos estudiar bajo la integral de Riemann, para ello pensemos en espacio medida, sea,

$$f_n(x) := \frac{1+nx}{(1+x^2)^n}$$

Vemos que f_n es continua y por tanto es R-integrable. Notemos también que,

$$(1+x^2)^n \geq 1 + nx^2$$

Entonces $|f_n(x)| \leq 1 \in L^1[0,1]$, por otro lado,

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0, & x \neq 0 \\ 1, & x = 0 \end{cases} = \mathbb{1}_{\{0\}}$$

Luego por el teorema de la convergencia dominada se tiene que,

$$\begin{aligned} \lim_{n\to\infty} \int_{[0,1]} f_n d\lambda &= \int_{[0,1]} \lim_{n\to\infty} f_n d\lambda \\ &= \int_{[0,1]} \mathbb{1}_{\{0\}} d\lambda = 0 \end{aligned}$$

Ahora como f_n es R-integrable, se cumple que,

$$\int_{[0,1]} f_n d\lambda = \int_0^1 f_n(x) dx$$

Y por tanto podemos relacionas ambas integrables y concluir que,

$$\lim_{n\to\infty} \int_0^1 \frac{1+nx}{(1+x^2)^n} dx = 0$$

...

Medida Producto

Vamos a estudiar el producto de dos σ-álgebras. Consideremos dos espacios de medida $(\Omega_1, \mathcal{F}_1, \mu_1), (\Omega_2, \mathcal{F}_2, \mu$
Ahora necesitamos considerar los rectángulos conformados por los elementos de los σ-álgebras $\mathcal{F}_1, \mathcal{F}_2$ dada por:

$$\mathcal{R} := \{A \times B : A \in \mathcal{F}_1, B \in \mathcal{F}_2\} \subseteq \mathcal{F}_1 \times \mathcal{F}_2$$

Ahora tomaremos el menor σ-álgebra que contiene a $\mathcal{R}$, es decir, trabajaremos con.

$$\mathcal{F}_1 \otimes \mathcal{F}_2 := \sigma(\mathcal{R})$$

De esta forma construimo el espacio medible $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$. Ahora queremos definir una medida sobre este nuevo espacio medible, pero antes necesitamos algunas herramientas.

Definición 1 (Secciones): *Sea $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ espacio medible. Sea $E \in \mathcal{F}_1 \otimes \mathcal{F}_2$ y definimos,*

$$\begin{aligned} E^1(y) &:= \{x \in \Omega_1 : (x, y) \in E\} \subseteq \Omega_1 \\ E^2(x) &:= \{y \in \Omega_1 : (x, y) \in E\} \subseteq \Omega_2 \end{aligned}$$

Estos "filamentos" son buenos debido a que son medibles bajo sus respectivos σ-álgebra, es decir, $E^1(y) \in \mathcal{F}_1$ para todo y y análogamente con $E^2(x)$.

Por otro lado existe una propiedad similar sobre las funciones.

Proposición 2: *Sea $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ un espacio medible. Sea $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ una función $\mathcal{F}_1 \otimes \mathcal{F}_2$-medible. Entonces,*

(a) $y \mapsto f(x, y)$ es $\mathcal{F}_2$-medible para todo $x \in \Omega_1$,

(b) $x \mapsto f(x, y)$ es $\mathcal{F}_2$-medible para todo $y \in \Omega_2$.

Aquí también podemos construir $\mathcal{F}_1 \otimes \mathcal{F}_2$ a partir de un álgebra.

Proposición 3: *Consideremos el conjunto cartesiano $\Omega_1 \times \Omega_2$ con Ω_1, Ω_2 no vacíos. Sea $\mathcal{A}$ la colección de todas las uniones finitas de elementos de $\mathcal{R}$, entonces,*

(a) $\mathcal{A}$ es un álgebra,

(b) $\sigma(\mathcal{A}) = \mathcal{F}_1 \otimes \mathcal{F}_2$.

Ahora, para definir una medida existen varios caminos, pero una forma que usaremos, es considerar un espacio medible $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ determinada por los espacios de medida $(\Omega_1, \mathcal{F}_1, \mu_1), (\Omega_2, \mathcal{F}_2, \mu_2)$ y considerar una función $\mu : \mathcal{F}_1 \otimes \mathcal{F}_2 \to [0, \infty]$ tal que,

$$\mu(A \times B) = \mu_1(A)\mu_2(B)$$

para todo $A \times B \in \mathcal{R}$. Luego ciertas propiedades de forma que sea una medida sobre $\mathcal{A}$ definida anteriormente y luego aplicar Caratheodory para obtener una medida en $\mathcal{F}_1 \otimes \mathcal{F}_2$.

De aquí podemos obtener una medida, pero no intentarmeos buscar una, solo daremos una función que satisface las condiciones privas, pero antes un resultado previo.

Lema 4: *Sean $(\Omega_1, \mathcal{F}_1, \mu_1), (\Omega_2, \mathcal{F}_2, \mu_2)$ espacios de medida. Supongamos que μ_1, μ_2 son σ-finitas, entonces,* *(a)* $y \mapsto \mu_1(E^1(y))$ *es* $\mathcal{F}_2$*-medible,*

(b) $x \mapsto \mu_2(E^2(x))$ *es* $\mathcal{F}_1$*-medible.*

Propocisión 5: *Sean* $(\Omega_1, \mathcal{F}_1, \mu_1), (\Omega_2, \mathcal{F}_2, \mu_2)$ *espacios de medida. Supongamos que* μ_1, μ_2 *son* σ*-finitas. Sea la función,*

$$\mu(E) := \int_{\Omega_2} \mu_1(E^1(y))\mu_2(dy)$$

para todo $E \in \mathcal{F}_1 \otimes \mathcal{F}_2$*. Entonces* μ *es una medida sobre* $\mathcal{F}_1 \otimes \mathcal{F}_2$ *tal que* $\mu(A \times B) = \mu_1(A)\mu_2(B)$ *para todo* $A \in \mathcal{F}_1, B \in \mathcal{F}_2$*.*

Observación 6: Podemos definir la medida,

$$\mu^*(E) := \int_{\Omega_1} \mu_2(E^1(x))\mu_2(dx)$$

y está cumple las implicancias de la propocisión 5, es más, $\mu^* = \mu$, de forma que podemos definir una medida buena en $\mathcal{F}_1 \otimes \mathcal{F}_2$.

Definición 7 (Medida Producto): *Sea* $(\Omega_1, \mathcal{F}_1, \mu_1), (\Omega_2, \mathcal{F}_2, \mu_2)$ *espacios de medida* σ*-finitas. Definimos la medida* $\mu_1 \otimes \mu_2$ *sobre el espacio medible* $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ *por:*

$$(\mu_1 \otimes \mu_2)(E) = \int_{\Omega_2} \mu_1(E^1(y))\mu_2(dy) = \int_{\Omega_1} \mu_2(E^2(x))\mu_1(dx)$$

El producto de espacios de medida se puede extender hasta n-ésimos espacios. Pero todas las propiedades se pueden entender estudiando el caso $n = 3$. Consideremos $(\Omega_i, \mathcal{F}_i, \mu_i)$ espacios de medida para $i = 1, 2, 3$. Luego se cumple que,

$$\mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3 = (\mathcal{F}_1 \times \mathcal{F}_2) \times \mathcal{F}_3 = \mathcal{F}_1 \times (\mathcal{F}_2 \times \mathcal{F}_3)$$

En general sin ninguna condición adicional. Por otro lado, definimos las medida producto igual que antes, y se cumple que,

$$\mu_1 \otimes \mu_2 \otimes \mu_3 = (\mu_1 \otimes \mu_2) \otimes \mu_3 = \mu_1 \otimes (\mu_2 \otimes \mu_3)$$

solamente si los espacios son σ-finito. Y así recursivamente.

Observación 8: Consideremos los espacios de medida $(\mathbb{R}, \mathcal{B}(\mathbb{R}^n), \lambda^n), (\mathbb{R}, \mathcal{B}(\mathbb{R}^m), \lambda^m)$, por lo que son σ-finitos y se cumple que,

$$\mathcal{B}(\mathbb{R}^n) \times \mathcal{B}(\mathbb{R}^m) = \mathcal{B}(\mathbb{R}^{n+m})$$
$$\lambda^n \otimes \lambda^n = \lambda^{n+m}$$

Uno de los resultado más importante de la medida producto, es poder trabajar la integral de una función medible real de dos variable, es integrarlo de forma independiente por cada variable, siendo esto, conocido como el teorema de Fubini-Tonelli.

Teorema 9 (Fubini-Tonelli): *Sean* $(\Omega_1, \mathcal{F}_1, \mu_1), (\Omega_2, \mathcal{F}_2, \mu_2)$ *espacios de medida* σ*-finitos. Sea* $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ *una función* $\mathcal{F}_1 \times \mathcal{F}_1$*-medible. Si las integrables:*

$$\begin{aligned} I &= \int_{\Omega_1 \times \Omega_2} f(x,y)(\mu_1 \otimes \mu_2)(dxdy) \\ I_1 &= \int_{\Omega_1} \left(\int_{\Omega_2} f(x,y)\mu_2(dy) \right) \mu_2(dx) \\ I_2 &= \int_{\Omega_2} \left(\int_{\Omega_1} f(x,y)\mu_1(dx) \right) \mu_2(dy) \end{aligned}$$

están bien definidas (toman valores $\overline{\mathbb{R}}$*), entonces,*

(a) ***Tonelli:*** *Si* f *es no negativa, entonces* $I = I_1 = I_2$*,*

(b) ***Fubini:*** *Si* $f \in L^1(\mu_1 \otimes \mu_2)$*, entonces* $I = I_1 = I_2$*.*

Observación 10: El resultado de Fubini es algo más complicado de determinar, pero tiene un resultado potente. Por otro lado Tonelli toma una hipótesis más fuerta y sencillo de estudiar, el cual su resultado es igual de potente que el de Fubini.

Ayudantía 1

P1: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y consideremos una secuecnias de eventos $\{A_n\} \subseteq \mathcal{F}$. Se define el límite superior de los eventos por:

$$\limsup_{n\to\infty} A_n := \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k$$

Y definimos el límite inferior de los eventos por:

$$\liminf_{n\to\infty} A_n := \bigcup_{n\in\mathbb{N}} \bigcap_{k\geq n} A_k$$

Entonces se cumple las siguientes propiedades:

(a) la siguiente inclusión:

$$\liminf_{n\to\infty} A_n \subseteq \limsup_{n\to\infty} A_n$$

(b) las siguientes desigualdades:

$$\mathbb{P}\left(\liminf_{n\to\infty} A_n\right) \leq \liminf_{n\to\infty} \mathbb{P}(A_n)$$
$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) \geq \limsup_{n\to\infty} \mathbb{P}(A_n)$$

(c) las siguientes identidades:

$$\mathbb{1}_{\liminf_{n\to\infty} A_n} = \liminf_{n\to\infty} \mathbb{1}_{A_n}$$
$$\mathbb{1}_{\limsup_{n\to\infty} A_n} = \limsup_{n\to\infty} \mathbb{1}_{A_n}$$

Sol: Por comodidad diremos denotaremos el límite superior y el límite inferior por:

$$A^+ = \limsup_{n\to\infty} A_n$$
$$A_- = \liminf_{n\to\infty} A_n$$

respectivamente.

(a) Por definición, sea $\omega \in A_-$, entonces existe un $n \in \mathbb{N}$ tal que para todo $k \geq n$ se tiene que $\omega \in A_k$. Sea $n \in \mathbb{N}$ arbitrario, podemos tomar k suficientemente grande de forma que $\omega \in A_k$ como $\omega \in A_-$, pero esto equivale a decir que para todo $n \in \mathbb{N}$ existe un $k \geq n$ tal que $\omega \in A_k$, que es equivalente a decir que $\omega \in A^+$, por tanto,

$$A_- \subseteq A^+$$

(b) Notemos lo siguiente con respecto al límite inferior de los A_n,

$$\bigcap_{k\geq 1} A_k \subseteq \bigcap_{k\geq 2} A_k \subseteq \cdots \subseteq \bigcap_{k\geq n} A_k \subseteq \dots$$

Es decir, $\left\{\bigcap_{k\geq n} A_k\right\}_{n\in\mathbb{N}}$ es una secuencia de eventos crecientes. Entonces por el lema 1 podemos concluir que,

$$\mathbb{P}\left(\liminf_{n\to\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcap_{k\geq n} A_k\right)$$

Ahora notemos que,

$$\mathbb{P}\left(\bigcap_{k\geq n} A_k\right) \leq \mathbb{P}(A_k)$$

para todo $k \geq n$, es decir, el lado derecho es una cota inferior de los $\mathbb{P}(A_k)$ sobre los k mayores o iguales de n, dicho de otra forma,

$$\mathbb{P}\left(\bigcap_{k\geq n} A_k\right) \leq \inf_{k\geq n} \mathbb{P}(A_k)$$

Finalmente se concluye,

$$\mathbb{P}\left(\liminf_{n\to\infty} A_n\right) \leq \lim_{n\to\infty} \inf_{k\geq n} \mathbb{P}(A_k) = \liminf_{n\to\infty} \mathbb{P}(A_n)$$

Probemos la desigualdad con respecto al límite superior. Notemos que la secuencia de eventos $\left\{\bigcup_{k\geq n} A_k\right\}_{n\in\mathbb{N}}$ es decreciente, donde $\mathbb{P}(A_1) < \infty$, entonces por lema 1 se tiene que,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}\left(\bigcup_{k\geq n} A_k\right)$$

Podemos ver que,

$$\mathbb{P}\left(\bigcup_{k\geq n} A_k\right) \geq \mathbb{P}(A_k)$$

para todo $k \geq n$, es decir, el lado izquierdo es cota superior de los $\mathbb{P}(A_k)$ para todo $k \geq n$, por tanto,

$$\mathbb{P}\left(\bigcup_{k\geq n} A_k\right) \geq \sup_{k\geq n} \mathbb{P}(A_k)$$

Finalmente obtenemos,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) \geq \lim_{n\to\infty} \sup_{k\geq n} \mathbb{P}(A_k) = \limsup_{n\to\infty} \mathbb{P}(A_n)$$

(c) Vamos a probarlo por inspección. Sea $\omega \in A_-$, entonces es claro que,

$$\mathbb{1}_{A_-}(\omega) = 1$$

Ahora si $\omega \in A_-$, entonces existe un $n \in \mathbb{N}$ tal que $\omega \in A_k$ para todo $k \geq n$, esto implica que $\mathbb{1}_{A_k}(\omega) = 1$ para todo $k \geq n$, luego,

$$1 = \lim_{n\to\infty} 1 = \lim_{n\to\infty} \inf_{k\geq n} \mathbb{1}_{A_k} = \liminf_{n\to\infty} \mathbb{1}_{A_n}$$

Por otro lado, si $\omega \notin A_-$, entonces,

$$\mathbb{1}_{A_-}(\omega) = 0$$

Además, para todo $n \in \mathbb{N}$, existe un $k \geq n$ tal que $\omega \notin A_k$, es decir,

$$\inf_{k\geq n} \mathbb{1}_{A_k} = 0$$

Tomando $n \to \infty$ obteniendo que,

$$0 = \lim_{n\to\infty} \inf_{k\geq n} \mathbb{1}_{A_k} = \liminf_{n\to\infty} \mathbb{1}_{A_n}$$

Para la otra identidad se procede de forma ánaloga. Sea $\omega \in A^+$ por lo que,

$$\mathbb{1}_{A^+}(\omega) = 1$$

Luego para todo $n \in \mathbb{N}$ existe un $k \geq n$ tal que $\omega \in A_k$,

$$\sup_{k\geq n} \mathbb{1}_{A_k}(\omega) = 1$$

Aplicando $n \to \infty$ llegamos a que,

$$1 = \lim_{n\to\infty} \sup_{k\geq n} \mathbb{1}_{A_k} = \limsup_{n\to\infty} \mathbb{1}_{A_n}$$

Ahora consideremos $\omega \notin A^+$ por lo que,

$$\mathbb{1}_{A^+}(\omega) = 0$$

Además, existe un $n \in \mathbb{N}$ tal que para todo $k \geq n$ se tiene que $\omega \notin A_k$, y esto implica que,

$$\sup_{k\geq n} \mathbb{1}_{A_k} = 0$$

para tal n. Tomando $n \to \infty$ se llega a que,

$$0 = \lim_{n\to\infty} \sup_{k\geq n} \mathbb{1}_{A_k} = \limsup_{n\to\infty} \mathbb{1}_{A_n}$$

Probando las identidades.

P2: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sea $\{X_n\}_{n\in\mathbb{N}}$ funciones $\mathcal{F}|\mathcal{B}(\mathbb{R})$-medibles. Sea $T : (\Omega, \mathcal{F}) \to (\mathbb{N}, \mathcal{P}(\mathbb{N}))$ una función medible y definimos la función:

$$X_T : \Omega \to \mathbb{R}$$

dada por $(X_T)(\omega) = X_{T(\omega)}(\omega)$. Muestre que X_T es medible.

Nota: Cuando hablamos de variables aleatorias, nos referimos a funciones medibles.

Nota: Sea $X : \Omega_1 \to \Omega_2$ una variable aleatoria (función medible), en teoría de probabilidad escribimos de la siguiente manera las preimagenes de conjuntos: Dado $B \in \mathcal{F}_2$ se tiene que,

$$X^{-1}(B) = \{\omega \in \Omega_1 : X(\omega) \in B\} = \{X \in B\}$$

Sol: Vamos a probar por definición. Sea $B \in \mathcal{B}(\mathbb{R})$, vemos que,

$$(X_T)^{-1}(B) = \{X_T \in B\}$$

Ahora vamos a usar un truco. Notemos lo siguiente,

$$\Omega = \bigcup_{n\in\mathbb{N}} \{T = n\}$$

Entonces,

$$\begin{aligned}\{X_T \in B\} &= \{X_T \in B\} \cap \Omega \\ &= \{X_T \in B\} \cap \left(\bigcup_{n\in\mathbb{N}} \{T = n\} \right) \\ &= \bigcup_{n\in\mathbb{N}} \underbrace{\{X_n \in B\}}_{\in\mathcal{F}}\end{aligned}$$

Por tanto $\{X_T \in B\} \in \mathcal{F}$, es decir, X_T es $\mathcal{F}|\mathcal{B}(\mathbb{R})$-medible.

P3: Sean $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$ espacios medibles. Sea $\mathcal{C} \subseteq \mathcal{F}_2$ tal que $\sigma(\mathcal{C}) = \mathcal{F}_2$. Sea $X : \Omega_1 \to \Omega_2$ tal que,

$$X^{-1}(C) \in \mathcal{F}_1$$

para todo $C \in \mathcal{C}$. Entonces X es medible.

Sol: Consideremos el siguiente conjunto:

$$\mathcal{L} = \{L \in \mathcal{F}_2 : X^{-1}(A) \in \mathcal{F}_1\} \subseteq \mathcal{F}_2$$

Está claro que $\mathcal{C} \subseteq \mathcal{L}$, entonces,

$$\mathcal{F}_2 = \sigma(\mathcal{C}) \subseteq \sigma(\mathcal{L}) \subseteq \mathcal{F}_2$$

$(\sigma(\cdot)$ preserva la inclusión). Por lo que $\sigma(\mathcal{L}) \subseteq \mathcal{F}_2$. Probemos que $\mathcal{L}$ es un σ-álgebra, para ello debemos probar los tres axiomas:

(a) Vemos que $\emptyset \in \mathcal{L}$, puesto que $\emptyset = X^{-1}(\emptyset) \in \mathcal{F}_1$.

(b) Sea $L \in \mathcal{L}$, luego $X^{-1}(L) \in \mathcal{F}_1$, entonces,

$$(X^{-1}(L))^c = X^{-1}(L^c) \in \mathcal{F}_1$$

Luego $L^c \in \mathcal{L}$.

(c) Sea $\{L_n\}_{n\in\mathbb{N}}$ una colección numerable de eventos de $\mathcal{L}$. Notemos que,

$$\bigcup_{n\in\mathbb{N}}(X^{-1})(L_n) = X^{-1}\left(\bigcup_{n\in\mathbb{N}} L_n\right) \in \mathcal{F}_1$$

Por tanto,

$$\bigcup_{n\in\mathbb{N}} L_n \in \mathcal{L}$$

De esta forma $\mathcal{L}$ es un σ-álgebra, concluyendo que $\mathcal{L} = \mathcal{F}_2$. Ahora esto significa que X es una función medible.

P4: Consideremos,

$$I_s = \{(a,b] : -\infty < a < b < \infty, a, b, \in \mathbb{R}\}$$

Entonces $\sigma(I_s) = \mathcal{B}(\mathbb{R})$. Análogamente muestre que dado el conjunto,

$$I_- = \{(-\infty, x] : x \in \mathbb{R}\}$$

entonces $\sigma(I_-) = \mathcal{B}(\mathbb{R})$. ¿Qué semi-álgebra de $\mathbb{R}^n$ cumple el análogo?

Sol: Recordemos que se define el σ-álgebra generado por un conjunto C de la siguiente forma:

$$\sigma(C) = \bigcap_{\substack{A\ \sigma\text{-álg} \\ C\subseteq A}} A$$

Entonces para que $\sigma(I_s) \subseteq \mathcal{B}(\mathbb{R})$ debemos probar que $I_s \subseteq \mathcal{B}(\mathbb{R})$. Sea $(a,b]\imath I_s$, notemos la siguiente identidad:

$$(a,b] = \bigcap_{n\in\mathbb{N}} \underbrace{\left(a, b+\frac{1}{n}\right)}_{\in\mathcal{B}(\mathbb{R})}$$

es decir, $(a,b]$ es intersección numerable de elementos del conjunto borel $\mathcal{B}(\mathbb{R})$, por tanto $I_S \subseteq \mathcal{B}(\mathbb{R})$. Para la otra inclusión consideremos el siguiente lema:

Lema: *Todo abierto en $\mathbb{R}$ es unión disjunta de intervalos abiertos.*

Sea $O \in \{O : O \subseteq \mathbb{R}, \text{abierto}\}$, entonces por el lema O se escribe de la siguiente forma:

$$O = \bigcup_{n\in\mathbb{N}} I_n$$

donde I_n es un intervalo abierto. Ahora notemos que un intervalo abierto (p, q) se puede escribir de la siguiente forma:

$$(p, q) = \bigcup_{n\in\mathbb{N}} \underbrace{\left(p, q - \frac{1}{n}\right]}_{\in I_s}$$

Por tanto O es unión numerable de conjuntos de la colección I_s. Por lo tanto,

$$\sigma(I_s) = \mathcal{B}(\mathbb{R})$$

Probemos la otra igualdad. Sea $(-\infty, x] \in I_-$, notemos que,

$$(-\infty, x]^c = (x, \infty) \in \mathcal{B}(\mathbb{R})$$

Entonces $\sigma(I_-) \subseteq \mathcal{B}(\mathbb{R})$. Sea $I = (p, q)$ un intervalo abierto, luego,

$$I^c = (-\infty, p] \cup [q, \infty)$$

Claramente $(-\infty, p] \in \sigma(I_-)$, falta verificar el otro intervalo, notemos que,

$$[q, \infty)^c = (-\infty, q) = \bigcup_{n\in\mathbb{N}} \underbrace{\left(-\infty, q - \frac{1}{n}\right]}_{\in I_-}$$

Por tanto $[q, \infty) \in \sigma(I_-)$ y por lo tanto $I \in \sigma(I_-)$. Finalmente aplicamos el lema anterior y concluimos que $\sigma(I_-) = \mathcal{B}(\mathbb{R})$.

Determinemos un semiálgebra que genere a $\mathcal{B}(\mathbb{R}^n)$. Recordemos que un semiálgebra, es una colección $\mathcal{S}$ tal que la intersección es cerrada y para todo $S \in \mathbf{S}$, el complementos es la unión finita disjunta de elementos de la colección.

Para ello trabajaremos en el caso $n = 2$. Consideremos la colección,

$$I_s^2 = \{(-\infty, x_1] \times (-\infty, x_2] : x_1, x_2 \in \mathbb{R}\}$$

Entonces I_s^2 es un semiálgebra con σ-álgebra generado los Boreleanos de $\mathbb{R}^2$.

- **Semiálgebra:** Observemos lo siguiente,

$$(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$$

Entonces es evidente que dados $I_1, I_2 \in I_s^2$, se tiene que $I_1 \cap I_2 \in I_s^2$.

Ahora consideremos $I \in I_s^2$, digamos que,

$$I = (-\infty, x_1] \times (-\infty, x_2]$$

Luego,

$$I^c = (x_1, \infty) \times (-\infty, x_2] \cup (-\infty, x_1] \times (x_2, \infty) \cup (x_1, \infty) \times (x_2, \infty)$$

Pero no sabemos si $(x_1, \infty) \times (-\infty, x_2]$, sin embargo, sabemos por construcción que $(x_1, \infty) = (-\infty, x_1]$ y de por si I_s es un semiálgebra, de forma que tenemos que I^c es unión finita de elementos de $\mathcal{S}$.

- **σ-álgebra Generado:** Por definición $\mathcal{B}(\mathbb{R}^2)$ es el σ-álgebra generado por los abiertos de $\mathbb{R}^2$, además, todo abierto $O \subseteq \mathbb{R}^2$ **terminar.0**

Ayudantía 2

Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Los objetos de interés en probabilidad, son las aplicaciones de la forma $X : (\Omega, \mathcal{F}) \to (T, \mathcal{B}(T))$ donde T en general es un espacio topológico (en algunos casos metrizable). Usualmente trabajamos con $\mathbb{R}$ pero también se puede trabajar con $\mathbb{R}^d$ $d \geq 1$ o bien espacios de funciones como $C[0,1]$.

Pregunta: ¿Se pueden construir funciones aleatorias o vectores aleatorios a partir de variables aleatorias? Esto se responderá en esta ayudantía.

P1: Sea Ω_1 un conjunto, $(\Omega_\alpha, \mathcal{F}_\alpha)_{\alpha\in\Gamma}$ una familia de espacios medibles y $X_\alpha : \Omega_1 \to \Omega_\alpha$ para cada $\alpha \in \Gamma$. Sea $\mathcal{G} = \sigma(X_\alpha : \alpha \in \Gamma)$ la menor σ-álbegra sobre Ω_1 para la cual, las funciones X_α son medibles. Sea $(\Omega, \mathcal{F})$ un espacio medible y sea $Y : \Omega \to \Omega_1$. Entonces:

(a) Muestre que Y es $\mathcal{G}$-medible si y sólo si $X_\alpha \circ Y$ es $\mathcal{F}_\alpha$-medible para todo $\alpha \in \Gamma$,

(b) Sea $(\Omega, \mathcal{F})$ un espacio de probabilida y sea $\{X_\alpha\}_{\alpha\in\Gamma}$ una colección de variables aleatorias definidas en este espacio (lo que suele llamarse proceso estocástico). Demuestre que el mapa $X : (\Omega, \mathcal{F}) \to \mathbb{R}^\Gamma$ definida por $X(\omega) := \{X_\alpha(\omega)\}_{\alpha\in\Gamma}$ es medible (lo será si y sólo si cada coordenada es una variable aleatoria **verificar**)

La información de la hipótesis del enunciado el primer problema, se puede resumir en un diagrama conmutativo.

Diagrama conmutativo ayudanti a2

Sol:

(a) Supongamos que Y es $\mathcal{G}$-medible, sea $B \in \mathcal{F}_\alpha$, luego se tiene que,

$$(X_\alpha \circ Y)^{-1}(B) = Y^{-1}(X_\alpha^{-1}(B))$$

Si X_α es $\mathcal{G}$-medible, entonces $X_\alpha^{-1}(B) \in \mathcal{G}$, ahora por hipótesis se tiene que $= Y^{-1}(X_\alpha^{-1}(B)) \in \mathcal{F}$ como queriamos probar.

Supongamos que $X_\alpha \circ Y$ es $\mathcal{F}_\alpha$-medible, entonces para todo $B \in \mathcal{F}_\alpha$ se tiene que,

$$Y^{-1}(X_\alpha^{-1}(B)) \in \mathcal{F}$$

Queremos probar que para todo $C \in \mathcal{G}$ se tiene que $Y^{-1}(C) \in \mathcal{F}$. Para ello queremos podeer descomponer todo C es algo compuesto por X_α^{-1}.

Afirmación: *Sea* $\mathcal{G}_\alpha := \sigma(X_\alpha)$, *entonces,*

$$\mathcal{G} = \sigma\left(\bigcup_{\alpha\in\Gamma} \mathcal{G}_\alpha\right)$$

Dem: Debemos probar que el σ-álgebra generado por la unión de los $\mathcal{G}_\alpha$ es la menor σ-álgebra tal que los X_α son medibles.

- **σ-álgebra:** Claramente es σ-álgebra por definición.

- G **la menor σ-álgebra:** Sea $\mathcal{L}$ un σ-álgebra tal que todos los X_α son medibles. Luego para,

$$A \in \sigma\left(\bigcup_{\alpha\in\Gamma} \mathcal{G}_\alpha\right)$$

Se tiene que existe un $\alpha \in \Gamma$ y un $B \in \mathcal{F}_\alpha$ tal que $A = X_\alpha^{-1}(B)$, por definición de $\mathcal{L}$ se tiene que $A \in \mathcal{L}$ probando que es la menor.

■

De esta forma, para todo $A \in \mathcal{G}$ se puede escribir como combinaciones numerables de X_α^{-1}, esto implica que,

$$Y^{-1}(A) \in \mathcal{F}$$

como queriamos probar.

(b) Con respecto a $\mathbb{R}^\Gamma$ **terminar**

P2: Sea $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}), \lambda)$ una variable no negativa. Muestre que:

$$\mathbb{E}(X) := (\mathbb{P} \otimes \lambda)(\{(\omega, t) \in \Omega \times \mathbb{R}_+ : X(\omega) \geq t \geq 0\})$$

donde $\mathbb{P} \otimes \lambda$ es la medida producto. ¿Qué interpretación se le puede dar al lado derecho?

Sol: Debemos probar antes que el conjunto:

$$A = \{(\omega, t) \in \Omega \times \mathbb{R}_+ : X(\omega) \geq t \geq 0\}$$

es medible, es decir, $A \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R})$. Observemos lo siguiente:

$$A = \bigcup_{n\in\mathbb{N}} \{(\omega, t) \in \Omega \times \mathbb{R}_+ : X(\omega) \geq t \geq 0, n \geq t \geq n-1\})$$

La inclusión $\supseteq$) es evidente. Para la otra notemos que para $(\omega, t) \in A$ se tiene que $X(\omega) \geq t \geq 0$, pero de forma independiente existe un $n \in \mathbb{N}$ tal que $n \geq t \geq n-1$, luego se concluye la otra inclusión.

Ahora, notemos que,

$$\{(\omega, t) \in \Omega \times \mathbb{R}_+ : X(\omega) \geq t \geq 0, n \geq t \geq n-1\}) = (X^{-1}([t, \infty)) \times \mathbb{R}_+) \cap (\Omega \times [n-1, n])$$

Por tanto $A \in \mathcal{F} \otimes \mathcal{B}(\mathbb{R})$.

Dado que A es medible, podemos demostrar el enunciado. Notemos que la función,

$$\mathbb{1}_A : \Omega \times \mathbb{R}_+ \to \mathbb{R}$$

es medible dado que A es medible, y es no negativa, por lo tanto por Fubini se tiene que,

$$\begin{aligned}(\mathbb{P}\otimes\lambda)(A) &= \int_{\Omega\times\mathbb{R}_+} \mathbb{1}_A d(\mathbb{P}\otimes\lambda)\\ &= \int_\Omega \left(\int_{\mathbb{R}_+} \mathbb{1}_A d\lambda\right) d\mathbb{P}\\ &= \int_\Omega \left(\int_{[0,X(\omega)]} 1 d\lambda\right) d\mathbb{P}\\ &= \int_\Omega X(\omega) d\mathbb{P} = \mathbb{E}(X)\end{aligned}$$

Usando que 1 es R-integrable en el intervalo $[0, X(\omega)]$ por lo que coincide con la integral de Lebesgue en el mismo intervalo con la medida λ de Lebesgue.

P3: Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida finita y sea $X : \Omega \to \mathbb{R}$ una función medible. Demuestre que para todo $q \leq r$ tal que $\mathbb{E}(|X|^r) < \infty$, entonces $\mathbb{E}(|X|^q) < \infty$.

Sol: Queremos acotar la integral:

$$\mathbb{E}(|X|^q) = \int_\Omega |X(\omega)|^q d(\mu(\omega))$$

El problema es que no necesariamente $|X(\omega)|^q \leq |X((\omega)|^r$ para todo $\omega \in \Omega$, puede pasar que la desigualdad se invierta cuando ω es tal que $X(\omega) < 1$. Para arreglat esto vamos a separar $\Omega = \{\omega \in \Omega : |X(\omega)| \leq 1\} \cup \{\omega \in \Omega : |X(\omega)| > 1\}$ (por comodidad denotaremos $\{\omega \in \Omega : |X(\omega)| \leq 1\} = \{|X| \leq 1\}$), entonces se tiene que,

$$\begin{aligned}\mathbb{E}(|X|^q) &= \int_\Omega |X(\omega)|^q d\mu\\ &= \int_\Omega |X|^q \mathbb{1}_{\{|X|\leq 1\}} + |X|^q \mathbb{1}_{\{|X|>1\}} d\mu\\ &\leq \int_\Omega |X|^q \mathbb{1}_{\{|X|\leq 1\}} d\mu + \int_\Omega |X|^q \mathbb{1}_{\{|X|>1\}} d\mu\\ &\leq \mu(\{|X|\leq 1\}) + \int_{\{|X|<1\}} |X|^r d\mu\\ &\leq \mu(\{|X|\leq 1\}) + \mathbb{E}(|X|^r)\\ &< \infty\end{aligned}$$

Probando que $\mathbb{E}(|X|^q) < \infty$.

Ayudantía 3

En esta ayudantía estudiaremos la independencia y algunas propiedades implícitas de los valores centralisados como la esperanza. Pero antes haremos un pequeño estudio sobre los elementos aleatorios.

Elementos Aleatorios

En probabilidad los elementos aleatorios los entendemos como aplicaciones que toman un espacio de probabilidad y lo manda a un espacio de probabilidad con ciertas características. Un ejemplo sencillo es una variable aleatoria $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, esta variable aleatoria induce una medida en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ conocida como la distribución de X definida como:

$$\mathbb{P}_X(B) := \mathbb{P}(X \in B)$$

para todo $B \in \mathcal{B}(\mathbb{R})$. También podemos definir un vector aleatorio que es de la forma:

$$X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$$

De forma análoga podemos definir una medida en $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ dada por:

$$\mathbb{P}_X(B) := \mathbb{P}(X \in B)$$

para todo $B \in \mathcal{B}(\mathbb{R}^n)$.

Queremos generalizar para funciones aleatorias, cosas de la forma:

$$X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^\Gamma, \mathcal{G})$$

donde $\mathbb{R}^\Gamma$ son las sucesiones inducidas por el conjunto Γ que puede ser infinito numerable o no numerable. Supongamos que X está compuesta de la siguiente forma:

$$X(\omega) = (X_\alpha(\omega))_{\alpha\in\Gamma}$$

para todo $\omega \in \Omega$ donde $\{X_\alpha\}_{\alpha\in\Gamma}$ son variables aleatorias. Por ejemplo, podemos considerar $\Gamma = [0, 1] \subset \mathbb{R}^+$ y construir una función aleatoria como la siguiente figura: **Figura ayudantia 3**

Para que todo funcione necesitamos que los X_α sea medibles, y para ello debemos entender $\mathcal{G}$ y claramente pedir que sea σ-álgebra. Vamos a definir $\mathcal{G}$ como el σ-álgebra generado por los cilindros finitos dimensionales.

Definición: *Sea $\{(\Omega_i, \mathcal{F}_i)\}_{i\in\Gamma}$ espacios medibles. Definimos,*

$$\Omega := \prod_{i\in\Gamma} \Omega_i$$

Sea $j \subset\subset \Gamma$ (conjunto finito) y sea mapa natural,

$$\prod_J : \Omega \to \Omega_J =: \prod_{i\in J} \Omega_i$$

$$\prod_J ((\omega_i)_{i\in I}) := (\omega_i)_{i\in J}$$

Los cilindros finitos dimensionales son los conjuntos de la forma:

$$\begin{aligned} C_J &= \left\{ \left(\prod_J \right)^{-1} (B) : B \in \bigotimes_{j \in J} \mathcal{F}_j \right\} \\ &= B \times \prod_{i \in \Gamma \setminus J} \Omega_i \end{aligned}$$

También definimos:

$$\mathcal{Z} := \left\{ \left(\prod_J \right)^{-1} (B) : B \in \bigotimes_{j \in J} \mathcal{F}_j, \text{ para todo } J \subset\subset \Gamma \right\}$$

y por tanto $\mathcal{G} = \sigma(\mathcal{Z})$.

Hecho: *Si* $\Gamma = \mathbb{N}$ *entonces* $\sigma(\mathcal{Z}) = \mathcal{B}(\mathbb{R}^{\mathbb{N}})$.

Hecho: *Se cumple que* $\sigma(\mathcal{Z}) = \sigma\left(\prod_\alpha : \alpha \in \Gamma\right)$.

Es decir, $\sigma(\mathcal{Z})$ es la menor σ-álgebra donde las proyecciones finitas son medibles.

P1: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sean $X, Y : \Omega \to \mathbb{R}$ variables aleatorias.

(a) Demostrar que (X, Y) es una aplicación $\mathcal{F}|\mathcal{B}(\mathbb{R}^2)$-medible,

(b) Probas que las siguientes afirmaciones son equivalentes:

  (i) Las variables aleatorias X, Y son independientes,

  (ii) $F_{(X,Y)} = F_X \cdot F_y$,

  (iii) $\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y$,

  (iv) En $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}), \mathbb{P}_{(X,Y)})$ las proyecciones canónicas son independientes, con distribución $\mathbb{P}_X$ y $\mathbb{P}_Y$ respectivamente.

Sol:

(a) Tenemos un vector aleatorio,

$$(X, Y) : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2))$$

Consideremos las proyecciones canónicas dadas por:

$$\begin{aligned} \pi_x : \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto x \\ \pi_y : \mathbb{R}^2 &\to \mathbb{R} \\ (x, y) &\mapsto y \end{aligned}$$

Entonces para todo $A \subseteq \mathbb{R}^2$ se puede escribir de la siguiente forma:

$$A = (\pi_x(A) \times \mathbb{R}) \cap (\mathbb{R} \times \pi_y(A)) = \pi_x(A) \times \pi_y(A)$$

También notemos que π_x, π_y son funciones abiertas, es decir, si $A \subseteq \mathbb{R}^2$ es abierto, entonces $\pi_x(A), \pi_y(A) \subseteq \mathbb{R}$ también lo son, y esto se cumple debido a que, para $a \in \pi_x(A)$ se tiene que existe un b tal que $(a, b) \in A$ luego existe un $r > 0$ de forma que,

$$B := B((a, b), r) \in A$$

Aplicando π_x a B obtenemos lo siguiente,

$$\pi_x(B) = (a - r, a + r)$$

que es abierto en $\mathbb{R}$, y por tanto,

$$a \in (a - r, a + r) \subseteq \pi_x(A)$$

Análogamente se puede probar para π_y. Observemos también que para $A \in \mathcal{B}(\mathbb{R}^2)$ se cumple que,

$$(X, Y)^{-1}(A) = X^{-1}(\pi_x(A)) \cap Y^{-1}(\pi_y(A))$$

por lo que para probar que (X, Y) es medible, debemos probar que $\pi_x(A), \pi_y(A) \in \mathcal{B}(\mathbb{R}^2)$ para todo $A \in \mathcal{B}(\mathbb{R}^2)$. Supongamos que A es abierto, entonces se puede escribir de la siguiente forma:

$$A = \bigcup_{x \in A} B(x, r_x)$$

donde $B(x, r_x) \subseteq \mathcal{B}(\mathbb{R}^2)$, y consideremos la siguiente colección:

$$\mathcal{L} := \{A \in \mathcal{B}(\mathbb{R}^2) : \pi_x(A) \in \mathcal{B}(\mathbb{R})\} \subseteq \mathcal{B}(\mathbb{R}^2)$$

Observemos que todo abierto A está en $\mathcal{L}$, ya que,

$$\pi_x(A) = \bigcup_{n \geq 1} I_n \in \mathcal{B}(\mathbb{R})$$

donde I_n son invervalos disjuntos abiertos (la imagen de abierto es abierto), probemos que $\mathcal{L}$ es un σ-álgebra.

(a) $\emptyset$ es abierto, por lo que $\emptyset \in \mathcal{L}$,

(b) Si $A \in \mathcal{L}$, entonces,

$$\pi_x(A^c) = (\pi_x(A))^c \in \mathcal{B}(\mathbb{R})$$

Por lo que $A^c \in \mathcal{L}$,

(c) Sea $\{A_n\}_{n \in \mathbb{N}}$ una secuencia de eventos en $\mathcal{L}$, luego,

$$\pi_x\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \bigcup_{n \in \mathbb{N}} \pi_x(A_n) \in \mathcal{B}(\mathbb{R})$$

Por lo tanto $\mathcal{L}$ es un σ-álgebra tal que,

$$\mathcal{B}(\mathbb{R}^2) = \mathcal{L}$$

Es decir, $\pi_x(A) \in \mathcal{B}(\mathbb{R})$ para todo $A \in \mathcal{B}(\mathbb{R}^2)$. Finalmente podemos concluir que,

$$X^{-1}(\pi_x(A)) \cap Y^{-1}(\pi_y(A)) \in \mathcal{F}$$

probando que (X, Y) es medible.

De forma inductiva podemos concluir que si $X_1, \ldots, X_n$ son medibles, entonces $(X_1, \ldots, X_n)$ es medible.

(b) Probemos las equivalencias.

- **(i) implica (ii):** Supongamos que las variables aleatorias son independientes. Entonces para todo $A, B \in \mathcal{B}(\mathbb{R})$ se cumple que,

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

  Tomando $A = (-\infty, x], B = (-\infty, y]$ con $x, y \in \mathbb{R}$ obtenemos,

$$F_{(X,Y)}(x, y) = \mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y) = F_X(x)F_Y(y)$$

  Probando (ii).

- **(ii) implica (iii):** Vamos a demostrar el siguiente resultado:

  Proposición: *Sea $(\Omega, \mathcal{F})$ un espacio medible y sean $\mathbb{P}_1, \mathbb{P}_2$ medidas de probabilidad tales que,*

$$\mathbb{P}_1(C) = \mathbb{P}_2(C)$$

  con $C \in \mathcal{C}$ un π-sistema tal que $\sigma(\mathcal{C}) = \mathcal{F}$, entonces $\mathbb{P}_1 = \mathbb{P}_2$.

  Dem: Vamos a demostar la proposición usando el teorema $\pi - \lambda$. Consideremos la colección:

$$\mathcal{L} := \{C \in \mathcal{F} : \mathbb{P}_1(C) = \mathbb{P}_2(C)\} \subseteq \mathcal{F}$$

  Claramente $\mathcal{C} \subseteq \mathcal{L}$, probemos que $\mathcal{L}$ es un λ-sistema.

  - Sean $A, B \in \mathcal{L}$ tales que $A \subset B$, notemos que,

$$\begin{aligned} \mathbb{P}_1(B \setminus A) &= \mathbb{P}_1(B) - \mathbb{P}_1(A) \\ &= \mathbb{P}_2(B) - \mathbb{P}_2(A) \\ &= \mathbb{P}_2(B \setminus A) \end{aligned}$$

    Luego $B \setminus A \in \mathcal{L}$.

- Consideremos $\{A_n\}_{n\in\mathbb{N}}$ una sucesión de elementos de $\mathcal{B}$ tal que $A_n \uparrow A$. Entonces,

$$\begin{aligned}\mathbb{P}_1(A) &= \mathbb{P}_1\left(\bigcup_{n\in\mathbb{N}} A_n\right)\\ &= \lim_{n\to\infty} \mathbb{P}_1(A_n)\\ &= \lim_{n\to\infty} \mathbb{P}_2(A_n)\\ &= \mathbb{P}_2\left(\bigcup_{n\in\mathbb{N}} A_n\right)\\ &= \mathbb{P}_2(A)\end{aligned}$$

  Por tanto $A \in \mathcal{L}$.

Por lo tanto $\mathcal{L}$ es un λ-sistema que contiene a $\mathcal{C}$, luego se tiene que $\sigma(\mathcal{C}) = \mathcal{F} \subseteq \mathcal{L}$, es decir, $\mathcal{F} = \mathcal{L}$. Probando que $\mathbb{P}_1 = \mathbb{P}_2$ en $\mathcal{F}$. ■

Queremos probar que para $A \in \mathcal{B}(\mathbb{R}^2)$ se tiene que,

$$\mathbb{P}_{(X,Y)}(A) = (\mathbb{P}_X \otimes \mathbb{P}_Y)(A)$$

Observemos que para $A = \{(-\infty, x] \times (-\infty, y]\}$ se cumple que,

$$\begin{aligned}\mathbb{P}_{(X,Y)}(A) &= \mathbb{P}((X,Y) \in A)\\ &= \mathbb{P}(\{\omega \in \Omega : X(\omega) \le x, Y(\omega) \le y\})\\ &= F_{(X,Y)}(x,y)\\ &= F_X(x)F_Y(y)\\ &= \mathbb{P}_X((-\infty, x])\mathbb{P}_Y((-\infty, y])\\ &= (\mathbb{P}_X \otimes \mathbb{P}_Y)(A)\end{aligned}$$

Dado que los intervalos $(a, b]$ se puede escribir como unión o intersección de (c, d) y viceversa, se tiene que el σ-álgebra generado por la colección de los conjunto de la forma $\{(-\infty, x] \times (-\infty, y]\}$ con $x, y \in \mathbb{R}$ es igual a $\mathcal{B}(\mathbb{R}^2)$, y como la colección en si es un π-sistema, por la proposición se tiene que,

$$\mathbb{P}_{(X,Y)}(A) = (\mathbb{P}_X \otimes \mathbb{P}_Y)(A)$$

se cumple para todo $A \in \mathcal{B}(\mathbb{R}^2)$.

- **(iii) implica (iv):** Las proyecciones canónicas se definen por:

$$\begin{aligned}\pi_i : (\mathbb{R}^2, \mathcal{B}(\mathbb{R}^2)) &\to (\mathbb{R}, \mathcal{B}(\mathbb{R}))\\ (x,y) &\mapsto i\end{aligned}$$

donde $i = x, y$. Veamos que son medibles. Sea $A \in \mathcal{B}(\mathbb{R})$, entonces,

$$(\pi_x)^{-1}(A) = A \times \mathbb{R} \in \mathcal{B}(\mathbb{R}^2)$$

por lo que π_x es medible. Análogamente con π_y. Por lo tanto π_x, π_y son variablea aleatorias, de forma que debemos probar que,

$$\mathbb{P}(\pi_x \in A, \pi_y \in B) = \mathbb{P}(\pi_x \in A)\pi_\frown \in \mathbb{B}$$

para todo $A, B \in \mathcal{B}(\mathbb{R})$. Por definición,

$$\begin{aligned}\mathbb{P}(\pi_x \in A, \pi_y \in B) &= \mathbb{P}_{(\pi_x,\pi_y)}(A \times B)\\ &= (\mathbb{P}_{\pi_x} \otimes \mathbb{P}_{\pi_y})(A \times B)\\ &= \mathbb{P}_{\pi_x}(A)\mathbb{P}_{\pi_y}(B)\\ &= \mathbb{P}(\pi_x \in A)\mathbb{P}(\pi_y \in B)\end{aligned}$$

Para todo $A, B \in \mathcal{B}(\mathbb{R})$. Es decir, las proyecciones canónicas son independientes.

- **(iv) implica (i):** Debemos probar que,

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$$

para todo $A, B \in \mathcal{B}(\mathbb{R})$. Observemos que,

$$\begin{aligned}\pi_i(X, Y) : \Omega &\to \mathbb{R}\\ \omega &\mapsto x_i(\omega)\end{aligned}$$

donde $x_i = X, Y$ con $i = x, y$ de forma evidente. Notemos que $\pi_x(X, Y)$ es medible por composición de medibles, lo mismo se concluye para $\pi_y(X, Y)$, de forma que es una variable aleatoria. Luego,

$$\begin{aligned}\mathbb{P}(X \in A, Y \in B) &= \mathbb{P}(\pi_x(X, Y) \in A, \pi_y(X, Y) \in B)\\ &= \mathbb{P}(\pi_x(X, Y) \in A)\mathbb{P}(\pi_y(X, Y) \in B)\\ &= \mathbb{P}(X \in A)\mathbb{P}(Y \in B)\end{aligned}$$

Probando que las afirmaciones son equivalentes.

P2: Sean $(\Omega_1, \mathcal{F}_1), (\Omega_2, \mathcal{F}_2)$ espacios medibles con $\mathbb{P}$ una medida de probabilidad en $(\Omega_1, \mathcal{F}_1)$ y sea $X : \Omega_1 \to \Omega_2$ una función medible. En $(\Omega_2, \mathcal{F}_2)$ definimos la medida:

$$\mathbb{P}_X(A) := \mathbb{P}(X^{-1}(A))$$

Muestre que para $f : (\Omega_2, \mathcal{F}_2) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ se tiene que,

$$f \text{ es } \mathbb{P}_X\text{-integabrable} \iff f(X) \text{ es } \mathbb{P}\text{-integabrable}$$

y que, en cuyo caso, vale la igualdad:

$$\int_{\Omega_2} f d\mathbb{P}_X = \int_{\Omega_1} f(X) d\mathbb{P}$$

Sol: Supongamos que $f = \mathbb{1}_A$ donde $A \in \mathcal{F}_2$. Luego se tiene que,

$$\begin{aligned}
\int_{\Omega_1} f \circ X d\mathbb{P} &= \int_{\Omega_1} \mathbb{1}_{X^{-1}(A)} d\mathbb{P} \\
&= \mathbb{P}(X^{-1}A) \\
&= \mathbb{P}_X(A) \\
&= \int_{\Omega_2} \mathbb{1}_A d\mathbb{P}_X = \int_{\Omega_2} f d\mathbb{P}_X
\end{aligned}$$

Ahora si consideramos f una función simple no negativa, se concluye de igual de forma análoga que,

$$\int_{\Omega_2} f d\mathbb{P}_X = \int_{\Omega_1} f(X) d\mathbb{P}$$

Supongamos que f es no negativa, entonces existe una sucesión $f_n \uparrow f$, luego por el teorema de convergencia monótona se tiene que,

$$\begin{aligned}
\int_{\Omega_2} f d\mathbb{P}_X &= \int_{\Omega_2} \lim_{n\to\infty} f_n d\mathbb{P}_X \\
&= \lim_{n\to\infty} \int_{\Omega_2} f_n d\mathbb{P}_X \\
&= \lim_{n\to\infty} \int_{\Omega_1} f_n \circ X d\mathbb{P} \\
&= \int_{\Omega_1} \lim_{n\to\infty} f_n \circ X d\mathbb{P} \\
&= \int_{\Omega_1} f \circ X d\mathbb{P}
\end{aligned}$$

Para concluir el resultado debemos recordear que toda función real puede se expresedado como la resta de dos funciones no negativas, de forma que aplicar lo anterior y obtenemos el resultado pedido.

P3: Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias. Definimos la σ-álgebra de la cola por:

$$\mathcal{T} := \bigcap_{n\in\mathbb{N}} \sigma(X_n, X_{n+1}, \dots)$$

Muestre que,

$$\left\{ \lim_{n\to\infty} X_n \text{ existe} \right\} \in \mathcal{T}$$

y que si $\{X_n\}_{n\in\mathbb{N}} \sim \mathcal{U}[0,1]$, entonces,

$$\left\{ \prod_{i=1}^{\infty} X_n = 0 \right\} \notin \mathcal{T}$$

Ayudantía 4

P1: Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias. Recordemos la definición de la σ-álgebra de la cola,

$$\mathcal{T} = \bigcap_{n\in\mathbb{N}} \sigma(X_n, X_{n+1}, \dots)$$

Muestre que $\{\lim_{n\to\infty} X_n \text{ existe }\} \in \mathcal{T}$ y que si $\{X_n\}_{n\in\mathbb{N}} \sim \mathcal{U}[0,1]$ son las proyecciones canónicas de $([0,1]^{\mathbb{N}}, \sigma(\mathcal{Z}), \mathbb{P})$, entonces $\{\prod_{i=1}^{\infty} X_n = 0\} \notin \mathcal{T}$ donde $\sigma(\mathcal{Z})$ es la σ-álgebra de los cilindros-dimensionales y $\mathbb{P}$ la medida producto. Como ejercicio, verifique que los siguientes eventos pertenecen a la σ-álgebra de la cola:

(a) $\{\limsup_{n\to\infty} X_n \text{ existe }\}$,

(b) $\{\liminf_{n\to\infty} X_n \text{ existe }\}$,

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

Sol: Sabemos que si X es una variable aleatoria X en $(\Omega, \mathcal{F}, \mathbb{P})$. Entonces $\sigma(X) = \{X^{-1}(A) : A \in \mathcal{B}(\mathbb{R})\}$. Ahora consideremos $X_1, \dots, X_n$ variables aleatorias, entonces,

$$\sigma(X_1, \dots, X_n) = \{(X_1, \dots, X_n)^{-1}(B) : B \in \mathcal{B}(\mathbb{R}^n)\}$$

En general, para una colección $\{X_i\}_{i\in\mathbb{N}}$ de variables aleatorias se define el menor σ-álgebra tales que X_i son medibles para todo $i \geq 1$ es,

$$\sigma(X_1, X_2, \dots) = \{(X_1, X_2, \dots)^{-1}(B) : B \in \sigma(\mathcal{Z})\}$$

donde $\sigma(\mathcal{Z})$ es el σ-álgebra generado por la colección de los cilindros finitos dimensionales. En particular, $\sigma(mathcalZ) = \mathcal{B}(\mathbb{R}^{\mathbb{N}})$. Ahora definimos la σ-álgebra cola por:

$$\mathcal{T} = \bigcap_{n\in\mathbb{N}} \sigma(X_n, X_{n+1}, \dots)$$

Probemos que,

$$\left\{\lim_{n\to\infty} X_n \text{ existe}\right\} \in \mathcal{T}$$

Si para un $\omega \in \Omega$ se tiene que,

$$\lim_{n\to\infty} X_n(\omega)$$

existe, entonces se tiene que Cauchy, es decir, para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que para todo $n, m \geq N$ se tiene que,

$$\|X_n(\omega) - X_m(\omega)\| < \varepsilon$$

Vamos a expandir esto, esto significa que,

$$\left\{\lim_{n\to\infty} X_n \text{ existe}\right\} = \bigcap_{\varepsilon>0} \bigcup_{N\in\mathbb{N}} \bigcap_{n,m\geq N} \{\|X_n - X_m\| < \varepsilon\}$$

Observemos que $\|X_n - X_m\|$ es una variable aleatoria no negativa, si consideramos $\varepsilon = 1/n$ obtenemos,

$$\bigcap_{1/n>0} \bigcup_{N\in\mathbb{N}} \bigcap_{n,m\geq N} \{\|X_n - X_m\| < \frac{1}{n}\}$$

Y claramente $\{\|X_n - X_m\| < 1/n\} = (\|X_m - X_n\|)^{-1}(-\infty, 1/n) \in \sigma(X_n, X_m)$. Luego,

$$\bigcap_{n,m\geq N} \{\|X_n - X_m\| < \varepsilon\} \in \sigma(X_n, X_{n+1}, \dots)$$

Por definición de la menor σ-álgebra. Como estamos uniendo con respecto a N de forma numerable, se tiene que,

$$\bigcup_{N\in\mathbb{N}} \bigcap_{n,m\geq N} \{\|X_n - X_m\| < \varepsilon\} \in \sigma(X_k, X_{k+1}, \dots)$$

Ahora observemos que la intersección numerable con respecto a $1/n > 0$ se puede ignorar y pensar todo con respecto a k solamente, es eecir,

$$\bigcap_{1/n>0} \bigcup_{N\in\mathbb{N}} \bigcap_{n,m\geq N} \{\|X_n - X_m\| < \frac{1}{n}\} \in \sigma(X_k, X_{k+1}, \dots)$$

para todo $k \geq 1$, es decir,

$$\left\{ \lim_{n\to\infty} X_n \text{ existe} \right\} \in \mathcal{T}$$

Como queriamos probar.

1.

P2: Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes. Para cada $n \in \mathbb{N}$ se definen las σ-álgebras:

$$\begin{aligned} \mathcal{F}_n &:= \sigma(X_i : 1 \leq i \leq n) \\ \mathcal{G}_n &:= \sigma(X_i : i > n) \end{aligned}$$

Sean además $\mathcal{F}_\infty = \sigma\left(\bigcup_{n\in\mathbb{N}} \mathcal{F}_n\right)$ y $\mathcal{G}_\infty \bigcap_{n\in\mathbb{N}} \mathcal{G}_n$.

(a) Demuestre que todo elemento de $\mathcal{G}_\infty$ es independiente de $\mathcal{F}_n$ para cada $n \in \mathbb{N}$,

(b) Deducir que todo elemento de $\mathcal{G}_\infty$ es independiente de los elementos de $\mathcal{F}_\infty$,

(c) Concluir que para todo $\mathcal{A} \in \mathcal{G}_\infty$, se tiene que $P(A) \in \{0, 1\}$.

P3: Sean $X \sim \mathcal{N}(\mu_X, \sigma_X^2), Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ variables aleatorias independientes. Muestre que $X + Y \sim \mathcal{N}(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$.

Sol:

Ayudantía 5

P1: Sean $\{X_k\}_{k=1}^n$ variables aleatorias tales que el vector aleatorio posee una densidad $f(x_1, \dots, x_n)$ y que tal densidad puede expresarse como:

$$f(x_1, \dots, x_n) = \prod_{i=1}^{n} g_i(x_i)$$

con cada $g_i(x_i)$ funciones medibles no negativas. Muestre que en tal caso las variables aleatorias son independientes.

Sol: Se tiene que,

$$\mathbb{P}_{(X_1,\dots,X_n))}(A) = \int_A f(x_1, \dots, x_n) d\lambda$$

para todo $A \in \mathcal{B}(\mathbb{R}^n)$. Queremos probar que,

$$\mathbb{P}(X_1 \leq x_1, \dots, X_n \leq x_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq x_i)$$

para todo $x_i \in \mathbb{R}$. Esto se puede hacer debido a que la colección $\{(-\infty, x] : x \in \mathbb{R}\}$ es un π-sistema contenido en $\mathcal{B}(\mathbb{R})$ un λ-sistema, luego por el teorema $\pi - \lambda$, se tiene que,

$$\sigma(\{(-\infty, x] : x \in \mathbb{R}\}) = \mathcal{B}(\mathbb{R})$$

Observemos que,

$$\begin{aligned}
\mathbb{P}_{(X_1,\dots,X_n)} \underbrace{\left(\prod_{i=1}^{n} (-\infty, x_i] \right)}_{=A} &= \int_A f(x_1, \dots, x_n) d\lambda \\
&= \int_A g_1(x_1) \dots g_n(x_n) d\lambda \\
&= \prod_{i=1}^{n} \int_{(-\infty, x_i]} g_i(x_i) d\lambda
\end{aligned}$$

usando fubini. Notemos que $g_i(x_i)$ son integrables, puesto que si no lo fueran, entonces f no sería integrable, siendo una contradicción ya que por hipótesis la integral de f sobre $\mathbb{R}^n$ debe valer 1. Entonces cada integral de $g_i(x)$ está bien definida y es finita, sin embargo, no necesariamente son la densidad de los X_i, y para resolver este problema haremos un truco. Como g_i es integrable, definimos,

$$C_i := \int_{\mathbb{R}} g(x_i) d\lambda < \infty$$

Notemos que $C_i \neq 0$ porque si fuera $C_i = 0$ para algún $i = 1, \dots, n$; entonces,

$$1 = \mathbb{P}_{(X_1,\dots,X_n)}(\mathbb{R}^n) = \prod_{i=1}^{n} C_i = 0$$

siendo imposible. Multiplicando por un 1 conveniente, obtenemos que,

$$\begin{aligned}\mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) &= \frac{\prod_{i=1}^n C_i}{\prod_{i=1}^n C_i}\left(\prod_{i=1}^n \int_{(-\infty,x_i]} g_i(x_i)d\lambda\right) \\ &= \prod_{i=1}^n C_i \left(\prod_{i=1}^n \int_{(-\infty,x_i]} \frac{g_i(x_i)}{C_i}d\lambda\right)\end{aligned}$$

Pero como vimos, necesariamente,

$$\prod_{i=1}^n C_i = 1$$

Por tanto,

$$\mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) \prod_{i=1}^n \int_{(-\infty,x_i]} \frac{g_i(x_i)}{C_i}d\lambda$$

Para concluir que X_i tiene densidad g_i/C_i vamos a realizar un truco. Tomando $x_2 = \cdots = x_n = \infty$ (considerando que X_i no lo alcanza), se tiene que,

$$\begin{aligned}\mathbb{P}(X_1 \leq x_1) &= \mathbb{P}(X_1 \leq x_1, X_2 \in \mathbb{R}, \ldots, X_n \in \mathbb{R}) \\ &= \int_{(-\infty,x_1]} \frac{g_1(x_1)}{C_1}d\lambda\end{aligned}$$

Probando que X_1 tiene densidad g_1/C_1, de forma análoga se puede hacer para todo $i = 1, \ldots, n$. Finalmente concluimos que para todo $x_i \in \mathbb{R}$ con $i = 1, \ldots, n$; se tiene que,

$$\mathbb{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) = \prod_{i=1}^n \mathbb{P}(X_i \leq x_i)$$

Como queriamos probar.

P2: Sean f una función integrable Lebesgue en $[0, 1]$ y $\{U_k\}_{k\in\mathbb{N}}$ una sucesión de variables aleatorias independientes con distribución $\mathcal{U}([0, 1])$. Definamos para cada $n \in \mathbb{N}$

$$I_n := \frac{f(U_1) + \cdots + f(U_n)}{n}$$

Demostrar que $I_n \to \int_0^1 f(x)dx$ en probabilidad.

Sol: Usaremos la ley débil de los grandes números. Notemos algunas cosas. Que U_k tenga distribución $\mathcal{U}([0, 1])$ significa que su densidad es de la forma,

$$f_k(x) = \begin{cases} 0, & x < 0 \\ 1, & x \in [0, 1] \\ 0, & x > 1 \end{cases}$$

Es decir, $f_k : \mathbb{R} \to [0,1]$ es Riemann-integrable, por lo que su integral de Lebesgue está bien definida y se cumple que,

$$\int_{\mathbb{R}} f_k d\lambda = \int_0^1 f_k(x)dx$$

De aquí vemos que,

$$\begin{aligned}\mathbb{P}_{U_i}((-\infty,y]) &= \int_{(-\infty,y]} f_k d\lambda \\ &= \int_0^y f_k(x)dx = \begin{cases} 0, & y < 0 \\ y, & y \in [0,1] \\ 1, & y > 1 \end{cases}\end{aligned}$$

Lo mejor es que $\mathbb{P}_{U_i}$ coincide con el espacio de probabilidad $([0,1], \mathcal{B}([0,1],\lambda))$ **se podrá hacer esto¿?**(pensando en que fuera de $[0,1]$ se anula λ), por lo que,

$$\mathbb{P}_{U_i}((-\infty,y]) = y = \lambda([0,y])$$

Ahora, para usar la ley débil de los grandes números, necesitamos ver que $f(U_k) : \mathbb{R} \to \mathbb{R}$ forma una colección de variables aleatorias independientes e identicamente distribuidas.

Lema: *Sobre un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Sea $\{X_n\}_{n\in\mathbb{N}}$ variables aleatorias independientes e identicamente distribuidas. Sea $f : \mathbb{R} \to \mathbb{R}$ medible, entonces $\{f(X_n)\}_{n\in\mathbb{N}}$ es una colección de variables aleatorias independientes e identicamente distribuidas.*

Dem: Debemos probar independencia y que son identicamente distribuidas.

- **Independencia:** Probemos para dos variables aleatorias. Sean $X_1, X_2 : \Omega \to \mathbb{R}$ variables aleatorias independientes, entonces para todo $A_1, A_2 \in \mathcal{B}(\mathbb{R})$ se tiene que $f(X_i)$ es medible (composició de medibles) y que.

$$\begin{aligned}\mathbb{P}(f(X_1) \in A_1, f(X_2) \in A_2) &= \mathbb{P}(X_1 \in f^{-1}(A_1), X_2 \in f^{-1}(A_2)) \\ &= \mathbb{P}(X_1 \in f^{-1}(A_1))\mathbb{P}(X_2 \in f^{-1}(A_2)) \\ &= \mathbb{P}(f(X_1) \in A_1)\mathbb{P}(f(X_2) \in A_2)\end{aligned}$$

  donde claramente $f^{-1}(A_1), f^{-1}(A_2) \in \mathcal{B}(\mathbb{R})$ al ser medible. Probando que $f(X_1), f(X_2)$ son independientes.

- **Igual distribución:** Consideremos X_1, X_2 identicamente distribuidos. Sea $A \in \mathcal{B}(\mathbb{R})$, luego por definición de distribución,

$$\begin{aligned}\mathbb{P}_{f(X_1)}(A) &= \mathbb{P}(f(X_1) \in A) \\ &= \mathbb{P}(X_1 \in f^{-1}(A)) \\ &= \mathbb{P}_{X_1}(f^{-1}(A)) \\ &= \mathbb{P}_{X_2}(f^{-1}(A)) \\ &= \mathbb{P}(X_2 \in f^{-1}(A)) \\ &= \mathbb{P}(f(X_2) \in A) \\ &= \mathbb{P}_{f(X_2)}(A)\end{aligned}$$

Por tanto $f(X_1), f(X_2)$ son identicamente distribuidos.

Probando el lema. ■

Estudiemos la esperanza, por definición sabemos que,

$$\mathbb{E}(X) = \int_\Omega X(\omega) d\mathbb{P}$$

Sin embargo, del problema P2 de la ayudantía 2 demostramos que si $f : \mathbb{R} \to \mathbb{R}$ es medible, entonces,

$$\mathbb{E}(f(X)) = \int_\Omega f \circ X d\mathbb{P} = \int_\mathbb{R} f d\mathbb{P}_X$$

Esto nos dice que podemos ignorar el espacio Ω y solamente trabajar en $\mathbb{R}$ con una medida en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Veamos que $\mathbb{E}(|f(U_k)|) < \infty$ y en efecto, como f es $L^1(\mathbb{R})$ se tiene que,

$$\mathbb{E}(|f(U_k)|) = \int_\mathbb{R} |f(u_k)| d\lambda(u_k) < \infty$$

Obteniendo todos los requisitos para aplicar la ley débil. Por tanto se tiene que,

$$\frac{S_n}{n} \longrightarrow \mathbb{E}(f(U_1))$$

en probabilidad, donde $S_n = f(U_1) + \cdots + f(U_n)$. Observemos que,

$$\begin{aligned} \mathbb{E}(f(U_1)) &= \int_\mathbb{R} f d\lambda \\ &= \int_0^1 f(x) dx \end{aligned}$$

Como queriamos probar, (ver que $f_k = f$ al ser indeticamente distribuidos).

P3: Sea $\{X_{(n,k)}\}_{n \geq 1, 1 \leq k \leq n}$ un arreglo triangular de variables aleatorias en $L^1(\Omega, \mathcal{F}, \mathbb{P})$. Pruebe que las proposiciones siguientes son equivalentes:

a)

$$\frac{1}{n} \sum_{k=1}^n (X_{n,k} - \mathbb{E}(X_{n,k})) \xrightarrow{P} 0$$

b)

$$\lim_{n \to \infty} \mathbb{E}\left[\frac{(\sum_{k=1}^n (X_{n,k} - \mathbb{E}(X_{n,k}))^2)}{n^2 + (\sum_{k=1}^n (X_{n,k} - \mathbb{E}(X_{n,k}))^2} \right] = 0$$

Ayudantía 6

P1: Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes. Muestre que,

$$\sup_{n\in\mathbb{N}}\{X_n\} < \infty \iff \sum_{n\in\mathbb{N}} \mathbb{P}(X_n > A) < \infty, \text{ para algún } A \in \mathbb{R}$$

Sol: Supongamos que la serie las probabilidades $\mathbb{P}(X_n > A)$ es finita. Entonces por el primer lema de Borel-Cantelli se cumple que,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 0$$

donde $A_n := \{X_n > A\}$. Entonces se cumple que,

$$\mathbb{P}\left(\liminf_{n\to\infty} A_n\right) = 1$$

Ahora si $w \in \liminf A_n$, entonces $X_k(w) \leq A$ para todo $k \geq n$ para algún n, esto implica que,

$$X_n(\omega) \leq C$$

para todo $n \in \mathbb{N}$ para alguna constante C, por tanto $w \in \{\sup X_n < \infty\}$, esto implica que,

$$\mathbb{P}\left(\sup_{n\in\mathbb{N}} X_n < \infty\right) = 1$$

Por lo tanto,

$$\sup_{n\in\mathbb{N}} X_n < \infty$$

Casi seguramente.

Supongamos que,

$$\sum_{n\in\mathbb{N}} \mathbb{P}(X_n > A) = \infty$$

Para todo $A \in \mathbb{R}$. Observemos que el evento $A_n := \{X_n > A\}$ es independiente, entonces por el segundo lema de Borel-Cantelli se cumple que,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 1$$

Ahora, definimos,

$$B_m := \limsup_{n\to\infty}\{X_m > m\}$$

Entonce,

$$B := \bigcap_{m\in\mathbb{N}} B_m \subseteq \{\sup X_n = \infty\}$$

Entonces,

$$\mathbb{P}\left(\sup_{n\in\mathbb{N}} X_n = \infty\right) = 1$$

Luego por contrarecíproba se cumple el resultado.

P2: Consideremos el espacio de secuencias $\Omega = \{0,1\}^{\mathbb{N}}$, o sea, el conjunto cuyos elementos son de la forma $\omega = \{\omega_n\}_{n\in\mathbb{N}}$ con $\omega_n \in \{0,1\}$. Definimos la σ-álgebra de producto:

$$\mathcal{F} = \sigma(\omega_n : n \in \mathbb{N})$$

Muestre que:

- El conjunto A =obtener infinitas caras. Pertenece a la σ-álgebra producto,
- Se tiene que $\mathbb{P}(A) = 1$.

P3: Sean $\{X_n\}_{n\in\mathbb{N}} \longrightarrow X, Y$ en probabilidad. Muestre que $X = Y$ casi seguramente.

Sol: Se cumple que para todo $\varepsilon_1, \varepsilon_2 > 0$ se tiene que,

$$\mathbb{P}(|X_n - X| > \varepsilon_1) \overset{n\to\infty}{\longrightarrow} 0$$
$$\mathbb{P}(|X_n - Y| > \varepsilon_2) \overset{n\to\infty}{\longrightarrow} 0$$

Observemos lo siguiente,

$$\begin{aligned}\mathbb{P}(X = Y) &= \mathbb{P}(|X_n - X| > \varepsilon, X = Y) + \mathbb{P}(|X_n - X| \leq \varepsilon, X = Y)\\ &= \mathbb{P}(|X_n - X| > \varepsilon, X = Y) + \mathbb{P}(|X_n - Y| \leq \varepsilon, X = Y)\\ &= \mathbb{P}(|X_n - X| > \varepsilon, X = Y) + 1 - \mathbb{P}(|X_n - Y| > \varepsilon, X \neq Y)\end{aligned}$$

Si,

$$\{|X_n - X| > \varepsilon, X = Y\} \subseteq \{|X_n - X| > \varepsilon\}$$

Entonces,

$$\mathbb{P}(|X_n - X| > \varepsilon, X = Y) \overset{n\to\infty}{\longrightarrow} 0$$

De forma análoga para el conjunto $\{|X_n - Y| > \varepsilon, X \neq Y\}$, entonces,

$$\lim_{n\to\infty} \mathbb{P}(X = Y) = 1$$

Por tanto $\mathbb{P}(X = Y) = 1$, es decir, $X = Y$ casi seguramente.

Algunas Convergencias

En esta parte veremos algunos resultado interesantes del libro **A course in probability theory** por Kai Lai Chung. Estudiartemos el capítulo 4.

Definición 1: *Sea $\{X_n\}$ una sucesión de variables aleatorias. Decimos que la sucesión converge casi seguramente si y sólo si existe un conjunto Ω_0 tal que,*

$$\lim_{n\to\infty} X_n(\omega) = X(\omega)$$

finito para todo $\omega \in \Omega_0^c$ y tal que $\mathbb{P}(\Omega_0) = 0$.

Teorema 2: *La secuencia $\{X_n\}$ casi seguramente a X si y sólo si para todo $\varepsilon > 0$ se tiene que,*

$$\lim_{m\to\infty} \mathbb{P}(|X_n - X| < \varepsilon, \text{ para todo } n \geq m) = 1$$

o equivalente,

$$\lim_{m\to\infty} \mathbb{P}(|X_n - X| > \varepsilon, \text{ para algún } n \geq m) = 0$$

Dem: Supongamos que $\{X_n\}$ converge casi seguramente y sea $\Omega_0 = \Omega \setminus N$ donde N es despreciable. Para $m \geq 1$ definimos,

$$A_m(\varepsilon) := \bigcap_{n=m}^{\infty} \{|X_n - X| \leq \varepsilon\}$$

Entonces es creciente. Para cada ω_0, la sucesión $\{X_n(\omega_0)\}$ converge a $X(\omega_0)$, esto implica que para todo $\varepsilon > 0$ existe $m(\omega_0, \varepsilon)$ tal que si $n \geq m(\omega_0, \varepsilon)$, entonces,

$$|X_n(\omega_0) - X(\omega_0)| \leq \varepsilon$$

En partícular se cumple,

$$\Omega_0 \subset \bigcup_{m=1}^{\infty} A_m(\varepsilon)$$

Por tanto,

$$\lim_{m\to\infty} \mathbb{P}(A_m(\varepsilon)) = 1$$

que es equivalente a lo que queriamos probar.

Probemos la equivalencia. Supongamos que,

Interrogación 1

P1: Sea $\{X_i\}_{i\in\mathbb{N}}$ una sucesión de variables aleatroias i.i.d. con $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = 0) = 1/2$. Definamos I_n como la cantidad de índices para los cuales las variables aleatorias asumen el valor 1 apartir de n. Es decir $I_n = k$ si y sólo si $X_n = \cdots = X_{n+k-1} = 1$ y $X_{n+k} = 0$. Sea $r_n \in \mathbb{R}$ con $\sum_{n\in\mathbb{N}} 2^{-r_n} < \infty$. Demuestre que $\mathbb{P}(I_n \geq r_n, i.o) = 0$.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Observemos I_n se puede escribir en función de las variables aleatorias, en particular,

$$I_n(\omega) = \sum_{i=0}^{S(n,\omega)} X_{n+i}(\omega)$$

donde $S(n, \omega) = \text{ínf}\{l \geq 0, X_{n+l}(\omega) = 0\}$. De esta forma, si $S(n, \omega) = 0$ entonces,

$$I_n(\omega) = \sum_{i=0}^{0} X_{n+i}(\omega) = X_n(\omega)$$

Y si $S(n, \omega) = k$, entonces,

$$\begin{aligned} I_n(\omega) &= X_n(\omega) + \cdots + X_{n+k-1}(\omega) + X_{n+k}(\omega) \\ &= k \cdot 1 + 0 \end{aligned}$$

Por tanto $I_n : \Omega \to \mathbb{R}$ es una función medible al ser suma de medibles. Definimos el conjunto,

$$A_n := \{\omega \in \Omega : I_n(\omega) \geq r_n\}$$

Observemos que A_n es medible puesto que $A_n = I_n^{-1}([r_n, \infty))$. Por definición se tiene que,

$$\mathbb{P}(A_n) = \mathbb{P}\left(\left\{\omega \in \Omega : \sum_{i=1}^{S(n,\omega)} X_{n+i}(\omega) \geq r_n\right\}\right)$$

Aquí necesariamente $S(n, \omega) \geq [r_n] + 1$ para que $I_n(\omega) = [r_n] + 1 \geq r_n$. Entonces tenemos la siguiente descomposición,

$$\begin{aligned} \mathbb{P}(A_n) &= \mathbb{P}\left(\left\{\bigcup_{k=[r_n]}^{\infty} \{\omega \in \Omega : X_n = 1, X_{n+1} = 1, \ldots, X_{n+k} = 1\right\}\right) \\ &\leq \sum_{k=[r_n]}^{\infty} \mathbb{P}(X_n = 1) \ldots \mathbb{P}(X_{n+k} = 1) \\ &= \sum_{k=[r_n]}^{\infty} \frac{1}{2^{k-1}} \\ &= \frac{1}{2^{[r_n]-2}} \end{aligned}$$

Luego,

$$\begin{aligned}\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) &\leq \sum_{n\in\mathbb{N}} \frac{1}{2^{[r_n]-2}} \\ &= 4\sum_{n\in\mathbb{N}} \frac{1}{2^{[r_n]}} < \infty\end{aligned}$$

Luego por el teorema de Borel-Cantelli, se tiene que,

$$0 = \mathbb{P}(\limsup_{n\to\infty} A_n) = \mathbb{P}(I_n \geq r_n, i.o)$$

Como queriamos probar.

P2: Demuestre que si $X_n \longrightarrow X$ casi seguramente si para todo $\varepsilon > 0$ existe $n \in \mathbb{N}$ tal que, $\mathbb{P}(|X_k - X| < \varepsilon \text{ para } n \leq k \leq m) > 1 - \varepsilon$ para todo $m > n$.

Sol: Sea

P3: Sea $f : [0,1] \to \mathbb{R}$ una función medible con $\int_0^1 |f(x)|dx < \infty$. Sea $\{U_i\}_{i\in\mathbb{N}}$ una sucesión de variables aleatorias i.i.d. uniformemente distribuidas en $[0,1]$. Demuestre que $\frac{1}{n}\sum_{i=1}^n f(U_i)$ converge a $\int_0^1 f(x)dx$ en probabilidad.

Sol: Vamos a usar la ley débil de los grandes números sobre la sucesión $\{f(U_i)\}_{i\in\mathbb{N}}$. Debemos ver que es una sucesión de variables aleatorias i.i.d. con esperanza finita.

- **Variables aleatorias:** Tenemos que $f(U_i) : \Omega \to \mathbb{R}$ es una composición de medibles, entonces $f(U_i)$ es una variable aleatorias.
- **Independientes:** Debemos probar que para cualquier subcolección finita $I = \{1, \ldots, n\} \subseteq \mathbb{N}$ se tiene que,

$$\mathbb{P}(f(U_1) \in A_1, \ldots, f(U_n) \in A_n) = \prod_{i=1}^n \mathbb{P}(f(U_i) \in A_i)$$

para todo $A_i \in \mathcal{B}(\mathbb{R})$ con $i \in I$. Para ello basta estudiar el caso con dos variables aleatorias. Para $f(U_1), f(U_2)$ y para $A_1, A_2 \in \mathcal{B}(\mathbb{R})$ se tiene que,

$$\begin{aligned}\mathbb{P}(f(U_1) \in A_1, f(U_2) \in A_2) &= \mathbb{P}(\{\omega \in \Omega : f(U_1(\omega)) \in A_1, f(U_2(\omega)) \in A_2\}) \\ &= \mathbb{P}(\{\omega \in \Omega : U_1(\omega) \in f^{-1}(A_1), U_2(\omega) \in f^{-1}(A_2)\}) \\ &= \mathbb{P}(U_1 \in f^{-1}(A_1), U_2 \in f^{-1}(A_2)) \\ &= \mathbb{P}(U_1 \in f^{-1}(A_1))\mathbb{P}(U_2 \in f^{-1}(A_2)) \\ &= \mathbb{P}(f(U_1) \in A_1)\mathbb{P}(f(U_2) \in A_2)\end{aligned}$$

Esto se puede hacer para cualquier subcolección $I \subseteq \mathbb{N}$, entonces $\{f(U_i)\}_{i\in\mathbb{N}}$ son independientes.

- **Identicamente distribuidos:** Debemos probar que para cualquier subcolección finito $I = \{1, \ldots, n\} \subseteq \mathbb{N}$ se tiene que $\{f(U_i)\}_{i \in I}$ son identicamente distribuidos. Probemos el caso para dos variables aleatorias. Para $A \in \mathcal{B}(\mathbb{R})$ se tiene que,

$$\begin{aligned} \mathbb{P}_{f(U_1)}(A) &= \mathbb{P}(f(U_1) \in A) \\ &= \mathbb{P}(U_1 \in f^{-1}(A)) \\ &= \mathbb{P}_{U_1}(f^{-1}(A)) \\ &= \mathbb{P}_{U_2}(f^{-1}(A)) \\ &= \mathbb{P}(U_2 \in f^{-1}(A)) \\ &= \mathbb{P}(f(U_2) \in A) \\ &= \mathbb{P}_{f(U_2)}(A) \end{aligned}$$

  Luego $f(U_1), f(U_2)$ son identicamente distribuidos. Luego para cualquier subcolección finita $I \subseteq \mathbb{N}$ se tiene que son identicamente distribuidos.

- **Esperanza finita:** Como $f \in L^1(\Omega)$ entonces,

$$\begin{aligned} \mathbb{E}(|f(U_i)|) &= \int_{\mathbb{R}} |f(u_i)| d\mathbb{P}_{U_i}(u_i) \\ &= \int_{\mathbb{R}} |f(u_i)| f_i(u_i) d\lambda(u_i) \end{aligned}$$

  donde f_i es la función densidad de U_i definido de la siguiente forma:

$$f_i(x) = \begin{cases} 0, & x \notin [0,1] \\ 1, & x \in [0,1] \end{cases} = \mathbb{1}_{[0,1]}(x)$$

  Luego,

$$\mathbb{E}(|f(U_i)|) = \int_0^1 |f(u_i)| du_i < \infty$$

  Es decir, la esperanza de $f(U_i)$ es finita.

Luego por la ley débil de los grandes números se tiene que,

$$\frac{f(U_1) + \ldots f(U_n)}{n} \longrightarrow E(f(U_i))$$

en probabilidad. Determinemos $E(f(U_i))$. Por definición,

$$E(f(U_i)) = \int_0^1 f(x) dx$$

Por tanto,

$$\frac{f(U_1) + \ldots f(U_n)}{n} \longrightarrow \int_0^1 f(x) dx$$

en probabilidad.

P4: Sean X_i variables aleatorias independientes para $i = 1, \dots, k$. Con distribución exponencial de parámetro 1, es decir, con densindad $f(x) = \exp(-x)\mathbb{1}_{\{x \geq 0\}}(x)$. Encunetra la densidad de $X_1 + \dots + X_k$.

Sol: Sea $(\Omega, \mathcal{F}, \shortmid)$ un espacio de probabilidad donde están definidas las variavbles aleatorias X_i. Determinemos la densidad de $X_1 + X_2$. Observemos que,

$$f(x) = e^{-x}\mathbb{1}_{[0,\infty)}(x)$$

Denotemos f_i la densidad de X_i. Como X_1, X_2 tienen densidad y son independientes, entonces $X_1 + X_2$ tiene densidad dada por:

$$\begin{aligned} f_{(1,2)}(z) &= \int_{\mathbb{R}} f_1(z-x)f_2(x)d\lambda(x) \\ &= \int_{\mathbb{R}} e^{-(z-x)}\mathbb{1}_{[0,\infty)}(z-x)\mathbb{1}_{[0,\infty)}(x)d\lambda(x) \\ &= e^{-z}\int_{[0,\infty)} \mathbb{1}_{[0,\infty)}(z-x)d\lambda(x) \end{aligned}$$

Observemos que,

$$\mathbb{1}_{[0,\infty)}(z-x) = \begin{cases} 0, & z-x<0 \\ 1, & z-x \geq 0 \end{cases} = \begin{cases} 0, & z<x \\ 1, & z \geq x \end{cases} = \mathbb{1}_{[0,z]}(x)$$

Luego,

$$\begin{aligned} \int_{[0,\infty)} \mathbb{1}_{[0,\infty)}(z-x)d\lambda(x) &= \int_{[0,z]} d\lambda(x) \\ &= \int_0^z dx = z \end{aligned}$$

como la constante 1 es R-integrable, por lo tanto,

$$f_{(1,2)}(z) = ze^{-z}\mathbb{1}_{[0,\infty)}$$

Determinemos $X_1 + X_2 + X_3$. Sabemos que si X_1, X_2, X_3 son independientes, entonces $X_1 + X_2, X_3$ también lo son, luego podemos aplicar el argumento anterior, de forma que,

$$\begin{aligned} f_{(1,2,3)}(z) &= \int_{\mathbb{R}} f_3(z-x)f_{(1,2)}(x)d\lambda(x) \\ &= \int_{\mathbb{R}} e^{-(z-x)}xe^{-x}\mathbb{1}_{[0,\infty)}(x)\mathbb{1}_{[0,\infty)}(z-x)d\lambda(x) \\ &= e^{-z}\int_0^z xdx \\ &= z^2e^{-z}\mathbb{1}_{[0,\infty)} \end{aligned}$$

Luego por inducción se puede ver que la densidad de $X_1 + \dots + X_n$ es,

$$\widetilde{f}(z) = z^{n-1}e^{-z}\mathbb{1}_{[0,\infty)}$$

Ayudantía 6

P1: Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes que toman valores en $\mathbb{R}$. Probar que:

$$\mathbb{P}\left(\sup_{n\in\mathbb{N}} X_n < \infty\right) \in \{0,1\}$$

Sol: Usaremos la ley 0,1 de Kolmogorov. Sea la σ-álgebra cola dada por:

$$\mathcal{T} = \bigcap_{k\geq 1} \sigma(X_k, X_{k+1}, \dots)$$

Luego si $A \in \mathcal{T}$, entonces $\mathbb{P}(A) \in \{0,1\}$. Dicho de otra forma, debemos probar que,

$$\left\{\sup_{n\in\mathbb{N}} X_n < \infty\right\} \in \mathcal{T}$$

Antes que nada notemos que,

$$\begin{aligned}
\left\{\omega\in\Omega : \sup_{n\in\mathbb{N}} X_n < \infty\right\} &= \{\omega\in\Omega : \text{existe } C_\omega \in \mathbb{N} \text{ tal que para todo } n\in\mathbb{N}\ X_n < C_\omega\} \\
&= \bigcup_{C_\omega\in\mathbb{N}} \bigcap_{n\in\mathbb{N}} \{X_n < C_\omega\} \in \sigma(X_1, X_2, \dots)
\end{aligned}$$

En particular,

$$\left\{\sup_{n\in k} X_n < \infty\right\} \in \sigma(X_k, X_{k+1}, \dots)$$

Ahora notemos que lo siguiente,

$$\begin{aligned}
\left\{\sup_{n\in\mathbb{N}} X_n < \infty\right\} &= \left\{\sup_{n\geq 2} X_n < \infty\right\} \cap \{X_1 < \infty\} \\
&= \left\{\sup_{n\geq 2} X_n < \infty\right\} \cap \Omega \\
&= \left\{\sup_{n\geq 2} X_n < \infty\right\}
\end{aligned}$$

Inductivamente podemos hacer este proceso. Por lo tanto, el conjunto de los ω con supremo X_n finito está en la σ-cola. Luego por el teorema 0,1 de Kolmogorov se cumple que.

$$\mathbb{P}\left(\left\{\sup_{n\geq 2} X_n < \infty\right\}\right) \in \{0,1\}$$

P2: Demuestre utilizando la ley fuerte de los grandes números, la desigualdad de Jense: Si $\varphi : \mathbb{R} \to \mathbb{R}$ es una función convexa y $X : \Omega \to \mathbb{R}$ variable aleatoria tales que $\mathbb{E}(|X|), \mathbb{E}(|\varphi(X)|) < \infty$. Entonces,

$$\varphi(\mathbb{E}(X)) \leq \mathbb{E}(\varphi(X))$$

Sol: Una función conexa es una función que cumple la siguiente desigualdad:

$$\varphi((1-t)a+tb) \leq (1-t)\varphi(a) + t\varphi(b)$$

para todo $t \in [0,1]$ $a < b$ con $a, b \in \mathbb{R}$.

Sol¿?: Una función convexa es una función dos veces diferenciable con segunda deriva positiva. Usaremos ley fuerte de los grandes números. Si $\{X_n\}_{n\in\mathbb{N}}$ son variables aleatorias i.i.d y $X_1 \in L^1(\Omega)$, entonces,

$$\frac{X_1 + \cdots + X_n}{n} \longrightarrow \mathbb{E}(X_1)$$

casi seguramente. Por el teorema de extensión de Kolgomorov, existe $(X, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad tal que concedible $\{X_n\}_{n\in\mathbb{N}}$ tal que son independientes y $\mathbb{P}_{X_n} = \mathbb{P}_X$ para todo $n \in \mathbb{N}$ de tal forma que,

$$\frac{X_1 + \cdots + X_n}{n} \longrightarrow \mathbb{E}(X_1)$$

casi seguramente.

P3: Sea $X = X_+ - X_-$ una variable aleatoria tal que $\mathbb{E}(X_+) = \infty$ y $\mathbb{E}(X_-) < \infty$. Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes con la distribución de X. Muestre que,

$$\mathbb{P}\left(\frac{1}{n}\sum_{k=1}^{n} X_k = \infty\right) = 1$$

Sol (Revisar): Estudiar Fatou y truncar.

Sea $M \in \mathbb{N}$, definimos,

$$X_{i,M} := \text{mín}\{X_i, M\} = X_i \wedge M \leq M$$

Como estamos en un espacio de probabilidad, se tiene que toda constante es integrable, por tanto,

$$\int_\Omega |X_{i,M}| d\mathbb{P} \leq M$$

Es decir, $X_{i,M} \in L^1(\Omega)$. Además son i.i.d por hipótesis. Sea $H : \mathbb{R} \to \mathbb{R}$ tal que $H(X_i) = X_{i,M}$ con H medible. Luego por la ley fuerte de los grandes números se tiene que,

$$\frac{S_{n,m}}{m} \longrightarrow \mathbb{E}(X_{i,M})$$

casi seguramente. Además,

$$X_{iM} \leq X_i$$

Luego,

$$\begin{aligned}\liminf_{n\to\infty} \frac{S_n}{n} &\geq \liminf_{n\to\infty} \frac{S_{n,m}}{n} \\ &\geq \mathbb{E}(X_{i,M})\end{aligned}$$

Entonces,

$$\liminf_{n\to\infty} \frac{S_n}{n} \geq \mathbb{E}(X_{i,M}^+) - \mathbb{E}(X_{i,M}^-)$$

Tomando $M \to \infty$ obtenemos que,

$$\liminf_{n\to\infty} \frac{S_n}{n} \geq \lim_{M\to\infty} \mathbb{E}(X_{i,M}^+) - \lim_{M\to\infty} \mathbb{E}(X_{i,M}^-)$$

Por el teorema de convergencia monótona se tiene que,

$$\lim_{M\to\infty} \mathbb{E}(X_{i,M}^-) \to \mathbb{E}(X_M^-)$$

Análogamente con $X_{i,M}^+$. Por tanto,

$$\liminf_{n\to\infty} \frac{S_n}{n} \geq \infty$$

Finalmente tenemos que,

$$\frac{S_n}{n} \longrightarrow \infty$$

Casi seguramente, que es lo que queriamos probar.

Ayudantía 8

Vamos a estudiar la generalización del teorema de extensión de Kolmogorov.

Teorema de Kolmogorov: Sea $\{S_\alpha\}_{\alpha\in J}$ espacios polacos (completos y separables) donde J es un conjunto de índices arbitrario. Sea $\{Q_I\}_{I\subset\subset J}$ una familias de medidas definidas en,

$$\left(\prod_{\alpha\in I} S_\alpha, \prod_{\alpha\in I} \mathcal{B}_\alpha\right)$$

(producto de σ-álgebras), donde $\mathcal{B}_\alpha$ es el conjunto Boreleano de S_α pensando que S_α es topológico o métrico. Supongamos que para todo $I \subset\subset J$ se tiene que si $I_1 \subset I_2$ y para todo $B \in \prod_{\alpha\in I_1} \mathcal{B}_\alpha$ se cumple que,

$$Q_{I_2}\left(\left(\prod_{I_2,I_1}\right)^{-1}(B)\right) = Q_{I_1}(B)$$

donde

$$\begin{aligned}\prod_{I_2,I_1} : \prod_{\alpha\in I_2} S_\alpha &\to \prod_{\alpha\in I_1} S_\alpha\\ (\omega_\alpha)_{\alpha\in I_2} &\mapsto (\omega_\alpha)_{\alpha\in I_1}\end{aligned}$$

es la proyección canóninca que va de los índices de I_2 a los de I_1. Por definición de proyección se tiene que,

$$\left(\prod_{I_2,I_1}\right)^{-1}(B) = B \times \prod_{\alpha\in I_2\setminus I_1} S_\alpha$$

Por lo que,

$$Q_{I_2}\left(B \times \prod_{\alpha\in I_2\setminus I_1} S_\alpha\right) = Q_{I_1}(B)$$

Esta condición se le conoce como consistencia, ya que hay consistencia en perspectivas. Entonces el teorema de Kolmogorov nos dice que todo lo anterior se cumple, entonces existe una única medida de probabilidad $\mathbb{P}$ definida en,

$$\left(\prod_{\alpha\in J} S_\alpha, \bigotimes_{\alpha\in J} \mathcal{B}_\alpha\right)$$

tal que,

$$\mathbb{P}\left(\left(\prod_{J,I}\right)^{-1}(B)\right) = Q_I(B)$$

para todo I subconjunto finito de J y para todo $B \in \prod_{\alpha \in I} \mathcal{B}_\alpha$.

P1: En este problema la idea es dar una construcción formal que demuestre la existen de un Random Walk.

Sol: Antes de probar la existencia de un Random Walk, necesitamos definilar. Sea $\{S_n\}_{n\in\mathbb{N}_0}$ una colección de puntos cualesquieras en $\mathbb{Z}^d$ definidos en algún espacio de probabilidad. La colección $\{S_n\}_{n\in\mathbb{N}_0}$ junto a una medida de probabilidad $\mathbb{P}_0$ se le dice un Random Walk en $\mathbb{Z}^d$ que empieza en $0 \in \mathbb{Z}^d$ si,

- Los incrementos $\{S_n - S_{n-1}\}_{n\in\mathbb{N}}$ son independientes.
- $\mathbb{P}_0(S_0 = 0) = 1$.
- $\mathbb{P}_0(S_n - S_{n-1} = \pm e_j) = \frac{1}{2d}$, para todo $j = 1, \ldots, d$.

Vamos a probar que existe en $\mathbb{Z}$ ya que el resto de casos es similar. Sea X_0 una variable aleatoria que toma valor en 0 y tal que $\mathbb{P}(X_0 = 0) = 1$. Podemos definir X_1 una variable aleatoria donde de X_0 se mueve una unidad a la izquierda o la derecha, es decir, $X_1 \in \{-1, 1\}$. Definimos una sucesión de incrementos $\{X_i\}_{i\in\mathbb{N}_0}$ i.i.d tales que,

$$\mathbb{P}_{X_i} = \frac{1}{2}\delta_1 + \frac{1}{2}_{-1}$$

Definimos la colección de elementos en $\mathbb{Z}$ por,

$$S_n := \sum_{i=0}^{n} X_i$$

para $n \in \mathbb{N}_0$. Pero ¿existe un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ talque $\{S_n\}_{n\in\mathbb{N}_0}$ existe y $\mathbb{P}(S_0 = 0) = 1$? Y existe en virtud del teorema de extensión de Kolmogorov. Sea μ una medida de probabilidad en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ y sea $S_\alpha = \mathbb{R}$ polaco. Sea el conjunto de índices $J = \mathbb{N}$, entonces para todo $\alpha \in \mathbb{N}$ consideramos la medida producto dada por,

$$Q_I \bigotimes_{\alpha \in I} \mu$$

para todo $I \subset\subset \mathbb{N}$ medida en $(\mathbb{R}^I, \mathcal{B}(\mathbb{R}^I))$. Probemos que las medidas son consistentes. Sean I_1, I_2 subconjuntos finitos de $\mathbb{N}$ tal que $I_1 \subset I_2$ y sea $\mathcal{B}(\mathbb{R}^I)$. Por la independiencia de X_i se tiene que la medidas productos sobre una cantidad finita de medidas se puede separar, entonces tenemos que,

$$\begin{aligned}
Q_{I_2}\left(\left(\prod_{I_2, I_1}\right)^{-1} B\right) &= \left(\bigotimes_{\alpha \in I_2} \mu\right)\left(B \times \mathbb{R}^{I_2 \setminus I_1}\right) \\
&= \left(\bigotimes_{\alpha \in I_1} \mu\right)(B) \cdot \left(\bigotimes_{\alpha \in I_2 \setminus I_1} \mu\right)\left(\mathbb{R}^{I_2 \setminus I_1}\right) \\
&= \left(\bigotimes_{\alpha \in I_1} \mu\right)(B)
\end{aligned}$$

Luego $\{Q_I\}_{I\subset\subset\mathbb{N}}$ es un sistema de medidas consistentes. Luego por la extensión de Kolmogorov existe una única medida $\mathbb{P}$ tal que está definido en $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$ y que,

$$\mathbb{P}\left(B \times \prod_{\alpha J\setminus I}\right) = \left(\bigotimes_{\alpha\in I}\mu\right)(B)$$

para todo $B \in \mathcal{B}(\mathbb{R}^I)$ para todo I subconjunto finito de $\mathbb{N}$.

Si $I = \{\alpha\}$ con $\alpha \in \mathbb{N}$ entonces tenemos que,

$$\mathbb{P}\left(\prod_{\alpha}^{-1} B\right) = \mu(B) \Longleftrightarrow \mathbb{P}_{\pi_\alpha}(B) = \mu(B)$$

para todo $\alpha \in \mathbb{N}$ y para todo $B \in \mathcal{B}(\mathbb{R})$, además, $\{\pi_\alpha\}_{\alpha\in\mathbb{N}}$ son variables aleatorias i.i.d (dado que la distribución son independientes).

P2: Se define la función característica de una variable aleatoria X como la función $\varphi_X : \mathbb{R} \to \mathbb{C}$ tal que,

$$\varphi_X(t) := \mathbb{E}(e^{itX})$$

Muestre que si $\varphi(t)$ es una función característica, entonces $|\varphi(t)|^2$ también lo es.

Sol: Observemos que la función característica está bien definido puesto que,

$$\mathbb{E}(|e^{itX}|) = \int_\Omega d\mathbb{P} = 1$$

dado que $|e^{itX}| = 1$. Por cambio de variables se cumple que,

$$\varphi_X(t) = \int_{\mathbb{R}} e^{itx} d\mathbb{P}_X(x)$$

Observación: La función característica se puede definir en base a una medida μ en $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, puesto μ induce a una variable aleatoria X', luego,

$$\varphi_\mu(t) = \int_{\mathbb{R}} e^{itx} d_\mu(x) = \mathbb{E}(e^{itX'}) = \varphi_{X'}(t)$$

De forma similar, puede existir un espacio de probabilidad $(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \widetilde{P})$ donde hay una variable aleatoria $\widetilde{X}$ tal que,

$$\varphi_{\widetilde{X}} = \varphi_X$$

Por lo que los espacios de probabilidas pueden ser irrelevantes.

Probemos que $|\varphi(t)|^2$ es una función característica. Sabemos que,

$$\varphi_X(t) = \int_{\mathbb{R}} e^{itx} d\mathbb{P}_X(x)$$

Sea $\mu = \mathbb{P}_X$ y sea $(\Omega', \mathcal{F}', \mathbb{P}')$ un espacio de probabilidad tal que existan X_1, X_2 variables aleatorias tales que,

$$\mathbb{P}_{X_1} = \mathbb{P}_{X_2} = \mu$$

con X_1, X_2 independientes. Definimos $Z = X_2 - X_1$ variable aleatoria, entonces,

$$\begin{aligned}
\varphi_Z(t) &= \mathbb{E}(e^{itX_2 - itX_1}) \\
&= \mathbb{E}(e^{itX_2})\mathbb{E}(e^{-tX_1}) \\
&= \varphi_{X_1}(t)\overline{\varphi}_{X_2}(t) \\
&= |\varphi_{X_1}|^2 = |\varphi_X|^2
\end{aligned}$$

Dado que X_1, X_2 son independientes y por hipótesis se cumple lo de la conjugadas. Por tanto $|\varphi_X|^2$ es una función característica.

P3: Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias tales que $X_n \xrightarrow{w} X$. Muestre que si $f : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ es una función continua, entonces $f(X_n) \xrightarrow{w} f(X)$.

Ayudantía 9

P1 (Teorema de Slutsky): Sean $\{X_n\}_{n\in\mathbb{N}}$ y $\{Y_n\}_{n\in\mathbb{N}}$ secuencias de vairables aleatorias tales que $X_n \Rightarrow X$ e $Y_n \Rightarrow c$ con c constante. Demuestre que $X_n + Y_n \Rightarrow X + c$.

Sol: Consideremos el siguiente lema.

Lema: *Sean $\{X_n\}, \{Y_n\}$ dos secuencias de variables aleatorias tales que $X_n \Rightarrow X$ y $|X_n - Y_n| \to 0$ en probabilidad. Entonces $Y_n \Rightarrow X$.*

Dem: La idea es usar Pormantoau. En particular, probaremos que para todo F cerrado se cumple que,

$$\limsup_{n\to\infty} \mathbb{P}(Y_n \in F) \leq \mathbb{P}(X \in F)$$

Definimos,

$$F_\varepsilon := \{x \in \mathbb{R} : d(x, F) \leq \varepsilon\}$$

Claramente es cerrado tal que $F \subseteq F_\varepsilon$. Luego se cumple que,

$$\begin{aligned}
\mathbb{P}(Y_n \in F) &= \mathbb{P}(|X_n - Y_n| > \varepsilon, Y_n \in F) + \mathbb{P}(|X_n - Y_n| \leq \varepsilon, Y_n \in F) \\
&\leq \mathbb{P}(|X_n - Y_n| > \varepsilon) + \mathbb{P}(X_n \in F_\varepsilon) \\
&\xrightarrow{n\to\infty} \mathbb{P}(X \in F_\varepsilon)
\end{aligned}$$

tomando $\varepsilon \downarrow 0$ se concluye que,

$$\limsup_{n\to\infty} \mathbb{P}(Y_n \in F) \leq \mathbb{P}(X \in F)$$

Por tanto $Y_n \Rightarrow X$. ■

Afirmación: *Sea $\{(X_n, Y_n)\}$ vectores aleatorios en $\mathbb{R}^2$ y sea $g(x, y) = x + y$, entonces si $(X_n, Y_n) \Rightarrow (X, Y)$ entonces $g(X_n, Y_n) \Rightarrow (X, Y)$.*

P2. Sea $(\Omega, \mathcal{B}(\Omega))$ un espacio medible y $\{\mathbb{P}_n\}_{n\in\mathbb{N}}$ una secuencia de medidas de probabilidad en este espacio. Demuestre que, dado $\mathcal{A} \subset \mathcal{F}$, si se cumplen:

a) $\mathcal{A}$ es π-sistema.

b) Todo abierto es unión de elementos de $\mathcal{A}$.

Entonces si $\mathbb{P}_n(A) \to \mathbb{P}(A)$ para todo $A \in \mathcal{A}$, podemos concluir que $\mathbb{P}_n \Rightarrow \mathbb{P}$.

P3: Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias tales que $X_n \xrightarrow{w} X$. Muestre que si $f : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ es una función continua, entonces $f(X_n) \xrightarrow{w} f(X)$.

Sol: Vamos a generalizar el teorema del mapeo.

Teorema del mapeo: *Sean $\mathbb{P}_n, \mathbb{P}$ medidas de probabilidad definidos sobre $(S, \mathcal{B}(S))$ donde S es un espacio métrico. Se cumple que si $\mathbb{P}_n \Rightarrow \mathbb{P}$, entonces si $h : (S, \mathcal{B}(S)) \to (S', \mathcal{B}(S'))$ una* *función continua sobre espacios métricos. Entonces* $\mathbb{P}_n h^{-1} \Rightarrow \mathbb{P}h^{-1}$*, donde* $\mathbb{P}_n h^{-1}$ *es una medida de probabilidad definido en* $(S', \mathcal{B}(S'))$ *donde,*

$$\mathbb{P}_n h^{-1}(A) = \mathbb{P}_n(h^{-1}(A))$$

para todo $A \in \mathcal{B}(S')$*.*

Dem: Demostraremos mediante la equivalencia de las esperanza sobre funciones continuas acotadas. Sea $f : (S', \mathcal{B}(S')) \to (S', \mathcal{B}(S'))$ una función continua acotada, entonces, por el cambio de variable se tiene que,

$$\int_{S'} f d\mathbb{P}_n h^{-1} = \int_S f(h) d\mathbb{P}_n$$

Observemos que $f(h)$ es una función continua acotada, y si $\mathbb{P}n \Rightarrow \mathbb{P}$, entonces se cumple que,

$$\int_S f(h) d\mathbb{P}_n \overset{n\to\infty}{\longrightarrow} \int_S f(h) d\mathbb{P}$$

Por lo tanto,

$$\int_{S'} f d\mathbb{P}_n h^{-1} \overset{n\to\infty}{\longrightarrow} \int_{S'} f d\mathbb{P}h^{-1}$$

Es decir, $\mathbb{P}_n h^{-1} \Rightarrow \mathbb{P}h^{-1}$. ■

Existe otra forma de probar esto. Se tiene que $\mathbb{P}_n h^{-1} \Rightarrow \mathbb{P}h^{-1}$ si y sólo si,

$$\mathbb{P}h^{-1}(O) \leq \liminf_{n\to\infty} \mathbb{P}_n h^{-1}(O)$$

para todo O abierto. Pero esto último es equivalente a,

$$\mathbb{P}(h^{-1}(O)) \leq \liminf_{n\to\infty} \mathbb{P}_n(h^{-1}(O))$$

Pero $h^{-1}(O)$ es abierto y como $\mathbb{P}_n \Rightarrow \mathbb{P}$ llegamos al resultado pedido. Ahora para resolver el problema notemos que si $f : \mathbb{R} \to \mathbb{R}$ es una función continua y si $X_n \Rightarrow X$ entonces,

$$\mathbb{P}_n f^{-1} \Rightarrow \mathbb{P}f^{-1}$$

donde $\mathbb{P}_n, \mathbb{P}$ son medidas de probabilidad inducidas por X_n, X respectivamente. Ahora observemos que si $B \in \mathcal{B}(\mathbb{R})$,

$$\begin{aligned}
\mathbb{P}_n f^{-1}(B) &= \mathbb{P}_n(f^{-1}(B)) \\
&= \mathbb{P}(X_n \in f^{-1}(B)) \\
&= \mathbb{P}(f(X_n) \in B) \\
&\Rightarrow \mathbb{P}(f(X) \in B)
\end{aligned}$$

para todo B. Esto es equivalente a decir que $f(X_n) \Rightarrow f(X)$.

Ayudantía 10

P1: Sean $\{X_n\}_{n\in\mathbb{N}}, \{Y_n\}_{n\in\mathbb{N}}$ variables aleatorias positivas independientes e identicamente distribuidas tales que $\mathbb{E}(X_1) = 2, \mathbb{E}(Y_1) = 3$. Probar que:

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n Y_i} = \frac{2}{3}$$

Casi seguramente.

Sol: Observemos lo siguiente,

$$\frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n Y_i} = \frac{\frac{1}{n}\sum_{i=1}^n X_i}{\frac{1}{n}\sum_{i=1}^n Y_i}$$

Ahora, por la ley fuerte de los grandes números se tiene que,

$$\frac{X_1 + \cdots + X_n}{n} \to 2$$
$$\frac{Y_1 + \cdots + Y_n}{n} \to 3$$

Casi seguramente. Supongamos que los límites convergen en $\Omega \setminus \Omega_i$ donde $i = X, Y$ de forma que Ω_i es despreciable. Entonces,

$$\begin{aligned} \Omega \setminus \Omega_X \cap \Omega \setminus \Omega_Y &= \Omega \cap (\Omega_X^c \cap \Omega_Y^c) \\ &= \Omega \setminus (\Omega_X \cup \Omega_Y) \end{aligned}$$

Donde,

$$\mathbb{P}(\Omega_0 := \Omega_X \cup \Omega_Y) \leq \mathbb{P}(\Omega_X) + \mathbb{P}(\Omega_Y) = 0$$

Entonces en Ω_0^c se cumple la convergencia y por tanti,

$$\frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n Y_i} \to \frac{2}{3}$$

Casi seguramente.

Teorema: *Si $X = (X_1, \ldots, X_n)$ con X_n independientes, es un vector aleatoria y $f : \mathbb{R}^n \to \mathbb{R}$ una función integrable, entonces se cumple,*

$$\begin{aligned} \mathbb{E}(f(X_1, \ldots, X_n)) &= \int_{\mathbb{R}^n} f(x_1, \ldots, x_n) d\mathbb{P}(X_1, \ldots, X_n) \\ &= \int_{\mathbb{R}^n} f(x_1, \ldots, x_n) \bigotimes_{i=1}^n d\mathbb{P}_{X_i} \end{aligned}$$

Lema de Slutsky: *Sean $\{X_n\}_{n\in\mathbb{N}}$ variables aleatorias tales que $X_n \Rightarrow X$ y $Y_n \Rightarrow c$. Entonces $X_n + Y_n \Rightarrow X + c$.*

Lema: *Si* $\{X_n\}_{n\in\mathbb{N}}, \{Y_n\}_{n\in\mathbb{N}}$ *son variables aleatorias tales que* $X_n \Rightarrow X$ *y* $d(X_n, Y_n) \to 0$ *en probabilidad, entonces* $Y_n \Rightarrow X$.

Dem Idea Lema de Slutsky: Si mostramos que $(X_n, Y_n) \Rightarrow (X, c)$ y definimos $f(x, y) = x+y$ es continua y por el teorema del mapea,

$$\mathbb{P}(X_n, Y_n)f^{-1} \Rightarrow \mathbb{P}_{(X,c)}f^{-1}$$

Que es equivalente,

$$\mathbb{P}_{f(X_n,Y_n)} \Rightarrow \mathbb{P}_{f(X,c)}$$

Probemos algo más sencillo, probemos que $(X_n, c) \Rightarrow (X, c)$. Entonces por la definición alterna de convergencia débil se cumple $\mathbb{P}_{(X_n,c)} \Rightarrow \mathbb{P}_{(X,c)}$. Sea $f : \mathbb{R}^2 \to \mathbb{R}$ continua y acotada, entonces observemos que,

$$\begin{aligned}
\int_{\mathbb{R}^2} f(x, y) d\mathbb{P}_{(X_n,c)} &= \int_{\mathbb{R}^2} f(x, c) d\mathbb{P}_{(X_n,c)} \\
&= \int_{\mathbb{R}\times\{c\}} f(x, c) d\mathbb{P}_{X_n} \otimes d\delta_c \\
&= \int_{\mathbb{R}} f(x, c) d\mathbb{P}_{X_n} \\
&\xrightarrow{n\to\infty} \int_{\mathbb{R}} f(x, c) d\mathbb{P}_X \\
&= \int_{\mathbb{R}^2} f(x, y) d\mathbb{P}_{(X,c)}
\end{aligned}$$

Por lo tanto $(X_n, c) \Rightarrow (X, c)$.

Finalmente se observa que,

$$\|(X_n, Y_n) - (X_n, c)\| = \|(0, Y_n - c)\| = |Y_n - c| \to 0$$

en probabilidad, dado que $Y_n \Rightarrow c$. Por el lema se concluye que,

$$(X_n, Y_n) \Rightarrow (X, c)$$

Interrogación 2

P1: Para funciones de distribución F y G, definimos,

$$\rho(F,G) = \text{ínf}\{\varepsilon > 0 : F(x-\varepsilon) - \varepsilon \leq G(x) \leq F(x+\varepsilon) + \varepsilon,\ \forall x \in \mathbb{R}\}$$

Demuestre que $\text{lím}_{n\to\infty} \rho(F_n, G) = 0$ implica que F_n converge débilmente a G (sabiendo que ρ es una métrica).

P2: Sean $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias y X un variable aleatorias con densidades $\{f_n\}_{n\in\mathbb{N}}$ y f respectivamente, Demuestre que $\text{lím}_{n\to\infty} f_n(x) = f(x)$ para todo $x \in \mathbb{R}$ implica que $X_n \Rightarrow X$.

De un ejemplo de variables aleatorias X_n, X con densidades $\{f_n\}_{n\in\mathbb{N}}$ y f tales que $X_n \Rightarrow X$ pero que $f_n(x)$ no converge a $f(x)$ para ningún $x \in \mathbb{R}$.

P3: Sea X_n una variable aleatoria binomial con parámetros n y $p = \lambda/n$. Demuestre que X_n converge débilmente a una variable aleatoria Poisson de parámetro λ. **Hint:** Demuestre que $\text{lím}_{n\to\infty} \frac{n!}{(n-k)!n^k} = 1$ para todo k.

P4: Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e identicamente distribuidas a una función de distribución,

$$F(x) = \begin{cases} 0, & x < 1 \\ 1 - x^{-\alpha}, & x \geq 0 \end{cases}$$

con $\alpha > 0$. Sea $M_n = \text{máx}_{i=1,\ldots,n} X_i$. Demuestre que $M_n/n^{1/\alpha}$ converge en distribución a una variable aleatoria con función de distribución.

$$F_0(x) = \begin{cases} 0, & x \leq 0 \\ \exp(-x^{-\alpha}), & x > 0 \end{cases}$$

P5: Calcule,

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 \text{sen}(\sqrt[n]{x_1 x_2 \ldots x_n})dx_1 \ldots dx_n$$

Ayudantía 11

Repaso: Hagamos un pequeño repaso de las funciones caracterísitca. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sea X una variable aleatoria sobre este espacio. Sabemos que la función característica se define por:

$$\varphi_X(t) = \mathbb{E}(e^{itX})$$

Pero no es necesario definir una función característica sobre una variable aleatoria. Si $e^{itx} \in L^1$ entonces por cambio de variable, se tiene que,

$$\varphi_X(t) = \int_{\mathbb{R}} e^{itx} d\mathbb{P}_X$$

donde $\mathbb{P}_X$ es la distribución de X. Observemos que la función característica depende de la medida $\mathbb{P}_X$, de forma que podemos definir sobre una medida μ sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, observando que la función característica a partir de la medida μ es,

$$\varphi_\mu(t) = \int_{\mathbb{R}} e^{itx} d\mu$$

Y si μ induce una variable aleatoria X, se obtiene que,

$$\varphi_\mu(t) = \varphi_X$$

¿Pero, para qué sirve una función característica? Consideremos X una variable aleatoria, entonces se puede definir una función distribución acumulada F_x, una medida de distribución de X $\mathbb{P}_X$ y la función caracterísitca. Lo Interesante es que estas tres característica de X son equivalentes en el sentido que todos se inducen entre sí.

Veamos que la distribución acumulada es equivalente a la distribución. Si tenemos la distribución, es claro que obtenemos a la acumulada y para la otra basta recordar que el conjunto $\mathcal{I} = \{(-\infty, x] : x \in \mathbb{R}\}$ es un semiálgebra que genera $\mathcal{B}(\mathbb{R})$.

Ahora para otra equivalencia entre la característica y la distribución. Es claro que si tenemos una distribución, entonces se puede obtener la función característica y, para obtener la otra dirección basta aplicar el teorema de la inversa y usar unos pequeños trucos. Supongamos que conocemos solamente la función característica, entonces,

$$\mu(a,b) + \frac{1}{2}\mu(\{a,b\}) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t) dt$$

Tomando $a \to \infty$ y $b \to y^+$ se obtiene que,

$$\lim_{a\to\infty} \lim_{b\to y^+} \mu(a,b) + \frac{1}{2}\mu(\{a,b\}) = \mu((a,b])$$

De forma que si conocemos la función característica, podemos determinar la función característica.

P1: Encuentre una constante $a \in \mathbb{R}$ tal que,

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 \arctan\left(\frac{1}{\sqrt{n}}(x_1^2 + \cdots + x_n^2) - a\sqrt{n}\right) dx_1 \dots dx_n$$

Sol: Recordemos el teorema central del límite. Dado $\{X_n\}_{n\in\mathbb{N}}$ una secuencoa de variables aleatorias independientes e identicamente distribuidas tal que $\mathbb{E}(X_1) = \mu$ y $\operatorname{Var}(X_1) = \sigma^2 < \infty$. Entonces,

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \Rightarrow \mathcal{X}$$

Donde $\mathcal{X}$ una variable aleatoria con la distribución normal estándar.

Por extensión de Kolmogorov existe un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ en donde existe una secuencia de variables aleatorias $\{X_n\}_{n\in}\mathbb{N}$ i.i.d para cualquier medida de distribución. Supongamos que tiene distribución la uniforme en $(0,1)$. Entonces $\mathbb{E}(X_1), \operatorname{Var}(X_1) < \infty$. Notemos que,

$$\begin{aligned}
&\mathbb{E}\left(\arctan\left(\frac{1}{\sqrt{n}}(X_1^2 + \cdots + X_n^2) - a\sqrt{n}\right)\right) \\
&= \int_{[0,1]^n} \arctan\left(\frac{1}{\sqrt{n}}(x_1^2 + \cdots + x_n^2) - a\sqrt{n}\right) d\mathbb{P}_{(X_1,\dots,X_n)}
\end{aligned}$$

En particular, arcotangente es integrable en $[0,1]$, de forma que,

$$\begin{aligned}
&\mathbb{E}\left(\arctan\left(\frac{1}{\sqrt{n}}(X_1^2 + \cdots + X_n^2) - a\sqrt{n}\right)\right) \\
&\int_0^1 \cdots \int_0^1 \arctan\left(\frac{1}{\sqrt{n}}(x_1^2 + \cdots + x_n^2) - a\sqrt{n}\right) dx_1 \dots dx_n
\end{aligned}$$

También debido a la R-integrabilidad. Observemos que el argumento del arcotangente no hay un σ, de hecho usaremos el siguiente teorema central del límite variante,

$$\frac{S_n - n\mu}{\sqrt{n}} \Rightarrow N(0, \sigma^2)$$

Aquí podemos determinar $a = 1/3$ y por otro lado, arcotangente es continua y acotada, por lo que,

$$\mathbb{E}\left(\arctan\left(\frac{1}{\sqrt{n}}(X_1^2 + \cdots + X_n^2) - a\sqrt{n}\right)\right) \xrightarrow{n\to\infty} \mathbb{E}(\arctan(N(0,1)))$$

Finalmente,

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 \arctan\left(\frac{1}{\sqrt{n}}(x_1^2 + \cdots + x_n^2) - a\sqrt{n}\right) dx_1 \dots dx_n = \int_0^1 \arctan(x) \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx$$

P2: Demuestre que si $\varphi(t)$ es una función característica, entonces,

$$1 - |\varphi(2t)|^2 \leq 4(1 - |\varphi(t)|^2)$$

para todo $t \in \mathbb{R}$.

Sol: Probaremos de forma general que si φ es una función característica real, entonces,

$$1 - \varphi(2t) \leq 4(1 - \varphi(t))$$

para todo $t \in \mathbb{R}$. Consideremos las siguientes dos desigualdades trigonométrica,

$$\begin{aligned} 1 - (\cos(t)) &= 2\operatorname{sen}^2\left(\frac{t}{2}\right) \\ |\operatorname{sen}(2t)| &= 2|\operatorname{sen}(t)\cos(t)| \leq 2|\operatorname{sen}(t)| \end{aligned}$$

Luego si φ es real, se tiene que,

$$\begin{aligned} 1 - \varphi(t) &= \int_{\mathbb{R}} (1 - \cos(tx))d\mu \\ &\geq \frac{1}{4}\int_{\mathbb{R}} (1 - \cos(2tx))d\mu \\ &= \frac{1}{4}(1 - \varphi(2t)) \end{aligned}$$

Esta última desiguladad se verifica usando las desigualdades trigonométricas, ya que,

$$\begin{aligned} 1 - \cos(tx) = 2\operatorname{sen}^2\left(\frac{tx}{2}\right) && \geq \frac{1}{2}\operatorname{sen}^2(tx) \\ = \frac{1}{4}(1 - \cos(2tx)) \end{aligned}$$

Probando para el caso real. Ahora, para el caso complejo observemos que si $\varphi : \mathbb{R} \to \mathbb{C}$ es una función caracteristica, entonces $|\varphi|^2 : \mathbb{R} \to \mathbb{R}$ es otra función característica, por lo tanto,

$$1 - |\varphi(2t)|^2 \leq 4(1 - |\varphi(t)|^2)$$

P3: Muestre que la esperanza condicional es una contracción en $L^p(\Omega)$.

Sol: Usaremos la desigualdad de Jensen para la esperanza condicional, la cual dice que si $\varphi : \mathbb{R} \to \mathbb{R}$ es convexo y $\varphi(X), X \in L^1$, entonces,

$$\varphi(\mathbb{E}(X|\mathcal{F})) \leq \mathbb{E}(\varphi(X)|\mathcal{F})$$

Queremos mostrar que para todo p se tiene que,

$$|\mathbb{E}(X|\mathcal{F})|^p \leq \mathbb{E}(|X|^p|\mathcal{F})$$

Si $|\cdot|^p$ es convexo, luego por Jensen se obtiene la desigualdad.

P4: Supongamos que $X \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. Demuestre que la esperanza condicional coincide con la proyección ortogonal de X al sub-espacio $L^2(\Omega, \mathcal{F}, \mathbb{P})$.

Ayudantía 11.5

P1:

P2:

Interrogación 3

P1: Sean X e Y variables aleatorias independientes e identicamente distribuidas con $\mathbb{E}(|X|) < \infty$. Demuestre que,

$$\mathbb{E}(X|X+Y) = \mathbb{E}(Y|X+Y) = \frac{X+Y}{2}$$

Sol: Observemos que,

$$\mathbb{E}(X+Y|X+Y) = X+Y$$

debido a que $X+Y$ es $\sigma(X+Y)$-medible. Probemos que.

$$\mathbb{E}(X|X+Y) = \mathbb{E}(Y|X+Y)$$

Claramente ambos son $\sigma(X+Y)$-medible, luego para $A \in \sigma(X+Y)$ se tiene que,

$$\begin{aligned}
\int_A \mathbb{E}(X|X+Y)d\mathbb{P} &= \int_A Xd\mathbb{P} \\
&= \int_{\widetilde{A}} xd\mathbb{P}_X \otimes d\mathbb{P}_Y \\
&= \int_{\widetilde{A}} yd\mathbb{P}_X \otimes d\mathbb{P}_Y \\
&= \int_A Yd\mathbb{P} \\
&= \int_A \mathbb{E}(Y|X+Y)d\mathbb{P}
\end{aligned}$$

Por tanto $\mathbb{E}(X|X+Y) = \mathbb{E}(Y|X+Y)$. Para concluir el resultado basta ver que,

$$\begin{aligned}
\frac{X+Y}{2} &= \frac{1}{2}\mathbb{E}(X+Y|X+Y) \\
&= \frac{1}{2}(\mathbb{E}(X|X+Y) + \mathbb{E}(Y|X+Y)) \\
&= \mathbb{E}(X|X+Y)
\end{aligned}$$

Lo mismo se puede concluir para Y.

P2: Encontrar explícitamente $a \in \mathbb{R}$ tal que el siguiente límite,

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 \cos\left(\frac{x_1^3 + \cdots + x_n^3}{\sqrt{n}} - \sqrt{n}a\right) dx_1 \dots dx_n$$

existe y obtenga tal límite.

Sol: Por la extensión de Kolmogorov podemos asegurar la existencia de la secuencia $\{X_n\}_{n\in\mathbb{N}}$ de variables aleatorias independentes e identicamente distribuidas con distribución Unif$(0,1)$.

Sea $f : \mathbb{R} \to \mathbb{R}$ dada por $f(x) = x^3$, entonces f es medible y luego $\{f(X_n)\}_{n\in\mathbb{N}}$ es una secuencia de variables aleatorias independientes e identicamente distribuidas. En particular,

$$\mathbb{E}(|X_n^3|) = \int_0^1 x^3 dx = \frac{1}{4}$$
$$\mathrm{Var}(|X_n^3|) = \int_0^1 \left(x^3 - \frac{1}{4}\right)^2 dx = \frac{9}{16\cdot 7}$$

Ambos momentos son finitos, por lo que por el teorema central del límite se cumple que,

$$\frac{S_n - n\mu}{\sqrt{n}} \Rightarrow N(0, \sigma^2)$$

donde $S_n = X_1^3 + \cdots + X_n^3$. Observemos lo siguiente,

$$\mathbb{E}\left(\cos\left(\frac{S_n - n\mu}{\sqrt{n}}\right)\right) = \int_0^1 \cdots \int_0^1 \cos\left(\frac{x_1^3 + \cdots + x_n^3}{\sqrt{n}} - \sqrt{n}a\right) dx_1 \ldots dx_n$$

de forma que para que el límite converga, necesariamente $a = \mu = 1/4$. Ahora, por la convergencia débil se tiene que,

$$\mathbb{E}\left(\cos\left(\frac{S_n - n\mu}{\sqrt{n}}\right)\right) \xrightarrow{n\to\infty} \mathbb{E}(\cos(N(0,\sigma^2))) = \int_0^1 \cos(x)\frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx$$

Finalmente,

$$\lim_{n\to\infty} \int_0^1 \cdots \int_0^1 \cos\left(\frac{x_1^3 + \cdots + x_n^3}{\sqrt{n}} - \sqrt{n}a\right) dx_1 \ldots dx_n = \int_0^1 \cos(x)\frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} dx$$

P3: Suponga que $X_n \Rightarrow X$ donde X_n es normal con esperanza 0 y varianza σ_n^2. Demuestre que σ_n^2 es convergente.

P4: Sea X una variable aleatoria y ϕ su función característica. Demuestre que si X admite una densidad, entonces $\phi(t) \to 0$ cuando $|t| \to \infty$.

Sol: Recordemos que para toda función $f : \mathbb{R} \to \mathbb{R}^+$ y para cada $\varepsilon > 0$ existe una función g escalonada tal que,

$$\int_{\mathbb{R}} |f(t) - g(t)| dt < \varepsilon$$

Si X admite una densidad $f(x)$, entonces la función densidad es una función positiva. Sea $\varepsilon > 0$, luego existe una función característica g que satisface la condición anterior, luego,

$$\begin{aligned}
|\phi(t)| &= \left|\int_{\mathbb{R}} e^{itx} f(x) dx\right| \\
&= \left|\int_{\mathbb{R}} e^{itx} f(x) dx - \int_{\mathbb{R}} e^{itx} g(x) dx + \int_{\mathbb{R}} e^{itx} g(x) dx\right| \\
&\leq \int_{\mathbb{R}} |f(x) - g(x)| dx + \left|\int_{\mathbb{R}} e^{itx} g(x) dx\right| \\
&< \varepsilon + \left|\int_{\mathbb{R}} e^{itx} g(x) dx\right|
\end{aligned}$$

Si g es una escalonada, entonces es de la forma,

$$g(x) = \sum_{i=1}^{n} a_i \mathbb{1}_{[n_i, m_i]}$$

con $A \in \mathcal{F}$. Luego,

$$\begin{aligned}
\int_{\mathbb{R}} e^{itx} g(x) dx &= \sum_{i=1}^{n} a_i \int_{n_i}^{m_i} e^{itx} dx \\
&= \sum_{i=1}^{n} a_i \frac{e^{itm_i} - e^{itn_i}}{it}
\end{aligned}$$

Finalmente,

$$|\varphi(t)| \leq \varepsilon + \left| \sum_{i=1}^{n} a_i \frac{e^{itm_i} - e^{itn_i}}{it} \right| \xrightarrow{|t| \to \infty} \varepsilon$$

Como se puede hacer para todo $\varepsilon > 0$ se cumple que $|\varphi(t)| \to 0$ cuando $|t| \to \infty$.

Ayudantía 12

P1: Sea $\{\xi_m\}_{m\in\mathbb{N}}$ una secuencita de variables aleatorias independientes tales que $\mathbb{E}(\xi_m) = 0$ para todo $m \in \mathbb{N}$ y $\mathrm{Var}(\xi_m) = \sigma_m^2 < \infty$. Definimos,

$$s_n := \sum_{m=0}^{n} \sigma_m^2$$

muestre que $S_n^2 - s_n^2$ es una martingala.

P2: Sea $\{Y_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e identicamente distribuidas no negativas con $\mathbb{E}(Y_1) = 1, \mathbb{P}(Y_1 = 1) < 1$.

a Muestre que $X_n := \prod_{m=0}^{n} Y_m$ es una martingala.

b Muestre que $X_n \to 0$ casi seguramente.

P3: (Teorema de Glivenko-Cantelli): Caso Continuo Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e identicamente con distribución μ y definidas sobre $(\Omega, \mathcal{F}, \mathbb{P})$. Considere las siguientes medidas aleatorias:

$$\mu_n(\omega) := \sum_{i=1}^{n} \frac{1}{n} \cdot \delta_{X_i(\omega)}$$

Muestre que:

a) La función de distribución asociada a μ_n viene dada por:

$$F_{\mu_n}(x) = \sum_{i=1}^{n} \frac{1}{n} \cdot \mathbb{1}_{X_i \leq x}$$

b) Demuestre que para cada $x \in \mathbb{R}$ se tiene que:

$$\lim_{n\to\infty} F_{\mu_n}(x) = F_\mu(x)$$

c) Concluya que si F_μ es continua en todo punto, entonces $F_{\mu_n} \Rightarrow F_\mu$ casi seguramente.

Sol:

a)

Ayudantía Final

P1:

Se cumple las siguientes proposiciones:

(a) Si $X_n \xrightarrow{L_1} X$, entonces X_n converge a X en probabilidad.

(b) Si $X_n \xrightarrow{P} X$, entonces existe una subsucesión $\{X_{n_k}\}_{k\in\mathbb{N}}$ tal que X_{n_k} converge a X casi seguramente.

(c) Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia que converge en $L^1(\Omega)$ a X. Muestre que existe una subsucesión $\{X_{n_j}\}_{j\in\mathbb{N}}$ tal que la convergencia es casi segura.

P2: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad ty sea $Y : \Omega \to \mathbb{R}_{\geq 0}$ con $\mathbb{E}(Y) = 1$. Se define la medida de probabilidad sobre $(\Omega, \mathcal{F})$,

$$P^Y(A) := \int_A Y d\mathbb{P}$$

Muestre que si $X : \Omega \to \mathbb{R}$ es una variable aleatoria, entonces X es P^Y-integrable si y sólo si XY es $\mathbb{P}$-integrable y en cuso caso se cumple la desigualdad,

$$\int_\Omega X d\mathbb{P}^Y = \int_\Omega XY d\mathbb{P}$$

P3: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ y sean $X, Y : \Omega \to \mathbb{R}$ variables aleatorias.

(a) Demostrar que (X, Y) es una aplicación $\mathcal{F}|\mathcal{B}(\mathbb{R}^2)$ medible.

(b) Probar que las siguientes afirmaciones son equivalentes:

- Las variables X, Y son independientes.
- $F_{(X,Y)} = F_X \cdot F_Y$.
- $\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y$.
- En $(\mathbb{R}^2, \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}), \mathbb{P}_{(X,Y)})$ las proyecciones canónicas son independientes, con distribución $\mathbb{P}_X$ y $\mathbb{P}_Y$ respectivamente.w

P4: Sea $\{X_n\}_{n\in\mathbb{N}}$ una sucesión de variables aleatorias no correlacionadas tales que $\frac{\mathrm{Var}(X_n)}{n} \to 0$. Mostrar que, si $S_n = X_1 + \cdots + X_n$ y $\mu_n = \mathbb{E}(S_n)$, entonces,

$$\frac{S_n - \mu_n}{n} \longrightarrow 0$$

en probabilidad.

P5: Demuestre que si $\{X_n\}_{n\in\mathbb{N}}$ es una secuencia de variables tales que cada X_n es uniforme en $\{1, \ldots, n\}$, es decir, $\mathbb{P}(X_n = i) = 1/n$ para cada $1 \leq i \leq n$, entonces $X_n/n \xrightarrow{d} X \sim \mathcal{U}[0, 1]$.

P6: Considere $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e identicamente distribuidas a una normal estándar. Muestre que para cualquier $\delta > 0$ se tiene que,

$$\lim_{n\to\infty} \frac{\log\left(P\left(\left|\frac{S_n}{n}\right| \geq \delta\right)\right)}{n} - \frac{\delta^2}{2}$$

¿Qué quiere decir el eunciado? Justifique vía la ley fuerte de los grandes números la intuición del resultado.

P7: Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e identicamente distribuidas con $\mathbb{E}(X_1) = 0$ y $X_1 \in L^2(\Omega)$. Muestre que,

$$\limsup_{n\to\infty} \frac{\sum_{i=1}^{n} X_i}{\sqrt{n}} \overset{\text{c.s}}{=} \infty$$

P8: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad $\mathcal{G} \subseteq \mathcal{F}$ una sub σ-álgebra y $X : (\Omega, \mathcal{F}) \to (\Omega_1, \mathcal{F}_1)$ y sea $Y : (\Omega, \mathcal{F}) \to (\Omega_2, \mathcal{F}_2)$ elementos aleatorios con X independiente de $\mathcal{G}$. Muestre que en este caso, para cualquier $f : (\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ no negativa tenemos que,

$$\mathbb{E}(f(X,Y)|\mathcal{G}) = \int_{\Omega_1} f(x,Y)dP_X(x)$$

P9: Sea $(\Omega, \mathcal{F}, \mathcal{P})$ un espacio de probabilidad, $\mathcal{G} \subseteq \mathcal{F}$ una sub σ-álgebra y X una variable aleatroia no negatica definida sobre este espacio.

(a) Probar que $\{X > 0\}$ está contenido $\{\mathbb{E}(X|\mathcal{G}) \leq 0\}$ salvo probabilidad nula, es decir, $\mathbb{P}(X > 0, \mathbb{E}(X|\mathcal{G}) \leq 0) = 0$.

(b) Mostrar que $\{\mathbb{E}(X|\mathcal{G}) > 0\}$ es, salvo probabilidad nula, el evento $\mathcal{G}$-medible más pequeño que contiene al evento$\{X > 0\}$ salvo probabilidad nula, es decir, para todo $G \in \mathcal{G}$ se tiene que si $\mathbb{P}(X > 0, G^c) = 0$, entonces,

$$\mathbb{P}(\mathbb{E}(X|\mathcal{G}) > 0, G^c) = 0$$

(c) Deducir que si $a < X < b$ con $a, b \in \mathbb{R}$, entonces $\mathbb{P}(a < \mathbb{E}(X|\mathcal{G}) < b) = 1$.

P10: Demostrar que si $\{X_n\}_{n\in\mathbb{N}}$ es una martingala con respecto a la filtración $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$, entonces también es martingala con respecto a su filtración natural.

P11 (Teorema de descomposición de Doob): Sea $\{X_n\}_{n\in\mathbb{N}_0}$ una submartingala. Luego, existe una única descomposición de la forma $X_n = M_n + A_n$, donde $\{M_n\}_{n\in\mathbb{N}_0}$ es una martingala y $\{A_n\}_{n\in\mathbb{N}}$ es un proceso predecible creciente tal que $A_0 = 0$.

6. Guías

En esta sección se compilarán ejercicios sacados directamente de Durret.

Guía 1

P1: Suponga que $(X_1, \dots, X_n)$ tiene desindad $f(x_1, \dots, f_n)$ dada por:

$$\mathbb{P}((X_1, \dots, X_n) \in A) = \int_A f(x)dx$$

para todo $A \in \mathcal{R}^n$. Si $f(x)$ puede ser escrito como $g_1(x_1) \dots g_n(x_n)$ donde $g_m \geq 0$ son medible, entonces $X_1, \dots, X_n$ son independientes. Notar que los g_m no se asumen con densidad.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de medida. Supongamos que $f(x) = g_1(x_1) \dots g_n(x_n)$, notemos que f está bien definido puesto que g_i son medibles y el producto de medibles es un medible. Debemos probar que,

$$\mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \in A_i)$$

para todo $A_i \in \mathcal{R}$ con $i = 1, \dots, n$. Para ello vamos a probar el caso cuando $n = 2$. Tenemos las variables aleatorias $X_1, X_2 : \Omega \to \mathbb{R}$, entonces $f : \Omega^2 \to \mathbb{R}$ es una función que se puede escribir como $f(x, y) = g_1(x)g_2(y)$ donde $g_1, g_2 \geq 0$, es decir, f es no negativa. Por constucción, las integrales:

$$\int_{\Omega^2} f(x, y)\mathbb{P} \otimes \mathbb{P}(dxdy)$$
$$\int_{\Omega} f(x, y)\mathbb{P}(dx)$$
$$\int_{\Omega} f(x, y)\mathbb{P}(dy)$$

Son finitas **justificar**. Luego para todo $A \subseteq \Omega$ las integrables sobre A están bien definidas. Ahora Si g_1, g_2 son no negativas, se cumple que el producto de g_1, g_2 con cualquier indicatriz medibles, es no negativa. Entonces por el teorema de Fubini-Tonelli, para $A_1, A_2 \in \mathcal{R}$ se tiene que,

$$\begin{aligned}
\mathbb{P}((X_1, X_2) \in (A_1, A_2)) &= \mathbb{P}(X_1 \in A_1, X_2 \in A_2) \\
&= \int_{A_1 \times A_2} f(x, y)\mathbb{P} \otimes \mathbb{P}(dxdy) \\
&= \int_{\Omega^2} f(x, y)\mathbb{1}_{A_1 \times A_2}(x, y)\mathbb{P} \otimes \mathbb{P}(dxdy) \\
&= \int_{\Omega} \left(\int_{\Omega} g_1(x)g_2(y)\mathbb{1}_{A_1}(x)\mathbb{1}_{A_2}(y)\mathbb{P}(dx) \right) \mathbb{P}(dy)
\end{aligned}$$

Lo importante es que $g_2(y)\mathbb{1}_{A_2}(y)$ es independiente de x, por lo que podemos sacarlo de la integral y obtener:

$$\int_\Omega g_1(x)\mathbb{1}_{A_1}(x)\mathbb{P}(dx)\int_\Omega g_2(y)\mathbb{1}_{A_2}(y)\mathbb{P}(dy) = \mathbb{P}(X_1 \in A_1)\mathbb{P}(X_2 \in A_2)$$

Es decir, X_1, X_2 son variables aleatoiras independientes. Ahora para n variables aleatorias debemos pensar que todo está bien definido y en que $g_i(x_i)\mathbb{1}_{A_i}(x_i)$ es independiente de las variables x_j con $j \neq i$ y por tanto,

$$\begin{aligned}
\mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n) &= \mathbb{P}((X_1, \dots, X_n) \in \underbrace{(A_1, \dots, A_n)}_{=:A}) \\
&= \int_A f(x)\mathbb{P}(dx) \\
&= \int_A g_1(x_1)\dots g_n(x_n)\mathbb{P}(dx) \\
&= \left(\int_{A_1} g_1(x_1)\mathbb{P}(dx_1)\right)\dots\left(\int_{A_n} g_n(x_n)\mathbb{P}(dx_n)\right) \\
&= \prod_{i=1}^n \mathbb{P}(X_i \in A_i)
\end{aligned}$$

Probando que las variables aleatorias $X_1, \dots, X_n$ son independientes.

P2: Suponga que $X_1, \dots, X_n$ son variables aleatorias que toman valotres en los conjuntos numerables $S_1, \dots, S_n$. Entonces para decir que $X_1, \dots, X_n$ son independientes, vasta decir que para todo $x_i \in S_i$ se tiene que,

$$\mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^n \mathbb{P}(X_i = x_i)$$

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sean $X_1, \dots, X_n : \Omega \to \mathbb{R}$ variables aleatorias. Como X_i toma valores en S_i con S_i numerable, podemos pensar que para todo $A \subseteq \mathbb{R}$ se tiene que $X_i(\omega) \in A$ si y sólo si $X_i(\omega) \in A \cap S_i$. Además, se tiene que,

$$A \cap S_i = \widetilde{S_i}$$

donde $\widetilde{S_i}$ es un subconjunto de S_i. Esto es evidente pero servirá para concluir el resultado. Estudiemos el caso cuando $n = 2$, por lo que debemos probar que para todo $A_1, A_2 \in \mathcal{R}$ se tiene que,

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2) = \mathbb{P}(X_1 \in A_1, X_2 \in A_2)$$

Si $A_i \cap S_i = \emptyset$ para cualquier $i = 1, 2$, entonces,

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2) = 0 = \mathbb{P}(X_1 \in A_1)\mathbb{P}(X_2 \in A_2)$$

Supongamos que $A_i \cap S_i \neq \emptyset$ para todo $i = 1, 2$, entonces,

$$\begin{aligned}\mathbb{P}(X_1 \in A_1, X_2 \in A_2) &= \mathbb{P}(\{\omega \in \Omega : X_1(\omega) \in A_1, X_2(\omega) \in A_2\}) \\ &= \mathbb{P}(\{\omega \in \Omega : X_1(\omega) \in \widetilde{S}_1, X_2(\omega) \in \widetilde{S}_2\})\end{aligned}$$

Supongamos que,

$$\widetilde{S}_i = \{x_{i,j_i} \in S_i : j_i \in I_i \subseteq \mathbb{N}\}$$

para $i = 1, 2$. Entonces,

$$\begin{aligned}\mathbb{P}(X_1 \in A_1, X_2 \in A_2) &= \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = x_{1,j_1}, X_2(\omega) = x_{1,j_2}, i_j \in I_i, i = 1, 2\}) \\ &= \sum_{j_1 \in I_1} \sum_{j_2 \in I_2} \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = x_{1,j_1}, X_2(\omega) = x_{2,j_2}\}) \\ &= \sum_{j_1 \in I_1} \sum_{j_2 \in I_2} \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = x_{1,j_1}\})\mathbb{P}(\{\omega \in \Omega : X_2(\omega) = x_{2,j_2}\}) \\ &= \left(\sum_{j_1 \in I_1} \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = x_{1,j_1}\})\right) \mathbb{P}(\{\omega \in \Omega : X_2(\omega) = x_{2j_2}, j_2 \in I_2\}) \\ &= \mathbb{P}(\{\omega \in \Omega : X_1(\omega) = x_{1j_2}, j_1 \in I_1\})\mathbb{P}(\{\omega \in \Omega : X_2(\omega) = x_{2j_2}, j_2 \in I_2\} \\ &= \mathbb{P}(X_1 \in \widetilde{S}_1, X_2 \in \widetilde{S}_2) \\ &= \mathbb{P}(X_1 \in A_1, X_2 \in A_2)\end{aligned}$$

Probando que X_1, X_2 son independientes. Para probar el caso de n, basta ver que podemos jugar con las sumatorias y como trabajamos con conjunto I_i numerable todo funciona bien. Por lo tanto, para todo $A_i \in \mathcal{R}$ se cumple que,

$$\mathbb{P}(X_1 \in A_1, \ldots, X_n \in A_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \in A_i)$$

P3: Pruebe directamente de la definición que si X e Y son independientes y f y g son funciones medibles, entonces $f(X)$ y $g(Y)$ son independientes.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean $X, Y : \Omega \to \mathbb{R}$ variables aleatorias independientes y sean $f, g : \mathbb{R} \to \widetilde{\Omega}$ funciones medibles sobre el espacio medible $(\widetilde{\Omega}, \widetilde{\mathcal{F}})$. Debemos probar que para todo $A, B \in \widetilde{\mathcal{F}}$ se tiene que,

$$\mathbb{P}(f(X) \in A, g(Y) \in B) = \mathbb{P}(f(X) \in A)\mathbb{P}(g(Y) \in B)$$

Observación 1: Notemos que $f(X) : \Omega \to \widetilde{\Omega}$ es una función medible, porque para todo $A \in \widetilde{\mathcal{F}}$ se tiene que,

$$(f(X))^{-1}(A) = X^{-1}(f^{-1}(A)) \in \mathcal{F}$$

puesto que $f^{-1}(A) \in \mathcal{R}$. Análogamente se puede concluir que $g(Y)$ es una función medible.

Observación 2: La probabilidad $\mathbb{P}(f(X) \in A)$ está bien definida, dado que $f(X) \in A \equiv \{\omega \in \Omega : f(X(\omega)) \in A\} \in \mathcal{F}$.

Por definición notemos que,

$$\begin{aligned}\mathbb{P}(f(X) \in A, g(Y) \in B) &= \mathbb{P}(\{\omega \in \Omega : f(X(\omega)) \in A, g(Y(\omega)) \in B\}) \\ &\overset{\star}{=} \mathbb{P}(\{\omega \in \Omega : X(\omega) \in f^{-1}(A), Y(\omega) \in g^{-1}(B)) \\ &= \mathbb{P}(\{\omega \in \Omega : X(\omega) \in f^{-1}(A)\})\mathbb{P}(\{\omega \in \Omega : Y(\omega) \in g^{-1}(B)\}) \\ &= \mathbb{P}(X \in f^{-1}(A))\mathbb{P}(Y \in g^{-1}(A)) \\ &= \mathbb{P}(f(X) \in A)\mathbb{P}(g(Y) \in B)\end{aligned}$$

Usando independencia de variables aleatorias. Por tanto $f(X), g(Y)$ son independientes.

Nota: La igualdad $\star$ se cumple debido a que,

$$\{\omega \in \Omega : f(X(\omega)) \in A\} = \{\omega \in \Omega : X(\omega) \in f^{-1}(A)\}$$

Para ver esto notemos que $f(X(\omega)) \in A$ es equivalente a decir que $X(\omega) \in f^{-1}(A)$ que además, es equivalente a decir que $\omega \in X^{-1}(f^{-1}(A))$.

P4: Sea $K \geq 3$ un primo y sean X e Y variables aleatorias independientes que son uniformemente distribuidas en $\{0, 1, \ldots, K-1\}$. Para todo $0 \leq n < K$ sea $Z_n := X + nY$ módulo K. Muestre que $Z_0, \ldots, Z_{K-1}$ son independientes a pares.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sea $K \geq 3$ primo. Que una variable aleatoria X sea uniformemente distribuida en $\{0, 1 \ldots, K-1\}$ significa que,

$$\mathbb{P}(X = n) = \frac{1}{K}$$

para todo $n = 0, \ldots, K-1$.

P5: Encuentre cuatro variables aleatorias que toman valores en $\{-1, 1\}$ tales que para cualquier triple de estos sean independientes pero que los cuatros no lo sean. **Hint:** Considerar el producto de las variables aleatorias.

Sol: Sea $(\{\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean $X_1, X_2, X_3 : \mathbb{R} \to \mathbb{R}$ variables aleatorias independientes con distribución $\frac{1}{2}\mathbb{1}_1 + \frac{1}{2}\mathbb{1}_{-1}$. Sea $X_4 = X_1X_2X_3$, verifiquemos que no se cumple la independencia de las cuatros variables aleatorias.

Por definición se tiene que,

$$\mathbb{P}(X_n = i) = \frac{1}{2}\mathbb{1}_1(i) + \frac{1}{2}\mathbb{1}_{-1}(i)$$

para $i = -1, 1$ e $n = 1, 2, 3$. Entonces,

$$\begin{aligned}\mathbb{P}(X_4 = i) &= \mathbb{P}(X_1X_2X_3 = i) \\ &= \frac{1}{2}\mathbb{1}_1(i) + \frac{1}{2}\mathbb{1}_{-1}(i)\end{aligned}$$

Sin embargo,

$$\mathbb{P}(X_1 = X_2 = X_3 = X_4 = 1) = \mathbb{P}(X_1 = 1, X_2 = 1, X_3 = 1) = \frac{1}{8}$$

Pero si las cuatros variales aleatorias fueran independientes, entonces,

$$\mathbb{P}(X_1 = X_2 = X_3 = X_4 = 1) = \frac{1}{2}$$

Tomando otro valor, por tanto las cuatros no son independientes.

P6: Muestre que si X e Y son variables aleatorias independientes (integer-valued), entonces,

$$P(X + Y = n) = \sum_{m} P(X = m)P(Y = n - m)$$

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sean $X, Y : \Omega \to \mathbb{R}$ variables aleatorias independientes. Claramente $X + Y : \Omega \to \mathbb{R}$ es una variable aleatoria. Notemos lo siguiente:

$$\{\omega \in \Omega : X(\omega) + Y(\omega) = n\} = \{\omega \in \Omega : X(\omega) = m, Y(\omega) = n - m, m = \mathbb{Z}\}$$

debido a que X, Y toman valores enteros. Entonces,

$$\begin{aligned}\mathbb{P}(\{\omega \in \Omega : X(\omega) + Y(\omega) = n\}) &= \sum_{m \in \mathbb{Z}} \mathbb{P}(\{\omega \in \Omega : X(\omega) = m, Y(\omega) = n - m\}) \\ &= \sum_{m \in \mathbb{Z}} \mathbb{P}(X = m)\mathbb{P}(Y = n - m)\end{aligned}$$

Usando independencia. Probando el resultado.

P7: Una variable aleatoria X tiene distribución Poisson con parámetro λ si,

$$\mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$$

para $k = 0, 1, 2, \ldots$. Usando el **problema 6** para $X = \text{Poisson}(\lambda), Y = \text{Poisson}(\mu)$ independientes, entonces $X + Y = \text{Poisson}(\lambda + \mu)$.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sean X, Y variables aleatorias independientes con distribución Poisson(λ) y Poisson(μ) respectivamente. Entonces para $s \in \mathbb{Z}$ se cumple que,

$$\mathbb{P}(X + Y = s) = \sum_{m \in \mathbb{Z}} \mathbb{P}(X = m)\mathbb{P}(Y = s - m)$$

pero por hipóstesis se tiene que $m, s = 0, 1, \ldots$; en particular $s \geq m$. Luego, reemplazando por la distribución Poisson, obtenemos,

$$\begin{aligned}
\mathbb{P}(X+Y=s) &= \sum_{m=0}^{s} \frac{e^{-\lambda}\lambda^m e^{-\mu}\mu^{(s-m)}}{(m!)(s-m)!} \\
&= \frac{e^{-(\lambda+\mu)}}{s!}\mu^s \sum_{m=0}^{s} \binom{s}{m}\left(\frac{\lambda}{\mu}\right)^m \\
&= \frac{e^{-(\lambda+\mu)}}{s!}\mu^s \left(\frac{\lambda}{\mu}+1\right)^s \\
&= \frac{e^{-(\lambda+\mu)}(\lambda+\mu)^s}{s!}
\end{aligned}$$

para todo $s = 0, 1, \ldots$. Es decir, $X+Y$ tiene distribución Poisson$(\lambda+\mu)$

P8: Sea X variable aleatoria. Se dice que tiene distribución Binomial(n,p) si,

$$P(X=m) = \binom{n}{m} p^m (1-p)^{n-m}$$

(i) Muestre que si $X = \text{Binomial}(n,p)$ e $Y = \text{Binomial}(m,p)$ son independientes, entonces $X+Y = \text{Binomial}(n+m,p)$.

(ii) Observe el **ejemplo 1.6.12** y use induccion para concluir que la suma de n independientes Bernoulli(p) variables aleatorias es una Binomial(n,p)

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sea $X : \Omega \to \mathbb{R}$ una variable aleatoria.

(i) Sean X, Y con distribución Binomial(n,p) y Binomial(m,p) respectivamente independientes. Sea $X+Y : \Omega \to \mathbb{R}$ una función, es más, es variable aleatoria debido a que es la suma de dos funciones medibles y luego es medible. Notemos que $X+Y$ toma valores $\{0, \ldots, n+m\}$. Sea $s \in \{0, \ldots, n+m\}$, entonces,

$$\begin{aligned}
\mathbb{P}(X+Y=s) &= \mathbb{P}(\{\omega \in \Omega : X(\omega)+Y(\omega)=s\}) \\
&\overset{\star}{=} \mathbb{P}(\{\omega \in \Omega : X(\omega)=l, Y(\omega)=s-l, l=0,1,\ldots,s\}) \\
&= \sum_{l=0}^{\text{mín}(n,s)} \mathbb{P}(\{\omega \in \Omega : X(\omega)=l\})\mathbb{P}(\{\omega \in \Omega : Y(\omega)=s-l\}) \\
&= \sum_{l=0}^{\text{mín}(n,s)} \binom{n}{l} p^l (1-p)^{n-l} \binom{m}{s-l} p^{s-l}(1-p)^{m-(s-l)} \\
&= \sum_{l=0}^{\text{mín}(n,s)} \binom{n}{l}\binom{m}{s-l} p^s (1-p)^{m+n-s} \\
&= p^s(1-p)^{m+n-s} \sum_{l=0}^{s} \binom{n}{l}\binom{m}{s-l}
\end{aligned}$$

Probemos ahora que,

$$\sum_{l=0}^{\text{mín}(n,s)} \binom{n}{l}\binom{m}{s-l} = \binom{n+m}{s}$$

Para ello necesitamos el uso de la distribución hipergeométria. Sea X una variable aleatoria con distribución hipergeometría, es decir, es de la forma,

$$\mathbb{P}(X = x) = \frac{\binom{K}{x}\binom{N-K}{a-x}}{\binom{N}{a}}$$

La esperanza de la hipergeométrica es,

$$\mathbb{E}(X) = ap$$

donde $p = K/N$. Ahora en nuestro contexto tenemos que, $n = K, N-K = m, l = x, s = a$, entonces $N = n + m$ y por tanto,

$$\sum_{l=0}^{\text{mín}(n,s)} \binom{n}{l}\binom{m}{s-l} = ap\binom{n+m}{s}$$

donde $p = n/s$, obteniendo que,

$$\sum_{l=0}^{\text{mín}(n,s)} \binom{n}{l}\binom{m}{s-l} = s$$

faltan detalles. Por lo tanto,

$$\mathbb{P}(X + Y = s) = \binom{n+m}{s} p^s (1-p)^{n+m-s}$$

Es decir, la variable aleatoria $X + Y$ tiene distribución Binomial$(n + m, p)$.

(ii) Debemos probar la siguiente afirmación:

Afirmación: *Sea* $(\Omega, \mathcal{F}, \mathbb{P})$ *un espacio de probabilidad. Sean* $X_1, \ldots, X_n$ *variables aleatorias independientes. Entonces* $S_n := X_1 + \cdots + X_{n-1}$ *es una variable aleatoria independiente de* X_n.

Dem: Vamos a probar por inducción. Consideremos $S_2 = X_1 + X_2$, veamos que es indpendiente de X_3 y en efecto, sean $x, y \in \mathbb{R}$, luego se tiene que debemos estudiar,

$$\mathbb{P}(X_1 + X_2 \leq x, X_3 \leq y)$$

Por otro lado notemos que,

$$\begin{aligned}
\{X_1 + X_2 \leq x\} &= \{\omega \in \Omega : X_1(\omega) + X_2(\omega) \leq x\} \\
&= \{\omega \in \Omega : X_1(\omega) \leq x - X_2(\omega)\} \cap \Omega \\
&= \{\omega \in \Omega : X_1(\omega) \leq x - X_2(\omega)\} \cap \bigcup_{n\in\mathbb{Z}} \{\omega \in \Omega : n-1 \leq X_2(\omega) \leq n\} \\
&= \bigcup_{n\in\mathbb{Z}} \{\omega \in \Omega : X_1(\omega) \leq x - (n-1)\} \cap \{\omega \in \Omega : n-1 \leq X_2(\omega) \leq n\} \\
&= \bigcup_{n\in\mathbb{Z}} X_1^{-1}(-\infty, x-(n-1)] \cap X_2^{-1}[n-1, n]
\end{aligned}$$

Es decir, $\{X_1 + X_2 \leq x\}$ es unión disjuntas de cosas medibles, por tanto tenemos,

$$\begin{aligned}
&\mathbb{P}\left(\bigcup_{n\in\mathbb{Z}} X_1^{-1}(-\infty, x-(n-1)] \cap X_2^{-1}[n-1,n] \cap X_3^{-1}(-\infty, y]\right) \\
&= \sum_{n\in\mathbb{Z}} (X_1^{-1}(-\infty, x-(n-1)] \cap X_2^{-1}[n-1,n] \cap X_3^{-1}(-\infty, y]) \\
&= \left(\sum_{n\in\mathbb{Z}} \mathbb{P}(X_1^{-1}(-\infty, x-(n-1)] \cap X_2^{-1}[n-1,n])\right) \mathbb{P}(X_3 \leq y) \\
&= \mathbb{P}(X_1 + X_2 \leq x)\mathbb{P}(X_3 \leq y)
\end{aligned}$$

Siendo $X_1 + X_2, X_3$ independientes.

Supongamos que se cumple para n, debemos probar que S_n es independiente de X_{n+1}, pero basta notar que,

$$\{X_1 + \cdots + X_n \leq x\} = \bigcup_{n\in\mathbb{N}} (X_1 + \cdots + X_{n-1})^{-1}(-\infty, x-(n-1)] \cap X_n^{-1}[n-1, n]$$

Luego aplicando la hipótesis inductiva de que S_{n-1} es independiente de X_n, se tiene que,

$$\begin{aligned}
&\mathbb{P}(X_1 + \cdots + X_n \leq x, X_{n+1} \leq y) \\
&= \left(\sum_{n\in\mathbb{Z}} \mathbb{P}((X_1 + \cdots + X_{n-1})^{-1}(-\infty, x-(n-1)] \cap X_n^{-1}[n-1,n])\right) \mathbb{P}(X_{n+1} \leq y) \\
&= \mathbb{P}(X_1 + \cdots + X_n \leq x)\mathbb{P}(X_{n+1} \leq y)
\end{aligned}$$

Por la independiente de los $X_1, \ldots, X_{n+1}$. Probando que S_n es independiente de X_{n+1} para todo $n \in N$. ■

Para concluir el resultado consideramos $X_1, \ldots, X_n$ variables aleatorias iid, con distribución Binomial$(1,p)$. Claramente $X_1 + X_2$ tiene distribución Binomial$(2,p)$, supongamos que $X_1 + \cdots + X_{n-1}$ tiene distribución Binomial$(n-1,p)$. Entonces por la afirmación anterior se tiene que S_{n-1} y X_n son independientes, entonces S_n tiene distribución Binomial(n,p).

P9: Sean $X, Y \geq 0$ variables aleatorias independientes con funciones distribución F y G respectivamente. Encuentre la función distribución de XY.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $X, Y \geq 0$ variables aleatorias independientes no negativas. Sea $s \geq 0$, entonces,

$$
\begin{aligned}
\mathbb{P}(XY \leq s) &= \mathbb{E}(\mathbb{1}_{\{XY \leq s\}}) \\
&= \int_{\Omega^2} \mathbb{1}_{\{XY \leq s\}}(\omega) d\mathbb{P}(\omega) \\
&= \int_{\mathbb{R}^2} \mathbb{1}_{\{xy \leq z\}}(x, y) d\mathbb{P}_X \otimes \mathbb{P}_Y \\
&= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{1}_{\{xy \leq z, y \neq 0\}} + \mathbb{1}_{\{0 \leq z, y=0\}} dF(x) \right) dG(y) \\
&= \int_{\mathbb{R}} \mathbb{P}(x \leq z/y) dG(y) + \mathbb{P}(Y = 0) I(z \geq 0) \\
&= \int_{\mathbb{R} \setminus \{0\}} F_(y/z) dG(y) + \mathbb{P}(Y = 0) I(z \geq 0)
\end{aligned}
$$

Guía 2

P1: Sea $X_1, \dots, X_n$ no correlacionados tales que $\mathbb{E}(X_i) = \mu_i$ y $\mathrm{Var}(X_i)/i \to 0$ cuando $i \to \infty$. Sea $S_n = X_1 + \cdots + X_n$ y $\nu_n = \mathbb{E}(S_n/n)$, entonces cuando $n \to \infty$ se tiene que $S_n/n - \nu_n \to 0$ en L^2 y en probabilidad.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sean $X_i : \Omega \to \mathbb{R}$ variables aleatorias no correlacionadas. Sea $\varepsilon > 0$, tenemos que,

$$\begin{aligned}
\mathbb{P}\left(\left|\frac{S_n}{n} - \nu_n\right| \geq \varepsilon\right) &= \mathbb{P}\left(\left|\frac{S_n}{n} - \nu_n\right|^2 \geq \varepsilon^2\right) \\
&\leq \frac{1}{\varepsilon^2}\mathbb{E}\left(\left(\frac{S_n}{n} - \nu_n\right)^2\right) \\
&= \frac{1}{\varepsilon^2 n^2}\mathrm{var}(S_n) \overset{n\to\infty}{\longrightarrow} 0
\end{aligned}$$

Por tanto $S_n/n - \nu_n \longrightarrow 0$ en probabilidad y en L^2.

P2: Integración de Monte Carlo:

(i) Sea f una función medible en $[0,1]$ conm

$$\int_0^1 |f(x)|dx < \infty$$

Sea $U_1, U_2, \dots$ independientes y uniformemente distribuidos en $[0,1]$ y sea,

$$I_n = n^{-1}(f(U_1) + \cdots + f(U_n))$$

Muestre que $I_n \to I \equiv \int_0^1 f dx$ en probabilidad.

(ii) Supongamos que,

$$\int_0^1 |f(X)|^2 dx < \infty$$

Use la desigualdad de Chebyshev para estimar,

$$\mathbb{P}\left(|I_n - I| > \frac{a}{n^{1/2}}\right)$$

P3: Sean $X_1, \dots$ variables aleatorias independientes e identicamente distribuidas con $\mathbb{P}(X_i = -(1^k)k) = C/k^2\log(k)$ para $k \geq 2$ donde C es tal que la suma de las probabilidades es 1. Muetre que $\mathbb{E}(|X_i|) = \infty$, pero hay una constante finita μ tal que $S_n/n \to \mu$ en probabilidad.

Sol: Notemos la siguiente caracterización de la esperanza. Sabemos que,

$$|X| = X\mathbb{1}_{\{|X|=2\}} + X\mathbb{1}_{\{|X|=3\}} + \dots$$

Luego,

$$\begin{aligned}
\mathbb{E}(|X_i|) &= \sum_{k=2}^{\infty} \mathbb{E}\left(X\mathbb{1}_{|X|=k}\right) \\
&= \sum_{k=2}^{\infty} \int_{\Omega} X\mathbb{1}_{\{|X|=k\}} d\mathbb{P} \\
&= \sum_{k=2}^{\infty} k\mathbb{P}(|X|=k) \\
&= \sum_{k=2}^{\infty} \frac{C}{k\log(k)}
\end{aligned}$$

P4: Sean $X_1, X_2, \ldots$ variables aleatorios independientes e identicamente distribuidos con $\mathbb{P}(X_i > x) = e/x\log(x)$ con $x \geq e$. Muestre que $\mathbb{E}(|X_i|) = \infty$, pero hay una secuencia de constantes $\mu_n \to \infty$ tal que $S_n/n - \mu \to 0$ en probabilidad.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean $X_i : \Omega \to \mathbb{R}$ variables aleatorias iid. Observemos que si,

$$\mathbb{P}(X_i > x) = \frac{e}{x\log x}$$

para todo $x \geq e$, entonces,

$$x\mathbb{P}(X_i > x) = \frac{e}{\log x} \overset{x\to\infty}{\longrightarrow} 0$$

Por otro lado,

$$\begin{aligned}
\mathbb{E}(|X_i|) &= \int_{\Omega} |X(\omega)| d\mathbb{P} \\
&= \int_{\mathbb{R}} |x| d\mathbb{P}_{X_i}
\end{aligned}$$

P5:

(i) Muestre que si $X \geq 0$ toma valores enteros, entonces,

$$\mathbb{E}(X) = \sum_{n\geq 1} \mathbb{P}(X \geq n)$$

(ii) Encuentre una expresión similar para $\mathbb{E}(X^2)$.

Sol:

(a) Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y sea $X : \Omega \to \mathbb{R}$ una variable aleatoria no nevativa que toma valores enteros, es decir, $X = k$ para todo $k \in \mathbb{N}_0$. Observemos lo siguiente:

$$X(\omega) = \sum_{n\in\mathbb{N}} \mathbb{1}_{\{X\geq n\}}(\omega)$$

Verifiquemos esta igualdad. Si ω es tal que $X(\omega) = i$, entonces,

$$\begin{gathered} X(\omega) \geq 1 \\ X(\omega) \geq 2 \\ \vdots \\ X(\omega) \geq i \end{gathered}$$

Pero $X(\omega) \ngeq i+1$, entonces,

$$\sum_{n\in\mathbb{N}} \mathbb{1}_{\{X\geq n\}}(\omega) = i$$

Coincidiendo con $X(\omega)$. Ahora aplicando la linealidad de la esperanza, obtenemos,

$$\begin{aligned} \mathbb{E}(X) &= \sum_{n\in\mathbb{N}} \mathbb{E}\left(\mathbb{1}_{\{X\geq n\}}\right) \\ &= \sum_{n\in\mathbb{N}} \mathbb{P}(X \geq n) \end{aligned}$$

(b) El procedimiento es análogo. Notemos que,

$$\begin{aligned} X^2(\omega) &= \left(\mathbb{1}_{\{X\geq n\}}\right)^2 \\ &= \sum_{n_1,n_2\in\mathbb{N}} \mathbb{1}_{\{X\geq n_1\}} \cdot \mathbb{1}_{\{X\geq n_2\}} \end{aligned}$$

Notemos que,

$$\mathbb{1}_{\{X\geq n_1\}} \cdot \mathbb{1}_{\{X\geq n_2\}} = \mathbb{1}_{\{X\geq \text{máx}\{n_1,n_2\}\}}$$

Entonces debemos hacer un arreglo. Notemos que,

$$\mathbb{1}_{\{X\geq 3\}}$$

Puede ser generado por 5 casos $n_1 = 1, n_2 = 3$ o $n_1 = 2, n_2 = 3$ y asi hasta tener lo cinco casos, en general si $n_1 = n$, entonces hay $2n-1$ casos, de forma que,

$$\sum_{n_1,n_2\in\mathbb{N}} \mathbb{1}_{\{X\geq n_1\}} \cdot \mathbb{1}_{\{X\geq n_2\}} = \sum_{n\in\mathbb{N}} (2n-1)\mathbb{1}_{\{X\geq n\}}$$

Finalmente aplicando la esperanza se obtiene,

$$\mathbb{E}(X^2) = \sum_{n\in\mathbb{N}} (2n-1)\mathbb{P}(X\geq n)$$

P6: An unfair "fair game" Consideremos,

$$p_k = \frac{1}{2^k k(k+1)}, \ \ k = 1, , 2 \dots$$
$$p_0 = 1 - \sum_{k \geq 1} p_k$$

De fomar que,

$$\sum_{k=1}^{\infty} 2^k p_k = 1$$

Si $X_1, \dots, X_n$ son independientes e identicamente distribuidas con $\mathbb{P}(X_n = -1) = p_0$ y $\mathbb{P}(X_n = 2^k - 1) = p_k$ para todo $k \geq 1$. Entonces $\mathbb{E}(X_n) = 0$. Sea $S_n = X_1 + \cdots + X_n$, usando el teorema **blabla** con $b_n = 2^{m(n)}$ donde $m(n) = \text{mín}\{m : 2^{-m} m^{-3/2} \leq n^{-1}\}$, para concluir que

$$\frac{S_n}{\frac{n}{\log_2 n}} \to -1$$

en probabilidad.

Guía 3

P1: Sea A_n una secuencia de eventos independientes con $\mathbb{P}(A_n) < 1$ para todo n. Muestre que si,

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = 1$$

implica que,

$$\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = \infty$$

y por tanto $\mathbb{P}(\limsup A_n) = 1$.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Consideremos $B_{n+1} := A_{n+1} \setminus A_n$ con $B_1 = A_1$, claramente $B_{n+1} \in \mathcal{F}$ por lo que es medible. Observemos que se cumple que,

$$\begin{aligned}
\mathbb{P}(B_{n+1}) &= \mathbb{P}(A_{n+1} \cap A_n^c) \\
&= \mathbb{P}(A_{n+1})\mathbb{P}(A_n^c) \\
&= \mathbb{P}(A_{n+1}) - \mathbb{P}(A_n)
\end{aligned}$$

dado que si A_{n+1}, A_n son independientes, entonces A_{n+1}, A_n^c también lo son. Por tanto,

$$\begin{aligned}
\sum_{i=1}^{n} \mathbb{P}(B_i) &= \mathbb{P}(A_n) - \mathbb{P}(A_{n-1}) + \ldots \mathbb{P}(A_2) - \mathbb{P}(A_1) + \mathbb{P}(A_1) \\
&= \mathbb{P}(A_n)
\end{aligned}$$

Por otro lado observemos que si definimos $S_n := \mathbb{P}(B_n) + \cdots + \mathbb{P}(B_1)$ observamos que $S_n \geq 0$ y S_n es creciente. Si ocurre que $S_n = 0$ para todo $n \in \mathbb{N}$, entonces, $\mathbb{P}(A_n) = 0$ para todo $n \in \mathbb{N}$, pero entonces,

$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) \leq \sum_{n\in\mathbb{N}} \mathbb{P}(A_n) = 0$$

siendo una contradicción, por lo tanto exite un $N \in \mathbb{N}$ al que $S_N > 0$, digamos que $S_N = C > 0$, luego se tiene que $S_n \geq C$ para todo $n \geq N$ y esto implica que,

$$\begin{aligned}
\sum_{n\in\mathbb{N}} \mathbb{P}(A_n) &= \sum_{n\in\mathbb{N}} S_n \\
&\geq \sum_{n\geq N} C = \infty
\end{aligned}$$

Luego por el lema de Borel-Cantelli II se cumple que,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 1$$

P2 (Lema de Kochen-Stone): Supongamos que,

$$\sum_{k\in\mathbb{N}} \mathbb{P}(A_k) = \infty$$

Use los ejercicios 1.6.6 y 2.3.1 para mostrar que si,

$$\limsup_{n\to\infty} \frac{\left(\sum_{k=1}^{n} \mathbb{P}(A_k)\right)^2}{\left(\sum_{1\leq j,k\leq n} \mathbb{P}(A_j \cap A_k)\right)} = \alpha > 0$$

entonces $\mathbb{P}(\limsup A_n) \geq \alpha$. El caso $\alpha = 1$ contiene al teorema 2.3.7.

Sol:

`https://proofwiki.org/wiki/Borel-Cantelli_Lemma_10_to_Kochen-Stone_Theorem`

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad.

P3: Sean $X_1, X_2, \ldots$; variables aleatorias i.i.d con distribución F, sea $\lambda_n \uparrow \infty$ y sea,

$$A_n := \left\{ \max_{1\leq m\leq n} X_m > \lambda_n \right\}$$

Muestre que $\mathbb{P}(\limsup A_n) = 0$ o 1 acorde si,

$$\sum_{n\geq 1} (1 - F(\lambda_n))$$

es finita o infinito.

Sol: Observemos algunas cosas. Si $\{X_n\}$ son variables aleatorias independientes e identicamente distribuidas, entonces se cumple que,

$$\{\{X_n > \lambda_n\}\}_{n\in\mathbb{N}}$$

es una secuencia de eventos independientes, puesto que,

$$\begin{aligned}\mathbb{P}(X_n > \lambda_n, X_m > \lambda_n) &= \mathbb{P}(X_n \in (\lambda_n, \infty), X_m \in (\lambda_m, \infty)) \\ &= \mathbb{P}(X_n \in (\lambda_n, \infty))\mathbb{P}(X_m \in (\lambda_m, \infty))\end{aligned}$$

Y esto se puede hacer para cualquier subcolección finita de $\mathbb{N}$. Supongamos que la serie,

$$\sum_{n\in\mathbb{N}} (1 - F(\lambda_n)) = \infty$$

Entonces por definición de la distribución acumulada, se tiene que,

$$(1 - F(\lambda_n)) = \mathbb{P}(X_n > \lambda_n)$$

Luego si son independientes los eventos $\{X_n > \lambda_n\}$, por el segundo lema de Bore-Cantelli, se cumple que,

$$\mathbb{P}\left(\limsup_{n\to\infty}\{X_n > \lambda_n\}\right) = 1$$

Por otro lado,

$$\{X_n > \lambda_n\} \leq A_n$$

Por tanto,

$$1 = \mathbb{P}\left(\limsup_{n\to\infty}\{X_n > \lambda_n\}\right) \leq \mathbb{P}\left(\limsup_{n\to\infty} A_n\right)$$

Es decir,

$$\mathbb{P}\left(\limsup_{n\to\infty} A_n\right) = 1$$

Supongamos ahora que la serie,

$$\sum_{n\in\mathbb{N}}(1 - F(\lambda_n)) < \infty$$

Por el primer lema de Borel-Cantelli se cumple que,

$$\mathbb{P}\left(\limsup_{n\to\infty}\{X_n > \lambda_n\}\right) = 1$$

Por tanto,

$$\mathbb{P}\left(\liminf_{n\to\infty}\{X_n \leq \lambda_n\}\right) = 0$$

Sea $\Omega_0 = \liminf\{X_n \leq \lambda_n\}$ de forma que,

$$\Omega_0 = \{\omega \in \Omega : \text{existe un } n \in \mathbb{N} \text{ tal que para todo } k \geq n \text{ se tiene que } X_n(\omega) \leq \lambda_n\}$$

Nuestro objetivo es probar la siguiente inclusión,

$$\Omega_0 \subseteq \left\{\liminf_{n\to\infty} \max_{1\leq m\leq n} X_m \leq \lambda_n\right\}$$

Sea $\omega \in \Omega_0$, entonces existe $N_1(\omega) \in \mathbb{N}$ tal que $X_n(\omega) \leq \lambda_n$ para todo $n \geq N_1(\omega)$. Observemos que podemos separar el máximo de los $X_m(\omega)$ con $m = 1, \ldots, n$ de la siguiente forma:

$$\max_{1\leq m\leq n} X_m(\omega) = \max\left\{\max_{1\leq m\leq N_1(\omega)} X_m(\omega), \max_{N_1(\omega)<m\leq n} X_m(\omega)\right\}$$

Si $\lambda_n \uparrow \infty$ (creciente no acotado), entonces existe $N_2(\omega) \in \mathbb{N}$ tal que,

$$\max_{1 \leq m \leq N_1(\omega)} X_m(\omega) \leq \lambda_n$$

Para todo $n \geq N_2(\omega)$. De forma general se tiene que si $N_1(m) < m \leq n$, entonces $X_m(\omega) \leq \lambda_m \leq \lambda_n$. Por tanto, sea $N(\omega) := \max\{N_1(\omega), N_2(\omega)\}$, obteniendo que si $n \geq N(\omega)$, se tiene que,

$$\max_{1 \leq m \leq n} X_m(\omega) \leq \lambda_n$$

Encontrando un $N(\omega)$ tal que si $n \geq N(\omega)$ se cumple la desigualdad anterior, es decir, se cumple la inclusión que queriamos probar. Ahora, para concluir observamos que,

$$1 = \mathbb{P}(\Omega_0) \leq \mathbb{P}\left(\liminf_{n\to\infty} \max_{1\leq m\leq n} X_m \leq \lambda_n\right)$$

Por lo tanto,

$$\mathbb{P}\left(\limsup_{n\to\infty} \max_{1\leq m\leq n} X_m \leq \lambda_n\right) = 0$$

Como queriamos probar.

P4: Si X_i tiene distribución Poisson con media 1, entonces S_n tiene distribución Poisson con media n, es decir,

$$\mathbb{P}(S_n = k) = \frac{e^{-n}n^k}{k!}$$

Use la fórmula de Stirling para mostrar que si $(k-n)/\sqrt{n} \longrightarrow x$, entonces,

$$\sqrt{2\pi n}\,\mathbb{P}(S_n = k) \longrightarrow \exp(-x^2/2)$$

En los siguientes dos problemas se recomienda considerar inicialmente $\mathbb{P}(S_n = k)$ y ver $k/n \longrightarrow a$ y entonces relacionar $\mathbb{P}(S_n = j+1)$ con $\mathbb{P}(S_n = j)$ para mostrar que $\mathbb{P}(S_n \geq k) \leq C\mathbb{P}(S_n = k)$.

P5: Supongamos que $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = -1) = 1/2$. Muestre que si $a \in (0,1)$, entonces,

$$\frac{1}{2n}\log \mathbb{P}(S_{2n} \geq 2na) \longrightarrow -\gamma(a)$$

donde $\gamma(a) = \frac{1}{2}[(1+a)\log(1+a) + (1-a)\log(1-a)]$.

P6: Supongamos que $\mathbb{P}(X_i = k) = e^{-1}/j!$ con $k = 0, 1, \ldots$. Muestre que si $a > 1$, entonces,

$$\frac{1}{n}\log \mathbb{P}(S_n \geq na) \longrightarrow a - 1 - a\log a$$

P7: De un ejemplo de variables aleatorias X_n con densidades f_n tales que X_n converge débilmente a una distribución uniformente en $(0,1)$ pero $f_n(x)$ no converga a 1 para todo $x \in [0,1]$.

Sol: Consideremos el espacio de probabilidad $([0,1], \mathcal{B}([0,1]), \lambda)$ donde λ es la medida de Lebesbue. Abusaremos el comportamiento de la integral de Lesbesgue. Definimos,

$$f_n(x) := \mathbb{1}_{(0,1)}$$
$$f(x) := \mathbb{1}_{[0,1]}$$

Podemos observar claramente que f_n no converge puntualmente a f, puesto que,

$$\lim_{n\to\infty} f_n(0) = 0 \not\to 1 = f(0)$$

Sin embargo,

$$F_n(x) = \int_{-\infty}^{x} \mathbb{1}_{(0,1)}(t)dt = \begin{cases} 1, & x > 1 \\ x, & x \in [0,1] \\ 0, & x < 0 \end{cases}$$

(Usamos directamente la integral Riemann puesto que trabajamos con funciones Riemann integrables y luego se puede asociar con la integral de Lebesgue). Recordemos que en la integral de Rieman los puntos son despreciables, también ocurre en algunas medidas, pero en la medida de Lebesgue esto también se cumple. Por otro lado,

$$F(x) = \int_{-\infty}^{x} \mathbb{1}_{[0,1]}(t)dt = \begin{cases} 1, & x > 1 \\ x, & x \in [0,1] \\ 0, & x < 0 \end{cases}$$

Se puede observar que F es continua en todo $\mathbb{R}$ y además,

$$\lim_{n\to\infty} F_n(x) = F(x)$$

para todo $x \in \mathbb{R}$.

P8 (Convergenxia del Máximo): Sean $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes con disitribución F, y sea $M_n := \max_{m\leq n} X_m$. Entonces $\mathbb{P}(M_n \leq x) = F(x)^n$. Pruebe las siguientes leyes para M_n:

(i) Si $F(x) = 1 - x^{-\alpha}$ para $x \geq 1$ donde $\alpha > 0$, entonces para $y > 0$,

$$\mathbb{P}(M_n/n^{1/\alpha} \leq y) \longrightarrow \exp(-y^{-\alpha})$$

(ii) Si $F(x) = 1 - |x|^{\beta}$ para $-1 \leq x \leq 0$ donde $\beta > 0$, entonces para $y < 0$,

$$\mathbb{P}(n^{1/\beta} M_n \leq y) \longrightarrow \exp(-|y|^{\beta})$$

(iii) Si $F(x) = 1 - e^{-x}$ para $x \geq 0$, entonces para todo $y \in (-\infty, \infty)$,

$$\mathbb{P}(M_n - \log n \leq y) \longrightarrow \exp(-e^{-y})$$

Los límites anteriores son llamados los valor extremos de las distribuciones. En particular, el último es llamado la **distribución doble exponencial** o **distribución de Gumbel**.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ espacio de probabilidad. Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e ideticamente distribuidos con distribución F. Observemos que M_n es una variable aleatoria, puesto que para $a \in \mathbb{R}$ se tiene que,

$$\begin{aligned}\{\omega \in \Omega : M_n(\omega) \leq a\} &= \{\omega \in \Omega : \text{ para todo } m = 1, \ldots, n \text{ se tiene que, } X_n(\omega) \leq a\} \\ &= \bigcap_{m=1}^{n} \{\omega \in \Omega : X_n(\omega) \leq a\} \in \mathcal{F}\end{aligned}$$

Luego M_n es variable aleatoria. Determinemos la distribución de M_n, de lo anterior y sabiendo que los X_n son independientes, se cumple que,

$$\begin{aligned}F_{M_n}(x) &= \mathbb{P}(M_n \leq x\}) \\ &= \prod_{m=1}^{n} \mathbb{P}(X_m \leq x) \\ &= \prod_{m=1}^{n} F(x) = (F(x))^n\end{aligned}$$

(i) Sea $F(x) = 1 - x^{-\alpha}$ con $x \geq 1$ y $\alpha > 0$. Claramente F es una función de distribución. Sea $y > 0$ fijo, entonces se cumple que para n suficientemente grande se cumple que $n^{1/\alpha}y \geq 1$. Entonces tomando n suficientemente grande se tiene que,

$$\begin{aligned}\mathbb{P}(M_n/n^{1/\alpha} \leq y) &= \mathbb{P}(M_n \leq n^{1/\alpha}y) \\ &= F(n^{1/\alpha}y)^n \\ &= (1 - n^{-1}y^{-\alpha})^n \\ &= \left(1 - \frac{y}{n}\right)^n \overset{n\to\infty}{\longrightarrow} \exp(-y^{-\alpha})\end{aligned}$$

Como se queria verificar.

(ii) Sea $F(x) = 1 - |x|^{\beta}$. Observemos que F es una función distribución (se puede arreglar). Sea $y < 0$ fijo, entonces $-1 \leq n^{-1/\beta}y < 0$ para n suficientemente grande. Entonces tomando n suficientemente grande se tiene que,

$$\begin{aligned}\mathbb{P}(n^{1/\beta}M_n \leq y) &= \mathbb{P}(M_n \leq n^{-1/\beta}y) \\ &= (1 - n^{-1}|y|^{\beta})^n \\ &= \left(1 - \frac{|y|^{\beta}}{n}\right)^n \overset{n\to\infty}{\longrightarrow} \exp(-|y|^{\beta})\end{aligned}$$

Como se queria probar.

(iii) Sea $F(x) = 1-e^{-x}$ una función distribución. Sea $y \in \mathbb{R}$, por lo que claramente $y+\log n \geq 0$ para n suficientemente grande. Entonces tomando n suficientemente grande se tiene que,

$$\begin{aligned}
\mathbb{P}(M_n - \log n \leq y) &= \mathbb{P}(M_n \leq y + \log n) \\
&= (1 - e^{-y-\log n})^n \\
&= (1 - e^{-y} n^{-1})^n \\
&= \left(1 - \frac{e^{-y}}{n}\right)^n \overset{n\to\infty}{\longrightarrow} \exp(-e^{-y})
\end{aligned}$$

Como se queria probar.

P9: Sean $\{X_n\}_{n\in\mathbb{N}}$ una secuancia de variables aleatorias independientes e identicamente distribuidas con distribución normal.

(i) Del teorema 1.2.6 sabemos que,

$$\mathbb{P}(X_i > x) \sim \frac{1}{\sqrt{2\pi}x} e^{-x^2/2}$$

cuando $x \to \infty$. Use esto para concluir que para cualquier número real θ, se cumple,

$$\mathbb{P}(X_i > x + (\theta/x))/\mathbb{P}(X_i > x) \longrightarrow e^{-\theta}$$

(ii) Muestre que si definimos b_n por $\mathbb{P}(X_i > b_n) = 1/n$, entonces,

$$\mathbb{P}(b_n(M_n - b_n) \leq x) \longrightarrow \exp(-e^{-x})$$

(iii) Muestre que $b_n \sim (2\log n)^{1/2}$ y concluya $M_n/(2\log n)^{1/2} \to 1$ en probabilidad.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean $\{X_n\}_{n\in\mathbb{N}}$ variables aleatorias independientes con distribución normal estándar (media 0 y varianza 1). Por definición podemos pensar la integral de Lebesgue usando la integral de Riemann, dado que la distribución normal tiene una densidad y esta es Riemann integrable, luego,

$$\begin{aligned}
\sqrt{2\pi}\mathbb{P}(X_n > x) &= \int_x^\infty e^{-\frac{t^2}{y}} dt \\
&= -\frac{1}{t} e^{-\frac{t^2}{2}} \bigg|_x^\infty - \int_x^\infty \frac{1}{t^2} e^{-\frac{t^2}{2}} dt \\
&= -\frac{1}{t} e^{-\frac{t^2}{2}} \bigg|_x^\infty - \left(-\frac{1}{t^3} e^{-\frac{t^2}{2}} \bigg|_x^\infty - \int_x^\infty \frac{3}{t^4} e^{-\frac{t^2}{2}} dt \right)
\end{aligned}$$

De la segunda igualdad obtenemos que,

$$\mathbb{P}(X_n > x) \leq \frac{1}{\sqrt{2\pi}x} e^{-\frac{x^2}{2}}$$

Y de la tercera igualdad obtenemos que,

$$\frac{1}{\sqrt{2\pi}}\left(\frac{1}{x}-\frac{1}{x^3}\right)e^{-\frac{x^2}{2}} \leq \mathbb{P}(X_n > x)$$

Por tanto, tomando x suficientemente grande obtenemos que,

$$\mathbb{P}(X_n > x) \sim \frac{1}{\sqrt{2\pi x}}e^{-\frac{x^2}{2}}$$

(i) De lo anterior se tiene que,

$$\frac{\mathbb{P}(X_n > x+\theta/x)}{\mathbb{P}(X_n > x)} \sim \frac{\frac{1}{\sqrt{2\pi(x+\theta/x)}}\exp\left(-\frac{1}{2}(x+\theta/x)^2\right)}{\frac{1}{\sqrt{2\pi x}}\exp\left(-\frac{1}{2}x^2\right)} \xrightarrow{x\to\infty} e^{-\theta}$$

(ii) Probemos que $b_n \to \infty$. Supongamos que $b_n \geq b_{n+1}$ para algún $n \in \mathbb{N}$, luego,

$$\{X_n > b_{n+1}\} \supseteq \{X_n > b_n\}$$

Entonces,

$$\frac{1}{n+1} = \mathbb{P}(X_n > b_{n+1}) \geq \mathbb{P}(X_n > b_n) = \frac{1}{n}$$

Pero entonces $n \geq n+1$ siendo imposible. Luego $b_n < b_{n+1}$ para todo $n \in \mathbb{N}$. **terminar**. Sea $M_n := \text{máx}_{m\leq n} X_m$, luego M_n es una variable aleatoria con distribución F^n donde F es la distribución normal estándar. De esta forma se obtiene que,

$$\begin{aligned}
\mathbb{P}(b_n(M_n - b_n) \leq x) &= \mathbb{P}\left(M_n \leq b_n + \frac{x}{b_n}\right)\\
&= \left(\mathbb{P}\left(X_m \leq b_n + \frac{x}{b_n}\right)\right)^n\\
&= \left(1-\mathbb{P}\left(X_m > b_n + \frac{x}{b_n}\right)\right)^n\\
&= \left(1-\mathbb{P}(X_m > b_n)\frac{\mathbb{P}\left(X_m > b_n + \frac{x}{b_n}\right)}{\mathbb{P}(X_m > b_n)}\right)^n\\
&= \left(1-\frac{1}{n}\frac{\mathbb{P}\left(X_m > b_n + \frac{x}{b_n}\right)}{\mathbb{P}(X_m > b_n)}\right)^n \xrightarrow{n\to\infty} \exp(-e^{-x})
\end{aligned}$$

(iii) Probemos que perfectamente $b_n \sim \sqrt{2\log n}$. Observemos que si tomamos $x = \sqrt{2\log n}$, entonces por el item (i) se obtiene que,

$$\mathbb{P}(X_n > \sqrt{2\log n}) \sim \frac{1}{n\sqrt{4\pi\log n}}$$

Luego para n suficientemente grande $b_n \leq \sqrt{2\log n}$. Mediante un argumenta similar tomando $x = \sqrt{2\log n - 2\log(\log n)}$ se concluye la otra desigualdad y por tanto $b_n \sim \sqrt{2\log n}$.

Probemos la convergencia en probabilidad. Sean $\{x_n\}_{n\in\mathbb{N}}, \{y_n\}_{n\in\mathbb{N}}$ sucesiones reales tales que $x_n \to \infty$ e $y_n \to -\infty$, entonces,

$$\mathbb{P}\left(M_n - b_n \leq \frac{x_n}{b_n}\right) \to 1$$
$$\mathbb{P}\left(M_n - b_n \leq \frac{y_n}{b_n}\right) \to 0$$

Luego,

$$\mathbb{P}\left(\frac{y_n}{b_n} \leq M_n - b_n \leq \frac{x_n}{b_n}\right) \to 1$$

Por lo tanto,

$$\mathbb{P}\left(\frac{y_n}{b_n\sqrt{2\log n}} \leq \frac{M_n - b_n}{\sqrt{2\log n}} \leq \frac{x_n}{b_n\sqrt{2\log n}}\right) \to 1$$

Tomando x_n, y_n como $\mathcal{O}(b_n^2)$ el error, se tiene que,

$$\frac{M_n}{\sqrt{2\log n}} \longrightarrow 1$$

en probabilidad.

Guía 4

P1 (Lema de Fatou): Sea $g \geq 0$ función continua. Si $X_n \Rightarrow X$, entonces,

$$\liminf_{n\to\infty} \mathbb{E}(g(X_n)) \geq \mathbb{E}(g(X))$$

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad con X_n, X variables aleatorias tales que $X_n \Rightarrow X$. Si $g \geq 0$ es continua, entonces es medible. Enunciemos el lema de Fatou.

Lema de Fatou: *Sea $(\Omega, \mathcal{F}, \mu)$ un espacio de medida. Sea $f_n : \Omega \to [0, \infty]$ medible para todo $n \in \mathbb{N}$. Entonces,*

$$\int_\Omega \liminf_{n\to\infty} f_n d\mu \leq \liminf_{n\to\infty} \int_\Omega f_n d\mu$$

Dem: Por definición,

$$\liminf_{n\to\infty} f_n = \lim_{n\to\infty} \inf_{k\geq n}\{f_n\}$$

Tomando $g_n := \inf_{k\geq n}\{f_n\}$, entonces se cumple que $g_n \geq 0, g_n \leq g_{n+1}, g_n \leq f_n$ para todo $n \in \mathbb{N}$. Entonces por el teorema de la convergencia monótona se cumple que,

$$\begin{aligned}
\int_\Omega \lim_{n\to\infty} g_n d\mu &= \lim_{n\to\infty} \int_\Omega g_n d\mu \\
&= \liminf_{n\to\infty} \int_\Omega g_n d\mu \\
&\leq \liminf_{n\to\infty} \int_\Omega f_n d\mu
\end{aligned}$$

Probando el lema de Fatou. ■

Si $X_n \Rightarrow X$, entonces existe un espacio de probabilidad con variables aleatorias Y_n, Y tales que tienen igual distribución con X_n, X respectivamente y Y_n converge a Y casi seguramente. Luego se cumple que $g(Y_n)$ converge a $g(Y)$ casi seguramente por la continuidad de g. Luego,

$$\begin{aligned}
\int_\Omega g(Y) d\mu &= \int_\Omega \liminf_{n\to\infty} g(Y_n) d\mu \\
&\leq \liminf_{n\to\infty} \int_\Omega g(Y_n) d\mu
\end{aligned}$$

Usando que Y_n, Y comparte distribución con X_n, X, se cumple que,

$$\mathbb{E}(X) = \mathbb{E}(Y) \leq \liminf_{n\to\infty} \mathbb{E}(g(Y_n)) = \liminf_{n\to\infty} \mathbb{E}(X_n)$$

Probando lo que queriamos probar.

P2 (La métrica de Lévy): Muestre que,

$$\rho(F, G) = \inf\{\varepsilon : F(x - \varepsilon) - \varepsilon \leq G(x) \leq F(x + \varepsilon) + \varepsilon \text{ para todo } x\}$$

Define una métrica en el espacio de probabilidad y $\rho(F_n, F) \to 0$ si y sólo si $F_n \Rightarrow F$.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ espacio de probabilidad. Sea $\mathcal{C}$ la colección de funciones $F : \mathbb{R} \to \mathbb{R}$ que satisfacen las condiciones de una función distribución acumulada. Definimos la función,

$$\rho : \mathcal{C} \times \mathcal{C} \to [0, \infty)$$

Probaremos que es una métrica. Probemos los axiomas de métrica.

(a) Supongamos que $F = G$, observemos que para todo $\varepsilon > 0$ se cumple,

$$F(x - \varepsilon) - \varepsilon \leq F(x) \leq F(x + \varepsilon) + \varepsilon$$

para todo $x \in \mathbb{R}$ donde F es creciente. Es decir, podemos tomar una sucesión $\{\varepsilon_n\}_{n\in\mathbb{N}}$ tal que $\varepsilon_n \to 0$, es decir, $\rho(F, F) = 0$.

Ahora supongamos que F, G son tales que $\rho(F, G) = 0$, es decir, existe una sucesión $\{\varepsilon_n\}_{n\in\mathbb{N}}$ tal que,

$$F(x - \varepsilon_n) - \varepsilon_n \leq G(x) \leq F(x + \varepsilon_n) + \varepsilon_n$$

para todo $x \in \mathbb{R}$ donde $\varepsilon_n \to 0$. Tomando $n \to \infty$ se llega a que,

$$F(x^-) \leq G(x) \leq F(x)$$

para todo $x \in \mathbb{R}$. Esto implica que $G(x) \leq F(x)$ pero no necesariamente $F(x) \leq G(x)$. Si x es tal que F es continua, entonces $F(x^-) = F(x) = G(x)$. Ahora observemos que $D(F)$ es a lo más numerable, entonces podemos tomar una sucesión $\{x_n\}_{n\in\mathbb{N}}$ de puntos donde F es continua que converge por arriba de $x \in D(F)$, luego $F(x_n) = G(x_n)$ para todo $n \in \mathbb{N}$ y además,

$$F(x) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} G(x_n) = G(x)$$

Probando que $F = G$.

(b) Debemos probar la simetría. Pero si considermos $\varepsilon > 0$ tal que,

$$F(x - \varepsilon) - \varepsilon \leq G(x) \leq F(x + \varepsilon) + \varepsilon$$

para todo $x \in \mathbb{R}$. Tomando $x = y - \varepsilon$, obtenemos que para todo $y \in \mathbb{R}$ se cumple que,

$$G(x) = G(y - \varepsilon) \leq F(y) + \varepsilon \iff G(y - \varepsilon) - \varepsilon \leq F(y)$$

Ahora tomando $x = y + \varepsilon$, se concluye que para todo $y \in \mathbb{R}$ se cumple que,

$$G(x) = G(y + \varepsilon) \geq F(y) - \varepsilon \iff G(y + \varepsilon) + \varepsilon \geq F(y)$$

Por tanto para todo $y \in \mathbb{R}$ se cumple que,

$$G(y - \varepsilon) - \varepsilon \leq F(y) \leq G(y + \varepsilon) + \varepsilon$$

Esto implica que $\rho(F, G) = \rho(G, F)$.

(c) Debemos probar la desigualdad triangular. Sean F, G, T funciones distribuciones acumuladas, y sean $\varepsilon_1, \varepsilon_2 > 0$ tales que,

$$\begin{aligned} F(x-\varepsilon_1)-\varepsilon_1 \leq G(x) \leq F(x+\varepsilon_1)+\varepsilon_1 \\ G(x-\varepsilon_2)-\varepsilon_2 \leq T(x) \leq G(x+\varepsilon_2)+\varepsilon_2 \end{aligned}$$

para todo $x \in \mathbb{R}$. Probemos que $\varepsilon_1+\varepsilon_2 \geq \rho(F,T)$. Notemos que,

$$\begin{aligned} F(x-(\varepsilon_1+\varepsilon))_2-\varepsilon_1 &\leq G(x-\varepsilon_2) \\ &= T(x)+\varepsilon_2 \end{aligned}$$

Y por otro lado,

$$\begin{aligned} T(x) &\leq G(x+\varepsilon_2)+\varepsilon_2 \\ &\leq F(x+\varepsilon_1+\varepsilon_2)+\varepsilon_1+\varepsilon_2 \end{aligned}$$

Para todo $x \in \mathbb{R}$. Juntando estas dos partes obtenemos,

$$F(x-(\varepsilon_1+\varepsilon_2))-(\varepsilon_1+\varepsilon_2) \leq T(x) \leq F(x+(\varepsilon_1+\varepsilon_2))+(\varepsilon_1+\varepsilon_2)$$

Por lo tanto $\rho(F,T) \leq \varepsilon_1+\varepsilon_2$. Para concluir el resultado observemos que,

$$\rho(F,T)-\varepsilon_1 \leq \varepsilon_2$$

y podemos tomar el ínfimo con respecto a G, T, obteniendo,

$$\rho(F,T)-\varepsilon_1 \leq \rho(G,T)$$

Haciendo lo mismo para ε_1 se concluye que,

$$\rho(F,T) \leq \rho(F,G)+\rho(G,T)$$

Probando la desigualdad triangular.

Luego ρ es una métrica.

Probemos la segunda parte. Supongamos que $\rho(F_n, F) \to 0$ cuando $n \to \infty$. Entonces existe un $\varepsilon > 0$ tal que,

$$F(x-\varepsilon)-\varepsilon \leq F_n(x) \leq F(x+\varepsilon)+\varepsilon$$

para todo $x \in \mathbb{R}$. Tomando los límites inferiores y superioes obtenemos,

$$F(x-\varepsilon)-\varepsilon \leq \liminf_{n\to\infty} F_n(x) \leq \limsup_{n\to\infty} F_n(x) \leq F(x+\varepsilon)+\varepsilon$$

Y para concluir se toma x donde F es continua y tomando $\varepsilon \downarrow 0$ se concluye que,

$$\lim_{n\to\infty} F_n(x) = F(x)$$

para todo x donde F es continua, es decir, $F_n \Rightarrow F$.

P3: Si $F_n \Rightarrow F$ y F es continua, entonces,

$$\sup_x |F_n(x) - F(x)| \to 0$$

Sol: Por contradicción. Supongamos que,

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \neq 0$$

Sea k fijo. Definimos,

$$x_{j,k} = F^{-1}\left(\frac{j}{k}\right)$$

y también definimos $x_{0,k} = -\infty, x_{k,k} = \infty$. Para cada $j = 1, \ldots, k-1$ existe un $N_k(j)$ tal que,

$$|F_n(x_{j,k}) - F(x_{j,k})| < \frac{1}{k}$$

para todo $n \geq N_k(j)$ en virtud de que $F_n \Rightarrow F$ con F continua. Ahora para todo $n \geq$ $\max\{N_k(1), \ldots, N_k(k-1)\}$ y para todo $x \in (x_{j-1,k}, x_{j,k})$ se cumple que,

$$F(x) - \frac{2}{k} \leq F(x_{j-1,k}) - \frac{1}{k}\mathbf{gflaw}$$

Por tanto,

$$|F_n(x) - F(x)| < \frac{1}{k}$$

para todo $x \in (x_{j-1,k}, x_{j,k})$ para todo $j = 0, \ldots, k$. Finalmente para n suficientemente grande se cumple que,

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| < \frac{1}{k}$$

Luego tomando $k \to \infty$ se concluye el resultado.

P4: Si F es una función distribución, hay una sucesión de distribuciones de la forma,

$$F_n(x) = \sum_{m=1}^{n} a_{n,m} \mathbb{1}_{\{x_{n,m} \leq x\}}$$

con $F_n \Rightarrow F$. 0 **P5:** Sea X_n $1 \leq n \leq \infty$ variables aleatorias enteros. Muestre que $X_n \Rightarrow X$ si y sólo si $\mathbb{P}(X_n = m) \to \mathbb{P}(X = m)$ para todo m.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ espacio de probabilidad. Se tiene que $X_n \in \mathbb{Z}$ para todo $n \in \mathbb{N}$. Supongamos que $X_n \Rightarrow X$. Sean F_n, F las funciones distribuciones de X_n, X respectivamente. Por definición de distribución acumulada se cumple que,

$$F(t) = \sum_{\substack{y\leq m \\ m\in\mathbb{Z}}} \mathbb{P}(X \leq m)$$

para todo $t \in \mathbb{R}$. Análogamente para F_n. Si graficamos las funciones distribuciones, entonces se puede ver que F es discontinua en todo los enteros $m \in Z$, pero continua en todo $x \in \mathbb{R} \setminus \mathbb{Z}$. Entonces,

$$\lim_{n\to\infty} F_n(x) = F(x)$$

para todo $x \in \mathbb{R} \setminus \mathbb{Z}$. Ahora observemos lo siguiente,

$$\mathbb{P}(X_n = m) = \mathbb{P}(X_n \le m + 1/2) - \mathbb{P}(X_n \le m - 1/2) = F_n(m + 1/2) - F_n(m - 1/2)$$

Por el comportamiento de X_n. Tomando $n \to \infty$ se cumple que,

$$\mathbb{P}(X_n = m) = \lim_{n\to\infty} (F_n(m + 1/2) - F_n(m - 1/2)) = F(m + 1/2) - F(m - 1/2) = \mathbb{P}(X = m)$$

Para todo $m \in \mathbb{Z}$ (esto se cumple debido a que tenemos convergencia finita).

P6: Muestre que si $X_n \to X$ en probabilidad, entonces $X_n \Rightarrow X$. Por otro lado, muestre que si si $X_n \Rightarrow c$, donde c es una constante, entonces $X_n \to c$ en probabilidad.

Sol: Supongamos que $X_n \to X$ en probabilidad, entonces para todo $\varepsilon > 0$ se cumple que,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0$$

Sea C cerrado, definimos

$$C_\varepsilon := \{x \in \mathbb{R} : d(x, C) \le \varepsilon\}$$

Observemos que $C \subseteq C_\varepsilon$, luego notemos que,

$$\begin{aligned} \mathbb{P}(X_n \in C) &= \mathbb{P}(|X_n - X| \le \varepsilon, X_n \in C) + \mathbb{P}(|X_n - X| > \varepsilon, X_n \in C) \\ &\le \mathbb{P}(|X_n - X| \le \varepsilon, X_n \in C) + \mathbb{P}(|X_n - X| > \varepsilon) \\ &\le \mathbb{X}_{\ltimes} \in \mathbb{C}_\varepsilon + \mathbb{P}(|X_n - X| > \varepsilon) \end{aligned}$$

Tomando el límite superior se obtiene que,

$$\limsup_{n\to\infty} \mathbb{P}(X_n \in C) \le \limsup_{n\to\infty} \mathbb{P}(X \in C_\varepsilon)$$

Esto para tood $\varepsilon > 0$, ahora tomando $\varepsilon \downarrow 0$ se tiene que $F_\varepsilon \downarrow F$, de forma que,

$$\limsup_{n\to\infty} \mathbb{P}(X_n \in C) \le \mathbb{P}(X \in C)$$

Es decir, $X_n \Rightarrow X$.

Supongamos que $X_n \Rightarrow c$, sea $\varepsilon > 0$ fijo, luego se tiene que,

$$\mathbb{P}(|X_n - c| > \varepsilon) = \mathbb{P}(X_n > c + \varepsilon) + \mathbb{P}(X_n < c - \varepsilon)$$

Ahora observemos que $\mathbb{P}(c \in \{c+\varepsilon\}) = \mathbb{P}(c \in \{c-\varepsilon\}) = 0$, por lo que,

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}(|X_n - c| > \varepsilon) &= \lim_{n\to\infty} \mathbb{P}(X_n > c+\varepsilon) + \mathbb{P}(X_n < c-\varepsilon) \\
&= \mathbb{P}(c > c+\varepsilon) + \mathbb{P}(c < c-\varepsilon) \\
&= 0
\end{aligned}$$

Y esto se puede hacer para todo $\varepsilon > 0$. Por tanto $X_n \to c$ en probabilidad.

P7 (Lema ...): Si $X_n \Rightarrow X$ y $Y_n \Rightarrow c$, donde c es una constante, entonces $X_n + Y_n \Rightarrow X + c$. Se puede concluir que si $X_n \Rightarrow X$ y $Z_n - X_n \Rightarrow 0$, entonces $Z_n \Rightarrow X$.

Sol:

P8: Muestre que si $X_n = (X_n^1, \ldots, X_n^n)$ es uniformemente distribuida sobre la superficie de la esfera de radio $\sqrt{n}$ en $\mathcal{R}^n$ entonces $X_n^1 \Rightarrow$ a una normal. **Hint:** Sean $Y_1, Y_2, \ldots$ i.i.d con distribución normal y sea $X_n^i = Y_i(n/\sum_{m=1}^n Y_m^2)^{1/2}$.

Guía 5

P1: Muestre que si φ es una función característica, entonces $\text{Re}(\varphi)$ y $|\varphi|^2$ también lo son.

Sol: Usaremos un pequeño truco. Sea $\widetilde{X}$ una variable aleatoria independiente de X e identicamente distribuida a $-X$. Luego afirmamos que $(X+\widetilde{X})/2$ es una variable aleatoria que genera $\text{Re}(\varphi)$ y $X+\widetilde{X}$ genera $|\varphi|^2$.

Verificar que la parte real es una función característica es complicado. Un argumento es mediante transformación de Fourier

Para la segunda parte es también fácil, ya que de la parte anterior se concluye que,

$$\varphi_{X+\widetilde{X}}(t) = |\varphi_X(t)|^2$$

P2:

(a) Recrée la demostración del **teorema 3.3.11** para mostrar que,

$$\mu(\{a\}) = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} e^{-ita}\varphi(t)dt$$

(b) Si $\mathbb{P}(X \in h\mathbb{Z}) = 1$ donde $h > 0$, entonces su función característica $\varphi(2\pi/h + t) = \varphi(t)$, por lo que,

$$\mathbb{P}(X = x) = \frac{h}{2\pi}\int_{-\pi/h}^{\pi/h} e^{-itx}\varphi(t)dt, \ \ x \in h\mathbb{Z}$$

(c) Si $X = Y + b$ entonces,

$$\mathbb{E}(e^{itX}) = e^{itb}\mathbb{E}(e^{itY})$$

Por lo que si $\mathbb{P}(X \in b + h\mathbb{Z}) = 1$, la fórmula de inversión en (b) es válido para $x \in b + h\mathbb{Z}$.

Sol:

(a) Se tiene que,

$$\frac{1}{2T}\int_{-T}^{T} e^{-ita}\left(\int_{\mathbb{R}} e^{itx}d\mu(x)\right)dt = \frac{1}{2T}\int_{-T}^{T}\int_{\mathbb{R}} e^{-ita}e^{itx}d\mu(x)dt$$

Claramente la función que se integra es L^1, de forma que por Fubini obtenemos que,

$$\frac{1}{2T}\int_{-T}^{T} e^{-ita}\left(\int_{\mathbb{R}} e^{itx}d\mu(x)\right)dt = \int_{\mathbb{R}}\left(\frac{1}{2T}\int_{-T}^{T} e^{it(x-a)}dt\right)d\mu(x)$$

Ahora observemos que,

$$\left|\frac{1}{2T}\int_{-T}^{T} e^{it(x-a)}dt\right| \leq \frac{1}{2|T|}\int_{-T}^{T}\left|e^{it(x-a)}\right|dt \leq 1 \in L1$$

Estudiemos el límite,

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} e^{it(x-a)}dt$$

Si $x \neq a$, entonces,

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} e^{it(x-a)}dt = \lim_{T\to\infty} \frac{\operatorname{sen}(T(x-a))}{T(x-a)} = 0$$

dado que sen es acotado.

Si $x = a$, entonces,

$$\lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} e^{it(x-a)}dt = 1$$

Por tanto el límite converge y luego por convergencia dominada se cumple que,

$$\begin{aligned}\lim_{T\to\infty}\int_{\mathbb{R}}\left(\frac{1}{2T}\int_{-T}^{T} e^{it(x-a)}dt\right)d\mu(x) &= \int_{\mathbb{R}}\left(\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} e^{it(x-a)}dt\right)d\mu(x)\\ &= \int_{\mathbb{R}}(1\cdot \mathbb{1}_{\{a\}} + 0\cdot \mathbb{1}_{\{a\}^c})d\mu(x)\\ &= \mu(\{a\})\end{aligned}$$

Probando lo que queriamos probar.

(b) Sea μ la distribución de X. Entonces se cumple que,

$$\begin{aligned}\varphi\left(\frac{2\pi}{h}+t\right) &= \int_{\mathbb{R}} e^{i(2\pi/h+t)x}d\mu(x)\\ &= \int_{h\mathbb{Z}} e^{i(2\pi/h+t)x}d\mu(x)\\ &= \int_{h\mathbb{Z}} e^{itx}d\mu(x)\\ &= \int_{\mathbb{R}} e^{itx}d\mu(x) = \varphi(t)\end{aligned}$$

Por lo anterior se tiene que,

$$\mathbb{P}(X = x = \lim_{T\to\infty})$$

P3: Supongamos que X e Y son variables aleatorias independientes y tienen función característica φ y distribución μ. Usando el problema anterior a $X-Y$ y usando el ejercio 2.1.5 para obtener que,

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|\varphi(t)|^2dt = \mathbb{P}(X-Y=0) = \sum_x \mu(\{x\})^2$$

Sol: Probemos que $X - Y$ tiene función característica $|\varphi(t)|^2$, y en efecto, se cumple que,

$$\begin{aligned}\varphi_{X-Y}(t) &= \mathbb{E}(e^{itX})\mathbb{E}(e^{-itY}) \\ &= \mathbb{E}(e^{itX})\overline{\mathbb{E}(e^{itX})} \\ &= |\varphi(t)|^2\end{aligned}$$

Ahora por el punto (a) del problema anterior se tiene que,

$$\begin{aligned}\mathbb{P}(X - Y = 0) &= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} e^{-it0}|\varphi(t)|^2 dt \\ &= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} |\varphi(t)|^2 dt\end{aligned}$$

Probemos ahora la igualdad de la derecha. Observemos lo siguiente,

$$\begin{aligned}\mathbb{P}(X - Y = 0) &= \mathbb{P}(\{\omega \in \Omega : X - Y = 0\}) \\ &= \mathbb{P}(\{\omega \in \Omega : X = x, Y = x, x \in \mathbb{R}\}) \\ &= \sum_{x\in\mathbb{R}} \mathbb{P}(X = x, Y = x) \\ &= \sum_{x\in\mathbb{R}} (\mathbb{P}(X = x))^2\end{aligned}$$

Siendo la parte derecha. Como queriamos probar. FInalmente podemos concluir que,

$$\sum_{x\in\mathbb{R}} (\mathbb{P}(X = x))^2 = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} |\varphi(t)|^2 dt$$

P4: De un ejemplo de una medida μ con densidad pero el cual $\int |\varphi(t)|dt = \infty$. **Hint:** Dos de los ejemplos anteriores tienen esta propiedad.

Sol: Vamos a considerar la distribución exponencial la cual es una medida μ con densidad $f(y) = e^{-y}\mathbb{1}_{(0,\infty)}$. Determinemos su función característica.

$$\begin{aligned}\varphi(t) &= \int_{\mathbb{R}} e^{itx} d\mu(x) \\ &= \int_{(0,\infty)} e^{itx}e^{-x} d\lambda \\ &= \int_0^{\infty} e^{itx-x} dx \\ &= \frac{e^{x(it-1)}}{it-1}\Big|_0^{\infty}\end{aligned}$$

Observemos que.

$$e^{x(it-1)} = e^{itx}e^{-x} \xrightarrow{x\to\infty} 0$$

dado que e^{itx} es acotado y controlado, mientras que e^{-x} se va a 0 cuando $x \to \infty$. Por lo que,

$$\varphi(t) = -\frac{1}{it-1} = \frac{1}{1-it}$$

Ahora estudiemos la integral de la norma de la función característica. Por lo que se tiene que,

$$\begin{aligned}\int_{\mathbb{R}} |\varphi(t)|dt &= \int_{\mathbb{R}} \frac{1}{|1-it|} \\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{1+t^2}} = \infty\end{aligned}$$

P5: Use el resultado del **ejemplo 3.3.16** para concluir que si $\{X_n\}_{n\in\mathbb{N}}$ son independiente con distribución de Cauchy, entonces $(X_1 + \cdots + X_n)/n$ tiene la misma distribución de X_1.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad. Sean $\{X_n\}_{n\in\mathbb{N}}$ independientes e identicamente distribuidas con distribución de Cauchy, es decir, su función característica es,

$$\varphi_{X_i}(t) = e^{-|t|}$$

para todo $i \in \mathbb{N}$. Sea $S_n = (X_1 + \cdots + X_n)/n$, luego se tiene que,

$$\begin{aligned}\varphi_{S_n}(t) &= \mathbb{E}(e^{itS_n}) \\ &= (\mathbb{E}(e^{itX_1/n}))^n \\ &= \left(\varphi_{X_1}\left(\frac{t}{n}\right)\right)^n \\ &= (e^{-|t|/n})^n = e^{-|t|}\end{aligned}$$

Dado que los $X_1, \ldots, X_n$ son independientes. Por lo tanto, S_n tiene distribución Cauchy.

P5: Suponga que $X_n \Rightarrow X$ donde X_n tiene distribución normal con media 0 y varianza σ_n^2. Pruebe que $\sigma_n^2 \to \sigma^2 \in [0, \infty)$.

Sol: Sea φ_{X_n} la función característica de X_n. Si X_n es una normal con media 0 y varianza σ^2, entonces,

$$\varphi_{X_n}(t) = e^{-\frac{t^2\sigma_n^2}{2}}$$

Si $X_n \Rightarrow X$ entonces $\varphi_{X_n}(t) \to \varphi_X(t)$ para todo $t \in \mathbb{R}$. Si $t = 1$, entonces,

$$\lim_{n\to\infty} \varphi_{X_n}(1) = \varphi_X(1) < \infty$$

Es decir, σ_n^2 converge, por lo que existe un σ^2 tal que $\sigma_n^2 \to \sigma^2$. Si $\sigma_n^2 \in (0, \infty)$ entonces $\sigma^2 \in [0, \infty]$. Si $\sigma^2 = \infty$, entonces,

$$\lim_{n\to\infty} \varphi_{X_n}(t) = 0$$

para todo $t \in \mathbb{R}$, sin embargo $\varphi_{X_n}(0) = 1$, es contradice la continuidad de φ. Luego $\sigma^2 \in [0, \infty)$.

P6: Muestre que si X_n e Y_n son independientes para todo $1 \leq n \leq \infty$ donde $X_n \Rightarrow X_\infty$ e $Y_n \Rightarrow Y_\infty$, entonces $X_n + Y_n \Rightarrow X_\infty + Y_\infty$.

Sol: Si $X_n \Rightarrow X_\infty, Y_n \Rightarrow Y_\infty$, entonces se cumple que las funciones características convergen de la siguiente forma:

$$\varphi_{X_n}(t) \overset{n\to\infty}{\longrightarrow} \varphi_{X_\infty}(t)$$
$$\varphi_{Y_n}(t) \overset{n\to\infty}{\longrightarrow} \varphi_{Y_\infty}(t)$$

Luego si X_n, Y_n son independientes, se cumple que,

$$\varphi_{X_n+Y_n}(t) = \varphi_{X_n}(t)\varphi_{Y_n}(t) \overset{n\to\infty}{\longrightarrow} \varphi_{X_\infty}(t)\varphi_{Y_\infty}(t) = \varphi_{X_\infty+Y_\infty}(t)$$

dado que X_n, Y_n son independientes para todo $1 \leq n \leq \infty$. Observemos que $\varphi_{X_\infty+Y_\infty}(t)$ es continua en $t = 0$ al ser funciones característica, por lo tanto $X_n + Y_n \Rightarrow X_\infty + Y_\infty$.

P7: Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes y sea $S_n = X_1 + \cdots + X_n$. Sea φ_j la función característica de X_j y supongamos qu $S_n \to S_\infty$ casi seguramente. Entonces S_∞ tiene función característica,

$$\prod_{j=1}^{\infty} \varphi_j(t)$$

Sol: Si $S_n \to S_\infty$ casi seguramente, entonces se cumple que $S_n \Rightarrow S_\infty$. Por lo que,

$$\lim_{n\to\infty} \varphi_{S_n}(t) = \varphi_{S_\infty}(t)$$

Dicho de otra forma,

$$\prod_{n=1}^{\infty} \varphi_n(t) = \lim_{n\to\infty} \prod_{i=1}^{n} \varphi_i(t) = \lim_{n\to\infty} \varphi_{S_n}(t) = \varphi_{S_\infty}(t)$$

Probando lo que queriamos probar.

P8: Use el **teorema 3.3.18** y la expansión de la serie $e^{-t^2/2}$ para mostrar que la distribución normal estándar tiene,

$$\mathbb{E}(X^{2n}) = \frac{(2n)!}{2^n n!}$$

Sol: Para usar el **teorema 3.3.18** se necesita ver que,

$$\int_{\mathbb{R}} |x^n| d\mu(x) < \infty$$

Y en efecto. Si μ tiene densidad, entonces,

$$\int_{\mathbb{R}} |x^n| \frac{1}{\sqrt{2\pi}} e^{-x^2/2} d\lambda(x) = \frac{2}{\sqrt{2\pi}} \int_0^\infty x^n e^{-x^2/2} dx < \infty$$

Sabemos que la última integral es finita porque es la función Gamma. Luego se cumple que,

$$\frac{d^{2n}}{dt^{2n}} \varphi(0) = i^{2n} \mathbb{E}(X^{2n}) = (-1)^n \mathbb{E}(X^{2n})$$

Por otro lado, si X es la normal estándar, entonces,

$$\varphi(t) = e^{-t^2/2} = \sum_{m=0}^{\infty} \frac{(-1)^m t^{2m}}{2^m m!}$$

La $2n$-ésima derivada de φ es de la forma,

$$\sum_{k=n}^{\infty} \frac{(-1)^m (2k)(\dots)(2k-(2n-1)) t^{2k-2k}}{2^k k!}$$

dado que t está al cuadrado. Luego para $t = 0$ se obtiene que,

$$\varphi^{(n)}(0) = \frac{(-1)^n (2n)!}{2^n n!}$$

Finalmente,

$$\mathbb{E}(X^{2n}) = \frac{(2n)!}{2^n n!}$$

P9:

(i) Suponga que la familia de medidas $\{\mu_i\}_{i \in I}$ es tensa. Use (d) del **teorema 3.3.1** y (3.3.3) con $n = 0$ para muestre que sus funciones características φ_i son equicontinuas.

(ii) Suponga que $\mu_n \Rightarrow \mu_\infty$. Use el **teorema 3.3.17** y la equicontinuidad para concluir que $\varphi_n \to \varphi_\infty$ uniformemente en conjuntos compactos.

(iii) De un ejemplo que la convergencia no necesita ser uniformemente en toda la recta real.

Guía 6

P1: Sea $\{X_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e identicamente distribuidas con $\mathbb{E}(X_i) = 0, 0 < \mathrm{Var}(X_i) < \infty$. Sea $S_n = X_1 + \cdots + X_n$.

(a) Use el teorema central del límite y la ley 0-1 de Kolmogorov para concluir que,

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} = \infty$$

casi seguramente.

(b) Use un argumento por contradicción para mostrar que,

$$\frac{S_n}{\sqrt{n}}$$

no converge en probabilidad. **Hint:** Considere $n = m!$.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad donde están definidos las variables aleatorias.

(a) Sea $\mathcal{T}$ la σ-álgebra cola con respecto a los $\{X_n\}_{n\in\mathbb{N}}$. Observemos que el conjunto,

$$\left\{\limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} = \infty\right\}$$

es un evento de $\mathcal{T}$. Puesto que,

$$\left\{\limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} = \infty\right\} = \bigcap_{M>0} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \left\{\sup_{k\geq n} \frac{S_k}{\sqrt{k}} > M\right\} \in \sigma(X_l, \dots)$$

para todo $l \geq 1$. Por lo que o bien, el evento tiene probabilidad 0 o 1. Por lo que basta probar que la probabilidad no puede ser 0.

Por el teorema central del límite se tiene que,

$$\frac{S_n}{\sigma\sqrt{n}} \Rightarrow \mathcal{X}$$

Entonces se tiene que,

$$\begin{aligned}\limsup_{n\to\infty} \mathbb{P}\left(\frac{S_n}{\sqrt{n}} > M\right) &= \limsup_{n\to\infty} \mathbb{P}\left(\frac{S_n}{\sigma\sqrt{n}} > \frac{M}{\sigma}\right)\\ &= \mathbb{P}\left(\mathcal{X} > \frac{M}{\sigma}\right) \neq 0\end{aligned}$$

Por tanto, la única opción posible, es que,

$$\limsup_{n\to\infty} \frac{S_n}{\sqrt{n}} = \infty$$

casi seguramente.

(b)

P2: Sea $\{X_n\}_{n\in\mathbb{N}}$ variables aleatorias independientes e identicamente distribuidas no negativas tales que $\mathbb{E}(X_i) = 1$ y $\text{Var}(X_i) = \sigma^2 \in (0, \infty)$. Muestre que $2(\sqrt{S_n} - \sqrt{n}) \Rightarrow \sigma\mathcal{X}$.

Sol: Sea $(\Omega, \mathcal{F}, \mathbb{P})$ espacio de probabilidad donde están definidos $\{X_n\}_{n\in\mathbb{N}}$. Observemos que por el teorema central del límite se cumple que,

$$\frac{S_n - n}{\sigma\sqrt{n}} \Rightarrow \mathcal{X}$$

Además, $S_n \geq 0$. Sea x tal que $\sqrt{S_n} \leq x/2 + \sqrt{S_n}$ se cumple. Luego,

$$\begin{aligned}
\mathbb{P}\left(\sqrt{S_n} \leq \frac{x}{2} + \sqrt{S_n}\right) &= \mathbb{P}\left(S_n \leq \frac{x^2}{2} + x\sqrt{n} + n\right) \\
&= \mathbb{P}\left(\frac{S_n - n}{\sigma n} \leq \frac{s^2}{2\sigma\sqrt{n}} + x\right)
\end{aligned}$$

Si la función distribución de $\mathcal{X}$ es continua en todos lados y aplicando que se puede meter el límite dentro de la medida se cumple que,

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - n}{\sigma n} \leq \frac{s^2}{2\sigma\sqrt{n}} + x\right) &= \mathbb{P}\left(\mathcal{X} \leq \frac{x}{\sigma}\right) \\
&= \int_{-\infty}^{x/\sigma} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx
\end{aligned}$$

Tomando el cambio de variable $s = \sigma x$ se obtiene que,

$$\int_{-\infty}^{x/\sigma} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = \int_{-\infty}^{s} \frac{1}{\sqrt{2\pi}\sigma} e^{-s^2/2\sigma^2} ds$$

Por lo tanto $2(\sqrt{S_n} - \sqrt{n}) \Rightarrow \sigma\mathcal{X}$, (la normal estándar con media 0 y varianza σ^2).

P3 (Fórmula de Bayes): Sea $G \in \mathcal{G}$. Muestre que,

$$\mathbb{P}(G|A) = \int_G \mathbb{P}(A|\mathcal{G}) d\mathbb{P} \bigg/ \int_\Omega \mathbb{P}(A|\mathcal{G}) d\mathbb{P}$$

cuando $\mathcal{G}$ es un σ-álgebra generado por particiones, la fórmula se reduce a la fórmula usual de Bayes,

$$\mathbb{P}(G_i|A) = \mathbb{P}(A|G_i)\mathbb{P}(G_i) \bigg/ \sum_j \mathbb{P}(A|G_j)\mathbb{P}(G_j)$$

Sol: Sabemos que $\mathbb{P}(G|A) = \mathbb{P}(G\cap A)/\mathbb{P}(A)$. Probaremos que el lado derecho es el lado izquierdo. Observemos que,

$$\begin{aligned}
\int_G \mathbb{P}(A|\mathcal{G})d\mathbb{P} &= \int_G \mathbb{E}(\mathbb{1}_A|\mathcal{G})d\mathbb{P} \\
&= \int_\Omega \mathbb{1}_G \mathbb{E}(\mathbb{1}_A|\mathcal{G})d\mathbb{P} \\
&= \int_\Omega \mathbb{E}(\mathbb{1}_{A\cap G}|\mathcal{G})d\mathbb{P} \\
&= \int_\Omega \mathbb{1}_{A\cap G}d\mathbb{P} \\
&= \mathbb{P}(A\cap G)
\end{aligned}$$

esto se cumple debido a que por definición A debe ser $\mathcal{G}$-medible, luego $A \cap G$ es $\mathcal{G}$-medible. Ahora por el otro lado,

$$\int_\Omega \mathbb{P}(A|\mathcal{G})d\mathbb{P} = \mathbb{P}(A)$$

De esta forma se prueba la igualdad.

P4: Pruebe la desigualdad de Chebyshev para la esperanza condicional, es decir, si $a > 0$, entonces,

$$\mathbb{P}(|X| \geq a|\mathcal{F}) \leq a^{-2}\mathbb{E}(X^2|\mathcal{F})$$

Sol: Observemos que se cumple la siguiente desigualdad,

$$\begin{aligned}
\int_A \mathbb{P}(|X| \geq a|\mathcal{F}) &= \int_A \mathbb{E}(\mathbb{1}_{|X|\geq a}|\mathcal{F}) \\
&\int_A \mathbb{1}_{|X|\geq a}d\mathbb{P} \\
&\leq \int_A \frac{|X|^2}{a^2}d\mathbb{P} \\
&= \int_A a^{-2}\mathbb{E}(X^2|\mathcal{F})d\mathbb{P}
\end{aligned}$$

para todo $A \in \mathcal{F}$. Si consideramos $A_\varepsilon = \{\omega \in \Omega : \mathbb{P}(|X| \geq a|\mathcal{F}) - a^{-2}\mathbb{E}(X^2|\mathcal{F}) \geq \varepsilon\}$ para $\varepsilon > 0$. Entonces se cumple que,

$$\varepsilon\mathbb{P}(A_\varepsilon) \leq 0$$

Es decir, $\mathbb{P}(A_\varepsilon) = 0$ para todo $\varepsilon > 0$ y tomando $\varepsilon \downarrow 0$ se obtiene que $\mathbb{P}(A) = 0$, por lo que $\mathbb{P}(A^c) = 1$, es decir,

$$\mathbb{P}(|X| \geq a|\mathcal{F}) \leq a^{-2}\mathbb{E}(X^2|\mathcal{F})$$

casi seguramente.

P5: De un ejemplo sobre $\Omega = \{a, b, c\}$ el cual,

$$\mathbb{E}(\mathbb{E}(X|\mathcal{F}_1)|\mathcal{F}_2) \neq \mathbb{E}(\mathbb{E}(X|\mathcal{F}_2)|\mathcal{F}_1)$$

Sol:

P6: Sea $\{Y_n\}_{n\in\mathbb{N}}$ una secuencia de variables aleatorias independientes e identicamente distribuidas con esperanza μ y varianza σ^2. Sea N independiente positivo **¿?** con $\mathbb{E}(N^2) < \infty$ y sea $X = Y_1 + \cdots + Y_N$. Muestre que $\text{Var}(X) = \sigma^2\mathbb{E} + \mu^2\text{Var}(N)$. Estudie que pasa cuando N o Y son constantes.

Guía 7

P1: Generalize el **teorema 4.2.7** mostrando que si $\{X_n\}_{n\in\mathbb{N}}, \{Y_n\}_{n\in\mathbb{N}}$ son submartingala con respecto a la filltración $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$, entonces $X_n \vee Y_n$ también es una submartingala.

P2: Sea $\{X_n\}_{n\in\mathbb{N}_0}$ una submartingala con $\sup_{n\in N} X_n < \infty$. Sea $\xi_n = X_n - X_{n-1}$ y suponga que $\mathbb{E}(\sup_{n\in\mathbb{N}} \xi_n^+) < \infty$. Muestre que X_n converge casi seguramente.

P3: De un ejemplo de una martingala X_n tal que $X_n \to -\infty$ casi seguramente. **Hint:** Considere $X_n = \xi_1 + \cdots + \xi_n$ donde ξ_i son independiente aunque no identicamente distribuida tal que $\mathbb{E}(\xi_i) = 0$.

P4: Sea $\{Y_n\}_{n\in\mathbb{N}}$ variables aleatorias independientes e identicamente distribuidas no negativas tal que $\mathbb{E}(Y_m) = 1$ y $\mathbb{P}(Y_m = 1) < 1$. Por el **ejemplo 4.2.3** $X_n = \prod_{m\leq n} Y_m$ define una martingala.

(i) Use el **teorema 4.2.12** y argumente por contradicción para mostrar que $X_n \to 0$ casi seguramente.

(ii) Use la ley fuerte de los grandes números para concluir que $(1/n)\log X_n \to c < 0$.

P5: Sean $\{X_n\}_{n\in\mathbb{N}}, \{Y_n\}_{n\in\mathbb{N}}$ variables aleatorias positiva integrable y adaptadas a $\mathcal{F}_n$. Suponga que,

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) \leq (1 + Y_n)X_n$$

tal que $\sum Y - n < \infty$ casi seguramente. Pruebe que X_n converge casi seguramente a un límite finite **terminar 4.2.8**

P6: Sea $\{X_n^1\}_{n\in\mathbb{N}}, \{X_n^2\}_{n\in\mathbb{N}}$ supermartingalas con respecto a $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$, y sea N el tiempo de parada tal que $X_N^1 \geq X_N^2$. Entonces,

$$\begin{aligned} Y_n &= X_n^1 \mathbb{1}_{N>n} + X_n^2 \mathbb{1}_{N\leq n} \\ Z_n &= X_n^1 \mathbb{1}_{N\geq n} + X_n^2 \mathbb{1}_{N<n} \end{aligned}$$

son supermartingalas.

P7: Muestre que si $\mathbb{P}(\text{lím}\, Z_n/\mu^n = 0) < 1$ entonceses igual a ρ y se tiene que,

$$\left\{ \lim_{n\to\infty} \frac{Z_n}{\mu^n} > 0 \right\} = \{Z_n > 0\, \text{para todo}\, n\}$$

casi seguramente.

P8: Sea Z_n un proceso **¿' 4.3.12**

P9: terminar