Probabilidad

Matías Carrasco y Victor Ortega
Facultad de Economía, UNAM

18 de junio de 2025

Índice general

Capítulo 1

Fundamentos

Las siguientes son notas de clase del curso de Probabilidad del Act. Servando Valdés Cruz de la Facultad de Ciencias de la UNAM. Estas notas fueron originalmente escritas por Magali Díaz.

1.1. Espacio de Probabilidad

Definición 1.1. Un espacio de probabilidad es la terna $(\Omega, \mathcal{F}, P)$ donde $\Omega \neq \varnothing$ es el espacio muestral, $\mathcal{F}$ es una σ-álgebra de subconjuntos de Ω, y P es una medida de probabilidad.

Definición 1.2. Se dice que $\mathcal{F}$ es una σ-álgebra de subconjuntos de $\Omega \neq \varnothing \Leftrightarrow$

$(\mathcal{F}_1)$ $\Omega \in \mathcal{F}$ — Universalidad

$(\mathcal{F}_2)$ $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$ — Cerrado Bajo Complementos

$(\mathcal{F}_3)$ $A_n \in \mathcal{F}\ \forall\, n \in \mathbb{N} \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{F}$ — Cerrado Bajo Uniones

Teorema 1.1 (Propiedades de σ-álgebra)**.** Sea $\mathcal{F}$ una σ-álgebra de subconjuntos de $\Omega \neq \varnothing \Rightarrow$

$(\mathcal{F}_1)$ $A_1, ..., A_n \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{F}\ \forall n \in \mathbb{N}$ — Cerrada Bajo Uniones Finitas

$(\mathcal{F}_2)$ $\varnothing \in \mathcal{F}$ — Conjunto Vacio

$(\mathcal{F}_3)$ $A_n \in \mathcal{F}\ \forall n \in \mathbb{N} \Rightarrow \bigcap_{n\in\mathbb{N}} A_n \in \mathcal{F}$ — Cerrada Bajo Intersecciones Numerables

$(\mathcal{F}_4)$ $A_1, ..., A_n \in \mathcal{F} \Rightarrow \bigcap_{i=1}^{n} A_i \in \mathcal{F}\ \forall n \in \mathbb{N}$ — Cerrada Bajo Intersecciones Finitas

$(\mathcal{F}_5)$ $A_1, A_2 \in \mathcal{F} \Rightarrow A_2 \setminus A_2 \in \mathcal{F}$. — Diferencia de Conjuntos

Proof. Probemos el Teorema 1.1

$(\mathcal{F}_1)$ Sea $A_1, \ldots, A_n \in \mathcal{F}$. Definimos una sucesión:

$$B_k = \begin{cases} A_k & \text{si } k \leqslant n \\ \varnothing & \text{si } k > n \end{cases}$$

$$\Rightarrow \bigcup_{k=1}^{\infty} B_k = \bigcup_{i=1}^{n} A_i$$

Como cada $B_k \in \mathcal{F}$ por $(\mathcal{F}_2)$

$$\Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{F}$$

esto por cerradura bajo uniones numerables.

$(\mathcal{F}_2)$ Como $\Omega \in \mathcal{F}$ y $\mathcal{F}$ es cerrada bajo complementos $\Rightarrow \Omega^c = \varnothing \in \mathcal{F}$.

$(\mathcal{F}_3)$ Sean $\{A_n\}_{n\in\mathbb{N}} \subseteq \mathcal{F} \Rightarrow \forall\, n \Rightarrow A_n^c \in \mathcal{F}$, y por $(\mathcal{F}_1) \Rightarrow$

$$\bigcup_{n\in\mathbb{N}} A_n^c \in \mathcal{F} \Rightarrow \left(\bigcup_{n\in\mathbb{N}} A_n^c\right)^c = \bigcap_{n\in\mathbb{N}} A_n \in \mathcal{F}.$$

$(\mathcal{F}_4)$ Sean $A_1, \ldots, A_n \in \mathcal{F}$, y por $(\mathcal{F}_1) \Rightarrow$

$$\bigcap_{i=1}^{n} A_i = \left(\bigcup_{i=1}^{n} A_i^c\right)^c \in \mathcal{F}.$$

$(\mathcal{F}_5)$ Sean $A_1, A_2 \in \mathcal{F}$. Entonces $A_2^c \in \mathcal{F}$ y, como $\mathcal{F}$ es cerrada bajo intersecciones finitas:

$$A_1 \setminus A_2 = A_1 \cap A_2^c \in \mathcal{F}.$$

$\therefore$ Teorema 1.1 es cierto. □

Definición 1.3 (Evento)**.** Se dice que A es un **evento** $\Leftrightarrow A \in \mathcal{F} \subseteq \Omega$

Definición 1.4 (Medida de Probabilidad)**.** Se dice que P es una medida de probabilidad $\Leftrightarrow$

(P_1) $P(\Omega) = 1$

(P_2) $\forall\, A \in \mathcal{F} \Rightarrow P \geqslant 0$

(P_3) Si $\{A_i\}_i^\infty \in \mathcal{F} \mid A_i \cap A_j = \varnothing\ \forall\, j \neq i \Rightarrow P(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty P(A_i)$

Teorema 1.2. Sea P una medida de probabilidad $\Rightarrow$

(P_1) $P(\varnothing) = 0$.

(P_2) Sean $A_1, ..., A_n$ eventos ajenos $\Rightarrow P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$.

(P_3) $\forall\, A \in \mathcal{F} \Rightarrow P(A) = 1 - P(A^c)$.

(P_4) Sean $A, B \in \mathcal{F} \subseteq \Omega$ y $A \in B \Rightarrow P(B \setminus A) = P(B) - P(A)$ y $P(A) \subseteq P(B)$.

(P_5) Sean $A, B \in \mathcal{F} \subseteq \Omega \Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

(P_6) Sean $A, B, C \in \mathcal{F} \subseteq \Omega \Rightarrow$
$P(A\cup B\cup C) = P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$

(P_7) Se cumple la fórmula de inclusión y exclusión.

(P_8) Sean A, B dos eventos cualesquiera $\Rightarrow P(A) = P(A \cap B) + P(A \cap B^c)$.

(P_9) Se cumple la desigualdad de Boole.

Proof. Probemos el Teorema 1.2

(P_1) Notemos que podemos expresar

$$\varnothing = \varnothing \cup \varnothing \cup \varnothing \cup ... \cup \varnothing \quad \text{y ademas} \quad \varnothing \cap \varnothing = \varnothing$$

Por la Definición 1.4

$$P(\varnothing \cup ... \cup \varnothing) = \sum_{i=1}^{\infty} P(\varnothing) \Leftrightarrow P(\varnothing) = \sum_{i=1}^{\infty} P(\varnothing)$$

Ya que

$$P(A) \geqslant 0 \ \forall A \in \mathcal{F} \quad \text{y ademas} \quad \varnothing \in \mathcal{F} \Rightarrow P(\varnothing) = 0$$

(P_2) $\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{\infty} A_i, A_m = \varnothing$ Notemos que $\forall\, m \geqslant n+1 \Rightarrow$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) = \sum_{i=1}^{n} P(A_i) + \sum_{j=n+1}^{\infty} P(A_j) =$$

$$\sum_{i=1}^{n} P(A_i) + \sum_{j=n+1}^{\infty} P(\varnothing) \ = \sum_{i=1}^{n} P(A_i)$$

(P_3) Note usted que $\Omega = A + A^c \Rightarrow$

$$P(\Omega) = P(A) + P(A^c) \Rightarrow P(A \cup A^c) = P(A) + P(A^c) = 1 \Rightarrow P(A) = 1 - P(A^c)$$

(P_4) Es fácil ver que $B = A \cup (B \setminus A)$ y que $A \cap (B \setminus A) = \varnothing$

Se sigue por la Definición 1.4 que

$$\Rightarrow P(B) = P(A \cup B \setminus A) = P(A) + P(B \setminus A) \Rightarrow P(B \setminus A) = P(B) - P(A)$$

Ahora, como $P(B \setminus A) = P(B) - P(A)$ y $B \setminus A \in \mathcal{F}$, usando la Definición 1.4

$$\Rightarrow P(B \setminus A) \geqslant 0 \Rightarrow P(B) = P(A) + P(B \setminus A)$$

(P_5) Notemos que $A \cup B = A \cup B \setminus A$ y que $A \cap (B \setminus A) = \varnothing$

$$\Rightarrow P(A \cup B) = P(A) + P(B \setminus A)$$

Note usted que $B \setminus A = B \setminus A \cap B$ y que $(A \cap B) \leqslant 1 \Rightarrow$

$$P(B \setminus A) = P(B) - P(A \cap B) \Rightarrow P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

(P_6) Por (P_5) sabemos lo siguiente

$$P(A \cup B \cup C) = P((A \cup B) \cup C) = P(A \cup B) + P(C) - P((A \cup B) \cap C) \Rightarrow$$

Se sigue que

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

(P_7) Sean $\{A_i\}_{i=1}^n \in \mathcal{F}$ cualesquiera n eventos $\Rightarrow$

$$P(A_1 \cup ... \cup A_n) = P\left(\bigcup_{i=1}^n A_i\right) =$$

$$\sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) + ... + (-1)^{n-1} P\left(\bigcap_{i=1}^n A_i\right)$$

(P_8) Note que $A = A \cap \Omega = A \cap (B \cup B^c) = (A \cap B) \cup (A \cap B^c)$, y por (P_5) s.t.q.

$$P(A) = P(A \cap B) + P(A \cap B^c)$$

(P_9) Para $n = 1$ tenemos que $P(A_1) \leqslant P(A_1)$.

Se sigue que $P(\bigcup_{i=1}^n A_i) \leqslant \sum_{i=1}^n P(A_i)$.

Por (P_5) tenemos que

$$P\left(\bigcup_{i=1}^{n+1} A_i\right) = P\left(\bigcup_{i=1}^n A_i\right) + P(A_{n+1}) - P\left(\bigcup_{i=1}^n A_i \cap A_{n+1}\right)$$

Como $P(\bigcup_{i=1}^n A_i \cap A_{n+1}) \geqslant 0 \Rightarrow$

$$P\left(\bigcup_{i=1}^{n+1} A_i\right) \leqslant P\left(\bigcup_{i=1}^n A_i\right) + P(A_{n+1}) = \sum_{i=1}^{n+1} P(A_i)$$

$\therefore$ el Teorema 1.2 es cierto. $\square$

Notación (Función). Una función f es una regla que asigna a cada elemento $x \in A$ un único elemento $f(x) \in B$, lo que denotamos:

$$f : A \to B, \quad x \mapsto f(x)$$

En este caso, A es el dominio, B el codominio, y $f(x)$ la imagen de x bajo f.

Observación. P es una función conjuntista.

Definición 1.5 (Monotonía). Sea $\{A_n\}_{n \geqslant 1} \subseteq \mathcal{F}$.

Decimos que es **monótona creciente** si $A_1 \subseteq A_2 \subseteq \cdots$, en cuyo caso

$$\lim_{n\to\infty} A_n = \bigcup_{n=1}^\infty A_n.$$

Decimos que es **monótona decreciente** si $A_1 \supseteq A_2 \supseteq \cdots$, en cuyo caso

$$\lim_{n\to\infty} A_n = \bigcap_{n=1}^\infty A_n.$$

Teorema 1.3 (Teorema de Continuidad). Sea $\{A_n\}_{n \geqslant 1}$ una sucesión monótona de eventos de un espacio de probabilidad $(\Omega, \mathcal{F}, P) \Rightarrow P(A) = P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$

Proof. Caso 1:

$\{A_n\}$ es una sucesión creciente, es decir, $A_1 \subseteq A_2 \subseteq \cdots$. Definimos:

$$A := \lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n.$$

Sea $B_1 := A_1$ y para $n \geqslant 2$, definimos $B_n := A_n \setminus A_{n-1}$. Entonces los B_n son eventos disjuntos dos a dos, y se cumple que:

$$\bigcup_{n=1}^{\infty} A_n = \bigsqcup_{n=1}^{\infty} B_n,$$

donde $\bigsqcup$ denota unión disjunta.

Por la σ-aditividad de la medida de probabilidad:

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

Además, como $A_n = \bigcup_{k=1}^{n} B_k$, se tiene que:

$$P(A_n) = \sum_{k=1}^{n} P(B_k),$$

y por lo tanto:

$$\lim_{n\to\infty} P(A_n) = \sum_{k=1}^{\infty} P(B_k) = P(A).$$

Caso 2:

$\{A_n\}$ es una sucesión decreciente, es decir, $A_1 \supseteq A_2 \supseteq \cdots$. Definimos:

$$A := \lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n.$$

Sea $C_n := A_1 \setminus A_n$, que define una sucesión creciente de eventos (ya que $A_n \supseteq A_{n+1} \Rightarrow C_n \subseteq C_{n+1}$), y:

$$\bigcup_{n=1}^{\infty} C_n = A_1 \setminus \bigcap_{n=1}^{\infty} A_n = A_1 \setminus A.$$

Por el caso creciente, sabemos que:

$$\lim_{n\to\infty} P(C_n) = P(A_1) - P(A).$$

Por lo tanto:

$$P(A) = P(A_1) - \lim_{n\to\infty} P(C_n) = \lim_{n\to\infty} P(A_n).$$

En ambos casos se cumple que:

$$P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n).$$

$\therefore$ el Teorema 1.3 es cierto. $\square$

1.2. Análisis Combinatronico

Definición 1.6 (Principio fundamental del conteo). Si hay n caminos para llegar de A a B y m caminos para llegar de B a $C \Rightarrow$ hay $m \times n$ caminos para llegar de A a C.

Definición 1.7 (Ordenación sin repetición). Se desea ordenar sin repetición k elementos tomados de una población n. Donde $0 \leqslant k \leqslant n < \infty$ y $n, k \in \mathbb{Z}^+ \Rightarrow$ en total hay

$$n(n-1)(n-1)...(n-k+1) = \frac{n!}{(n-k)!}$$

Definición 1.8 (Ordenación con repetición). Se desea ordenar con repetición k elementos tomados de una población n. Donde $0 \leqslant k \leqslant n < \infty \ \wedge n, k \in \mathbb{Z}^+ \Rightarrow$ en total hay

$$\frac{n}{1} \times \frac{n}{2} \times ... \times \frac{n}{k} = n^k$$

Definición 1.9. Sea $n \in \mathbb{Z}^+$, y sean $n_1, n_2, ..., n_k \in \mathbb{Z}^+$, tales que $\sum_{i=1}^{k} n_i = n$, entonces definimos la **coeficiente multinomial** como

$$\binom{n}{n_1, n_2, ..., n_k} = \frac{n!}{n_1! n_2! ... n_k!}$$

Definición 1.10. Supongamos que se tiene una población con n elementos y tomamos de ella una muestra de k elementos sin reemplazo (es decir, sin repetición), entonces el número de **combinaciones** a través de la cual podemos extraer dicha muestra es igual a

$$x = \frac{n!}{(n-k)!k!} = \binom{n}{k}$$

Teorema 1.4 (Identidad de Vandermonde). Sean $m, n, r \in \mathbb{N}_0$. Entonces

$$\sum_{k=0}^{r} \binom{m}{k} \binom{n}{r-k} = \binom{m+n}{r}$$

Proof. Consideramos un conjunto de $m+n$ elementos, donde m son de tipo A y n son de tipo B. Queremos contar de cuántas maneras podemos escoger r elementos en total.

Por un lado, directamente:

$$\text{Total de maneras de escoger } r \text{ elementos} = \binom{m+n}{r}$$

Se puede hacer por casos según cuántos de los r elementos provienen del grupo A: si tomamos k elementos de A $\Rightarrow$ tomamos $r-k$ de B. Esto se puede hacer de:

$$\binom{m}{k} \cdot \binom{n}{r-k}$$

maneras, y sumando sobre todos los posibles k:

$$\sum_{k=0}^{r} \binom{m}{k} \binom{n}{r-k}$$

$\therefore$ Teorema 1.4 es cierto. □

1.3. Probabilidad Condicional

Notación. La probabilidad condicional de A dado B es denotada como $P(A|B)$

Definición 1.11 (Ley General de la Probabilidad Condicional)**.** Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y sean $A, B \in \mathcal{F}$, donde $P(B) > 0 \Rightarrow P(A|B)$ se define como

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

Definición 1.12 (Ley General de la Independencia)**.** Se dice que dos eventos $A, B \in \mathcal{F}$ son independientes si se cumple alguna de las siguientes identidades

(I_1) $P(A|B) = P(A)$

(I_2) $P(B|A) = P(B)$

(I_3) $P(A \cap B) = P(A)P(B)$

Corolario. La probabilidad de intersección de dos eventos dependientes $A, B \in \mathcal{F}$ es

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

Para eventos independientes tenemos que $P(A \cap B) = P(A)P(B)$.

Proof. Se sigue directamente de la Definición 1.11 y 1.12. □

Teorema 1.5 (Teorema de Bayes)**.** Sean $A, B \in \mathcal{F}$ con $P(B) \neq 0$. Entonces:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Proof. Por definición de probabilidad condicional:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{y} \quad P(B \mid A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A \cap B)}{P(A)}$$

Sustituimos el Corolario anterior en la primera ecuación

$$\Rightarrow P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

∴ Teorema 1.5 es cierto. □

Teorema 1.6. Sean $\{B_1, ..., B_n\}$ particiones de Ω disjuntas tales que $B_1 \cup ... \cup B_n = \Omega$ para algún $n \in \mathbb{N}$ y $P(B_i) > 0 \, \forall \, i = 1, ..., n \Rightarrow \; \forall \, A \in \mathcal{F} \subseteq \Omega \Rightarrow$

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i)P(B_i)$$

Proof.

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i)P(B_i) = P(A \cap B_1) + ... + P(A \cap B_n)$$

$$= P((A \cap B_1) \cup ... \cup (A \cap B_n)) = P(A \cap (B_1 \cup ... \cup B_n)) = P(A \cap \Omega) = P(A)$$

□

Teorema 1.7 (Teorema General de Bayes). Sean $\{B_1, ..., B_n\}$ particiones de Ω disjuntas tales que $B_1 \cup, ..., \cup B_n = \Omega$ para algún $n \in \mathbb{N}$ y $P(B_i) > 0 \, \forall \, i = 1, ..., n \Rightarrow \; \forall \, A \in \mathcal{F} \subseteq \Omega \Rightarrow$

$$P(B_j \mid A) = \frac{P(A \mid B_j)P(B_j)}{\sum\limits_{i=1}^{n} P(A \mid B_i)P(B_i)}$$

Un evento B_j se llama hipótesis, $P(B_j)$ se llama probabilidad a priori, y $P(B_j|A)$ es una probabilidad a posteriori.

Proof. Por el Teorema 1.6 sabemos que

$$P(B_j|A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{n} P(A|B_i)P(B_i)}$$

□

Observación. El Teorema 1.5 aplica a dos eventos arbitrarios. El Teorema 1.7 requiere una partición de Ω. El primero es un caso particular del segundo.

Teorema 1.8. Sean $A_1, A_2, ..., A_n \in \mathcal{F}$ tal que $P\left(\bigcap_{i=1}^{n-1} A_i\right) > 0 \Rightarrow$

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P\left(A_i | \bigcap_{j=1}^{i-1} A_j\right)$$

Proof. Procederemos por inducción matemática sobre n.

(*Base inductiva*). Tenemos que:

$$P(A_1 \cap A_2) = P(A_1) \cdot P(A_2 \mid A_1)$$

que es la definición de probabilidad condicional, siempre que $P(A_1) > 0$. Así, el resultado se cumple para $n = 2$.

(Paso inductivo). Supongamos que para cierto $k \geqslant 2$, se cumple:

$$P\left(\bigcap_{i=1}^{k} A_i\right) = \prod_{i=1}^{k} P\left(A_i \,\middle|\, \bigcap_{j=1}^{i-1} A_j\right)$$

Queremos probar que entonces también se cumple para $k + 1$, es decir:

$$P\left(\bigcap_{i=1}^{k+1} A_i\right) = \prod_{i=1}^{k+1} P\left(A_i \,\middle|\, \bigcap_{j=1}^{i-1} A_j\right)$$

Notamos que:

$$P\left(\bigcap_{i=1}^{k+1} A_i\right) = P\left(\left(\bigcap_{i=1}^{k} A_i\right) \cap A_{k+1}\right)$$

Usando la definición de probabilidad condicional:

$$P\left(\bigcap_{i=1}^{k+1} A_i\right) = P\left(\bigcap_{i=1}^{k} A_i\right) \cdot P\left(A_{k+1} \,\middle|\, \bigcap_{i=1}^{k} A_i\right)$$

Aplicando la hipótesis inductiva al primer término:

$$= \left(\prod_{i=1}^{k} P\left(A_i \,\middle|\, \bigcap_{j=1}^{i-1} A_j\right)\right) \cdot P\left(A_{k+1} \,\middle|\, \bigcap_{i=1}^{k} A_i\right)$$

que equivale a:

$$= \prod_{i=1}^{k+1} P\left(A_i \,\middle|\, \bigcap_{j=1}^{i-1} A_j\right)$$

$\square$

Teorema 1.9. Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y sea $B \in \mathcal{F}$ tal que $P(B) > 0$. Sea $Q(\cdot) = P(\cdot|B)$ Entonces $Q(\cdot)$ es una medida de probabilidad donde

(Q_1) $Q(A) \geqslant 0 \ \forall A \in \mathcal{F}$

(Q_2) $Q(\Omega) = 1$

(Q_3) Si $\{A_i\}_{i=1}^{\infty}$ tal que $A_i \cap A_j = \emptyset \ \forall i \neq j \Rightarrow Q(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} Q(A_i)$

Proof. Veamos que es cierto

(Q_1) $Q(A) = P(A|B) = \frac{P(A\cap B)}{P(B)} \Rightarrow P(A \cap B) \geqslant 0$ y $P(B) > 0 \Rightarrow Q(A) \geqslant 0$.

(Q_2) $Q(\Omega) = P(\Omega|B) = \frac{P(\Omega\cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$

(Q_3) Note usted que

$$Q(\bigcup_{i=1}^{\infty} A_i) = P(\bigcup_{i=1}^{\infty} A_i|B) = \frac{P[(\bigcup_{i=1}^{\infty} A_i) \cap B]}{P(B)} = \frac{P[\bigcup_{i=1}^{\infty}(A_i \cap B)]}{P(B)}$$

$$\Rightarrow \frac{\sum_{i=1}^{\infty}(A_i \cap B)}{P(B)} = \sum_{i=1}^{\infty} \frac{P(A_i \cap B)}{P(B)} = \sum_{i=1}^{\infty} P(A_i|B) = \sum_{i=1}^{\infty} Q(A_i)$$

$\therefore Q$ es una medida de probabilidad. $\square$

Definición 1.13 (Independencia de dos eventos). Sean $A, B \in \mathcal{F}$, se dice que A y B son estrictamente independientes $\Leftrightarrow$

$$P(A|B) = P(A) \Leftrightarrow A \perp B \Rightarrow \frac{P(A \cap B)}{P(B)} = P(A) \therefore A \perp B \Leftrightarrow P(A \cap B) = P(A) \times P(B)$$

Teorema 1.10. Sean $A, B \in \mathcal{F}$ tal que $A \perp B \Rightarrow$

(c_1) $A \perp B^c$

(c_2) $A^c \perp B$

(c_3) $A^c \perp B^c$.

Proof. Veamos que es cierto

(c_1) Se sigue que

$$P(A \cap B^c) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)[1 - P(B)] = P(A)P(B^c)$$

(c_2) De forma análoga a (c_1)

(c_3) Desarrollando, tenemos que

$$\begin{aligned} P(A^c \cap B^c) &= P[(A \cup B)^c] = 1 - P(A \cup B) \\ &= 1 - P(A) - P(B) + P(A \cap B) \\ &= 1 - P(A) - P(B) + P(A)P(B) \\ &= [1 - P(A)][1 - P(B)] = P(A^c)P(B^c) \end{aligned}$$

□

Capítulo 2

Variable Aleatoria

Definición 2.1 (Variable Aleatoria). Dado un experimento en un espacio muestral Ω, una variable aleatoria (VA) X es una función que va de Ω a $\mathbb{R}$, es decir

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

Definición 2.2 (VA Discreta). Decimos que una variable aleatoria X es discreta si hay una lista finita de valores $a_1, a_2, ...a_n$ o una lista infinita numerable $a_1, a_2, ...$ tales que

$$\sum_i P(X = a_i \text{ para alguna } i = 1, 2, ...) = 1$$

Definición 2.3 (PMF). La función de masa de probabilidad (PMF) de una VA discreta es la función p_X dada por $p_X(X) = P(X = x)$. Formalmente:

$$p(X = x) = P\left(\left\{s \in \Omega \mid X(s) = x\right\}\right) = P(X^{-1}(x))$$

Definición 2.4 (CDF). Decimos que la función de distribución acumulada (CDF) de una VA llamada X cualquiera es la función F_X dada por $F_X(x) = P(X \leqslant x)$ que cumple

($\mathcal{C}_1$) Si $x \leqslant y \Rightarrow F_X(x) \leqslant F_X(y)$

($\mathcal{C}_2$) $\forall\, a \Rightarrow F_X(a) = \lim\limits_{x \to a^+} F_X(x)$, la función es continua por la derecha.

($\mathcal{C}_3$) $\lim\limits_{x \to -\infty} F_X(x) = 0$ y $\lim\limits_{x \to \infty} F_X(x) = 1$

Teorema 2.1. Sea X una VA discreta. Su PMF p_X cumple las siguientes propiedades:

(P_1) $\forall\, j \in 1, ... \Rightarrow p_X(x_j) > 0$. Las VA X con probabilidad 0 no se enlistan.

(P_2) $\sum\limits_{i=1}^{\infty} p_X(x_i) = 1$

Proof. La primera es trivial por (P_2) de Definición 1.4. Para la segunda, se tiene que:

$$\sum_{i=1}^{\infty} P(X = x_i) = P\left(\bigcup_{i=1}^{\infty} \{X = x_i\}\right) = P(X = x_1 \text{ o } X = x_2 \text{ o } ...) = 1$$

$\square$

Definición 2.5. Decimos que una VA X es continua si su CDF es diferenciable.

Observación. No todas las variables aleatorias continuas tienen función de densidad de probabilidad (PDF) en el sentido clásico, ya que la función de distribución acumulada (CDF) puede no ser diferenciable en todos los puntos.

Definición 2.6 (PDF). Sea $F_X(x)$ la CDF de una VA continua $X \Rightarrow$ denotamos $f(x)$ como

$$f(x) = \frac{dF(x)}{dx} = F'(x)$$

y le llamamos a $f(x)$ la PDF de la VA X.

Notación. Por convenencia, usamos f_X para referirnos tanto a PMF y PDF.

Corolario. Notemos que podemos escribir a la CDF como

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

Definición 2.7. Si una VA continua X tiene una PDF $f(x)$, y se tiene que $a < b$, entonces, la probabilidad de que X caiga en el intervalo $[a, b]$ es

$$P(a \leqslant x \leqslant b) = \int_{a}^{b} f(x)dx$$

Teorema 2.2. Sea X una VA continua. Su PDF f_X cumple las siguientes propiedades:

(P_1) $f(x) \geqslant 0$

(P_2) $\int_{-\infty}^{\infty} f(x)dx = 1$

Proof. Notemos que es análogo a una VA discreta, cambiando la suma por la integral. □

Observación. La PMF da probabilidades exactas en puntos porque la variable es discreta. La PDF no da probabilidades puntuales, sino densidades; la probabilidad se obtiene integrando la PDF en un intervalo. Por eso, la PDF no es una PMF para variables continuas.

Notación. Sea X una variable aleatoria con CDF $F_X(x) = P(X \leqslant x)$. La función de supervivencia $S_X(x)$ se define como la probabilidad de que X sea mayor que x, es decir,

$$S_X(x) = P(X > x) = 1 - F_X(x).$$

Definición 2.8. La σ-álgebra de Borel en $\mathbb{R}$, denotada $\mathcal{B}(\mathbb{R})$, es la σ-álgebra generada por los intervalos abiertos de $\mathbb{R}$ i.e.

$$\mathcal{B}(\mathbb{R}) = \sigma\left(\{(a, b) : a, b \in \mathbb{R}, a < b\}\right)$$

donde $\sigma(\cdot)$ denota la operación de generar la sigma-álgebra más pequeña que contiene a ese conjunto.

Notación (Espacio Medible). Un espacio medible es un par $(\Omega, \mathcal{F})$ donde Ω es un conjunto y $\mathcal{F}$ es una σ-álgebra sobre Ω.

Observación. Todo espacio de probabilidad es un espacio medible.

Definición 2.9 (Función Medible)**.** Sea $(\Omega, \mathcal{F})$ un espacio medible y $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ el espacio real con su σ-álgebra de Borel. Una función $f : \Omega \to \mathbb{R}$ se dice medible si

$$\forall\, B \in \mathcal{B}(\mathbb{R}) \Rightarrow f^{-1}(B) \in \mathcal{F}$$

Es decir, la preimagen de cualquier conjunto de Borel es un evento.

Definición 2.10 (Variable Aleatoria Medible)**.** Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad y $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ el espacio real con su sigma-álgebra de Borel. Una función

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

se llama variable aleatoria medible si $\forall\, B \in \mathcal{B}(\mathbb{R})$ se cumple que

$$X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}$$

Es decir, la preimagen de cualquier conjunto de Borel es un evento en Ω.

Definición 2.11 (Vector Aleatorio Discreto)**.** Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Un vector aleatorio discreto en $\mathbb{R}^n$ es una función medible

$$\mathbf{X} = (X_1, X_2, \ldots, X_n) : \Omega \to \mathbb{R}^n$$

tal que $\mathbf{X}$ toma valores en un subconjunto numerable de $\mathbb{R}^n$ y se cumple

$$\sum_{\mathbf{x} \in \mathbb{R}^n} P(\mathbf{X} = \mathbf{x}) = 1$$

Definición 2.12 (Vector Aleatorio Absolutamente Continuo)**.** Sea $\mathbf{X} = (X_1, X_2, \ldots, X_n) : \Omega \to \mathbb{R}^n$ un vector aleatorio. Decimos que $\mathbf{X}$ es absolutamente continuo si existe una función $f_{\mathbf{X}} : \mathbb{R}^n \to \mathbb{R}$ tal que para todo conjunto medible $A \subseteq \mathbb{R}^n$ se cumple:

$$P(\mathbf{X} \in A) = \int_A f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}$$

A esta función $f_{\mathbf{X}}$ se le llama función de densidad conjunta de $\mathbf{X}$.

Capítulo 3

Momentos de VA

3.1. $\mathbb{E}[X]$ y $\mathrm{Var}[X]$

Definición 3.1 (Valor Esperado)**.** Sea X una VA discreta con PMF $f_X(x) = P(X = x)$ y soporte $S \subseteq \mathbb{R}$ tal que $S = \{x \in \mathbb{R} \mid f_X(x) > 0\}$. Supongamos que $\sum_{x \in S} |x| f_X(x) < \infty \Rightarrow$ se define la esperanza de la VA X como

$$\mathbb{E}[X] := \sum_{x \in S} x f_X(x) \in \mathbb{R}$$

En caso de que X sea una VA continua tal que $\int_{-\infty}^{\infty} |x| f_X(x)\, dx < \infty \Rightarrow$

$$\mathbb{E}[X] := \int_{-\infty}^{\infty} x f_X(x)\, dx \in \mathbb{R}$$

Observación. Sea $0 < p < 1 \Rightarrow$

$$\sum_{x=0}^{\infty} x p^x = \sum_{x=1}^{\infty} x p^{x-1} p = p \sum_{x=1}^{\infty} x p^{x-1} = p \sum_{x=1}^{\infty} \frac{d}{dp} p^x$$

$$= p \frac{d}{dp} \sum_{x=1}^{\infty} p^x = p \frac{d}{dp} \left(\frac{p}{1-p} \right) = p \frac{(1-p)p}{(1-p)^2} = \frac{p}{(1-p)^2}$$

Lema 3.1. Sea $Y \geqslant 0$ una VA no negativa y continua $\Rightarrow \mathbb{E}[Y] = \int_0^{\infty} (1 - F_Y(y)) dy$

Proof. Como $F_Y(y) = P(Y \leqslant y) \Rightarrow$

$$\int_0^{\infty} (1 - F_Y(y)) dy = \int_0^{\infty} P(Y > y) dy$$

$$= \int_0^{\infty} \left(\int_y^{\infty} f_Y(t) dt \right) dy = \int_0^{\infty} \int_0^t (f_Y(t) dy) dt$$

$$= \int_0^{\infty} f_Y(t) \left(\int_0^t dy \right) dt = \int_0^{\infty} t f_Y(t) dt = \mathbb{E}[Y]$$

□

Observación. Recordemos que F_Y es la función de distribución, f_Y su derivada: la densidad.

Teorema 3.1. Sea X una VA continua. Sea $g(x)$ cualquier función no negativa real de X

$$\mathbb{E}[g(x)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

Proof. Por el Lema 3.1, tenemos que

$$\mathbb{E}[g(x)] = \int_0^{\infty} P(g(x) > y) dy = \int_0^{\infty} \left[\int_{\{x \in \mathbb{R} | g(x) > y\}} f_X(x) dx \right] dy$$

Integramos en la región de $\mathbb{R}^2 = \{(x, y) \mid 0 < y < g(x) < \infty\}$

$$\Rightarrow \int_{\{x \in \mathbb{R} | g(x) > y\}} \left(\int_0^{g(x)} f_X(x) dy \right) dx = \int_{\{x \in \mathbb{R} | g(x) > y\}} f_X(x) \left(\int_0^{g(x)} dy \right) dx$$

$$= \int_{\{x \in \mathbb{R} | g(x) > y\}} g(x) f_X(x) dx$$

□

Corolario. En caso de ser una VA discreta tenemos que

$$\mathbb{E}[g(x)] = \sum_{X \in S} g(x) f_X(x)$$

Proof. Análogo al Teorema 3.1. □

Ejemplo 3.1. Sea X una VA con FDP

$$f_X(x) = \frac{c}{x!} \qquad x = 0, 1, 2, \ldots$$

Proof. Encontremos $\mathbb{E}[X]$

$$\mathbb{E}[X] = \sum_{x=0}^{\infty} x f_X(x) = \sum_{x=1}^{\infty} x \frac{e^{-1}}{x!}$$

Sea $y = x - 1$

$$= \sum_{x=1}^{\infty} \frac{e^{-1}}{(x-1)!} = e^{-1} \sum_{y=0}^{\infty} \frac{1}{y!} = e^1 \cdot e = \frac{e}{e} = 1$$

Sea $g(x) = x^2$.

Ahora calculemos $\mathbb{E}[X^2]$

$$\mathbb{E}[X^2] = \sum_{x=0}^{\infty} x^2 f_X(x) = \sum_{x=1}^{\infty} x^2 \frac{e^{-1}}{x!}$$

Sea $y = x - 1$

$$= e^{-1} \sum_{x=1}^{\infty} \frac{x}{(x-1)!} = e^{-1} \sum_{y=0}^{\infty} \frac{y+1}{y!} = e^{-1} \left[\sum_{y=0}^{\infty} \frac{y}{y!} + \sum_{y=0}^{\infty} \frac{1}{y!} \right]$$

$$= \frac{e + e}{e} = \frac{2e}{e} = 2$$

□

Observación. $\mathbb{E}[X^2] \neq \mathbb{E}[X]^2$

Teorema 3.2 (Propiedades del Valor Esperado). $\mathbb{E}[\cdot]$ cumple con lo siguiente

($\mathbb{E}_1$) $\forall\, \lambda \in \mathbb{R} \Rightarrow \mathbb{E}[\lambda] = \lambda$

($\mathbb{E}_2$) Para cualesquiera dos funciones $g_1, g_2 : \mathbb{R} \to \mathbb{R}$ y $\forall\, \lambda_1 \lambda_2, \lambda_3 \in \mathbb{R} \Rightarrow$

$$\mathbb{E}[\lambda_1 g_1 \pm \lambda_2 g_2 + \lambda_3] = \lambda_1 \mathbb{E}[g_1] \pm \lambda_2 \mathbb{E}[g_2] + \lambda_3$$

($\mathbb{E}_3$) Sea X una VA no negativa $P(X \geqslant 0) = 1 \Rightarrow \mathbb{E}[X] \geqslant 0$

($\mathbb{E}_4$) Sean $g_1, g_2 : \mathbb{R} \to \mathbb{R}$. Si $g_1(X) \geqslant g_2(X) \Rightarrow$

$$\mathbb{E}[g_1(X)] \geqslant \mathbb{E}[g_2(X)]$$

($\mathbb{E}_5$) Sea X una VA cualquiera con esperanza $\mathbb{E}[X] \Rightarrow \mathbb{E}[|X|] \geqslant \big|\mathbb{E}[X]\big|$

($\mathbb{E}_6$) Sea X una VA con FDP $f_X(x)$ continua y soporte $S = \mathbb{R} \Rightarrow$

$$E[X] = \int_0^{\infty} (1 - F_X(x))dx - \int_{-\infty}^{0} F_X(x)dx$$

($\mathbb{E}_7$) $\forall\, \omega \in \Omega \Rightarrow \big|X(\omega)\big| < M$ t.q $M > 0 \Rightarrow \mathbb{E}[X] \leqslant M < \infty$

Proof. Probemos el Teorema 3.2

($\mathbb{E}_1$) $\mathbb{E}[\lambda] = \int_{-\infty}^{\infty} \lambda f_X(x) = \lambda \int_{-\infty}^{\infty} f_X(x) = \lambda \cdot 1 = \lambda$

($\mathbb{E}_2$) $\mathbb{E}[\lambda_1 g_1 \pm \lambda_2 g_2 + \lambda_3] = \int_{-\infty}^{\infty} (\lambda_1 g_1 \pm \lambda_2 g_2 + \lambda_3) f_X(x)dx = \lambda_1 \mathbb{E}[g_1] \pm \lambda_2 \mathbb{E}[g_2] + \lambda_3$

($\mathbb{E}_3$) $P(X \geqslant 0) = 1 \Rightarrow x f_X(x) \geqslant 0 \Rightarrow \int x f_X(x)dx \geqslant \int 0dx \Rightarrow \mathbb{E}[X] \geqslant 0$

($\mathbb{E}_4$) Notemos que

$$g_1(X) \geqslant g_2(X) \Rightarrow g_1(X) - g_2(X) \geqslant 0$$

Por el inciso ($\mathbb{E}_3$) s.t.q.

$$\mathbb{E}[g_1(X) - g_2(X)] \geqslant 0$$

Y ahora, por ($\mathbb{E}_2$) s.t.q.

$$\mathbb{E}[g_1(X)] - \mathbb{E}[g_2(X)] \geqslant 0 \Rightarrow \mathbb{E}[g_1(X)] \geqslant \mathbb{E}[g_2(X)]$$

($\mathbb{E}_5$) $\mathbb{E}[-|X|] \leqslant \mathbb{E}[X] \leqslant \mathbb{E}[|X|] \Rightarrow -\mathbb{E}[|X|] \leqslant \mathbb{E}[X] \leqslant \mathbb{E}[|X|] \Rightarrow \mathbb{E}[|X|] \geqslant \big|\mathbb{E}[X]\big|$

($\mathbb{E}_6$) Notemos que

$$\int_{-\infty}^{0} x f_X(x)dx + \int_0^{\infty} x f_X(x)dx$$
$$= \int_0^{\infty} (1 - F_X(x))dx - \int_{-\infty}^{0} F_X(x)dx$$

$\therefore$ es cierto el Teorema 3.2. $\square$

Observación. Una VA X no tiene esperanza finita cuando $\mathbb{E}[X] = \infty$

Definición 3.2 (Varianza). Sea X una VA tal que $\mathbb{E}[X^2] < \infty$. Definimos la varianza de X, denotada como $\text{Var}[X]$ o σ^2, como

$$\text{Var}[X] = \sigma^2 = \mathbb{E}[(X-\mu)^2] \text{ donde } \mu = \mathbb{E}[X]$$

Observación. Esto significa que $\text{Var}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$

Notación. Escribimos en el caso discreto y en el caso continuo respectivamente

$$\sum_{\forall X}(X-\mathbb{E}[X])^2 f_X(x) \qquad \int_{-\infty}^{\infty}(X-\mathbb{E}[X])^2 f_X(x)dx$$

Teorema 3.3 (Propiedades de la Varianza). Sea X una VA con $\mathbb{E}[X^2] < \infty \Rightarrow \text{Var}[X]$ cumple con lo siguiente

(σ_1^2) $\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}^2[X]$

(σ_2^2) $\forall\, \lambda \in \mathbb{R} \Rightarrow \text{Var}[\lambda \cdot X] = \lambda^2 \cdot \text{Var}[X]$

(σ_3^2) $\forall\, \lambda \in \mathbb{R} \Rightarrow \text{Var}[X+\lambda] = \text{Var}[X]$

(σ_4^2) $\text{Var}[X] \geqslant 0$

(σ_5^2) $\forall\, \lambda \in \mathbb{R} \Rightarrow \text{Var}[X] = 0 \Leftrightarrow X = \lambda$

Proof. Probemos el Teorema 3.3

(σ_1^2) $\text{Var}[X] = \mathbb{E}[X-\mu]^2 = \mathbb{E}\left[X^2 - 2\mu X + \mu^2\right] = \mathbb{E}\left[X^2\right] - 2\mu\mathbb{E}[X] + \mu^2$

$= \mathbb{E}\left[X^2\right] - 2\mu^2 + \mu^2 = \mathbb{E}\left[X^2\right] - \mu^2 = \mathbb{E}[X^2] - \mathbb{E}^2[X]$

(σ_2^2) $\text{Var}[\lambda \cdot X] = \mathbb{E}[(\lambda X - \mathbb{E}[\lambda X])^2] = \mathbb{E}[(\lambda X - \lambda\mathbb{E}[X])^2] = \mathbb{E}[\lambda^2(X-\mathbb{E}[X])^2]$

$\lambda^2 \cdot \mathbb{E}[(X-\mathbb{E}[X])^2] = \lambda^2 \cdot \text{Var}[X]$

(σ_3^2) $\text{Var}[X+\lambda] = \mathbb{E}[(X+\lambda-\mathbb{E}[X+\lambda])^2] = \mathbb{E}[(X-\mathbb{E}[X])^2] = \text{Var}[X]$

(σ_4^2) Trivial por ($\mathbb{E}_3$) del Teorema 3.2.

(σ_5^2) $\Longleftarrow$ Supongamos que $X = \lambda$

$$\text{Var}[\lambda] = \mathbb{E}[(\lambda - \mathbb{E}[\lambda])^2] = \mathbb{E}[(\lambda-\lambda)^2] = \mathbb{E}[0] = 0$$

$\Longrightarrow$ Supongamos que $\text{Var}[X] = 0$

$$0 = \text{Var}[X] = \sum_{\forall\, X}(X-\mathbb{E}[X])^2 f_X(x) = 0$$

$$\Rightarrow \sum_{\forall\, X}(X-\mathbb{E}[X])^2 = 0 \Rightarrow X = \mathbb{E}[X]$$

$$\Rightarrow \forall\, \omega \in \Omega \Rightarrow X(\omega) = \mathbb{E}[X] = \lambda$$

$\therefore$ es cierto el Teorema 3.3. □

Definición 3.3 (Desviación Estandar). Se define como la raíz postiva de la Definición 3.2

$$\text{SD}[X] = \sigma = \sqrt{\text{Var}[X]}$$

3.2. Momentos y FGM

Definición 3.4 (r-ésimo momento). Sea X una VA y $r \in \mathbb{Z}^+$ tal que $\mathbb{E}[X^r] < \infty \Rightarrow$ se define al r-esimo momento de la VA X como $\mathbb{E}[X^r]$

Ejemplo 3.2 (r-ésimo momento alrededor de λ). Sea X una VA tal que $\mathbb{E}[X^r] < \infty$, se define al r-ésimo momento alrededor de λ como

$$\mathbb{E}[(X-\lambda)^r]$$

Ejemplo 3.3 (r-ésimo momento central). Del Ejemplo 3.2 cuando tomamos a $\lambda = \mathbb{E}[X]$ s.t.q.

$$\mathbb{E}[(X-\mathbb{E}[X])^r]$$

Observación. Notemos que la varianza es el segundo momento central de X.

Notación. Cuando $r = 3$ llamamos al tercer momento el coeficiente de asimetría α.

$$\alpha = \frac{\mathbb{E}[(X-\mu)^3]}{\sigma^3}$$

Cuando $r = 4$ llamamos al cuarto momento el coeficiente de kurtosis $\mathcal{K}$.

$$\mathcal{K} = \frac{\mathbb{E}[(X-\mu)^4]}{\sigma^4}$$

Definición 3.5 (Función Generadora de Momentos). Sea X una VA con FDP $f_X(x)$ y $h > 0$. Supongamos que $\forall\, t \in (-h, h)$ s.t.q. $\mathbb{E}[e^{tX}] < \infty$. Se define a la FGM como

$$m_X(t) := \mathbb{E}[e^{tX}]$$

donde $m_X(t)$ es una función de valor real.

Notación. Escribimos a la FGM en el caso discreto y en el caso continuo respectivamente

$$m_X(t) = \sum_{x\in X} e^{tx} f_X(x) \qquad\qquad m_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx$$

Lema 3.2. Sea X una VA con n-ésimo momento finito.

$$\Rightarrow \mathbb{E}[|X^n|] = n\int_0^\infty x^{n-1}(1-F_X(x)) + n\int_0^\infty |x|^{n-1}F_X(-x)$$

Proof. Notemos que

$$\mathbb{E}[|X^n|] = \int_{-\infty}^{\infty} |x|^n f_X(x) = \int_{-\infty}^{0} |x|^n f_X(x)dx + \int_0^\infty |x|^n f_X(x)dx$$

Simplificaremos estas dos expresiones para llegar al resultado

$$\int_{-\infty}^{0} |x|^n f_X(x)dx = \int_{-\infty}^{0} -x^n f_X(x)dx$$

$$= (-x^n)F_X(x)\mid_{-\infty}^{n} + n\int_{-\infty}^{0} (-x)^{n-1}F_X(x)dx = 0 + n\int_0^\infty |x|^{n-1}F_X(-x)$$

Lo cual consiste en la segunda parte de la igualdad.

$$\int_0^\infty |x|^n f_X(x)dx = \int_0^\infty x^n f_X(x)dx$$

$$= -x^n(1 - F_X(x)) \mid_0^\infty + n\int_0^\infty x^{n-1}(1 - F_X(x)) = n\int_0^\infty x^{n-1}(1 - F_X(x))$$

Lo cual compreba al Lema 3.2 □

Lema 3.3. Sea $h > 0$ y X una VA cuya FGM $m_X(t)$ existe para $t \in (-h, h)$

$$\Rightarrow m_X(t) = 1 + t\left[\int_0^\infty (1 - F_X(x))e^{tx}dx - \int_{-\infty}^0 e^{tx}F_X(x)dx\right]$$

Proof. Notemos lo siguiente

$$m_X(t) = \int_{-\infty}^\infty e^{tx}f_X(x)dx = \int_{-\infty}^0 e^{tx}f_X(x)dx + \int_0^\infty e^{tx}f_X(x)dx$$

Simplificamos, primero a la primera porción de la suma.

$$\int_{-\infty}^0 e^{tx}f_X(x)dx = e^{tx}F_X(x) \mid_{-\infty}^0 - t\int_{-\infty}^0 e^{tx}F_X = F_X(0) - t\int_{-\infty}^0 e^{tx}F_X$$

Ahora, simplificamos la segunda.

$$\int_0^\infty e^{tx}f_X(x)dx = -e^{tx}(1 - F_X(x)) \mid_0^\infty + t\int_0^\infty e^{tx}dx(1 - F_X(x))$$

$$= 1 - F_X(0) + t\int_0^\infty e^{tx}dx(1 - F_X(x))$$

Ahora combinamos ambos resultados

$$F_X(0) - t\int_{-\infty}^0 e^{tx}F_X + 1 - F_X(0) + t\int_0^\infty e^{tx}dx(1 - F_X(x))$$

$$= 1 + t\left[\int_0^\infty (1 - F_X(x))e^{tx}dx - \int_{-\infty}^0 e^{tx}F_X(x)dx\right]$$

$\therefore$ el Lema 3.3 es cierto. □

Teorema 3.4 (Propiedades de FGM)**.** Sea X una VA $\Rightarrow m_X(t)$ cumple con lo siguiente

(m_1) $m_X(t) = \sum_{k=0}^\infty \frac{t^k \mathbb{E}[X^k]}{k!}$

(m_2) $m_X(0) = 1$

(m_3) $\frac{d}{dt}m_X(t) \mid_{t=0} = \mathbb{E}[X]$

(m_4) $\frac{d^n}{d^n t}m_X(t) \mid_{t=0} = \mathbb{E}[X^n]$

(m_5) Sea X una VA con FGM $m_X(t)$. Sea $Y = ax + b$ con $a, b \in \mathbb{R}$ y $a \neq 0$

$$\Rightarrow m_Y(t) = e^{bt}m_X(at)$$

(m_6) Sean X, Y VAs $\Rightarrow m_X(t) = m_Y(t) \Leftrightarrow f_X = f_Y$

(m_7) Sea $h > 0$ y X VA con FGM $m_X(t) < \infty$ para $-h < t < h \Rightarrow \mathbb{E}[X^n]$ existe $\forall\, n \geqslant 1$

Proof. Probemos el Teorema 3.4

(m_1) Note que, por Taylor, por teorema de convergencia denominada y por el Teorema 3.1

$$m_X(t) = \mathbb{E}[e^{tX}] = \mathbb{E}\left[\sum_{k=0}^{\infty} \frac{(tk)^k}{k!}\right] = \sum_{k=0}^{\infty} \mathbb{E}\left[\frac{(tk)^k}{k!}\right]$$

$$= \sum_{k=0}^{\infty} \frac{t^k \mathbb{E}[X^k]}{k!} = 1 + t\mathbb{E}[X] + \frac{t^2 \mathbb{E}[X^2]}{2} + \frac{t^3 \mathbb{E}[X^3]}{6} + \dots$$

(m_2) $\mathbb{E}[e^{0x}] = \mathbb{E}[e^0] = \mathbb{E}[1] = 1$

(m_3) Tomemos la primera derivada de la FGM, y por (m_1) s.t.q.

$$m_X'(t) = \frac{d}{dt}\left[1 + t\mathbb{E}[X] + \frac{t^2}{2}\mathbb{E}[X^2] + \dots\right] = \mathbb{E}[X] + t\mathbb{E}[X^2] + \frac{t^2}{2}\mathbb{E}[X^3] + \dots$$

Evaluando en $t = 0$

$$\mathbb{E}[X] + 0 \cdot \mathbb{E}[X^2] + \frac{0^2}{2}\mathbb{E}[X^3] + \dots = \mathbb{E}[X]$$

(m_4) Por inducción para $n \geq 1$.

(m_5) $m_Y(t) = \mathbb{E}[e^{tY}] = \mathbb{E}[e^{t(ax+b)}] = \mathbb{E}[e^{atx+tb}] = \mathbb{E}[e^{atx}e^{tb}] = e^{tb}\mathbb{E}[e^{atx}] = e^{bt}m_X(at)$

(m_6) $m_X(t) = m_Y(t) \Leftrightarrow \mathbb{E}[e^{tX}] = \mathbb{E}[e^{tY}] \Leftrightarrow f_X = f_Y$

(m_7) Ya que $X < |X|$, basta probar que $\mathbb{E}[|X|^n] < \infty$. Por el Lema 3.2 basta probar:

$$\int_0^{\infty} x^{n-1}(1 - F_X(x))dx < \infty \qquad \text{y} \qquad \int_{-\infty}^{0} |x|^{n-1} F_X(x)dx < \infty$$

Primero, notemos que para $t \in (-h, h)$

$$m_X(t) = \int_{-\infty}^{\infty} e^{tx} f_X(x)dx < \infty \Rightarrow \int_0^{\infty} e^{tx}(1 - F_X(x))dx < \infty$$

Tomamos a $t \in (0, h) \Rightarrow \frac{(tx)^{n-1}}{(n-1)!} < e^{tx} \Rightarrow$

$$x^{n-1} < \frac{(n-1)! \cdot e^{tx}}{t^{n-1}} \Rightarrow x^{n-1}(1 - F_X(x)) < \frac{(n-1)! \cdot e^{tx}}{t^{n-1}}(1 - F_X(x))$$

$$\Rightarrow \int_0^{\infty} x^{n-1}(1 - F_X(x)) < \int_0^{\infty} \frac{(n-1)! \cdot e^{tx}}{t^{n-1}}(1 - F_X(x))$$

Viendo el comienzo de la desigualdad, y por el Lema 3.3, s.t.q.

$$= \frac{(n-1)!}{t^{n-1}} \int_0^{\infty} e^{tx}(1 - F_X(x)) < \infty$$

Segundo, como lo pasado viene de $m_X(t) < \infty$, notemos que si $t \in (-h, 0)$

$$t\int_{-\infty}^{0} e^{tx} F_X(x)dx < \infty \Rightarrow tx > 0$$

Ahora, notemos qué

$$\frac{|tx|^{n-1}}{(n-1)!} \leqslant e^{|tx|} = e^{tx} \Rightarrow |x|^{n-1} < \frac{e^{tx}(n-1)!}{|t|^{n-1}}$$

Ahora, como anteriormente, aplicamos esta integral.

$$\int_{-\infty}^{0} |x|^{n-1} F_X(x) dx < \int_{-\infty}^{0} \frac{(n-1)! \cdot e^{tx}}{|t|^{n-1}} F_X(x) < \infty$$

Hemos probado ambas desigualdades $\therefore \forall\, n \geqslant 1 \Rightarrow \mathbb{E}[|X|^n] < \infty$

$\therefore$ el Teorema 3.4 es cierto. □

Definición 3.6 (Función Cumulante). Sea X una VA con FGM $m_X(t) < \infty$ para $t \in (-h, h)$. Se define a la función $\Psi_X(t)$ como

$$\Psi_X(t) := \ln[m_X(t)]$$

Teorema 3.5 (Propiedades de Ψ). Sea X una VA con FGM $m_X(t) < \infty$ para $t \in (-h, h) \Rightarrow$ $\Psi_X(t)$ cumple con lo siguiente

(Ψ_1) $\Psi_X(0) = 0$

(Ψ_2) $\frac{d}{dt}\Psi_X(t) \mid_{t=0} = \mathbb{E}[X]$

(Ψ_3) $\frac{d^2}{d^2t}\Psi_X(t) \mid_{t=0} = \mathrm{Var}[X]$

Proof. Probemos el Teorema 3.5

(Ψ_1) $\Psi_X(0) = \ln[m_X(0)] = \ln[1] = 0$

(Ψ_2) $\frac{d}{dt}\Psi_X(t) = \frac{d}{dt}\ln[m_X(t)] = \frac{m'_X(t)}{m_X(t)} \mid_{t=0} = \frac{\mathbb{E}[X]}{1} = \mathbb{E}[X]$

(Ψ_3) Notemos, por (Ψ_2) que

$$\frac{d^2}{d^2t}\Psi_X(t) = \frac{d^2}{d^2t}\ln[m_X(t)] = \frac{m'_X(t)}{m_X(t)} \mid_{t=0} = \frac{m_X(t)m''_X(t) - m'_X(t)m'_X(t)}{(m'_X(t))^2}$$

$$= m''_0(t) - (m'_X(0))^2 = \mathbb{E}[X^2] - \mathbb{E}^2[X] = \mathrm{Var}[X]$$

$\therefore$ es cierto el Teorema 3.5. □

Definición 3.7 (Percentiles). Sean X una VA y $0 < p < 1$. Se define el percentil $p100\,\%$ como $\pi_p \in \mathbb{R}$ tal que

$$P(x \leqslant \pi_p) = p \qquad \text{y} \qquad P(x > \pi_p) = 1 - p$$

Definición 3.8 (Mediana). La mediana es el percentil $50\,\%$ oara X i.e. $\mathrm{Me} = \pi_{0,5}$

Definición 3.9 (Cuartiles). Aquellos valores que dividen la distribución en cuatro partes.

Observación. Los mediana es el segundo cuartil.

Definición 3.10 (Moda)**.** Sea X una VA con FDP $f_X(x)$. La moda de X es el valor $\mathrm{Mo} \in \mathbb{R}$ que maximiza a $f_X(x)$.

En el caso continuo, es el máximo de $f_X(x)$

En el caso discreto, es $\mathrm{Mo} = x$ tal que

$$\frac{f(x_j)}{f(x_{j-1})} \geqslant 1 \qquad \text{y} \qquad \frac{f(x_j)}{f(x_{j+1})} \leqslant 1$$

Observación. En caso en el que Mo es unico, se dice que es una distribución unimodal. Si tiene dos, es bimodal. Si tiene más, se le llama multimodal.

Definición 3.11 (Covarianza)**.** Sean X y Y dos variables aleatorias con esperanza finita. La covarianza entre X y Y se define como

$$\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$$

Equivalentemente, si $\mathbb{E}[XY]$ existe,

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Definición 3.12 (Matriz de Covarianza)**.** Sea $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ un vector aleatorio con medias finitas. La matriz de covarianza de $\mathbf{X}$ es la matriz simétrica definida por

$$\Sigma = \mathrm{Cov}(\mathbf{X}) = \left[\mathrm{Cov}(X_i, X_j)\right]_{i,j=1}^{n}$$

Es decir, Σ es una matriz $n \times n$ donde la entrada (i, j) es la covarianza entre X_i y X_j

Definición 3.13 (Vector Aleatorio Gaussiano)**.** Un vector aleatorio $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ en $\mathbb{R}^n$ se dice gaussiano si para todo $\mathbf{a} \in \mathbb{R}^n$, la combinación lineal $Y = \mathbf{a}^\top \mathbf{X}$ es una variable aleatoria normal unidimensional, i.e.

$$Y \sim \mathcal{N}(\mu, \sigma^2).$$

Equivalentemente, $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$ con vector de medias $\boldsymbol{\mu} \in \mathbb{R}^n$ y matriz de covarianza $\Sigma \in \mathbb{R}^{n \times n}$ simétrica definida positiva.

Definición 3.14 (Esperanza Condicional)**.** Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, X una variable aleatoria integrable, y $\mathcal{G} \subseteq \mathcal{F}$ una sub-σ-álgebra. La esperanza condicional de X dado $\mathcal{G}$, denotada $\mathbb{E}[X \mid \mathcal{G}]$, es cualquier variable aleatoria $\mathcal{G}$-medible que satisface

$$\forall\, G \in \mathcal{G} \Rightarrow \int_G \mathbb{E}[X \mid \mathcal{G}]\, dP = \int_G X\, dP$$

Definición 3.15 (Varianza Condicional)**.** Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad, X una variable aleatoria cuadrado integrable, y $\mathcal{G} \subseteq \mathcal{F}$ una sub-σ-álgebra. La varianza condicional de X dado $\mathcal{G}$, denotada $\mathrm{Var}(X \mid \mathcal{G})$, es la variable aleatoria $\mathcal{G}$-medible definida por

$$\mathrm{Var}(X \mid \mathcal{G}) = \mathbb{E}\left[(X - \mathbb{E}[X \mid \mathcal{G}])^2 \mid \mathcal{G}\right]$$

Además, cumple

$$\forall\, G \in \mathcal{G} \Rightarrow \int_G \mathrm{Var}(X \mid \mathcal{G})\, dP = \int_G (X - \mathbb{E}[X \mid \mathcal{G}])^2\, dP$$

Capítulo 4

Familias Paramétricas

4.1. Discretas

Definición 4.1 (Distribución Uniforme Discreta). Decimos que una VA X tiene distribución uniforme discreta en los valores $x_1, x_2, \dots, x_n$ (es decir, toma valores finitos con igual probabilidad)

$$X \sim \mathcal{U}(x_1, x_2, \dots, x_n)$$

si y solo si su PMF es la siguiente

$$f(x) = \mathcal{U}(x; x_1, \dots, x_n) = \begin{cases} \frac{1}{n} & \text{si } x \in \{x_1, \dots, x_n\} \\ 0 & \text{si } x \notin \{x_1, \dots, x_n\} \end{cases}$$

donde $x_1, \dots, x_n \in \mathbb{R}$ y $x_i \neq x_j$ para $i \neq j$

Observación. En algunos casos, los parámetros de $\mathcal{U}$ son $a, b \in \mathbb{R}$ tal que $n = b - a + 1$

Teorema 4.1 (PMF de Uniforme Discreta). Sea $X \sim \mathcal{U}(x_1, \dots, x_n)$ su PMF $f(x)$ cumple

($\mathcal{U}_1$) $\forall\, x \in \mathbb{R} \Rightarrow f(x) \geqslant 0$

($\mathcal{U}_2$) $\sum_{x \in \mathbb{R}} f(x) = 1$

Proof. Veamos que Teorema 4.1 es cierto.

($\mathcal{U}_1$) $\frac{1}{n} \geqslant 0 \Rightarrow f(x) \geqslant 0$

($\mathcal{U}_2$) $\sum_{x \in \{x_1, \dots, x_n\}} \frac{1}{n} = \frac{n}{n} = 1$

$\therefore$ Teorema 4.1 es cierto. □

Teorema 4.2 (Uniforme Discreta). Sea $X \sim \mathcal{U}(x_1, \dots, x_n) \Rightarrow$

($\mathcal{U}_1$) $m_X(t) = \frac{e^t(1-e^{tn})}{n(1-e^t)}$ FGM

($\mathcal{U}_2$) $\mathbb{E}[X] = \frac{n+1}{2}$ Esperanza

($\mathcal{U}_3$) $\mathrm{Var}[X] = \frac{n^2-1}{12}$ Varianza

Proof. Sean $X \sim \mathcal{U}(\{1, 2, \dots, n\})$, es decir, $P(X = k) = \frac{1}{n}$ para $k = 1, \dots, n$.

($\mathcal{U}_1$) Tomamos la FGM

$$m_X(t) = \mathbb{E}[e^{tX}] = \sum_{k=1}^{n} \frac{1}{n} e^{tk} = \frac{1}{n} \sum_{k=1}^{n} e^{tk} = \frac{1}{n} \cdot e^t \cdot \frac{1 - e^{tn}}{1 - e^t} = \frac{e^t(1 - e^{tn})}{n(1 - e^t)}$$

($\mathcal{U}_2$) Para la esperanza s.t.q.

$$\mathbb{E}[X] = \sum_{k=1}^{n} \frac{1}{n} k = \frac{1}{n} \sum_{k=1}^{n} k = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}$$

($\mathcal{U}_3$) Para la esperanza s.t.q.

$$\mathbb{E}[X^2] = \frac{1}{n} \sum_{k=1}^{n} k^2 = \frac{1}{n} \cdot \frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

$$\begin{aligned} \operatorname{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 &= \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 \\ &= \frac{(n+1)}{6}\left(2n + 1 - \frac{3(n+1)}{2}\right) = \frac{n^2 - 1}{12} \end{aligned}$$

$\therefore$ Teorema 4.2 es cierto. □

Definición 4.2 (Distribución Bernoulli)**.** Decimos que una VA X tiene distribución Bernoulli con parámetro p (es decir, probabilidad de éxito p)

$$X \sim \operatorname{Bern}(p)$$

si y solo si su PMF es la siguiente

$$f(x) = \operatorname{Bern}(x; p) = \begin{cases} p & \text{si } x = 1 \\ q = 1 - p & \text{si } x = 0 \end{cases}$$

donde $p \in [0, 1]$ y $P(X = 1) = p$ y $P(X = 0) = 1 - p$

Observación. La distribución Bernoulli modela eventos discontinuos, es decir, con dos posibles resultados: los cuales se denominan éxito o fracaso.

Teorema 4.3 (PMF de Bernoulli)**.** Sea $X \sim \operatorname{Bern}(p)$ su PMF $f(x)$ cumple

(B_1) $\forall\, x \in \mathbb{R} \Rightarrow f(x) \geqslant 0$

(B_2) $\sum_{x \in \mathbb{R}} f(x) = 1$

Proof. Veamos que Teorema 4.3 es cierto.

(B_1) $p \in [0, 1] \Rightarrow 1 - p \geqslant 0$ y también $p \geqslant 0 \Rightarrow f(x) \geqslant 0$

(B_2) $\sum_{x \in \{0,1\}} f(x) = (1 - p) + p = 1$

$\therefore$ Teorema 4.3 es cierto. □

Teorema 4.4 (Bernoulli)**.** Sea $X \sim \text{Bern}(p) \Rightarrow$

(B_1) $m_X(t) = (1-p) + pe^t$ FGM

(B_2) $\mathbb{E}[X] = p$ Esperanza

(B_3) $\text{Var}[X] = p(1-p)$ Varianza

Proof. Sea $X \sim \text{Bern}(p)$, es decir, $P(X=1) = p$ y $P(X=0) = 1-p$.

(B_1) Tomamos la FGM:

$$m_X(t) = \mathbb{E}[e^{tX}] = e^{t\cdot 0}(1-p) + e^{t\cdot 1}p = (1-p) + pe^t$$

(B_2) Para la esperanza, notemos que:

$$\mathbb{E}[X] = 0(1-p) + 1\cdot p = p$$

(B_3) Finalmente, para la varianza:

$$\mathbb{E}[X^2] = 0^2(1-p) + 1^2\cdot p = p$$

$$\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = p - p^2 = p(1-p)$$

$\therefore$ Teorema 4.4 es cierto. $\square$

Definición 4.3 (Distribución Binomial)**.** Decimos que una VA X tiene distribución binomial con parámetros $n \in \mathbb{N}$ y $p \in [0,1]$

$$X \sim \text{Bin}(n,p)$$

si y solo si su PMF es la siguiente

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$$

donde n es el número de ensayos y p la probabilidad de éxito en cada ensayo.

Teorema 4.5 (PMF de Binomial)**.** Sea $X \sim \text{Bin}(n,p)$ su PMF $f(k)$ cumple

(B_1) $\forall\, k \in \mathbb{N} \Rightarrow f(k) \geqslant 0$

(B_2) $\sum_{k=0}^{n} f(k) = 1$

Proof. Veamos que Teorema 4.5 es cierto.

(B_1) $\binom{n}{k} p^k (1-p)^{n-k} \geqslant 0 \Rightarrow f(k) \geqslant 0$

(B_2) $\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p+1-p)^n = 1^n = 1$

Esto por el teorema del binomio, que postula

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} y^k \cdot x^{n-k}$$

$\therefore$ Teorema 4.5 es cierto. $\square$

Observación. Notemos que si $n = 1 \Rightarrow X \sim \text{Bern}$

Teorema 4.6. Sea $X \sim \text{Bin}(n,p)$ y definamos a $q = 1 - p \Rightarrow n - X \sim \text{Bin}(n,q)$

Proof. Veamos que $n - X$ tiene la PDF de la binomial. Sea $Y = n - x$

$$P(Y = k) = P(X = n - k) = \binom{n}{n-k} p^{n-k}(q)^k = \binom{n}{k} q^k (p)^{n-k}$$

□

Teorema 4.7 (Binomial). Sea $X \sim \text{Bin}(n,p) \Rightarrow$

(B_1) $m_X(t) = (pe^t + 1 - p)^n$ FGM

(B_2) $\mathbb{E}[X] = np$ Esperanza

(B_3) $\text{Var}[X] = np(1-p)$ Varianza

Proof. Sea $X \sim \text{Bin}(n,p)$, es decir, $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ para $k = 0, \ldots, n$.

(B_1) Tomamos la FGM:

$$m_X(t) = \mathbb{E}[e^{tX}] = \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k (1-p)^{n-k} = (pe^t + 1 - p)^n$$

(B_2) Por Teorema 3.4 derivamos $m_X(t)$ y evaluamos en $t = 0$

$$m'_X(t) = n(pe^t + 1 - p)^{n-1} pe^t$$

$$\Rightarrow \mathbb{E}[X] = m'_X(0) = n(p + 1 - p)^{n-1} p = np$$

(B_3) Derivamos dos veces $m_X(t)$ y evaluamos en $t = 0 \Rightarrow$

$$m''_X(t) = n(n-1)(pe^t + 1 - p)^{n-2}(pe^t)^2 + n(pe^t + 1 - p)^{n-1} pe^t$$

$$\Rightarrow m''_X(0) = n(n-1)p^2 + np = np((n-1)p + 1)$$

$$\Rightarrow \text{Var}[X] = m''_X(0) - (m'_X(0))^2 = np((n-1)p + 1) - (np)^2 = np(1-p)$$

$\therefore$ Teorema 4.7 es cierto. □

Observación. En esta distribución X contabiliza el número de éxitos u ocurrencias que suceden en n ensayos de Bernoulli independientes.

Definición 4.4 (Distribución Poisson). Decimos que una VA X tiene distribución Poisson con parámetro $\lambda > 0$

$$X \sim \text{Pois}(\lambda)$$

si y solo si su función de masa de probabilidad es la siguiente

$$P(X = k) = \text{Pois}(k; \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k \in \mathbb{N}_0$$

donde $\lambda \in \mathbb{R}^+$

Observación. Esta distribución es comúnmente utilizada para modelar eventos de extraña ocurrencia. También se usa en el conteo de eventos en un periodo de tiempo determinado.

Teorema 4.8 (PMF de Poisson). Sea $X \sim \text{Poisson}(\lambda)$ su PMF $f(k)$ cumple

(P_1) $\forall\, k \in \mathbb{N}_0 \Rightarrow f(k) \geqslant 0$

(P_2) $\sum_{k=0}^{\infty} f(k) = 1$

Proof. Veamos que Teorema 4.8 es cierto.

(P_1) $f(k) = \frac{\lambda^k e^{-\lambda}}{k!} \geqslant 0$ para todo $k \in \mathbb{N}_0$ pues $\lambda > 0$ y factoriales son positivos.

(P_2)

$$\sum_{k=0}^{\infty} \frac{\lambda^k e^{-\lambda}}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1.$$

$\therefore$ Teorema 4.8 es cierto. $\square$

Teorema 4.9 (Poisson). Sea $X \sim \text{Poisson}(\lambda) \Rightarrow$

(P_1) $m_X(t) = \exp\big(\lambda(e^t - 1)\big)$ FGM

(P_2) $\mathbb{E}[X] = \lambda$ Esperanza

(P_3) $\text{Var}[X] = \lambda$ Varianza

Proof. Sea $X \sim \text{Poisson}(\lambda)$, es decir, $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$ para $k = 0, 1, 2, \ldots$

(P_1) Tomamos la FGM:

$$m_X(t) = \mathbb{E}[e^{tX}] = \sum_{k=0}^{\infty} \frac{(\lambda e^t)^k e^{-\lambda}}{k!} = e^{-\lambda} \cdot e^{\lambda e^t} = \exp\big(\lambda(e^t - 1)\big)$$

(P_2) Usamos la función cumulante $\Psi_X(t) = \ln m_X(t) = \lambda(e^t - 1)$ y (Ψ_2) del Teorema 3.5

$$\Psi'_X(t) = \lambda e^t \mid_{t=0} \Rightarrow \Psi'_X(0) = \lambda = \mathbb{E}[X]$$

(P_3) Derivando de nuevo y evaluando en $t = 0$:

$$\Psi''_X(t) = \lambda e^t \mid_{t=0} \Rightarrow \text{Var}[X] = \Psi''_X(0) = \lambda = \text{Var}[X]$$

$\therefore$ Teorema 4.9 es cierto. $\square$

Observación. La moda de la Poisson es $\lfloor \lambda \rfloor$ si $\lambda \notin \mathbb{Z}^+$, y tanto λ y $\lambda - 1$ si $\lambda \in \mathbb{Z}^+$

Definición 4.5 (Distribución Geométrica). Decimos que una VA Y tiene distribución geométrica con parámetro $p \in (0, 1]$

$$Y \sim \text{Geo}(p)$$

si y solo si su PMF es la siguiente

$$P(Y = k) = \text{Gem}(k; p) = (1 - p)^k p, \quad k \in \mathbb{N}_0$$

donde $p \in (0, 1]$ representa la probabilidad de éxito en un solo ensayo, y $k \in \mathbb{N}_0$ contabiliza el número de fracasos que se presentan antes del primer éxito en ensayos $\sim \text{Bern}(p)$

Teorema 4.10 (PMF de Geométrica). Sea $Y \sim \text{Geom}(p)$ su PMF $f(k)$ cumple

(G_1) $\forall\, k \in \mathbb{N}_0 \Rightarrow f(k) \geqslant 0$

(G_2) $\sum_{k=0}^{\infty} f(k) = 1$

Proof. Veamos que Teorema 4.10 es cierto.

(G_1) $f(k) = (1-p)^k p \geqslant 0$ para todo $k \in \mathbb{N}_0$ pues $p \in (0,1]$

(G_2)

$$\sum_{k=0}^{\infty}(1-p)^k p = p\sum_{k=0}^{\infty}(1-p)^k = p \cdot \frac{1}{1-(1-p)} = p \cdot \frac{1}{p} = 1.$$

Esto por la fórmula de la suma de una serie geométrica.

$\therefore$ Teorema 4.10 es cierto. □

Teorema 4.11 (Geométrica). Sea $Y \sim \text{Geom}(p) \Rightarrow$

(G_1) $m_Y(t) = \frac{p}{1-(1-p)e^t}$, para $t < -\ln(1-p)$ FGM

(G_2) $\mathbb{E}[Y] = \frac{1-p}{p}$ Esperanza

(G_3) $\text{Var}[Y] = \frac{1-p}{p^2}$ Varianza

Proof. Sea $Y \sim \text{Geom}(p)$, es decir, $P(Y=k) = (1-p)^k p$ para $k = 0, 1, 2, \ldots$

(G_1) Tomamos la FGM:

$$m_Y(t) = \mathbb{E}[e^{tY}] = \sum_{k=0}^{\infty} e^{tk}(1-p)^k p = p\sum_{k=0}^{\infty}\left[(1-p)e^t\right]^k$$

$$= \frac{p}{1-(1-p)e^t}, \quad \text{para } t < -\ln(1-p)$$

(G_2) Usamos la función cumulante $\Psi_Y(t) = \ln m_Y(t) = \ln p - \ln\big(1-(1-p)e^t\big)$ y (Ψ_2) del Teorema 3.5

$$\Psi_Y'(t) = \frac{(1-p)e^t}{1-(1-p)e^t}\,|_{t=0} = \frac{1-p}{p} = \mathbb{E}[Y]$$

(G_3) Derivando de nuevo y evaluando en $t=0$:

$$\Psi_Y''(t) = \frac{(1-p)e^t[1-(1-p)e^t] + (1-p)^2 e^{2t}}{[1-(1-p)e^t]^2} \Rightarrow \Psi_Y''(0) = \frac{1-p}{p^2} = \text{Var}[Y]$$

$\therefore$ Teorema 4.11 es cierto. □

Definición 4.6 (Binomial Negativa). Una VA X es binomial negativa con parámetros $r \in \mathbb{N}$ y $p \in (0,1]$

$$X \sim \text{NegBin}(r,p)$$

si y solo si su PMF es la siguiente

$$P(X=k) = \text{NegBin}(k; r, p) = \binom{k+r-1}{k}(1-p)^k p^r, \quad k \in \mathbb{N}_0$$

donde X cuenta el número de fracasos antes del r-ésimo éxito en ensayos i.i.d $\sim \text{Bern}(p)$.

Teorema 4.12 (PMF de Binomial Negativa). Sea $X \sim \text{NegBin}(r, p)$ su PMF $f(k)$ cumple

(B_1) $\forall\, k \in \mathbb{N}_0 \Rightarrow f(k) \geqslant 0$

(B_2) $\sum_{k=0}^{\infty} f(k) = 1$

Proof. Veamos que Teorema 4.12 es cierto.

(B_1) $f(k) = \binom{k+r-1}{k}(1-p)^k p^r \geqslant 0$ para todo $k \in \mathbb{N}_0$, pues $r \in \mathbb{N}$ y $p \in (0, 1]$

(B_2) Para $0 < p < 1$

$$\sum_{k=0}^{\infty} \binom{k+r-1}{k}(1-p)^k p^r = p^r \sum_{k=0}^{\infty} \binom{k+r-1}{k}(1-p)^k$$

$$= p^r \cdot \left(\frac{1}{1-(1-p)}\right)^r = p^r \cdot \left(\frac{1}{p}\right)^r = 1$$

Esto por la fórmula del desarrollo de la serie binomial negativa.

$\therefore$ Teorema 4.12 es cierto. □

Teorema 4.13 (Binomial Negativa). Sea $X \sim \text{NegBin}(r, p) \Rightarrow$

(B_1) $m_X(t) = \left(\frac{p}{1-(1-p)e^t}\right)^r$, para $t < -\ln(1-p)$ FGM

(B_2) $\mathbb{E}[X] = \frac{r(1-p)}{p}$ Esperanza

(B_3) $\text{Var}[X] = \frac{r(1-p)}{p^2}$ Varianza

Proof. Sea $X \sim \text{NegBin}(r, p)$, es decir, $P(X = k) = \binom{k+r-1}{k}(1-p)^k p^r$ para $k = 0, 1, 2, \ldots$

(B_1) Tomamos la FGM:

$$m_X(t) = \mathbb{E}[e^{tX}] = \sum_{k=0}^{\infty} e^{tk} \binom{k+r-1}{k}(1-p)^k p^r = p^r \sum_{k=0}^{\infty} \binom{k+r-1}{k} \left[(1-p)e^t\right]^k$$

Y notemos que, para $t < -\ln(1-p)$

$$= \left(\frac{p}{1-(1-p)e^t}\right)^r,$$

(B_2) Usamos la función cumulante $\Psi_X(t) = \ln m_X(t) = r \ln p - r \ln\big(1-(1-p)e^t\big)$ y (Ψ_2) del Teorema 3.5

$$\Psi'_X(t) = \frac{r(1-p)e^t}{1-(1-p)e^t} \,|_{t=0}= \frac{r(1-p)}{p} = \mathbb{E}[X]$$

(B_3) Derivando de nuevo y evaluando en $t = 0$:

$$\Psi''_X(t) = \frac{r(1-p)e^t[1-(1-p)e^t] + r(1-p)^2 e^{2t}}{[1-(1-p)e^t]^2} \Rightarrow \Psi''_X(0) = \frac{r(1-p)}{p^2} = \text{Var}[X]$$

$\therefore$ Teorema 4.13 es cierto. □

Definición 4.7 (Distribución Hipergeométrica). Una VA X es hipergeométrica con parámetros $N, K, n \in \mathbb{N}$ tales que $K \leq N$ y $n \leq N$

$$X \sim \text{HGeom}(N, K, n)$$

si y solo si su PMF es la siguiente

$$P(X = k) = \text{HGeom}(k; N, K, n) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}, \quad \text{máx}(0, n - N + K) \leq k \leq \text{mín}(n, K)$$

donde X cuenta el número de éxitos en una muestra aleatoria de tamaño n tomada sin reemplazo de una población de tamaño N con K elementos exitosos.

Teorema 4.14 (PMF de Hipergeométrica). Sea $X \sim \text{HGeom}(N, K, n)$ su PMF $f(k)$ cumple

(H_1) $\forall\, k \in [\text{máx}(0, n - N + K), \text{mín}(n, K)] \Rightarrow f(k) \geqslant 0$

(H_2) $\sum_{k=\text{máx}(0,n-N+K)}^{\text{mín}(n,K)} f(k) = 1$

Proof. Veamos que el Teorema 4.14 es cierto.

(H_1) $f(k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}} \geqslant 0$ pues todos los binomiales son no negativos.

(H_2) Notemos que, por el Teorema 1.3

$$\sum_{k=\text{máx}(0,n-N+K)}^{\text{mín}(n,K)} \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{k} \binom{K}{k}\binom{N-K}{n-k} = 1$$

$\therefore$ el Teorema 4.14 es cierto. $\square$

Teorema 4.15 (Hipergeométrica). Sea $X \sim \text{HGeom}(N, K, n) \Rightarrow$

(H_1) $\mathbb{E}[X] = \frac{nK}{M}$ Esperanza

(H_2) $\text{Var}[X] = n \cdot \frac{K}{M} \cdot \left(1 - \frac{K}{M}\right) \cdot \frac{M-n}{M-1}$ Varianza

Proof. Sea $X \sim \text{HGeom}(N, K, n)$

(H_1) Sabemos que

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n}{k} \frac{(n-1)!}{(k-1)!(n-1-(k-1))!} = \frac{n}{k}\binom{n-1}{k-1}$$

Y como $x = 0$ no aporta nada, de la Definición 3.1

$$\mathbb{E}[X] = \sum_{x=1}^{n} \frac{x\binom{K}{x}\binom{M-K}{n-x}}{\binom{M}{n}} = \frac{nK}{M} \sum_{x=1}^{n} \frac{\binom{K-1}{x-1}\binom{M-1-(K-1)}{n-1-(x-1)}}{\binom{M-1}{n-1}}$$

Y si $l = x - 1$, del Teorema 4.14 se tiene a la PMF

$$\Rightarrow \mathbb{E}[X] = \frac{nK}{M}$$

(H_2) Para la varianza, usamos la Definición 3.2 y que

$$\mathbb{E}[X^2] = \mathbb{E}[X(X-1)] + \mathbb{E}[X]$$

Primero calculamos

$$\mathbb{E}[X(X-1)] = \sum_{x=0}^{n} x(x-1)\frac{\binom{K}{x}\binom{M-K}{n-x}}{\binom{M}{n}}$$

$$\Rightarrow \mathbb{E}[X(X-1)] = \frac{n(n-1)K(K-1)}{M(M-1)}.$$

$$\mathbb{E}[X^2] = \mathbb{E}[X(X-1)] + \mathbb{E}[X] = \frac{n(n-1)K(K-1)}{M(M-1)} + \frac{nK}{M}.$$

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = n\frac{K}{M}\left(1-\frac{K}{M}\right)\frac{M-n}{M-1}.$$

$\therefore$ Teorema 4.15 es cierto. □

4.2. Continuas

Definición 4.8 (Distribución Uniforme Continua). Decimos que una VA X tiene distribución uniforme continua en el intervalo $[a, b]$ con $a < b$

$$X \sim \mathcal{U}(a, b)$$

si y solo si su PDF es la siguiente

$$f(x) = \mathcal{U}(x; a, b) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases}$$

donde $a, b \in \mathbb{R}$ y $a < b$

Observación. Esta distribución modela espacios equiprobables que toman valores en un intervalo de longitud finita de $\mathbb{R}$, digamos en $\mathrm{I} = (a, b)$ con $-\infty < a < b < \infty$

Teorema 4.16 (PDF de Uniforme Continua). Sea $X \sim \mathcal{U}(a, b)$ con $a < b$, su PDF $f(x)$ cumple

$(\mathcal{U}_1)$ $\forall\, x \in \mathbb{R} \Rightarrow f(x) \geqslant 0$

$(\mathcal{U}_2)$ $\int_{-\infty}^{\infty} f(x)\, dx = \int_a^b \frac{1}{b-a}\, dx = 1$

Proof. Veamos que el Teorema 4.16 es cierto.

$(\mathcal{U}_1)$ $f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases} \geqslant 0$ pues es constante no negativa o cero.

$(\mathcal{U}_2)$ Notemos que

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_a^b \frac{1}{b-a}\, dx = \frac{b-a}{b-a} = 1$$

$\therefore$ el Teorema 4.16 es cierto. □

Definición 4.9 (CDF Uniforme Continua). Sea $X \sim \mathcal{U}(a,b)$ con $a < b$. La CDF de X es

$$F(x) = P(X \leq x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}$$

Teorema 4.17 (Uniforme Continua). Sea $X \sim \mathcal{U}(a,b) \Rightarrow$

($\mathcal{U}_1$) $m_X(t) = \frac{e^{bt}-e^{at}}{t(b-a)} \quad t \neq 0$ FGM

($\mathcal{U}_2$) $\mathbb{E}[X^r] = \frac{b^{r+1}-a^{r+1}}{(r+1)(b-a)}$ r-ésimo momento

($\mathcal{U}_3$) $\mathbb{E}[X] = \frac{a+b}{2}$ Esperanza

($\mathcal{U}_4$) $\mathrm{Var}[X] = \frac{(b-a)^2}{12}$ Varianza

Proof. Sea $X \sim \mathcal{U}(a,b)$

($\mathcal{U}_1$) Sea $t \neq 0$. Usamos la definición de función generadora de momentos:

$$m_X(t) = \mathbb{E}[e^{tX}] = \int_a^b e^{tx} \cdot \frac{1}{b-a}\, dx = \frac{1}{b-a}\int_a^b e^{tx}\, dx$$

$$= \frac{1}{b-a}\left[\frac{e^{tx}}{t}\right]_a^b = \frac{1}{t(b-a)}(e^{bt} - e^{at})$$

($\mathcal{U}_2$) Sea $r \in \mathbb{N}$, usamos la definición de esperanza para funciones continuas:

$$\mathbb{E}[X^r] = \int_a^b x^r \cdot \frac{1}{b-a}\, dx = \frac{1}{b-a}\int_a^b x^r\, dx = \frac{1}{b-a}\left[\frac{x^{r+1}}{r+1}\right]_a^b$$

$$= \frac{b^{r+1} - a^{r+1}}{(r+1)(b-a)}$$

($\mathcal{U}_3$) Para la esperanza, tomamos $r = 1$ en el resultado anterior:

$$\mathbb{E}[X] = \frac{b^2 - a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}$$

($\mathcal{U}_4$) Para la varianza, usamos que

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

Tomamos $r = 2$:

$$\mathbb{E}[X^2] = \frac{b^3 - a^3}{3(b-a)} = \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} = \frac{b^2+ab+a^2}{3}$$

$$(\mathbb{E}[X])^2 = \left(\frac{a+b}{2}\right)^2 = \frac{a^2+2ab+b^2}{4}$$

$$\Rightarrow \mathrm{Var}[X] = \frac{b^2+ab+a^2}{3} - \frac{a^2+2ab+b^2}{4} = \frac{(b-a)^2}{12}$$

$\therefore$ Teorema 4.17 es cierto. $\square$

Definición 4.10 (Distribución Exponencial). Decimos que una VA X tiene distribución exponencial con parámetro $\lambda > 0$

$$X \sim \exp(\lambda)$$

si y solo si su PDF es la siguiente

$$f(x) = \exp(x; \lambda) = \begin{cases} \lambda e^{-\lambda x}, & x \geqslant 0 \\ 0, & x < 0 \end{cases}$$

donde $\lambda \in \mathbb{R}^+$

Teorema 4.18 (PDF de Exponencial). Sea $X \sim \exp(\lambda)$ con $\lambda > 0$, su PDF $f(x)$ cumple

(e_1) $\forall\, x \in \mathbb{R} \Rightarrow f(x) \geqslant 0$

(e_2) $\int_{-\infty}^{\infty} f(x)\, dx = \int_0^{\infty} \lambda e^{-\lambda x}\, dx = 1$

Proof. Veamos que el Teorema 4.18 es cierto.

(e_1) $f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \geqslant 0$ pues $\lambda > 0$ y la exponencial es positiva.

(e_2) Notemos que

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_0^{\infty} \lambda e^{-\lambda x}\, dx = \lambda \int_0^{\infty} e^{-\lambda x}\, dx = \lambda \left[-\frac{1}{\lambda} e^{-\lambda x} \right] = \lambda \left[-0 + \frac{1}{\lambda} \right] = 1$$

$\therefore$ el Teorema 4.18 es cierto. $\square$

Definición 4.11 (CDF Exponencial). Sea $X \sim \exp(\lambda)$ con $\lambda > 0$. La CDF de X es

$$F(x) = P(X \leq x) = \begin{cases} 1 - e^{-\lambda x}, & x \geqslant 0 \\ 0, & x < 0 \end{cases}$$

Teorema 4.19 (Exponencial). Sea $X \sim \exp(\lambda) \Rightarrow$

(e_1) $m_X(t) = \frac{\lambda}{\lambda - t}, \quad t < \lambda$ FGM

(e_2) $\mathbb{E}[X] = \frac{1}{\lambda}$ Esperanza

(e_3) $\mathrm{Var}[X] = \frac{1}{\lambda^2}$ Varianza

Proof. Sea $X \sim \exp(\lambda)$, es decir, $f(x) = \lambda e^{-\lambda x}$ para $x \geq 0$.

(e_1) Tomamos la FGM, y para $t < \lambda$ s.t.q.

$$m_X(t) = \mathbb{E}[e^{tX}] = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x}\, dx = \lambda \int_0^{\infty} e^{-(\lambda - t)x}\, dx = \frac{\lambda}{\lambda - t}$$

(e_2) Usamos la función $\Psi_X(t) = \ln m_X(t) = \ln \lambda - \ln(\lambda - t)$ y (Ψ_2) del Teorema 3.5

$$\Psi'_X(t) = \frac{1}{\lambda - t} \,|_{t=0} = \frac{1}{\lambda} = \mathbb{E}[X]$$

(e_3) Derivando de nuevo:

$$\Psi_X''(t) = \frac{1}{(\lambda - t)^2} \Rightarrow \Psi_X''(0) = \frac{1}{\lambda^2} = \mathrm{Var}[X]$$

$\therefore$ Teorema 4.19 es cierto. $\square$

Definición 4.12 (Función Gamma). Sea $\alpha \in \mathbb{R}^+$. Se define a la función Gamma como

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$$

Ejemplo 4.1. $\Gamma(1) = 1$

Proof.

$$\Gamma(1) = \int_0^\infty t^0 e^{-t} dt = \int_0^\infty e^{-t} dt = \int_{-\infty}^0 e^t dt = 1$$

$\square$

Teorema 4.20. $\forall\, \alpha \in \mathbb{R}^+ \Rightarrow \Gamma(\alpha+1) = \alpha \cdot \Gamma(\alpha)$

Proof. Nótese que

$$\Gamma(\alpha) + 1 = \int_0^\infty t^\alpha e^{-t} dt = t^\alpha e^{-t} \mid_0^\infty + \alpha \int_0^\infty t^{\alpha-1} e^{-t} dt = 0 + \alpha \cdot \Gamma(\alpha)$$

$\therefore \Gamma(\alpha+1) = \alpha \cdot \Gamma(\alpha)$ $\square$

Corolario. $\forall\, \alpha \in \mathbb{Z}^+ \Rightarrow \Gamma(\alpha+1) = \alpha!$

Proof. Suponemos que se cumple para $\alpha = k$ i.e. $\Gamma(k+1) = k!$

$$\Gamma(k+1+1) = \int_0^\infty t^{k+1} e^{-t} dt$$

$$= t^{k+1} e^{-t} \mid_0^\infty + (k+1) \int_0^\infty t^k e^{-t} dt = (k+1)\Gamma(k+1) = (k+1)k! = (k+1)!$$

$\therefore \Gamma(\alpha+1) = \alpha!$ $\square$

Teorema 4.21. $\forall\, \alpha \in \mathbb{R}^+ \Rightarrow \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

Proof. Sea $c = \Gamma(\frac{1}{2}) = \int_0^\infty t^{\frac{1}{2}-1} e^{-t} dt = \int_0^\infty \frac{1}{\sqrt{t}} e^{-t} dt$. Ahora tomamos $u^2 = t$ y $2udu = dt$

$$\Rightarrow \int_0^\infty \frac{1}{u} e^{-u^2} 2udu \Rightarrow c = 2 \int_0^\infty u \frac{e^{-u^2}}{u} du = 2 \int_0^\infty e^{-u^2} du$$

Notemos que

$$c \cdot c = 4 \int_0^\infty e^{-u^2} du \int_0^\infty e^{-v^2} dv$$

Por el Teorema de Fubbini

$$4 \int_0^\infty \int_0^\infty e^{-(u^2+v^2)} dudv$$

Sea $u = r\cos\theta$ y $v = r\sin\theta$

$$[\mathbb{J}] = \begin{bmatrix} \frac{\partial u}{\partial r} & \frac{\partial u}{\partial \theta} \\ \frac{\partial v}{\partial r} & \frac{\partial v}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} = r\cos^2\theta + r\sin^2\theta = r$$

$$\Rightarrow 4\int_0^\infty \int_0^{\frac{\pi}{2}} (e^{-r^2\sin^2\theta - r^2\cos^2\theta}) r d\theta dr = 4\int_0^\infty \int_0^{\frac{\pi}{2}} e^{-r^2} r d\theta dr$$

$$= 4\int_0^\infty re^{-r^2}\left(\int_0^{\frac{\pi}{2}} d\theta\right) dr = 4\frac{\pi}{2}\int_0^\infty re^{-r^2} dr$$

Sea $w = r^2$ y $dw = 2rdr$

$$= \frac{2\pi}{2}\int_0^\infty e^{-w} dw = \pi\int_0^\infty e^{-w} dw = \pi \Rightarrow c^2 = \pi \Rightarrow c = \sqrt{\pi}$$

$\therefore \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ $\square$

Definición 4.13 (Distribución Gamma)**.** Decimos que una VA X tiene distribución Gamma con parámetros $r > 0$ y $\lambda > 0$

$$X \sim \text{Gamma}(r, \lambda)$$

si y solo si su PDF es la siguiente

$$f(x) = \text{Gamma}(x; r, \lambda) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x} \qquad x > 0$$

donde $r, \lambda \in \mathbb{R}^+$ y $\Gamma(r)$ es la Definición 4.12.

Teorema 4.22 (PDF de Gamma)**.** Sea $X \sim \text{Gamma}(r, \lambda)$ con $r, \lambda \in \mathbb{R}^+$, su PDF $f(x)$ cumple

(Γ_1) $\forall\, x \in \mathbb{R} \Rightarrow f(x) \geqslant 0$

(Γ_2) $\int_{-\infty}^{\infty} f(x)\, dx = \int_0^\infty \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}\, dx = 1$

Proof. Veamos que el Teorema 4.22 es cierto.

(Γ_1) $f(x) = \begin{cases} \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}, & x > 0 \\ 0, & x \leqslant 0 \end{cases} \geqslant 0$

Esto como todos los factores son no negativos para $x > 0$.

(Γ_2) Notemos que

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_0^\infty \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}\, dx = \frac{\lambda^r}{\Gamma(r)} \cdot \int_0^\infty x^{r-1} e^{-\lambda x}\, dx.$$

Sea $u = \lambda x$ y $dx = \frac{1}{\lambda} du$

$$= \frac{\lambda^r}{\Gamma(r)} \cdot \int_0^\infty \left(\frac{u}{\lambda}\right)^{r-1} e^{-u} \cdot \frac{1}{\lambda}\, du = \frac{\lambda^r}{\Gamma(r)} \cdot \frac{1}{\lambda^r} \int_0^\infty u^{r-1} e^{-u}\, du = 1.$$

$\therefore$ el Teorema 4.22 es cierto. $\square$

Observación. Esta distribución tiene usos importantes en la econometría.

Definición 4.14 (CDF Gamma). Sea $X \sim \text{Gamma}(r, \lambda)$ con $r > 0$ y $\lambda > 0$. La CDF de X es

$$F(x) = P(X \leq x) = \begin{cases} \displaystyle\int_0^x \frac{\lambda^r}{\Gamma(r)} t^{r-1} e^{-\lambda t}\, dt, & x \geqslant 0 \\ 0, & x < 0 \end{cases}$$

Teorema 4.23 (Gamma). Sea $X \sim \text{Gamma}(r, \lambda) \Rightarrow$

(Γ_1) $m_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^r \quad t < \lambda$ FGM

(Γ_2) $\mathbb{E}[X^n] = \frac{\Gamma(r+n)}{\lambda^n \Gamma(r)}$ n-ésimo momento

(Γ_3) $\mathbb{E}[X] = \frac{r}{\lambda}$ Esperanza

(Γ_4) $\text{Var}[X] = \frac{r}{\lambda^2}$ Varianza

Proof. Sea $X \sim \text{Gamma}(r, \lambda)$

(Γ_1) Sea $t < \lambda$. Por definición de función generadora de momentos:

$$m_X(t) = \mathbb{E}[e^{tX}] = \int_0^\infty e^{tx} \cdot \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}\, dx = \frac{\lambda^r}{\Gamma(r)} \int_0^\infty x^{r-1} e^{-(\lambda - t)x}\, dx$$

$$= \frac{\lambda^r}{\Gamma(r)} \cdot \frac{\Gamma(r)}{(\lambda - t)^r} = \left(\frac{\lambda}{\lambda - t}\right)^r$$

(Γ_2) Para $n \in \mathbb{N}$ usamos la definición de momento:

$$\mathbb{E}[X^n] = \int_0^\infty x^n \cdot \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}\, dx = \frac{\lambda^r}{\Gamma(r)} \int_0^\infty x^{n+r-1} e^{-\lambda x}\, dx$$

Hacemos el cambio $u = \lambda x$, $du = \lambda dx$, $x = u/\lambda$:

$$= \frac{\lambda^r}{\Gamma(r)} \cdot \int_0^\infty \left(\frac{u}{\lambda}\right)^{n+r-1} e^{-u} \cdot \frac{1}{\lambda}\, du = \frac{\lambda^r}{\Gamma(r)} \cdot \frac{1}{\lambda^{n+r}} \int_0^\infty u^{n+r-1} e^{-u}\, du$$

$$= \frac{\Gamma(n+r)}{\lambda^n \Gamma(r)}$$

(Γ_3) Para la esperanza, tomamos $n = 1$:

$$\mathbb{E}[X] = \frac{\Gamma(r+1)}{\lambda \Gamma(r)} = \frac{r\Gamma(r)}{\lambda \Gamma(r)} = \frac{r}{\lambda}$$

(Γ_4) Para la varianza usamos que $\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$ y tomamos $n = 2$:

$$\mathbb{E}[X^2] = \frac{\Gamma(r+2)}{\lambda^2 \Gamma(r)} = \frac{(r+1) r \Gamma(r)}{\lambda^2 \Gamma(r)} = \frac{r(r+1)}{\lambda^2}$$

$$(\mathbb{E}[X])^2 = \left(\frac{r}{\lambda}\right)^2 = \frac{r^2}{\lambda^2}$$

$$\Rightarrow \text{Var}[X] = \frac{r(r+1)}{\lambda^2} - \frac{r^2}{\lambda^2} = \frac{r}{\lambda^2}$$

$\therefore$ Teorema 4.23 es cierto. $\square$

Observación. Notemos que si $r = 1 \Rightarrow X \sim \text{Gamma}(1, \lambda) = X \sim \exp(\lambda)$

Ejemplo 4.2 (Erlang)**.** Si $X \sim \text{Gamma}(r, \lambda)$ con $r \in \mathbb{Z}^+ \Rightarrow$

$$X \sim \text{Erlang}(r, \lambda)$$

si y solo si su PDF es la siguiente

$$f(x) = \text{Erlang}(x; r, \lambda) = \frac{\lambda^r}{(r-1)!} x^{r-1} e^{-\lambda x}, \quad x \geqslant 0$$

donde se usa que $\Gamma(r) = (r-1)!$ cuando $r \in \mathbb{Z}^+$.

Ejemplo 4.3 (Distribución Ji-Cuadrada)**.** Si $X \sim \text{Gamma}(r, \lambda)$ con $r = \frac{k}{2}$ y $\lambda = \frac{1}{2}$, para algún $k \in \mathbb{Z}^+ \Rightarrow X$ tiene una distribución Ji-Cuadrada con k grados de libertad

$$X \sim \chi^2(k)$$

si y solo si su PDF es la siguiente

$$f(x) = \chi^2(x; k) = \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} e^{-x/2}, \quad x \geqslant 0$$

donde $k \in \mathbb{Z}^+$ representa el número de grados de libertad.

Proof. Sea $r = \frac{k}{2}$ y $\lambda = \frac{1}{2}$

Heredamos de Teorema 4.22 las pruebas de la PDF y de Teorema 4.23 podemos sacar

(χ_1) $m_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^r \quad t < \lambda = \frac{1}{2} = \left(\frac{\frac{1}{2}}{\frac{1}{2} - t}\right)^{\frac{k}{2}} = \frac{1}{1-2t}^{\frac{k}{2}} = (1-2t)^{-\frac{k}{2}}$

(χ_2) $\mathbb{E}[X] = \frac{r}{\lambda} = \frac{\frac{k}{2}}{\frac{1}{2}} = k$

(χ_3) $\text{Var}[X] = \frac{r}{\lambda^2} = \frac{\frac{k}{2}}{\frac{1}{2}^2} = 2k$

$\therefore$ la Ji-Cuadrada es una distribución. $\square$

Definición 4.15 (Función Beta)**.** Sean $\alpha, \beta \in \mathbb{R}^+$. Se define a la función Beta como

$$\text{B}(\alpha, \beta) = \int_0^1 u^{\alpha-1}(1-u)^{\beta-1} du$$

Teorema 4.24. La relación entre la Función Gama y Beta es la siguiente

$$\text{B}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

Proof. Consideramos el cambio de variable en la integral doble de dos funciones Gamma

$$\Gamma(\alpha)\Gamma(\beta) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx \cdot \int_0^\infty y^{\beta-1} e^{-y}\, dy = \iint_{\mathbb{R}^2_+} x^{\alpha-1} y^{\beta-1} e^{-(x+y)}\, dx\, dy$$

Sea $u = x + y$ y $t = \frac{x}{x+y} \Rightarrow x = ut$ y $y = u(1-t)$

El jacobiano de esta transformación es $[J] = u$, entonces

$$\Gamma(\alpha)\Gamma(\beta) = \iint_{u>0,\, t\in(0,1)} (ut)^{\alpha-1}(u(1-t))^{\beta-1}e^{-u}u\,dt\,du$$

$$= \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\,dt \cdot \int_0^\infty u^{\alpha+\beta-1}e^{-u}\,du = \mathrm{B}(\alpha,\beta)\cdot\Gamma(\alpha+\beta)$$

Luego,

$$\Rightarrow \mathrm{B}(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$\therefore$ el teorema es cierto. □

Definición 4.16 (Distribución Beta). Decimos que una VA X tiene distribución Beta con parámetros $\alpha > 0$ y $\beta > 0$

$$X \sim \mathrm{Beta}(\alpha,\beta)$$

si y solo si su PDF es la siguiente

$$f(x) = \mathrm{Beta}(x;\alpha,\beta) = \frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}, \quad x\in(0,1)$$

donde $\mathrm{B}(\alpha,\beta)$ es la Definición 4.15 y $\alpha,\beta\in\mathbb{R}^+$

Teorema 4.25 (PDF de Beta). Sea $X\sim\mathrm{Beta}(\alpha,\beta)$ con $\alpha,\beta\in\mathbb{R}^+$, su PDF $f(x)$ cumple

(B_1) $\forall\, x\in\mathbb{R} \Rightarrow f(x)\geqslant 0$

(B_2) $\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 \frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}\,dx = 1$

Proof. Veamos que el Teorema 4.25 es cierto.

(B_1) $f(x) = \begin{cases} \frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}, & x\in(0,1) \\ 0, & x\notin(0,1)\end{cases} \geqslant 0$

Esto ya que todos los factores son no negativos para $x\in(0,1)$ y $\alpha,\beta>0$.

(B_2) Notemos que

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_0^1 \frac{1}{\mathrm{B}(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}\,dx = \frac{1}{\mathrm{B}(\alpha,\beta)}\cdot\mathrm{B}(\alpha,\beta) = 1.$$

$\therefore$ el Teorema 4.25 es cierto. □

Definición 4.17 (CDF Beta). Sea $X\sim\mathrm{Beta}(\alpha,\beta)$ con $\alpha>0$ y $\beta>0$. La CDF de X es

$$F(x;\alpha,\beta) = \frac{\mathrm{B}(x;\alpha,\beta)}{\mathrm{B}(\alpha,\beta)} = I_x(\alpha,\beta)$$

Teorema 4.26 (Beta). Sea $X\sim\mathrm{Beta}(\alpha,\beta)$ con $\alpha,\beta\in\mathbb{R}^+ \Rightarrow$

(B_1) $\mathbb{E}[X^r] = \frac{\mathrm{B}(\alpha+r,\beta)}{\mathrm{B}(\alpha,\beta)} = \prod_{r=0}^{n-1}\frac{\alpha+r}{\alpha+\beta+r}$ r-ésimo momento

(B_2) $\mathbb{E}[X] = \frac{\alpha}{\alpha+\beta}$ Esperanza

(B_3) $\mathrm{Var}[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$ Varianza

Proof. Sea $X \sim \text{Beta}(\alpha, \beta)$

(B_1) Para $r \in \mathbb{N}$, usamos la definición del momento:

$$\mathbb{E}[X^r] = \int_0^1 x^r \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\text{B}(\alpha,\beta)}\,dx = \frac{1}{\text{B}(\alpha,\beta)}\int_0^1 x^{r+\alpha-1}(1-x)^{\beta-1}\,dx = \frac{\text{B}(\alpha+r,\beta)}{\text{B}(\alpha,\beta)}$$

(B_2) Para la esperanza tomamos $r = 1$:

$$\mathbb{E}[X] = \frac{\text{B}(\alpha+1,\beta)}{\text{B}(\alpha,\beta)} = \frac{\alpha}{\alpha+\beta}$$

(B_3) Para la varianza usamos $\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$ con $r = 2$:

$$\mathbb{E}[X^2] = \frac{\text{B}(\alpha+2,\beta)}{\text{B}(\alpha,\beta)} = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}$$

$$(\mathbb{E}[X])^2 = \left(\frac{\alpha}{\alpha+\beta}\right)^2 = \frac{\alpha^2}{(\alpha+\beta)^2}$$

$$\Rightarrow \text{Var}[X] = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \frac{\alpha^2}{(\alpha+\beta)^2} = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

$\therefore$ el Teorema 4.26 es cierto. □

Definición 4.18 (Distribución Pareto). Decimos que una VA X tiene distribución Pareto con parámetros $\alpha > 0$ y $\theta > 0$

$$X \sim \text{Pareto}(\alpha, \theta)$$

si y solo si su PDF es la siguiente

$$f(x) = \text{Pareto}(x; \alpha, \theta) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}} \cdot \mathbb{I}_{\{x \geqslant \theta\}}$$

donde α es el parámetro de forma y θ el parámetro de escala.

Teorema 4.27 (PDF de Pareto). Sea $X \sim \text{Pareto}(\alpha, \theta)$ con $\alpha, \theta \in \mathbb{R}^+$, su PDF $f(x)$ cumple

(P_1) $\forall\, x \in \mathbb{R} \Rightarrow f(x) \geqslant 0$

(P_2) $\int_{-\infty}^{\infty} f(x)\,dx = \int_\theta^\infty \frac{\alpha\theta^\alpha}{x^{\alpha+1}}\,dx = 1$

Proof. Veamos que el Teorema 4.27 es cierto.

(P_1) $f(x) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}} \cdot \mathbb{I}_{\{x \geq \theta\}} \geqslant 0$ ya que $\alpha, \theta > 0$ y $x^{\alpha+1} > 0$ para $x \geq \theta$.

(P_2) Notemos que

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_\theta^\infty \frac{\alpha\theta^\alpha}{x^{\alpha+1}}\,dx = \alpha\theta^\alpha \int_\theta^\infty x^{-\alpha-1}\,dx.$$

Como $\alpha > 0$, la integral converge:

$$\alpha\theta^\alpha \left[\frac{x^{-\alpha}}{-\alpha}\right]_\theta^\infty = \alpha\theta^\alpha \cdot \left(0 - \frac{\theta^{-\alpha}}{-\alpha}\right) = 1.$$

$\therefore$ el Teorema 4.27 es cierto. □

Definición 4.19 (CDF Pareto)**.** Sea $X \sim \text{Pareto}(\alpha, \theta)$ con $\alpha > 0$ y $\theta > 0$. La CDF de X es

$$F(x; \alpha, \theta) = 1 - \left(\frac{\theta}{x}\right)^{\alpha}, \quad x \geqslant \theta$$

Teorema 4.28 (Pareto)**.** Sea $X \sim \text{Pareto}(\alpha, \theta)$ con $\alpha, \theta \in \mathbb{R}^+ \Rightarrow$

(P_1) $\mathbb{E}[X^r] = \frac{\alpha\theta^r}{\alpha - r} \quad r < \alpha$ — r-ésimo momento

(P_2) $\mathbb{E}[X] = \frac{\alpha\theta}{\alpha - 1} \quad \alpha > 1$ — Esperanza

(P_3) $\text{Var}[X] = \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)} \quad \alpha > 2$ — Varianza

Proof. Sea $X \sim \text{Pareto}(\alpha, \theta)$

(P_1) Para $r < \alpha$:

$$\mathbb{E}[X^r] = \int_{\theta}^{\infty} x^r \cdot \alpha\theta^{\alpha} x^{-\alpha-1}\, dx = \alpha\theta^{\alpha} \int_{\theta}^{\infty} x^{r-\alpha-1}\, dx$$

$$= \alpha\theta^{\alpha} \cdot \left[\frac{x^{r-\alpha}}{r-\alpha}\right]_{\theta}^{\infty} = \frac{\alpha\theta^{\alpha}}{\alpha - r} \cdot \theta^{r-\alpha} = \frac{\alpha\theta^r}{\alpha - r}$$

(P_2) Para la esperanza tomamos $r = 1$, requiere $\alpha > 1$:

$$\mathbb{E}[X] = \frac{\alpha\theta}{\alpha - 1}$$

(P_3) Para la varianza usamos que $\text{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$ y requiere $\alpha > 2$:

$$(\mathbb{E}[X])^2 = \left(\frac{\alpha\theta}{\alpha - 1}\right)^2 \Rightarrow \text{Var}[X] = \frac{\alpha\theta^2}{\alpha - 2} - \frac{\alpha^2\theta^2}{(\alpha-1)^2} = \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}$$

$\therefore$ el Teorema 4.28 es cierto. $\square$

Definición 4.20 (Distribución Normal)**.** Decimos que una VA X tiene distribución normal con media μ y varianza σ^2 (es decir, de parámetros μ y σ)

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

si y solo si su PDF es la siguiente

$$f(x) = \mathcal{N}(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

donde $\mu \in \mathbb{R}$ y $\sigma^2 > 0$

Teorema 4.29 (PDF de Normal)**.** Sea $X \sim \mathcal{N}(\mu, \sigma^2)$ con $\mu \in \mathbb{R}$ y $\sigma^2 > 0 \Rightarrow$

($\mathcal{N}_1$) $\forall\, x \in \mathbb{R} \Rightarrow f(x) \geqslant 0$

($\mathcal{N}_2$) $\int_{-\infty}^{\infty} f(x)\, dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx = 1$

Proof. Veamos que el Teorema 4.29 es cierto.

($\mathcal{N}_1$) Claramente $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] \geq 0$ para todo $x \in \mathbb{R}$, ya que $\exp(z) > 0$ para todo z real y $\sigma > 0$.

($\mathcal{N}_2$) Consideramos el cambio de variable $z = \frac{x-\mu}{\sigma} \Rightarrow dx = \sigma\, dz$. Entonces

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right] dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$

Definimos $I := \int_{-\infty}^{\infty} e^{-z^2/2}\, dz$. Entonces

$$I^2 = \left(\int_{-\infty}^{\infty} e^{-z^2/2} dz\right)^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-(z^2+y^2)/2}\, dz\, dy = \int_{\mathbb{R}^2} e^{-(x^2+y^2)/2}\, dx\, dy$$

Ahora usamos coordenadas polares: $x = r\cos\theta$, $y = r\sin\theta$

$\Rightarrow dx\, dy = r\, dr\, d\theta$ y $x^2 + y^2 = r^2$

$$I^2 = \int_0^{2\pi}\int_0^{\infty} e^{-r^2/2} r\, dr\, d\theta$$

Sea $u = \frac{r^2}{2} \Rightarrow du = r\, dr$, entonces

$$I^2 = \int_0^{2\pi}\left(\int_0^{\infty} e^{-u}\, du\right) d\theta = \int_0^{2\pi} 1 \cdot d\theta = 2\pi$$

Por tanto, $I = \sqrt{2\pi}$, así que

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz = \frac{1}{\sqrt{2\pi}} \cdot \sqrt{2\pi} = 1$$

$\therefore$ el Teorema 4.29 es cierto. □

Definición 4.21 (CDF Normal)**.** Sea $X \sim \mathcal{N}(\mu, \sigma^2)$ con $\mu \in \mathbb{R}$ y $\sigma^2 > 0$. La CDF de X es

$$F(x; \mu, \sigma^2) = \Phi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{x-\mu}{\sigma}} e^{-t^2/2}\, dt$$

Teorema 4.30 (Normal)**.** Sea $X \sim \mathcal{N}(\mu, \sigma^2)$ con $\mu \in \mathbb{R}$ y $\sigma^2 > 0 \Rightarrow$

($\mathcal{N}_1$) $m_X(t) = \mathbb{E}[e^{tX}] = \exp\left(\mu t + \frac{\sigma^2 t^2}{2}\right)$ FGM

($\mathcal{N}_2$) $\mathbb{E}[X] = \mu$ Esperanza

($\mathcal{N}_3$) $\text{Var}[X] = \sigma^2$ Varianza

Proof. Sea $X \sim \mathcal{N}(\mu, \sigma^2)$

($\mathcal{N}_1$) Calculemos la función generadora de momentos:

$$m_X(t) = \mathbb{E}[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f_X(x)\, dx = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

completamos el cuadrado en el exponente,

$$tx - \frac{(x-\mu)^2}{2\sigma^2} = -\frac{1}{2\sigma^2}\Big((x-\mu)^2 - 2\sigma^2 t(x-\mu)\Big) + t\mu$$
$$= -\frac{1}{2\sigma^2}\left(x-\mu-\sigma^2 t\right)^2 + \frac{\sigma^2 t^2}{2} + t\mu$$

Notemos que

$$m_X(t) = \frac{e^{t\mu + \frac{\sigma^2 t^2}{2}}}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-\mu-\sigma^2 t)^2}{2\sigma^2}\right) dx$$

Hacemos el cambio de variable $y = \frac{x-\mu-\sigma^2 t}{\sigma}$, entonces $dx = \sigma dy$ y

$$m_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$

$(\mathcal{N}_2)$ Usamos la función cumulante $\Psi_X(t) = \ln m_X(t) = \mu t + \frac{\sigma^2 t^2}{2}$ y (Ψ_2) del Teorema 3.5

$$\Psi_X'(t) = \mu + \sigma^2 t \mid_{t=0} = \mu = \mathbb{E}[X]$$

$(\mathcal{N}_3)$ Derivando de nuevo y evaluando en $t=0$:

$$\Psi_X''(t) = \sigma^2 \Rightarrow \Psi_X''(0) = \sigma^2 = \mathrm{Var}[X]$$

$\therefore$ Teorema 4.30 es cierto. $\square$

Observación. La normal (μ, σ^2) cumple $\mathbb{E}[X] = \mathrm{Mo} = \mathrm{Me}$

Teorema 4.31. Sea $X \sim \mathcal{N}(\mu, \sigma^2)$. Sea $a, b \neq 0 \in \mathbb{R}$. Sea $Y = aX + b$.

$$\Rightarrow Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$$

Proof. Como $m_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}} = \mathbb{E}[e^{tX}]$ sustiyamos por Y

$$m_Y(t) = \mathbb{E}[e^{tY}] = \mathbb{E}[e^{t(aX+b)}] = e^{tb}\mathbb{E}[e^{taX}]$$

$$e^{tb}\left(e^{\mu at + \frac{\sigma^2 at^2}{2}}\right) = e^{\mu at + tb + \frac{\sigma^2 at^2}{2}} = e^{(\mu a + b)t + \frac{\sigma^2 a^2 t^2}{2}}$$

$\therefore Y \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$ $\square$

Corolario. Si $X \sim \mathcal{N}(\mu, \sigma^2) \Rightarrow Y = X - \mu \sim \mathcal{N}(0, \sigma^2)$

Teorema 4.32. Sea $X \sim \mathcal{N}(\mu, \sigma^2)$ y $Y = X - \mu$

$$\Rightarrow \mathbb{E}[Y^r] = \mathbb{E}[(X-\mu)^r] = \begin{cases} 0 & \text{si r es impar} \\ \frac{2k!\sigma^{2k}}{k!2^k} & \text{si r es par} \end{cases}$$

Proof. **Caso 1:** r es impar

$$\mathbb{E}[Y^r] = \int_{-\infty}^{\infty} y^r \frac{1}{\sqrt{2\pi}\cdot\sigma} \cdot \exp\left[-\frac{y^2}{2r^2}\right] dy$$

$$= \frac{1}{\sqrt{2}\sqrt{\pi}\cdot\sigma} \lim_{t\to\infty} \int_{-\infty}^{\infty} y^r \exp\left[-\frac{y^2}{2r^2}\right] dy = \frac{1}{\sqrt{2}\sqrt{\pi}\cdot\sigma} \lim_{t\to\infty} 0 = 0$$

Caso 2: r es par

Recordemos que

$$m_Y(t) = 1 + t\mathbb{E}[Y] + \frac{t^2}{2!}\mathbb{E}[Y^2] + \frac{t^3}{3!}\mathbb{E}[Y^3] = \ldots = \sum_{k=0}^{\infty} \frac{t^{2k}\mathbb{E}[Y^{2k}]}{2k!}$$

Por otro lado si $Y \sim \mathcal{N}(0, \sigma^2) \Rightarrow$

$$m_Y(t) = \exp\left[\frac{\sigma^2 t^2}{2}\right] = \sum_{k=0}^{\infty} \left(\frac{\frac{\sigma^2 t^2}{2}}{k!}\right) = \sum_{k=0}^{\infty} \frac{\sigma^{2k} t^{2k}}{k! 2^k}$$

Estas dos ecuaciones son iguales si los coeficientes de $t^2 k$ i.e.

$$\frac{\mathbb{E}[y^{2k}]}{2k!} = \frac{\sigma^{2k}}{k! 2^k}$$

Despejando a $\mathbb{E}[Y^{2k}]$ s.t.q.

$$\mathbb{E}[Y^r] = \frac{2k!\sigma^{2k}}{k! 2^k}$$

$\therefore$ el Teorema 4.32 es cierto. □

Corolario. Sea $X \sim \mathcal{N}(\mu, \sigma^2) \Rightarrow$ el coeficiente de asimetría $\alpha = 0$ y de kurtosis $\mathcal{K} = 3$

Proof. Por el Teorema 4.32, como 3 es impar

$$\alpha = \frac{\mathbb{E}[(X-\mu)^3]}{\sigma^3} = \frac{\mathbb{E}[Y^3]}{\sigma^3} = 0$$

Ahora, como 4 es par

$$\mathcal{K} = \frac{\mathbb{E}[(X-\mu)^4]}{\sigma^4} = \frac{\mathbb{E}[Y^4]}{\sigma^4} = \frac{\mathbb{E}[Y^{2\cdot 2}]}{\sigma^4} =$$

$$\frac{4!\sigma^4}{2! 2^2 \sigma^4} = \frac{24}{8} = 3$$

□

Ejemplo 4.4 (Distribución Normal Estandar). Decimos que una VA X tiene distribución normal estandarizada si sigue la Definición 4.20 con $\mu = 0$ y $\sigma = 1$.

$$X \sim \mathcal{N}(0, 1)$$

es decir, su PDF es

$$\varphi(x) = \mathcal{N}(x; 0, 1) = \frac{1}{\sqrt{2\pi}} \cdot \exp\left[-\frac{x^2}{2}\right]$$

La CDF se denota como

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t) dt$$

Proof. Todas las propiedades de la PDF se heredan del Teorema 4.29. □

Teorema 4.33 (Teorema de la Transformación). Sea X una VA continua con densidad $f_X(x)$ y soporte S. Sea $Y = g(X)$, con $g : S \to \mathbb{R}$ continua, biyectiva y derivable, con inversa $x = g^{\leftarrow}(y)$ tal que $\frac{d}{dy}g^{\leftarrow}(y) \neq 0$.

$$\Rightarrow f_Y(y) = \left|\frac{d}{dy}g^{\leftarrow}(y)\right| f_X(g^{\leftarrow}(y))\, \mathbb{I}_{S^*}(y)$$

donde $S^* = g(S)$ es el soporte de Y.

Proof. **Caso 1:**

Supongamos que $y = g(x)$ es monótona creciente. En ese caso

$$F_Y(y) = P(Y \leqslant y) = P(g(x) \leqslant y) = P(X \leqslant g^{\leftarrow}(y)) = F_X(g^{\leftarrow}(y))$$

Como g es continua, derivable, y biyectiva $\Rightarrow$

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_Y(g^{\leftarrow}(y)) = f_X(g^{\leftarrow}(y))\frac{d}{dy}g^{\leftarrow}(y)$$

$$\Rightarrow f_Y(y) = \left|\frac{d}{dy}g^{\leftarrow}(y)\right| f_X(g^{\leftarrow}(y))\, \mathbb{I}_{S^*}(y)$$

Caso 2:

Supongamos que $y = g(x)$ es monótona decreciente. En ese caso

$$F_X(y) = P(Y \leqslant y) = P(g(x) \leqslant y) = P(X \geqslant g^{\leftarrow}(y)) = 1 - F_X(g^{\leftarrow}(y))$$

Esto como $g(x)$ es monótona decreciente. Luego

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \frac{d}{dy}[1 - F_X(g^{\leftarrow}(y))] = -f_X(g^{\leftarrow}(y))\frac{d}{dy}g^{\leftarrow}(y)$$

$$\Rightarrow f_Y(y) = \left|\frac{d}{dy}g^{\leftarrow}(y)\right| f_X(g^{\leftarrow}(y))\, \mathbb{I}_{S^*}(y)$$

$\therefore$ el Teorema 4.33 es cierto. □

Ejemplo 4.5 (Distribución Log-Normal). Decimos que una VA X tiene distribución Log-Normal si $\log X \sim \mathcal{N}(\mu, \sigma^2)$, es decir,

$$X \sim \log\mathcal{N}(\mu, \sigma^2)$$

Entonces, su PDF es

$$f(x) = \log\mathcal{N}(x; \mu, \sigma^2) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right], \quad x > 0$$

donde $\mu \in \mathbb{R}$ y $\sigma^2 > 0$ son los parámetros de la normal subyacente.

Proof. Sea $Y = \log X \sim \mathcal{N}(\mu, \sigma^2)$

Usamos el Teorema 4.33 con $X = e^Y \Rightarrow Y = \log X$, que es biyectiva, continua y derivable.

$$\Rightarrow f_X(x) = \left|\frac{d}{dx}\log x\right| \cdot f_Y(\log x) = \frac{1}{x} \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right]$$

$$\Rightarrow f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right], \quad x > 0$$

Además, del Teorema 4.30 podemos obtener:

(L$_1$) $m_X(t) = \mathbb{E}[e^t X] = e^{r\mu + \frac{\sigma^2 r^2}{2}}$

(L$_2$) $\mathbb{E}[X] = \exp\left(\mu + \frac{\sigma^2}{2}\right)$

(L$_3$) $\text{Var}[X] = \left(e^{\sigma^2} - 1\right) e^{2\mu + \sigma^2}$

$\therefore$ la Log-Normal es una distribución. $\square$

Capítulo 5

Teoremas de Límite

5.1. De Moivre-Laplace

Notación. Denotamos $\theta(n)$ a una función tal que $\theta(x) \to 0$ rápidamente cuando $n \to 0$.

Teorema 5.1 (Teorema de De Moivre–Laplace). Sea $X \sim \text{Bin}(n,p)$ con $p \neq 0$ o 1

Cuando $n \to \infty$

$$X \sim \mathcal{N}(\mu = np, \sigma^2 = npq)$$

Es decir, para k en un entorno de np, se puede aproximar

$$\binom{n}{k} p^k q^{n-k} \simeq \frac{1}{\sqrt{2\pi npq}} \exp\left(-\frac{(k-np)^2}{2npq}\right), \quad p+q=1, \quad p,q>0.$$

Proof. Notemos que esto es equivalente a demostrar que $Z = \frac{x-np}{\sqrt{npq}} \sim \mathcal{N}(0,1)$

Busquemos la FGM de Z

$$m_Z(t) = \exp\left(\frac{-npt}{\sqrt{npq}}\right) \cdot m_X\left(\frac{t}{\sqrt{npq}}\right) = \exp\left(\frac{-npt}{\sqrt{npq}}\right)\left(q + p \cdot \exp\left(\frac{t}{\sqrt{npq}}\right)\right)^n$$

$$= \left(\exp\left(\frac{-pt}{\sqrt{npq}}\right)\left(q + p \cdot \exp\left(\frac{t}{\sqrt{npq}}\right)\right)\right)^n = \left(q \cdot \exp\left(\frac{-pt}{\sqrt{npq}}\right) + p \cdot \left(\exp\frac{t-pt}{\sqrt{npq}}\right)\right)^n$$

$$= \left(q \cdot \exp\left(\frac{-pt}{\sqrt{npq}}\right) + p \cdot \exp\left(\frac{tq}{\sqrt{npq}}\right)\right)^n$$

Ahora, notemos que $e^u = 1 + u + \frac{u^2}{2!} + \frac{u^3}{3!}$

$$\Rightarrow \exp\left(\frac{-pt}{\sqrt{npq}}\right) = 1 - \frac{pt}{\sqrt{npq}} + \underbrace{\frac{p^2t^2}{\sqrt{2npq}} - \frac{p^3t^3}{\sqrt{3npq}} + \dots}_{\theta(n)\to 0}$$

$$q \cdot e^{\frac{-pt}{\sqrt{npq}}} = q - \frac{qpt}{\sqrt{npq}} + \frac{qp^2t^2}{\sqrt{2npq}} + \theta(n) \quad \text{y} \quad q \cdot e^{\frac{-pt}{\sqrt{npq}}} = p - \frac{qpt}{\sqrt{npq}} + \frac{pq^2t^2}{\sqrt{2npq}} + \theta(n)$$

Sumando las dos expresiones s.t.q.

$$= \underbrace{q}_{1-p} + p - \frac{qpt}{\sqrt{npq}} + \frac{qpt}{\sqrt{npq}} + \frac{pt^2}{2n} + \frac{qt^2}{2n} + \theta(n)$$

Sustutuyendo en la FGM de Z

$$m_Z(t) = \left[q \cdot e^{\frac{-pt}{\sqrt{npq}}} + p \cdot e^{\frac{tq}{\sqrt{npq}}}\right]^n = \left[1 + \frac{t^2}{2n} + \theta(n)\right]^n$$

$$= \lim_{n\to\infty} \left[1 + \frac{t^2}{2n} + \theta(n)\right]^n = \lim_{n\to\infty} \left[1 + \frac{t^2}{2n}\right]^n = e^{\frac{t^2}{2}}$$

Y recordamos del Ejemplo 4.4 que esto es la FGM $\sim \mathcal{N}(0,1)$

$$Z = \frac{x - np}{\sqrt{npq}} \to_{n\to\infty} \mathcal{N}(0,1)$$

$\therefore X \sim \mathcal{N}(np, npq)$ □

5.2. Desigualdades

Teorema 5.2 (Desigualdad de Markov). Sea X una VA no negativa con media finita y $t > 0$

$$\Rightarrow P(X \geqslant t) \leqslant \frac{\mathbb{E}[X]}{t}$$

Proof. Recordemos de la Definición 3.1 que

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

Con $t > 0$ s.t.q

$$\int_{x\geqslant t} x f_X(x) dx + \int_{x<t} x f_X(x) dx \geqslant \int_{x\geqslant t} x f_X(x) dx$$

$$\Rightarrow \underbrace{x f_X(x)}_{x\geqslant t} \geqslant t f_X(x)$$

$$\Rightarrow \int_{x\geqslant t} x f_X(x) dx \geqslant \int_{x\geqslant t} t f_X(x) dx = \int_t^{\infty} t f_X(x) dx$$

$$\Rightarrow \mathbb{E}[X] \geqslant t \int_t^{\infty} f_X(x) dx \Rightarrow \mathbb{E}[X] \geqslant t P(X \geqslant t)$$

$\therefore P(X \geqslant t) \leqslant \frac{\mathbb{E}[X]}{t}$ □

Teorema 5.3 (Desigualdad de Tchebyschev). Sea X una VA con varianza $\sigma^2 \in \mathbb{R}^+$ y media μ. Sea $t > 0$

$$\Rightarrow P(|X - \mu| \geqslant) \frac{\sigma^2}{t^2}$$

Proof. Sea $Y = (X - \mu)^2 \geqslant 0$

$$\mathbb{E}[Y] = \mathbb{E}[(X - \mu)^2] = \sigma^2$$

Aplicando Teorema 5.2 con $t^2 > 0$

$$P(Y \geqslant t^2) \leqslant \frac{\mathbb{E}[Y]}{t^2} \Leftrightarrow P((X-\mu)^2 \geqslant t^2) \leqslant \frac{\sigma^2}{t^2}$$

$$P(|X-\mu| \geqslant)\frac{\sigma^2}{t^2}$$

$\therefore$ es cierto el Teorema 5.3 □

Corolario. $P(|X-\mu| \leqslant t) \geqslant 1 - \frac{\sigma^2}{t^2}$

Corolario. $P(|X-\mu| \geqslant t\sigma) \leqslant \frac{\sigma^2}{(t\sigma)^2}$

Corolario. $P(|X-\mu| < t\sigma) \geqslant 1 - \frac{1}{t^2}$

Observación (Taylor). Si $f_X(x)$ es dos veces derivable en $c \Rightarrow$

$$f_X(x) = f(c) + \frac{f'(c)(x-c)}{1!} + \frac{f''(c)(x-c)^2}{2!} + \dots$$

Por el Teorema de Valor Intermedio $\exists\, 0 < \varphi < c$ tal que

$$f(x) = f(c) + f'(c)(x-c) + \frac{f''(\varphi)(x-\varphi)}{2}$$

Teorema 5.4 (Desigualdad de Jensen). Sea X una VA con$\mathbb{E}[X] < \infty$ y $g : \mathbb{R} \to \mathbb{R}$ con $\mathbb{E}[g(X)] < \infty \Rightarrow$

(J_1) Si $g(x)$ es cóncava $\Rightarrow \mathbb{E}[g(X)] \geqslant g(\mathbb{E}[X])$

(J_2) Si $g(x)$ es convexa $\Rightarrow \mathbb{E}[g(X)] \leqslant g(\mathbb{E}[X])$

Proof. Aplicando Taylor con $g(x)$ alrededor de μ

$$g(X) = g(\mu) + \frac{g'(\xi)(x-\mu)}{1!} + \frac{g''(\xi)(x-\mu)^2}{2!}$$

Tomando $\mathbb{E}[\cdot]$ de ambos lados

$$\mathbb{E}[g(X)] = \mathbb{E}[g(\mu)] + \underbrace{g'(\mu)(x-\mu)}_{=0} + \frac{g''(\xi)}{2!}\mathbb{E}[(x-\xi)^2]$$

$$\Rightarrow \mathbb{E}[g(X)] = g(\mathbb{E}[X]) + \frac{g''(\xi)}{2!}\mathbb{E}[(x-\xi)^2]$$

Si g es cóncava $g''(\xi) > 0 \Rightarrow$

$$\mathbb{E}[g(X)] \geqslant g(\mathbb{E}[X])$$

Si g es convexa $g''(\xi) > 0 \Rightarrow$

$$\mathbb{E}[g(X)] \leqslant g(\mathbb{E}[X])$$

$\therefore$ el Teorema 5.6 es cierto. □

Corolario. Sea X una VA con media $\mathbb{E}[X] < \infty$ y $g(X) = X^2$

$$\mathbb{E}[X^2] \geqslant \mathbb{E}[X]^2 \Rightarrow \mathbb{E}[X^2] \geqslant \mathbb{E}^2[X] \geqslant 0$$

Teorema 5.5 (Desigualdad de Chernoff). Sea X una variable aleatoria y $t > 0$

$$\forall\, a \in \mathbb{R} \Rightarrow P(X \geqslant a) \leqslant e^{-ta}\, \mathbb{E}\left[e^{tX}\right]$$

Esta cota es válida siempre que $\mathbb{E}\left[e^{tX}\right] < \infty$ De manera análoga, si $t < 0$,

$$P(X \leqslant a) \leqslant e^{-ta}\, \mathbb{E}\left[e^{tX}\right]$$

5.3. LGN y TCL

Definición 5.1 (Convergencia Puntual). Sea un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Diremos que la sucesión de VA (que son funciones) X_n converge puntualmente a X si

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1$$

Definición 5.2 (Convergencia de Probabilidad). Sea un espacio de probabilidad $(\Omega, \mathcal{F}, P)$. Diremos que la sucesión de VA X_n converge a la VA X si $\forall\, \varepsilon > 0$ s.t.q.

$$\lim_{n\to\infty} P\left(|X_n - X| > \varepsilon\right) = 0$$

i.e. cuando $n \to \infty$ la probabilidad de que la sucesión de VA X_n este *lejos* de X es nula.

Notación. Podemos denotar la convergencia en probabilidad como $X_n \xrightarrow{p} X$

Teorema 5.6 (Ley de Kolmogorov). La media muestral $\overline{X}_n$ converge puntualmente a la esperanza, o media poblacional $\mathbb{E}[\overline{X}_n] = \mu$. Recordando que las VA son funciones de $\Omega \to \mathbb{R}$

Cuando $n \to \infty$

$$\forall s \in \Omega \Rightarrow \overline{X}_n(s) \to \mu$$

Es decir

$$P\left(\lim_{n\to\infty} \overline{X}_n = \mu\right) = 1$$

Teorema 5.7 (Ley de Khinchin). Cuando $n \to \infty$

$$\overline{X}_n \xrightarrow{p} \mu$$

Es decir, $\forall\, \varepsilon > 0$

$$\lim_{n\to\infty} P\left(\left|\overline{X}_n - \mu\right|\right) = 0$$

Proof. Usamos el Teorema 5.3. Supongamos un $\varepsilon > 0$ fijo pero arbitrario

$$P\left(\left|\overline{X}_n - \mu\right| > \varepsilon\right) \leqslant \frac{\sigma^2}{n \cdot \varepsilon^2}$$

Cuando $n \to \infty$ el lado derecho tiende a cero. □

Notación. Nos referimos al Teorema 5.6 como la Ley de Grandes Números (LGN) fuerte al Teorema 5.7 como la LGN débil.

Teorema 5.8 (Teorema Central del Límite)**.** Sea $(\Omega, \mathcal{F}, P)$ un espacio muestral. Supongamos que tenemos $X_1, X_2, ... \in \mathcal{F} \subseteq \Omega$ variables aleatorias independientes e identicamente distribuidas con media poblacional μ y varianza σ^2, y media muestral $\overline{X}_n$

Cuando $n \to \infty$

$$\sqrt{n} \cdot \left(\frac{\overline{X}_n - \mu}{\sigma} \right) \to \mathcal{N}(0, 1)$$

Observación. Es decir, cuando n se hace muy muy muy pero muy muy grande, al hacer la estandarización de la distribución de la VA $\overline{X}_n$ se acerca a una distribución normal estándar.

Teorema 5.9 (Aprox. de TCL)**.** Para grandes n, la distribución de $\overline{X}_n$ es aproximadamente

$$\mathcal{N} \left(\mu, \frac{\sigma^2}{n} \right)$$

Lema 5.1 (Lema de Borel-Cantelli)**.** Sea $(A_n)_{n \geq 1}$ una sucesión de eventos en un espacio de probabilidad $(\Omega, \mathcal{F}, P)$.

(B$_1$) Si $\sum_{n=1}^{\infty} P(A_n) < \infty$, entonces

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

es decir, la probabilidad de que ocurran infinitos A_n es cero.

(B$_2$) Si los eventos A_n son independientes y $\sum_{n=1}^{\infty} P(A_n) = \infty$, entonces

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

es decir, ocurren infinitos A_n casi seguramente.

Definición 5.3 (Función Característica)**.** Sea X una variable aleatoria con distribución en $\mathbb{R}^n$. La función característica de X es la función $\varphi_X : \mathbb{R}^n \to \mathbb{C}$ definida por

$$\varphi_X(t) = \mathbb{E} \left[e^{i\langle t, X \rangle} \right]$$

donde $\langle t, X \rangle$ denota el producto interno usual en $\mathbb{R}^n$ e $i = \sqrt{-1}$.

Teorema 5.10 (Teorema de Continuidad de Lévy)**.** Sea $(X_n)_{n \geq 1}$ una sucesión de variables aleatorias con funciones características φ_{X_n}, y sea X una variable aleatoria con función característica φ_X. Entonces,

$$X_n \xrightarrow{d} X \quad \Leftrightarrow \forall\, t \in \mathbb{R} \Rightarrow \varphi_{X_n}(t) \to \varphi_X(t)$$

Es decir, la convergencia en distribución de X_n a X es equivalente a la convergencia puntual de sus funciones características.