# UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO

## FACULTAD DE ESTUDIOS SUPERIORES ACATLÁN

CONCEPTOS FUNDAMENTALES DE PROBABILIDAD

T E S I N A

QUE PARA OBTENER EL TITULO DE

**LICENCIATURA EN ACTUARIA**

**PRESENTA**

**MONTIEL TOVAR MANUEL FERNANDO**

**ASESORA: ACT. LUZ MARÍA LAVÍN ALANÍS**

**OCTUBRE 2012**

**INTRODUCCIÓN:**

La probabilidad mide la posibilidad con la que se obtiene un resultado (o conjunto de resultados) al llevar a cabo un experimento o fenómeno aleatorio, del que se conocen todos los resultados posibles, bajo condiciones específicas. La teoría de la probabilidad se usa extensamente en áreas como la estadística, la física, la matemática, la ciencia y la actuaría para sacar conclusiones sobre la ocurrencia de sucesos potenciales y la mecánica subyacente de sistemas complejos.

La probabilidad constituye un importante parámetro en la determinación de las diversas casualidades obtenidas tras una serie de eventos esperados dentro de un rango estadístico.

Constituye un tema de estudio fundamental para la carrera de Actuaria y se incluye en la formación curricular por la importancia que representa, pues los conocimientos de probabilidad y estadística son las bases indispensables y de aplicación práctica en el ámbito laboral.

En la expectativa teórica de la presente tesina se relaciona con el desarrollo de estrategias de aprendizaje mediante la utilización de medios electrónicos de consulta de información que ayudarán en la asimilación significativa.

El propósito principal es elaborar un material didáctico de apoyo para la impartición de la asignatura de Probabilidad II del plan de estudios de Actuaria vigente, en la Facultad de Estudios Superiores Acatlán, así que el material didáctico aportará información significativa de cada unidad y así brindar al lector un complemento del contenido de aprendizaje e integrar al alumno en un conocimiento global de la materia, este material se apega al plan de estudios vigente, se basa principalmente en la Bibliografía que aparece en el temario, recopilando los aspectos y conocimientos más importantes de parte de todos los autores citados en el programa de estudios, también enfoca todos los puntos de vista de los distintos autores, para que la comprensión sea mucho más accesible para los alumnos, profundizando en los temas que merecen mayor atención, sin olvidar que todos los temas tienen importancia para la construcción del conocimiento significativo y fueron seleccionados para formar parte del temario. Por tal motivo, en el presente trabajo desarrollé temas trascendentes en la comprensión y aplicación de la Probabilidad.

Este material será un apoyo para que los estudiantes puedan fortalecer su conocimiento, resolver sus dudas o consultar información. La ventaja que se puede encontrar es que contiene una sola notación, ya que toda la bibliografía en la cual se basa el curso utiliza distintas notaciones, así que se analiza toda la información bibliográfica para unificar la notación, también sucede que algunos autores dan por hecho que el lector ya tiene conocimientos previos y además que ciertas demostraciones teóricas son explicadas de manera muy trivial, pero algunos alumnos se pierden en el procedimiento de solución por falta de comprensión en los pasos para obtener el resultado. Así que se retoman las bases y se explica de manera sencilla los conceptos, los procedimientos y como se estructuran estos para lograr la solución.

Algo que se busca de manera muy importante es disminuir los altos índices de reprobación de la materia en la carrera de Actuaria. Otro problema existente para el aprendizaje es que en el contenido de la bibliografía en la que se basa el curso contiene libros sin traducción al español o libros inexistentes en la Biblioteca de la Facultad.

Por tal motivo, el presente trabajo desarrolla los temas de la Probabilidad de la siguiente forma:

CAPITULO 1: El estudiante reconocerá los principios y fundamentos de la Teoría de las Probabilidades a partir de la Axiomática de Kolmogorov. Así que se enuncian de manera resumida los diferentes conceptos de probabilidad y estadística necesarios para la construcción de experimentos aleatorios, así como también los resultados principales de la teoría de probabilidad. Se propone una amplia serie de ejemplos de espacios que no sean de probabilidad y ejemplos de eventos independientes y dependientes así como de transformaciones medibles y no medibles.

CAPITULO 2: El alumno reafirmará sus conocimientos con respecto a las Variables Aleatorias su función de distribución y la obtención de sus momentos para extenderlo formalmente al concepto de vectores aleatorios, se analizan experimentos y fenómenos aleatorios que involucren a dos o más variables aleatorias, sucesiones de variables aleatorias a fin de determinar su límite respectivo, se plantean y resuelven problemas de vectores aleatorios así como ejercicios para determinar funciones de probabilidad marginal y condicional a partir de funciones de distribución y densidad de probabilidad conjunta y se calcula el grado de dependencia entre componentes de un vector aleatorio.

CAPITULO 3: El estudiante identificará los diferentes tipos de convergencia Estocástica, y su relación con las Leyes de los Grandes Números. Así que se proponen ejemplos específicos de convergencia en probabilidad, se plantean ejemplos de eventos aleatorios  que se repiten sucesivamente, con el propósito de encontrar su probabilidad frecuentista y se dan ejemplos concretos para demostrar que en general la cota superior dada por la desigualdad de Chebyshev es óptima.

CAPITULO 4: El alumno identificará las propiedades básicas e implicaciones del Teorema de Limite Central, así como sus bondades y limitaciones en su aplicación a fenómenos reales. Se ven aplicaciones del Teorema de Inversión para Distribuciones específicas de probabilidad, se ejemplifica como encontrar la distribución de probabilidad de sumas de variables aleatorias independientes para algunos casos particulares, así como también se estudian las condiciones que establece el Teorema del Limite Central en el caso especifico de la Distribución de Cauchy y se plantea problemas para calcular y aproximar probabilidades mediante la Distribución Normal.

## ÍNDICE

# Conceptos Fundamentales de Probabilidad.

Objetivo General.

Analizar la Ley de los Grandes Números, la Función Característica y el Teorema del Límite Central, así como diferentes ideas acerca de las variables aleatorias y de la probabilidad y leyes relacionadas con ellas.

## Capítulo 1 ESPACIOS DE PROBABILIDAD.

Objetivo: Modelación de los fenómenos o experimentos aleatorios, reconocer los principios y fundamentos de la teoría de las probabilidades a partir de la axiomática de Kolmogorov.

### 1.1 Conjuntos y σ-algebras.

**Definición:** Un conjunto es una colección de objetos con una característica en común.

Sea $\Omega$ un conjunto arbitrario y sean $\mathrm{A}, \mathrm{B} \subset \Omega$. Entonces

$(\mathrm{A} \subset \Omega, \mathrm{B} \subset \Omega)$.

1) $\mathrm{A} \subset \mathrm{B}$ si $\forall x \in \mathrm{A} \Rightarrow x \in \mathrm{B}$.

(A esta contenido en B, es decir, A es subconjunto de B).

2) Al conjunto que no tiene elementos, se le llama el conjunto vacio, denotado por $\emptyset$.

   Nota: $\forall \mathrm{A} \subset \Omega, \emptyset \subset \mathrm{A}$.

Podemos definir “operaciones” sobre conjuntos

3) Intersección: $\mathrm{A} \cap \mathrm{B} \wedge \forall x \in \mathrm{A}, x \in \mathrm{B}$,
   es decir, $\mathrm{A} \cap \mathrm{B} = \{x \in \mathrm{A} \wedge x \in \mathrm{B}\}$.

4) Unión: $A \cup B \wedge x \in A$ ó $x \in B$,
 es decir, $A \cup B = \{x | x \in A \vee x \in B\}$.

5) Dos conjuntos de $\Omega$ son mutuamente excluyentes o disconjuntos si: $A \cap B = \emptyset$.

6) Sean $A$ y $B$ subconjuntos de $\Omega$, la diferencia:
 $A \setminus B$ o $(A - B) = \{x \mid x \in A \wedge x \notin B\}$.

7) Complemento $A^c = \{x \mid x \in (\Omega - A)\}$ así que por definición:
 $A - B = \{x \mid x \in A \wedge x \notin B\} \Rightarrow \{x \mid x \in A \wedge x \in B^c\} = A \cap B^c$.
 entonces $A - B = A \cap B^c$.

8) Diferencia geométrica ( $A \Delta B$ ).
 Sean $A$ y $B$ subconjuntos de $\Omega$.
 $\Rightarrow A \Delta B = \{x \mid x \in [(A \cup B) - (A \cap B)]\}$.

**Proposición 1:**

Sean $A, B$ y $C$ subconjuntos de $\Omega$.

1) $A \subseteq A \cup B$.
 $B \subseteq A \cup B$.
2) $A \cup B = B \cup A$.
3) $(A \cup B) \cup C = A \cup (B \cup C)$.
4) $A \cap B \subseteq A$.
 $A \cap B \subseteq B$.
5) $A \cap B = B \cap A$.
6) $A \cap (B \cap C) = (A \cap B) \cap C$.
7) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ .
8) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ .
9) $(A^c)^c = A$.
10) $A \cup A^c = \Omega$.
11) $A \cap A^c = \emptyset$.
12) $A \cup B = (A \Delta B) \Delta (A \cap B)$.
13) $A - B = A \Delta (A \cap B)$.
14) $(A \cup B)^c = A^c \cap B^c$.
15) $(A \cap B)^c = A^c \cup B^c$.
16) $A - (B \cap C) = (A - B) \cup (A - C)$.
17) $A - (B \cup C) = (A - B) \cap (A - C)$.

Las demostraciones de los incisos de la proposición anterior se dejan como ejercicios para el lector.

**Definición**:

$\mathrm{C}$ es una clase de subconjuntos de $\Omega$ si es un conjunto cuyos elementos son subconjuntos de $\Omega$.

Ejemplo:

Sea $\mathrm{E} \subset \Omega$ si $\mathrm{C} = \{ \mathrm{E}, \mathrm{E}^{\mathrm{c}}, \emptyset, \Omega \}$.

Entonces $\mathrm{C}$ es un conjunto de subconjuntos de $\Omega$, ya que

$\mathrm{E} \subset \Omega$, $\mathrm{E}^{\mathrm{c}} \subset \Omega$, $\emptyset \subset \Omega$, $\Omega \subset \Omega$ $\therefore$ $\mathrm{C}$ es una clase de $\Omega$.

**Definición**:

Sea $\mathrm{A} \subset \Omega$. El conjunto potencia de A (denotado por $\prod(\mathrm{A})$ o $2^{\mathrm{A}}$) se define como,

$$\prod(\mathrm{A}) = \{\mathrm{B} \subset \Omega : \mathrm{B} \subset \mathrm{A}\}.$$

Ejemplo:

Sea $\mathrm{A} = \{\mathrm{a}, \mathrm{b}, \mathrm{c}\}$.

$\prod(\mathrm{A}) = \{ \emptyset, (\mathrm{a}), (\mathrm{b}), (\mathrm{c}), (\mathrm{a}, \mathrm{b}), (\mathrm{a}, \mathrm{c}), (\mathrm{b}, \mathrm{c}), (\mathrm{a}, \mathrm{b}, \mathrm{c}) \}$.

Nota: El conjunto potencia $\prod(\cdot)$ es una clase y es la mayor clase de cualquier conjunto.

**Definición**:

Cardinalidad de un conjunto es el número de elementos de un conjunto $(\#)$.

$$\mathrm{A}^{\#}, \#\mathrm{A}, n(\mathrm{A}).$$

Ejemplo:

$$A = \{a, b, c\} \qquad A^{\#} = 3.$$

$$\prod(A)^{\#} = 8 = 2^3$$

Si $A^{\#} = n$, entonces

$$\prod(A)^{\#} = 2^n = 2^{A^{\#}}.$$

Nota: Las operaciones de conjuntos tales como la unión, la intersección y la diferencia geométrica, las podemos extender a una clase o a una sucesión de eventos o conjuntos.

**Definición**:

Sucesión de conjuntos.

Sean $A_1, A_2, A_3, \dots$ subconjuntos de $\Omega \Rightarrow A_1, A_2, A_3, \dots$ es una sucesión y se denotara por $\{A_n\}$ con $n \in \mathbb{N}$.

$$\bigcup_{n \in N} A_n = \{x \mid x \in A_1 \text{ o } x \in A_2 \text{ o} \dots \text{ o } x \in A_n \text{ o} \dots\}$$

$$\bigcap_{n \in N} A_n = \{x \mid x \in A_1 \text{ y } x \in A_2 \text{ y } \dots \text{ y } x \in A_n \text{ y} \dots\}$$

**Definición**:

Sea $\{A_n\}_{n \in \mathbb{N}}$ $\forall\ n \geq 1$ una sucesión de subconjuntos de $\Omega$.

1) Se dice que la sucesión es no-decreciente ó expansiva $(A_n \uparrow)$ si $A_n \subset A_{n+1}$.
2) Se dice que la sucesión es no-creciente ó contráctil $(A_n \downarrow)$ si $A_n \supset A_{n+1}$.
3) Se dice que la sucesión es monótona si es no-decreciente o no-creciente.

**Definición**:

El límite de una sucesión monótona de conjuntos es.

$$\text{A)} \quad \lim_{n\to\infty} \mathrm{A_n} = \bigcup_{k=1}^{\infty} \mathrm{A_k} \qquad \text{si } \mathrm{A_n} \uparrow.$$

$$\text{B)} \quad \lim_{n\to\infty} \mathrm{A_n} = \bigcap_{k=1}^{\infty} \mathrm{A_k} \qquad \text{si } \mathrm{A_n} \downarrow.$$

Ejemplo:

Sea

$$\mathrm{A_n} = \left\{ x \,\middle|\, 0 \le x \le \frac{1}{\mathrm{n}} \right\} \quad \mathrm{n} \ge 1 \quad \mathrm{n} \in \mathbb{N}, \qquad x \in \mathbb{R}.$$

Entonces

$$\mathrm{A_1} = \{ x \mid 0 \le x \le 1 \}.$$

$$\mathrm{A_n} = \left\{ x \,\middle|\, 0 \le x \le \frac{1}{\mathrm{n}} \right\}.$$

Sea $x \in \mathrm{A_{n+1}} \Rightarrow 0 \le x \le \frac{1}{\mathrm{n+1}}$ pero $0 \le x \le \frac{1}{\mathrm{n+1}} \le 1 \Rightarrow 0 \le x \le 1 \Rightarrow x \in \mathrm{A_n}$.

Asi que $\mathrm{A_1} \supset \mathrm{A_2} \supset \ldots \supset \mathrm{A_n}$.

Entonces $\mathrm{A_{n-1}} \supset \mathrm{A_n} \Rightarrow \mathrm{A_n} \downarrow$ y por lo tanto se considera no-creciente.

$$\lim_{\mathrm{k}\to\infty} \mathrm{A_k} = \bigcap_{\mathrm{k}=1}^{\infty} \mathrm{A_k} = \bigcap_{\mathrm{k}=1}^{\infty} \left\{ x \,\middle|\, 0 \le x \le \frac{1}{\mathrm{k}} \right\} = \emptyset \qquad .$$

**Proposición 2:**

Sea $\{A_n\}_{n\geq 1}$ una sucesión de subconjuntos de $\Omega$. Entonces

1. Si $A_n \uparrow \Rightarrow A_n^c \downarrow$.
2. Si $A_n \downarrow \Rightarrow A_n^c \uparrow$.

Demostración:

Sea $\{A_n\}_{n\geq 1}$ es no-decreciente,

$$A_n \subseteq A_{n+1} \ (\forall\ n \in \mathbb{N}).$$

Entonces

$$A_{n+1}^c \subseteq A_n^c \ (\forall\ n \in \mathbb{N}).$$

Así, la sucesión $\{A_n^c\}_{n\geq 1}$ es no-decreciente.

∎

Un resultado muy importante de las leyes de De Morgan es:

$$A^c = \bigcap_{k=1}^{\infty} A_k^c \ .$$

Que puede ser útil para algunas demostraciones.

**Proposición 3:**

Sea $\{A_n\}_{n\geq 1}$ una sucesión de subconjuntos de $\Omega$. Entonces

$$\text{Si } A_n \downarrow \;\Rightarrow\; A_n^c \uparrow \;\Rightarrow\; \lim_{n\to\infty} A_n^c = \bigcup_{k=1}^{\infty} A_k^c = A^c .$$

Demostración:

Por las definiciones anteriores sabemos que si $A_n \downarrow$ .

$$A = \lim_{n\to\infty} A_n = \bigcap_{k=1}^{\infty} A_k \quad \Rightarrow \quad A^c = \left(\bigcap_{k=1}^{\infty} A_k\right)^c .$$

Aplicando las leyes de De Morgan.

$$A^c = \bigcup_{k=1}^{\infty} A_k^c .$$

∎

**Definición**:

Sea $\{A_k\}$ con $k \in \mathbb{N}$ una sucesión de subconjuntos de $\Omega$.

1) Se define al supremo de la sucesión al conjunto.

$$C_n = \sup_{k\geq n} A_k = \bigcup_{k=n}^{\infty} A_k .$$

2) Se define al ínfimo de la sucesión al conjunto.

$$B_n = \inf_{k\geq n} A_k = \bigcap_{k=n}^{\infty} A_k .$$

**Definición**:

1) Se define al límite superior de la sucesión al conjunto.

$$C = \lim_{n\to\infty} \sup_{k\geq n} A_k = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k .$$

Sea.

$$\mathrm{C_n} = \bigcup_{k=n}^{\infty} \mathrm{A_k} \qquad \mathrm{C_{n+1}} = \bigcup_{k=n+1}^{\infty} \mathrm{A_k} \ .$$

$$\mathrm{C_n} \supset \mathrm{C_{n+1}} \implies \mathrm{C_n} \downarrow \mathrm{C}\,.$$

$$\mathrm{C} = \lim_{\mathrm{n}\to\infty} \mathrm{C_n}\ .$$

Entonces

$$\mathrm{C} = \lim_{\mathrm{n}\to\infty} \mathrm{C_n} = \bigcap_{n=1}^{\infty} \mathrm{C_n} = \bigcap_{n=1}^{\infty}\left(\bigcup_{k=n}^{\infty} \mathrm{A_k}\right) .$$

2) Se define al límite inferior de la sucesión al conjunto.

$$\mathrm{B} = \lim_{\mathrm{n}\to\infty} \inf_{\mathrm{k}\geq\mathrm{n}} \mathrm{A_k} = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} \mathrm{A_k}\ .$$

$$\mathrm{B_n} = \bigcap_{k=n}^{\infty} \mathrm{A_k} \qquad \mathrm{B_{n+1}} = \bigcap_{k=n+1}^{\infty} \mathrm{A_k}\ .$$

$$\bigcap_{k=n}^{\infty} \mathrm{A_k} \subset \bigcap_{k=n+1}^{\infty} \mathrm{A_k} \implies \mathrm{B_n} \subset \mathrm{B_{n+1}} \implies \mathrm{B_n} \uparrow.$$

Entonces

$$\mathrm{B} = \lim \mathrm{B_n} = \bigcup_{n=1}^{\infty} \mathrm{B_n} = \bigcup_{n=1}^{\infty}\left(\bigcap_{k=n}^{\infty} \mathrm{A_k}\right) .$$

**Definición**:

Si la sucesión de conjuntos $\{\mathrm{A_k}\}_{\mathrm{k}\in\mathbb{N}}$ cumple con $\mathrm{C} = \mathrm{B} = \mathrm{A}$ , se dice que $\{\mathrm{A_k}\}_{\mathrm{k}\in\mathbb{N}}$ converge a $\mathrm{A}$.

**Definición:**

Toda sucesión monótona de conjuntos es convergente. Tenemos 2 casos:

1) Sea $A_n \uparrow \Longrightarrow A_n \subset A_{n+1}$ .

2) Sea $A_n \downarrow \Longrightarrow A_n \supset A_{n+1}$ .

Demostración:

Sea una sucesión no-decreciente.

Entonces.

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

Entonces y de acuerdo con las definiciones anteriores

$$B = \bigcap_{n=1}^{\infty}\left(\bigcup_{k=n}^{\infty} A_k\right) = \bigcap_{n=1}^{\infty} A_n = A \, .$$

Por otro lado.

$$A_* = \bigcup_{n=1}^{\infty}\left(\bigcap_{k=n}^{\infty} A_k\right) \ .$$

Como.

$$A_n \uparrow \Longrightarrow A_n \subset A_{n+1} \Longrightarrow \bigcap_{k=n}^{\infty} A_k = A_n \, .$$

Entonces.

$$C = \bigcup_{n=1}^{\infty} (A_n) = A \, .$$

Por la definición Anterior.

$$\therefore \ B = C = A \, .$$

■

**Definición**:

Si $\mathcal{F}$ es una clase de conjuntos y si realizamos sobre ella cierta operación en $\mathcal{F}$ y obtenemos un elemento que pertenece a $\mathcal{F}$, diremos que la clase es cerrada bajo dicha operación que se utilizo.

Ejemplo:

Si $\mathcal{F} \in \mathrm{A} \Rightarrow \mathrm{A}^{c} \in \mathcal{F} \quad \forall \ \mathrm{A} \in \mathcal{F}$ , $\mathcal{F}$ es cerrada bajo el complemento.

Para cada $\mathrm{A}, \mathrm{B} \in \mathcal{F} \Rightarrow \mathrm{A} \cup \mathrm{B} \in \mathcal{F}$ la unión es cerrada.

$\mathcal{F}$ es cerrada bajo la unión.

**Definición**:

Si una clase de conjuntos $\mathcal{F}$ es tal que.

1) $\Omega \in \mathcal{F}$ .
2) Si $\mathrm{A} \in \mathcal{F} \Rightarrow \mathrm{A}^{c} \in \mathcal{F}$.
3) Si $\mathrm{A}_1 \wedge \mathrm{A}_2 \in \mathcal{F} \Rightarrow$

$$\bigcup_{k=1}^{n} \mathrm{A}_{\mathrm{k}} \in \mathcal{F}.$$

Entonces se dice que $\mathcal{F}$ es una álgebra.

**Proposición 4**:

Sea $\{\mathrm{A}_{\mathrm{n}}\}_{\mathrm{n} \in \mathbb{N}}$ una sucesión de eventos y sea $\mathrm{B}_1 = \mathrm{A}_1$ y para $\mathrm{n} \geq 2$.

$$\mathrm{B}_{\mathrm{n}} = \mathrm{A}_{\mathrm{n}} - \bigcup_{k=1}^{n-1} \mathrm{A}_{\mathrm{k}} \ .$$

Entonces $\{\mathrm{B}_{\mathrm{n}}\}_{\mathrm{n} \in \mathbb{N}}$ es una sucesión de eventos con las siguientes propiedades.

$\forall \ \mathrm{n}, \mathrm{m}$ tal que $\mathrm{n} \neq \mathrm{m}$

$\mathrm{i})\ \mathrm{B}_{\mathrm{n}} \subseteq \mathrm{A}_{\mathrm{n}}$ .
$\mathrm{ii})\ \mathrm{B}_{\mathrm{n}} \cap \mathrm{B}_{\mathrm{m}} = \emptyset$ .
$\mathrm{iii})\ \bigcup_{n=1}^{\infty} \mathrm{B}_{\mathrm{n}} = \bigcup_{n=1}^{\infty} \mathrm{A}_{\mathrm{n}}$ .

Demostración:

ii) $B_n \cap B_m =$

$$= \left(A_n - \bigcup_{k=1}^{n-1} A_k\right) \cap \left(A_m - \bigcup_{k=1}^{n-1} A_k\right)$$

$$= \left(A_n \bigcap_{k=1}^{n-1} A_k^c\right) \cap \left(A_m \bigcap_{k=1}^{n-1} A_k^c\right) \qquad \text{si } m < n \Rightarrow A_m \subset A_n$$

$$\Rightarrow \left(A_n \bigcap_{k=1}^{n-1} A_k^c\right) \cap \left(A_m \bigcap_{k=1}^{n-1} A_k^c\right) = \emptyset\,.$$

iii) por i) $B_n \subseteq A_n$ .

$$\text{si } B_n = A_n \Rightarrow \bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n\,.$$

∎

**Definición**:

El espacio de probabilidad se define como una terna $(\Omega, \mathcal{F}, \mathcal{P})$.

$\Omega$ Es un conjunto arbitrario (espacio muestral).

$\mathcal{F}$ σ-algebra de subconjuntos de $\Omega$.

$\mathcal{P}$ Es una medida de probabilidad definida sobre $\mathcal{F}$ correspondiente.

**Definición**:

Espacio Muestral $(\Omega)$ conjunto de todos los resultados posibles de un experimento aleatorio, o el conjunto de todas las observaciones que puede presentar un fenómeno ó experimento aleatorio. Y se llama experimento o fenómeno aleatorio a aquél que es susceptible de dar varios resultados, no pudiéndose predecir de antemano cuál de ellos va a producirse en una experiencia concreta.

A la pareja ordenada $(\Omega, \mathcal{F})$ se le llama un espacio medible. Donde la σ-álgebra $(\mathcal{F})$ clase o colección no vacía de ese espacio muestral $(\Omega)$ cerrado bajo las operaciones de complemento y uniones numerables, a los elementos de $\mathcal{F}$ se les llama eventos ó conjuntos medibles. Y estos son simples (un elemento de $\Omega$) o compuestos (2 ó más elementos de $\Omega$).

**Definición**:

Sea $(\Omega, \mathcal{F})$ un espacio medible y sea $\mathcal{P}\colon f \to [0,1]$.

Decimos que $\mathcal{P}$ es una medida de probabilidad si satisface las siguientes propiedades:

1) $\mathcal{P}(\Omega) = 1.$
2) $\mathcal{P}(\mathrm{A}) \geq 0 \ \forall \, \mathrm{A} \in \mathcal{F}$ .
3) Si $\{\mathrm{A_n}\}_{n \geq 1} \in \mathcal{F}$ .

Con $\mathrm{A_n} \cap \mathrm{A_m} = \emptyset \ \ y \ \ \mathrm{n} \neq \mathrm{m}$ .

Entonces

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

$\mathcal{P}(\mathrm{A})$ Es la medida de ocurrencia del evento $\mathrm{A}$.

**Definición**:

Una colección $\mathcal{F}$ de subconjuntos de $\Omega$ es una σ-algebra si cumple con las siguientes condiciones:

1) $\Omega \in \mathcal{F}.$
2) $\text{Si } \mathrm{A} \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}.$
3) $$\text{Si } \mathrm{A_1}, \dots, \mathrm{A_n} \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} \mathrm{A}_i \in \mathcal{F}.$$

Ejemplos:

Sea $\Omega$ un conjunto cualquiera no vacio, las siguientes colecciones son σ-algebras de subconjuntos de $\Omega$.

1) $\mathcal{F}_1 = \{\Omega,\ \emptyset\}$ .

i) $\Omega \in \mathcal{F}$.
ii) $\emptyset = \Omega^c$ .
iii) $\Omega \cup \emptyset = \Omega \in \mathcal{F}_1$.
Entonces $\mathcal{F}_1$ es σ-algebra.

Nota: $\mathcal{F}_1$ es la σ-algebra más pequeña que podemos asociar a un conjunto cualquiera $\Omega$.

2) $\mathcal{F}_2 = \{\Omega,\ \emptyset,\ A,\ A^c\}$ donde $A \subseteq \Omega$.

i) $\Omega \in \mathcal{F}_2$.
ii) $\emptyset^c,\ A^c,\ (A^c)^c,\ \Omega^c \in \mathcal{F}_2$ .
iii) Si $A \subseteq \Omega \Rightarrow A \cup A^c = \Omega \Rightarrow \emptyset \cup A \cup A^c \cup \Omega = \Omega \in \mathcal{F}_2$.
Entonces $\mathcal{F}_2$ es σ-algebra.

Ejercicio:

Demostrar que $\mathcal{F}_3 = (2^\Omega)$ (conjunto potencia) es una σ-algebra y que es la más grande asociada a $\Omega$, también se puede encontrar como $\prod(\Omega)$ .

Ejemplo:

Sean $A$ y $B$ subconjuntos de $\Omega$ tal que $A \subseteq B$ la colección.

$$\mathcal{F} = \{\ \emptyset, A, B, A^c, B^c, (B-A), (B-A)^c, \Omega\ \} .$$

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

ii) $\emptyset^c = \Omega \in \mathcal{F}$.
$A^c \in \mathcal{F}$.
$B^c \in \mathcal{F}$.
$A \in \mathcal{F}$.
$B \in \mathcal{F}$.
$(B-A)^c \in \mathcal{F}$.
$(B-A) \in \mathcal{F}$.

iii) Si $A \subseteq B \Rightarrow (B-A) \subseteq B \subseteq \Omega$
$\Rightarrow (B-A)^c \subseteq \Omega$
$\emptyset \cup A \cup B \cup A^c \cup B^c \cup (B-A) \cup (B-A)^c \cup \Omega = \Omega$
$\Omega \in \mathcal{F}$ Entonces $\mathcal{F}$ es una σ-algebra.

**Proposición 5**:

Sea $\mathcal{F}$ una σ-algebra de subconjuntos de $\Omega$ entonces.

1) $\emptyset \in \mathcal{F}$ .
2) Si

$$\mathrm{A}_1,\ \mathrm{A}_2,\ldots \in \mathcal{F} \Rightarrow \bigcap_{n=1}^{\infty} \mathrm{A}_n \in \mathcal{F}.$$

3) Si $\mathrm{A}$ y $\mathrm{B} \in \mathcal{F} \Rightarrow (\mathrm{A} - \mathrm{B}) \in \mathcal{F}$.
4) Si $\mathrm{A}$ y $\mathrm{B} \Rightarrow (\mathrm{A}\,\Delta\,\mathrm{B}) \in \mathcal{F}$.

Demostración:

1) Por definición de σ-álgebra sabemos que

$$\Omega \in \mathcal{F} \Rightarrow \Omega^c = \emptyset \in \mathcal{F}.$$

2) Si $A_1,\ A_2, \ldots \in \mathcal{F}$ entonces
   Entonces por definición de σ-algebra.

$$\mathrm{A}_1^c,\ \mathrm{A}_2^c, \ldots \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} \mathrm{A}_n \in \mathcal{F}.$$

   Entonces

$$\bigcup_{n=1}^{\infty} \mathrm{A}_n^c \in \mathcal{F}.$$

   Utilizando las Leyes de De Morgan.

$$\bigcup_{n=1}^{\infty} \mathrm{A}_n = \left(\bigcap_{n=1}^{\infty} \mathrm{A}_n\right)^c .$$

   Entonces por ser σ-algebra su complemento también pertenece a $\mathcal{F}$ y por definición

$$\left(\left(\bigcap_{n=1}^{\infty} \mathrm{A}_n\right)^c\right)^c \in \mathcal{F} \Rightarrow \bigcap_{n=1}^{\infty} \mathrm{A}_n \in \mathcal{F}.$$

3) Sabemos que $\mathrm{A}$ y $\mathrm{B} \in \mathcal{F}$ .

   Entonces $\mathrm{A}^{\mathrm{C}} \cap \mathrm{B}^{\mathrm{C}} \in \mathcal{F}$.
   Entonces $\mathrm{A} \cap \mathrm{B}^{\mathrm{C}} \in \mathcal{F} \wedge (\mathrm{A} \cap \mathrm{B})^c = (\mathrm{A} - \mathrm{B}) \in \mathcal{F}$.

4) Sabemos que

$$(A \,\Delta\, B) = (A - B) \cup (B - A).$$

Sabemos también que $A \in \mathcal{F}$ y $B \in \mathcal{F}$, y por definición σ-algebra $A \cup B \in \mathcal{F}$ y por 2) $A \cap B \in \mathcal{F}$.

Y como $A \cup B \in \mathcal{F}$ y $A \cap B \in \mathcal{F}$ ⇒ por 3) $(A \cup B) - (A \cap B) \in \mathcal{F}$ ∴ $(A \,\Delta\, B) \in \mathcal{F}$.

∎

Nota: Las σ-álgebras son estructuras también cerradas bajo las operaciones de diferencia e intersecciones numerables.

**Proposición 6**:

La intersección de 2 σ-álgebras es una σ-álgebra.

Demostración:

Sean $\mathcal{F}_1$ y $\mathcal{F}_2$ dos σ-álgebras de $\Omega$.

Sea $\{A_n\}_{n \geq 1} \in \mathcal{F}_1 \cap \mathcal{F}_2$ .

Entonces $\{A_n\}_{n \geq 1} \in \mathcal{F}_1$ y $\{A_n\}_{n \geq 1} \in \mathcal{F}_2$.

Como $\mathcal{F}_1$ y $\mathcal{F}_2$ son σ-álgebras de $\Omega$ entonces

Sea $A$ un elemento de $\mathcal{F}_1 \cap \mathcal{F}_2$ .

Entonces $A \in \mathcal{F}_1$ y $A \in \mathcal{F}_2$ por ser σ-álgebra.

Entonces $A^c \in \mathcal{F}_1$ y $A^c \in \mathcal{F}_2$.

∴ $A^c \in (\mathcal{F}_1 \cap \mathcal{F}_2)$

$$\bigcup_{n\geq 1}^{\infty} A_n \in \mathcal{F}_1 \quad \text{y} \quad \bigcup_{n\geq 1}^{\infty} A_n \in \mathcal{F}_2 \;\Rightarrow\; \bigcup_{n\geq 1}^{\infty} A_n \in \mathcal{F}_1 \cap \mathcal{F}_2$$

∎

Entonces Sean $\mathcal{F}_1$ y $\mathcal{F}_2$ son dos σ-álgebras de un mismo conjunto $\Omega$ , entonces $\mathcal{F}_1 \cap \mathcal{F}_2$ es también una σ-álgebra de subconjuntos de $\Omega$ más pequeña que $\mathcal{F}_1$ y $\mathcal{F}_2$.

Dado que $\mathcal{F}_1 \cap \mathcal{F}_2 \subseteq \mathcal{F}_2 \wedge \mathcal{F}_1 \cap \mathcal{F}_2 \subseteq \mathcal{F}_2$.

Nota: La unión de 2 σ-álgebras no es necesariamente una σ-álgebra.

**Proposición 7**:

La intersección finita, infinita numerable o bien arbitraria de σ-álgebras es nuevamente una σ-álgebra.

Demostración:
Sea $\mathrm{T}$ un conjunto arbitrario de índices.

1) Como $\mathcal{F}_1$ es una σ-álgebra $(\forall\ t \in \mathrm{T})\ \ \Omega \in \mathcal{F}_\mathrm{T}$.

Entonces $\Omega \in \bigcap_{t \in \mathrm{T}} \mathcal{F}_\mathrm{T}$.

2) Sea $\mathrm{A}$ un elemento tal que

$$\mathrm{A} \in \bigcap_{t \in \mathrm{T}} \mathcal{F}_\mathrm{T} \Rightarrow \mathrm{A}^c \in \mathcal{F}_\mathrm{T} \quad \forall\ t \in \mathrm{T}.$$

Entonces pertenece a

$$\bigcap_{t \in \mathrm{T}} \mathcal{F}_\mathrm{T}\ .$$

3) Sean $\mathrm{A}_1, \mathrm{A}_2, \ldots$ elementos tales que

$$\mathrm{A}_1, \mathrm{A}_2, \ldots \in \bigcap_{t \in \mathrm{T}} \mathcal{F}_\mathrm{T}.$$

Entonces $\mathrm{A}_1, \mathrm{A}_2, \ldots \in \mathcal{F}_\mathrm{T} \quad \forall\ t \in \mathrm{T}$.
Entonces

$$\bigcup_{n=1}^{\infty} \mathrm{A}_n \in \mathcal{F}_\mathrm{T} \quad \forall\ t \in \mathrm{T} \Rightarrow \bigcup_{n=1}^{\infty} \mathrm{A}_n \in \bigcap_{t \in \mathrm{T}} \mathcal{F}_\mathrm{T}.$$

∎

Por lo tanto

$$\bigcap_{t \in T} \mathcal{F}_T.$$

Es una σ-álgebra donde es un conjunto arbitrario.

Por lo tanto la intersección arbitraria de σ-álgebras es una σ-álgebra.

**Definición**:

Sea $C$ una colección no vacía de subconjuntos de $\Omega$, la σ-álgebra generada por $C$, denotada por $\sigma(C)$, es la colección.

$\sigma(C) = \cap \{\mathcal{F} \mid \mathcal{F}$ es una σ-álgebra y $C \subseteq \mathcal{F}\}$.

Recordando que $\sigma(C)$ es la mínima σ-álgebra generada por $C$.

Ejemplo:

Sea $A \wedge B \subset \Omega$ con $A \cap B = \emptyset$.

Defina la colección $C = \{A, B\}$.
En general $C$ no es σ-álgebra, $C$ sería σ-algebra añadiendo algunos subconjuntos de Ω, que sería la σ-algebra generada por $C$.

Entonces $C = \{\emptyset, A, B, A^c, B^c, (A \cup B), (A \cup B)^c, \Omega\}$.

Demostración:

i) $\Omega \in \sigma(C)$ .
ii) $\{\emptyset, A, B, A^c, B^c, (A \cup B), (A \cup B)^c, \Omega\} \in \sigma(C)$ .
iii) $\emptyset \cup A \cup B \cup A^c \cup B^c \cup (A \cup B) \cup (A \cup B)^c \cup \Omega = \Omega$.

$\Omega \in \sigma(C) \therefore \sigma(C)$ es σ-álgebra.

∎

Como $A^c = \Omega - A \subseteq \Omega$.

$A^c \subseteq \Omega$.

Entonces $\sigma(C)$ es la más pequeña (la mínima) de las σ-álgebra que podemos generar.

$$\sigma(C) = \{\emptyset, A, B, A^c, B^c, (A \cup B), (A \cup B)^c, \Omega\}.$$

**Proposición 8**:

Sea $C_1 \wedge C_2$ 2 colecciones de subconjuntos de $\Omega$ tal que $C_1 \subseteq C_2$.

Entonces $\sigma(C_1) \subseteq \sigma(C_2)$.

Demostración:

Si $C_1 \subseteq \sigma(C_1) \wedge C_2 \subseteq \sigma(C_2)$ y si $C_1 \subseteq C_2 \Rightarrow C_1 \subseteq C_2 \subseteq \sigma(C_2) \Rightarrow C_1 \subseteq \sigma(C_2)$.

Por definición de σ-álgebra $\sigma(C_1)$ es la mínima σ-álgebra $\Rightarrow \sigma(C_1) \subseteq \sigma(C_2)$.

∎

**Proposición 9**:

Si $\mathcal{F}$ es una σ-álgebra $\Rightarrow \sigma(\mathcal{F}) = \mathcal{F}$.

Demostración:

Sabemos que $\mathcal{F} \subseteq \sigma(\mathcal{F})$.

Por otro lado, como $\mathcal{F}$ es una σ-álgebra.

Entonces $\sigma(\mathcal{F}) \subseteq \mathcal{F} \Rightarrow \sigma(\mathcal{F}) = \mathcal{F}$.

∎

**1.2 Conjuntos de Borel.**

Considere toda la colección de todos los intervalos abiertos $(a, b) \in \mathbb{R}$ donde $a \leq b$.
A la mínima σ-álgebra generada por esta colección se le llama σ-álgebra de Borel de $\mathbb{R}$ y se le denota como $\mathfrak{B}(\mathbb{R})$.

**Definición**:

$\mathfrak{B}(\mathbb{R}) = \sigma\{(a\,,b) \subseteq \ \mathbb{R} \mid a \leq b\}$.

A los elementos de $\mathfrak{B}(\mathbb{R})$ se les puede llamar: borelianos, conjuntos de borel, conjuntos medibles ó conjuntos de Borel medibles.

Es posible asociar la σ-álgebra $\mathfrak{B}(\mathbb{R})$ al conjunto de números reales y obtener el espacio medible $(\mathbb{R}\,,\mathfrak{B}(\mathbb{R}))$.

**Proposición 10**:

Para cualesquiera dos números reales $a \leq b$, los elementos.

$$[a,b],(a,\infty),(-\infty,b),[a,b),(a,b],\{a\}.$$

Son todos elementos de $\mathfrak{B}(\mathbb{R})$.

Demostración:

$$[a,b] = \bigcap_{n=1}^{\infty}\left(a - \frac{1}{n}\,,b + \frac{1}{n}\right) \ \in \ \mathfrak{B}(\mathbb{R}).$$

Los intervalos cerrados $[a,b]$ son conjuntos Borelianos ya que son igual a la intersección numerable de intervalos abiertos.

Nota: Cada elemento de la intersección anterior es un conjunto Boreliano, ya que $\mathfrak{B}(\mathbb{R})$ es σ-álgebra, la intersección infinita es un elemento de $\mathfrak{B}(\mathbb{R})$.

Por lo tanto el elemento $[a,b]$ que nos representa un intervalo cerrado $\in \mathfrak{B}(\mathbb{R})$.

$$(a,\infty) = \bigcup_{n=1}^{\infty} (a,a+n) \ \in \ \mathfrak{B}(\mathbb{R}).$$

$$(-\infty,b) = \bigcup_{n=1}^{\infty} (b-n,b) \ \in \ \mathfrak{B}(\mathbb{R}).$$

$$[a,b) = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n}\,,b\right) \ \in \ \mathfrak{B}(\mathbb{R}).$$

$$(a,b] = \bigcap_{n=1}^{\infty} \left(a,b + \frac{1}{n}\right) \ \in \ \mathfrak{B}(\mathbb{R}).$$

$$\{a\} = \bigcap_{n=1}^{\infty} \left(a - \frac{1}{n},b + \frac{1}{n}\right) \ \in \ \mathfrak{B}(\mathbb{R}).$$

Por lo tanto los elementos de un solo valor también son $\mathfrak{B}(\mathbb{R})$.

Nota: Los complementos, intersecciones y uniones de esos conjuntos numerables también son $\mathfrak{B}(\mathbb{R})$.

**Proposición 11**:

Las siguientes σ-álgebras son todas idénticas a $\mathfrak{B}(\mathbb{R})$.

a) $\sigma\{\, [a,b] \mid a \leq b\}$.
b) $\sigma\{\, (a,b] \mid a \leq b\}$.
c) $\sigma\{\, [a,b) \mid a \leq b\}$.
d) $\sigma\{\, (a,\infty) \mid a \in \mathbb{R}\}$.
e) $\sigma\{\, (-\infty,b) \mid b \in \mathbb{R}\}$.

Demostración:

$\sigma\{\, [a,b] \mid a \leq b\} = \mathfrak{B}(\mathbb{R})$.

$[a,b] \subseteq \mathfrak{B}(\mathbb{R})$ por las proposiciones antes mencionadas,

Implican que $\sigma\{\, [a,b] \mid a \leq b\} \subseteq \sigma\big(\mathfrak{B}(\mathbb{R})\big) = \mathfrak{B}(\mathbb{R})$.

$$(a,b) \subseteq [a,b].$$

$$\mathfrak{B}(\mathbb{R}) = \sigma\{\, (a,b) \mid a \leq b\} \subseteq \sigma\{\, [a,b] \mid a \leq b\}.$$

∎

Asimismo, es posible definir a $\mathfrak{B}(\mathbb{R})$ como la mínima σ-álgebra generada por la colección de todos los subconjuntos abiertos de $\mathbb{R}$, llegando a la conclusión de que la σ-álgebra generada es $\mathfrak{B}(\mathbb{R})$.

**Definición**:

Sea $A \in \mathfrak{B}(\mathbb{R})$ la σ-álgebra de Borel de $A$, se denota por $\mathfrak{B}(A)$ ó $A \cap \mathfrak{B}(\mathbb{R})$ y se define como:

$$\mathfrak{B}(A) = \{\, A \cap B \mid B \in \mathfrak{B}(\mathbb{R})\}.$$

Entonces $B(A)$ es una σ-álgebra de subconjuntos de $A$.

**Definición**:

Sea $\mathrm{C}$ la colección de todos los rectángulos abiertos de $\mathbb{R}^2$, es decir,

$$\mathrm{C} = \{(\mathrm{a},\mathrm{b}) \times (\mathrm{c},\mathrm{d}) \mid \mathrm{a} \leq \mathrm{b},\ \ \mathrm{c} \leq \mathrm{d}\}.$$

Se definen los conjuntos de Borel de $\mathbb{R}^2$ como los elementos de la mínima σ-álgebra generada por $\mathrm{C}$, es decir,

$$\mathfrak{B}(\mathbb{R}^2) = \sigma(\mathrm{C}).$$

ó

$$\mathfrak{B}(\mathbb{R}^2) = \{\mathfrak{B}(\mathbb{R}) \times \mathfrak{B}(\mathbb{R})\}.$$

Generalizando

$$\mathfrak{B}(\mathbb{R}^{\mathrm{n}}) = \{\mathfrak{B}(\mathbb{R}) \times \dots \times \mathfrak{B}(\mathbb{R})\}_{(\mathrm{n-veces})}.$$

## 1.3 Espacios de Probabilidad.

El espacio de probabilidad es una terna $(\Omega, \mathcal{F}, \mathcal{P})$ donde:

$(\Omega)$ Es el espacio muestral, conjunto arbitrario.

$(\mathcal{F})$ σ-álgebra formada por subconjuntos de $\Omega$.

$(\mathcal{P})$ Medida de probabilidad asociada $\mathcal{F}$.

**Definición**:

Sea $(\Omega, \mathcal{F})$ es un espacio medible.

Una medida de probabilidad $\mathcal{P}$ es una función.

$\mathcal{P}: \mathcal{F} \to [0, 1]$ que satisface.

(Axiomas de Kolmogorov).

i) $\mathcal{P}(\Omega) = 1.$

ii) $\mathcal{P}(\mathrm{A}) \geq 0 \quad \forall \ \mathrm{A} \in \mathcal{F}.$

iii) Si $\mathrm{A}_1, \mathrm{A}_2, \dots \in \mathcal{F}$ son eventos disjuntos dos a dos, es decir, $\mathrm{A}_i \cap \mathrm{A}_j = \emptyset \quad \forall \ i \neq j \geq 1.$

$$\mathcal{P}(\mathrm{A}_1 \cup \mathrm{A}_2 \dots \cup \mathrm{A}_\mathrm{n} \dots) = \mathcal{P}\left(\bigcup_{\mathrm{n}=1}^{\infty} \mathrm{A}_\mathrm{n}\right) = \sum_{i=1}^{\infty} \mathcal{P}\,(\mathrm{A}_i).$$

Si solo si $\mathrm{A}_i \cap \mathrm{A}_j = \emptyset \quad \forall \quad i \neq j \geq 1.$

$\mathrm{A}_i \in \mathcal{F}.$

Entonces toda función $\mathcal{P}$ definida sobre una σ-álgebra $\mathcal{F}$ con valores en $[0,1]$ y que cumple con los axiomas de Kolmogorov se le llama “medida de probabilidad”.

Ejemplo:

Un experimento aleatorio con espacio muestral finito $\Omega$ con Cardinalidad $\Omega^{\#}$ asociado a la σ-álgebra del conjunto potencia $\prod(\Omega)$.

Entonces para toda $\mathrm{A} \subset \Omega$ se define

$$\mathcal{P}(\mathrm{A}) = \frac{\mathrm{A}^{\#}}{\Omega^{\#}}.$$

Entonces $\mathcal{P}$ es una medida de probabilidad y se le conoce como la definición clásica de probabilidad.

i) $\mathcal{P}(\mathrm{A}) = \frac{\Omega^{\#}}{\Omega^{\#}} = 1.$

ii) $\forall \; \mathrm{A} \in \Omega \qquad \mathcal{P}(\mathrm{A}) \geq 0.$

$\mathrm{A} \subseteq \Omega.$

$\Rightarrow \mathrm{A}^{\#} \leq \Omega^{\#} \quad \Rightarrow \quad \frac{\mathrm{A}^{\#}}{\Omega^{\#}} \leq 1.$

$\Omega^{\#} > 0$ y como $\mathrm{A} \subset \Omega \;\Rightarrow\; \mathrm{A}^{\#} \geq 0.$

$\Rightarrow \mathcal{P}(\mathrm{A}) = \frac{\mathrm{A}^{\#}}{\Omega^{\#}} \geq 0.$

iii) Sean $\mathrm{A}_1, \mathrm{A}_2, \dots, \mathrm{A}_\mathrm{n}$ una partición de $\Omega$.

$$\bigcup_{i=1}^{\mathrm{n}} \mathrm{A}_i = \Omega.$$

$$\Rightarrow \Omega^{\#} = \sum_{i=1}^{\mathrm{n}} \mathrm{A}_i^{\#}.$$

$$\Rightarrow \mathcal{P}\left(\bigcup_{i=1}^{\mathrm{n}} \mathrm{A}_i\right) = \frac{\left(\bigcup_{i=1}^{\mathrm{n}} \mathrm{A}_i\right)^{\#}}{\Omega^{\#}} = \frac{\sum_{i=1}^{\mathrm{n}} \mathrm{A}_i^{\#}}{\Omega^{\#}} = \frac{\mathrm{A}_1^{\#}}{\Omega^{\#}} + \dots + \frac{\mathrm{A}_\mathrm{n}^{\#}}{\Omega^{\#}} = \sum_{i=1}^{\mathrm{n}} \mathcal{P}(\mathrm{A}_i).$$

Ejemplo:

Considere un evento aleatorio cuyo espacio muestral es $\mathbb{N}$, asocie a este espacio muestral la σ-álgebra correspondiente al conjunto potencia $2^{\mathrm{n}}$.

Entonces para toda $A \subset \mathbb{N}$ Defina

$$\mathcal{P}(\mathrm{A_n}) = \sum_{\mathrm{n}=1}^{\infty} \frac{1}{2^{\mathrm{n}}} \qquad \forall \;\; \mathrm{n} \in \mathbb{N}, \quad \mathrm{n} \in \mathrm{A}.$$

i)

$$\mathcal{P}(\mathbb{N}) = \sum_{\mathrm{n}=1}^{\infty} \frac{1}{2^{\mathrm{n}}} = \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \dots$$

$$= \frac{1}{2}\left(1 + \frac{1}{2^2} + \frac{1}{2^3} + \dots\right) = \frac{1}{2}\left(1 + \frac{1}{1 - \frac{1}{2}}\right) = \frac{1}{2}\,(2) = 1.$$

ii)

$$\mathcal{P}(\mathrm{A_n}) = \sum_{\mathrm{n}=1}^{\infty} \frac{1}{2^{\mathrm{n}}}.$$

como

$$2^{\mathrm{n}} > 0 \quad \Rightarrow \quad \frac{1}{2^{\mathrm{n}}} > 0 \quad \therefore \quad \sum_{\mathrm{n}=1}^{\infty} \frac{1}{2^{\mathrm{n}}} > 0.$$

iii)

$$\mathcal{P}\left(\bigcup_{\mathrm{n}=1}^{\infty} \mathrm{A}_i\right) = \sum_{\mathrm{n}=\bigcup_{\mathrm{n}=1}^{\infty} \mathrm{A}_i}^{\infty} \frac{1}{2^{\mathrm{n}}}.$$

Por ser mutuamente excluyentes

$$= \sum_{i=1}^{\infty} \sum_{\mathrm{n} \in \mathrm{A}_i} \frac{1}{2^{\mathrm{n}}} = \sum_{\mathrm{n}=1}^{\infty} \mathcal{P}(\mathrm{A}_i).$$

∎

Ejemplo:

Considere al espacio medible $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ una función no negativa e integrable, $f_x(x)$ cuya integral en el intervalo $(-\infty, \infty)$, es igual a 1.

Entonces para cualquier $\mathrm{A}$ que pertenezca a $(\mathbb{R}, \mathfrak{B}(\mathbb{R}))$ específicamente $\mathrm{A} \subset \mathfrak{B}(\mathbb{R})$, se define:

$$\mathcal{P}(\mathrm{A}) = \int_{\mathrm{A}} f_x(x)\, dx.$$

i) Por definición de la función.

$$\mathcal{P}(\mathbb{R}) = \int_{-\infty}^{\infty} f_x(x)\, dx = 1.$$

ii)

$$\mathcal{P}(\mathrm{A}) \geq 0 \quad \Rightarrow \quad \mathcal{P}(\mathrm{A}) = \int_{a_1}^{a_2} f_x(x)\, dx \geq 0.$$

iii)

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

$$\bigcup_{\mathrm{n}=1}^{\infty} \mathrm{A_n} = \mathrm{A_1} \cup \mathrm{A_2} \cup \dots$$

$$\mathcal{P}\left(\bigcup_{\mathrm{n}=1}^{\infty} \mathrm{A_n}\right) = \mathcal{P}(\mathrm{A_1}) + \mathcal{P}(\mathrm{A_2}) + \dots$$

$$= \int_{\mathrm{A_1}} f_x(x)\, dx + \int_{\mathrm{A_2}} f_x(x)\, dx + \dots$$

Esto implica que $\mathcal{P}(\mathrm{A})$ sea una medida de probabilidad.

■

**Proposición 12**:

Sea $\mathcal{P}$ una medida de probabilidad.

1) $\mathcal{P}(\mathrm{A})^{c} = 1 - \mathcal{P}(\mathrm{A})$.

Demostración:

$$\mathrm{A} \cup \mathrm{A}^{c} = \Omega.$$
$$\mathrm{A} \cap \mathrm{A}^{c} = \emptyset.$$

Entonces $\mathcal{P}(\mathrm{A} \cup \mathrm{A}^{c}) = \mathcal{P}(\Omega) = 1$.
Entonces $\mathcal{P}(\mathrm{A}) \cup \mathcal{P}(\mathrm{A}^{c}) = 1$.

Entonces $\mathcal{P}(\mathrm{A}^{c}) = 1 - \mathcal{P}(\mathrm{A})$.

∎

2) $\mathcal{P}(\emptyset) = 0$.

Demostración:

$$\Omega^{c} = \emptyset \quad \wedge \quad \emptyset^{c} = \Omega.$$

$\mathcal{P}(\emptyset) = \mathcal{P}(\Omega^{c}) = 1 - \mathcal{P}(\Omega) = 1 - 1 = 0.$

∎

3) Si $\mathrm{A} \subseteq \mathrm{B} \Rightarrow \mathcal{P}(\mathrm{B} - \mathrm{A}) = \mathcal{P}(\mathrm{B}) - \mathcal{P}(\mathrm{A})$.

Demostración:

Si $\mathrm{A} \subseteq \mathrm{B} \Rightarrow \mathrm{B} = \mathrm{A} \cup (\mathrm{B} - \mathrm{A})$.

Entonces $\mathcal{P}(\mathrm{B}) = \mathcal{P}(\mathrm{A}) + \mathcal{P}(\mathrm{B} - \mathrm{A})$.

$\mathcal{P}(\mathrm{B} - \mathrm{A}) = \mathcal{P}(\mathrm{B}) - \mathcal{P}(\mathrm{A})$.

∎

4) Si $A \subseteq B \Rightarrow \mathcal{P}(A) \leq \mathcal{P}(B)$.

Demostración:

$$\mathcal{P}(A) = \mathcal{P}(B) + \mathcal{P}(B - A).$$

Por el axioma 2 $\quad \mathcal{P}(B - A) \geq 0 \quad$ y $\quad \mathcal{P}(A) \geq 0$.

Entonces $\mathcal{P}(A) \leq \mathcal{P}(B)$.

$\blacksquare$

5) $0 \leq \mathcal{P}(A) \leq 1$.

Demostración:

Sabemos que $A \subseteq \Omega$.

$$\mathcal{P}(A) \leq \mathcal{P}(\Omega) = 1.$$

Por Axioma 2 $\quad \mathcal{P}(A) \geq 0 \quad \Rightarrow \quad 0 \leq \mathcal{P}(A) \leq 1$.

$\blacksquare$

6) Para 2 eventos cualesquiera $A \wedge B$ en $\Omega$.
$\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B)$.

Demostración:

$$(A \cup B) = (B - A) \cup (A \cap B) \cup (A - B).$$

$$(A \cup B) = \big(B - (A \cap B)\big) \cup (A \cap B) \cup \big(A - (A \cap B)\big).$$

Por axioma 3.

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

$$\mathcal{P}(A \cup B) = \mathcal{P}(A) + \mathcal{P}(B) - \mathcal{P}(A \cap B).$$

$\blacksquare$

**Propiedades de las medidas de probabilidad.**

**Proposición 13**:

Desigualdad de Boole.

Sea $\{A_n |\ n \in \mathbb{N}\}$ una sucesión de eventos.

1)

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

2)

$$\mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right) \leq 1 - \sum_{n=1}^{\infty} \mathcal{P}(A_n^c).$$

Demostración:

La probabilidad de la unión (de que al menos uno de estos ocurra) es menor o igual a la suma de las probabilidades de los eventos individuales.

1)

Sea

$$B_n = A_n - \bigcup_{k=1}^{n-1} A_k .$$

Con $A_1 = B_1$.

donde $\{A_n |\ n \in \mathbb{N}\} \Rightarrow \{B_n |\ n \in \mathbb{N}\}$.

i) $B_n \subseteq A_n$.
ii) $B_n \cap B_n = \emptyset \quad \forall\, n \neq m \qquad n,\ m \in \mathbb{N}$.
iii)

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

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

$$= \sum_{n=1}^{\infty} \mathcal{P}(A_n) - \sum_{n=1}^{\infty} \mathcal{P}\left(\bigcup_{k=1}^{n-1} A_k\right) \geq 0.$$

Por el axioma 2 de Kolmogorov

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

∎

2)

$$\mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = 1 - \mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right)^{c}.$$

Y con Kolmogorov esto es igual a

$$1 - \mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n^{c}\right).$$

Y como

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

Entonces

$$\mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right) \geq 1 - \sum_{n=1}^{\infty} \mathcal{P}(A_n^{c})$$

∎

**Proposición 14**:

Sea $\{A_n | \, n \in \mathbb{N}\}$ una sucesión de eventos.

1) Si

$$\mathcal{P}(A_n) = 1 \quad \forall \; n \in \mathbb{N} \quad \Rightarrow \; \mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = 1.$$

Demostración:

$$\mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = 1 - \mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n^c\right).$$

Y

$$\mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right) \geq 1 - \sum_{n=1}^{\infty} \mathcal{P}(A_n^c).$$

Entonces si $\mathcal{P}(A_n) = 1 \quad \Rightarrow \quad \mathcal{P}(A_n^c) = 0.$
Entonces

$$\mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = 1.$$

∎

2) Si $\mathcal{P}(A_n) = 1$ para alguna $n$ entonces

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

Demostración:

Sabemos que

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

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

Como $\mathcal{P}(A_n) = 1$ por Proposiciones anteriores, entonces

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

∎

3) Si $\mathcal{P}(A_n) = 1$ para alguna $n$ entonces

$$\mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = 0.$$

Demostración:

Por proposiciones anteriores

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

Entonces

$$0 \leq \mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right) \leq \mathcal{P}(A_n) = 0.$$

Entonces

$$\mathcal{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = 0.$$

∎

4) Si

$$\mathcal{P}(A_n) = 0 \;\; \forall \; n \in \mathbb{N} \Rightarrow \mathcal{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = 0.$$

Demostración:

Por la desigualdad de Boole.

$$0 \leq \mathcal{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathcal{P}(A_n) = \sum_{n=1}^{\infty}(0) = 0.$$

Por axioma 2 de Kolmogorov, entonces

$$\mathcal{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = 0.$$

∎

Recordando que los Axiomas de Kolmogorov son los siguientes.

1) $\mathcal{P}(A) \geq 0.$
2) $\mathcal{P}(\Omega) = 1.$
3) $\mathcal{P}(A_1 \cup A_2 \cup \dots) = \Sigma\, \mathcal{P}(A_i).$

**Continuidad.**

Las funciones de probabilidad son funciones continuas, para sucesiones crecientes.

**Proposición 15**:

Sea $\{A_n |\, n \in \mathbb{N}\}$ una sucesión no decreciente de eventos, es decir,

$$A_1 \subseteq A_2 \subseteq A_3 \dots$$

Entonces

$$\Rightarrow \mathcal{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mathcal{P}(A_n).$$

Demostración:

Como $A_n \subseteq A_{n+1}$.

Entonces por las proposiciones anteriores tenemos que $\mathcal{P}(A_n) \leq \mathcal{P}(A_{n+1})$.

Por lo tanto la sucesión numérica $\{\mathcal{P}(A_n) \,|\, n \in \mathbb{N}\}$ es NO decreciente y acotada superiormente $\Rightarrow \exists$ el límite.

Sea

$$B_1 = A_1 \qquad B_m = A_m - A_{m-1}.$$

$$B_n \cap B_m = \emptyset \quad m \neq n.$$

Es decir que $\{B_n | \, n \in \mathbb{N}\}$ es una colección de eventos disjuntos 2 a 2 tales que:

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

Entonces

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

$$= \mathcal{P}(B_1) + \sum_{n=2}^{\infty} \mathcal{P}(B_n)$$

$$= \mathcal{P}(A_1) + \sum_{n=2}^{\infty} [\mathcal{P}(A_n) - \mathcal{P}(A_{n-1})]$$

$$= \mathcal{P}(A_1) + \lim_{m \to \infty} \sum_{n=2}^{m} [\mathcal{P}(A_n) - \mathcal{P}(A_{n-1})]$$

$$= \mathcal{P}(A_1) + \lim_{m \to \infty} [\mathcal{P}(A_m) - \mathcal{P}(A_{m-1})]$$

$$= \mathcal{P}(A_1) + \lim_{m \to \infty} \mathcal{P}(A_m) - \mathcal{P}(A_1) = \lim_{m \to \infty} \mathcal{P}(A_m).$$

∎

Para sucesiones no crecientes (decrecientes)

**Proposición 16**:

Sea $\{A_n | \, n \in \mathbb{N}\}$ una sucesión no decreciente de eventos, es decir,

$$A_1 \supseteq A_2 \supseteq A_3 \ldots$$

Entonces

$$\mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \to \infty} \mathcal{P}(A_n).$$

Demostración:

Sabemos que si $A_1 \supseteq A_2 \supseteq A_3 \ldots$

$$A_n \supseteq A_{n+1} \Rightarrow A_n^c \subseteq A_{n+1}^c.$$

Por la proposición anterior tenemos que

$$\mathcal{P}\left(\bigcup_{n=1}^{\infty} A_n^c\right) = \lim_{n\to\infty} \mathcal{P}(A_n^c).$$

Aplicando las Leyes de De Morgan.

$$1 - \mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} [1 - \mathcal{P}(A_n)].$$

$$\lim_{n\to\infty} 1 - \lim_{n\to\infty} \mathcal{P}(A_n).$$

Entonces

$$\mathcal{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} A_n.$$

∎

Ejemplo:

El problema del mono.
Un mono escribe caracteres al azar en una máquina de escribir sin ninguna tendencia ¿Cuál es la probabilidad de que eventualmente el mono obtenga exactamente y sin ningún error las obras completas de Shakespeare?.

Demostramos a continuación que la probabilidad de este raro evento es uno. Imagine entonces que un mono escribe caracteres al azar en una máquina de escribir y que lo hace de manera continua generando una sucesión lineal de caracteres.

Sea $m =$ al número de caracteres disponibles en la máquina de escribir.

$N =$ es el total de caracteres que tienen las obras de Shakespeare, es decir el tamaño de la muestra.

Todas estas muestras son de tamaño $N$.

1. $AAA \ldots$
2. $ABA \ldots$
3. $AAB \ldots$

$\vdots$

Para cada número natural $k$, defina $A_k \rightarrow k - \text{esimo}$ bloque que contiene exactamente y sin errores las obras completas, donde $A_k$ son eventos independientes (no importa la ocurrencia de uno de ellos para que se presente otro).

$$\Rightarrow \mathcal{P}(A_k) = \left(\frac{1}{m}\right)^N = p.$$

Defina $B_k = A_1^c \cap A_2^c \cap \ldots \cap A_k^c$.

Evento en el que el mono falló los primeros $k$ bloques.

$\Rightarrow B_k \supseteq B_{k+1} \Rightarrow B_k$ es una sucesión decreciente.

De la proposición anterior tenemos que

$$\mathcal{P}\left(\bigcap_{k=1}^{\infty} B_k\right) = \lim_{k\to\infty} \mathcal{P}(B_k) = \lim_{k\to\infty} (1-p)^k = 0.$$

Esto es la probabilidad de que el mono falle siempre.

Por lo tanto la probabilidad de que escriba el resultado planteado es $= 1$.

**Proposición 17**:

Continuidad en probabilidad.

Sea $\{A_n | n \in \mathbb{N}\}$ una sucesión de eventos convergente al evento $A$.

$$\Rightarrow \lim_{n\to\infty} \mathcal{P}(A_n) = \mathcal{P}(A).$$

Tomando en cuenta la convergencia de eventos.

Sea $\{A_n | n \in \mathbb{N}\} \ \exists \ A$ tal que

$$\lim_{n\to\infty} \inf A_n = \lim_{n\to\infty} \sup A_n = A.$$

La sucesión tiene como límite el evento $A$

$$\lim_{n\to\infty} A_n = A.$$

Demostración:

Sea $\{A_n | n \in \mathbb{N}\}$ una sucesión de eventos, sabemos que

$$A_n \subseteq \bigcup_{k=n}^{\infty} A_k.$$

Entonces

$$\mathcal{P}(A_n) \leq \mathcal{P}\left(\bigcup_{k=n}^{\infty} A_k\right).$$

Donde

$$\left\{\bigcup_{k=n}^{\infty} A_k \middle| \forall\ n \in \mathbb{N}\right\}.$$

Es decreciente, considerando el límite superior.

$$\lim_{n\to\infty} \sup \mathcal{P}(A_n) \leq \lim_{n\to\infty} \sup \mathcal{P}\left(\bigcup_{k=n}^{\infty} A_k\right) = \lim_{n\to\infty} \mathcal{P}\left(\bigcup_{k=n}^{\infty} A_k\right)$$

$$= \mathcal{P}\left(\lim_{k\to\infty} \bigcup_{k=n}^{\infty} A_k\right) = \mathcal{P}\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k\right)$$

Entonces

$$\lim_{n\to\infty} \sup \mathcal{P}(A_n) \leq \mathcal{P}\left(\lim_{n\to\infty} \sup A_n\right). \qquad \ldots 1$$

Por otro lado

$$\bigcap_{k=n}^{\infty} A_k \subseteq A_n \Rightarrow \mathcal{P}\left(\bigcap_{k=n}^{\infty} A_k\right) \leq \mathcal{P}(A_n).$$

Donde

$$\left\{\bigcap_{k=n}^{\infty} A_k \middle| n \in \mathbb{N}\right\}.$$

Es una sucesión creciente.

Entonces considerando el límite inferior.

$$\lim_{n\to\infty} \inf \mathcal{P}(A_n) \geq \lim_{n\to\infty} \inf \mathcal{P}\left(\bigcap_{k=n}^{\infty} A_k\right).$$

$$= \lim_{n\to\infty} \mathcal{P}\left(\bigcap_{k=n}^{\infty} A_k\right) = \mathcal{P}\left(\lim_{n\to\infty} \bigcap_{k=n}^{\infty} A_k\right) = \mathcal{P}\left(\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k\right).$$

Entonces

$$\lim_{n\to\infty} \inf \mathcal{P}(A_n) \geq \mathcal{P}\left(\lim_{n\to\infty} \inf A_n\right). \qquad \ldots 2$$

Con base en los dos resultados anteriores conocemos a la sucesión de eventos.

$\{A_n | n \in \mathbb{N}\}$ convergente al evento $A$. Entonces

$$\lim_{n\to\infty} \sup A_n = \lim_{n\to\infty} \inf A_n = A.$$

De los resultados marcados como $\ldots 1$ y $\ldots 2$ obtenemos lo siguiente:

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

$$\lim_{n\to\infty} \sup \mathcal{P}(A_n) \leq \mathcal{P}(A).$$

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

Entonces

$$\lim_{n\to\infty} \inf \mathcal{P}(A_n) \geq \mathcal{P}(A).$$

$$\lim_{n\to\infty} \sup \mathcal{P}(A_n) \leq \mathcal{P}(A) \leq \lim_{n\to\infty} \mathcal{P}(A_n).$$

Entonces

$$\lim_{n\to\infty} \mathcal{P}(A_n) = \mathcal{P}(A).$$

∎

Ejemplo:

Se lanza un dado equilibrado, indefinidamente.

$A_n =$ Evento de obtener $\{2, 4, 6\}$ en cada uno de los primeros $n$ lanzamientos que se realicen.

$$A_n \supseteq A_{n+1} \quad \forall\ n \in \mathbb{N}.$$

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

Resultado en el que siempre se obtiene un número por una sucesión infinita de lanzamientos.

$$\lim_{n\to\infty} A_n = \lim_{n\to\infty} \left(\bigcap_{n=1}^{\infty} A_n\right) \Rightarrow \mathcal{P}\left(\lim_{n\to\infty} \left(\bigcap_{n=1}^{\infty} A_n\right)\right) = \mathcal{P}\left(\lim_{n\to\infty} A_n\right)$$

$$= \lim_{n\to\infty} \mathcal{P}(A_n) = \lim_{n\to\infty} \left(\frac{1}{2}\right)^n = 0$$

Esto es la probabilidad de que siempre salga par, por lo tanto $A_n$, es imposible

Si $A = \{1, 2, 3, 4, 5\}$.

$$\mathcal{P}\left(\lim_{n\to\infty} \left(\bigcap_{n=1}^{\infty} A_n\right)\right) = \lim_{n\to\infty} \left(\frac{5}{6}\right)^n = 0.$$

Si $A = \{1, 2, 3, 4, 5, 6\}$.

$$= \lim_{n\to\infty} \mathcal{P}(A_n) = 1.$$

**1.4 Independencia de eventos.**

**Definición**:

Sea $(\Omega, \mathcal{F}, \mathcal{P})$ un espacio de probabilidad, la probabilidad conjunta de dos eventos $A \wedge B \in \mathcal{F}$ es igual a $\mathcal{P}\,(A \cap B)$.

**Definición**:

Sea $(\Omega, \mathcal{F}, \mathcal{P})$ un espacio de probabilidad, y sean $A \wedge B \in \mathcal{F}$ con $\mathcal{P}\,(B) > 0 \Rightarrow$ $\mathcal{P}\,(A|B) = \frac{\mathcal{P}\,(A\cap B)}{\mathcal{P}\,(B)}$ (condicional) Para cada $A \in \mathcal{F}$.

**Proposición 18**:

Sea $(\Omega, \mathcal{F}, \mathcal{P})$ un espacio de probabilidad, y sean $A \wedge B \in \mathcal{F}$ si $A \wedge B$ son independientes.

$$\Rightarrow \mathcal{P}\,(A \cap B) = \mathcal{P}\,(A)\,\mathcal{P}\,(B).$$

Demostración:

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

Sabemos que si $\mathrm{A}$ y $\mathrm{B}$ son independientes.

Entonces $\mathcal{P}\ (\mathrm{A}|\mathrm{B}) = \mathcal{P}\ (\mathrm{A})$ y $\mathcal{P}\ (\mathrm{B}|\mathrm{A}) = \mathcal{P}\ (\mathrm{B})$.

Entonces $\mathcal{P}\ (\mathrm{A} \cap \mathrm{B}) = \mathcal{P}\ (\mathrm{A})\ \mathcal{P}\ (\mathrm{B})$.

∎

**Definición**:

Dos σ-algebra $\mathcal{F}_1$ y $\mathcal{F}_2$ son independientes si para cada $\mathrm{A}$ en $\mathcal{F}_1$ y cada $\mathrm{B}$ en $\mathcal{F}_2$ se cumple que:

$$\mathcal{P}\ (\mathrm{A}|\mathrm{B}) = \mathcal{P}\ (\mathrm{A}) \quad \wedge \ \mathcal{P}\ (\mathrm{B}|\mathrm{A}) = \mathcal{P}\ (\mathrm{B}).$$

$$\Rightarrow \mathcal{P}\ (\mathrm{A} \cap \mathrm{B}) = \mathcal{P}\ (\mathrm{A})\ \mathcal{P}\ (\mathrm{B}).$$

Teorema de probabilidad total:

Sea $(\Omega,\ \mathcal{F},\ \mathcal{P})$ un espacio de probabilidad, y sea $\{A_1, A_2, \ldots\}$ una partición de $\Omega$ tal que $\forall\ \mathrm{n} \geq 1$ el conjunto $\mathrm{A_n}$ es un elemento de $\mathcal{F}$ con $\mathcal{P}\ (\mathrm{A_n}) \geq 0$.

$$\mathcal{P}\ (\mathrm{B}) = \sum \mathcal{P}\ (\mathrm{B}|\mathrm{A_n})\ \mathcal{P}\ (\mathrm{A_n}).$$

Demostraremos que es para todo evento $\mathrm{B} \in \mathcal{F}$.

Demostración:

$$\bigcup_{\mathrm{n}=1}^{\infty} \mathrm{A_n} = \Omega \qquad \mathrm{A}_i \cap \mathrm{A}_j = \emptyset \qquad i \neq j, \quad i,j \geq 1.$$

Sabemos que

$$\mathrm{B} = \mathrm{B} \cap \Omega = \mathrm{B} \cap \left( \bigcup_{n=1}^{\infty} \mathrm{A_n} \right) = \bigcup_{n=1}^{\infty} (\mathrm{B} \cap \mathrm{A_n}).$$

Entonces

$$\mathcal{P}(\mathrm{B}) = \mathcal{P}\left( \bigcup_{n=1}^{\infty} (\mathrm{B} \cap \mathrm{A_n}) \right) = \sum_{n=1}^{\infty} \mathcal{P}(\mathrm{B} \cap \mathrm{A_n}) = \sum_{n=1}^{\infty} \mathcal{P}(\mathrm{A_n})\, \mathcal{P}(\mathrm{B}|\mathrm{A_n}).$$

∎

**Teorema de Bayes:**

Sea $(\Omega, \mathcal{F}, \mathcal{P})$ un espacio de probabilidad, y sea $\{A_1, A_2, \ldots\}$ una partición de $\Omega$ tal que $\forall\ n \geq 1$ el conjunto $\mathrm{A_n}$ es un elemento de $\mathcal{F}$ con $\mathcal{P}(\mathrm{A_n}) > 0$.

Demostrando que para cualquier evento $\mathrm{B} \in \mathcal{F}$ tal que $\mathcal{P}(\mathrm{B}) > 0$ y para cualquier $\mathrm{m} \geq 1$ fijo.

$$\mathcal{P}(\mathrm{A_m}|\mathrm{B}) = \frac{\mathcal{P}(\mathrm{B}|\mathrm{A_m})\, \mathcal{P}(\mathrm{A_m})}{\sum_{n=1}^{\infty} \mathcal{P}(\mathrm{B}|\mathrm{A_n})\, \mathcal{P}(\mathrm{A_n})}.$$

Demostración:

Sabemos que

$$\mathcal{P}(\mathrm{A_m}|\mathrm{B}) = \frac{\mathcal{P}(\mathrm{A_n} \cap \mathrm{B})}{\mathcal{P}(\mathrm{B})} = \frac{\mathcal{P}(\mathrm{B}|\mathrm{A_m})\, \mathcal{P}(\mathrm{A_m})}{\mathcal{P}(\mathrm{B})}.$$

Y por el teorema de probabilidad total.

$$\mathcal{P}(\mathrm{B}) = \sum_{n=1}^{\infty} \mathcal{P}(\mathrm{A_n})\, \mathcal{P}(\mathrm{B}|\mathrm{A_n}).$$

Sustituyendo obtenemos

$$\mathcal{P}(\mathrm{A_m}|\mathrm{B}) = \frac{\mathcal{P}(\mathrm{B}|\mathrm{A_m})\, \mathcal{P}(\mathrm{A_m})}{\sum_{n=1}^{\infty} \mathcal{P}(\mathrm{B}|\mathrm{A_n})\, \mathcal{P}(\mathrm{A_n})}.$$

∎

Ejemplo:

Una compañía de transporte público tiene 3 líneas en la zona metropolitana, el 60% de autobuses cubre el servicio de la línea 1, el 30% está en la línea 2 y el 10% sobre la línea 3, se conoce que la probabilidad de que diariamente un autobús se descomponga es del 2%, 4% y 1% respectivamente para cada línea, defina la probabilidad de que en un día un autobús se descomponga.

Por el teorema de Probabilidad Total:

$$\mathcal{P}\left(\mathrm{B}\right) = \sum \mathcal{P}\left(\mathrm{B}\middle|\mathrm{A_n}\right)\mathcal{P}\left(\mathrm{A_n}\right).$$

Tenemos que

| | $\mathcal{P}\left(\mathrm{L}_i\right)$ | $\mathcal{P}\left(\mathrm{Desc}\middle|\mathrm{L}_i\right)$ |
|---|---|---|
| Línea 1 $(\mathrm{L}_1)$ | 60% | 2% |
| Línea 2 $(\mathrm{L}_2)$ | 30% | 4% |
| Línea 3 $(\mathrm{L}_3)$ | 10% | 1% |

$$\mathcal{P}\left(\mathrm{Desc}\right) = \sum \mathcal{P}\left(\mathrm{Desc}\middle|\mathrm{L}_i\right)\mathcal{P}\left(\mathrm{L}_i\right).$$

$$\mathcal{P}\left(\mathrm{Desc}\right) = \mathcal{P}\left(\mathrm{Desc}\middle|\mathrm{L}_1\right)\mathcal{P}\left(\mathrm{L}_1\right) + \mathcal{P}\left(\mathrm{Desc}\middle|\mathrm{L}_2\right)\mathcal{P}\left(\mathrm{L}_2\right) + \mathcal{P}\left(\mathrm{Desc}\middle|\mathrm{L}_3\right)\mathcal{P}\left(\mathrm{L}_3\right).$$

$$\mathcal{P}\left(\mathrm{Desc}\right) = 0.6\left(0.2\right) + 0.3\left(0.4\right) + 0.1\left(0.1\right) = 0.25$$

La probabilidad de que en un día un autobús se descomponga es del $25\%$.

Ejemplo:

En una casa de bolsa el 20% de las inversiones se realiza vía internet. De los inversionistas que realizan operaciones vía internet. 80% consulta la pagina de *infobolsa.com* de los que no realizan operaciones vía internet solo un 20% consulta *infobolsa.com.*

a) Obtener la probabilidad de que un inversionista al azar consulte *infobolsa.com.*

   20% Opera internet y de los cuales únicamente el 80% consulta la página web.
   80% No opera internet y de los cuales únicamente el 20% consulta la página web.

$$\mathcal{P}\,(\text{Cons}) = \mathcal{P}\,(\text{Web}|\text{Int})\mathcal{P}\,(\text{Int}) + \mathcal{P}\,(\text{Web}|\text{NoInt})\mathcal{P}\,(\text{NoInt}).$$

$$\mathcal{P}\,(\text{Cons}) = 0.2(0.8) + 0.8(0.2) = 0.32$$

b) Si se elige al azar y consulta *infobolsa.com* ¿cuál es la probabilidad de que realice operaciones vía internet?

$$\mathcal{P}(\text{Int}|\text{Cons}) = \frac{\mathcal{P}\,(\text{Int} \cap \text{Web})}{\mathcal{P}\,(\text{Cons})} = \frac{\mathcal{P}(\text{Web}|\text{Int})\,\mathcal{P}\,(\text{Int})}{\mathcal{P}\,(\text{Cons})} = \frac{0.16}{0.32} = 0.5$$

**Lema Borel-Cantelli:**

Lema de Borel-Cantelli (resultado muy importante relacionado con la ley fuerte de los grandes números).

Sea $\{A_n |\, n \in \mathbb{N}\}$ una sucesión de eventos y defina

$$A = \lim_{n\to\infty} \sup A_n.$$

1) Si

$$\sum_{n=1}^{\infty} \mathcal{P}(A_n) < \infty \quad \Rightarrow \mathcal{P}\,(A) = 0.$$

Demostración:

Para cada $n \in \mathbb{N}$ sabemos que

$$\sum_{n=1}^{\infty} \mathcal{P}(A_n) < \infty \quad \Rightarrow \quad \lim_{n\to\infty} \sum_{k=n}^{\infty} \mathcal{P}\,(A_k).$$

$$A \subset \bigcup_{k=n}^{\infty} A_k \quad \forall\ n \in \mathbb{N} \quad \Rightarrow \mathcal{P}\,(A) \leq \mathcal{P}\left(\bigcup_{k=n}^{\infty} A_k\right) \leq \sum_{n=1}^{\infty} \mathcal{P}\,(A_n) \quad \forall\ n \in \mathbb{N}.$$

Entonces

$$\mathcal{P}\,(A) \leq \sum_{k=n}^{\infty} \mathcal{P}\,(A_k) = 0 \quad \Rightarrow \quad \mathcal{P}\,(A) = 0.$$

∎

Ejemplo:

Se lanza $\mathrm{n}$ –veces una moneda, hasta el momento en que caiga la cara que tiene el águila.

$$\sum_{n=1}^{\infty} A_n \leq \infty \quad \Rightarrow \quad A_n = \left(\frac{1}{2}\right)^n.$$

$$\sum_{n=1}^{\infty} \left(\frac{1}{2}\right)^n = 1 < \infty \quad \Rightarrow \quad \lim_{n\to\infty} \left(\frac{1}{2}\right)^n = 0.$$

2) Si $A_1,\ A_2, \ldots$ son independientes y

$$\sum_{n=1}^{\infty} \mathcal{P}(A_n) = \infty \quad \Rightarrow \mathcal{P}\ (A) = 1.$$

Demostración:

$$\mathcal{P}\ (A) = \lim_{n\to\infty} \mathcal{P}\left(\bigcup_{k=n}^{\infty} A_k\right) = 1 - \lim_{n\to\infty} \mathcal{P}\left(\bigcap_{k=n}^{\infty} A_k^c\right).$$

P.D. que

$$\lim_{n\to\infty} \mathcal{P}\left(\bigcap_{k=n}^{\infty} A_k^c\right) = 0.$$

$$\mathcal{P}\left(\bigcap_{k=n}^{\infty} A_k^c\right) = \lim_{m\to\infty} \mathcal{P}\left(\bigcap_{j=n}^{m} A_j^c\right) = \lim_{m\to\infty} \prod_{j=n}^{m} \mathcal{P}\ (A_j^c).$$

Como $A_i$ son independientes entonces $A_i^c$ también lo son

Por otro lado

$$\lim_{m\to\infty} \prod_{j=n}^{m} \mathcal{P}\left(A_j^c\right) = \lim_{m\to\infty} \prod_{j=n}^{m} \left[1 - \mathcal{P}\left(A_j\right)\right].$$

Sabemos que $1 - p \leq e^{-p}$

Así que

$$\lim_{m\to\infty} \prod_{j=n}^{m} \left[1 - \mathcal{P}\left(A_j\right)\right] \leq \lim_{m\to\infty} \prod_{j=n}^{m} e^{\left[-\mathcal{P}\left(A_j\right)\right]} .$$

$$\lim_{m\to\infty} \prod_{j=n}^{m} e^{\left[-\mathcal{P}\left(A_j\right)\right]} = \lim_{m\to\infty} e^{\left[-\sum \mathcal{P}\left(A_j\right)\right]}.$$

Como $-\sum \mathcal{P}\left(A_j\right) = 0.$

$$\lim_{m\to\infty} e^{\left[-\sum \mathcal{P}\left(A_j\right)\right]} = 0.$$

$$\Rightarrow \mathcal{P}\left(A\right) = 1 - 0 = 1.$$

∎

Variable aleatoria es el resultado que se obtiene de una muestra aleatoria.

## Capítulo 2 VARIABLES ALEATORIAS.

### 2.1 Variables Aleatorias como Función en el Espacio Muestral y Propiedades.

**Definición**:

Sea un espacio de probabilidad $(\Omega,\ \mathcal{F},\ \mathcal{P})$ una variable aleatoria real es una función.

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

Tal que para cualquier conjunto Boreliano $\mathfrak{B}$, se cumple ue el conjunto $\mathrm{X}^{-1}\mathfrak{B}$ es un elemento de $\mathcal{F}$.

Ejemplo:

Se lanzan 2 monedas,

Espacio muestral del espacio aleatorio.

$$\Omega = \{(a,a),(a,s),(s,s),(s,a)\}.$$

$\mathrm{X}:$ Número de águilas a obtener.

$$\Omega_{\mathrm{x}} = \{x|x = 0,1,2\}.$$

Entonces la variable aleatoria $\mathrm{X}$ es una función de $\Omega$ en $\mathbb{R}$ tal que $\mathrm{X}^{-1}\mathfrak{B}$ la imagen inversa de cualquier conjunto Boreliano es un elemento de $\mathcal{F}$**.**

Esta condición se conoce como Medibilidad en la Teoría de la Medida. $\Rightarrow$ $\mathcal{P}$ es una función medible respecto de las σ-álgebra $\mathcal{F}$ y $\mathfrak{B}(\mathbb{R})$.

Entonces $\mathcal{P}$ es una medida de probabilidad sobre el espacio medible $(\Omega,\ \mathcal{F})$.

Nota: Si $\mathrm{X}$ es una variable aleatoria, el espacio medible $(\Omega,\ \mathcal{F})$ se puede trasladar al espacio medible $(\mathbb{R},\mathfrak{B}(\mathbb{R}))$.

Entonces $\mathcal{P}_{\mathrm{X}}(\mathrm{B}) = \mathcal{P}(\mathrm{X}^{-1}\mathrm{B})$.

$\mathrm{X}^{-1}\mathrm{B}\ \in\ \mathcal{F}$ (Dominio para la función para $\mathcal{P}$).

Entonces a la función $\mathcal{P}_X : \mathfrak{B}(\mathbb{R}) \to [0, 1]$ se le llama "medida de probabilidad inducida" por $X$.

Se dice que construye $(\mathbb{R}, \mathfrak{B}(\mathbb{R}), \mathcal{P}_X)$.

De ahí $X^{-1}B = \{\omega \in \Omega \mid X(\omega) \in B\}$.

**Proposición 19**:

Una función $X : \Omega \to \mathbb{R}$ es una variable aleatoria $\Leftrightarrow$ el conjunto $X^{-1}(-\infty, x]$ es un elemento de $\mathcal{F}$ para cada $x \in \mathbb{R}$.

Demostración:

$\Rightarrow$ Si $X$ es variable aleatoria $\Rightarrow$ $\forall\ x \in \mathbb{R}$ el conjunto $X^{-1}(-\infty, x] \in \mathcal{F}$ por definición.

$\Leftarrow$ Supongamos que para cada real $x$, el conjunto $X^{-1}(-\infty, x]$ es un elemento de $\mathcal{F}$.

Sean:

$$\beta = \{B \in \mathfrak{B}(\mathbb{R}) \mid X^{-1}B \in \mathcal{F}\}.$$
$$C = \{(-\infty, x] \mid x \in \mathbb{R}\}.$$

$$C \subseteq \beta \subseteq \mathfrak{B}(\mathbb{R}).$$

$$\Rightarrow \sigma(C) \subseteq \sigma(\beta) \subseteq \mathfrak{B}(\mathbb{R}).$$

P.D. que $\beta$ es una σ-álgebra.

i) $\mathbb{R} \in \beta$ pues $\mathbb{R} \in \beta(\mathbb{R}) \wedge X^{-1}\mathbb{R} = \Omega \in \mathbb{R}$.
ii) Sea $B \in \beta \Rightarrow B \in \beta(\mathbb{R}) \wedge X^{-1}B \in \mathcal{F}$.
Entonces $B^c \in \beta(\mathbb{R}) \wedge X^{-1}B^c = (X^{-1}B)^c \in \mathcal{F}$, es decir, $B^c \in \beta$.
iii) Sean $B_1, B_2, \ldots$ una sucesión en $\beta$ es decir que para cada $n \in \mathbb{N}$.

$B_n \in \beta(\mathbb{R})$.
Entonces

$$\bigcup_{n=1}^{\infty} B_n \in \beta(\mathbb{R}) \wedge \bigcup_{n=1}^{\infty} X^{-1}B_n = X^{-1}\bigcup_{n=1}^{\infty} B_n \in \mathcal{F}.$$

Entonces

$$\bigcup_{n=1}^{\infty} B_n \in \beta.$$

∎

**Proposición 20**:

La función constante $X = c$ es una variable aleatoria.

Nota: Los eventos determinísticos los podemos interpretar como un subconjunto de los elementos aleatorios, donde la probabilidad de esos eventos es igual a 1 y la probabilidad de su complemento es igual a 0.

Demostración:

Sea $B$ un elemento cualquiera de $\mathfrak{B}(\mathbb{R})$ para la función constante $X = c$ se tiene que:

$X^{-1}(B) = \Omega$ si $c \in B$ $y$ $X^{-1}(B) = \emptyset \notin B.$

$$\Rightarrow \quad X^{-1}(B) \in \mathcal{F}.$$

Por lo tanto $X$ es una variable aleatoria.

∎

**Proposición 21**:

Si $X$ es variable aleatoria y $c$ una constante $cX$ es una variable aleatoria.

Demostración:

$\forall \ x \in \mathbb{R}$ P.D.

$$c(X)^{-1} \ (-\infty, \ x] \in \mathcal{F}.$$

3 casos

Si:

$X > 0 \Rightarrow (cX \leq x) = \left(X \leq \frac{x}{c}\right) \in \mathcal{F}.$

$X < 0 \Rightarrow (cX \leq x) = \left(X \geq \frac{x}{c}\right) \in \mathcal{F}.$

$X = 0 \Rightarrow$ por la proposición anterior es una variable aleatoria.

∎

**Proposición 22**:

Si $\mathrm{X}$ y $\mathrm{Y}$ son variables aleatorias entonces $(\mathrm{X}+\mathrm{Y})$ es también variable aleatoria.

Demostración:

$$(\mathrm{X}+\mathrm{Y})^{-1}\ (x,\ \infty) = [(\mathrm{X}+\mathrm{Y}) > x] \in \mathcal{F} \qquad \forall\, x \in \mathbb{R}.$$

Supongamos que $[(\mathrm{X}+\mathrm{Y}) > x]$ lo podemos expresar como:

$$(\mathrm{X}+\mathrm{Y}) > x = \bigcup_{\mathrm{r}\in\mathbb{Q}} (\mathrm{X} > \mathrm{r}) \cap (\mathrm{Y} > x-\mathrm{r})\,.$$

$(\subseteq)$ Sea $\omega \in \Omega$ tal que $\mathrm{X}(\omega)+\mathrm{Y}(\omega) > x \Rightarrow \mathrm{X}(\omega) > \mathrm{X}-\mathrm{Y}(\omega)$
$\Rightarrow \exists\ \mathrm{r} \in \mathbb{Q}$ tal que $\mathrm{X}(\omega) > r > X - Y(\omega)$. Por lo tanto

$$\mathrm{X}(\omega) > r \qquad X - Y(\omega) < \mathrm{r} \qquad \mathrm{Y}(\omega) < \mathrm{X}-\mathrm{r}.$$

Entonces

$$\omega \in \bigcup_{\mathrm{r}\in\mathbb{Q}} (\mathrm{X} > \mathrm{r}) \cap (\mathrm{Y} > x-\mathrm{r}).$$

$(\supseteq)$ Sea $\omega$ un elemento

$$\bigcup_{\mathrm{r}\in\mathbb{Q}} (\mathrm{X} > \mathrm{r}) \cap (\mathrm{Y} > x-\mathrm{r}).$$

Entonces existe un $\mathrm{r}_0 \in \mathbb{Q}$ tal que $\mathrm{X}(\omega) > \mathrm{r}_0 \wedge \mathrm{Y}(\omega) > x-\mathrm{r}_0$.

Entonces $\mathrm{X}(\omega)+\mathrm{Y}(\omega) > \mathrm{r}_0 + x - \mathrm{r}_0 = x \in \Omega$.

Por lo tanto $\mathrm{X}\ y\ \mathrm{Y}$ son variables aleatorias y las operaciones de una unión involucrada en la igualdad es numerable se cumple que $(\mathrm{X}+\mathrm{Y})$ es variable aleatoria.

∎

**Proposición 23**:

Si $\mathrm{X}\ y\ \mathrm{Y}$ son variables aleatorias entonces $\mathrm{XY}$ es una variable aleatoria.

Demostración:

Si $\mathrm{X} = \mathrm{Y}$.

P.D. $\{\mathrm{X}^2 \le x\} \in \mathcal{F}$.

$$\mathrm{X} < 0 \qquad \{\mathrm{X}^2 \le x\} = \emptyset \in \mathcal{F}.$$

$$\mathrm{X} \geq 0 \quad \{\mathrm{X}^2 \leq x\} = \{-\sqrt{\mathrm{X}} \leq x \leq \sqrt{\mathrm{X}}\}.$$

$$\{-\sqrt{\mathrm{X}} \leq x\} \cup \{x \leq \sqrt{\mathrm{X}}\} \in \mathcal{F}.$$

Entonces $(\mathrm{X}^2)^{-1}(-\infty, x] \in \mathcal{F}$ por lo tanto $\mathrm{X}^2$ es una variable aleatoria.

Si $\mathrm{X} \neq \mathrm{Y}$.

$$\mathrm{XY} = \frac{(\mathrm{X}+\mathrm{Y})^2 - (\mathrm{X}-\mathrm{Y})^2}{4}.$$

“Formula general de interpolación”.

Por las proposiciones anteriores.

$(\mathrm{XY})^{-1}(-\infty, x] \in \mathcal{F} \Rightarrow \mathrm{XY}$ es variable aleatoria.

Nota:

Generalizando.

$\mathrm{C}$ si $\mathrm{X}$ es variable aleatoria entonces $\mathrm{X}^{\mathrm{n}}$ también es variable aleatoria.

$\mathrm{C}$ si $\mathrm{X}$ es variable aleatoria entonces cualquier función polinomial de $\mathrm{X}$ es variable aleatoria.

∎

**Proposición 24**:

Sean $\mathrm{X}\ y\ \mathrm{Y}$ variables aleatorias con $\mathrm{Y} \neq 0$ entonces $\frac{\mathrm{X}}{\mathrm{Y}}$ es variable aleatoria.

Demostración:

P.D. $\frac{1}{\mathrm{Y}}$ es variable aleatoria para cualquier $y \in \mathbb{R}$.

$\mathrm{Y} > 0$.

$$\left(\frac{1}{\mathrm{Y}} \leq y\right) = \left(\frac{1}{\mathrm{Y}} \leq y \middle| \mathrm{Y} > 0\right) \cup \left(\frac{1}{\mathrm{Y}} \leq y \middle| \mathrm{Y} < 0\right)$$
$$= \left(\mathrm{Y} \leq \frac{1}{y} \middle| \mathrm{Y} > 0\right) \cup \left(\mathrm{Y} \leq \frac{1}{y} \middle| \mathrm{Y} < 0\right)$$
$$= (\mathrm{Y} \geq y) \cup (\mathrm{Y} < 0) \in \mathcal{F}.$$

Para $y \in \mathbb{R},\ y < 0$.

$$\left(\frac{1}{\mathrm{Y}} \leq y\right) = \left(\frac{1}{\mathrm{Y}} \leq y \middle| \mathrm{Y} > 0\right) \cup \left(\frac{1}{\mathrm{Y}} \leq y \middle| \mathrm{Y} < 0\right).$$

$$\left(\frac{1}{\mathrm{Y}} \leq y\right) = \emptyset \ \cup \ \left(\mathrm{Y} = \frac{1}{y} \middle| \mathrm{Y} < 0\right) \in \mathcal{F}.$$

$$\left(\frac{1}{\mathrm{Y}} \leq y\right) = (\mathrm{Y} < 0) \in \mathcal{F}.$$

Por lo tanto $\frac{1}{\mathrm{Y}}$ es variable aleatoria.

Y por la proposición anterior $\mathrm{X}\left(\frac{1}{\mathrm{Y}}\right)$ es variable aleatoria también.

∎

**Proposición 25**:

Si $\mathrm{X}\ y\ \mathrm{Y}$ son variables aleatorias entonces el $\max\{\mathrm{X},\ \mathrm{Y}\}$, $\min\{\mathrm{X},\ \mathrm{Y}\}$ también lo son.

Demostración:

$\forall\ x \in \mathbb{R}$ sabemos que

$$\max\{\mathrm{X},\ \mathrm{Y} \,|\leq x\} = (\mathrm{X} \leq x, \mathrm{Y} \leq x)$$
$$(\mathrm{X} \leq x) \cap (\mathrm{Y} \leq x) \in \mathcal{F}.$$

$$\min\{\mathrm{X},\ \mathrm{Y} \,|\geq x\} = (\mathrm{X} \geq x, \mathrm{Y} \geq x)$$
$$(\mathrm{X} \geq x) \cap (\mathrm{Y} \geq x) \in \mathcal{F}.$$

$\mathrm{C}$ Sea $\mathrm{X}^{+} = \max\{0, x\} \ \wedge \ \mathrm{X}^{-} = \min\{0, x\}$ entonces $\mathrm{X}^{+} \wedge \mathrm{X}^{-}$ son variables aleatorias.

∎

**Proposición 26**:

Si $\mathrm{X}$ es variable aleatoria entonces $|\mathrm{X}|$ es variable aleatoria $\forall \ \omega \in \Omega$.

Demostración:

Si
$\mathrm{X} \geq 0 \Rightarrow |\mathrm{X}|^{-1}(-\infty, \ x] = \{\omega | -x \leq \mathrm{X}(\omega) \leq x\} \in \mathcal{F}$.

$\mathrm{X} < 0 \Rightarrow |\mathrm{X}|^{-1}(-\infty, \ x] = \emptyset \in \mathcal{F}$.

Notación $|\mathrm{X}| = \mathrm{X}^{+} + \mathrm{X}^{-} \ \wedge$ por el C anterior $|\mathrm{X}|$ es variable aleatoria.

$\mathrm{X} : \Omega \to \mathbb{R}$ es una función tal que $|\mathrm{X}|$ es variable aleatoria.

Entonces NO necesariamente $\mathrm{X}$ es variable aleatoria.

Sea $\Omega = \{-1, 0, 1\}$ con $\mathcal{F} = \{\emptyset, \ \{0\}, \ \{-1,1\}, \ \Omega\}$.

Entonces $|\mathrm{X}|$ es una variable aleatoria para cualquier Boreliano $\mathrm{B}$.

$$|\mathrm{X}|^{-1}\mathrm{B} \begin{cases} \emptyset \text{ si } 0 \wedge 1 \notin \mathrm{B}. \\ \Omega \text{ si } 0 \wedge 1 \in \mathrm{B}. \\ 0 \text{ si } 0 \in \mathfrak{B} \wedge 1 \notin \mathrm{B}. \\ (-1,1) \text{ si } 1 \in \mathfrak{B} \wedge 0 \notin \mathrm{B}. \end{cases}$$

Entonces $|\mathrm{X}|^{-1}\mathrm{B} \in \mathcal{F}$.

$$\mathrm{X}^{-1}(-1) = -1 \notin \mathcal{F}.$$

Por lo tanto $\mathrm{X}$ no es variable aleatoria.

∎

Tipos de variables aleatorias:

1) Discretas:
   $F_x(x)$ un número numerable de saltos.
   $P_x(x) = P\{X = x\}$.
   i) $P_x(x) \geq 0$ función masa de probabilidad.
   ii) $\sum P_x(x) = 1$.

2) Continuas:
   $F_x(x)$ es absolutamente continua.
   $f_x(x)$ Teorema fundamental de cálculo $\frac{\partial F_x}{\partial x} = f_x(x)$.

3) Mixtas:
   $f_x(x) = \begin{cases} 1 - e^{-x}, & \text{si } x < 0. \\ 0 \quad , & \text{si } x \geq 0. \end{cases}$

**Variables Aleatorias Discretas.**

**Definición**:

Sea $X$ una variable aleatoria, si su correspondiente función de distribución $F_x(x)$ es una función continua por intervalos donde en los valores (puntuales) de su recorrido $x_1, x_2, \ldots$ encontramos discontinuidades y estas tienen una magnitud de $P_x(x) = P\{X = x\}$ entonces X es una variable aleatoria discreta.

Magnitud de salto es:

$P\{X = x\} = F_x(x) - F_x(x^-)$.

La función $P_x(x)$ nos indica los incrementos en $F_x(x)$.

Entonces

$$F_x(x_0) = \sum_{x=0}^{x_0} P_x(x).$$

**Variables Aleatorias Continuas.**

**Definición**:

Sea $X$ variable aleatoria con función de distribución $F_x(x)$ continua entonces cuando existe una función integrable $f > 0$ tal que $\forall\ x \in \mathbb{R}$.

$$F_x(x) = \int_{-\infty}^{x} f_x(x)\, dx.$$

Se dice que la variable aleatoria $\mathrm{X}$ es absolutamente continua y a $f_{\mathrm{X}}(x)$ se le llama función de densidad de probabilidad.

**Teorema Fundamental del Calculo.**

Sea $f$ integrable sobre $[a,b]$ y defínase $\mathrm{F}$ sobre $[a,b]$ por:

$$\mathrm{F}(x) = \int_a^x f.$$

Ejemplo:

$$\mathrm{F}_{\mathrm{X}}(x) = \begin{cases} 1 - e^{-x}, & \text{si } x > 0. \\ 0 \quad , & \text{si } x \leq 0. \end{cases}$$

$$f_{\mathrm{X}}(x) = \begin{cases} e^{-x}, & \text{si } x > 0. \\ 0 \ , & \text{en otro caso.} \end{cases}$$

i) $f_{\mathrm{X}}(x) \geq 0.$

ii)
$$\int_0^\infty e^{-x}\,dx = 1 \qquad (\mathrm{F}(1) = 1).$$

$$\mathrm{F}_{\mathrm{Y}}(y) = \begin{cases} 0 \quad , & \text{si } \mathrm{y} < 0. \\ 1 - e^{-y}, & \text{si } 0 \leq \mathrm{y} < M. \\ 1 \quad , & \text{si } \mathrm{y} \geq \mathrm{M}. \end{cases}$$

$$f_{\mathrm{Y}}(y) = \begin{cases} e^{-y}, & \text{si } 0 \leq y < M. \\ 0 \ , & \text{en otro caso.} \end{cases}$$

i) $f_{\mathrm{Y}}(y) \geq 0.$

ii)
$$\int_0^\infty e^{-y}\,d\mathrm{y} = -e^{-y}\big|_0^{\mathrm{M}} = -e^{-\mathrm{M}} + 1.$$

Por lo tanto no es continua.

Utilizando la propiedad de la integral Rieman-Stieljes.

Tenemos

$$\int_{a}^{b} h(x)\, dF(x).$$

Siendo $h(x)$ una función acotada $[a, b]$.

$F(x)$ debe ser:

- Monótona no decreciente.
- Continua por la derecha.
- $F(\infty) - F(-\infty) < M \quad \forall\ M > 0$.

Para variables aleatorias cuando $F_x(x)$ es diferenciable. Entonces

$$\int_{a}^{b} h(x)\, dF(x) = \int_{a}^{b} h(x)\, F'(x)\, dx$$

Cuando $F_x(x)$ es continua, excepto en los puntos $x_1, x_2, \ldots$ en donde $P_x(x_1),\ P_x(x_2), \ldots$ nos representan la magnitud de la discontinuidad en los mismos.

Si converge entonces

$$\int_{a}^{b} h(x)\, dF_x(x) = \sum_{i=0}^{\infty} h(x_i)\, P_x(x_i)\,.$$

### 2.1.1 Valor Esperado (Esperanza).

**Definición**:

Es el promedio ponderado de los valores que puede tomar la variable aleatoria con respecto a una probabilidad de ocurrencia.

**Definición**:

Sea $X$ variable aleatoria con función de distribución $F_x(x)$ y sea $g : \mathbb{R} \to \mathbb{R}$ una función Borel medible. El valor esperado (esperanza) de $g(x)$ se define como:

$$\mathrm{E}[\mathrm{g}(x)]\begin{cases}\displaystyle\sum_{i=0}^{\infty}\mathrm{g}(x_i)\ \mathrm{P_x}(x_i) & \text{variables aleatorias discretas.}\\ \displaystyle\int_{-\infty}^{\infty}\mathrm{g}(x_i)\ \mathrm{F'_x}(x_i)\,dx & \text{variables aleatorias continuas.}\end{cases}$$

Nota: La distribución de Cauchy no tiene valor esperado pues su integral no es absolutamente convergente.

$$f_{\mathrm{x}}(x)=\begin{cases}\dfrac{1}{\pi\,\beta\left[1+\left(\frac{x-\alpha}{\beta}\right)^2\right]}, & \text{si } -\infty\leq x\leq\infty \mid (\alpha,\ \beta)\in\mathbb{R}^2.\\ 0, & \text{en otro caso.}\end{cases}$$

Estandarizando $\alpha = 0$ , $\beta = 0.$

$$f_{\mathrm{x}}(x)=\begin{cases}\dfrac{1}{\pi\,(1+\ x)^2}, & \text{si } -\infty\leq x\leq\infty.\\ 0, & \text{en otro caso.}\end{cases}$$

$\mathrm{E}(\cdot)$ es un operador lineal.

**Proposición 27**:

Sean $\mathrm{X}$ y $\mathrm{Y}$ variables aleatorias, una esperanza finita y sea $\mathrm{c}$ constante, entonces,

1) $\mathrm{E(c) = c}$.

Demostración:

Sea $\mathrm{g}(x) = \mathrm{c}$

Entonces

$$\mathrm{E}\big(\mathrm{g}(x)\big) = \int_{-\infty}^{\infty} \mathrm{g}(x)\, d\mathrm{F_x}(x) = \int_{-\infty}^{\infty} \mathrm{c}\, d\mathrm{F_x}(x) = c \int_{-\infty}^{\infty} d\mathrm{F_x}(x) = c.$$

Por propiedad de función de densidad.

$$\int_{-\infty}^{\infty} d\mathrm{F_x}(x) = 1.$$

∎

2) $\mathrm{E(cX) = c\,E(X)}$.

Demostración:

$$\mathrm{E}\big(\mathrm{g}(x)\big) = \int_{-\infty}^{\infty} \mathrm{c}x\, d\mathrm{F_x}(x) = \mathrm{c} \int_{-\infty}^{\infty} x\, d\mathrm{F_x}(x) = \mathrm{c\,E(X)}.$$

∎

3) Si $\mathrm{X} \geq 0 \Rightarrow \mathrm{E(X)} \geq 0$.

Demostración:

$$x \geq 0.$$

$f_{\mathrm{x}}(x) \geq 0.$

$$\Rightarrow\ xF_{\mathrm{x}}(x) \geq 0 \ \Rightarrow \int_{-\infty}^{\infty} xF_{\mathrm{x}}(x) \geq 0 \quad \Rightarrow\ \mathrm{E(X)} \geq 0.$$

∎

4) $X \leq Y \;\Rightarrow\; E[X] \leq E[Y].$

Demostración:

$X \leq Y \;\Rightarrow\; Y - X \geq 0$ por el inciso anterior $E[Y - X] \geq 0.$

Entonces $E[Y] - E[X] \geq 0 \;\Rightarrow\; E[X] \leq E[Y].$

∎

Nota:

$$E[X + Y] = E[X] + E[Y].$$

$$E[X + Y] = \iint (x + y) f_{xy}(x, y)\, dxdy = \iint x f_{x}(x, y) dxdy + \iint y f_{xy}(x, y)\, dxdy$$

$$= \int x \left[f_{xy}(x, y) dy\right] dx + \int y \left[f_{xy}(x, y) dx\right] dy$$

$$= \int x\, f_{x}(x) dx + \int y\, f_{y}(y) dy \;= E[X] + E[Y].$$

### 2.1.2 Varianza.

Grado de dispersión de los valores que puede tomar una variable aleatoria con respecto a su $E(X) = \mu.$

$$E(X - \mu) = E(X) - E(\mu) = \mu - \mu = 0.$$

$$Var(X) = E(X - \mu)^2.$$

La varianza de una variable aleatoria se define como el valor de:

$$Var(X) = E\big(X - E(X)\big)^2.$$

Y siempre positivo cuando existe.

**Proposición 28**:

La $\mathrm{Var}(X)$ cumple con lo siguiente:

1) $\mathrm{Var}(X) \geq 0$.

Demostración:

Sabemos que si $X \geq 0 \Rightarrow E(X) \geq 0$ .

Y

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

$$\mathrm{Var}(X) \geq 0.$$

∎

2) $\mathrm{Var}(c) = 0$.

Demostración:

$$\mathrm{Var}(c) = E(c - \mu)^2.$$

$$= E(c - c)^2 = 0.$$

∎

3) $\mathrm{Var}(cx) = c^2\, \mathrm{Var}(x)$.

Demostración:

$$\mathrm{Var}(cx) = E(cx - c\mu)^2 = E\big(c(x - \mu)\big)^2$$

$$= E[c^2 (x - \mu)^2]$$

$$= c^2\, \mathrm{Var}(x).$$

∎

4) $\mathrm{Var}(x+\mathrm{c}) = \mathrm{Var}(x)$.

Demostración:

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

$$= \mathrm{E}[\mathrm{X}+\mathrm{c}]^2 = \mathrm{Var}(\mathrm{X}).$$

∎

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

Demostración:

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

$$= \mathrm{E}(\mathrm{X}^2) - 2\mathrm{E}(\mathrm{X})\mathrm{E}(\mathrm{X}) + \mathrm{E}(\mathrm{X})^2 = \mathrm{E}(\mathrm{X}^2) - \mathrm{E}(\mathrm{X})^2.$$

∎

### 2.1.3 Momentos.

Determinan de manera única a $\mathrm{F_x}(x)$ cuando existen.

**Definición**:

Sea $\mathrm{X}$ una variable aleatoria con $\mathrm{E}(\mathrm{X}) = \mu$ y sea $\mathrm{n} \in \mathbb{N}$ entonces cuando existen.

1) $\mathrm{E}(\mathrm{X}^\mathrm{n}) \rightarrow \mathrm{n} -$ esimo momento no central ó con respecto al origen.

2) $\mathrm{E}(\mathrm{X}-\mu)^\mathrm{n} \rightarrow \mathrm{n} -$esimo momento central ó con respecto a la media $\mu$.

Nota: Si $\mathrm{X}$ es variable aleatoria discreta tal que $\mathrm{E}(\mathrm{X}),\ \mathrm{E}(\mathrm{X}^2), \dots$ existen y son finitos, si se cumple que la serie.

$$\sum_{\mathrm{n}=0}^{\infty} \frac{\mathrm{t^n}}{\mathrm{n}!}\ \mathrm{E}(\mathrm{X}^\mathrm{n}).$$

Es absolutamente convergente para toda $t > 0$.

Entonces esta sucesión de momentos determina de manera única a $F_x(x)$, esta condición es suficiente pero no necesaria.

Momentos alrededor del origen $0$ (no centrales).

$$\mu'_n = \begin{cases} \text{discreta,} & \sum_x x^n P_x(x). \\ \text{continua,} & \int_x x^n f_x(x)\, dx = \int x^n\, dF_x(x). \end{cases}$$

$\mu'_1 = E(X) = \mu.$

$\mu'_2 = E(X^2).$

$\mu'_3 = E(X^3).$

$\mu'_4 = E(X^4).$

Momentos alrededor de la media $\mu$ (centrales).

$$\mu_n = \begin{cases} \text{discreta,} & \sum_x (x-\mu)^n\, P_x(x). \\ \text{continua,} & \int_x (x-\mu)^n f_x(x)\, dx. \end{cases}$$

$\mu_0 = E(x-\mu)^0 = E(1) = 1.$

$\mu_1 = E(x-\mu) = E(x) - E(\mu) = \mu - \mu = 0.$

$\mu_2 = E(x-\mu)^2 = E(x^2) - \mu^2 = \mu'_2 - (\mu'_1)^2.$

$\mu_3 = E(x-\mu)^3 = E(x^3 - 3x^2\mu + 3x\mu^2 - \mu^3) = \mu'_3 - 3\mu'_1\mu'_2 + 2(\mu'_1)^3.$

$\mu_4 = E(x-\mu)^4 = E(x^4 - 4x^3\mu + 6x^2\mu^2 - 4x\mu^3 + \mu^4) = \mu'_4 - 4\mu'_1\mu'_3 + 6(\mu'_1)^2\mu'_2 - 3(\mu'_1)^4.$

Coeficiente de asimetría:

$$\propto_3 = \frac{\mu_3}{(\mu_2)^{3/2}}.$$

$\propto_3 < 0$ Asimetría negativa.

$\propto_3 > 0$ Asimetría positiva.

$\propto_3 = 0$ Simétrica.

Coeficiente de curtosis:

$$\propto_4 = \frac{\mu_4}{(\mu_2)^2}.$$

$\propto_4 < 3$ Planticúrtica.

$\propto_4 > 3$ Leptocúrtica.

$\propto_4 = 3$ Mesocúrtica.

## 2.2 Sucesión de Variables Aleatorias.

**Proposición 29**:

Sea $\{X_n | n \in \mathbb{N}\}$ una sucesión de variables aleatorias.

Entonces $\sup_n \{X_n\} \wedge \inf_n \{X_n\}$ cuando existen son variables aleatorias.

Demostración:

Sabemos que $\forall\ x \in \mathbb{R}$.

$$(\sup\ X_n \leq x) = \bigcap_{n=1}^{\infty} (X_n \leq x) \in \mathcal{F}.$$

$$(\inf\ X_n \geq x) = \bigcap_{n=1}^{\infty} (X_n \geq x) \in \mathcal{F}.$$

Esto implica que son variables aleatorias.

∎

**Proposición 30**:

Sea $\{X_n | n \in \mathbb{N}\}$ una sucesión de variables aleatorias.

Entonces

$$\lim_{n\to\infty} \sup X_n \quad \wedge \quad \lim_{n\to\infty} \inf X_n.$$

Cuando existen son variables aleatorias.

Demostración:

$$\lim_{n\to\infty} (\sup\ X_n \leq x) = \inf_{k\ \ n\geq k} (\sup X_n) \in \mathcal{F}.$$

$$\lim_{n\to\infty} (\inf\ X_n \geq x) = \sup_{k\ \ n\geq k} (\inf X_n) \in \mathcal{F}.$$

Aplicando el límite a la igualdad de la proposición anterior.

Entonces $X$ es una variable aleatoria.

∎

**Proposición 31**:

Sea $\{X_n | n \in \mathbb{N}\}$ una sucesión de variables aleatorias tales que:

$$\lim_{x\to\infty} X_n(\omega)\ \exists \quad \forall\ \omega \in \Omega.$$

El límite de una variable aleatoria es variable aleatoria.

Función de distribución.

$$F_x(x) = P\{X \leq x\}.$$

i)
$\lim_{x \to -\infty} F_x(x) = 0.$

ii)
$\lim_{x \to \infty} F_x(x) = 1.$

iii) Monótona no decreciente.
Si $x_1 \leq x_2$.
$F_x(x_1) \leq F_x(x_2)$.

iv) Continua por la derecha.
$\lim_{h \to 0^+} F_x(x + h) = F_x(x)$.

## 2.3 Vectores Aleatorios.

**Definición**:

Un Vector Aleatorio $\underline{X}$ es una función $\underline{X}: \Omega \to \mathbb{R}^n$ tal que para cualquier $B\, B(\mathbb{R}^n)$, se cumple que $X^{-1}B$ es un elemento de $\mathcal{F}$.

$\underline{X} = (x_1, \ldots, x_n)$ donde $x_i$ son variables aleatorias $\Omega \to \mathbb{R}$.

Ejemplo:

Lanzamiento de dos dados:

$$\Omega = \{(x.y) | x, y = 1, 2, 3, 4, 5, 6\} \quad \Omega^{\#} = 36.$$

Sea:

$$\underline{X} = (x, y) \to x + y.$$

Nota: Si regreso a la imagen inversa, debo ver de dónde salen los resultados de $x + y$. Es decir que por ejemplo $3$ puede obtenerse de $1 + 2$ o de $2 + 1$.

**Proposición 32**:

Una función $\underline{X} = (X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ es un vector aleatorio si solo cada coordenada del mismo es una variable aleatoria.

Demostración:

Para que $\underline{X}$ sea un vector aleatorio, la imagen inversa de cualquier elemento del conjunto de Borel en $\mathbb{R}^n \in \mathcal{F}$ del espacio de probabilidad definido.

Supongamos el conjunto $B \times \Omega \times \ldots \times \Omega \in \mathcal{F}$.

Es decir el conjunto Boreliano en $\mathbb{R}^n$ para cualquier Boreliano en $\mathbb{R}$.

Para el primer componente implica que $X_1^{-1}B \in \mathcal{F}$ entonces $X_1$ es una variable aleatoria.

Así subsecuentemente, para el n componente implica que $X_n^{-1}B \in \mathcal{F}$ entonces $X_n$ es una variable aleatoria.

Entonces supongamos que la función $(X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ es una variable aleatoria ( para cada una de las coordenadas).

Y considérese la colección:

$$\beta = \{B \in \mathfrak{B}(\mathbb{R}^n) \big| (X_1, \ldots, X_n)^{-1} B \in \mathcal{F} \}.$$

Entonces, los conjuntos de Borel de la forma:

$$B_1 \times B_2 \times \ldots \times B_n.$$

Nos representan el conjunto correspondiente para el vector, siendo $B_i$ el conjunto de Borel asociado a $X_i$ , donde $B_i \in \beta$.

Entonces $\beta(\mathbb{R}) \times \beta(\mathbb{R}) \times \ldots \times \beta(\mathbb{R}) \subseteq \beta \subseteq \mathfrak{B}(\mathbb{R}^n)$.

Por lo tanto cada uno de los componentes de un vector aleatorio es una variable aleatoria.

Un vector aleatorio discreto es aquel que todos sus componentes son variables aleatorias discretas, mientras que un vector aleatorio continuo es aquel que sus componentes con variables aleatorias continuas.

Y no puede haber vectores aleatorios con ambos tipos de variables, pues no se cumplirían ciertas propiedades.

### 2.3.1 Función de Distribución Conjunta y Marginal y Función de Densidad Conjunta y Marginal de un Vector Aleatorio.

**Definición**:

La función de distribución de un vector $\underline{X} = (x, y)$ denotada por $F_{\underline{X}}(x, y): \mathbb{R}^2 \to [0, 1]$ se define como:

$$F_{\underline{X}}(x, y) = P\{X \leq x,\ \ Y \leq y\}.$$

**Proposición 33**:

La distribución $F_{\underline{X}}(x, y)$ satisface las siguientes propiedades:

1)

$$\lim_{x,y\to\infty} F_{\underline{X}}(x, y) = 1 \qquad P\{X \leq \infty,\ \ Y \leq \infty\} = 1.$$

2)

$$\lim_{x\text{ ó }y\to\infty} F_{\underline{X}}(x, y) = 0 \qquad P\{X \leq -\infty,\ \ Y \leq y\} = 0 \quad \text{ó} \quad P\{X \leq x,\ \ Y \leq -\infty\} = 0.$$

3) $F_{\underline{X}}(x, y)$ es no decreciente en cada componente del vector aleatorio.

4) Si $a_1 < b_1$ y $a_2 < b_2$ , entonces

$$F_{\underline{X}}(b_1, b_2) - F_{\underline{X}}(a_1, b_2) - F_{\underline{X}}(b_1, a_2) + F_{\underline{X}}(a_1, a_2) \geq 0.$$

**Definición**:

Una función cualquiera $F_{\underline{X}}(x, y): \mathbb{R}^2 \to [0, 1]$ , no necesariamente definida para un vector aleatorio, es función de distribución si cumple con las propiedades de la proposición anterior.

Generalizando el inciso 4) de la proposición anterior tenemos que una función $F_{\underline{X}}(x,y)\colon \mathbb{R}^n \to [0,1]$ es una función de distribución si:

$$\sum_{x_i \in \{a_i, b_i\}} (-1)^{\#a} \; F_{\underline{X}}(x_1, \dots, x_n) \geq 0.$$

Donde $\#a$ es el número de veces que alguna variable $x_i$ toma el valor $a$ en la evaluación de $F_{\underline{X}}(x)$.

**Proposición 34**:

Sea $F_{\underline{X}}(x_1, \dots, x_n) : \mathbb{R}^n \to \mathbb{R}$ una función de distribución. Entonces existe un espacio de probabilidad $(\mathbb{R}^n, \mathfrak{B}(\mathbb{R}^n), P^*)$ vector aleatorio cuya función de distribución de probabilidad es $F_{\underline{X}}(x_1, \dots, x_n)$.

**Definición**:

La densidad conjunta de un vector aleatorio discreto $(x,y)$ es la función $P_{\underline{X}}(x,y)\colon \mathbb{R}^2 \to \mathbb{R}$ , donde:

$P_{\underline{X}}(x,y) = P\{X = x, \; Y = y\}$ Función masa de probabilidad.

Con

i) $P_{\underline{X}}(x) \geq 0.$

ii)

$$\sum_x \sum_y P_{\underline{X}}(x,y) = 1.$$

Ejemplo:

Sea $\underline{X}$ vector aleatorio discreto tal que:

$$P_{\underline{X}}(x,y) = \begin{cases} \left(\frac{1}{2}\right)^{x+y}, & \forall \; x,y \in \mathbb{N}. \\ 0, & \text{en otro caso.} \end{cases}$$

i) $P_{\underline{X}}(x, y) \geq 0 \quad \longrightarrow \quad \frac{1}{2} > 0,\ \ x, y \in \mathbb{N} \Rightarrow x + y > 0.$

ii)

$$\sum_{x=1}^{\infty}\sum_{y=1}^{\infty}\left(\frac{1}{2}\right)^{x+y} = 1.$$

Demostración:

$$\sum_{x=1}^{\infty}\sum_{y=1}^{\infty}\left(\frac{1}{2}\right)^{x+y} = \sum_{x=1}^{\infty}\left[\left(\frac{1}{2}\right)^{x+1} + \left(\frac{1}{2}\right)^{x+2} + \ldots\right].$$

Factorizando $\left(\frac{1}{2}\right)^{x}$

$$= \sum_{x=1}^{\infty}\left(\frac{1}{2}\right)^{x}\left[\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^{2} + \left(\frac{1}{2}\right)^{3} + \ldots\right]$$

$$= \left[\sum_{x=1}^{\infty}\left(\frac{1}{2}\right)^{x}\right]^{2} = (1)^2 = 1.$$

■

Ahora demostrando que la serie converge a 1.

Tomamos los primeros $\mathrm{n}$ terminos:

$$\left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^{2} + \left(\frac{1}{2}\right)^{3} + \ldots + \left(\frac{1}{2}\right)^{\mathrm{n}}.$$

Factorizamos $\left(\frac{1}{2}\right)$.

$$\frac{1}{2}\left[1 + \left(\frac{1}{2}\right) + \left(\frac{1}{2}\right)^{2} + \left(\frac{1}{2}\right)^{3} + \ldots + \left(\frac{1}{2}\right)^{\mathrm{n}-1}\right].$$

Tenemos la razón r = $\frac{1}{2}$.

Aplicamos la fórmula

$$\frac{1-\left(\frac{1}{2}\right)^{n-1} r}{1-r} = \frac{1-\left(\frac{1}{2}\right)^{n}}{\frac{1}{2}}.$$

$$\lim_{n\to\infty} \frac{1-\left(\frac{1}{2}\right)^{n}}{\frac{1}{2}} = 2.$$

$$\frac{1}{2}\left[1+\left(\frac{1}{2}\right)+\left(\frac{1}{2}\right)^{2}+\left(\frac{1}{2}\right)^{3}+\ \dots+\left(\frac{1}{2}\right)^{n-1}\right] = \frac{1}{2}[2] = 1.$$

■

**Definición**:

Sea $\underline{\mathrm{X}}$ el vector aleatorio continuo con función de distribución de probabilidad $\mathrm{F}_{\underline{\mathrm{X}}}(x,y)$. Se dice que $\underline{\mathrm{X}}$ es "absolutamente continuo" si existe una función no negativa e integrable $f_{\underline{\mathrm{X}}}(x,y): \mathbb{R}^2 \to \mathbb{R}$ para todo $(x,y) \in \mathbb{R}^2$.

Si se cumple la igualdad

$$\mathrm{F}_{\underline{\mathrm{X}}}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{\underline{\mathrm{X}}}(u,v)\, du\, dv$$

$$= \mathrm{P}\,\{\mathrm{X} \le x,\ \ \mathrm{Y} \le y\}.$$

A la función $f_{\underline{\mathrm{X}}}(x,y)$ se le llama Función de Densidad Conjunta de $x$ e $y$.

Se debe cumplir:

i) $f_{\underline{\mathrm{X}}}(x,y) \ \ge 0.$

ii)

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\underline{\mathrm{X}}}(x,y)\, dx\, dy = 1.$$

Integral iterada.

Si tomamos en cuenta el Teorema de Fubinni el espacio debe ser producto cruz.

$$\int_{A\times B} |f(x,y)|\, d(x,y) < \infty.$$

Dando a entender que converge.

Debe tomarse en cuenta que ya no es necesario usar valor absoluto pues como se está trabajando con función densidad, es $\geq 0$.

La integral respecto al producto cartesiano de los dos intervalos en el espacio $(A \times B)$ es igual a:

$$\int_{A\times B} f(x,y)\, d(x,y) = \int_{A} \left[ \int_{B} f(x,y) dy \right] dx \quad = \quad \int_{B} \left[ \int_{A} f(x,y) dx \right] dy.$$

Por otro lado, toda función no negativa $f: \mathbb{R}^2 \to \mathbb{R}$ que integre a $1$, se llama función de densidad si $f_{\underline{X}}(x,y)$ es continua,

$$f_{\underline{X}}(x,y) = \frac{\delta^2}{\delta x \delta y} F_{\underline{X}}(x,y).$$

Teorema fundamental del calculo.

**Definición**:

La función de distribución de la variable aleatoria $X$ es la función $F_x(x) : \mathbb{R} \to \mathbb{R}$.

Definida como:

$$P\{X \leq x\} \quad \forall \ x \in \mathbb{R}.$$

**Proposición 35**:

Sea $F_x(x)$ función de distribución de probabilidad de la variable aleatoria $X$.

1)
$$\lim_{x\to\infty^+} F_x(x) = 1.$$

Demostración:

Sea $\{X_n | n \in \mathbb{N}\}$ una sucesión cualquiera de números reales creciente e infinita y sean los eventos.

$$A_n = (x \le X_n).$$

Entonces $A_n$ es también una sucesión creciente cuyo límite es $\Omega$.

Entonces
$$\lim_{n\to\infty} F_x(X_n) = \lim_{n\to\infty} P(A_n) = P(\Omega) = 1.$$

Como $\mathbb{R}$ es un espacio medible ⇒ $F_x(x)$ converge a 1 cuando $n \to \infty$.

∎

2)
$$\lim_{x\to\infty^-} F_x(x) = 0 .$$

Demostración:

Sea $\{X_n | n \in \mathbb{N}\}$ una sucesión cualquiera de números reales decreciente e infinita y sean los eventos.

$$A_n = (x \le X_n).$$

Entonces $A_n$ es también una sucesión decreciente cuyo límite es $\emptyset$.

Entonces
$$\lim_{n\to\infty} F_x(X_n) = \lim_{n\to\infty} P(A_n) = P(\emptyset) = 0.$$

∎

3) $F_x(x)$ Es una función monótona no decreciente.

Si $x_1 \le x_2 \quad \Rightarrow \quad F_x(x_1) \le F_x(x_2).$

Demostración:

Sabemos que:

$$F_x(x_1) \leq F_x(x_1) + P(x_1 \leq X \leq x_2).$$

$$F_x(x_1) \leq P(X \leq x_1) + P(x_1 \leq X \leq x_2).$$

$$F_x(x_1) \leq P[(X \leq x_1) \cup (x_1 \leq X \leq x_2)].$$

$$F_x(x_1) \leq P[X \leq x_2].$$

$$\therefore \ F_x(x_1) \leq F_x(x_2).$$

∎

4) $F_x(x)$ es continua por la derecha.

$F_x(x^+) = F_x(x)$.

Demostración:

Sea $\{X_n \mid n \in \mathbb{N}\}$ una sucesión cualesquiera de números reales no negativos y decreciente a $0$.

Entonces $F_x(x + x_n) = F_x(x) + P\{x < X \leq x + x_n\} = P(X \leq x) + P\{x < X \leq x + x_n\}$.

Donde $A_n = (x < X \leq x + x_n)$ sucesión decreciente a $\emptyset$.

Entonces $F_x(x + x_n) = P\{x < X \leq x + x_n\}$.

$$\lim_{n\to\infty} f_x(x + x_n) = \lim_{n\to\infty} P\{x < X \leq x + x_n\}.$$

$$F_x(x) \Rightarrow F_x(x^+) = F_x(x).$$

∎

**Definición**:

Una función $F_x(x) : \mathbb{R} \to \mathbb{R}$ es llamada función de distribución si cumple con los cuatro postulados de la proposición anterior.

**Proposición 36**:

Para cualesquiera números $x$, a, b $\in \mathbb{R}$ con $a \leq b$ tenemos:

1) $P\{X < x\} = F_x(x^-)$ limite por la izquierda de la función F en el punto $x$.

2) $F_x(x) - F_x(x^-) = P\{X = x\}$.

3) $P(x \in [a,b]) = F_x(b) - F_x(a)$.

4) $P(x \in [a,b]) = F_x(b) - F_x(a^-)$.

5) $P(x \in [a,b]) = F_x(b^-) - F_x(a)$.

6) $P(x \in [a,b]) = F_x(b^-) - F_x(a^-)$.

Por ser no decreciente 5) y 6) $P\{X \leq b\} - P\{X \leq a\}$ $F_x(b) - F_x(a)$.

Todos estos límites existen porque la función $F_x(x)$ es monótona no decreciente.

Demostración:

1) Sean $\{X_n \mid n \in \mathbb{N}\}$ una sucesión cualesquiera de números reales no negativos y decrecientes a $0$.

   Sea $A_n$ el evento $(x \leq a - X_n)$ $\Rightarrow$ $\{A_n \mid n \in \mathbb{N}\}$ es una sucesión decreciente al evento $(x < a)$ $\Rightarrow$ por la probabilidad de continuidad.

$$P(x < a) = \lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} P(a - X_n) = F_x(a^-).$$

∎

2)

$$P\{X = x\} = P\{X \leq x\} - P\{X < x\}.$$

Por el inciso anterior

$$= F_x(x) - F_x(x^-) = 0.$$

Si X es variable aleatoria continua.

Si $F_x(x) - F_x(x^-) > 0$.

La magnitud del salto es $P\{X = x\}$ si $X \sim$ variable aleatoria continua.

$P\{X = x\}$ Tamaño del salto o discontinuidad de $F_x(x)$ en el punto $x$.

∎

**Proposición 37**:

Toda función de distribución tiene a lo más un número numerable de saltos (discontinuidades).

Demostración:

Sea $\mathrm{D}$ el conjunto de puntos de discontinuidad de una función de distribución $\mathrm{F_x}(x) \Rightarrow$ Sea la sucesión $\{\mathrm{D_1}, \mathrm{D_2}, \ldots\} \quad \forall \ \mathrm{n} \in \mathbb{N}$.

$$\mathrm{D_n} = \left\{ \mathrm{X} \in \mathrm{D} \,\middle|\, \frac{1}{\mathrm{n}+1} \leq \mathrm{F}(x) - \mathrm{F}(x^-) \leq \frac{1}{\mathrm{n}} \right\}.$$

Conforme avanzamos a cero se va acotando la diferencia.

$\mathrm{D_n}$ tiene un número numerable de elementos.

$$\mathrm{D} = \bigcup_{\mathrm{n}=1}^{\infty} \mathrm{D_n}.$$

Es conjunto numerable.

∎

Dada la función de distribución conjunta $\mathrm{F}_{\underline{\mathrm{X}}}(x, y)$ de un vector aleatorio $\underline{\mathrm{X}}$, es posible obtener $\mathrm{F}_{\underline{\mathrm{X}}}(x)$ y $\mathrm{F}_{\underline{\mathrm{X}}}(y)$.

**Definición**:

Sea $\underline{\mathrm{X}} = (x, y)$ vector aleatorio con función de probabilidad continua $f_{\underline{\mathrm{X}}}(x, y)$.

- Caso continuo

$$f_{\underline{\mathrm{X}}}(x) = \int_{-\infty}^{\infty} f_{\underline{\mathrm{X}}}(x, y)\, dy.$$

$$f_{\underline{\mathrm{X}}}(y) = \int_{-\infty}^{\infty} f_{\underline{\mathrm{X}}}(x, y)\, dx.$$

- Caso discreto

$$\mathrm{P}_{\underline{\mathrm{X}}}(x) = \sum_{y} \mathrm{P}_{\underline{\mathrm{X}}}(x, y).$$

$$\mathrm{P}_{\underline{\mathrm{X}}}(y) = \sum_{x} \mathrm{P}_{\underline{\mathrm{X}}}(x, y).$$

**Definición**:

Sea $\underline{\mathrm{X}} = (x, y)$ vector aleatorio con función de probabilidad continua $f_{\underline{\mathrm{X}}}(x, y)$ con $y$ tal que $f_{\underline{\mathrm{X}}}(y) > 0$.

$$x \to f_{\underline{\mathrm{X}}}(x|y) = \frac{f_{\underline{\mathrm{X}}}(x, y)}{f_{\underline{\mathrm{X}}}(y)}.$$

A esta función se le conoce como Función de Densidad Condicional de $x$ dado $y$ toma el valor especifico de $y$.

Si $\underline{\mathrm{X}}$ es absolutamente continua, con función de probabilidad continua $f_{\underline{\mathrm{X}}}(x, y)$ y con $f_{\underline{\mathrm{X}}}(y) > 0$ a la función.

$$x \to \mathrm{F}_{\underline{\mathrm{X}}}(x|y) = \int_{-\infty}^{x} f_{\underline{\mathrm{X}}}(u|v)\, du.$$

Se le conoce como función de distribución de $x$ dado que $y$ toma el valor especifico de $y$.

Si $\underline{\mathrm{X}}$ es un vector aleatorio discreto entonces $\int$ se sustituye por una $\sum$.

Independencia de Variables Aleatorias (componentes de un vector aleatorio).

**Definición**:

Se dice que $x$ e $y$ son independientes si para $(x, y) \in \mathbb{R}^2$, se cumple que:

$\mathrm{F}_{\underline{\mathrm{X}}}(x|y) = \mathrm{F}_{\underline{\mathrm{X}}}(x)$ y $\mathrm{F}_{\underline{\mathrm{X}}}(y|x) = \mathrm{F}_{\underline{\mathrm{X}}}(y)$.

Entonces cuando existe

$f_{\underline{\mathrm{X}}}(x|y) = f_{\underline{\mathrm{X}}}(x)$ y $f_{\underline{\mathrm{X}}}(y|x) = f_{\underline{\mathrm{X}}}(y)$.

Nota: Esta definición se puede generalizar para vectores aleatorios con n componentes.

**Proposición 38**:

Si $x$ e $y$ componentes de $\underline{\mathrm{X}}$ son independientes, entonces:

$\mathrm{F}_{\underline{\mathrm{X}}}(x,y) = \mathrm{F}_{\underline{\mathrm{X}}}(x)\,\mathrm{F}_{\underline{\mathrm{X}}}(y)$ y $f_{\underline{\mathrm{X}}}(x,y) = f_{\underline{\mathrm{X}}}(x)\,f_{\underline{\mathrm{X}}}(y)$.

Demostración:

Sabemos que si $x$ e $y$ son independientes,

$\mathrm{F}_{\underline{\mathrm{X}}}(x|y) = \mathrm{F}_{\underline{\mathrm{X}}}(x)$ y $\mathrm{F}_{\underline{\mathrm{X}}}(y|x) = \mathrm{F}_{\underline{\mathrm{X}}}(y)$.

Y de acuerdo a la definición de función de distribución condicional.

$$\mathrm{F}_{\underline{\mathrm{X}}}(x|y) = \frac{\mathrm{F}_{\underline{\mathrm{X}}}(x,y)}{\mathrm{F}_{\underline{\mathrm{X}}}(y)}.$$

Igualando las definiciones anteriores tenemos:

$$\mathrm{F}_{\underline{\mathrm{X}}}(x) = \frac{\mathrm{F}_{\underline{\mathrm{X}}}(x,y)}{\mathrm{F}_{\underline{\mathrm{X}}}(y)} \Rightarrow \mathrm{F}_{\underline{\mathrm{X}}}(x,y) = \mathrm{F}_{\underline{\mathrm{X}}}(x)\mathrm{F}_{\underline{\mathrm{X}}}(y).$$

■

Ejemplo:

Sea

$$\mathrm{F}_{\underline{\mathrm{X}}}(x,y) = \begin{cases} 0, & (x,y) = 0. \\ 1 - e^{-x} - e^{-y} + e^{-x-y}, & (0,0) \leq (x,y) \leq (\infty,\infty). \\ 1, & (x,y) \geq (\infty,\infty). \end{cases}$$

Obteniendo Marginales $\left(\mathrm{F}_{\underline{\mathrm{X}}}(x), \mathrm{F}_{\underline{\mathrm{X}}}(y)\right)$, densidad conjunta y $f_{\underline{\mathrm{X}}}(x)$, $f_{\underline{\mathrm{X}}}(y)$.

$$\mathrm{F}_{\underline{\mathrm{X}}}(x) = \lim_{y\to\infty} \mathrm{F}_{\underline{\mathrm{X}}}(x,y) = \begin{cases} 0, & x < 0. \\ 1 - e^{-x}, & 0 \leq x < \infty. \\ 1, & x \geq \infty. \end{cases}$$

$$F_{\underline{X}}(y) = \lim_{x\to\infty} F_{\underline{X}}(x,y) = \begin{cases} 0, & y < 0. \\ 1 - e^{-y}, & 0 \le y < \infty. \\ 1, & y \ge \infty. \end{cases}$$

$$\frac{\delta^2}{\delta x \delta y} F_{\underline{X}}(x,y) = f_{\underline{X}}(x,y) = \begin{cases} e^{-(x+y)}, & (0,0) \le (x,y) \le (\infty,\infty). \\ 0, & \text{en otro caso.} \end{cases}$$

$$f_{\underline{X}}(x) = \int_0^{\infty} e^{-(x+y)} dy = -e^{-(x+y)}\Big|_0^{\infty} = (0 - e^{-x}) \begin{cases} e^{-x}, & 0 \le x < \infty. \\ 0, & \text{en otro caso.} \end{cases}$$

$$f_{\underline{X}}(y) = \int_0^{\infty} e^{-(x+y)} dx = -e^{-(x+y)}\Big|_0^{\infty} = -(0 - e^{-y}) \begin{cases} e^{-y}, & 0 \le y < \infty. \\ 0, & \text{en otro caso.} \end{cases}$$

Esperanza de la función de un vector aleatorio.

**Definición**:

Sea $\underline{X} = (x,y)$ un vector aleatorio y sea $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ una función de Borel medible, entonces:

$$E[\varphi(x,y)] = \int_{\mathbb{R}^2} \varphi(x,y)\ dF_{\underline{X}}(x,y).$$

Para el caso absolutamente continuo $dF_{\underline{X}}(x,y) = f_{\underline{X}}(x,y)dxdy$.

Nota: Las imágenes inversas de los Borelianos de $\mathbb{R}$ caen en la σ-algebra del espacio de probabilidad del vector aleatorio $\underline{X}$.

Si $\underline{X}$ es vector aleatorio discreto, entonces:

$$E[\varphi(x,y)] = \sum_{(x,y)} \varphi(x,y)\ P_{\underline{X}}(x,y).$$

**Proposición 39**:

$\mathrm{E}[x+y] = \mathrm{E}[x] + \mathrm{E}[y]$.

Demostración:

Para caso continuo sea $\varphi(x) = x + y$ , por la definición,

$$\mathrm{E}[(x+y)] = \int_{\mathbb{R}^2} (x+y)\ d\mathrm{F}_{\underline{\mathrm{X}}}(x,y).$$

Por Fubini (integrales iteradas),

$$= \int_{\mathbb{R}^2} x\ d\mathrm{F}_{\underline{\mathrm{X}}}(x,y) +\ y\, d\mathrm{F}_{\underline{\mathrm{X}}}(x,y) = \int_{\mathbb{R}^2} x\ d\mathrm{F}_{\underline{\mathrm{X}}}(x,y) +\ y\, d\mathrm{F}_{\underline{\mathrm{X}}}(x,y) = \mathrm{E}(x) + \mathrm{E}(y)$$

■

**Proposición 40**:

Sean $x$ e $y$ variables aleatorias y componentes del vector aleatorio $\underline{\mathrm{X}}$ y sean $g$ y $h$ dos funciones Borel medibles tal que $g(x)$ y $h(x)$ tienen esperanza finita.

Entonces $\mathrm{E}[g(x)\, h(x)] = \mathrm{E}[g(x)]\ \mathrm{E}[h(x)]$.

Demostración:

Como $x$ e $y$ son independientes, entonces,

$$\mathrm{F}_{\underline{\mathrm{X}}}(x,y) = \mathrm{F}_{\underline{\mathrm{X}}}(x)\, \mathrm{F}_{\underline{\mathrm{X}}}(y).$$

Entonces,

$$\mathrm{E}[g(x)\, h(x)] = \int_{\mathbb{R}^2} g(x)\, h(x)\ d\mathrm{F}_{\underline{\mathrm{X}}}(x,y)$$

$$= \int_{\mathbb{R}} g(x)\ d\mathrm{F}_{\underline{\mathrm{X}}}(x) \int_{\mathbb{R}} h(x)\ d\mathrm{F}_{\underline{\mathrm{X}}}(y) = \mathrm{E}[g(x)]\ \mathrm{E}[h(y)].$$

$\not\Leftarrow$ ) Contraejemplo.

Sea $\underline{X} = (x, y)$ un vector aleatorio discreto con función masa de probabilidad continua.

| $x \backslash y$ | $-1$ | $0$ | $1$ | $P_{\underline{X}}(x) \downarrow$ |
|---|---|---|---|---|
| $-1$ | $\frac{1}{5}$ | $0$ | $\frac{1}{5}$ | $\frac{2}{5}$ |
| $0$ | $0$ | $\frac{1}{5}$ | $0$ | $\frac{1}{5}$ |
| $1$ | $\frac{1}{5}$ | $0$ | $\frac{1}{5}$ | $\frac{2}{5}$ |
| $P_{\underline{X}}(y) \rightarrow$ | $\frac{2}{5}$ | $\frac{1}{5}$ | $\frac{2}{5}$ | |

$$\Rightarrow \mathrm{E}(xy) = \sum_{x=-1}^{1} \sum_{y=-1}^{1} xy \;\; \mathrm{P}_{\underline{X}}(x, y) = \sum_{x=-1}^{1} x \left[-\mathrm{P}_{\underline{X}}(x, -1) + \mathrm{P}_{\underline{X}}(x, 1)\right]$$

$$= \mathrm{P}_{\underline{X}}(-1, -1) - \mathrm{P}_{\underline{X}}(-1, 1) - \mathrm{P}_{\underline{X}}(1, -1) + \mathrm{P}_{\underline{X}}(1, 1)$$

$$= \frac{1}{5} - \frac{1}{5} - \frac{1}{5} + \frac{1}{5} = 0.$$

$$\mathrm{E}(x) = \sum_{x=-1}^{1} x \;\; \mathrm{P}_{\underline{X}}(x) = -1 \; \mathrm{P}_{\underline{X}}(-1) + 0 \; \mathrm{P}_{\underline{X}}(0) + 1 \; \mathrm{P}_{\underline{X}}(1) = -\frac{2}{5} + \frac{2}{5} = 0.$$

$$\mathrm{E}(y) = \sum_{y=1}^{1} x \;\; \mathrm{P}_{\underline{X}}(y) = 0.$$

Entonces

$$\mathrm{E}(xy) = \mathrm{E}(x)\mathrm{E}(y) = 0.$$

Ahora obtendremos la probabilidad condicional.

$$\mathrm{P}_{\underline{X}}(x | \; y = 0) = \frac{\mathrm{P}_{\underline{X}}(x, y = 0)}{\mathrm{P}_{\underline{X}}(y = 0)} = \frac{0}{\frac{1}{5}} = 0.$$

$$\Rightarrow \mathrm{P}_{\underline{X}}(x | \; y = 0) \; \mathrm{f(x)} = \begin{cases} 0, & x = -1. \\ \frac{\frac{1}{5}}{\frac{1}{5}} = 1, & x = 0. \\ 0, & x = 1. \end{cases}$$

Obtenemos ahora las marginales.

$$\mathrm{P}_{\underline{\mathrm{X}}}(x) = \begin{cases} \frac{2}{5}, & x = -1. \\ \frac{1}{5}, & x = 0. \\ \frac{2}{5}, & x = 1. \end{cases}$$

Como la probabilidad marginal y la condicional son diferentes con respecto a $x$.

Por lo tanto $x$ e $y$ no son independientes.

Es decir que $x$ depende de los valores que tome $y$.

■

Nota: $\mathrm{E}(x)$ no necesariamente es un valor del espacio muestral, ejemplo claro es tener un espacio muestral (9, 10) su promedio es 9.5, valor que no es parte del espacio muestral.

## 2.4 Covarianza.

La covarianza de $x$ e $y$ variables aleatorias y componentes del vector aleatorio $\underline{\mathrm{X}}$ se denota por $\mathrm{Cov}(x)$ y es igual a:

$$\mathrm{Cov}(x, y) = \mathrm{E}\big[\big(x - \mathrm{E}(x)\big)\big(y - \mathrm{E}(y)\big)\big].$$

Donde $\mathrm{E}(x)$ , $\mathrm{E}(y)$ existen y son finitas.

Covarianza: Mide el grado de dispersión con respecto tanto a la media de $x$ como a la media de $y$.

**Proposición 41**:

La covarianza satisface las siguientes propiedades:

1) $\operatorname{Cov}(x, y) = \mathrm{E}(xy) - \mathrm{E}(x)\mathrm{E}(y)$.

Demostración:

Por definición:

$$\operatorname{Cov}(x, y) = \mathrm{E}\big[\big(x - \mathrm{E}(x)\big)\big(y - \mathrm{E}(y)\big)\big].$$

Implica que por ser el valor esperado un operador lineal.

$$= \mathrm{E}(xy) - x\mathrm{E}(y) - y\mathrm{E}(x) + \mathrm{E}(x)\mathrm{E}(y)$$

$$= \mathrm{E}(xy) - \mathrm{E}(y)\mathrm{E}(x) - \mathrm{E}(x)\mathrm{E}(y) + \mathrm{E}(x)\mathrm{E}(y)$$

$$= \mathrm{E}(xy) - \mathrm{E}(x)\mathrm{E}(y).$$

■

2) $\operatorname{Cov}(x, y) = \operatorname{Cov}(y, x)$.

Demostración:

Por definición y conmutatividad:

$$\operatorname{Cov}(x, y) = \mathrm{E}\big[\big(x - \mathrm{E}(x)\big)\big(y - \mathrm{E}(y)\big)\big] = \mathrm{E}\big[\big(y - \mathrm{E}(y)\big)\big(x - \mathrm{E}(x)\big)\big] = \operatorname{Cov}(y, x).$$

■

3) $\operatorname{Cov}(x, x) = \operatorname{Var}(x)$.

Demostración:

Por definición

$$\operatorname{Cov}(x, y) = \mathrm{E}\big[\big(x - \mathrm{E}(x)\big)\big(x - \mathrm{E}(x)\big)\big] = \mathrm{E}[x - \mathrm{E}(x)]^2 = \operatorname{Var}(x).$$

■

4) $\operatorname{Cov}(a, x) = 0.$

Donde $a$ es una constante por lo tanto la covarianza entre una constante y una variable aleatoria es cero.

Demostración:

$$\operatorname{Cov}(a, x) = \mathrm{E}\big[\big(a - \mathrm{E}(a)\big)\big(x - \mathrm{E}(x)\big)\big] = \mathrm{E}\big[(a - a)\big(x - \mathrm{E}(x)\big)\big] = 0.$$

■

5) $\operatorname{Cov}(ax, y) = a\operatorname{Cov}(x, y).$

Demostración:

$$\operatorname{Cov}(ax, y) = \mathrm{E}\big[\big(ax - \mathrm{E}(ax)\big)\big(y - \mathrm{E}(y)\big)\big]$$

$$= \mathrm{E}\big[\big(ax - a\mathrm{E}(x)\big)\big(y - \mathrm{E}(y)\big)\big] = \mathrm{E}\big[a\big(x - \mathrm{E}(x)\big)\big(y - \mathrm{E}(y)\big)\big]$$

$$= a\,\mathrm{E}\big[\big(x - \mathrm{E}(x)\big)\big(y - \mathrm{E}(y)\big)\big] = a\operatorname{Cov}(x, y).$$

■

6) $\operatorname{Cov}(x_1 + x_2,\ y) = \operatorname{Cov}(x_1, y) + \operatorname{Cov}(x_2, y).$

7) Si $x$ e $y$ son componentes de $\underline{\mathrm{X}}$ vector aleatorio y son independientes entonces $\operatorname{Cov}(x, y) = 0.$

Demostración:

Por la propiedad 1) sabemos que $\operatorname{Cov}(x, y) = \mathrm{E}(xy) - \mathrm{E}(x)\mathrm{E}(y).$

Así que por proposiciones anteriores tenemos que:

$$\mathrm{E}(xy) \stackrel{\perp}{=} \mathrm{E}(x)\mathrm{E}(y).$$

Entonces $\operatorname{Cov}(x, y) = \mathrm{E}(x)\mathrm{E}(y) - \mathrm{E}(x)\mathrm{E}(y) = 0.$

■

8) Si $\text{Cov}(x,y) = 0 \not\Rightarrow x \wedge y$ sean independientes.

Demostración:

Contraejemplo:

Sea $\underline{\text{X}} = (x,y)$ vector aleatorio discreto con función masa de probabilidad continua.

$$\text{P}_{\underline{\text{X}}}(x,y) = \begin{cases} \frac{1}{8}, & (x,y) \in \{(-1,-1),(-1,1),(1,-1),(1,1)\}. \\ \frac{1}{2}, & (x,y) \in \{(0,0)\}. \\ 0, & \text{en otro caso.} \end{cases}$$

$$\text{P}_{\underline{\text{X}}}(x) = \sum_{y=-1}^{1} \text{P}_{\underline{\text{X}}}(x,y) = \begin{cases} \frac{1}{8} + \frac{1}{8} = \frac{1}{4}, & x = -1. \\ \boldsymbol{\frac{1}{2}}, & x = 0. \\ \frac{1}{8} + \frac{1}{8} = \frac{1}{4}, & x = 1. \\ 0, & \text{en otro caso.} \end{cases}$$

$\text{P}_{\underline{\text{X}}}(x = 0)$.

$$\text{P}_{\underline{\text{X}}}(y) = \sum_{x=-1}^{1} \text{P}_{\underline{\text{X}}}(x,y) = \begin{cases} \frac{1}{4}, & y = -1. \\ \frac{1}{2}, & y = 0. \\ \frac{1}{4}, & y = 1. \\ 0, & \text{en otro caso.} \end{cases}$$

$\text{Cov}(x,y) = 0$ ya que $\text{E}(xy) = 0$, $\text{E}(x) = 0$, $\text{E}(y) = 0$.

$$
P_{\underline{X}}(x|y=0) = \begin{cases} 0, & x=-1. \\ \dfrac{\frac{1}{2}}{\frac{1}{2}} = 1, & x = 0. \\ 0, & x = 1. \end{cases}
$$

Como son diferentes $x$ e $y$ entonces por lo tanto $x$ e $y$ no son independientes.

■

## 2.5 Coeficiente de Correlación.

**Definición**:

El Coeficiente de Correlación de $x$ e $y$ variables aleatorias y componentes del vector aleatorio $\underline{X}$ se denota por $\rho(x,y)$.

$$
\rho(x,y) = \frac{\mathrm{Cov}(x,y)}{\sqrt{\mathrm{Var}(x)\mathrm{Var}(y)}}.
$$

Donde $\mathrm{Var}(x)$ y $\mathrm{Var}(y)$ son estrictamente $> 0$ (positivas) y finitas.

El coeficiente de correlación mide el grado de dependencia lineal.

**Proposición 42**:

El coeficiente de correlación $\rho(x,y)$ satisface las siguientes propiedades.

1) Si $x$ e $y$ son independientes entonces $\rho(x,y)=0$ pero si $\rho(x,y)=0$ no implica que $x$ e $y$ sean independientes, salvo en la distribución normal.

Demostración:

Por definición

$$\rho(x,y) = \frac{\mathrm{Cov}(x,y)}{\sqrt{\mathrm{Var}(x)\mathrm{Var}(y)}}.$$

Pero sabemos que si $x$ e $y$ son independientes, entonces

$$\mathrm{Cov}(x,y) = 0.$$

Entonces

$$\rho(x,y) = \frac{0}{\sqrt{\mathrm{Var}(x)\mathrm{Var}(y)}} = 0.$$

■

2) $-1 \leq \rho(x,y) \leq 1.$

Demostración:

Supongamos que $x$ e $y$ son variables aleatorias con $\mathrm{E}(x) = \mathrm{E}(y) = 0$ y $\mathrm{Var}(x) = \mathrm{Var}(y) = 1$ $\forall \ \lambda \in \mathbb{R}$, tomando $x + \lambda y$.

$$\mathrm{E}(x + \lambda y) = \mathrm{E}(x) + \lambda \mathrm{E}(y) = 0 + \lambda 0 = 0.$$

$$\mathrm{Var}(x + \lambda y) = \mathrm{E}[x^2 + 2\lambda xy + \lambda^2 y^2] = \mathrm{E}(x^2) + 2\lambda \mathrm{E}(xy) + \lambda^2 \mathrm{E}(y^2) = 0.$$

Entonces

$$\mathrm{Var}(x + \lambda y) = 1 + 2\lambda \mathrm{E}(xy) + \lambda^2.$$

Si $\lambda = 1.$

$$\mathrm{Var}(x + y) = 2 + 2\mathrm{E}(xy) = 2[1 + \mathrm{E}(xy)] \geq 0.$$

Entonces

$$\mathrm{E}(xy) \geq -1.$$

Si $\lambda = -1.$

$$\mathrm{Var}(x + y) = 2 - 2\mathrm{E}(xy) = 2[1 - \mathrm{E}(xy)] \geq 0.$$

Entonces

$\mathrm{E}(xy) \leq -1.$ Por lo tanto $-1 \leq \mathrm{E}(xy) \leq 1.$

Ahora sean las variables (estandarización).

$$u = \frac{x - \mu_x}{\sigma_x} \quad \text{y} \quad v = \frac{y - \mu_y}{\sigma_y}.$$

Para cualesquiera dos variables aleatorias del espacio de probabilidad $(\Omega,\ \mathcal{F},\ \mathcal{P})$.

Entonces

$$\mathrm{E}(uv) = \mathrm{E}\left[\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right] = \frac{\mathrm{E}(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}$$

$$= \frac{\mathrm{Cov}(x,y)}{\sigma_x\sigma_y} = \rho(x,y).$$

Por la primera parte, como $\mathrm{E}(u)=\mathrm{E}(v)=0$ y $\mathrm{Var}(u)=\mathrm{Var}(v)=1$.

$$-1 \le \mathrm{E}(uv) \le 1.$$

Entonces

$$-1 \le \rho(x,y) \le 1.$$

∎

3) $|\rho(x,y)| = 1 \iff$ podemos expresar $y = ax + b$ con $a, b$ constantes (Cualquiera de las variables se puede expresar en función de la otra).

Demostración:

Si $x$ e $y$ son variables aleatorias tales que $y = ax + b$ con $a \neq b$.
Entonces

$$\rho(x,y) = \frac{\mathrm{Cov}(x, ax+b)}{\sqrt{\mathrm{Var}(x)\mathrm{Var}(ax+b)}}.$$

Por proposiciones anteriores tenemos que:

$$\mathrm{Cov}(x,ax) + \mathrm{Cov}(x,b).$$
$$a\mathrm{Cov}(x,x) + 0 = a\mathrm{Var}(x).$$

Ahora

$$\sqrt{\mathrm{Var}(x)\mathrm{Var}(ax+b)} = \sqrt{\mathrm{Var}(x)\mathrm{Var}(ax)}$$

$$= \sqrt{a^2\ \mathrm{Var}^2(x)}$$

$$= |a|\ \mathrm{Var}(x).$$

Entonces

$$\rho(x,y) = \frac{a\mathrm{Var}(x)}{|a|\ \mathrm{Var}(x)} = \frac{a}{|a|}.$$

∎

Expresiones en regresión lineal.

$$a = \rho(x, y)\frac{\sigma_y}{\sigma_x}.$$

$$b = \mu_y - \rho(x, y)\mu_x\ \frac{\sigma_y}{\sigma_x}.$$

Esperanza y Varianza de un vector aleatorio.

La esperanza de un vector aleatorio $\underline{\mathrm{X}} = (x_1, \dots, x_\mathrm{n})$ se define como el vector numérico $\big(\mathrm{E}(x_1), \dots, \mathrm{E}(x_\mathrm{n})\big)$ si y solo si existe $\mathrm{E}(x_i)$ para todo $i = \overline{1, \mathrm{n}}$.

La varianza de un vector aleatorio se define como la matriz cuadrada.

$$\mathrm{E}\{[x - \mathrm{E}(x)]^\mathrm{t}\,[x - \mathrm{E}(x)]\} \in \mathrm{M}_{\mathrm{n}\times\mathrm{n}}.$$

Nos resulta la matriz de Varianzas y Covarianzas, entonces

$$\mathrm{Var}(x) = \begin{bmatrix} \mathrm{Var}(x) & \cdots & \mathrm{Cov}(x_1, x_n) \\ \vdots & \ddots & \vdots \\ \mathrm{Cov}(x_\mathrm{n}, x_1) & \cdots & \mathrm{Var}(x_\mathrm{n}) \end{bmatrix}.$$

La matriz es simétrica dado que

$$\mathrm{Cov}\big(x_i, x_j\big) = \mathrm{Cov}\big(x_j, x_i\big) \quad \forall\, i \neq j \quad i, j = \overline{1, \mathrm{n}}.$$

# Capítulo 3 MODOS DE CONVERGENCIA.

## 3.1 Convergencia Puntual.

**Definición**:

La sucesión $x_1, x_2, \ldots$ converge puntualmente a $x$ si para cada $\omega \in \Omega$.

$$\mathrm{X}(\omega) = \lim_{n \to \infty} \mathrm{X}_n(\omega).$$

Sucesión infinita de $x_1, x_2, \ldots$

$$\underbrace{\mathrm{X}_1(\omega), \mathrm{X}_2(\omega), \ldots, \mathrm{X}_n(\omega)}.$$

Conjunto de números reales.

Converge puntualmente $\rightarrow \mathrm{X}(\omega)$.

Entonces $\mathrm{X} : \Omega \rightarrow \mathbb{R}$ definida por:

$$\mathrm{X}(\omega) = \lim_{n \to \infty} \mathrm{X}_n(\omega).$$

Por la proposición anterior

Ejemplo:

Sea el espacio medible $\big([0,1],\ \mathfrak{B}([0,1])\big)$.

Defina la sucesión $\mathrm{X}_n(\omega) = \omega^n$.

Entonces para cada $\omega \in [0,1)$, $\mathrm{X}_n(\omega) \rightarrow 0$ como $\omega$ esta entre $0$ y $1$ cuando $n \rightarrow \infty$ eso tiende a $0$.

Entonces $\mathrm{X}_n$ converge puntualmente a la variable aleatoria.

$$\mathrm{X}_n(\omega) = \begin{cases} 0, & \text{si } 0 \leq \omega < 1. \\ 1, & \text{si } \omega = 1. \end{cases}$$

Esto no es probabilidad, solo se define la variable aleatoria para cada punto.

Nota: Se requiere la convergencia en cada punto del espacio muestral.

Se requiere entonces que para cada $\omega \in \Omega$.

$$\lim_{n\to\infty} X_n(\omega) = X(\omega).$$

Esto significa que $\forall \varepsilon > 0$ existe $n_0$ tal que $\forall\ n > n_0$.

$$|X_n(\omega) - X(\omega)| < \varepsilon.$$

$\varepsilon$ está en función de $n$ y de $\omega$.

Nota: No depende de $\varepsilon$ y del elemento $\omega$ de la σ-algebra considerada.

## 3.2 Convergencia Uniforme.

Cuando $n_0$ depende solo de $\varepsilon$ y es el mismo $\varepsilon$ $\forall\ \omega \in \Omega$, entonces se dice que la convergencia es uniforme.

Nota: La anterior es la convergencia más estricta.

Convergencia Uniforme implica Convergencia Puntual.

Por definición y condiciones manejadas pero,

Convergencia Uniforme no implica Convergencia Puntual.

## 3.3 Convergencia casi Segura.

**Definición**:

La sucesión $x_1, x_2, \dots$ converge casi seguramente a $x$ si:

$$P\left\{\omega \in \Omega\colon \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\} = 1\,.$$

En la convergencia casi segura se permite que algunos valores de la sucesión $X_1(\omega), X_2(\omega), \dots$ puedan no converger, siendo el subconjunto de nuestros valores, un subconjunto con probabilidad $= 0$.

Notación:

$$X_n \overset{c.s.}{\rightarrow} X.$$

$$\lim_{n\to\infty} X_n = X_{c.s.}.$$

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1.$$

Nota:

Convergencia Puntual implica Convergencia Casi segura.

Convergencia Puntual no implica Convergencia Casi segura.

Convergencia Uniforme implica Convergencia Puntual implica Convergencia Casi segura.

Convergencia Uniforme no implica Convergencia Puntual no implica Convergencia Casi segura.

Ejemplo:

Convergencia Casi Segura.

Considere el espacio de probabilidad $([0,1],\ \mathfrak{B}([0,1]), P)$.

$$X_n(\omega) = \begin{cases} 1, & 0 \leq \omega \leq \frac{1}{n}. \\ 0, & \text{en otro caso.} \end{cases}$$

Tomando en cuenta que $X_n \sim \mathrm{Ber}\left(\frac{1}{2}\right)$.

$$X_1(\omega) = \begin{cases} 1, & 0 \leq \omega \leq 1. \\ 0, & \text{en otro caso} \end{cases}$$

$$X_2(\omega) = \begin{cases} 1, & 0 \leq \omega \leq \frac{1}{2} \\ 0, & \text{en otro caso.} \end{cases}$$

$$X_3(\omega) = \begin{cases} 1, & 0 \leq \omega \leq \frac{1}{3}. \\ 0, & \text{en otro caso.} \end{cases}$$

Lo que nos hace ver que el limite $\rightarrow 0$, que existe intersección solo en el cero y que es una sucesión decreciente, así como la intersección de todos sus elementos es $= 0$ y el valor que no converge a cero es $n = 0$.

Para demostrar que la convergencia casi segura existe.

Demostración:

$$P\,(X_n \rightarrow 0) = 1\,.$$

$$P\left\{\omega \in \Omega\colon \lim_{n\rightarrow\infty} X_n(\omega) = 0\right\} = (0,1].$$

Y en el punto $\omega = 0$ no hay convergencia.

$$P\,\{X_n(\omega) = 0\} = 0.$$

Casi igual a cero pero no necesariamente el conjunto vacio.

∎

Retomando el ejemplo anterior:

Considerando el espacio de probabilidad $([0,1],\ \mathfrak{B}([0,1]), P)$ Donde $P$ es medida uniforme para todo el intervalo.

Elementos de la sucesión:

$$X_1(\omega) = \begin{cases} 1, & 0 \leq \omega \leq 1. \\ 0, & \text{en otro caso.} \end{cases}$$

$$X_2(\omega) = \begin{cases} 1, & 0 \leq \omega \leq \frac{1}{2}. \\ 0, & \text{en otro caso.} \end{cases}$$

Entonces $X_1(\omega), X_2(\omega), \ldots$ es una sucesión decreciente.

P.d. que $P\,[X_n \rightarrow 0] = 1.$

Como $X_1(\omega), X_2(\omega), \ldots$ es una sucesión decreciente.

$$\lim_{n\to\infty} X_n(\omega) = \bigcap_{n=1}^{\infty} X_n(\omega) = 0.$$

$$P\left\{\lim_{n\to\infty} X_n(\omega) = 0\right\} = 1.$$

Por la definición de convergencia casi segura.

$\forall\, n \in \mathbb{N}\;\; X_n(0) = 1$, entonces $\omega = 0$ no converge a $0$, por lo tanto la convergencia es casi segura.

Como $P$ es una probabilidad uniforme para todo el intervalo, sean $a, b \in \mathbb{R}\,[0,1]$, $a \leq b$, entonces $P(b-a) = b-a \Rightarrow P(0) = 0$.

### 3.4 Convergencia en Probabilidad.

**Definición**:

Decimos que $X_n$ converge en probabilidad a la variable aleatoria $X$ $\left(X_n \xrightarrow{p} X\right)$ cuando para $\varepsilon > 0$.

$$\lim_{n\to\infty} P\{|X_n - X| > \varepsilon\} = 0.$$

Replanteando esta definición, tenemos que para cada $\varepsilon > 0$ dado, existe $N(\varepsilon)$ tal que

$P\{|X_n - X| > \varepsilon\} < \varepsilon$ si $n > N(\varepsilon)$.

Al escribir $\{|X_n - X|\} > \varepsilon$ como $\{|X_n(\omega) - X(\omega)|\} > \varepsilon$.

Se identifica en la expresión anterior que la diferencia dada entre $X_n(\omega)$ y $X(\omega)$ en el conjunto $\Omega$ (de las $\omega$‘s) si excede de $\varepsilon$, su probabilidad es igual a cero.

Nota: No se establece que $\{|X_n(\omega) - X(\omega)|\} > \varepsilon \longrightarrow \emptyset$.

La convergencia en probabilidad toma en cuenta que la convergencia de una sucesión de números reales.

$$P\{|X_n - X| > \varepsilon\}\;\; n \geq 1.$$

Y no la convergencia de una sucesión de eventos.

También se conoce como convergencia en medida.

## 3.5 Convergencia en Distribución.

**Definición**:

La sucesión $x_1, x_2, \dots$ converge en distribución a $x$ si:

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x).$$

$\forall x$ en donde $F_{X_n}(x)$ es continua.

Notación:

$$X_n \xrightarrow{d} X.$$

Esta es la convergencia más débil (la que presenta menos restricciones).

Ejemplo:

Sea $x_1, x_2, \dots$ una sucesión de variables aleatorias tales que $X_n \sim N\left(0, \frac{\sigma^2}{n}\right)$.

P.d. que $X_n \xrightarrow{d} 0$.

Si $X \sim N\left(0, \frac{\sigma^2}{n}\right)$ y si $n \to \infty$, Entonces $Var \to 0$.

## 3.5 Convergencia en r-media.

**Definición**:

La sucesión $x_1, x_2, \dots$ converge en media a $x$ si:

$\lim_{n\to\infty}(X_n - X) = 0,$ $E\,|X_n - X| \to 0$ Cuando $n \to \infty$.

Nota: En este tipo de convergencia se requiere que cada elemento de la sucesión y el límite mismo deben ser variables aleatorias con esperanzas finitas.

Notación:

$X_n \xrightarrow{m} X$ ó $X_n \xrightarrow{l_1} X$.

**Definición**:

La sucesión $x_1, x_2, \ldots$ converge en media cuadrática a $x$ si:

$$\lim_{n\to\infty} |X_n - X|^2 = 0.$$

Nota: En este tipo de convergencia, se requiere que cada elemento de la sucesión y el límite mismo deben ser variables aleatorias con segundos momentos finitos si $x = 0 \Rightarrow$ es el segundo momento no central.

Notación:

$X_n \overset{m.c.}{\longrightarrow} X$ ó $X_n \overset{l_2}{\longrightarrow} X.$

**Definición**:

La sucesión $x_1, x_2, \ldots$ converge en r-media a $x$ si:

$$\lim_{n\to\infty} |X_n - X|^r = 0.$$

Nota: En este tipo de convergencia se requiere que cada elemento de la sucesión y el límite mismo deben ser variables aleatorias con r-esimos momentos finitos.

Notación:

$X_n \overset{r.m.}{\longrightarrow} X$ ó $X_n \overset{l_r}{\longrightarrow} X.$

**Proposición 43**:

Convergencia casi segura implica convergencia en probabilidad pero convergencia casi segura no implica convergencia en probabilidad.

Demostración:

$\Rightarrow)$ Supongamos que $X_n \overset{c.s.}{\longrightarrow} X.$

Sea $\varepsilon > 0$ y definamos los siguientes eventos:

$$A_n = \bigcup_{k=n}^{\infty} |X_k - X| > \varepsilon.$$

Entonces

$$A_1 = \bigcup_{k=1}^{\infty} |X_k - X| > \varepsilon.$$

$$A_2 = \bigcup_{k=1}^{\infty} |X_k - X| > \varepsilon.$$

$$\vdots$$

Entonces como $A_1 \supset A_2 \supset \ldots \Rightarrow A_n$ es no creciente.

Entonces su límite es igual a la $\cap A_n$.

Por otro lado $(|X_n - X| > \varepsilon) \subseteq A_n$.

Dado que $(|X_n - X| > \varepsilon)$ es solo un elemento de la sucesión está contenido en $A_n$.

Por las proposiciones anteriores

$$P(|X_n - X| > \varepsilon) \leq P(A_n).$$

Entonces

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) \leq \lim_{n\to\infty} P(A_n).$$

Por la continuidad de la probabilidad $(P)$.

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) \leq P\left(\lim_{n\to\infty} A_n\right).$$

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) \leq P\left(\bigcap_{n=1}^{\infty} A_n\right).$$

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) \leq P(|X_n - X| > \varepsilon) \text{ para cada } n \geq 1.$$

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) \leq P\left(\lim_{n\to\infty} X_n \neq X\right) = 0.$$

Por definición de convergencia casi segura.

Implica Convergencia en Probabilidad y por lo tanto existe convergencia en probabilidad.

∎

$\not\Leftarrow$) Ejemplo de que convergencia en probabilidad no implica convergencia casi segura.

Considérese el espacio de probabilidad $([0,1], \mathfrak{B}([0,1]), P)$ donde $P$ es medida uniforme.

Defina los eventos:

$$\begin{cases} A_1 = \left(0, \dfrac{1}{2}\right) \ A_2 = \left(\dfrac{1}{2}, 1\right). \\ A_3 = \left(0, \dfrac{1}{3}\right) \ A_4 = \left(\dfrac{1}{3}, \dfrac{2}{3}\right) \ A_5 = \left(\dfrac{2}{3}, 1\right). \\ A_6 = \left(0, \dfrac{1}{4}\right) \ A_7 = \left(\dfrac{1}{4}, \dfrac{1}{2}\right) \ A_8 = \left(\dfrac{1}{2}, \dfrac{3}{4}\right) \ A_9 = \left(\dfrac{3}{4}, 1\right). \\ \vdots \\ \vdots \\ A_{55}\left(0, \dfrac{1}{11}\right) \ A_{56}\left(\dfrac{1}{11}, \dfrac{2}{11}\right) \ A_{57}\left(\dfrac{2}{11}, \dfrac{3}{11}\right) \ A_{58}\left(\dfrac{3}{11}, \dfrac{4}{11}\right) \ \ldots \end{cases}$$

Esta sucesión tiende a los intervalos $(0,0)$ cuando $n \to \infty$.

Sea $X_n = \begin{cases} 1, \ \omega \in A_n. \\ 0, \ \omega \notin A_n. \end{cases}$

Para $X_1$ Tomamos la sucesión $A_1$.

$$P_{X_1}\left\{0 < X_1 < \frac{1}{2}\right\} = \frac{1}{2}.$$

Ya que la probabilidad es uniforme.

$$P_{X_1}\left\{X_1 > \frac{1}{2} \quad ó \quad X_1 \leq 0\right\} = \frac{1}{2}.$$

Para $X_5$ Tomamos la sucesión $A_5$.

$$P_{X_5}\left\{\frac{2}{3} < X_5 < 1\right\} = \frac{1}{3}.$$

$$P_{X_5}\left\{X_5 \geq 1 \quad ó \quad X_5 \leq \frac{2}{3}\right\} = \frac{2}{3}.$$

Sabemos que convergencia en probabilidad.

$P(|X_n - X| > \varepsilon) \to 0$ y $X_n \to 0$.

Entonces $X_n \xrightarrow{p} X$ , pues para cualquier $\varepsilon > 0$.

$$\lim_{n\to\infty} |X_n - X| > \varepsilon \;=\; \lim_{n\to\infty} P(A_n) = 0.$$

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1.$$

Por lo tanto no existe él límite y por lo tanto no hay convergencia casi segura.

Entonces la sucesión no converge casi seguramente porque,

$$\left\{\omega \in \Omega\colon \lim_{n\to\infty} X_n(\omega) = X(\omega) \; \exists \right\} = \emptyset.$$

∎

Ejemplo (convergencia en media no implica convergencia casi segura).

Demostración:

Consideramos la sucesión del ejemplo anterior.

Si existe convergencia en media entonces $E\,|X_n - X| \to 0$.

$X_n \overset{c.m}{\to} X$ , pues $E\,|X_n - 0| = \lim_{n\to\infty} \frac{1}{k_n} \to 0$.

Así que implica que se cumple para la sucesión la convergencia en media.

$P\left(\lim_{n\to\infty} X_n = 0\right) = P(\emptyset) = 0$ por lo tanto no existe convergencia casi segura.

Tomando en cuenta que el límite nunca va a tender a cero porque solo toma valores $0$ ó $1$ entonces es el conjunto vacio.

∎

Un espacio de Hilbert es un espacio vectorial dotado de un producto interior $\langle\ ,\ \rangle$ que es completo para la norma.

$$\|V\| = \langle v, v\rangle^{\frac{1}{2}}.$$

**Teorema de Pitágoras.**

Sean $(x, y) \in (E, \langle\ ,\ \rangle)$ ortogonales, entonces

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

**Corolario. Desigualdad de Cauchy-Schwarz.**

Sean $(x, y) \in (\mathrm{E}, \langle\ ,\ \rangle)$.

Entonces

$$\|\langle x, y\rangle\| \leq \|x\|\,\|y\|.$$

Decimos que $\mathrm{H}$ es un espacio de Hilbert si es completo con respecto a la norma $\|\cdot\|$ ó $\langle\ ,\ \rangle$.

Completo en este contexto significa que cualquier sucesión de Cauchy, tal que $\forall\ \varepsilon > 0\,,\ \ \exists\, \mathrm{N} \in \mathbb{Z}$ tal que $\forall\ \ \mathrm{n}, \mathrm{m} > N$.

$$|\mathrm{X_m} - \mathrm{X_n}| < \varepsilon.$$

Sucesión convergente.

Ejemplo (convergencia casi segura no implica convergencia en media).

Considere el espacio de probabilidad $([0,1], \mathfrak{B}([0,1]), \mathrm{P})$ donde $\mathrm{P}$ es medida de probabilidad uniforme.

Defina la sucesión $\mathrm{X_n} = \mathrm{nI}\left(0, \frac{1}{\mathrm{n}}\right)$,

$$\mathrm{I}\left(0, \frac{1}{\mathrm{n}}\right) = \begin{cases} 1, & \omega \in \left(0, \dfrac{1}{\mathrm{n}}\right). \\ \\ 0, & \omega \notin \left(0, \dfrac{1}{\mathrm{n}}\right). \end{cases}$$

$$\mathrm{I}\left(0, \frac{1}{\mathrm{n}}\right) = \begin{cases} 1, & \omega \in \left(0, \dfrac{1}{\mathrm{n}}\right). \\ \\ 0, & \omega \notin \left(0, \dfrac{1}{\mathrm{n}}\right). \end{cases}$$

$\mathrm{X_n}$ converge casi seguramente a cero cuando $\mathrm{n} \to \infty$ ya que existen saltos en el intervalo abierto.

$$\lim_{\mathrm{n}\to\infty} \left(0, \frac{1}{\mathrm{n}}\right) \to (0, 0).$$

Entonces

$$\mathrm{P}\left(\lim_{\mathrm{n}\to\infty} \mathrm{X_n} = 0\right) = 1 = \mathrm{P}(\Omega).$$

$$\mathrm{E}\,|\mathrm{X_n} - 0| = \mathrm{E}\,|\mathrm{X_n}| = \mathrm{n}\left(\frac{1}{\mathrm{n}}\right) = 1 \nrightarrow\ 0.$$

Por lo tanto no existe Convergencia en media.

**Definición**:

Una función es convexa si y solo si la región sobre su grafo es un conjunto convexo.

**Definición**:

Una función convexa definida en el intervalo abierto $C$ , es continua en $C$ y diferenciable en todos los puntos numerables.

Nota: Una región es convexa por trayectoria, si todos los puntos de una línea que conectan dos puntos de la región están en dicha región.

**Proposición 44**:

Convergencia en media cuadrática implica Convergencia en media.

Demostración:

Por la desigualdad de Jensen.

Sabemos que si $f(x)$ es una función convexa, entonces

$$\mathrm{E}[f(x)] \geq f\ [\mathrm{E}(x)].$$

Si $f(x) = x^2 \;\Rightarrow\; \mathrm{E}(x^2) \geq \mathrm{E}^2(x)$ .

Trabajando convergencia en media cuadrática.

$\mathrm{E}^2(\mathrm{X_n} - \mathrm{X}) \leq \mathrm{E}(\mathrm{X_n} - \mathrm{X})^2$.

Si existe convergencia en media cuadrática, entonces $\mathrm{E}|\mathrm{X_n} - \mathrm{X}| \to 0$ cuando $\mathrm{n} \to \infty$.

Entonces $\mathrm{E}^2|\mathrm{X_n} - \mathrm{X}| = \mathrm{E}|\mathrm{X_n} - \mathrm{X}|\ \mathrm{E}|\mathrm{X_n} - \mathrm{X}| \to 0$ conforme $\mathrm{n} \to \infty$.

Entonces $\mathrm{E}|\mathrm{X_n} - \mathrm{X}| \to 0$ por lo tanto existe convergencia.

∎

Otro enfoque.

Vale la pena recordar la desigualdad de Cauchy-Schwarz.

Sean $x$ e $y$ variables aleatorias en el espacio vectorial $(\Omega,\ \mathcal{F},\ \mathcal{P})$.

$$\mathrm{E}[xy] \leq \mathrm{E}[x^2]\mathrm{E}[y^2].$$

Demostración:

Sea $Q(\lambda) = \mathrm{E}(x - \lambda y)^2$ donde $Q(\lambda) \geq 0\ \forall\ \lambda > 0$.

$$\frac{\delta Q(\lambda)}{\delta \lambda} = \frac{\delta}{\delta \lambda}\ \mathrm{E}(x^2 - 2\lambda xy + \lambda^2 y^2) = \mathrm{E}(-2xy + 2\lambda y^2).$$

$$\frac{\delta^2 Q(\lambda)}{\delta \lambda} = \frac{\delta}{\delta \lambda}\ \mathrm{E}(-2xy + 2\lambda y^2) = \mathrm{E}(2y^2) \geq 0.$$

Esto siempre será mayor o igual a cero debido a que esta elevado al cuadrado.

Entonces el valor que se obtenga para $\lambda$ es un mínimo (porque $> 0$ ) si igualamos $\mathrm{E}(-2xy + 2\lambda y^2)$ a cero para obtener el valor de $\lambda$.

$$\mathrm{E}(-2xy + 2\lambda y^2) = -2\mathrm{E}(xy) + 2\lambda\mathrm{E}(y^2) = 0.$$

$$\lambda_0 = \frac{\mathrm{E}(xy)}{\mathrm{E}(y^2)}.$$

Entonces

$$Q(\lambda_0) = \mathrm{E}\left[x - \frac{\mathrm{E}(xy)}{\mathrm{E}(y^2)}\ y\right]^2 = \mathrm{E}\left[x^2 - 2\frac{xy\ \mathrm{E}(xy)}{\mathrm{E}(y^2)} + \frac{\mathrm{E}^2(xy)}{\mathrm{E}^2(\mathrm{y}^2)}\ y^2\right]$$

$$= \mathrm{E}(x^2) - 2\frac{\mathrm{E}(xy)}{\mathrm{E}(y^2)}\ \mathrm{E}(xy) + \frac{\mathrm{E}^2(xy)}{\mathrm{E}^2(\mathrm{y}^2)}\ \mathrm{E}(y^2)$$

$$= \mathrm{E}(x^2) - 2\frac{\mathrm{E}^2(xy)}{\mathrm{E}(y^2)} + \frac{\mathrm{E}^2(xy)}{\mathrm{E}(\mathrm{y}^2)} = \mathrm{E}(x^2) - \frac{\mathrm{E}^2(xy)}{\mathrm{E}(\mathrm{y}^2)} \geq 0.$$

Entonces

$$\mathrm{E}(x^2) \geq \frac{\mathrm{E}^2(xy)}{\mathrm{E}(\mathrm{y}^2)}.$$

Entonces

$$\mathrm{E}^2(xy) \leq \mathrm{E}(x^2)\ \mathrm{E}(y^2).$$

∎

**Proposición 45**:

Convergencia en media implica Convergencia en probabilidad.

Demostración:

Para cada $\varepsilon > 0$ defina el evento.

$$A_n = (|X_n - X| > \varepsilon).$$

Y sean

$$I_{A_n}(\omega) = \begin{cases} 1, & \omega \in A_n. \\ 0, & \omega \notin A_n. \end{cases}$$

$$I^c_{A_n}(\omega) = \begin{cases} 1, & \omega \notin A_n. \\ 0, & \omega \in A_n. \end{cases}$$

Entonces $E|X_n - X| = E(|X_n - X| \cdot I_{A_n}) + E(|X_n - X| \cdot I^c_{A_n})$.

Entonces $E|X_n - X| \geq E(|X_n - X| \cdot I_{A_n})$.

Ahora, como se supone que $\exists$ convergencia en media.

Entonces

$$E|X_{\underset{n\to\infty}{n}} - X| \to 0.$$

Entonces por definición de valor esperado.

$$E|X_{\underset{n\to\infty}{n}} - X| \geq \varepsilon\, P\,[|X_n - X| > \varepsilon].$$

$$\varepsilon\, P\,[|X_n - X| > \varepsilon] \leq E|X_n - X|.$$

Entonces

$$\lim_{n\to\infty} P\,[|X_n - X| > \varepsilon] \to 0.$$

Entonces existe Convergencia en probabilidad.

∎

**Proposición 46**:

Convergencia en probabilidad implica Convergencia en distribución.

Demostración:

Sea $x$ un punto de continuidad de $\mathrm{F_X}(x)$.

Por definición

$$\mathrm{F_{X_n}}(x) = \mathrm{P}(\mathrm{X_n} \leq \mathrm{X}).$$

$$\int_{\mathrm{a}}^{\mathrm{b}} f_{\mathrm{X}}(x)dx = \mathrm{P}\{\mathrm{a} \leq x \leq \mathrm{b}\} = \mathrm{F_X}(\mathrm{a}) - \mathrm{F_X}(\mathrm{b}).$$

Entonces para cualquier $\varepsilon > 0$, es posible descomponer el conjunto $\{\mathrm{X_n} \leq x\}$ en dos conjuntos disjuntos, pues la unión de disjuntos es la suma de las probabilidades.

Entonces $\{\mathrm{X_n} \leq \mathrm{X}\} = \{(\mathrm{X_n} \leq \mathrm{X});\ |\mathrm{X_n} - \mathrm{X}| < \varepsilon\} + \{(\mathrm{X_n} \leq \mathrm{X});\ |\mathrm{X_n} - \mathrm{X}| \geq \varepsilon\}$.

Primer caso cuando $\mathrm{X_n} < X$.

Esta sucesión es creciente pues $\mathrm{X}$ es más grande que $\mathrm{X_n}$ el n-esimo termino de la sucesión.

Aplicando la medida de probabilidad:

$$\mathrm{P}\{\mathrm{X_n} \leq \mathrm{X}\} = \mathrm{P}\{\mathrm{X_n} \leq \mathrm{X} + \varepsilon\} + \mathrm{P}\{|\mathrm{X_n} - \mathrm{X}| > \varepsilon\}.$$

$$\mathrm{F_{X_n}}(x) = \mathrm{F_X}(x + \varepsilon) + \mathrm{P}\{|\mathrm{X_n} - \mathrm{X}| > \varepsilon\}.$$

Como por el supuesto inicial hay convergencia en probabilidad y esto implica que esto tiende a cero cuando $\mathrm{n} \to \infty$.

$$\limsup_{\mathrm{n}\to\infty} \mathrm{F_{X_n}}(x) \leq \mathrm{F_X}(x).$$

Segundo caso cuando $\mathrm{X_n} \geq \mathrm{X}$.

$$\{\mathrm{X} \leq \mathrm{X} - \varepsilon\}, \qquad |\mathrm{X_n} - \mathrm{X}| \leq \varepsilon \ \cup\ \{(\mathrm{X} \leq \mathrm{X} - \varepsilon);\ |\mathrm{X_n} - \mathrm{X}| > \varepsilon\}$$

$$= \{\mathrm{X_n} \leq \mathrm{X}\} \cup \{|\mathrm{X_n} - \mathrm{X}| > \varepsilon\}.$$

Aplicando la medida de Probabilidad.

$$\mathrm{F_X}(x - \varepsilon) = \mathrm{P}\ \{\mathrm{X} \leq \mathrm{X} - \varepsilon\} = \mathrm{P}\{\mathrm{X_n} \leq \mathrm{X}\} + \mathrm{P}\{|\mathrm{X_n} - \mathrm{X}| > \varepsilon\}.$$

$$\liminf_{n\to\infty} F_X(x-\varepsilon) \leq \liminf_{n\to\infty} P\{X_n \leq X\}.$$

$$F_X(x-\varepsilon) \leq \liminf_{n\to\infty} F_{X_n}(x).$$

Como $x$ es un punto de discontinuidad, $\varepsilon \to 0$.

Entonces $F_X(x) \leq \liminf_{n\to\infty} F_{X_n}(x)$.

Usando el resultado del primer caso tenemos que:

$$\limsup_{n\to\infty} F_{X_n}(x) \leq F_X(x) \leq \liminf_{n\to\infty} F_{X_n}(x).$$

Entonces $\limsup_{n\to\infty} F_{X_n}(x) = \liminf_{n\to\infty} F_{X_n}(x)$.

Por teoremas anteriores.

Entonces $\lim_{n\to\infty} F_{X_n}(x) = F_X(x)$.

Por lo tanto existe Convergencia en distribución.

∎

Ejemplo Convergencia en distribución no implica Convergencia en Probabilidad.

Sea $X$ variable aleatoria tal que $X \sim N(0,1)$.

$$X_n = \begin{cases} x, & \text{n es par.} \\ -x, & \text{n es impar.} \end{cases}$$

Entonces $X_n \sim N(0,1)$.

$\forall\ x \in \mathbb{R}$ con el supuesto de que $F_{X_n}(x) \to F_X(x)$ cuando $n \to \infty$.

$$X_n \xrightarrow{d} X.$$

Por demostrar que existe convergencia en probabilidad $X_n \xrightarrow{p} X$.

Entonces para este caso.

$$P\{|X_n - X| > \varepsilon\} = P\{2|X_n - 0| > \varepsilon\} \geq \frac{1}{2}.$$

$$P\{|X_n - 0| > \varepsilon\} \nrightarrow 0.$$

Por lo tanto no existe Convergencia en probabilidad.

∎

**Teorema de Convergencia Monótona.**

Sea $0 \leq X_1 \leq X_2 \leq \cdots$ una sucesión de variables aleatorias que converge casi seguramente a la variable aleatoria $X$, entonces

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

Demostración:

$$0 \leq X_n \leq X.$$

$$E(X_n) \leq E(X).$$

$$\lim_{n\to\infty} E(X_n) \leq \lim_{n\to\infty} E(X).$$

$$\lim_{n\to\infty} E(X_n) \leq E(X).$$

Sea $\varepsilon > 0$ para cada entero $k \geq 0$.

$$A_k = (k\varepsilon \leq X < (k+1)\varepsilon).$$

Donde $A_k$ es una colección de elementos disjuntos a pares y además su unión es igual a $\Omega$.

Definimos $Y(\omega) = k\varepsilon$.

Si $k\varepsilon \leq X(\omega) < (k+1)\varepsilon$.

$$Y \leq X \leq Y + \varepsilon \Rightarrow X - \varepsilon < Y \leq X.$$

$$E(X) - \varepsilon < E(Y) \leq E(X).$$

$$B_n = (X_n \geq Y).$$

$\nearrow$ (Significa que tiende al límite superior).

$$B_n \nearrow \Omega.$$

$$A_k \cap B_n \nearrow A_k \quad n \to \infty.$$

$$A_k \cap \Omega \nearrow A_k \quad n \to \infty.$$

$$Y\, I_{B_n}(\omega) = \begin{cases} Y(\omega), & \omega \in B_n. \\ 0, & \omega \notin B_n. \end{cases}$$

$$0 \leq Y\, I_{B_n}(\omega) \leq X_n.$$

Aplicando operador lineal esperanza.

$$E\left(Y\, I_{B_n}(\omega)\right) \le E(X_n).$$

$$\lim_{n\to\infty} E(X_n) \ge \lim_{n\to\infty} E\left(Y\, I_{B_n}(\omega)\right).$$

$A_k$ es una partición de $\Omega$, pues

$$\bigcup_{k=1}^{\infty} A_k = \Omega.$$

Además, cuando $n \to \infty$, $B_n \nearrow \Omega$, entonces

$$\lim_{n\to\infty} E(X_n) \ge \lim_{n\to\infty} E\left(Y\, I_{B_n}(\omega)\right) = \lim_{n\to\infty} \sum_{k=0}^{\infty} E\left(Y\, I_{B_n \cap A_k}(\omega)\right)$$

$$= \lim_{n\to\infty} \sum_{k=0}^{\infty} k\,\varepsilon\, P(B_n \cap A_k)$$

$$\ge \lim_{n\to\infty} \sum_{k=0}^{m} k\,\varepsilon\, P(B_n \cap A_k) = \sum_{k=0}^{m} k\,\varepsilon\, P(A_k)$$

$$\lim_{n\to\infty} E(X_n) \ge \sum_{k=0}^{m} k\,\varepsilon\, P(A_k) > E(X) - \varepsilon$$

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

∎

**Teorema de convergencia Dominada.**

Sea $X_1, X_2, \ldots$ una sucesión de variables aleatorias para la cual existe otra variable $Y$ integrable tal que $|X_n| \le Y$ para $n \ge 1$.

Si $\lim_{n\to\infty} X_n = X$ entonces $X \wedge X_n$ son integrables y

$$\lim_{n\to\infty} E(X_n) = E(x).$$

Demostración:

Sea $Y_n = \inf\{X_n, X_{n+1}, \ldots\}$, entonces

$$Y_n \nearrow X.$$

$$Y_n + Y \nearrow X + Y.$$

Entonces $E[Y_{n+Y}] \nearrow E[X+Y]$.

Estas esperanzas existen porque $Y$ es integrable.

$$\{Y_n + Y\} \geq 0.$$

$$Y_n + Y \leq Y_{n+1} + Y \leq \cdots$$

Es decreciente ya que $Y_n$ es el infimo.

Definiendo una nueva variable aleatoria para hacer una cota superior.

$$Z_n = \sup\{X_n, X_{n+1,} \dots\}.$$

$Z_n \searrow X$ cuando $n \to \infty$.

$$(Y - Z_n) \nearrow (Y - X).$$

Donde $Y - Z_n > 0,$ pues $X_n \leq Y$ por hipótesis.

$Z_n \leq Y$ Por el teorema de convergencia monótona.

$$E(Y - Z_n) \nearrow E(Y - X).$$

Donde

$$E(Z_n) \searrow E(X).$$

Notemos que $Y_n \leq X_n \leq Z_n$.

Entonces $E(Y_n) \leq E(X_n) \leq E(Z_n)$.

Y obteniendo el límite de las esperanzas.

$$\lim_{n\to\infty} E(Y_n) \leq \lim_{n\to\infty} E(X_n) \leq \lim_{n\to\infty} E(Z_n).$$

$$E(X) \leq \lim_{n\to\infty} E(X_n) \leq E(X).$$

Por la propiedad de Tricotomía, entonces

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

∎

**Desigualdad de Markov.**

Sea $X \geq 0$ una variable aleatoria con $E(X) < \infty$ , para cualquier $\varepsilon > 0$.

$$P(X \geq \varepsilon) \leq \frac{E(X)}{\varepsilon}.$$

Demostración:

$$E(X) = E\left(XI_{(X\geq\varepsilon)} + XI_{(X<\varepsilon)}\right)$$

$$\geq E\left(XI_{(X\geq\varepsilon)}\right) \geq E\left(\varepsilon I_{(X\geq\varepsilon)}\right)$$

$$= \varepsilon\, E\left(XI_{(X\geq\varepsilon)}\right) = \varepsilon\, P(X \geq \varepsilon).$$

Entonces

$$E(X) \geq \varepsilon\, P(X \geq \varepsilon).$$

Entonces

$$\frac{E(X)}{\varepsilon} \geq P(X \geq \varepsilon).$$

**3.7 Desigualdad del tipo Chebyshev.**

Sea $X$ variable aleatoria con media $\mu$ y variable finita $\sigma^2$. Para cualquier $\varepsilon > 0$.

$$P[|X - \mu| \geq \varepsilon] \leq \frac{\sigma^2}{\varepsilon^2}.$$

Demostración:

Sabemos que $\sigma^2 = E[(X - \mu)^2]$.

Particionado con indicadoras:

$$\sigma^2 = E\left[(X - \mu)^2\, I_{(X-\mu)\geq\varepsilon} + (X - \mu)\, I_{(X-\mu)<\varepsilon}\right]$$

$$\geq E\left[(X - \mu)^2\, I_{(X-\mu)^2\geq\varepsilon}\right]$$

$$\geq E\left[\varepsilon^2\, I_{(X-\mu)\geq\varepsilon}\right] = \varepsilon^2 E\left[\, I_{(X-\mu)\geq\varepsilon}\right]$$

$$= \varepsilon^2 P[\,(X - \mu) \geq \varepsilon].$$

$$\frac{\sigma^2}{\varepsilon^2} \geq P[\,(X - \mu) \geq \varepsilon].$$

∎

**Desigualdad de Chebyshev extendida.**

Sea $X$ variable aleatoria, $g > 0$ no decreciente tal que $E[g(X)] < \infty$, $\varepsilon > 0$, entonces

$$P(X \geq \varepsilon) \leq \frac{E[g(X)]}{g(\varepsilon)}.$$

Demostración:

$$\begin{aligned} E[g(X)] &= E\left[g(X)I_{(X\geq\varepsilon)} + g(X)I_{(X<\varepsilon)}\right] \\ &\geq E\left[g(X)I_{(X\geq\varepsilon)}\right] \\ &\geq E\left[g(\varepsilon)I_{(X\geq\varepsilon)}\right] \\ &= g(\varepsilon)P(X \geq \varepsilon). \end{aligned}$$

Por lo tanto

$$\frac{E[g(X)]}{g(\varepsilon)} \geq \ P(X \geq \varepsilon).$$

∎

**Desigualdad de Kolmogorov.**

Sean $X_1, \ldots, X_n$ variables aleatorias independientes idénticamente distribuidas con $E(X_i) = 0$ y $E(X_i^2) < \infty$, $\varepsilon > 0$. Entonces

$$P\left[\max_k\{|X_1 + \cdots + X_n| \geq \varepsilon\}\right] \leq \frac{1}{\varepsilon^2}\sum_{k=1}^{n} \mathrm{Var}(X_k).$$

Demostración:

Sea

$$S_k = \sum_{i=1}^{k} S_i \ \Rightarrow E(S_k) = 0.$$

$S_n \perp S_n - S_k$ con $n > k$.

Entonces $E[(S_k)(S_n - S_k)] = E(S_k)\, E(S_n - S_k) = 0$

Definiendo eventos (sucesión de eventos).

$$A_k = (|S_k| \geq \varepsilon) \cap \bigcap_{i=1}^{k-1} (|S_i| < \varepsilon); \quad A = \bigcup_{k=1}^{n} A_k.$$

$$E[S_n^2] \geq E\left[S_n^2\, I_{\{A\}}\right] = \sum_{k=1}^{n} E\left[S_n^2\, I_{\{A_k\}}\right]$$

$$= \sum_{k=1}^{\infty} E\left[[S_k + (S_n - S_k)]^2\, I_{\{A_k\}}\right]$$

$$= \sum_{k=1}^{n} E\left[S_k^2 + 2[S_k(S_n - S_k)] + (S_n - S_k)^2\, I_{\{A_k\}}\right]$$

$$\geq \sum_{k=1}^{n} E\left[S_k^2\, I_{\{A_k\}}\right]$$

$$\geq \sum_{k=1}^{n} E\left[\varepsilon^2\, I_{\{A_k\}}\right]$$

$$= \sum_{k=1}^{n} \varepsilon^2\, E\left[\, I_{\{A_k\}}\right]$$

$$= \sum_{k=1}^{n} \varepsilon^2\, P[\, A_k]$$

$$= \varepsilon^2\, P(A).$$

Tenemos ahora que:

$$V(S_n) = E(S_n^2) - \ [E(S_n)^2]$$

$$= \mathrm{Var}\,(X_1 + \cdots + X_n).$$

Y por independencia de las variables aleatorias

$$V(S_n) = V(X_1) + \cdots + V(X_n)$$

$$= \sum_{i=1}^{n} V(X_i).$$

Por lo tanto

$$\frac{1}{\varepsilon} \sum_{k=1}^{n} V(X_k) \ \geq P(A).$$

∎

## 3.8 Ley Débil y Fuerte de los Grandes Números.

**Teorema Ley Débil de los Grandes Números (Bernoulli).**

Sean $X_1, \dots, X_n$ variables aleatorias independientes idénticamente distribuidas con $E(X) = \mu$ , $V(X) = \sigma^2$. Entonces

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu.$$

Demostración:

Sea

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

Entonces

$$E(S_n) = \frac{1}{n} E\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n} n\mu = \mu.$$

$$V(S_n) = V\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2} V\left[\sum_{i=1}^{n} X_i\right] = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Ahora, por la desigualdad de Chebyshev,

$$P(|X - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

$$P(|S_n - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2}.$$

Ahora obteniendo el límite de ambos lados,

$$\lim_{n\to\infty} P(|S_n - \mu| \geq \varepsilon) \leq \lim_{n\to\infty} \frac{\sigma^2}{n\varepsilon^2} = 0.$$

Y como la probabilidad está acotada por $0 \leq P \leq 1$.

Entonces $0 \leq \lim_{n\to\infty} P(|S_n - \mu| \geq \varepsilon) \leq 0$.

Y por lo tanto

$$\lim_{n\to\infty} P\left(|S_n - \mu| \geq \varepsilon\right) = 0.$$

Entonces

$$S_n \xrightarrow{p} \mu.$$

Por lo tanto

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu.$$

∎

Ejemplo:

Sea $A$ un evento cualquiera

$$X_k = \begin{cases} 1, & \textit{si ocurre } A. \\ 0, & \text{si no ocurre A.} \end{cases}$$

Notamos que $X_k \sim \mathrm{Ber}(p)$.

$E[X_k] = p$. $\qquad V[X_k] = pq$.

$$\frac{1}{n}\sum_{k=1}^{n} X_k \to p.$$

Tendríamos entonces:

$$\frac{X_1 + \cdots + X_n}{n} = \frac{\text{casos favorables}}{\text{casos totales}} = p.$$

**Teorema Ley Fuerte de los Grandes Números**

Sean $X_1, \ldots, X_n$ variables aleatorias independientes idénticamente distribuidas con $E(X) = \mu$, entonces

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{\text{c.s.}} \mu.$$

Demostración:

(Suponemos que la varianza y el cuarto momento existen).

$$\mathrm{E}[(\mathrm{X}-\mu)^2] = \sigma^2.$$

$$\mathrm{E}(\mathrm{X}-\mu) = 0.$$

$$\mathrm{E}\left[\sum_{i=1}^{\mathrm{n}}(\mathrm{X}_i-\mu)^4\right] = \mathrm{n}\,\mathrm{E}[(\mathrm{X}_i-\mu)^4] + 3\mathrm{n}\,(\mathrm{n}-1)\sigma^4.$$

Podemos denotar a $\mathrm{X}_i = \mathrm{X}$.

Por Lema de Borel Cantelli.

$$\lim\sup \mathrm{A_n} = \mathrm{A}, \quad \text{si} \quad \sum_{\mathrm{n}=1}^{\infty}\mathrm{P}(\mathrm{A_n}) < \infty \;\Rightarrow\; \mathrm{P}(\mathrm{A}) = 0.$$

Por la desigualdad de Chebyshev extendida.

$\mathrm{g} > 0$ no decreciente tal que $\mathrm{E}[\mathrm{g}(\mathrm{X})] < \infty$ y para $\varepsilon > 0$. Entonces

$$\mathrm{P}\,[\mathrm{X} \geq \varepsilon] \leq \frac{\mathrm{E}[\mathrm{g}(\mathrm{X})]}{\mathrm{g}(\varepsilon)}.$$

Sea

$$\mathrm{g}(\mathrm{X}) = \mathrm{X}^4 \qquad \mathrm{y} \qquad \left|\sum_{i=1}^{\mathrm{n}}(\mathrm{X}_i-\mu)\right|.$$

$$\varepsilon = \mathrm{n}\varepsilon'.$$

Entonces

$$\mathrm{P}\left[\left|\sum_{i=1}^{\mathrm{n}}(\mathrm{X}_i-\mu)\right| \geq \mathrm{n}\varepsilon'\right] \leq \frac{1}{(\mathrm{n}\varepsilon')^4}\;\mathrm{E}\left[\left(\sum_{i=1}^{\mathrm{n}}(\mathrm{X}_i-\mu)\right)^4\right] = \frac{\mathrm{n}\,\mathrm{E}[(\mathrm{X}_i-\mu)^4] + 3\mathrm{n}\,(\mathrm{n}-1)\sigma^4}{\mathrm{n}^4\varepsilon'^4}.$$

Definimos

$$\mathrm{A_n} = \left(\frac{1}{\mathrm{n}}\left|\sum_{i=1}^{\mathrm{n}}(\mathrm{X}_i-\mu)\right| > \varepsilon'\right).$$

$$\lim_{\mathrm{n}\to\infty}\mathrm{A_n} = \mathrm{A} = \lim_{\mathrm{n}\to\infty}\frac{1}{\mathrm{n}}\left|\sum_{i=1}^{\mathrm{n}}(\mathrm{X}_i-\mu)\right| > \varepsilon'.$$

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

Entonces

$$P(A) = 0.$$

Entonces

$$P(A^c) = 1.$$

$$P\left(\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu) \leq \varepsilon\right) = 1.$$

Entonces

$$P\left(\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (X_i) \leq \varepsilon + \mu\right) = 1.$$

Entonces

$$P\left(\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} (X_i) = \mu\right) = 1.$$

Por lo tanto

$$\frac{1}{n} \sum_{i=1}^{n} X_i \overset{c.s.}{\longrightarrow} \mu.$$

∎

## Capítulo 4 FUNCIÓN CARACTERÍSTICA.

### 4.1 Funciones Generatrices.

**Definición**:

La Función Generadora de Probabilidad de una variable aleatoria discreta $\mathrm{X}$ es la función.

$$\mathrm{G_X(t) = E[t^{\mathit{x}}]}.$$

Definida para valores reales de $\mathrm{t}$ y tales que el valor esperado sea absolutamente convergente. Entonces

$$\mathrm{G_X(t)} = \sum_{\mathrm{k}=0}^{\infty} \mathrm{t^k\, P_X(k)}.$$

Para asegurar la convergencia de esta serie $|\mathrm{t}| < 1$, entonces

$$|\mathrm{G_X(t)}| \leq \sum_{\mathrm{k}=0}^{\infty} |\mathrm{t}|^{\mathrm{k}}\, \mathrm{P_X(k)} \leq \sum_{\mathrm{k}=0}^{\infty} \mathrm{P_X(k)} = 1.$$

Entonces el radio de convergencia de $\mathrm{G_X(t)}$ es $1$.

**Definición**:

$$\mathrm{P_X}(x) = \frac{\left.\frac{\delta^{\mathrm{k}} \mathrm{G_X(t)}}{\delta \mathrm{t^k}}\right|_{\mathrm{t}=0}}{\mathrm{k}!}.$$

Sabemos que

$$\mathrm{G_X(t) = E[t^{\mathit{x}}]}.$$

$$\frac{\delta^{\mathrm{k}} \mathrm{G_X(t)}}{\delta \mathrm{t^k}} = \frac{\delta}{\delta \mathrm{t}}\, \mathrm{E}[\mathrm{t}^{x}] = \mathrm{E}\left[\frac{\delta \mathrm{t}^{x}}{\delta \mathrm{t}}\right] = \left.\frac{\mathrm{E}[x\mathrm{t}^{x-1}]}{1!}\right|_{\lim\ \mathrm{t}\to 0}.$$

Ejemplo:

Sea $X$ variable aleatoria discreta tal que $X \sim P_0(\lambda)$. Obtener $G_X(t)$

$$P_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}.$$

$$G_X(t) = E[t^x] = \sum_{x=0}^{\infty} t^x \, \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(t\lambda)^x}{x!} = e^{-\lambda} e^{\lambda t} = e^{-\lambda(1-t)}.$$

Entonces

$$G_X(t) = e^{-\lambda(1-t)}.$$

$$G'_X(t) = e^{-\lambda} e^{\lambda t} \, \lambda = \left. \frac{e^{\lambda(t-1)}\lambda}{1!} \right|_0 = \frac{e^{-\lambda}\,\lambda}{1!} = P_X(1).$$

$$G''_X(t) = e^{-\lambda}\lambda e^{\lambda t} \, \lambda = \left. \frac{e^{\lambda(t-1)}\lambda^2}{2!} \right|_0 = \frac{e^{-\lambda}\,\lambda^2}{2!} = P_X(2).$$

**Proposición 47**:

Propiedades de la función generadora de probabilidad.

1) Sean $X$ e $Y$ variables aleatorias discretas tal que $G_X(t)$ e $G_Y(t)$ existen y coinciden en algún intervalo alrededor de $t = 0$. Entonces $X$ e $Y$ tienen la misma distribución de probabilidad.

Demostración:

Para cada $k \geq 0$, sea $a_k = P_X(k)$. $b_k = P_Y(k)$.

Si $G_X(t) = G_Y(t)$.

Entonces

$$\sum_{k=0}^{\infty} t^x a_k = \sum_{k=0}^{\infty} t^x b_k \Rightarrow a_k = b_k \quad \forall\, k.$$

Entonces $P_X(k) = P_Y(k)$ entonces $X$ e $Y$ tienen la misma distribución.

∎

2) Si el n-ésimo momento factorial de $X$ existe, entonces

$$\lim_{t\to 1} \frac{d^n}{dt^n} G_X(t) = E[(x)(x-1)(x-2)\ldots(x-n+1)].$$

Demostración:

Sabemos que

$$G_X(t) = \sum_{k=0}^{\infty} t^k P_X(k).$$

$$\frac{d}{dt} G_X(t) = \frac{d}{dt} \sum_{k=0}^{\infty} t^k P_X(k) = \sum_{k=0}^{\infty} \frac{d}{dt} t^k P_X(k)$$

$$= \sum_{k=1}^{\infty} t^{k-1} P_X(k).$$

$$\lim_{t\to 1} \frac{d}{dt} G_X(t) = \sum_{k=1}^{\infty} k\ P_X(k) = E(X).$$

$$\lim_{t\to 1} \frac{d^2}{dt^2} G_X(t) = \lim_{t\to 1} \left[\sum_{k=1}^{\infty} k\,(k-1) t^{k-2}\ P_X(k)\right]$$

$$= \lim_{t\to 1} \left[\sum_{k=2}^{\infty} k\,(k-1) t^{k-2}\ P_X(k)\right]$$

$$= \sum_{k=2}^{\infty} k\,(k-1)\ P_X(k) = E[(X)(X-1)].$$

Al seguir desarrollando,

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

∎

3) Sean $X$ e $Y$ variables aleatorias independientes, con función generadora de probabilidad $G_X(t)$ y $G_Y(t)$ respectivamente.

Entonces $G_{X+Y}(t) = G_X(t) G_Y(t)$.

Demostración:

Si $X$ e $Y$ son independientes, entonces por definición:

$$G_{X+Y}(t) = E[t^{X+Y}] = E[t^X t^Y] = E[t^X]E[t^Y] = G_X(t)G_Y(t).$$

∎

Ejemplo:

Sean $X$ e $Y$ variables aleatorias discretas e independientes tal que $X\sim Po(\lambda_1)$ y $X\sim Po(\lambda_2)$ ¿Cómo se distribuye $X+Y$?

$$G_{X+Y}(t) = G_X(t)G_Y(t) = e^{-\lambda_1(1-t)}e^{-\lambda_2(1-t)} = e^{-(\lambda_1+\lambda_2)(1-t)}$$

Entonces

$$(X+Y)\sim Po(\lambda_1+\lambda_2).$$

**Definición.**

La función generadora de momentos (no centrales o con respecto al origen) de la variable aleatoria $X$ es la función.

$$\mu_X(t) = E(e^{t_x}).$$

Definida para valores reales de $t$ tales que la esperanza es absolutamente convergente.

Nota: $\mu_X(t) = G_X(e^t)$.

**Función Característica:**

**Definición**:

La función característica de la variable aleatoria $X$ es la función.

$$\varphi_X(t) = E\left(e^{itx}\right).$$

Definida para cualquier número real $t$, $i = \sqrt{-1}$.

Nota: Siempre existe, a diferencia de la función generadora de probabilidad y de la función generadora de momentos.

Transformación.

$$X \to e^{it\mathrm{x}}.$$

Donde $X$ es el valor real.

$e^{it\mathrm{x}}$ variable aleatoria con números complejos en la forma:

$$e^{it\mathrm{x}} = \cos(\mathrm{tx}) + i\,\mathrm{sen}(\mathrm{tx}).$$

Ejemplos:

1) $X \sim \mathrm{Bin}(n, p)$ Obtener $\varphi_x(t)$.

$$\varphi_x(t) = E\left(e^{itx}\right) = \sum_{x=0}^{\infty} e^{itx} \binom{n}{x} p^x\, q^{n-x}$$

$$= \sum_{x=0}^{\infty} (e^{it}\, p)^x \binom{n}{x} q^{n-x}$$

Por el teorema del binomio.

$$= (e^{it}\, p + q)^n.$$

∎

2) $X \sim N(\mu, \sigma^2)$ Obtener $\varphi_x(t)$.

$$\varphi_x(t) = E\left(e^{itx}\right) = \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{itx}\, e^{-\frac{1}{2\sigma^2}\left(x^2-2x\mu+\mu^2\right)} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx-\frac{1}{2\sigma^2}(x-\mu)^2}\, dx$$

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{itx2\sigma^2-x^2+2x\mu-\mu^2}{2\sigma^2}} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}\left(x-\left(\mu+it\sigma^2\right)\right)^2} dx.$$

$$= e^{-\frac{\sigma^2 t^2}{2}+it\mu} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2\sigma^2}\left(x-\left(\mu+it\sigma^2\right)\right)^2} dx = e^{-\frac{\sigma^2 t^2}{2}+it\mu}.$$

Entonces $\varphi_{x}(t)$ de $N(0,1) = e^{-\frac{t^2}{2}}$ .

Entonces $X \sim N(\mu + it\sigma^2,\ \sigma^2)$.

Así que la media es variable aleatoria compleja.

∎

3) $X \sim Po$ Obtener $\varphi_{x}(t)$.

$$\varphi_{x}(t) = \sum_{x=0}^{\infty} \frac{e^{itx}\,\lambda^{x}\,e^{-\lambda}}{x!} = e^{-\lambda}\sum_{x=0}^{\infty} \frac{(e^{it}\lambda)^{x}}{x!} = e^{-\lambda}\,e^{\lambda e^{it}}.$$

$$e^{\lambda(e^{it}-1)}.$$

$$\varphi_{x}(t) = e^{-\lambda(1-e^{it})}.$$

4) $X \sim G(r,\theta)$ Obtener $\varphi_{x}(t)$.

$$\Gamma(r) = \int_{0}^{\infty} x^{n-1}e^{-x}\ dx.$$

$$\varphi_{x}(t) = \int_{0}^{\infty} \frac{e^{itx}\theta^{r}\,x^{r-1} - e^{-\theta x}}{\Gamma(r)} = \int_{0}^{\infty} \frac{e^{itx-\theta x}\theta^{r}\,x^{r-1}}{\Gamma(r)}$$

$$= \int_{0}^{\infty} \frac{e^{x(it-\theta)}\theta^{r}\,x^{r-1}}{\Gamma(r)}\ dx = \int_{0}^{\infty} \frac{(\theta - it)^{r}\ x^{r-1}\,e^{-x(it-\theta)}}{\Gamma(r)} = \frac{\theta^{r}}{(\theta - it)^{r}} = \left(\frac{\theta}{\theta - it}\right)^{r}.$$

∎

**Teorema de la existencia de la función característica.**

Para cualquier número real $t$, $|\varphi_x(t)| \leq 1$ y en particular, $\varphi_x(0) = 1$.

Demostración:

Para cualquier número real $t$.

$$|\varphi_x(t)| = \left| \int_{-\infty}^{\infty} e^{itx}\, dF_x(x) \right| \leq \int_{-\infty}^{\infty} |e^{itx}|\, dF_x(x).$$

La norma de $e^{itx} = \cos(tx) + i\,\mathrm{sen}(tx)$ esta entre $0$ y $1$.

$$\leq \int_{-\infty}^{\infty} dF_x(x) \leq 1.$$

Nota: $\varphi_x(t)$ es número complejo de módulo $\leq 1$ para cualquier $t \in \mathbb{R}$.

**Proposición 48**:

Sea $x$ variable aleatoria con $n$ –ésimo momento finito, entonces

1) $\frac{d^n}{dt^n}\, \varphi_x(t)\big|_{t=0} = i^n\, E(x^n)$.

Demostración:

Por el teorema de la convergencia dominada.

$$\frac{d^n}{dt^n}\, \varphi_x(t) = \frac{d}{dt}\, E\big(e^{itx}\big) = E\left(\frac{d}{dt}\, e^{itx}\right)$$

$$= E\left(ix\, e^{itx}\right)\Big|_{t=0} = E(ix) = i\, E(x).$$

Generalizando:

$$\frac{d^n}{dt^n}\, \varphi_x(t) = E\big((ix)^n\, e^{itx}\big)\Big|_{t=0} = E(i^n\, x^n) = i^n\, E(x^n).$$

2) Si $x$ e $y$ son variables aleatorias independientes.

$$\varphi_{x+y}(t) = \varphi_x(t)\,\varphi_y(t).$$

Demostración:

Por definición,

$$\varphi_{x+y}(t) = E\left[e^{it(x+y)}\right] = E\left[e^{itx}\,e^{ity}\right]$$

$$= E\left[e^{itx}\right]\,E\left[e^{ity}\right] = \varphi_x(t)\,\varphi_y(t).$$

Método de la función de distribución.

1) Definir la región de integración de la variable aleatoria, $x = y_1 + y_2 + \cdots + y_n$ en el espacio $(y_1,\ y_2, \dots, y_n)$.

2) Determinar la región $\{X \leq x\}$.

3) Se obtiene $P\{X \leq x\} = F_x(x)$.

4) Se obtiene $f_x(x) = \frac{dF_x(x)}{dx}$ Teorema fundamental del calculo.

## 4.2 Teorema de Inversión.

**Formula de inversión de Levy.**

Sea $x$ variable aleatoria con función de distribución $F_x(x)$ y función característica $\varphi_x(t)$. Si $x < y$ son puntos de continuidad de $F_x(x)$.

Entonces

$$P\{x \leq X \leq y\} = F_x(y) - F_x(x) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{it}\,\varphi_x(t)\,dt.$$

**Teorema de Unicidad:**

Si $x$ e $y$ son variables aleatorias, entonces si $\varphi_x(t) = \varphi_y(t)$ para todo $t$ en donde está definida la función característica, $x \wedge y$ tienen la misma distribución de probabilidad.

Demostración:

Si $\varphi_x(t) = \varphi_y(t)$.

Por la formula de Levy.

$$F_x(y) - F_x(x) = F_y(y) - F_y(x).$$

$$\lim_{x \to -\infty}[F_x(y) - F_x(x)] = \lim_{x \to -\infty}[F_y(y) - F_y(x)].$$

Entonces

$$F_x(y) = F_y(y)$$

∎

Nota: Para toda $y$ que sean puntos de continuidad de ambas distribuciones.

Teorema de inversión en el caso absolutamente continuo.

Sea $x$ variable aleatoria absolutamente continua con función de probabilidad $f_x(x)$ y función característica $\varphi_x(t)$. Entonces,

$$f_x(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \ \varphi_x(t) \, dt.$$

Demostración:

$$P\{x \leq X \leq y\} = F_x(y) - F_x(x).$$

Sabemos que $F_x(x)$ es una primitiva de $f_x(x)$ y que por la formula de Levy para que $x < y$ dos puntos de continuidad de $F$.

$$F_x(y) - F_x(x) = \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{it} \ \varphi_x(t) \, dt.$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\frac{e^{-itx} - e^{-ity}}{it}\right] \varphi_x(t) \, dt.$$

$$\int_{y}^{x} f_x(x) \ dx = F_x(y) - F_x(x) = = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\int_{y}^{x} e^{-itx} \ dx\right] \varphi_x(t) \, dt.$$

Y por el Teorema de Fubinni,

$$\int_{y}^{x}\left[\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\ \varphi_x(t)\ dt\right] dx.$$

$$f_x(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\ \varphi_x(t)\ dt.$$

∎

Nota: Solo es válido si $x$ es absolutamente continua.

Para establecer la equivalencia entre la convergencia en distribución y la convergencia puntual de la función característica, tenemos.

**Teorema de continuidad.**

Sean $x, x_1, x_2, x_3, \ldots$ variables aleatorias, entonces $X_n \overset{c.d}{\rightarrow} X$ si y solamente si $\varphi_{X_n}(t) \longrightarrow \varphi_x(t)$.

Demostración:

$\Rightarrow$) Supongamos que $X_n \overset{d}{\rightarrow} X$ , entonces por el teorema de convergencia dominada y con

$\varphi_x(t) = \cos\ tx + i\text{sen tx}.$

$\varphi_{X_n}(t) = E[\cos\ tX_n + i\text{sen } tX_n].$

$$\lim_{n\to\infty} \varphi_{X_n}(t) = \lim_{n\to\infty} E(\cos\ tX_n) + \lim_{n\to\infty} E(i\text{sen } tX_n).$$

$$\lim_{n\to\infty} \varphi_{X_n}(t) = E\left[\lim_{n\to\infty} \cos\ tX_n\right] + E\left[\lim_{n\to\infty} i\text{sen } tX_n\right].$$

$$= E[\cos\ tx] + i\ E[\text{sen tx}].$$

$$= E[\cos\ tx + i\ \text{sen tx}] = E\left[e^{itx}\right] = \varphi_x(t).$$

$\Leftarrow$) Supongamos que $\varphi_{X_n}(t) \longrightarrow \varphi_x(t)$

Entonces para dos puntos de continuidad $x < y$ comunes a $F_{X_n}$ y $F_x$, por la formula de Levy,

$$F_{X_n}(y) - F_{X_n}(x) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{it} \varphi_x(t)\, dt$$

$$= \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-itx} - e^{-ity}}{it} \lim_{n\to\infty} \varphi_{X_n}(t)\, dt.$$

$$\lim_{n\to\infty} \left[F_{X_n}(y) - F_{X_n}(x)\right] = F_x(y) - F_x(x).$$

$$\lim_{x\to-\infty} \left[\lim_{n\to\infty}\left[F_{X_n}(y) - F_{X_n}(x)\right]\right] = \lim_{x\to-\infty}\left[F_x(y) - F_x(x)\right].$$

Entonces

$$\lim_{n\to\infty} F_{X_n}(y) = F_x(y).$$

Por lo tanto existe convergencia en distribución.

∎

**4.3 Teorema del Límite Central.**

Sea $X_1, X_2, X_3, \ldots$ variables aleatorias independientes e idénticamente distribuidas tales que $E(X_i) = \mu$ y $\mathrm{Var}(X_i) = \sigma^2 < \infty$.

Para cualquier $x \in \mathbb{R}$.

$$\lim_{n\to\infty} P\left[\frac{(X_1 + X_2 + \ldots + X_n) - n\mu}{\sqrt{n}\ \sigma} \leq X\right] = P\,\{Z \leq X\}.$$

En donde $Z \sim N(0,1)$.

Si $(x_1 + x_2 + \ldots + x_n)$ es una variable aleatoria con media $n\mu$ y varianza $n\sigma^2$, entonces podemos expresar el teorema como:

$$\frac{\frac{1}{n}(X_1 + X_2 + \ldots + X_n) - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow{d} N(0,1).$$

$$E(X_1 + X_2 + \cdots + X_n) = \|\text{independencia}\| = \mu + \mu + \cdots + \mu = n\mu.$$

$$\text{Var}\,(X_1 + X_2 + \cdots + X_n) = \|\text{indep} \Rightarrow \text{Cov} = 0\|$$

$$= \text{Var}(X_1) + \text{Var}(X_2) + \cdots + \text{Var}(X_n) = \sigma^2 + \cdots + \sigma^2 = n\sigma^2.$$

Demostración:

Supóngase que todos los elementos de la sucesión tienen momentos finitos de cualquier orden. Entonces

$$\frac{(X_1 + X_2 + \ldots + X_n) - n\mu}{\sqrt{n}\ \sigma} = \frac{X_1 - \mu}{\sqrt{n}\ \sigma} + \frac{X_2 - \mu}{\sqrt{n}\ \sigma} + \cdots + \frac{X_n - \mu}{\sqrt{n}\ \sigma}.$$

$\frac{(X_i - \mu)}{\sigma}$ es la estandarización y esta multiplicada por $\frac{1}{\sqrt{n}}$.

Supóngase que cada $X_i$ tiene una distribución con $\mu = 0$ y $\text{Var} = 1$.

Entonces sea $Z_n = \frac{(X_1 + X_2 + \ldots + X_n)}{\sqrt{n}}$ variable aleatoria.

Basta con demostrar que $Z_n \xrightarrow{d} N(0,1)$.

Entonces

$$\varphi_{Z_n}(t) = e^{-\frac{t}{2}}.$$

$$\varphi_{Z_n}(t) = E\left[e^{it\frac{(X_1 + X_2 + \ldots + X_n)}{\sqrt{n}}}\right].$$

$$= E\left[e^{\frac{it\,X_1}{\sqrt{n}}}\, e^{\frac{it\,X_2}{\sqrt{n}}} \ldots e^{\frac{it\,X_n}{\sqrt{n}}}\right].$$

$$\varphi_{Z_n}(t) = \left[\varphi_X\left(\frac{t}{\sqrt{n}}\right)\right]^n.$$

$$\text{Ln}\ \varphi_{Z_n}(t) = n\ \text{Ln}\left[\varphi_X\left(\frac{t}{\sqrt{n}}\right)\right].$$

$$\text{Ln}\ \varphi_{Z_n}(t) = n\ \text{Ln}\left[1 + \frac{it\,E(X)}{\sqrt{n}} + \frac{i^2 t^2\ E(X^2)}{n\ 2!} + \cdots\right].$$

Sabemos que

$$\varphi_X\left(\frac{t}{\sqrt{n}}\right) = E\left[e^{\frac{itX}{\sqrt{n}}}\right].$$

Y desarrollando $e^{\frac{itX}{\sqrt{n}}}$ en series de potencias tenemos:

$$e^{\frac{itX}{\sqrt{n}}} = \frac{i^0 t^0 X^0}{\left(\sqrt{n}\right)^0 0!} + \frac{itX}{\sqrt{n}} + \frac{i^2 t^2 X^2}{n\ 2!} + \cdots$$

Entonces

$$E\left(e^{\frac{iXt}{\sqrt{n}}}\right) = 1 + \frac{it\ E(X)}{\sqrt{n}} + \frac{i^2 t^2 E(X^2)}{n\ 2!} + \cdots$$

Recordemos que el desarrollo de,

$$\mathrm{Ln}\ (1+t) = t - \frac{1}{2}t^2 + \frac{1}{3}t^3 - \ \ldots$$

Entonces

$$\mathrm{Ln}\ \varphi_{Z_n}(t) = \left[\frac{it\ E(X)}{\sqrt{n}} + \frac{i^2 t^2 E(X^2)}{n\ 2!} + \cdots\right] - \frac{1}{2}\left[\frac{it\ E(X)}{\sqrt{n}} + \frac{i^2 t^2 E(X^2)}{n\ 2!} + \cdots\right]^2 \ldots$$

$$\mathrm{Ln}\ \varphi_{Z_n}(t) = \frac{E(X)}{\sqrt{n}} it + [E(X^2) - E^2(X)]\frac{i^2 t^2}{2\ n} + \left(\frac{E(X^3)}{3!\ n} - \frac{E(X)}{2\ n} + \frac{E^2(X)}{3\ n}\right)\frac{i^3 t^3}{\sqrt{n}} + \cdots$$

Como $E(X) = 0$ y $\mathrm{Var}(X) = 1$, entonces

$$\mathrm{Ln}\ \varphi_{Z_n}(t) = \frac{i^2 t^2}{2\ n} + \left(\frac{E(X^3)}{3!\ n} - \frac{E(X)}{2\ n} + \frac{E^2(X)}{3\ n}\right)\frac{i^3 t^3}{\sqrt{n}} + (\ldots)\frac{i^4 t^4}{n} + \ \ldots$$

$$\lim_{n\to\infty} \mathrm{Ln}\ \varphi_{Z_n}(t) = \lim_{n\to\infty} \mathrm{Ln}\ \frac{i^2 t^2}{2\ n} + \left(\frac{E(X^3)}{3!n} - \frac{E(X)}{2\ n} + \frac{E^2(X)}{3\ n}\right)\frac{i^3 t^3}{\sqrt{n}} + (\ldots)\frac{i^4 t^4}{n} + \ \ldots = \frac{i^2 t^2}{2}.$$

Entonces

$$\lim_{n\to\infty} \mathrm{Ln}\ \frac{i^2 t^2}{2\ n} + \left(\frac{E(X^3)}{3!\ n} - \frac{E(X)}{2\ n} + \frac{E^2(X)}{3\ n}\right)\frac{i^3 t^3}{\sqrt{n}} + (\ldots)\frac{i^4 t^4}{n} + \ \ldots = \frac{i^2 t^2}{2}.$$

$$= \| i^2 = -1 \| = -\frac{t^2}{2}.$$

Y como $\mathrm{Ln}$ es una función continua, la podemos sacar del límite.

Entonces

$$\mathrm{Ln} \lim_{n\to\infty} \varphi_{Z_n}(t) = -\frac{t^2}{2}.$$

Entonces

$$\lim_{n\to\infty} \varphi_{Z_n}(t) = e^{\frac{-t^2}{2}}.$$

$$\Rightarrow Z_n \xrightarrow{d} N(0,1)$$

∎

**CONCLUSIONES.**

El propósito de este trabajo fue desarrollar material de apoyo didáctico actualizado para la impartición de la materia de Probabilidad II que se imparte en la carrera de Actuaría. Este trabajo aporta de manera significativa material de estudio, de complemento y apoyo para los estudiantes con dificultades dentro de la comprensión de los temas en la materia, inclusive retomando en el principio temas muy fundamentales y básicos, haciendo un repaso de temas de Probabilidad I, para así poder fundamentar los temas que se retoman en Probabilidad II.

Los principales temas de estudio en el campo de la Probabilidad que son de gran interés para el alumno, se pueden aprender fácil y rápidamente ya que este trabajo es una síntesis de lo más importante, esperando que sea de utilidad para futuras generaciones de actuarios en el estudio de la Probabilidad, los que accedan al presente proyecto identificarán que sin ser una guía exhaustiva, es una referencia o simplemente un punto de partida para la comunidad que este decidida a hacerlo durante su trayecto de estudios en probabilidad presentes o futuros.

El objetivo de la materia de Probabilidad II se integro en el desarrollo del presente documento puesto que se consideraron los temas descritos en el contenido del Programa de Estudios de Probabilidad II, complementándose con temas de interés actual. Por lo tanto sin decirlo en forma peyorativa, se puede concluir que se cumplió con el objetivo principal de este trabajo.

## FUENTES DE CONSULTA.


### BIBLIOGRAFIA BÁSICA.

1. ASH R.B. **Real Analysis and Probability.** Academic Press. USA; 1996.
2. BHAT B.R. **Modern probability theory.** Wiley. India; 1995.
3. CLARKE L.E. **Random Variables.** Longman. Hungary; 1995.
4. DOMINGUEZ M. J.I. **Diseño y Análisis de Modelos de Probabilidad.** Editorial Iberoamérica. México; 2001.
5. GNEDENKO B. V. **The Theory of probability.** Chelsea Publishing Company. USA; 1996.
6. HARRIS B. **Theory of Probability.** Addison-Wesley. USA; 1996.
7. HERNÁNDEZ A. F. M. **Cálculo de Probabilidades.** Sociededad Matématica Mexicana. Mexico; 2003.
8. PAPOULIS A. **Probability, random variables and stochastic processes.** McGraw-hill. USA; 1991.
9. RINCÓN S.L. **Curso Intermedio de Probabilidad.** Editorial Las Prensas de Ciencias UNAM; 2007.

### BIBLIOGRAFÍA COMPLEMENTARIA

1. APOSTOL T. **Análisis Matemático.** Editorial Reverté. España; 1996.
2. BOLAND P.J. **Statical and Probabilistic Methods in Actuarial Sciences.** CRC Press. Taylor and Francis Group; 2007.
3. BOYER C. **Historia de la matemática.** Editorial Alianza. España; 1996.
4. FELLER W. **Introducción a la Teoría de Probabilidades y sus aplicaciones.** Editorial Limusa. México; 1995.
5. HERNANDEZ-LERMA O. y HERNANDEZ DEL VALLE A. **Elementos de Probabilidad y Estadística.** Sociedad Matemática Mexicana. México; 2003.
6. LANDRO A. **Acerca de la Probabilidad.** Ediciones C.E.C.E. Universidad de Buenos Aires. Argentina; 1999.
7. LOÉVE M. **Teoría de la Probabilidad.** Tecnos. España; 1976.
8. MEYER P.L. **Probabilidad y Aplicaciones Estadísticas.** Addison-Wesley Iberoamericana. México; 1986.
9. OBREGÓN S.I. **Teoría de la Probabilidad.** Editorial Limusa. México; 1980.
10. PARDO L., QUESADA V. **Curso Superior de Probabilidades.** Editorial PPU. España; 1987.

11. PARZEN E, **Teoría Moderna de Probabilidades y sus Aplicaciones.** Limusa. México; 1997.
12. SARABIA A.J.M., GOMEZ D.E. y VAZQUEZ P.F.J. **Estadística Actuarial. Teoría y Aplicaciones.** Pearson Educación. España; 2007.

**REFERENCIAS ELECTRONICAS.**

1. Actuary and Actuarial Jobs, Science Information Actuary,

Última consulta: 9 de marzo de 2012.

http://www.actuary.com

2. Be An Actuary-Actuarial Examinations,

Última consulta: 9 de marzo de 2012.

http://www.beanactuary.org/exams/

3. Colegio Nacional de actuarios,

Última consulta: 9 de marzo de 2012.

http://www.conac.org.mx

4. International Actuarial Association (IAA),

Última consulta: 9 de marzo de 2012.

http://www.actuaries.org

5. Probability.com,

Última consulta: 9 de marzo de 2012.

http://www.probability.com/

6. Society of Actuaries (SOA),

Última consulta: 9 de marzo de 2012.

http://www.soa.org