▷ Si $A_1, A_2, \dots$ es una colección de conjuntos definida dentro de un conjunto Ω, entonces

$$\bigcup_{i=1}^{\infty} A_i = \{x \in \Omega : x \in A_i \text{ para algún } i\}$$
$$\bigcap_{i=1}^{\infty} A_i = \{x \in \Omega : x \in A_i \text{ para todo } i\}$$

Definición 1.4

Una secuencia $A_1, A_2, \dots$ de subconjuntos de un conjunto Ω se llama **partición** de Ω si

i) $\cup_{i=1}^{\infty} A_i = \Omega$ (exhaustivos), y

ii) $A_i \cap A_j = \emptyset \, \forall \, i \neq j$ (mutuamente excluyentes)

Definición 1.6

El límite de una secuencia monótona se define por:

i) Si $A_n \uparrow$, entonces $\lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n$

ii) Si $A_n \downarrow$, entonces $\lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n$

Definición 1.7

Una colección $\mathcal{A}$ de subconjuntos de un conjunto Ω (no vacío) constituye una σ - **algebra (sigma algebra)**, si satisface los tres siguientes axiomas:

A1) $\Omega \in \mathcal{A}$ (el conjunto Ω es un elemento de $\mathcal{A}$)

A2) Si $A \in \mathcal{A}$, entonces $A^c \in \mathcal{A}$ ($\mathcal{A}$ es cerrada bajo complemento)

A3) si $A_1, A_2, \ldots \in \mathcal{A}$, entonces $\cup_{i=1}^{\infty} A_i \in \mathcal{A}$ ($\mathcal{A}$ es cerrada bajo uniones contables)

Definición 1.8

Sea $\mathcal{A}$ una σ - algebra de subconjuntos de Ω:

▷ Al par $(\Omega, \mathcal{A})$ se le denomina **espacio medible** o **espacio de sucesos**.

▷ Si $A \in \mathcal{A}$, se dice que A es medible.

Ejemplo 1.4

1) $\mathcal{A} = \{\emptyset, \Omega\}$: σ - algebra trivial

2) $\mathcal{A} = \{\emptyset, A, A^c, \Omega\}$ es σ - algebra para todo subconjunto A de Ω

3) $\mathcal{A} = \{$ todos los subconjuntos de $\Omega\}$: $\mathcal{P}(\Omega)$ o 2^{Ω} σ - algebra

Ejemplo 1.5

Si $\Omega = \mathbb{R} = (-\infty, \infty)$ es la recta real, entonces eligimos $\mathcal{A}$ de modo que contenga a todos los intervalos de la forma,

$$[a, b], \quad (a, b], \quad (a, b), \quad [a, b) \quad -\infty < a < b < \infty.$$

Además, de las propiedades de una σ - algebra sigue que $\mathcal{A}$ contiene a todos los subconjuntos de $\mathbb{R}$ que se pueden formar tomando uniones e intersecciones (posiblemente infinitas) de los intervalos anteriores. En este caso, $\mathcal{A} \equiv \mathcal{B}$ se llama σ - algebra de Borel, y sus elemento se llaman Borelianos.

Nota: La extensión para los Borelianos en $\mathbb{R}^n$ es similar (remplazando los intervalos por rectángulos), y la σ - algebra correspondiente se denota como $\mathcal{B}_n$.

Teorema 1.1

Sea $\Omega \neq \emptyset$. Si $\mathcal{A}_1, \mathcal{A}_2, \cdots$ son σ−algebras sobre Ω, entonces $\bigcap_{i=1}^{\infty} \mathcal{A}_i$ también es una σ−algebra sobre Ω.

Definición 1.6

σ**−Algebra de Borel:** La σ−algebra más pequeña sobre $\mathbb{R}$ generada por la clase de intervalos de la forma $(-\infty, a]$, $a \in \mathbb{R}$, se llama σ−**algebra de Borel**, y se denota por $\mathcal{B}$ (o $\mathcal{B}(\mathbb{R})$).

Si $A \in \mathcal{B}$, se dice que A es un **boreliano** (o subconjunto de Borel de $\mathbb{R}$).

Ya que $\mathcal{B}$ es una σ−algebra, si tomamos $a, b \in \mathbb{R}$ con $a < b$, entonces los siguientes subconjuntos de $\mathbb{R}$ son borelianos:

$$(a, \infty) = \mathbb{R} \backslash (-\infty, a]; \quad (a, b] = (-\infty, b] \cap (a, \infty);$$
$$(-\infty, a) = \bigcup_{n=1}^{\infty} \left(-\infty, a - \frac{1}{n}\right];$$
$$[a, \infty) = \mathbb{R} \backslash (-\infty, a); \quad (a, b) = (-\infty, b) \cap (a, \infty);$$
$$[a, b] = \mathbb{R} \backslash ((-\infty, a) \cup (b, \infty)); \quad \{a\} = [a, a]$$

Definición 1.8

Espacio Medible: Sea $\Omega \neq \emptyset$ y $\mathcal{A}$ una σ−algebra sobre Ω. Al par $(\Omega, \mathcal{A})$ se le denomina **espacio medible o espacio de sucesos**.

De la definición de σ−algebra, queda claro que ambos Ω y $\emptyset$ son eventos ya que pertenecen a cualquier σ−algebra definida sobre Ω.

$\emptyset$ se llama **evento imposible**, Ω se llama **evento seguro**, y un evento de la forma $\{\omega\}$ con $\omega \in \Omega$ se llama **evento simple** o **elemental**.

Decimos que el evento A de Ω ocurre si al de realizar el experimento aleatorio, el resultado obtenido pertenece al conjunto A.

En este sentido, si A y B son dos eventos de Ω, entonces:

i) El evento $A \cup B$ ocurre si y solo si A o B o ambos ocurren,

ii) El evento $A \cap B$ ocurre si y solo si A y B ocurren,

iii) El evento A^c ocurre si y solo si A no ocurre,

iv) El evento $A \backslash B$ ocurre si y solo si A ocurre pero B no.

Definición 1.9

Eventos mutuamente excluyentes: Dos eventos A y B se dicen mutuamente excluyentes o incompatibles si $A \cap B = \emptyset$.

Definición 1.10

Frecuencia relativa: Para cada evento A, el número $fr(A) = \frac{n(A)}{n}$ se llama frecuencia relativa de A, donde $n(A)$ indica la cantidad de veces que ocurrió el evento A en n repeticiones del experimento.

Lamentablemente, para cada evento A fijo, $fr(A)$ no es constante: su valor depende de n; sin embargo, se ha observado que cuando un experimento aleatorio se repite en las mismas condiciones un gran número de veces, la frecuencia relativa $fr(A)$ se estabiliza alrededor de un valor específico entre 0 y 1.

Definición 1.11

Medida de Probabilidad Sea $(\Omega, \mathcal{A})$ un espacio medible. Una función real valorada P definida sobre $\mathcal{A}$ que satisface los siguientes axiomas:

A1) $P(A) \geq 0$ para todo $A \in \mathcal{A}$ (no-negativa)

A2) $P(\Omega) = 1$ (normalizada)

A3) Si $A_1, A_2, \cdots \in \mathcal{A}$ son eventos (2 a 2) disjuntos ($A_i \cap A_j = \emptyset \, \forall \, i \neq j$), entonces $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ (aditividad contable)

se llama **medida de probabilidad** sobre $(\Omega, \mathcal{A})$.

La terna $(\Omega, \mathcal{A}, P)$ se llama **modelo o espacio de probabilidad**, y el número $P(A)$ se denomina **probabilidad** del evento $A \in \mathcal{A}$.

Teorema 1.2

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad. Entonces:

P1) $P(\emptyset) = 0$

P2) Si $A, B \in \mathcal{A}$ y $A \cap B = \emptyset$, entonces $P(A \cup B) = P(A) + P(B)$

P3) Para cualquier $A \in \mathcal{A}$, $P(A^c) = 1 - P(A)$

P4) Si $A \subseteq B$, entonces $P(A) \leq P(B)$ y $P(B \backslash A) = P(B) - P(A)$. En particular $P(A) \leq 1 \overset{(A1)}{\Longrightarrow} 0 \leq P(A) \leq 1$ para todo $A \in \mathcal{A}$

P5) Para cualquier $A, B \in \mathcal{A}$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

P6a) Si $\{A_n\}_{n=1}^{\infty}$ es una secuencia creciente de eventos en $\mathcal{A}$, es decir, $A_n \in \mathcal{A}$ y $A_n \subseteq A_{n+1}$ para todo $n = 1, 2, \ldots$; entonces $P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$ donde $\lim_{n\to\infty} A_n := \bigcup_{n=1}^{\infty} A_n$

P6b) Si $\{A_n\}_{n=1}^{\infty}$ es una secuencia decreciente de eventos en $\mathcal{A}$, es decir, $A_n \in \mathcal{A}$ y $A_n \supseteq A_{n+1}$ para todo $n = 1, 2, \ldots$; entonces $P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$ donde $\lim_{n\to\infty} A_n := \bigcap_{n=1}^{\infty} A_n$

Teorema 1.3

Sea $\Omega = \{\omega_1, \ldots, \omega_{N(\Omega)}\}$ un espacio muestral finito. Sea $\mathcal{A}$ cualquier σ−algebra sobre Ω. Sean $p_1, \ldots, p_n$ números no negativos que suman 1. Para cualquier $A \in \mathcal{A}$ defina $P(A)$ por

$$P(A) = \sum_{\{i : \omega_i \in A\}} p_i.$$

Entonces P es una medidad de probabilidad sobre $\mathcal{A}$.

Esto sigue siendo cierto si $\Omega = \{\omega_1, \omega_2, \ldots\}$ es un conjunto contable y $p_1, p_2, \ldots$ son números no negativos tal que $\sum_{i=1}^{\infty} p_i = 1$.

En tales casos, $(\Omega, \mathcal{A}, P)$ se llama **modelo de probabilidad discreto**.

Teorema 1.2

Sea $(\Omega, \mathcal{A}, P)$ un modelo de probabilidad. Entonces:

P3) Para cualquier $A \in \mathcal{A}$, $P(A^c) = 1 - P(A)$

P4) Si $A \subseteq B$, entonces $P(A) \leq P(B)$ (P no es decreciente)

P5) Para cualquier $A, B \in \mathcal{A}$,

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

Notas: Del axioma A1) y la propiedad P4), se desprende que:

a) Si $A \subset B$, entonces

$$P(B \backslash A) = P(A^c \cap B) = P(B - A) = P(B) - P(A)$$

b) Para todo $A \in \mathcal{A}$, $0 \leq P(A) \leq 1$

Teorema 1.4

Si P es una medida de probabilidad en $(\Omega, \mathcal{A})$, entonces:

P7) $P(A) = \sum_{i=1}^{\infty} P(A \cap B_i)$ para cualquier partición $B_1, B_2, \ldots$ de Ω

P8) $P(\cup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$ para cualquier secuencia de eventos $A_1, A_2, \ldots$ (Desigualdad de Boole)

Demostración 1.4

P7) Como $B_i \cap B_j = \emptyset \, \forall \, i \neq j$, y $\Omega = \cup_{i=1}^{\infty} B_i$, basta notar que,

$$A = A \cap \Omega = A \cap \left(\bigcup_{i=1}^{\infty} B_i\right) = \bigcup_{i=1}^{\infty} (A \cap B_i).$$

Definición 1.1

Medida de Probabilidad: Dado un espacio medible $(\Omega, \mathcal{A})$, una **medida de probabilidad** es una función real valorada P definida sobre $\mathcal{A}$ que verifica los siguientes axiomas:

A1) $P(A) \geq 0$ para todo $A \in \mathcal{A}$,

A2) $P(\Omega) = 1$,

A3) $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$ para toda secuencia contable $A_1, A_2, \cdots \in \mathcal{A}$ de eventos dos a dos disjuntos, es decir, $A_i \cap A_j = \emptyset \, \forall \, i \neq j$.

Nota: $P : \mathcal{A} \longrightarrow [0, 1]$ es tal que a cada $A \in \mathcal{A}$ le asigna un número $P(A) \in [0, 1]$ llamado probabilidad de ocurra el evento A

Nota: La terna $(\Omega, \mathcal{A}, P)$ se llama **modelo de probabilidad** o **espacio de probabilidad**.

Teorema 1.2

Sea $(\Omega, \mathcal{A}, P)$ un modelo de probabilidad. Entonces:

P3) Para cualquier $A \in \mathcal{A}$, $P(A^c) = 1 - P(A)$

P4) Si $A \subseteq B$, entonces $P(A) \leq P(B)$ (P no es decreciente)

P5) Para cualquier $A, B \in \mathcal{A}$,

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

Notas: Del axioma A1) y la propiedad P4), se desprende que:

a) Si $A \subset B$, entonces

$$P(B \backslash A) = P(A^c \cap B) = P(B - A) = P(B) - P(A)$$

b) Para todo $A \in \mathcal{A}$, $0 \leq P(A) \leq 1$

			número de muestras
muestras	↗ ordenadas	↗ sin devolución	$(N)_n$
		↘ con devolución	N^n
	↘ no ordenadas	↗ sin devolución	$\binom{N}{n}$
		↘ con devolución	$\binom{N+n-1}{n}$

$$(N)_n := \frac{N!}{(N-n)!} = N(N-1)\cdots(N-n+1) \quad \text{(variaciones s/r)}$$

Muestreo ordenado sin reemplazo: En este caso el espacio muestral es $\Omega = \{(s_1, s_2, \cdots, s_n) : s_i \in \{1, 2, \cdots, N\}, s_i \neq s_j, \forall i, j = 1, 2, \cdots, N\}$

Aquí A_k consiste de todos los elementos de Ω que tienen exactamente k elementos menores o iguales a R (de tipo 1). Por lo tanto,

$$N(\Omega) = \binom{N}{n} \quad \text{y} \quad N(A_k) = \binom{R}{k}\binom{N-R}{n-k}$$

Así, se obtiene el **Modelo Hipergeométrico:**

$$P(A_k) = \frac{N(A_k)}{N(\Omega)} = \frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}$$

Nota: En muestreo sin reemplazo, el resultado para $P(A_k)$ no depende de si la muestra es ordenada o no-ordenada.

Muestreo ordenado con reemplazo: En este caso, cada bolita extraída se devuelve a la urna, y después de mezclarlas, se saca una nueva bolita al azar. El espacio muestral viene dado por,

$$\Omega = \{(s_1, s_2, \cdots, s_n) : s_j \in \{1, 2, \cdots, N\}, j = 1, 2, \cdots, N\}$$

El evento A_k consiste de todas las n−tuplas de Ω con exactamente k componentes inferiores o iguales a R (de tipo 1). Entonces,

$$N(\Omega) = N^n \quad \text{y} \quad N(A_k) = \binom{n}{k} R^k (N-R)^{n-k}$$

y, en consecuencia, se obtiene el **Modelo Binomial:**

$$P(A_k) = \frac{N(A_k)}{N(\Omega)} = \binom{n}{k} p^k q^{n-k} \quad \text{donde } p = \frac{R}{N} \quad \text{y} \quad q = 1 - p.$$

Definición 1.2

Probabilidad Condicional Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad. Si $A, B \in \mathcal{A}$ con $P(A) > 0$, entonces la probabilidad condicional del evento B dado A se define como,

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

Teorema 1.1

Medida de Probabilidad Condicional: Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad, y sea $A \in \mathcal{A}$ con $P(A) > 0$. Entonces, $P(\cdot|A)$ es una medidad de probabilidad sobre $(\Omega, \mathcal{A})$ centrada en A, es decir, $P(A|A) = 1$

Teorema 1.2

I) Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad, y sea $A \in \mathcal{A}$ con $P(A) > 0$. Entonces,

1) Si $A \cap B = \emptyset$, entonces $P(B|A) = 0$

2) $P(B \cap C|A) = P(B|A \cap C)P(C|A)$ si $P(A \cap C) > 0$

II) Si $A_1, A_2, \cdots, A_n \in \mathcal{A}$ con $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$, entonces se tiene la siguiente **regla multiplicativa:**

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)\, P(A_2|A_1)\, P(A_3|A_1 \cap A_2) \cdots \times P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

Teorema 1.3

Teorema de probabilidad total Sea $A_1, A_2, \cdots$, una partición finita o contable de Ω, es decir, $A_i \cap A_j = \emptyset$ para todo $i \neq j$ y $\bigcup_{i=1}^{\infty} A_i = \Omega$ tal que $P(A_i) > 0$ para todo $A_i \in \mathcal{A}$. Entonces, para cualquier $B \in \mathcal{A}$,

$$P(B) = \sum_i P(B|A_i)\, P(A_i).$$

Demostración 1.3

Observe que, $B = B \cap \Omega = B \cap (\cup_{i=1}^{\infty} A_i) = \cup_{i=1}^{\infty} (B \cap A_i)$; luego,

$$P(B) = \sum_i P(B \cap A_i) = \sum_i P(B|A_i)\, P(A_i),$$

Corolario 1.1

Regla de Bayes Sea $A_1, A_2, \cdots$ una partición finita o contable de Ω with $P(A_i) > 0$ para todo i; entonces, para cualquier $B \in \mathcal{A}$ con $P(B) > 0$

$$P(A_i|B) = \frac{P(A_i)\, P(B|A_i)}{\sum_j P(B|A_j)\, P(A_j)} \quad \text{para todo } i$$

Demostración 1.4

De la definición de probabilidad condicional,

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)\, P(B|A_i)}{P(B)} = \frac{P(A_i)\, P(B|A_i)}{\sum_j P(B|A_j)\, P(A_j)}.$$

Definición 1.4

Eventos Independientes Dos eventos A y B se dicen independientes si y solo si,

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

Por el contrario, si no se cumple esta condición, se dice que los eventos son dependientes.

Teorema 1.4

Si A y B son eventos independientes, entonces los siguientes pares también son independientes,

i) A y B^c

ii) A^c y B

iii) A^c y B^c.

Demostración 1.5

Para i) hay que demostrar que $P(A \cap B^c) = P(A)P(B^c)$. En efecto,

$$\begin{aligned} P(A \cap B^c) &= P(A) - P(A \cap B) \\ &= P(A) - P(A)P(B) \quad (A \text{ y } B \text{ son independientes}) \\ &= P(A)(1 - P(B)) \\ &= P(A)P(B^c). \end{aligned}$$

Definición 1.3

Función de distribución acumulada La función de distribución acumulada (o acumulativa) o fda de una variable aleatoria X, denotada por $F_X(x)$, se define por

$$F_X(x) = P_X((-\infty, x]) = P(X \leq x), \quad \text{para todo } x \in \mathbb{R}.$$

Teorema 1.2

Sea X una variable aleatoria definida sobre $(\Omega, \mathcal{A}, P)$. La fda F_X satisface las siguientes propiedades,

F1 Si $x \leq y$, entonces $F_X(x) \leq F_X(y)$ (F_X es no-decreciente)

F2 Cuando $x \to \infty$, entonces $F_X(x) \to 1$, y cuando $x \to -\infty, F_X(x) \to 0$ ($0 \leq F_X(x) \leq 1 \, \forall x \in \mathbb{R}$)

F3 Cuando $h \downarrow 0, F_X(x + h) \to F_X(x)$ (F_X es continua por la derecha)

Demostración 1.2

F1 Si $x \leq y$, defina $A = \{\omega : X(\omega) \leq x\}$ y $B = \{\omega : X(\omega) \leq y\}$. Luego se tiene que $A \subseteq B$, así $P_X(X \leq x) = P(A) \leq P(B) = P_X(X \leq y)$, y el resultado sigue. Es decir, $F_X(x)$ es una función no decreciente de x.

Nota 1.1

También se cumple la implicación inversa. En efecto, puede demostrarse que cualquier función $F(x)$ que satisfaga las propiedades F1, F2 y F3 del teorema anterior, es la fda de alguna variable aleatoria X. O sea el teorema puede establecerse como: La función $F(x)$ es una fda si y sólo si se cumplen las condiciones F1, F2, F3.

Corolario 1.1

Sea X una variable aleatoria definida sobre $(\Omega, \mathcal{A}, P)$, F_X su fda y $a, b \in \mathbb{R}$ con $a < b$; entonces,

1 $F_X(x^-) = \lim_{h\to 0^+} F_X(x - h) = P_X(X < x)$.

2 $P_X(a \leq X \leq b) = F_X(b) - F_X(a^-)$.

3 $P_X(a \leq X < b) = F_X(b^-) - F_X(a^-)$.

4 $P_X(X = a) = F_X(a) - F_X(a^-)$.

Definición 1.4

Una variable aleatoria X es *continua* si $F_X(x)$ es una función continua de x. Una variable aleatoria X se dice *discreta* si $F_X(x)$ es una función escalera de x.

La F_X determina completamente la distribución de probabilidad de una variable aleatoria X. Esto es cierto si $P(X \in B)$ se define solo para eventos $B \in \mathcal{B}$, no en general (dos variables aleatorias pueden tener la misma fda pero no la misma probabilidad para cada evento). En este curso, como en la mayoría de las aplicaciones estadísticas, solo nos interesan los eventos que son intervalos, uniones contables o intersecciones de intervalos, etc.

Definición 1.5

Se dice que las variables aleatorias X y Y tienen la misma distribución (o son igualmente distribuídas), si para cada conjunto $B \in \mathcal{B}, P(X \in B) = P(Y \in B)$.

Note que dos variables aleatorias que son igualmente distribuidas no son necesariamente iguales, la definición no dice que $X = Y$.

Teorema 1.3

Las siguientes dos afirmaciones son equivalentes,

a) Las variables aleatorias X e Y son identicamente distribuídas.

b) $F_X(x) = F_Y(x)$ para cada x.

variable aleatoria X tiene *distribución de Poisson* con parámetro λ, $(\lambda > 0)$

$$P(X = n) = \frac{\lambda^n}{n!} e^{-\lambda} \qquad (n = 0, 1, 2, \ldots).$$

Binomial Negativa.

$$P(X = n) = \binom{n-1}{k-1} p^k q^{n-k}.$$

Distribución Geométrica Para la distribución geométrica la fmp está dada por,

$$f_X(x) = P(X = x) = \begin{cases} (1-p)^{x-1} p & \text{para } x = 1, 2, \ldots \\ 0 & \text{en otros casos.} \end{cases}$$

Podemos usar la fmp para calcular probabilidades. Como ahora podemos medir la probabilidad de un solo punto, solo necesitamos sumar todos los puntos en el evento apropiado. Por lo tanto, para enteros positivos a y b, con $a \leq b$, tenemos,

$$P(a \leq X \leq b) = \sum_{k=a}^{b} f_X(x) = \sum_{k=a}^{b} (1-p)^{x-1} p.$$

El procedimiento análogo en el caso continuo es sustituir sumas por integrales, y obtenemos,

$$P(X \leq x) = F_X(x) = \int_{-\infty}^{x} f_X(t) dt.$$

Usando el teorema fundamental del cálculo, si $F_X(x)$ es continua, tenem que,

$$\frac{d}{dx} F_X(x) = f_X(x) \quad \text{(c.p.1).}$$

Definición 1.7

La función de densidad de probabilidad o fdp $f_X(x)$, de una variable aleatoria continua X es la función que satisface, $F_X(x) = \int_{-\infty}^{x} f_X(t) dt$ para todo x.

Definición 1.6

Función de masa de probabilidad (fmp) La fmp de una variable aleatoria discreta X está dada por,

$$f_X(x) = P(X = x) \quad \text{para todo } x.$$

Definición 1.7

La función de densidad de probabilidad o fdp $f_X(x)$, de una variable aleatoria continua X es la función que satisface, $F_X(x) = \int_{-\infty}^{x} f_X(t) dt$ para todo x.

Note que para el caso continuo, $P(X = x) = 0$. En efecto, ya que $\{X = x\} \subseteq \{x - \epsilon < X \leq x\}$ para cualquier $h > 0$, we have that

$$P(X = x) \leq P_X(x - h < X \leq x) = F_X(x) - F_X(x - h)$$

para cualquier $h > 0$. Por lo tanto,

$$0 \leq P(X = x) \leq \lim_{h \downarrow 0} [F_X(x) - F_X(x - h)] = 0$$

por continuidad de F_X. Luego, si X es una variable aleatoria continua, tenemos que,

$$P(a < X < b) = P_X(a < X \leq b) = P(a \leq X < b) = P(a \leq X \leq b).$$

Teorema 1.4

Una función $f_X(x)$ es una fdp (o fmp) de una variable aleatoria X si y sólo si,

a) $f_X(x) \geq 0$ para todo x;

b) $\sum_x f_X(x) = 1$ (fmp) o $\int_{-\infty}^{\infty} f_X(x) dx = 1$ (fdp).

Demostración 1.4

Si $f_X(x)$ es una fdp (o fmp), entonces las dos propiedades son inmediatas desde la definición. En particular, para una fdp tenemos,

$$1 = \lim_{x\to\infty} F_X(x) = \int_{-\infty}^{\infty} f_X(t) dt.$$

La implicación inversa no es difícil de probar. Una vez que tengamos $f_X(x)$, podemos obtener $F_X(x)$ y usar el Teorema 1.2.