

A. Probabilidad.

Un **experimento aleatorio** tiene asociados los siguientes elementos:

- **Espacio muestral.** Conjunto Ω de todos los resultados (conceptualmente) posibles.
- **Resultados.** Elementos ω del espacio muestral, también llamados puntos muestrales o realizaciones.
- **Sucesos.** Subconjuntos de Ω para los cuales esta definida la probabilidad.

OBSERVACIÓN. En la práctica la mayor parte de los espacios muestrales pertenecen a una de las siguientes categorías:

1. Conjunto finito: $\Omega = \{0, 1\}$, $\Omega = \{1, \dots, n\}$.
2. Conjunto infinito numerable: $\Omega = \mathbb{N}$, $\Omega = \mathbb{Z}$.
3. Conjunto no numerable: $\Omega = \mathbb{R}$, $\Omega = [0, 1]$, $\Omega = \mathbb{R}_+ = [0, \infty)$.
4. Conjunto finito de replicaciones: $\Omega = \Omega_0^n = \{\omega = (\omega_1, \dots, \omega_n) : \omega_i \in \Omega_0 \forall i\}$.
5. Conjunto infinito (numerable) de replicaciones: $\Omega = \Omega_0^{\mathbb{N}} = \{\omega = (\omega_1, \omega_2, \dots) : \omega_i \in \Omega_0 \forall i\}$.
6. Espacios de funciones: $\Omega = \mathbf{C}[0, 1]$

DEFINICIÓN. Una σ -álgebra sobre Ω es una familia, \mathcal{F} , de subconjuntos de Ω , tal que verifica:

1. $\Omega \in \mathcal{F}$.
2. Si $A \in \mathcal{F}$, entonces $A^c \in \mathcal{F}$.
3. Si $A_1, A_2, \dots \in \mathcal{F}$, entonces $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

OBSERVACIÓN. Dada una familia \mathcal{C} de subconjuntos de Ω , existe la menor σ -álgebra que la contiene (menor en el sentido de que cualquier σ -álgebra que contenga a \mathcal{C} también la contiene a ella). La denotaremos, en general, por $\sigma(\mathcal{C})$ y diremos que es la σ -álgebra generada por \mathcal{C} .

Un caso particular es la σ -álgebra de Borel sobre \mathbb{R} : $\mathcal{B}(\mathbb{R})$.

Asociados a un experimento aleatorio siempre tendremos un espacio muestral y una σ -álgebra, (Ω, \mathcal{F}) . A los elementos de \mathcal{F} los llamaremos sucesos.

DEFINICIÓN. Una **probabilidad** sobre (Ω, \mathcal{F}) es una función $P: \mathcal{F} \rightarrow \mathbb{R}$ tal que

1. $P(A) \geq 0$, $\forall A \in \mathcal{F}$.
2. $P(\Omega) = 1$.
3. (Numerablemente aditiva) Cualesquiera que sean $A_1, A_2, \dots \in \mathcal{F}$ disjuntos dos a dos,

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

A la tripleta (Ω, \mathcal{F}, P) la llamaremos **espacio de probabilidad**.

OBSERVACIÓN. Quizás la interpretación más plausible de $P(A)$ es como la tendencia de la frecuencia con la que ocurre A bajo replicaciones independientes del experimento aleatorio.

TEOREMA 1. (Propiedades de la Probabilidad). Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Entonces:

1. $P(\emptyset) = 0$.
2. Si $A_1, \dots, A_n \in \mathcal{F}$ son disjuntos dos a dos, entonces $P(\bigcup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i)$.
3. Si $A \in \mathcal{F}$, entonces $P(A^c) = 1 - P(A)$.
4. Si $A, B \in \mathcal{F}$ tal que $A \subseteq B$, entonces $P(B \setminus A) = P(B) - P(A)$.
5. P es monótona.
6. Si $A, B \in \mathcal{F}$, entonces $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
7. Si $A, B \in \mathcal{F}$, entonces $P(A \cup B) \leq P(A) + P(B)$.
8. Si $A_1, A_2, \dots \in \mathcal{F}$, entonces $P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i)$.

DEFINICIÓN. Sean A_1, A_2, \dots y A subconjuntos de Ω .

1. El **límite superior** de (A_n) es el conjunto de los ω tal que $\omega \in A_n$ para un número infinito de valores de n :

$$\limsup_n A_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} A_n \quad (\text{utilizaremos la notación } \{A_n, i.o.\})$$

2. El **límite inferior** de (A_n) es el conjunto de los ω tal que $\omega \in A_n$ para un todos los valores de n salvo a lo sumo un número finito:

$$\liminf_n A_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} A_n \quad (\text{utilizaremos la notación } \{A_n, ult.\})$$

3. La sucesión (A_n) converge a A , y escribiremos $A = \lim_{n \rightarrow \infty} A_n$ ó $A_n \rightarrow A$, si $\liminf_n A_n = \limsup_n A_n = A$.

TEOREMA 2. Sea (Ω, \mathcal{F}, P) un espacio de probabilidad. Entonces se verifica:

1. Si $A_n \uparrow A$ en \mathcal{F} (i.e. $A_1 \subseteq A_2 \subseteq \dots$, $A_n \rightarrow A = \bigcup_{k=1}^{\infty} A_k$), entonces $P(A_n) \uparrow P(A)$.
2. Si $A_n \downarrow A$ en \mathcal{F} (i.e. $A_1 \supseteq A_2 \supseteq \dots$, $A_n \rightarrow A = \bigcap_{k=1}^{\infty} A_k$), entonces $P(A_n) \downarrow P(A)$.
3. Si $A_n \rightarrow A$ en \mathcal{F} , entonces $P(A_n) \rightarrow P(A)$

TEOREMA 3. (Lema de Borel-Cantelli). Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y $A_1, A_2, \dots \in \mathcal{F}$. Si $\sum_{n=1}^{\infty} P(A_n) < +\infty$, entonces $P(\{A_n, i.o.\}) = 0$.

DEFINICIÓN. Un suceso A es **casi seguro (c.s.)** u ocurre casi seguramente si $P(A) = 1$, y **nulo** si $P(A) = 0$. Una propiedad de los resultados de un experimento aleatorio se verifica casi seguramente (c.s.) si existe un suceso casi seguro sobre el cual es satisfecha.

B. Probabilidad sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Consideraremos P una probabilidad sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

DEFINICIÓN. La función de distribución de P es la función $F_P: \mathbb{R} \rightarrow [0, 1]$ definida por $F_P(t) = P((-\infty, t])$.

OBSERVACIÓN. Si $F_P = F_{P'}$, entonces $P = P'$.

TEOREMA 4. Sea F_P la función de distribución de P . Entonces:

1. F_P es creciente.
2. F_P es continua por la derecha (i.e. $F_P(t+) = F_P(t)$, donde $F_P(t+) = \lim_{s \downarrow t} F_P(s)$).
3. $F_P(-\infty) = \lim_{t \rightarrow -\infty} F_P(t) = 0$; $F_P(+\infty) = \lim_{t \rightarrow \infty} F_P(t) = 1$

TEOREMA 5. Sea $F: \mathbb{R} \rightarrow \mathbb{R}$ una función creciente, continua por la derecha con $F(-\infty) = 0$ y $F(+\infty) = 1$. Entonces existe una única probabilidad P sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ tal que $F_P = F$.

Hay dos clases muy importantes de probabilidades sobre \mathbb{R} : probabilidades discretas y absolutamente continuas.

DEFINICIÓN. Una probabilidad sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, P , se dice **discreta** si existe un conjunto numerable C tal que $P(C) = 1$.

PROPOSICIÓN 6. Las siguientes afirmaciones son equivalentes para una probabilidad P sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$:

1. P es discreta.
2. Existe un conjunto numerable de números reales $(t_i)_i$ y valores p_i , con $p_i > 0$ para cada i y $\sum_i p_i = 1$, tal que $P = \sum_i p_i \mathcal{E}_{t_i}$, (siendo $\mathcal{E}_x(A) = 1$ si $x \in A$; 0 si $x \notin A$).
3. Existe un conjunto numerable de números reales $(t_i)_i$ y valores p_i , con $p_i > 0$ para cada i y $\sum_i p_i = 1$, tal que $F_P(t) = \sum_i p_i I_{(-\infty, t]}(t_i)$, para todo $t \in \mathbb{R}$.

DEFINICIÓN. Una probabilidad P sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ se dice **absolutamente continua** si existe una función positiva, f_P , sobre \mathbb{R} , que llamaremos función de densidad de P , tal que para cualquier intervalo $(a, b]$

$$P((a, b]) = \int_a^b f_P(t) dt. \tag{1}$$

OBSERVACIÓN.

1. En el caso más general las integrales son de Lebesgue. De todos modos en casi todos los casos utilizados las integrales son de Riemann.
2. El término "la" función de densidad de P no es del todo preciso, puesto que en un sentido técnico no es única. De todos modos cualesquiera dos funciones que satisfagan (1) difieren a lo sumo en un conjunto de medida de Lebesgue nula.

PROPOSICIÓN 7. La probabilidad P sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ es **absolutamente continua** si y sólo si existe una función positiva sobre \mathbb{R} con $\int_{-\infty}^{+\infty} f(s) ds = 1$ y

$$F_P(t) = \int_{-\infty}^t f(s) ds, \quad t \in \mathbb{R}.$$

OBSERVACIÓN. Una probabilidad sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ no necesariamente es discreta o absolutamente continua, puede tener componentes de los dos términos o de ninguno. En realidad existe un tercer tipo de distribuciones de probabilidad sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, las *probabilidades singulares*.

C. Probabilidad Condicionada.

Consideraremos que asociado a un experimento aleatorio tenemos un espacio de probabilidad $(\Omega, \mathcal{F}, \mathcal{P})$.

DEFINICIÓN. Sean A y B sucesos. Siempre que $P(A) > 0$, la probabilidad condicionada de B dado A se define como

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

Si $P(A) = 0$, consideraremos $P(B|A) = P(B)$.

OBSERVACIÓN. $(\Omega, \mathcal{F}, P_A)$, con $P_A(B) = P(B|A)$, es un espacio de probabilidad.

PROPOSICIÓN 8. (Regla de la Multiplicación). Sea A_1, \dots, A_n sucesos tales que $P(\bigcap_{i=1}^{n-1} A_i) > 0$. Entonces

$$P\left(\bigcap_{i=1}^n A_i\right) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \dots P(A_n | \bigcap_{i=1}^{n-1} A_i)$$

PROPOSICIÓN 9. (Teorema de la Probabilidad Total). Supongamos $\{A_1, A_2, \dots\}$ una partición numerable disjunta de Ω . Entonces, para cada suceso B ,

$$P(B) = \sum_{i=1}^{\infty} P(B|A_i)P(A_i).$$

PROPOSICIÓN 10. (Teorema de Bayes). Supongamos $\{A_1, A_2, \dots\}$ una partición numerable disjunta de Ω . Entonces, para cada suceso B con $P(B) > 0$ y cada j ,

$$P(A_j|B) = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{\infty} P(B|A_i)P(A_i)}.$$

OBSERVACIÓN. Podemos interpretar los A_j como estados de la naturaleza no observables, cada uno de los cuales tiene una probabilidad a priori, $P(A_j)$, de ocurrir, junto con una probabilidad $P(B|A_j)$ de "causar" que un suceso observable B ocurra. Entonces, dado que B ha ocurrido, la fórmula de Bayes nos permite calcular las probabilidades a posteriori (revisadas), $P(A_j|B)$, de los estados A_j .

A. Variables Aleatorias.

Consideremos un espacio de probabilidad (Ω, \mathcal{F}, P) asociado a un experimento aleatorio.

DEFINICIÓN. Una **variable aleatoria** (v.a.) es una función $X: \Omega \rightarrow \mathbb{R}$ tal que $X^{-1}(B) \in \mathcal{F}$ para todo $B \in \mathcal{B}(\mathbb{R})$

OBSERVACIÓN. Se verifica que $X: \Omega \rightarrow \mathbb{R}$ es v.a. si y sólo si $\{X \leq t\} = X^{-1}((-\infty, t]) \in \mathcal{F}$ para todo $t \in \mathbb{R}$.

DEFINICIÓN. Un **vector aleatorio d-dimensional** es una función $X = (X_1, \dots, X_d): \Omega \rightarrow \mathbb{R}^d$ tal que cada componente X_i , $i = 1, \dots, d$, es una v.a.

DEFINICIÓN. Dado un conjunto infinito T un **proceso estocástico** con conjunto índice T , es una colección de v.a. $\{X_t: t \in T\}$

OBSERVACIÓN. Habitualmente el índice de un proceso estocástico representa el tiempo:

- Proceso estocástico en tiempo discreto: $\{X_n\}_{n \geq 0}$ sucesión.
- Proceso estocástico en tiempo continuo: $\{X_t: t \geq 0\}$

DEFINICIÓN. Una función $Z: \Omega \rightarrow \mathbb{C}$ es una **v.a. con valores complejos** si $Z = X + iY$, donde X e Y son v.a.

PROPOSICIÓN 11. Dada una v.a. X , $\sigma(X) = \{X^{-1}(B): B \in \mathcal{B}(\mathbb{R})\}$ es una σ -álgebra sobre Ω (la σ -álgebra generada por X). Análogamente, si $X = (X_1, \dots, X_d)$ es un vector aleatorio d-dimensional, $\sigma(X_1, \dots, X_d) = \{X^{-1}(B): B \in \mathcal{B}(\mathbb{R}^d)\}$ es una σ -álgebra sobre Ω (la σ -álgebra generada por $X = (X_1, \dots, X_d)$).

PROPOSICIÓN 12. Sean X e Y v.a. Entonces:

1. $aX + bY$ es una v.a. $\forall a, b \in \mathbb{R}$.
2. $\max\{X, Y\}$ y $\min\{X, Y\}$ son v.a.
3. XY es una v.a.
4. Siempre que $Y(\omega) \neq 0$ para cada ω , X/Y es una v.a.

PROPOSICIÓN 13. Sean X_1, X_2, \dots v.a. Entonces $\sup_n X_n$, $\inf_n X_n$, $\limsup_n X_n$ y $\liminf_n X_n$ son v.a. Consecuentemente si $X(\omega) = \lim_{n \rightarrow \infty} X_n(\omega)$ existe para cada $\omega \in \Omega$, entonces X es una v.a.

PROPOSICIÓN 14. Sean X_1, \dots, X_d v.a. y sea $g: \mathbb{R}^d \rightarrow \mathbb{R}$ una función Borel medible (i.e. $g^{-1}(B) \in \mathcal{B}(\mathbb{R}^d)$, $\forall B \in \mathcal{B}(\mathbb{R})$). Entonces, $Y = g(X_1, \dots, X_d)$ es una v.a.

PROPOSICIÓN 15. Para una v.a. X , la función de conjunto $P^X: \mathcal{B}(\mathbb{R}) \rightarrow \mathbb{R}$ dada por $P^X(B) = P(X^{-1}(B))$ es una probabilidad sobre $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

DEFINICIÓN. Sea X una v.a.

1. Llamaremos **distribución de probabilidad** de X a la probabilidad P^X .
2. Llamaremos **función de distribución** de X a $F_X(t) = F_{P^X}(t) = P^X((-\infty, t]) = P(X \leq t)$, $\forall t \in \mathbb{R}$

OBSERVACIÓN.

- Una v.a. X se dice discreta o absolutamente continua si P^X lo es.
- Las v.a. X e Y son idénticamente distribuidas si $F_X = F_Y$ ($\Leftrightarrow P^X = P^Y$) y escribiremos $X \stackrel{d}{=} Y$.

DEFINICIÓN. Sea $X = (X_1, \dots, X_d)$ un vector aleatorio d -dimensional.

- Llamaremos **distribución de probabilidad** de X a la probabilidad P^X sobre $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$, definida por $P^X(B) = P(X^{-1}(B))$, $B \in \mathcal{B}(\mathbb{R}^d)$.
- Llamaremos **función de distribución** de X (o **función de distribución conjunta** de $X = (X_1, \dots, X_d)$) a $F_X: \mathbb{R}^d \rightarrow [0, 1]$ dada por $F_X(t_1, \dots, t_d) = P(X_1 \leq t_1, \dots, X_d \leq t_d)$, $\forall (t_1, \dots, t_d) \in \mathbb{R}^d$.

OBSERVACIÓN.

- Si $X = (X_1, \dots, X_d)$ es un vector aleatorio d -dimensional, entonces para cada i y t ,

$$F_{X_i}(t) = \lim_{t_j \rightarrow +\infty (j \neq i)} F_X(t_1, \dots, t_{i-1}, t, t_{i+1}, \dots, t_d).$$

- $X = (X_1, \dots, X_d)$ es un vector aleatorio d -dimensional discreto si existe un subconjunto numerable $C \subseteq \mathbb{R}^d$ tal que $P(X \in C) = 1$.

$X = (X_1, \dots, X_d)$ es un vector aleatorio d -dimensional discreto si y sólo si X_1, \dots, X_d son v.a. discretas.

- $X = (X_1, \dots, X_d)$ es un vector aleatorio d -dimensional absolutamente continuo si existe una función $f_X: \mathbb{R}^d \rightarrow [0, +\infty)$ (llamada función de densidad de X o función de densidad conjunta de (X_1, \dots, X_d)) tal que

$$P(X_1 \leq t_1, \dots, X_d \leq t_d) = \int_{-\infty}^{t_1} \cdots \int_{-\infty}^{t_d} f_X(y_1, \dots, y_d) dy_1 \cdots dy_d.$$

Si $X = (X_1, \dots, X_d)$ es un vector aleatorio d -dimensional absolutamente continuo, entonces cada X_i es una v.a. absolutamente continua y su función de densidad es

$$f_{X_i}(x) = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f_X(y_1, \dots, y_{i-1}, x, y_{i+1}, \dots, y_d) dy_1 \cdots dy_{i-1} dy_{i+1} \cdots dy_d.$$

PROPOSICIÓN 16. Sea X una v.a. con función de distribución F_X . Sea g una función Borel medible de \mathbb{R} en \mathbb{R} , continua y estrictamente creciente, y $h = g^{-1}$. Sea $Y = g(X)$. Entonces para cada $t \in \mathbb{R}$, $F_Y(t) = F_X(h(t))$.

OBSERVACIÓN.

- Si g es una función Borel medible cualquiera, $Y = g(X)$ también es v.a. y su distribución depende de la de X .
- Si X es una v.a. discreta, entonces $g(X)$ también lo es.
- Si X es una v.a. absolutamente continua, entonces $g(X)$ no necesariamente lo es.

TEOREMA 17. Sea X una v.a. absolutamente continua con función de densidad f_X . Sea g una función diferenciable para todo x , y supongamos que $g'(x)$ es continua y distinta de 0 en casi todos (resp. a P^X) los valores de x . Entonces para cualquier número real y se verifica, bien:

- Existe un entero positivo $n = n(y)$ y números reales $x_1(y), \dots, x_n(y)$ tales que:

$$g(x_k(y)) = y, \quad g'(x_k(y)) \neq 0, \quad k = 1, \dots, n(y)$$

o bien

- No existe ningún x tal que $g(x) = y$, $g'(x) \neq 0$, en cuyo caso consideramos $n = n(y) = 0$.

Entonces $Y = g(X)$ es una v.a. absolutamente continua cuya función de densidad viene dada por:

$$f_Y(y) = \sum_{k=1}^n f_X(x_k(y)) |g'(x_k(y))|^{-1} \text{ si } n > 0; \quad 0 \text{ si } n = 0.$$

TEOREMA 18. Sea $X = (X_1, \dots, X_d)$ un vector aleatorio absolutamente continuo con función de densidad f_X , y sea $g: \mathbb{R}^d \rightarrow \mathbb{R}^d$ una función Borel medible para la cual existe un conjunto abierto $U \subseteq \mathbb{R}^d$ tal que $P(X \in U) = 1$, g es biyectiva sobre U , y $Jg(x) \neq 0 \forall x \in U$, donde Jg es el Jacobiano de g , i.e. el determinante de la matriz de derivadas parciales, $\partial g_i / \partial x_j$. Sea $h = g^{-1}$. Entonces $Y = g(X)$ es un vector aleatorio absolutamente continuo con

$$f_Y(y) = f_X(h(y)) |Jh(y)| \quad y \in g(U).$$

B. Independencia.

Consideremos un espacio de probabilidad (Ω, \mathcal{F}, P) asociado a un experimento aleatorio.

DEFINICIÓN. Las v.a. X_1, \dots, X_n son **independientes** si $P(X_1 \in B_1, \dots, X_n \in B_n) = \prod_{i=1}^n P(X_i \in B_i)$, para todos los $B_i \in \mathcal{B}(\mathbb{R})$, $i = 1, \dots, n$

Un conjunto infinito de v.a. es independiente si cualquier subconjunto finito es independiente.

OBSERVACIÓN. Las v.a. que son independientes y tienen la misma función de distribución, se dice que son *independientes e idénticamente distribuidas* y se denota por *i.i.d.*

TEOREMA 19. Las v.a. X_1, \dots, X_n son independientes si y sólo si $F_{(X_1, \dots, X_n)}(t_1, \dots, t_n) = \prod_{i=1}^n F_{X_i}(t_i)$, cualesquiera que sean $t_1, \dots, t_n \in \mathbb{R}$.

TEOREMA 20. Sean X_1, \dots, X_n v.a. discretas que toman valores en el conjunto numerable C . Las v.a. X_1, \dots, X_n son independientes si y sólo si $P(X_1 = a_1, \dots, X_n = a_n) = \prod_{i=1}^n P(X_i = a_i)$, cualesquiera que sean $a_1, \dots, a_n \in C$.

TEOREMA 21. Sea $X = (X_1, \dots, X_n)$ un vector aleatorio absolutamente continuo. Las v.a. X_1, \dots, X_n son independientes si y sólo si $f_X(y_1, \dots, y_n) = \prod_{i=1}^n f_{X_i}(y_i)$, cualesquiera que sean $y_1, \dots, y_n \in \mathbb{R}$.

COROLARIO. Si X_1, \dots, X_n son v.a. independientes y absolutamente continuas, entonces $X = (X_1, \dots, X_n)$ es un vector aleatorio absolutamente continuo.

TEOREMA 22. (Teorema de los bloques disjuntos). Supongamos que X_1, \dots, X_n son v.a. independientes. Sean J_1, \dots, J_k subconjuntos disjuntos de $\{1, \dots, n\}$, y para cada ℓ , sea Y_ℓ una función Borel-medible g_ℓ de $X^{(J_\ell)} = \{X_i : i \in J_\ell\}$. Entonces, Y_1, \dots, Y_k son independientes.

DEFINICIÓN. Los sucesos A_1, \dots, A_n son independientes si sus funciones indicadoras son v.a. independientes. Una colección infinita de sucesos es independiente si cualquier subcolección finita es independiente.

TEOREMA 23. Los sucesos A_1, \dots, A_n son independientes si y sólo si $P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$, siendo I cualquier subconjunto de $\{1, \dots, n\}$.

COROLARIO. Los sucesos A_1, \dots, A_n son independientes si y sólo si A_1^c, \dots, A_n^c son independientes.

TEOREMA 24. (Lema de Borel-Cantelli) Sean A_1, A_2, \dots sucesos independientes tales que $\sum_{i=1}^{\infty} P(A_i) = +\infty$. Entonces, $P(\{A_n, i.o.\}) = 1$

A. Esperanza Matemática.

Consideremos un espacio de probabilidad (Ω, \mathcal{F}, P) asociado a un experimento aleatorio. En primer lugar definiremos el concepto de esperanza para una v.a. simple (i.e. aquella que toma solamente un número finito de valores), $X = \sum_{i=1}^n a_i I_{A_i}$, $a_i \in \mathbb{R}$, (A_i) una partición de Ω .

DEFINICIÓN. La esperanza de $X = \sum_{i=1}^n a_i I_{A_i}$ es $E[X] = \int_{\Omega} X dP = \sum_{i=1}^n a_i P(A_i)$

OBSERVACIÓN:

- $E[X]$ está bien definida no dependiendo de la representación considerada de X .
- $E[I_A] = P(A)$.
- $E[c] = c$.
- Si X, Y son v.a. simples y $a, b \in \mathbb{R}$, entonces $aX + bY$ es una v.a. simple y $E[aX + bY] = aE[X] + bE[Y]$.
- Si X, Y son v.a. simples tales que $X \leq Y$, entonces $E[X] \leq E[Y]$.

Para cualquier v.a. $X \geq 0$, existe una sucesión de v.a. simples $0 \leq X_1 \leq X_2 \leq \dots$ tales que $X_n(\omega) \uparrow X(\omega)$, $\omega \in \Omega$.

DEFINICIÓN. Sea X una v.a. positiva.

1. La esperanza de X es $E[X] = \int_{\Omega} X dP = \lim_{n \rightarrow \infty} E[X_n] \leq +\infty$, donde las v.a. X_n son simples y positivas con $X_n(\omega) \uparrow X(\omega)$, $\omega \in \Omega$.
2. La esperanza de X sobre un suceso A se define como $E[X : A] = \int_A X dP = E[X I_A]$.

OBSERVACIÓN:

- $0 \leq E[X_1] \leq E[X_2] \leq \dots$ implica que el límite definido como $E[X]$ existe en $\bar{\mathbb{R}}_+$.
- Si (X_n) y (\tilde{X}_n) son sucesiones de v.a. simples que convergen de manera creciente a X , entonces $\lim_{n \rightarrow \infty} E[X_n] = \lim_{n \rightarrow \infty} E[\tilde{X}_n]$.
- Si X, Y son v.a. positivas y $a, b \in \mathbb{R}_+$, entonces $E[aX + bY] = aE[X] + bE[Y]$.
- Si X, Y son v.a. positivas tales que $X \leq Y$, entonces $E[X] \leq E[Y]$.

TEOREMA 25 (LEMA DE FATOU). Para $X_n \geq 0$, se verifica que $E[\liminf_{n \rightarrow +\infty} X_n] \leq \liminf_{n \rightarrow +\infty} E[X_n]$.

TEOREMA 26 (DE LA CONVERGENCIA MONÓTONA). Sean X, X_1, X_2, \dots v.a. positivas tales que $X_n(\omega) \uparrow X(\omega)$, $\omega \in \Omega$. Entonces $E[X_n] \uparrow E[X]$.

Si X es una v.a. cualquiera, definimos $X^+ = \max\{X, 0\}$ y $X^- = -\min\{X, 0\}$. Así $X = X^+ - X^-$ y $|X| = X^+ + X^-$.

DEFINICIÓN. Sea X una v.a. cualquiera.

1. X es integrable si $E[|X|] < +\infty$.
2. Si X es integrable, la esperanza de X se define como $E[X] = \int_{\Omega} X dP = E[X^+] - E[X^-]$.
3. Para X integrable y A un suceso, la esperanza de X sobre A se define como $E[X : A] = \int_A X dP = E[X I_A]$.

A. Esperanza Matemática.

OBSERVACIÓN:

- $E[X]$ está bien definido.
- Esta definición generaliza las dos dadas anteriormente.
- El conjunto de v.a. integrables lo denotaremos por L^1 .

TEOREMA 27. Si $X, Y \in L^1$ y $a, b \in \mathbb{R}$, entonces $aX + bY \in L^1$ y $E[aX + bY] = aE[X] + bE[Y]$.

OBSERVACIÓN: L^1 es un espacio vectorial.

COROLARIO.

1. Para $X \in L^1$, se verifica que $|E[X]| \leq E[|X|]$.
2. Si $X, Y \in L^1$ tales que $X \leq Y$, entonces $E[X] \leq E[Y]$.

TEOREMA 28 (DE LA CONVERGENCIA DOMINADA). Sean X, X_1, X_2, \dots v.a. integrables con $X_n(\omega) \rightarrow X(\omega)$ para todo ω , y supongamos que existe $Y \in L^1$ tal que $|X_n| \leq Y$ para cada n . Entonces $\lim_{n \rightarrow +\infty} E[X_n] = E[X]$.

DEFINICIÓN. Una v.a. compleja $Z = X + iY$ es integrable si $E[|Z|] = E[\sqrt{X^2 + Y^2}] < +\infty$, y en este caso la esperanza de Z es $E[Z] = E[X] + iE[Y]$.

TEOREMA 29. Para $X \in L^1$, $E[X] = \int_{\mathbb{R}} x dP^X = \int_{\mathbb{R}} x dF_X(x)$ (integral de Lebesgue-Stielges). De manera más general, si g es positiva ó tal que $g(X) \in L^1$, entonces $E[g(X)] = \int_{\mathbb{R}} g(x) dP^X = \int_{\mathbb{R}} g(x) dF_X(x)$.

TEOREMA 30. Sean X e Y v.a. independientes, y sea $h: \mathbb{R}^2 \rightarrow \mathbb{R}_+$ una función Borel-medible. Entonces

$$E[h(X, Y)] = \int \left(\int h(x, y) dF_X(x) \right) dF_Y(y) = \int \left(\int h(x, y) dF_Y(y) \right) dF_X(x).$$

COROLARIO. Sean X e Y v.a. independientes y sean g_1, g_2 funciones positivas ó tales que $g_1(X), g_2(Y) \in L^1$. Entonces

$$E[g_1(X)g_2(Y)] = E[g_1(X)]E[g_2(Y)].$$

B. Desigualdades.

DEFINICIÓN: Para $1 \leq p < +\infty$, L^p denota el conjunto de v.a. X tales que $E[|X|^p] < +\infty$.

TEOREMA 31 (DESIGUALDAD DE HÖLDER): Supongamos $p, q > 1$ tal que $\frac{1}{p} + \frac{1}{q} = 1$. Sea $X \in L^p$ e $Y \in L^q$. Entonces $XY \in L^1$ y

$$E[|XY|] \leq E[|X|^p]^{1/p} E[|Y|^q]^{1/q}.$$

COROLARIO (DESIGUALDAD DE CAUCHY-SCHWARZ): Si X e Y pertenecen a L^2 , entonces $XY \in L^1$ y

$$E[|XY|] \leq (E[|X|^2]E[|Y|^2])^{1/2}.$$

COROLARIO (DESIGUALDAD DE LYAPUNOV): Para $1 \leq r \leq s$, $L^s \subseteq L^r$, y para $X \in L^s$ se verifica que

$$E[|X|^r]^{1/r} \leq E[|X|^s]^{1/s}.$$

TEOREMA 32 (DESIGUALDAD DE MINKOWSKI): Supongamos $1 \leq p < +\infty$ y $X, Y \in L^p$. Entonces $X + Y \in L^p$ y

$$E[|X + Y|^p]^{1/p} \leq E[|X|^p]^{1/p} + E[|Y|^p]^{1/p}.$$

TEOREMA 33 (DESIGUALDAD DE JENSEN): Sea g una función convexa y supongamos que X y $g(X)$ son integrables. Entonces

$$g(E[X]) \leq E[g(X)].$$

TEOREMA 34 (DESIGUALDAD DE CHEBYSHEV): Sea X una v.a. positiva y sea g una función sobre \mathbb{R}_+ creciente y positiva. Entonces para cada $a > 0$,

$$P(X \geq a) \leq \frac{E[g(X)]}{g(a)}.$$

OBSERVACIÓN: Casos especiales de la desigualdad anterior aparecen frecuentemente en probabilidad:

1. $X \in L^1 \implies P(|X| \geq a) \leq \frac{E[|X|]}{a}$.
2. $X \in L^p \implies P(|X| \geq a) \leq \frac{E[|X|^p]}{a^p}$.
3. $X \in L^2 \implies P(|X - E[X]| \geq a) \leq \frac{Var[X]}{a^2}$, siendo $Var[X] = E[(X - E[X])^2]$.
4. $X \geq 0 \implies P(X \geq a) \leq \frac{E[e^{tX}]}{e^{ta}}$.

C. Momentos.

DEFINICIÓN: Sea X una v.a.

1. Para k entero positivo y si $X \in L^k$, se define el momento ordinario de orden k de X como $E[X^k]$.
2. Si $X \in L^1$, se define la media de X como $E[X]$ (se suele denotar por μ ó μ_X).
3. Para k entero positivo y si $X \in L^k$, se define el momento central de orden k de X como $E[(X - E[X])^k]$.
4. Si $X \in L^2$, se define la varianza de X como $Var[X] = E[(X - E[X])^2]$, i.e. el momento central de orden 2 (se suele denotar por σ^2 ó σ_X^2).

OBSERVACIÓN:

- La varianza de $X \in L^2$ se puede expresar como $Var[X] = E[X^2] - E[X]^2$.
- Se define la desviación típica o estándar de $X \in L^2$ como $\sqrt{Var[X]}$.

DEFINICIÓN: Sean las v.a. $X, Y \in L^2$

1. Se define la covarianza de X e Y como $Cov(X, Y) = E[(X - E[X])(Y - E[Y])]$.
2. Siempre que $Var[X] > 0$ y $Var[Y] > 0$, se define el coeficiente de correlación entre X e Y como

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{Var[X]Var[Y]}}$$

(se suele denotar por ρ ó $\rho_{X,Y}$).

3. Diremos que X e Y son v.a. incorreladas si $Corr(X, Y) = 0$.

OBSERVACIÓN:

- La covarianza de $X, Y \in L^2$ se puede expresar como $Cov(X, Y) = E[XY] - E[X]E[Y]$.
- El coeficiente de correlación mide la asociación lineal entre v.a., pero sólo asociación lineal, i.e. X e Y pueden depender de manera funcional y ser incorreladas.
- La desigualdad de Cauchy-Schwarz garantiza que $|Corr(X, Y)| \leq 1$, dándose la igualdad si y sólo si existen $a, b \in \mathbb{R}$ tales que $Y = aX + b$ casi seguramente.
- Variables aleatorias independientes en L^2 son incorreladas.

DEFINICIÓN: Sea $X = (X_1, \dots, X_d)$ un vector aleatorio

1. Si $X_i \in L^1$, para cada i , se define la media de X como el vector $\mu_X = (E[X_1], \dots, E[X_d])$.
2. Si $X_i \in L^2$, para cada i , se define la matriz de covarianzas de X como la matriz $d \times d$ cuyo coeficiente (i, j) es $Cov(X_i, X_j)$.

OBSERVACIÓN:

- La matriz de covarianzas es simétrica, semidefinida positiva y su diagonal está formada por $Var[X_i]$, $i = 1, \dots, d$.
- $Var[\sum_{i=1}^d a_i X_i] = \sum_{i=1}^d a_i^2 Var[X_i] + 2a_i a_j \sum_{i < j} Cov(X_i, X_j)$

Función Característica de una variable aleatoria.

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad asociado a un experimento aleatorio.

DEFINICIÓN. Una función $Z: \Omega \rightarrow \mathbb{C}$ es una **v.a. con valores complejos** ó v.a. compleja si $Z = X + iY$, donde X e Y son v.a.

DEFINICIÓN. Una v.a. compleja, $Z = X + iY$ es integrable si $E[|Z|] = E[\sqrt{(X^2 + Y^2)}] < +\infty$, y en este caso su esperanza es $E[Z] = E[X] + iE[Y]$.

OBSERVACIÓN. La esperanza de una v.a. compleja, Z , verifica la linealidad, el teorema de la convergencia dominada y $|E[Z]| \leq E[|Z|]$.

DEFINICIÓN. Sea X una v.a. (real). La **función característica** de X es $\varphi_X: \mathbb{R} \rightarrow \mathbb{C}$ definida por

$$\varphi_X(t) = E[e^{itX}] = \int_{-\infty}^{+\infty} e^{itx} dF_X(x),$$

siendo F_X la función de distribución de X .

OBSERVACIÓN.

- La función característica de una v.a. siempre existe.
- En el caso discreto adoptará la forma

$$\varphi_X(t) = \sum_{x \in C} e^{itx} P(X = x) \quad t \in \mathbb{R}.$$

y en el caso absolutamente continuo

$$\varphi_X(t) = \int_{-\infty}^{+\infty} e^{itx} f_X(x) dx \quad t \in \mathbb{R}.$$

PROPOSICIÓN 1. Sea X una v.a. con función característica $\varphi_X(t)$, entonces:

1. $\varphi_X(t)$ es uniformemente continua.
2. $\varphi_X(-t) = \overline{\varphi_X(t)}$

PROPOSICIÓN 2. Sean X e Y v.a. independientes con funciones características $\varphi_X(t)$ y $\varphi_Y(t)$, respectivamente. Entonces $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$ para todo $t \in \mathbb{R}$.

Teoremas de Inversión y Unicidad.

TEOREMA 3. (Teorema de inversión de Levy). Sea X una v.a. con función de distribución F_X . Entonces, se verifica que cualesquiera sean a y b puntos de continuidad de F_X ,

$$P(a < X < b) = \lim_{T \rightarrow \infty} \frac{1}{2\pi} \int_{-T}^T \frac{e^{-ita} - e^{-itb}}{it} \varphi_X(t) dt.$$

TEOREMA 4. (Teorema de unicidad). Sean X e Y v.a. con funciones de distribución F_X y F_Y y funciones características $\varphi_X(t)$ y $\varphi_Y(t)$, respectivamente. Entonces $\varphi_X(t) = \varphi_Y(t)$ para todo $t \in \mathbb{R}$, implica que $F_X(x) = F_Y(x)$ para todo $x \in \mathbb{R}$.

TEOREMA 5. (Teorema de inversión de Fourier). Sea X una v.a. con función característica $\varphi_X(t)$. Si $\int_{-\infty}^{+\infty} |\varphi_X(t)| dt < +\infty$, entonces X es absolutamente continua con densidad

$$f_X(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \varphi_X(t) dt.$$

TEOREMA 6. Sea X una v.a. con función característica $\varphi_X(t)$. Entonces, para cada $x \in \mathbb{R}$,

$$P(X = x) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T e^{-itx} \varphi_X(t) dt.$$

COROLARIO. Sea X una v.a. entero valuada con función característica $\varphi_X(t)$. Entonces, para cada $n \in \mathbb{Z}$,

$$P(X = n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-itn} \varphi_X(t) dt.$$

Cálculo de momentos a través de la función característica.

TEOREMA 7. Sea X una v.a. con función característica $\varphi_X(t)$ y $k \geq 1, k \in \mathbb{Z}$. Si $X \in L^k$, i.e. $E[|X^k|] < +\infty$, entonces existe la derivada k -ésima de $\varphi_X(t)$, $\varphi_X^{(k)}$ y $E[X^k] = \varphi_X^{(k)}(0)/i^k$

TEOREMA 8. Sea X una v.a. con función característica $\varphi_X(t)$ y $k \in \mathbb{Z}$ un número par. Si existe la derivada k -ésima de $\varphi_X(t)$ en 0 , $\varphi_X^{(k)}(0)$, entonces $E[|X^k|] < +\infty$.

TEOREMA 9. Sea X una v.a. con función característica $\varphi_X(t)$ y $k \geq 1, k \in \mathbb{Z}$. Si $X \in L^k$, i.e. $E[|X^k|] < +\infty$, entonces para t en un entorno de 0 ,

$$\varphi_X(t) = \sum_{j=0}^k \frac{(it)^j}{j!} E[X^j] + o(t^k).$$

Función característica de vectores aleatorios.

DEFINICIÓN. Sea $X = (X_1, \dots, X_n)$ un vector aleatorio n -dimensional. La **función característica** del vector X es $\varphi_X: \mathbb{R}^n \rightarrow \mathbb{C}$ definida por

$$\varphi_X(t) = E[e^{i \sum_{j=1}^n t_j X_j}] = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{i \sum_{j=1}^n t_j x_j} dF_X(x_1, \dots, x_n), \quad t = (t_1, \dots, t_n),$$

siendo F_X la función de distribución del vector X .

OBSERVACIÓN.

- La función característica de un vector aleatorio siempre existe.
- En el caso discreto adoptará la forma

$$\varphi_X(t) = \sum_{x=(x_1, \dots, x_n) \in C} e^{i \sum_{j=1}^n t_j x_j} P(X = x), \quad t = (t_1, \dots, t_n) \in \mathbb{R}^n.$$

y en el caso absolutamente continuo

$$\varphi_X(t) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} e^{i \sum_{j=1}^n t_j x_j} f_X(x_1, \dots, x_n) dx_1 \dots dx_n, \quad t = (t_1, \dots, t_n) \in \mathbb{R}^n.$$

TEOREMA 10. Sea $X = (X_1, \dots, X_n)$ un vector aleatorio n -dimensional con función característica $\varphi_X(t)$, entonces:

1. $\varphi_X(t)$ es uniformemente continua.
2. $\varphi_X(-t) = \overline{\varphi_X(t)}$, siendo $-t = (-t_1, \dots, -t_n)$.
3. La función característica del vector (X_1, \dots, X_r) , ($r < n$), en el punto (t_1, \dots, t_r) es $\varphi_X(t_1, \dots, t_r, 0, \dots, 0)$
4. $X \stackrel{d}{=} Y$ si y sólo si $\varphi_X(t) = \varphi_Y(t)$, para todo $t \in \mathbb{R}^n$
5. X_1, \dots, X_n son independientes si y sólo si $\varphi_X(t_1, \dots, t_n) = \prod_{j=1}^n \varphi_{X_j}(t_j)$

Funciones generatrices.

DEFINICIÓN. Sea X una v.a. La **función generatriz de momentos** (f.g.m.) de la v.a. X es la función,

$$M_X(t) = E[e^{tX}] = \int_{-\infty}^{+\infty} e^{tx} dF_X(x),$$

siempre que la esperanza exista para todo t en un entorno de 0.

OBSERVACIÓN.

- La función generatriz de momentos de una v.a. no siempre existe.
- En el caso discreto (si existe) adoptará la forma

$$M_X(t) = \sum_{x \in C} e^{tx} P(X = x).$$

y en el caso absolutamente continuo (si existe)

$$M_X(t) = \int_{-\infty}^{+\infty} e^{tx} f_X(x) dx.$$

PROPOSICIÓN 11. Si para una v.a. X existe su f.g.m., entonces para cada entero $k \geq 1$, $E[|X|^k] < +\infty$ y

$$E[X^k] = M_X^{(k)}(0).$$

OBSERVACIÓN.

- La existencia de la f.g.m. implica la existencia de momentos de todos los órdenes.
- La distribución de cualquier v.a. cuya f.g.m. existe está unívocamente determinada por sus momentos.

PROPOSICIÓN 12. Sean X e Y v.a. con f.g.m. $M_X(t)$ y $M_Y(t)$, respectivamente. Se verifica:

1. Si $M_X(t) = M_Y(t) < +\infty$ para todo t en un intervalo abierto de \mathbb{R} que contenga a 0, entonces $X \stackrel{d}{=} Y$.
2. Si X e Y son v.a. independientes, entonces $M_{X+Y}(t) = M_X(t)M_Y(t)$ para t en un entorno de 0.

Funciones generatrices.

DEFINICIÓN. Sea X una v.a. entero-valuada positiva. Definimos la **función generatriz de probabilidad** (f.g.p.) de la v.a. X como,

$$\psi_X(s) = E[s^X] = \sum_{n=0}^{\infty} P(X = n)s^n,$$

para $s \in [-1, 1]$ (ó $|s| \leq 1$).

PROPOSICIÓN 13. Sea X una v.a. entero-valuada positiva. Entonces se verifica que

$$P(X = 0) = \psi_X(0) \quad \text{y} \quad P(X = n) = \frac{\psi_X^{(n)}(0)}{n!}, \quad n = 1, 2, \dots$$

PROPOSICIÓN 14. Sean X e Y v.a. entero-valuadas positivas con f.g.p. $\psi_X(s)$ y $\psi_Y(s)$, respectivamente. Se verifica:

1. Si $\psi_X(s) = \psi_Y(s)$ para todo s en un intervalo abierto de \mathbb{R} que contenga a 0 , entonces $X \stackrel{d}{=} Y$.
2. Si X e Y son v.a. independientes, entonces $\psi_{X+Y}(s) = \psi_X(s)\psi_Y(s)$ para s en un entorno de 0 .
3. Si $E[|X|^k] < +\infty$, entonces existe $\psi_X^{(k)}(1) (= \lim_{s \uparrow 1} \psi_X^{(k)}(s))$ y $\psi_X^{(k)}(1) = E[X(X-1)\dots(X-k+1)]$.

Modos de Convergencia.

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad, y sea $\{X_n, n \geq 1\}$ una sucesión de variables aleatorias definidas sobre él.

DEFINICIÓN. La sucesión de variables aleatorias $\{X_n\}$ se dice que **converge casi seguramente** a la variable aleatoria X si existe un conjunto $F \in \mathcal{F}$ con $P(F) = 0$, tal que para todo $\omega \in F^c$, $X_n(\omega) \rightarrow X(\omega)$ cuando $n \rightarrow \infty$. En tal caso escribiremos $X_n \xrightarrow{c.s.} X$.

OBSERVACIONES.

- Se trata de la versión probabilística de la convergencia puntual.
- Es claro que la variable límite, si existe, es única c.s.[P] (i.e. salvo un conjunto de probabilidad P nula).

DEFINICIÓN. La sucesión de variables aleatorias $\{X_n\}$ se dice que **converge en probabilidad** a la variable aleatoria X si para todo $\epsilon > 0$,

$$P(\{\omega \in \Omega: |X_n(\omega) - X(\omega)| \geq \epsilon\}) \rightarrow 0$$

cuando $n \rightarrow \infty$. En tal caso escribiremos $X_n \xrightarrow{P} X$.

OBSERVACIONES.

- Que $\{X_n\}$ converja en probabilidad significa que la probabilidad de que $|X_n - X|$ sea tan pequeño como queramos tiende a uno cuando n se hace suficientemente grande.
- La variable límite, si existe, es única c.s.[P].

Sea $L^1 = L^1(\Omega, \mathcal{F}, P)$ el conjunto de todas las variables aleatorias con esperanza finita.

DEFINICIÓN. Sea $\{X_n\}$ una sucesión de variables aleatorias tales que $X_n \in L^1$, $n \geq 1$. Diremos que $\{X_n\}$ **converge en media o en L^1** a la variable aleatoria $X \in L^1$ si $E[|X_n - X|] \rightarrow 0$ cuando $n \rightarrow \infty$. En tal caso escribiremos $X_n \xrightarrow{L^1} X$.

OBSERVACIÓN. La variable límite, si existe, es única c.s.[P].

Sea $1 \leq p < \infty$ y $L^p = L^p(\Omega, \mathcal{F}, P)$ el conjunto de todas las variables aleatorias X tales que $E[|X|^p] < \infty$.

DEFINICIÓN. Sea $X_n \in L^p$, $n = 1, 2, \dots$. Diremos que la sucesión $\{X_n\}$ converge en media p o en L^p a la variable aleatoria $X \in L^p$ si $E[|X_n - X|^p] \rightarrow 0$ cuando $n \rightarrow \infty$. En tal caso escribiremos $X_n \xrightarrow{L^p} X$.

OBSERVACIONES.

- La variable límite, si existe, es única c.s.[P].
- Si $p = 2$ este tipo de convergencia se suele denominar en media cuadrática.

DEFINICIÓN. Diremos que la sucesión de variables aleatorias $\{X_n\}$ **converge en distribución o en ley** a la variable X si

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = F_X(x)$$

para todo x punto de continuidad de F_X . En tal caso escribiremos $X_n \xrightarrow{d} X$.

Modos de Convergencia: Criterios Alternativos.

Convergencia Casi Segura.

PROPOSICIÓN 1. La sucesión de variables aleatorias $\{X_n\}$ converge c.s. a una variable aleatoria X si y sólo si para todo $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} P\left(\bigcup_{m=n}^{\infty} \{|X_m - X| \geq \epsilon\}\right) = 0.$$

OBSERVACIÓN. En la proposición realmente probamos que

$$X_n \xrightarrow{c.s.} X \text{ si y sólo si } \sup_{k \geq n} |X_k - X| \xrightarrow{P} 0.$$

COROLARIO. La sucesión $\{X_n\}$ converge c.s. a la variable aleatoria X si para todo $\epsilon > 0$,

$$\sum_{n=1}^{\infty} P(\{|X_n - X| \geq \epsilon\}) < +\infty.$$

OBSERVACIÓN. La sucesión $\{X_n\}$ se dice que *converge completamente* a la variable aleatoria X si

$$\sum_{n=1}^{\infty} P(\{|X_n - X| \geq \epsilon\}) < +\infty$$

para todo $\epsilon > 0$. El corolario nos indica que la convergencia completa implica la convergencia c.s.

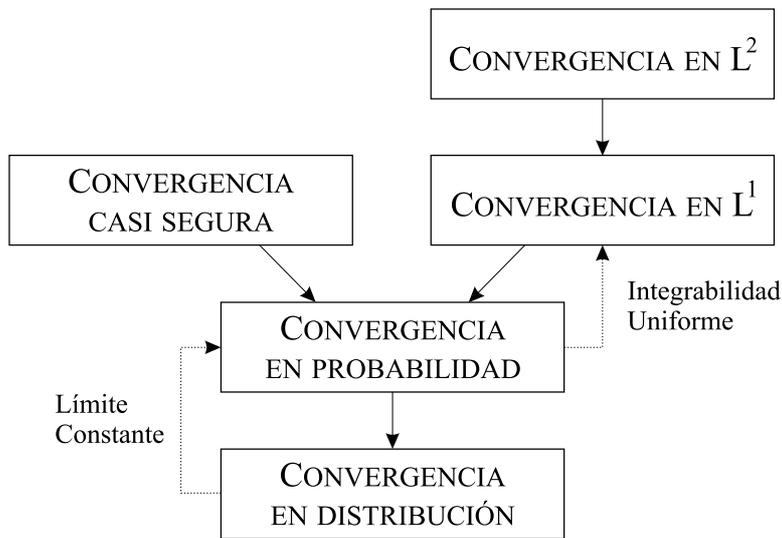
Convergencia en Distribución.

TEOREMA 2. (Teorema de continuidad de Lévy). Sea $\{X_n\}$ una sucesión de variables aleatorias con funciones características φ_{X_n} . Se verifican:

1. $X_n \xrightarrow{d} X$ si y sólo si $\varphi_{X_n}(t) \rightarrow \varphi_X(t)$ para todo $t \in \mathbb{R}$.
2. Si $\varphi(t) = \lim_{n \rightarrow \infty} \varphi_{X_n}(t)$ existe para todo $t \in \mathbb{R}$, y $\varphi(t)$ es continua en 0, entonces existe una variable aleatoria X con función característica $\varphi(t)$ y tal que $X_n \xrightarrow{d} X$.

OBSERVACIÓN. Caracterizaciones similares a la proporcionada por el teorema de continuidad de Lévy, aunque menos generales y de utilización más restringida, se pueden obtener a partir de las funciones generatrices de momentos y de probabilidad.

Relaciones entre los Modos de Convergencia.



PROPOSICIÓN 3. Sea $\{X_n\}$ una sucesión de variables aleatorias y X una variable aleatoria, todas ellas definidas sobre (Ω, \mathcal{F}, P) . Entonces, si $X_n \xrightarrow{c.s.} X$, se verifica que $X_n \xrightarrow{P} X$.

PROPOSICIÓN 4. Sea $\{X_n\}$ una sucesión de variables aleatorias tales que $X_n \in L^1, n \geq 1$, y $X \in L^1$. Entonces, si $X_n \xrightarrow{L^1} X$ se verifica que $X_n \xrightarrow{P} X$.

PROPOSICIÓN 5. Sea $1 \leq r < s, \{X_n\}$ una sucesión de variables aleatorias tales que $X_n \in L^s, n \geq 1$, y $X \in L^s$. Entonces, si $X_n \xrightarrow{L^s} X$ se verifica que $X_n \xrightarrow{L^r} X$.

TEOREMA 6. (Kolmogorov). Sea $\{X_n\}$ una sucesión de variables aleatorias tales que $X_n \xrightarrow{P} X$. Entonces $X_n \xrightarrow{d} X$.

PROPOSICIÓN 7. Sea $\{X_n\}$ una sucesión de variables aleatorias tales que $X_n \xrightarrow{d} c$. Entonces $X_n \xrightarrow{P} c$, donde c es una constante.

Convergencia bajo Transformaciones.

Convergencia bajo Transformaciones: Operaciones Algebraicas.

PROPOSICIÓN 8. Sean $\{X_n\}$ e $\{Y_n\}$ dos sucesiones de variables aleatorias y X e Y dos variables aleatorias todas ellas definidas sobre (Ω, \mathcal{F}, P) . Entonces se verifica que:

1. $X_n \xrightarrow{c.s.} X, Y_n \xrightarrow{c.s.} Y \implies X_n + Y_n \xrightarrow{c.s.} X + Y.$
2. $X_n \xrightarrow{P} X, Y_n \xrightarrow{P} Y \implies X_n + Y_n \xrightarrow{P} X + Y.$
3. Para $p \geq 1$, $X_n \xrightarrow{L^p} X, Y_n \xrightarrow{L^p} Y \implies X_n + Y_n \xrightarrow{L^p} X + Y.$

OBSERVACIÓN. La relación anterior no se verifica, en general, para la convergencia en distribución aunque sí es cierta cuando el límite de alguna de las sucesiones es constante.

PROPOSICIÓN 9. (Slutsky). Sean $\{X_n\}$ e $\{Y_n\}$ dos sucesiones de variables aleatorias y X una variable aleatoria, tales que $X_n \xrightarrow{d} X$ e $Y_n \xrightarrow{d} c$. Entonces se verifica que $X_n + Y_n \xrightarrow{d} X + c$.

PROPOSICIÓN 10. Sean $\{X_n\}$ e $\{Y_n\}$ dos sucesiones de variables aleatorias y X e Y dos variables aleatorias todas ellas definidas sobre (Ω, \mathcal{F}, P) . Entonces se verifica que:

1. $X_n \xrightarrow{c.s.} X, Y_n \xrightarrow{c.s.} Y \implies X_n Y_n \xrightarrow{c.s.} XY.$
2. $X_n \xrightarrow{P} X, Y_n \xrightarrow{P} Y \implies X_n Y_n \xrightarrow{P} XY.$
3. $X_n \xrightarrow{L^2} X, Y_n \xrightarrow{L^2} Y \implies X_n Y_n \xrightarrow{L^1} XY.$

PROPOSICIÓN 11. (Slutsky). Sean $\{X_n\}$ e $\{Y_n\}$ dos sucesiones de variables aleatorias y X una variable aleatoria, tales que $X_n \xrightarrow{d} X$ e $Y_n \xrightarrow{d} c$. Entonces se verifica que $X_n Y_n \xrightarrow{d} cX$.

Convergencia bajo Transformaciones: Funciones Continuas.

PROPOSICIÓN 12. Sea $g: \mathbb{R} \rightarrow \mathbb{R}$ una función continua. Entonces se verifica:

1. $X_n \xrightarrow{c.s.} X \implies g(X_n) \xrightarrow{c.s.} g(X).$
2. $X_n \xrightarrow{P} X \implies g(X_n) \xrightarrow{P} g(X).$
3. $X_n \xrightarrow{d} X \implies g(X_n) \xrightarrow{d} g(X).$

Ley Débil de los Grandes Números.

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad, y sea $\{X_n, n \geq 1\}$ una sucesión de variables aleatorias definidas sobre él. Sea $S_n = \sum_{i=1}^n X_i$, $n = 1, 2, \dots$, y sea $\{B_n\}_n$ una sucesión de constantes tales que $B_n > 0 \forall n$ y $B_n \rightarrow \infty$ cuando $n \uparrow \infty$

DEFINICIÓN. La sucesión de variables aleatorias $\{X_n\}$ se dice que obedece la **Ley Débil de los Grandes Números** respecto de la sucesión de constantes normalizadoras $\{B_n\}_n$, si existe una sucesión de constantes centralizadoras $\{A_n\}_n$ tal que

$$\frac{S_n - A_n}{B_n} \xrightarrow{P} 0$$

OBSERVACIÓN. Es habitual considerar $B_n = n$, $n = 1, 2, \dots$

TEOREMA 1. Sea $\{X_n\}_n$ una sucesión de variables aleatorias incorreladas dos a dos, con $\mu_i = E[X_i]$ y $\sigma_i^2 = \text{Var}[X_i] < \infty$, $i = 1, 2, \dots$. Si $\frac{1}{n^2} \sum_{i=1}^n \sigma_i^2 \rightarrow 0$ cuando $n \uparrow \infty$, entonces

$$\frac{S_n - \sum_{i=1}^n \mu_i}{n} \xrightarrow{P} 0$$

COROLARIO. (Chebyshev). Sea $\{X_n\}_n$ una sucesión de variables aleatorias incorreladas dos a dos e idénticamente distribuidas, con $\mu = E[X_i]$ y $\sigma^2 = \text{Var}[X_i] < \infty$, $i = 1, 2, \dots$. Entonces

$$\frac{S_n - n\mu}{n} \xrightarrow{P} 0$$

o equivalentemente

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

TEOREMA 2. (Khinchine). Sea $\{X_n\}_n$ una sucesión de variables aleatorias independientes e idénticamente distribuidas, con $\mu = E[X_i] < \infty$, $i = 1, 2, \dots$. Entonces

$$\frac{S_n - n\mu}{n} \xrightarrow{P} 0$$

o equivalentemente

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

Ley Fuerte de los Grandes Números.

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad, y sea $\{X_n, n \geq 1\}$ una sucesión de variables aleatorias definidas sobre él. Sea $S_n = \sum_{i=1}^n X_i$, $n = 1, 2, \dots$, y sea $\{B_n\}_n$ una sucesión de constantes tales que $B_n > 0 \forall n$ y $B_n \rightarrow \infty$ cuando $n \uparrow \infty$

DEFINICIÓN. La sucesión de variables aleatorias $\{X_n\}$ se dice que obedece la **Ley Fuerte de los Grandes Números** respecto de la sucesión de constantes normalizadoras $\{B_n\}_n$, si existe una sucesión de constantes centralizadoras $\{A_n\}_n$ tal que

$$\frac{S_n - A_n}{B_n} \xrightarrow{\text{c.s.}} 0$$

OBSERVACIÓN. Es habitual considerar $B_n = n$, $n = 1, 2, \dots$

TEOREMA 3. (Kolmogorov). Sea $\{X_n\}_n$ una sucesión de variables aleatorias independientes con $\mu_i = E[X_i]$ y $\sigma_i^2 = \text{Var}[X_i] < \infty$, $i = 1, 2, \dots$. Si $\sum_{n=1}^{\infty} \frac{\sigma_n^2}{n^2} < \infty$, entonces

$$\frac{S_n - \sum_{i=1}^n \mu_i}{n} \xrightarrow{\text{c.s.}} 0$$

COROLARIO. Sea $\{X_n\}_n$ una sucesión de variables aleatorias independientes con $\mu_i = E[X_i]$ y $\sigma_i^2 = \text{Var}[X_i] \leq A$, $i = 1, 2, \dots$ (i.e. varianzas uniformemente acotadas). Entonces

$$\frac{S_n - \sum_{i=1}^n \mu_i}{n} \xrightarrow{\text{c.s.}} 0$$

COROLARIO. Sea $\{X_n\}_n$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con $\mu = E[X_i]$ y $\sigma^2 = \text{Var}[X_i] < \infty$, $i = 1, 2, \dots$. Entonces

$$\frac{S_n - n\mu}{n} \xrightarrow{\text{c.s.}} 0$$

o equivalentemente

$$\frac{S_n}{n} \xrightarrow{\text{c.s.}} \mu$$

TEOREMA 4. Sea $\{X_n\}_n$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con $\mu = E[X_i] < \infty$, $i = 1, 2, \dots$. Entonces

$$\frac{S_n - n\mu}{n} \xrightarrow{\text{c.s.}} 0$$

o equivalentemente

$$\frac{S_n}{n} \xrightarrow{\text{c.s.}} \mu$$

Ley Fuerte de los Grandes Números.

RESULTADO DE BOREL. Sea $\{X_n\}_n$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con distribución $Be(p)$. Entonces

$$\frac{S_n}{n} \xrightarrow{c.s.} p$$

SIMULACIÓN DE n REPETICIONES INDEPENDIENTES DE UNA DISTRIBUCIÓN $Be(p)$.

```
borel<-function(p,n){  
  x<-numeric()  
  for(i in 1:n) x<-c(x,rbinom(1,1,p))  
  frec<-cumsum(x)/(1:n)  
  frec  
}
```

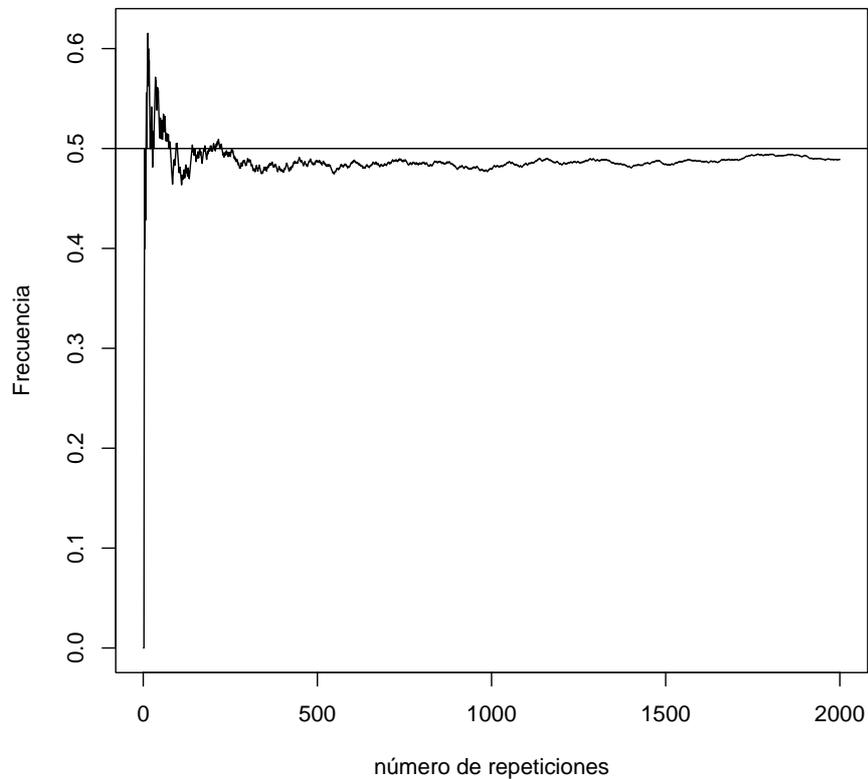


Fig. Simulación realizada para $p = 0.5$ y $n = 2000$

Teorema Central del Límite: Planteamiento Clásico.

Caso de variables aleatorias independientes e idénticamente distribuidas.

TEOREMA 1. Sea $\{X_n, n \geq 1\}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas, con distribución común $Be(p)$, $0 < p < 1$. Sea $S_n = \sum_{i=1}^n X_i$, $n = 1, 2, \dots$. Entonces se verifica:

1. (Bernoulli)

$$\frac{S_n}{n} \xrightarrow{d} p$$

2. (De Moivre-Laplace)

$$\frac{S_n - np}{\sqrt{np(1-p)}} \xrightarrow{d} Z, \quad Z \sim N(0, 1).$$

TEOREMA 2. (Poisson) Sean X_{n1}, \dots, X_{nn} variables aleatorias independientes e idénticamente distribuidas, con distribución común $Be(p_n)$, $0 < p_n < 1$, $n = 1, 2, \dots$. Sea $S_{nn} = \sum_{i=1}^n X_{ni}$, $n = 1, 2, \dots$. Entonces, si $np_n \rightarrow \lambda \in (0, +\infty)$ cuando $n \uparrow +\infty$, se verifica que

$$S_{nn} \xrightarrow{d} Y, \quad Y \sim P(\lambda).$$

TEOREMA 3. (Lindeberg-Levy) Sea $\{X_n, n \geq 1\}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas, con $E[X_i] = \mu$ y $Var[X_i] = \sigma^2 < +\infty$, $i = 1, 2, \dots$. Sea $S_n = \sum_{i=1}^n X_i$, $n = 1, 2, \dots$. Entonces se verifica:

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{d} Z, \quad Z \sim N(0, 1)$$

o equivalentemente

$$\frac{\bar{X}_n - \mu}{\sqrt{\frac{\sigma^2}{n}}} \xrightarrow{d} Z, \quad Z \sim N(0, 1),$$

siendo $\bar{X}_n = S_n/n$, $n = 1, 2, \dots$

Caso de variables aleatorias independientes.

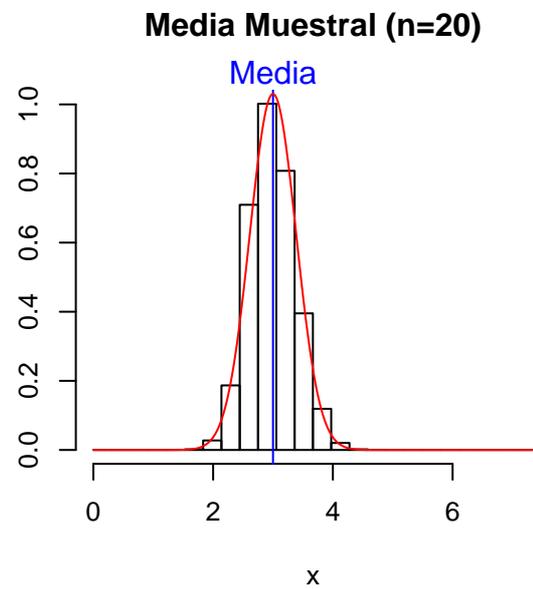
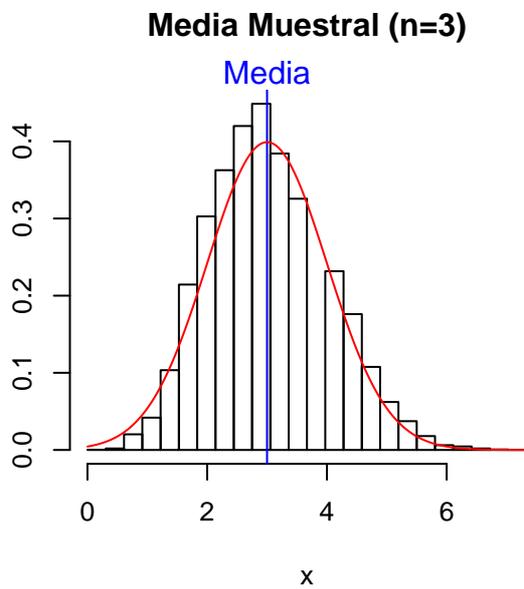
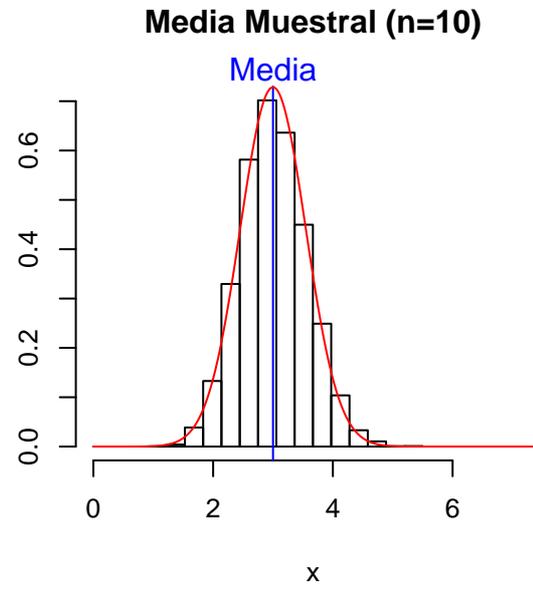
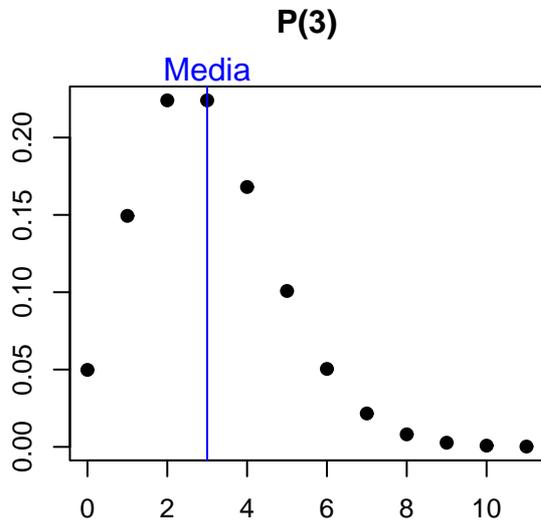
TEOREMA 4. (Lindeberg-Feller) Sea $\{X_n, n \geq 1\}$ una sucesión de variables aleatorias independientes con F_i la función de distribución de X_i , $E[X_i] = \mu_i$ y $Var[X_i] = \sigma_i^2 < +\infty$, $i = 1, 2, \dots$. Sea $S_n = \sum_{i=1}^n X_i$, $n = 1, 2, \dots$. Entonces se verifica:

$$\frac{S_n - E[S_n]}{\sqrt{Var[S_n]}} \xrightarrow{d} Z, \quad Z \sim N(0, 1) \quad y \quad \max_{1 \leq k \leq n} P\left(\frac{|X_k - \mu_k|}{\sqrt{Var[S_n]}} \geq \epsilon\right) \rightarrow 0 \text{ si } n \uparrow \infty \quad \forall \epsilon$$

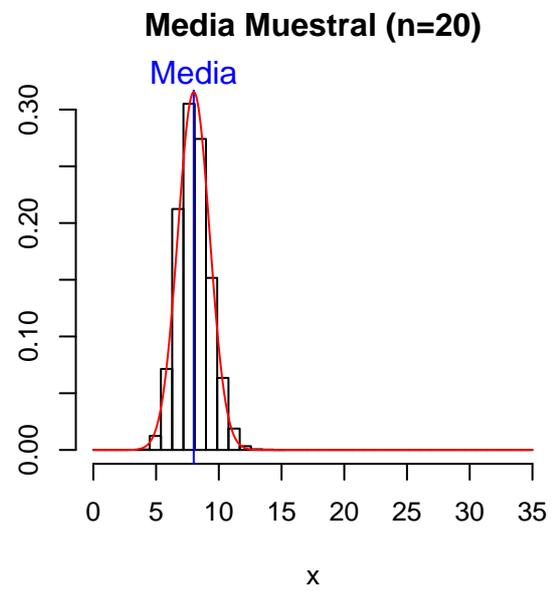
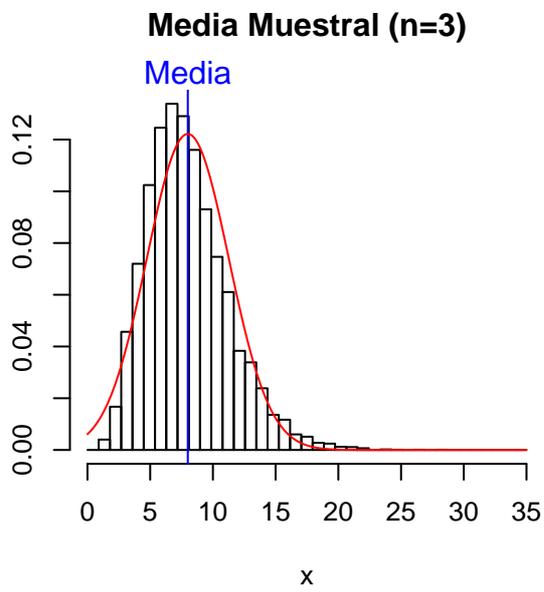
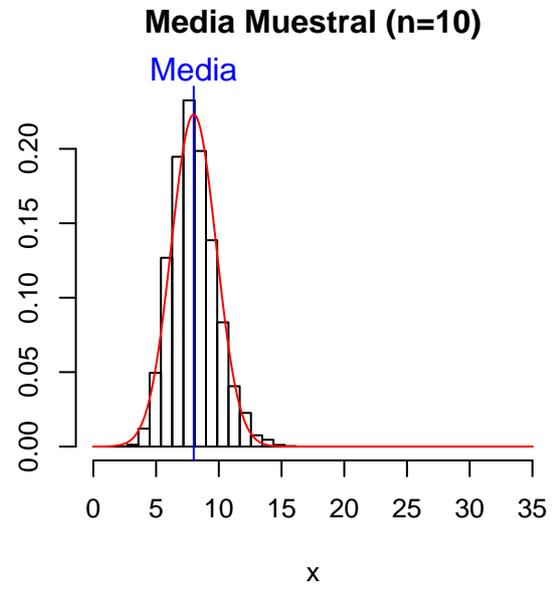
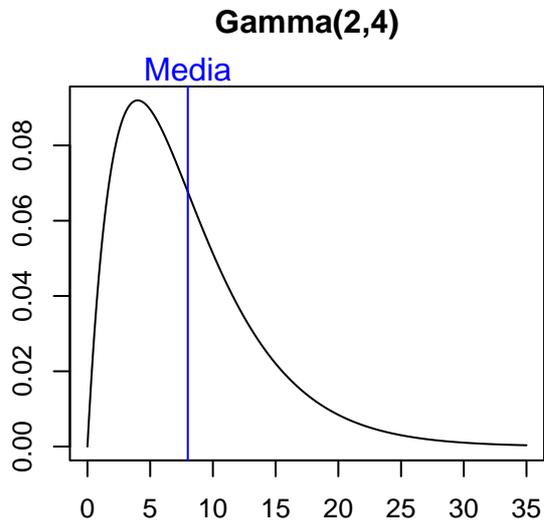
si y sólo si

$$\frac{1}{Var[S_n]} \sum_{k=1}^n \int_{\{x: |x - \mu_k| \geq \epsilon \sqrt{Var[S_n]}\}} (x - \mu_k)^2 dF_k(x) \rightarrow 0 \text{ si } n \uparrow \infty \quad \forall \epsilon$$

Teorema Central del Límite



Teorema Central del Límite



Procesos Estocásticos. Generalidades.

DEFINICIÓN. Dado un conjunto infinito T un **proceso estocástico** con conjunto índice T , es una familia de v.a. $X = \{X_t: t \in T\}$ definidas sobre el mismo espacio probabilístico (Ω, \mathcal{F}, P) y tomando valores en un espacio medible (S, \mathcal{A}) .

OBSERVACIÓN:

- El conjunto de índices T puede ser infinito numerable (habitualmente \mathbb{N}), en cuyo caso hablaremos de **proceso estocástico en tiempo discreto**; o un conjunto infinito no numerable (usualmente $T = \mathbb{R}$, $T = [0, \infty)$ ó $T = [0, a)$, $0 < a < \infty$), y hablaremos entonces de **procesos estocásticos en tiempo continuo**.
- El conjunto S , usualmente \mathbb{R} o un subconjunto suyo, se denomina **espacio de estados** del proceso, y dependiendo de su naturaleza discreta o continua hablaremos de **procesos estocásticos con espacio de estados discreto o continuo**, respectivamente. Cuando el espacio de estados sea \mathbb{R}^n o algún subconjunto suyo diremos que el proceso es multidimensional o n-dimensional.

Un proceso estocástico $X = \{X_t: t \in T\}$ puede ser considerado como una aplicación que depende de dos argumentos:

$$\mathcal{X}: T \times \Omega \rightarrow S \quad (t, \omega) \rightarrow \mathcal{X}(t, \omega) = X_t(\omega)$$

Considerando t fijo obtenemos $\mathcal{X}(t, \cdot) = X_t(\cdot)$ que es una v.a. definida sobre (Ω, \mathcal{F}, P) y tomando valores en (S, \mathcal{A}) . Si fijamos ω y dejamos variable t lo que obtenemos es una aplicación

$$\mathcal{X}(\cdot, \omega): T \rightarrow S \quad t \rightarrow X_t(\omega)$$

que denominaremos **trayectoria** o **realización muestral** asociada a ω .

DEFINICIÓN: Sea $\{X_t: t \in T\}$ un proceso estocástico y $\{t_1, \dots, t_n\}$ un subconjunto finito cualquiera de T . La distribución multivariante del vector aleatorio $(X_{t_1}, \dots, X_{t_n})$ se denomina una **distribución finito dimensional del proceso**.

Distribución de un Proceso Estocástico en Tiempo Discreto.

Dado un espacio probabilístico (Ω, \mathcal{F}, P) , se entiende por Proceso Estocástico en Tiempo Discreto cualquier sucesión de v.a. $\{X_n\}_{n \in \mathbb{N}}$ definidas todas ellas sobre dicho espacio. Consideraremos v.a. reales, i.e. $X_n: (\Omega, \mathcal{F}, P) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Para cada $\omega \in \Omega$, el proceso estocástico $\{X_n\}_{n \in \mathbb{N}}$ nos da una sucesión de números reales que se denomina trayectoria del proceso asociada a ω . Así podemos considerar la aplicación $X: \Omega \rightarrow \mathbb{R}^{\mathbb{N}}$, $\omega \rightarrow \{X_n(\omega)\}_n$. Para estudiar a fondo esta aplicación dotamos a $\mathbb{R}^{\mathbb{N}}$ de una σ -álgebra, $\mathcal{B}(\mathbb{R}^{\mathbb{N}})$, generada por los conjuntos cilíndricos, es decir conjuntos de la forma:

$$B = \{\{x_i\}_{i \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}: x_1 \in B_1, \dots, x_n \in B_n\}$$

para n arbitrario y $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$. En estas condiciones X es una v.a. (debido a que lo son las X_n).

DEFINICIÓN: Llamaremos **distribución del proceso** a la probabilidad P^X inducida por X en $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}(\mathbb{R}^{\mathbb{N}}))$, es decir

$$P^X(B) = P(X^{-1}(B)) \quad B \in \mathcal{B}(\mathbb{R}^{\mathbb{N}}).$$

DEFINICIÓN: Dos procesos estocásticos en tiempo discreto $X = \{X_n\}_n$ sobre (Ω, \mathcal{F}, P) e $Y = \{Y_n\}_n$ sobre $(\Omega', \mathcal{F}', P')$ se dicen **equivalentes** si dan lugar a la misma distribución del proceso.

Proceso de Bernoulli.

DEFINICIÓN: Llamaremos **proceso de Bernoulli** de parámetro p , ($0 < p < 1$), a una sucesión, $\{X_n\}_n$, de v.a. independientes, cada una con distribución $Be(p)$.

OBSERVACIONES: Asociados a este proceso podemos construir otros dos de interés:

1. Proceso de conteo de éxitos: $\{S_n\}_n$, donde $S_n = \sum_{i=1}^n X_i$, $n = 1, 2, \dots$
2. Proceso del tiempo de ocurrencia de éxitos: $\{T_k\}_k$, donde $T_k = \min\{n: S_n = k\}$, $k = 1, 2, \dots$ (en relación a este último es interesante el proceso $\{U_k\}_k$, donde $U_1 = T_1$, $U_k = T_k - T_{k-1}$, $k = 2, 3, \dots$ (tiempo transcurrido entre éxitos consecutivos)).

PROPOSICIÓN: Para cada k , T_k tiene distribución binomial negativa de parámetros k y p , i.e.

$$P(T_k = m) = \binom{m-1}{k-1} p^k (1-p)^{m-k} \quad m \geq k.$$

En particular $T_1 = U_1$ tiene una distribución geométrica de parámetro p .

PROPOSICIÓN: El proceso $\{S_n\}_n$ tiene incrementos independientes y estacionarios:

1. Para $0 < n_1 < \dots < n_k$, las v.a. $S_{n_1}, S_{n_2} - S_{n_1}, \dots, S_{n_k} - S_{n_{k-1}}$ son independientes.
2. Para un j fijo, la distribución de $S_{k+j} - S_k$ es la misma para todo k .

TEOREMA: Las v.a. U_1, U_2, \dots son independientes e idénticamente distribuidas.

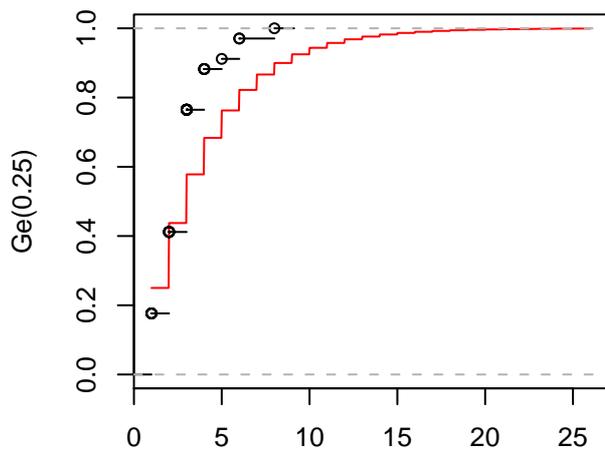
Proceso de Bernoulli

```

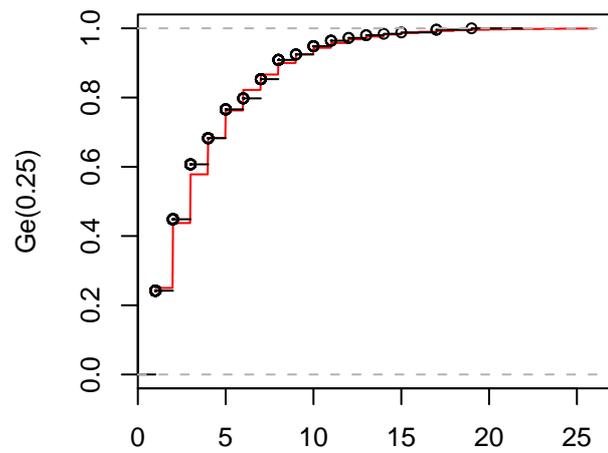
pe.bernoulli<-function(n,p){
  x<-rbinom(n,1,p)
  tk<-(1:n)[x==1]
  uj<-c(tk[1],tk[2:length(tk)]-tk[1:(length(tk)-1)])
  frec<-round(table(uj)/length(uj),4)
  list(uj=uj, frec=frec)
}

```

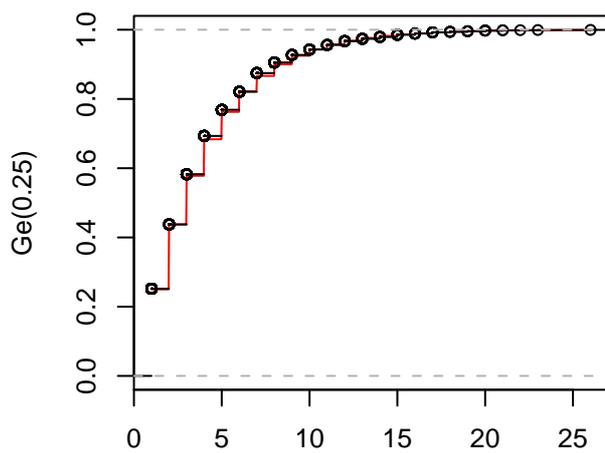
n=100



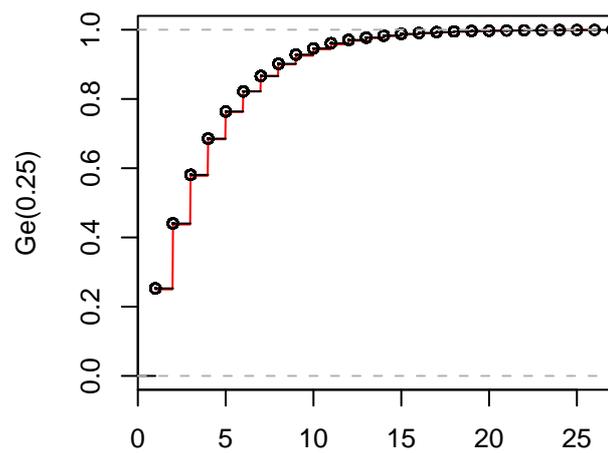
n=1000



n=10000



n=100000



Proceso de Bernoulli

```

> dgeom(0:25, .25)
 [1] 0.2500000000 0.1875000000 0.1406250000 0.1054687500 0.0791015625
 [6] 0.0593261719 0.0444946289 0.0333709717 0.0250282288 0.0187711716
 [11] 0.0140783787 0.0105587840 0.0079190880 0.0059393160 0.0044544870
 [16] 0.0033408653 0.0025056489 0.0018792367 0.0014094275 0.0010570706
 [21] 0.0007928030 0.0005946022 0.0004459517 0.0003344638 0.0002508478
 [26] 0.0001881359

> b.100$frec
uj
  1    2    3    4    5    6    8
0.1765 0.2353 0.3529 0.1176 0.0294 0.0588 0.0294

> b.1000$frec
uj
  1    2    3    4    5    6    7    8    9    10    11
0.2421 0.2063 0.1587 0.0754 0.0833 0.0317 0.0556 0.0556 0.0159 0.0238 0.0159
 12    13    14    15    17    19
0.0079 0.0079 0.0040 0.0040 0.0079 0.0040

> b.10000$frec
uj
  1    2    3    4    5    6    7    8    9    10    11
0.2520 0.1856 0.1451 0.1105 0.0755 0.0521 0.0541 0.0306 0.0219 0.0155 0.0131
 12    13    14    15    16    17    18    19    20    21    22
0.0111 0.0068 0.0052 0.0064 0.0036 0.0036 0.0016 0.0016 0.0020 0.0008 0.0004
 23    26    30
0.0004 0.0004 0.0004

> b.100000$frec
uj
  1    2    3    4    5    6    7    8    9    10    11
0.2529 0.1874 0.1405 0.1046 0.0784 0.0581 0.0444 0.0353 0.0265 0.0180 0.0148
 12    13    14    15    16    17    18    19    20    21    22
0.0095 0.0068 0.0058 0.0048 0.0029 0.0023 0.0019 0.0013 0.0008 0.0004 0.0006
 23    24    25    26    27    28    30    31    33    35    37
0.0003 0.0005 0.0002 0.0002 0.0001 0.0001 0.0001 0.0000 0.0000 0.0000 0.0000

```

Proceso de Poisson.

Un proceso de contaje de llegadas es un proceso estocástico $N = \{N_t\}_{t \geq 0}$ tal que para cada ω ,

1. $N_0(\omega) = 0$
2. La trayectoria $t \rightarrow N_t(\omega)$, es una función de \mathbb{R}^+ en \mathbb{R} , creciente y continua por la derecha, que se incrementa en saltos de magnitud 1, y tiene sólo un número finito de saltos en cualquier intervalo de tiempo acotado.

La interpretación es que N_t es el número de "llegadas" en algún sistema (p. ej. una cola), o el número de sucesos de algún tipo que han ocurrido, durante el intervalo $(0, t]$. De manera más general, para $s < t$, $N_t - N_s$ es el número de "llegadas" ocurridas en el intervalo de tiempo $(s, t]$.

Asociadas con un proceso de contaje de "llegadas" hay dos sucesiones importantes de v.a., representando "tiempo de llegada" y "tiempo entre llegadas". Para cada k , $T_k = \inf\{t: N_t = k\}$ es el tiempo de la k -ésima llegada. Las v.a. $U_k = T_k - T_{k-1}$ (con $T_0 = 0$) son tiempos entre sucesivas llegadas.

DEFINICIÓN: Un proceso de contaje de llegadas $N = \{N_t\}_t$ es un **proceso de Poisson con tasa $\lambda > 0$** si

1. N tiene incrementos independientes, i.e. para $0 < t_1 < \dots < t_n$, las v.a. $N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_n} - N_{t_{n-1}}$ son independientes.
2. N tiene incrementos estacionarios, i.e. para un s fijo, la distribución de $N_{t+s} - N_t$ es la misma para todo t .
3. Para cada t , N_t sigue distribución de Poisson de parámetro λt .

OBSERVACIÓN: En realidad la tercera condición es consecuencia de las dos primeras.

PROPOSICIÓN: Para $0 < t_1 < \dots < t_n$ y $k_1 \leq \dots \leq k_n$,

$$P(N_{t_1} = k_1, \dots, N_{t_n} = k_n) = \prod_{m=1}^n \exp\{-\lambda(t_m - t_{m-1})\} \frac{(\lambda(t_m - t_{m-1}))^{k_m - k_{m-1}}}{(k_m - k_{m-1})!}$$

PROPOSICIÓN: Para cada k , T_k es absolutamente continua con

$$f_{T_k}(t) = \frac{1}{(k-1)!} \lambda^k e^{-\lambda t} t^{k-1}, \quad t > 0.$$

PROPOSICIÓN: Las v.a. U_1, U_2, \dots son independientes e idénticamente distribuidas con distribución común exponencial de parámetro $1/\lambda$.

Movimiento Browniano.

DEFINICIÓN: Un proceso estocástico $X = \{X_t\}_{t \geq 0}$ es un **Movimiento Browniano** si

1. $X_0 = 0$
2. X tiene incrementos independientes, i.e. para $0 < t_1 < \dots < t_n$, las v.a. $X_{t_1}, X_{t_2} - X_{t_1}, \dots, X_{t_n} - X_{t_{n-1}}$ son independientes.
3. X tiene incrementos estacionarios, i.e. para un s fijo, la distribución de $X_{t+s} - X_t$ es la misma para todo t .
4. Para cada t , X_t sigue distribución $N(0, c^2 t)$.

OBSERVACIONES:

- El Movimiento Browniano también se denomina **Proceso de Wiener**.
- Si $s < t$, entonces $X_t - X_s \sim N(0, c^2(t - s))$.
- Cuando $c = 1$ el proceso se denomina Movimiento Browniano Estándar.
- Las trayectorias de un movimiento browniano son continuas.

PROPOSICIÓN: Sea $\{X_t\}_t$ un movimiento browniano estándar. Cualesquiera que sean $0 < t_1 < \dots < t_n$ la distribución del vector aleatorio $(X_{t_1}, \dots, X_{t_n})$ viene dada por la función de densidad conjunta:

$$f(x_1, \dots, x_n) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left\{-\frac{1}{2} \frac{(x_i - x_{i-1})^2}{t_i - t_{i-1}}\right\}$$

con $t_0 = x_0 = 0$ y $x_1, \dots, x_n \in \mathbb{R}$

Definición de Cadena de Markov en Tiempo Discreto

Sea (Ω, \mathcal{F}, P) un espacio de probabilidad y $\{X_n\}_{n \geq 0}$ un proceso estocástico en tiempo discreto construido sobre dicho espacio. Si todas las variables son discretas, la unión de todos los valores posibles de todas las X_n será el espacio de estados S , un conjunto numerable tal que $i \in S$ si y sólo si existe $n \geq 0$ tal que $P(X_n = i) > 0$.

DEFINICIÓN: Una **Cadena de Markov en tiempo discreto** es una sucesión de v.a. discretas $\{X_n\}_{n \geq 0}$ tales que $\forall n \geq 2, \forall t_1, \dots, t_n \in \mathbb{N}_0$ tales que $0 \leq t_1 < \dots < t_n$ y $\forall i_1, \dots, i_n \in S$ se verifica:

$$P(X_{t_n} = i_n | X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}) = P(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1})$$

siempre que el miembro de la izquierda esté bien definido.

OBSERVACIÓN: La condición anterior se conoce como **propiedad de Markov**. Es equivalente a la siguiente condición aparentemente más débil: $\forall n \geq 1$ y $\forall i_0, \dots, i_n \in S$ se verifica

$$P(X_n = i_n | X_0 = i_0, \dots, X_{n-1} = i_{n-1}) = P(X_n = i_n | X_{n-1} = i_{n-1})$$

PROPOSICIÓN: Para cualquier $m \geq 0$ y $0 \leq t_1 < \dots < t_n < \dots < t_{n+m}$ y cualesquiera $i_1, \dots, i_n, \dots, i_{n+m} \in S$ se verifica

$$P(X_{t_r} = i_r, n \leq r \leq n+m | X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}) = P(X_{t_r} = i_r, n \leq r \leq n+m | X_{t_{n-1}} = i_{n-1})$$

Probabilidades de Transición.

Sea $\{X_n\}_{n \geq 0}$ una cadena de Markov con espacio de estados S . Consideraremos fijada una enumeración cualquiera en S . A la distribución de probabilidad de la variable X_0 la llamaremos **distribución inicial** del proceso y vendrá dada por la sucesión $\{p_i, i \in S\}$, siendo $p_i = P(X_0 = i), i \in S$.

PROPOSICIÓN: Cualesquiera que sean $n \geq 0$ e $i_0, \dots, i_n \in S$ se verifica:

$$P(X_r = i_r, 0 \leq r \leq n) = p_{i_0} \prod_{t=1}^n p_{i_{t-1} i_t}(t)$$

donde $p_{i_{t-1} i_t}(t) = P(X_t = i_t | X_{t-1} = i_{t-1})$ (probabilidades de transición)

DEFINICIÓN: Diremos que la cadena de Markov $\{X_n\}_{n \geq 0}$ tiene **probabilidades de transición estacionarias** o es **homogénea** si cualesquiera que sean $n \geq 1$ e $i, j \in S$ se verifica:

$$P(X_n = j | X_{n-1} = i) = p_{ij}$$

independiente del valor de n . A la matriz $P = (p_{ij})_{i,j \in S}$ la llamaremos **matriz de transición de un paso**.

OBSERVACIONES:

- Si $\{X_n\}_n$ es una cadena de Markov con probabilidades de transición estacionarias entonces

$$P(X_r = i_r, 0 \leq r \leq n) = p_{i_0} \prod_{t=1}^n p_{i_{t-1} i_t}$$

- La matriz de transición de un paso P verifica

$$p_{ij} \geq 0 \quad \sum_{j \in S} p_{ij} = 1 \quad \forall i \in S$$

Cualquier matriz que verifique las dos condiciones anteriores se denomina **matriz estocástica**.

DEFINICIÓN: Definamos de modo recursivo las **probabilidades de transición en n-pasos** $p_{ij}^{(n)}$ para una cadena de Markov con probabilidades de transición estacionarias. En primer lugar definamos $p_{ij}^{(0)} = \delta_{ij}$. Denotemos $p_{ij}^{(1)} = p_{ij}$ y para $n \geq 0$ definamos

$$p_{ij}^{(n+1)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj}^{(1)}$$

PROPOSICIÓN: Cualesquiera que sean $i, j \in S$ y $n \geq 0$, se verifica que $p_{ij}^{(n)} = P(X_{r+n} = j | X_r = i), \forall r \geq 0$.

PROPOSICIÓN: (**Ecuaciones de Chapman-Kolmogorov**). Cualesquiera que sean $i, j \in S$, se verifica que $\forall n \geq 0, \forall m \geq 0$

$$p_{ij}^{(n+m)} = \sum_{k \in S} p_{ik}^{(n)} p_{kj}^{(m)}$$

OBSERVACIÓN:

- Si $P^{(n)} = (p_{ij}^{(n)})_{i,j \in S}$ es la matriz de transición en n -pasos ($P^{(1)} = P$) las ecuaciones de Chapman-Kolmogorov nos dicen que:

1. $P^{(n)} = P^n, \forall n \geq 1$.
2. $P^{(n+m)} = P^{(n)} P^{(m)} = P^{(m)} P^{(n)}, \forall m, n \geq 0$.

- La distribución de probabilidad de cada X_n viene dada por $p_i^{(n)} = P(X_n = i) = \sum_{j \in S} p_j p_{ji}^{(n)}, n \geq 0, i \in S$

Clasificación de los Estados.

DEFINICIÓN: Si $\{X_n\}_n$ es una cadena de Markov con probabilidades de transición estacionarias definimos

$$f_{ij}^{(m)} = \begin{cases} 0 & \text{si } m = 0 \\ P(X_{n+1} = j | X_n = i) & \text{si } m = 1 \\ P(X_{n+m} = j, X_{n+r} \neq j \ 0 < r < m | X_n = i) & \text{si } m = 2, 3, \dots \end{cases}$$

Consideremos, sin pérdida de generalidad, que $P(X_0 = i) = 1$, para cierto $i \in S$. Denotaremos

$$f_{ii} = \sum_{k=1}^{\infty} f_{ii}^{(k)} = P(X_{n+k} = i \text{ para algún } k | X_n = i)$$

DEFINICIÓN: El estado i se dice **transitorio** si $f_{ii} < 1$. Por el contrario lo llamaremos **recurrente** si $f_{ii} = 1$. En tal caso $\{f_{ii}^{(k)} : k = 1, 2, \dots\}$ es una distribución de probabilidad llamada tiempo de retorno al estado i , cuyo momento de orden uno será el tiempo medio de recurrencia al estado i , $\mu_i = \sum_{k=1}^{\infty} k f_{ii}^{(k)} \leq \infty$. Si $\mu_i < \infty$ diremos que i es un estado **recurrente positivo**. Por el contrario si $\mu_i = \infty$ diremos que i es estado **recurrente nulo**.

DEFINICIÓN: Cuando $p_{ii}^{(k)} > 0$ para algún k , llamaremos $\lambda_i = m.c.d.\{k : p_{ii}^{(k)} > 0\}$. Si $\lambda_i > 1$, diremos que i es un estado **periódico** de periodo λ_i . Si $\lambda_i = 1$, diremos que i es un estado **aperiódico**.

DEFINICIÓN: Un estado recurrente positivo y aperiódico diremos que es un estado **ergódico**.

TEOREMA: Un estado i es transitorio si y sólo si $\sum_{k=0}^{\infty} p_{ii}^{(k)} < \infty$

PROPOSICIÓN:

1. Un estado i es transitorio si y sólo si $g_{ii} = P(\limsup\{X_n = i\} | X_0 = i) = 0$
2. Un estado i es recurrente si y sólo si $g_{ii} = P(\limsup\{X_n = i\} | X_0 = i) = 1$

Clasificación de las Cadenas.

DEFINICIÓN: Diremos que un estado i **conduce** a un estado j ($i \rightsquigarrow j$) si existe $n \geq 0$ tal que $p_{ij}^{(n)} > 0$. Si $i \rightsquigarrow j$ y $j \rightsquigarrow i$ diremos que los estados son **comunicantes** ($i \leftrightarrow j$).

OBSERVACIÓN: \leftrightarrow define una relación de equivalencia sobre el espacio de estados S .

TEOREMA: Si en una cadena de Markov con probabilidades de transición estacionarias un estado recurrente i conduce a cualquier otro estado j , entonces dichos estados son comunicantes y del mismo tipo, i.e. ambos son recurrentes positivos o recurrentes nulos, y en caso de periodicidad ambos tienen el mismo periodo.

DEFINICIÓN: Un conjunto de estados, C , se dice

1. **cerrado** si $p_{ij}^{(n)} = 0, \forall n = 1, 2, \dots, \forall i \in C, \forall j \notin C$.
2. **irreducible** si $i \leftrightarrow j, \forall i, j \in C$.

OBSERVACIONES:

1. Si un conjunto cerrado de estados contiene un sólo elemento, a dicho estado se le llama **absorbente**.
2. Llamaremos a un conjunto de estados irreducible recurrente nulo (ó recurrente positivo ó transitorio, etc.) si todos los estados que lo constituyen son de este tipo.
3. Diremos que la cadena de Markov es irreducible si el espacio de estados, S , es irreducible. Una cadena de Markov irreducible se dice ergódica si todos sus estados son ergódicos, i.e. recurrentes positivos y aperiódicos.

TEOREMA: El espacio de estados S se puede dividir, de manera única en subconjuntos disjuntos

$$S = T \cup C_1 \cup C_2 \cup \dots,$$

donde T es el conjunto de estados transitorios, y los C_i son conjuntos cerrados e irreducibles de estados recurrentes.

Distribuciones Estacionarias.

DEFINICIÓN: Una cadena de Markov $\{X_n: n \geq 0\}$ tal que $p_i^{(n)} = P(X_n = i) = P(X_0 = i) = p_i, \forall i \in S, \forall n$, se denomina **cadena de Markov estacionaria**.

OBSERVACIÓN: Una Cadena de Markov homogénea con distribución inicial p y matriz de transición P es estacionaria si y sólo si $p = pP$

DEFINICIÓN: Una distribución de probabilidad $\pi = \{\pi_i\}_{i \in S}$ sobre S es una **distribución estacionaria** respecto a una cadena de Markov de matriz de transición P o simplemente una distribución estacionaria respecto a P , si $\pi P = \pi$, i.e.

$$\pi_j = \sum_{i \in S} \pi_i p_{ij} \quad \forall j \in S.$$

PROPOSICIÓN: Sea $\{X_n: n \geq 0\}$ una cadena de Markov con probabilidades de transición estacionarias, irreducible y con matriz de transición P . Existe una distribución estacionaria respecto a P si y sólo si $\{X_n: n \geq 0\}$ es recurrente positiva. En tal caso $\pi_j = 1/\mu_j \forall j \in S$.

OBSERVACIONES:

1. Cualquier distribución estacionaria, π , respecto a P verifica que $\sum_{i \in T} \pi_i = 0$.
2. Una condición necesaria y suficiente para que una cadena tenga distribuciones estacionarias es que tenga alguna subclase irreducible y cerrada recurrente positiva.
3. Si una cadena tiene varias subclases recurrentes positivas, las distribuciones estacionarias vendrán dadas por las combinaciones lineales convexas de las distribuciones estacionarias asociadas a cada una de las subclases recurrentes positivas.
4. Si el conjunto de estados de una cadena de Markov con probabilidades de transición estacionarias irreducible es finito, entonces es recurrente positiva.

Comportamiento Asintótico de una Cadena de Markov.

TEOREMA. Para una cadena de Markov con probabilidades de transición estacionarias se verifica:

1. Si i es un estado recurrente positivo y aperiódico

$$\lim_{n \rightarrow \infty} p_{ii}^{(n)} = \pi_i = 1/\mu_i, \quad \lim_{n \rightarrow \infty} p_{ji}^{(n)} = f_{ji}\pi_i \quad \forall j \in S \quad (f_{ji} = \sum_{m=1}^{\infty} f_{ji}^{(m)})$$

2. Si i es un estado recurrente positivo y de periodo λ_i

$$\lim_{n \rightarrow \infty} p_{ii}^{(n\lambda_i)} = \lambda_i\pi_i = \lambda_i/\mu_i, \quad (p_{ii}^{(m)} = 0, m \neq n\lambda_i)$$

$$\lim_{n \rightarrow \infty} p_{ji}^{(n\lambda_i+r)} = \lambda_i\pi_i \sum_{n=0}^{\infty} f_{ji}^{(n\lambda_i+r)} \quad \forall r = 1, \dots, \lambda_i, \quad \forall j \in S$$

3. Si i es un estado recurrente nulo o transitorio

$$\lim_{n \rightarrow \infty} p_{ji}^{(n)} = 0 \quad \forall j \in S$$

TEOREMA. Para una cadena de Markov con probabilidades de transición estacionarias irreducible y aperiódica se verifica:

1. Si los estados son recurrentes positivos, entonces $\lim_{n \rightarrow \infty} p_j^{(n)} = \pi_j, \quad \forall j \in S$.

2. Si los estados son recurrentes nulos o transitorios, entonces $\lim_{n \rightarrow \infty} p_j^{(n)} = 0, \quad \forall j \in S$.

En general para una cadena de Markov con probabilidades de transición estacionarias se verifica:

1. Si i es un estado recurrente positivo aperiódico, entonces

$$\lim_{n \rightarrow \infty} p_i^{(n)} = \pi_i \left(\sum_{j \in C} p_j + \sum_{j \in T} f_{ji} p_j \right),$$

siendo C la clase irreducible y cerrada a la que pertenece i y T la clase de estados transitorios.

2. Si i es un estado recurrente nulo o transitorio, entonces $\lim_{n \rightarrow \infty} p_i^{(n)} = 0$.

Proceso de Ramificación de Galton-Watson.

DEFINICIÓN. Sea $\{X_{ni}: i = 1, 2, \dots, n = 0, 1, \dots\}$ una sucesión de variables aleatorias i.i.d. sobre los enteros no negativos, \mathbb{Z}_+ , con ley de probabilidad $\{p_k\}_{k \in \mathbb{Z}_+}$. Definimos de manera recursiva el proceso $\{Z_n: n = 0, 1, \dots\}$ del siguiente modo:

$$Z_0 = 1 \quad Z_{n+1} = \sum_{i=1}^{Z_n} X_{ni} \quad n = 0, 1, \dots$$

Al proceso $\{Z_n: n = 0, 1, \dots\}$ lo llamaremos **proceso de ramificación de Galton-Watson**.

PROPOSICIÓN. $\{Z_n: n = 0, 1, \dots\}$ es una cadena de Markov con probabilidades de transición estacionarias.

Supongamos, para evitar trivialidades, que $p_0 + p_1 < 1$, entonces

PROPOSICIÓN. 0 es un estado absorbente y cualquier $k \neq 0$ es un estado transitorio.

Definición de Cadena de Markov en Tiempo Continuo

DEFINICIÓN: Sea $T = [0, +\infty)$. Diremos que un proceso estocástico $\{X_t: t \in T\}$ con espacio de estados numerable, S , es una **Cadena de Markov en tiempo continuo** si $\forall n \geq 2, \forall t_1, \dots, t_n \in T$ tales que $0 \leq t_1 < \dots < t_n$ y $\forall i_1, \dots, i_n \in S$ se verifica:

$$P(X_{t_n} = i_n | X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}) = P(X_{t_n} = i_n | X_{t_{n-1}} = i_{n-1})$$

siempre que el miembro de la izquierda esté bien definido.

OBSERVACIONES:

- Que S sea discreto (numerable) no queda garantizado por el hecho de que todas las variables X_t sean discretas. Para garantizar la numerabilidad de S consideraremos, de ahora en adelante, que el proceso tiene trayectorias continuas por la derecha, i.e. para todo $\omega \in \Omega$ y $t \in T$ existe $\varepsilon > 0$ tal que $X_t(\omega) = X_s(\omega)$ para $t \leq s \leq t + \varepsilon$.
- Para cualquier $m \geq 0$ y $0 \leq t_1 < \dots < t_n < \dots < t_{n+m}$ y cualesquiera $i_1, \dots, i_n, \dots, i_{n+m} \in S$ se verifica

$$P(X_{t_r} = i_r, n \leq r \leq n+m | X_{t_1} = i_1, \dots, X_{t_{n-1}} = i_{n-1}) = P(X_{t_r} = i_r, n \leq r \leq n+m | X_{t_{n-1}} = i_{n-1})$$

- En general, la definición anterior es equivalente a que si $t \in T$

$$P(M | \sigma(X_s: s \leq t)) = P(M | X_t) \text{ c.s. } \forall M \in \sigma(X_s: s > t)$$

Probabilidades de Transición. Matrices de Transición.

DEFINICIÓN. Diremos que la cadena de Markov $\{X_t: t \in T\}$ es **homogénea** o que las **probabilidades de transición son estacionarias** si $P(X_{s+t} = j | X_s = i)$ no depende de s , cualesquiera que sean $i, j \in S$. A la función $p_{ij}(t) = P(X_{s+t} = j | X_s = i)$, $t \geq 0$ la llamaremos **función de probabilidad de transición** desde el estado i al estado j . Al valor de la función en el punto t lo llamaremos **probabilidad de transición del estado i al estado j en un tiempo t** .

PROPOSICIÓN: Sea $\{X_t: t \in T\}$ una cadena de Markov con probabilidades de transición estacionarias. Si $0 \leq t_0 < t_1 < \dots < t_n \in T$ se verifica que

$$P(X_{t_1} = i_1, \dots, X_{t_n} = i_n | X_{t_0} = i_0) = \prod_{\nu=1}^n p_{i_{\nu-1}i_\nu}(t_\nu - t_{\nu-1}) \quad i_0, \dots, i_n \in S$$

PROPOSICIÓN: (**Ecuaciones de Chapman-Kolmogorov**). Sea $p_{ij}(0) = \delta_{ij}$, $i, j \in S$. Entonces se verifica $\forall t, s \geq 0$,

$$p_{ij}(t+s) = \sum_{k \in S} p_{ik}(t)p_{kj}(s) \quad \forall i, j \in S$$

OBSERVACIÓN:

• Si $\{X_t: t \in T\}$ una cadena de Markov con probabilidades de transición estacionarias y espacio de estados S , llamaremos matrices de transición a $P(t) = (p_{ij}(t))_{i,j \in S}$, $t \geq 0$. Dichas matrices tienen las siguientes propiedades:

1. $p_{ij}(t) \geq 0 \forall i, j \in S, \forall t \geq 0$.
2. $\sum_{j \in S} p_{ij}(t) = 1 \forall i \in S, \forall t \geq 0$
3. $P(s+t) = P(s)P(t)$.

• Si la distribución de X_t viene dada por $p(t) = \{p_i(t): i \in S\}$, con $p_i(t) = P(X_t = i)$, $t \geq 0$ (a $p(0)$ se le conoce como distribución inicial), entonces se verifica:

$$p(t) = p(0)P(t) \quad \forall t \geq 0$$

Q-matriz Asociada al Proceso.

PROPOSICIÓN: Se verifica que:

$$\lim_{t \downarrow 0} p_{ij}(t) = \delta_{ij} \quad i, j \in S$$

(matricialmente $\lim_{t \downarrow 0} P(t) = I$).

PROPOSICIÓN: Los límites

1. $\lim_{t \downarrow 0} \frac{p_{ii}(t) - 1}{t} = p'_{ii}(0), \forall i \in S$
2. $\lim_{t \downarrow 0} \frac{p_{ij}(t)}{t} = p'_{ij}(0), \forall i, j \in S, i \neq j$

existen, si bien el primero de ellos puede ser $-\infty$.

OBSERVACIÓN: Denotaremos $q_{ij} = p'_{ij}(0)$ y $q_i = -q_{ii} = -p'_{ii}(0)$ (supondremos siempre que $q_i < +\infty, \forall i \in S$). Obtenemos así la matriz

$$Q = (q_{ij})_{i,j \in S} = \lim_{t \downarrow 0} \frac{P(t) - I}{t}$$

conocida como **Q-matriz** asociada a la cadena de Markov. Las cantidades q_{ij} y q_i se denominan intensidad de transición e intensidad de paso, respectivamente.

PROPOSICIÓN: Se verifica que:

$$P'(t) = P(t)Q \text{ (ec. adelantada de Kolmogorov)}$$

$$P'(t) = QP(t) \text{ (ec. atrasada de Kolmogorov)}$$

Imponiendo las condiciones iniciales $p_{ij}(0) = \delta_{ij}$, las ecuaciones diferenciales anteriores tienen como solución, respectivamente

$$p_{ij}(t) = \delta_{ij} \exp(-q_j t) + \sum_{k \neq j} \int_0^t p_{ik}(s) q_{kj} \exp(-q_j(t-s)) ds$$

$$p_{ij}(t) = \delta_{ij} \exp(-q_i t) + \sum_{k \neq i} \int_0^t \exp(-q_i s) q_{ik} p_{kj}(t-s) ds$$

PROPOSICIÓN: La Q-matriz verifica que $\sum_{j \in S} q_{ij} = 0, \forall i \in S$ (i.e. $\sum_{j \neq i} q_{ij} = q_i, \forall i \in S$)

DEFINICIÓN: Llamaremos sucesión de tiempos de salto de $\{X_t: t \in T\}$ a la sucesión $\{J_n\}_{n \geq 0}$ definida recursivamente por $J_0 = 0$,

$$J_{n+1} = \inf\{t \geq J_n: X_t \neq X_{J_n}\} \quad n = 0, 1, \dots$$

(donde $\inf \emptyset = \infty$). Llamaremos sucesión de tiempos de permanencia de $\{X_t: t \in T\}$ a la sucesión $\{S_n\}_{n \geq 1}$ definida por

$$S_n = \begin{cases} J_n - J_{n-1} & \text{si } J_{n-1} < +\infty \\ \infty & \text{si } J_{n-1} = +\infty \end{cases}$$

Finalmente definimos también el proceso o cadena de saltos, $Y_n = X_{J_n}, n = 0, 1, \dots$ (si $J_{n+1} = \infty$ para algún n definimos $X_\infty = X_{J_n}$, en otro caso X_∞ queda sin definir).

PROPOSICIÓN: Sea $i \in S$ tal que $q_i > 0$. Se verifica que S_{n+1} , condicionado a que $Y_n = i$, sigue una distribución exponencial de parámetro $1/q_i$.

OBSERVACIÓN: Cuando $q_i > 0$ diremos que el estado i es **estable**. Si $q_i = 0$, entonces $p_{ii}(t) = 1 \forall t$, y se dice que el estado i es **absorbente**.

PROPOSICIÓN: Sea $i \in S$ tal que $q_i > 0$. Se verifica que $P(Y_{n+1} = j | Y_n = i) = q_{ij}/q_i, j \neq i$.

Clasificación de los Estados.

DEFINICIÓN: Si $\{X_t\}_{t \in T}$ es una cadena de Markov con probabilidades de transición estacionarias definimos

$$f_{ii} = P(X_{J_1+s} = i \text{ para algún } s > 0 | X_0 = i)$$

DEFINICIÓN: El estado i se dice **transitorio** si $f_{ii} < 1$. Por el contrario lo llamaremos **recurrente** si $f_{ii} = 1$.

DEFINICIÓN: Si $\{X_t\}_{t \in T}$ es una cadena de Markov con probabilidades de transición estacionarias definimos el tiempo de primera pasada por el estado $i \in S$ como

$$T_i = \inf\{t \geq J_1 : X_t = i\}$$

OBSERVACIÓN: Si F_{ii} es la función de distribución asociada a T_i condicionada a que $X_0 = i$, entonces $f_{ii} = F_{ii}(+\infty)$. Así F_{ii} será la función de distribución de una distribución de probabilidad si y sólo si i es un estado recurrente.

DEFINICIÓN: Si i es un estado recurrente, llamaremos tiempo medio de recurrencia al estado i a $\mu_i = E[T_i | X_0 = i] = \int_0^\infty t dF_{ii}(t) (\leq \infty)$. Si $\mu_i < \infty$ diremos que i es un estado **recurrente positivo**. Por el contrario si $\mu_i = \infty$ diremos que i es estado **recurrente nulo**.

DEFINICIÓN: Un estado recurrente positivo y aperiódico diremos que es un estado **ergódico**.

TEOREMA: Un estado i es transitorio si y sólo si $\int_0^\infty p_{ii}(t) dt < \infty$

Clasificación de las Cadenas.

DEFINICIÓN: Diremos que un estado i **conduce** a un estado j ($i \rightsquigarrow j$) si existe $t \geq 0$ tal que $p_{ij}(t) > 0$. Si $i \rightsquigarrow j$ y $j \rightsquigarrow i$ diremos que los estados son **comunicantes** ($i \leftrightarrow j$).

OBSERVACIÓN: \leftrightarrow define una relación de equivalencia sobre el espacio de estados S .

TEOREMA: Si en una cadena de Markov con probabilidades de transición estacionarias un estado recurrente i conduce a cualquier otro estado j , entonces dichos estados son comunicantes y del mismo tipo, i.e. ambos son recurrentes positivos o recurrentes nulos.

DEFINICIÓN: Un conjunto de estados, C , se dice

1. **cerrado** si $p_{ij}(t) = 0, \forall t > 0, \forall i \in C, \forall j \notin C$.
2. **irreducible** si $i \leftrightarrow j, \forall i, j \in C$.

OBSERVACIONES:

1. Si un conjunto cerrado de estados contiene un sólo elemento, a dicho estado se le llama **absorbente**. Se verifica que i es absorbente si y sólo si $q_i = 0$.
2. Llamaremos a un conjunto de estados irreducible recurrente nulo (ó recurrente positivo ó transitorio) si todos los estados que lo constituyen son de este tipo.
3. Diremos que la cadena de Markov es irreducible si el espacio de estados, S , es irreducible. Una cadena de Markov irreducible se dice **ergódica** si todos sus estados son ergódicos, i.e. recurrentes positivos y aperiódicos.

TEOREMA: El espacio de estados S se puede dividir, de manera única en subconjuntos disjuntos

$$S = T \cup C_1 \cup C_2 \cup \dots,$$

donde T es el conjunto de estados transitorios, y los C_i son conjuntos cerrados e irreducibles de estados recurrentes.

Distribuciones Estacionarias.

DEFINICIÓN: Una cadena de Markov $\{X_t : t \geq 0\}$ tal que $p_i(t) = P(X_t = i) = P(X_0 = i) = p_i(0)$, $\forall t, \forall i \in S$, se denomina **cadena de Markov estacionaria**.

DEFINICIÓN: Una distribución de probabilidad $\pi = \{\pi_i\}_{i \in S}$ sobre S es una **distribución estacionaria** para a una cadena de Markov en tiempo continuo con matrices de transición $P(t)$, $t \geq 0$, si $\pi P(t) = \pi \forall t \geq 0$, i.e.

$$\pi_j = \sum_{i \in S} \pi_i p_{ij}(t) \quad \forall t \geq 0, \quad \forall j \in S.$$

PROPOSICIÓN: Sea Q la Q -matriz asociada a la cadena de Markov $\{X_t : t \geq 0\}$ y sea $\pi = \{\pi_i\}_{i \in S}$ una distribución de probabilidad sobre S . La distribución π es estacionaria para $\{X_t\}_{t \geq 0}$ si y sólo si $\pi Q = 0$.

PROPOSICIÓN: Sea $\{X_t : t \geq 0\}$ una cadena de Markov con probabilidades de transición estacionarias, irreducible y con matrices de transición $P(t)$, $t \geq 0$. Existe una distribución estacionaria respecto a $\{X_t : t \geq 0\}$ si y sólo si $\{X_t : t \geq 0\}$ es recurrente positiva. En tal caso $\pi_j = 1/q_j \mu_j \forall j \in S$.

OBSERVACIONES:

1. Una condición necesaria y suficiente para que una cadena tenga distribuciones estacionarias es que tenga alguna subclase recurrente positiva.
2. Si una cadena tiene varias subclases recurrentes positivas, las distribuciones estacionarias vendrán dadas por las combinaciones lineales convexas de las distribuciones estacionarias asociadas a cada una de las subclases recurrentes positivas.
3. Si el conjunto de estados de una cadena de Markov con probabilidades de transición estacionarias irreducible es finito, entonces es recurrente positiva.

Comportamiento Asintótico de una Cadena de Markov en tiempo continuo.

TEOREMA. Para una cadena de Markov con probabilidades de transición estacionarias se verifica:

1. Si i es un estado recurrente positivo

$$\lim_{t \rightarrow \infty} p_{ii}(t) = \pi_i = 1/q_i \mu_i, \quad \lim_{t \rightarrow \infty} p_{ji}(t) = f_{ji} \pi_i \quad \forall j \in S \quad (f_{ji} = P(X_{t+s} = i \text{ para algún } s > 0 | X_t = j))$$

2. Si i es un estado recurrente nulo o transitorio

$$\lim_{t \rightarrow \infty} p_{ji}(t) = 0 \quad \forall j \in S$$

TEOREMA. Para una cadena de Markov con probabilidades de transición estacionarias irreducible se verifica:

1. Si los estados son recurrentes positivos, entonces $\lim_{t \rightarrow \infty} p_j(t) = \pi_j, \forall j \in S$.
2. Si los estados son recurrentes nulos o transitorios, entonces $\lim_{t \rightarrow \infty} p_j(t) = 0, \forall j \in S$.

Procesos de Nacimiento y Muerte.

DEFINICIÓN. Llamaremos **proceso de nacimiento y muerte** a una cadena de Markov en tiempo continuo con espacio de estados $\{0, 1, 2, \dots\}$ tal que:

1. $p_{ii+1}(t) = \lambda_i t + o(t), i \geq 0.$
2. $p_{ii-1}(t) = \mu_i t + o(t), i \geq 1.$
3. $p_{ii}(t) = 1 - (\lambda_i + \mu_i)t + o(t), i \geq 0.$
4. $\mu_0 = 0, \lambda_0 > 0, \mu_i, \lambda_i > 0, i = 1, 2, \dots$

OBSERVACIÓN: Para que se verifiquen las ecuaciones adelantadas de Kolmogorov asumiremos que $\sup_{i \geq 0} (\lambda_i + \mu_i) < \infty.$ Esta condición garantizará que existe solución única de los sistemas de ecuaciones atrasadas y adelantadas de Kolmogorov.

PROPOSICIÓN. Los sistemas de ecuaciones atrasadas y adelantadas de Kolmogorov son, respectivamente:

1. $p'_{0j}(t) = -\lambda_0 p_{0j}(t) + \lambda_0 p_{1j}(t)$
 $p'_{ij}(t) = \mu_i p_{i-1j}(t) - (\lambda_i + \mu_i) p_{ij}(t) + \lambda_i p_{i+1j}(t), i \geq 1.$
2. $p'_{i0}(t) = -\lambda_0 p_{i0}(t) + \mu_1 p_{i1}(t)$
 $p'_{ij}(t) = \lambda_{j-1} p_{ij-1}(t) - (\lambda_j + \mu_j) p_{ij}(t) + \mu_{j+1} p_{ij+1}(t), j \geq 1.$

PROPOSICIÓN. Un proceso de nacimiento y muerte tiene distribución estacionaria si y sólo si

$$\sum_{n=1}^{\infty} \frac{\lambda_0 \dots \lambda_{n-1}}{\mu_1 \dots \mu_n} < \infty.$$

Cuando existe, dicha distribución estacionaria viene dada por

$$\pi_n = \frac{\lambda_0 \dots \lambda_{n-1}}{\mu_1 \dots \mu_n} \pi_0, \quad n = 1, 2, \dots,$$

siendo $\pi_0 = \left(1 + \sum_{n=1}^{\infty} \frac{\lambda_0 \dots \lambda_{n-1}}{\mu_1 \dots \mu_n}\right)^{-1}$