

1. VECTORES ALEATORIOS

Definición 1 (σ -álgebra de Boole). Sea X un conjunto. Una σ -álgebra de Boole sobre X es una colección $\mathcal{A} \subseteq \mathcal{P}(X)$ tal que:

1. $X \in \mathcal{A}$.
2. Si $A \in \mathcal{A}$, entonces $X \setminus A \in \mathcal{A}$.
3. Si $\{A_n\}_{n=1}^{\infty} \subseteq \mathcal{A}$, entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Definición 2 (Vector aleatorio). Un vector aleatorio $X = (X_1, \dots, X_n)$ de un espacio de probabilidad (Ω, \mathcal{A}, P) se define como una función medible:

$$X : (\Omega, \mathcal{A}, P) \rightarrow (\mathbb{R}^n, \mathcal{B}^n)$$

tal que se cumple que:

$$X^{-1}(B) \in \mathcal{A}, \quad \forall B \in \mathcal{B}^n.$$

Es decir:

$$X^{-1}([-\infty, x]) = \{\omega \in \Omega \mid X_1(\omega) \leq x_1, \dots, X_n(\omega) \leq x_n\} \in \mathcal{A} \quad \forall x \in \mathbb{R}^n.$$

Además, considerando cada una de las componentes por separado, como cada componente de una función medible es medible, se tiene de forma directa:

Teorema 1. Sea $X = (X_1, \dots, X_n) : (\Omega, \mathcal{A}, P) \rightarrow (\mathbb{R}^n, \mathcal{B}^n)$. Entonces:

$$X \text{ es un vector aleatorio} \iff X_i \text{ es una variable aleatoria } \forall i = 1, \dots, n.$$

Definición 3 (Distribución de probabilidad). Sea X un vector aleatorio. La distribución de probabilidad (que será la función de densidad o función masa de probabilidad en el caso unidimensional) de X es la medida de probabilidad en \mathbb{R}^n definida por:

$$\begin{aligned} P_X : \mathcal{B}^n &\longrightarrow [0, 1] \\ B &\longmapsto P_X(B) = P(X^{-1}(B)), \quad \forall B \in \mathcal{B}^n = P[X \in B] \end{aligned}$$

Intuitivamente, qué tan probable es cada resultado.

Cumple las tres propiedades de la Axiomática de Kolmogorov:

1. No negatividad: $P_X(B) = P(X^{-1}(B)) \geq 0, \quad \forall B \in \mathcal{B}^n$
2. Suceso seguro: $P_X(\mathbb{R}^n) = P(X^{-1}(\mathbb{R}^n)) = P(\Omega) = 1$.
3. σ -aditividad: Sean $B_1, B_2, \dots \in \mathcal{B}^n$ disjuntos dos a dos. Entonces:

$$P_X\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P_X(B_i)$$

Definición 4 (Función de distribución). Sea X un vector aleatorio. La función de distribución de X es la función:

$$\begin{aligned} F_X : \mathbb{R}^n &\longrightarrow [0, 1] \\ x &\longmapsto F_X(x) = P_X([-\infty, x]) \end{aligned}$$

Si $X = (X_1, \dots, X_n)$, entonces denotaremos:

$$F_X(x) = P[X \leq x] = P[X_1 \leq x_1, \dots, X_n \leq x_n].$$

Intuitivamente, nos da una vista global de la probabilidad acumulada hasta cualquier valor.

Algunas de las propiedades que la función de distribución cumple son:

1. **Monotonía:** F_X es no decreciente en cada componente:

$$x_i \leq y_i \implies F_X(\dots, x_i, \dots) \leq F_X(\dots, y_i, \dots)$$

2. **Continuidad por la derecha:** para cada i ,

$$\lim_{x_i \rightarrow x_i^+} F_X(x_1, \dots, x_i, \dots, x_n) = F_X(x_1, \dots, x_i, \dots, x_n)$$

3. **Límites:**

$$\lim_{x_i \rightarrow -\infty} F_X(\dots, x_i, \dots) = 0, \quad \lim_{x_1, \dots, x_n \rightarrow +\infty} F_X(x_1, \dots, x_n) = 1$$

4. **Probabilidades no negativas en rectángulos:** para $\varepsilon_i > 0$,

$$P\left(\prod_{i=1}^n [x_i, x_i + \varepsilon_i]\right) \geq 0$$

o, en términos de F_X , usando sumas alternadas de los vértices del rectángulo.

Propiedades del cálculo de probabilidades en intervalos:

- $P([a, b]) = P[X \leq b] - P[X < a]$
- $P((a, b]) = P[X \leq b] - P[X \leq a]$
- $P([a, b)) = P[X < b] - P[X < a]$
- $P((a, b)) = P[X < b] - P[X \leq a]$
- $P(X > a) = 1 - P[X \leq a]$

1.1. Clasificación de vectores aleatorios

	Variables discretas	Variables Continuas
Distribución de probabilidad: Probabilidad de que X tome un valor específico. (Propiedades dichas)	<u>Función masa</u> ($P(X = x) = p_i \in [0, 1]$):	<u>Función de densidad</u> ($f_X(x)$): $P[X \in B] = \int_B f_X(x) dx$
Función de distribución: probabilidad acumulada de que X tome un valor menor o igual a X . (Propiedades dichas)	$F(x) = P(X \leq x) = \sum_{x_i \leq x} p_i$ Por consecuencia: $p_i = P(X = x_i) = F(x_i) - F(x_{i-1})$	$F(x) = \int_{-\infty}^x f_X(t) dt$ Por Consecuencia: $f_X(x) = F'(x)$
Distribuciones marginales: Dado X la distribución de una de sus componentes por separado.	Se obtiene sumando la probabilidad con x_i fija: $P(X_i = x_i) = \sum_{x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n \in E_X} P[X = x]$	Se obtiene integrando respecto al resto de componentes del vector: $f_{X_i}(x_i) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_X(x_1, \dots, x_n) dx_1 \dots dx_{i-1} \cdot dx_{i+1} \dots dx_n$
Distribuciones condicionadas: Dado X la distribución de una de sus componentes condicionada a que otra de sus componentes tome un valor concreto	$P[X_1 = x_1, \dots, X_{i-1} = x_{i-1}, X_{i+1} = x_{i+1}, \dots, X_n = x_n X_i = x_i^*]$ $= \frac{P[X_1 = x_1, \dots, X_{i-1} = x_{i-1}, X_i = x_i^*, X_{i+1} = x_{i+1}, \dots, X_n = x_n]}{P[X_i = x_i^*]} = \frac{\text{función masa}}{\text{función marginal}}$	$f_{X_1 = x_1, \dots, X_{i-1} = x_{i-1}, X_{i+1} = x_{i+1}, \dots, X_n = x_n X_i = x_i^*}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) =$ $= \frac{f_X(x_1, \dots, x_{i-1}, x_i^*, x_{i+1}, \dots, x_n)}{f_{X_i}(x_i^*)} = \frac{\text{función densidad}}{\text{función marginal}}$
Moda. El valor que maximiza la distribución de probabilidad, es decir, el que tiene más probabilidad de salir. No tiene por qué ser única, cuando no lo es se llama <i>multimodal</i> .	$Mo_X = \{x \in E_X P(x) \geq P(y) \quad \forall y \in E_X\}$	$Mo_X = \{x \in R^n f_X(x) \geq f_X(y) \quad \forall y \in R^n\}$
Percentiles. Una vez ordenados los datos de menor a mayor, el valor de la variable por debajo del cual se encuentra un porcentaje dado en fracción q . La <i>mediana</i> es el percentil 50: $Me = X_{50}$	$P_X([-\infty, X_q]) \geq q$ $P_X([X_q, +\infty]) \geq 1 - q$	$P_X([-\infty, X_q]) = q$ $P_X([X_q, +\infty]) = 1 - q$

1.1 Esperanza

Definición 5 (Esperanza). Sea $X = (X_1, \dots, X_n)$ un vector aleatorio. Entonces:

- $\exists E[X] \iff \exists E[X_i] \quad \forall i \in \{1, \dots, n\}.$
- En tal caso, la esperanza de X es:

$$E[X] := (E[X_1], \dots, E[X_n]).$$

Según la naturaleza de la Variable distinguimos:

- **Esperanza Discreta:** Sea X una variable aleatoria discreta con valores en el conjunto $Re_X = \{x_1, x_2, \dots\} \subset \mathbb{R}$ y sea P la función masa de probabilidad de X . Entonces, se define la esperanza matemática, media o valor esperado de X , denotado $E[X]$ como:

$$E[X] = \sum_{x \in Re_X} x P_X(x), \quad \text{si } \sum |x| P_X(x) < \infty$$

- **Esperanza Continua:** Sea X una variable aleatoria continua y f_X su función de densidad, se define la esperanza matemática de X como:

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx, \quad \text{si } \int |x| f_X(x) dx < \infty$$

1.1.1 Propiedades de la Esperanza

1. **Linealidad de la esperanza:** Sean X e Y dos variables aleatorias definidas sobre el mismo espacio de probabilidad. Entonces:

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

Presenta los siguientes Corolarios:

- (a) **Linealidad general:** Sea X_i una variable aleatoria $\forall i = 1, \dots, n$. Si $\exists E[X_i]$, entonces:

$$E\left[\sum_{i=1}^n a_i X_i\right] = \sum_{i=1}^n a_i E[X_i]$$

- (b) g, h funciones reales tales que $\exists E[g(X)], E[h(X)]$. Entonces:

$$E[\alpha g(X) + \beta h(X)] = \alpha E[g(X)] + \beta E[h(X)], \quad \forall \alpha, \beta \in \mathbb{R}$$

- (c) **Monotonía:** Sean X_1, X_2 variables aleatorias tales que $X_1 \leq X_2$. Entonces:

$$E[X_1] \leq E[X_2]$$

Más en general: si $g(X) \leq h(X)$, entonces $E[g(X)] \leq E[h(X)]$.

2. **Constante:** La esperanza de una variable aleatoria constante es la propia constante. Es decir, dado $c \in \mathbb{R}$ y $X = c \Rightarrow E[X] = c$
3. **Acotamiento:** Sea X una variable aleatoria acotada, es decir, $\exists M \in \mathbb{R}$ tal que $|X| \leq M \Rightarrow |E[X]| \leq M$
4. **No negatividad:** Sea X una variable aleatoria. Si $X \geq 0$ y $\exists E[X] \Rightarrow E[X] \geq 0$
En particular $E[X] = 0 \iff X = 0$
5. **Simetría:** Se dice que una variable aleatoria X es simétrica respecto de un valor c si $X - c$ y $c - X$ tienen la misma distribución.
 - Discreto: $P[X = c + x] = P[X = c - x]$ para todo $x \in \mathbb{R}$.
 - Continuo: $P[X \leq c - x] = P[X \geq c + x]$ para todo $x \in \mathbb{R}$.

Si X es simétrica respecto de c y $\exists E[X] \Rightarrow E[X] = c$

6. **Homogeneidad:** Sea X una variable aleatoria y $c \in \mathbb{R} \Rightarrow E[cX] = cE[X]$
7. **Minimización del error cuadrático medio:** La esperanza es el mejor estimador constante de X en el sentido del error cuadrático medio:

$$E[X] = \arg \min_{a \in \mathbb{R}} E[(X - a)^2]$$

1.2 Momentos, Varianza y Covarianza

1.2.1 Momentos

Sea X una variable aleatoria:

- **Momento de orden k centrado en a :** ${}_a m_k = E[(X - a)^k]$
- **Momento no centrado (en el origen):** $m_k = E[X^k]$
- **Momento centrado (en la media):** $\mu_k = E[(X - E[X])^k]$

Casos particulares:

$$\begin{aligned} m_0 &= 1, & m_1 &= E[X], & m_2 &= E[X^2] \\ \mu_0 &= 1, & \mu_1 &= 0, & \mu_2 &= \text{Var}[X] \end{aligned}$$

1.2.2 Varianza

- Definición: intuitivamente, mide qué tanto se dispersan o se alejan los valores de una variable respecto a su esperanza.

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

- Propiedades:

$$\begin{aligned} \text{Var}[aX + b] &= a^2 \text{Var}[X] \\ \text{Var}[X] &= 0 \iff X \text{ es constante} \end{aligned}$$

1.2.3 Covarianza y Matriz de Covarianzas

- Definición: Para dos variables X, Y : intuitivamente “correlación entre dos variables”: ¿se mueven juntas o en direcciones opuestas?

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

- Propiedades:

- $\text{Cov}(X, X) = \text{Var}[X]$
- $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- $\text{Cov}(aX + b, cY + d) = ac \text{Cov}(X, Y)$

- Varianza de suma:

$$\text{Var} \left[\sum_{i=1}^n X_i \right] = \sum_{i=1}^n \text{Var}[X_i] + \sum_{i \neq j} \text{Cov}(X_i, X_j)$$

1.3 Función Generatriz de Momentos (FGM)

1.3.1 Definición

Para X :

$$M_X(t) = E[e^{tX}], \quad t \in]-t_0, t_0[$$

- Caso discreto: $M_X(t) = \sum e^{tx_i} P(X = x_i)$
- Caso continuo: $M_X(t) = \int e^{tx} f(x) dx$

1.3.2 Propiedades

- Si existe M_X , determina la distribución de X
- Desarrollo en serie:

$$M_X(t) = \sum_{n=0}^{\infty} \frac{t^n E[X^n]}{n!}$$

- Relación con momentos: $M_X^{(n)}(0) = E[X^n]$
- Transformación lineal: si $Y = aX + b$,

$$M_Y(t) = e^{bt} M_X(at)$$

1.4 Cambio de Variable

Dado un vector aleatorio $X = (X_1, \dots, X_n)$, podemos estudiar la distribución de otro vector aleatorio $Y = (Y_1, \dots, Y_m)$ que depende de X mediante una transformación.

Proposición 1. Sea X un vector aleatorio n -dimensional, y una función medible

$$\begin{array}{ccc} g : (\mathbb{R}^n, \mathcal{B}^n) & \longrightarrow & (\mathbb{R}^m, \mathcal{B}^m) \\ X & \longmapsto & Y \end{array}$$

¿Cómo se transforma la distribución de probabilidad?

$$P_Y(B) = P_X(g^{-1}(B)), \quad \forall B \in \mathcal{B}^m.$$

Como resultado inmediato, tenemos que la función de distribución de Y es:

Fórmula general del cambio de variable

$$F_Y(y) = P[Y \leq y] = P[g(X) \leq y] = P[X \in g^{-1}((-\infty, y])] \quad y \in \mathbb{R}^m$$

1.4.1 Discreto \rightarrow Discreto

- Si X es discreto y $Y = g(X)$, entonces Y también es discreto.
- $Y = g(X)$ tiene con valores en $g(E_X)$ cuya función masa de probabilidad puede hallarse a partir de la de X como:

$$P[Y = y] = P[X \in g^{-1}(y)] = \sum_{x \in g^{-1}(y)} P[X = x], \quad x \in E_X \quad y \in g(E_X) = E_Y$$

1.4.2 Continuo \rightarrow Discreto

- Si X es continuo y $Y = g(X)$ resulta discreto:

$$P[Y = y] = \int_{g^{-1}(y)} f_X(x) dx.$$

1.4.3 Continuo \rightarrow Continuo

Sea $X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow (\mathbb{R}^n, \mathcal{B}^n, \mathbb{P}_X)$ un vector aleatorio con función de densidad $f_X(x)$ y $g : (\mathbb{R}^n, \mathcal{B}^n) \rightarrow (\mathbb{R}^n, \mathcal{B}^n)$ una función medible tal que:

1. $g(g_1, \dots, g_n)$ admite inversa $g^{-1}(g_1^*, \dots, g_n^*)$,
2. La inversa es derivable en todos los argumentos:

$$\forall i, j = 1, \dots, n \quad \exists \quad \frac{\partial g_i^*(y_1, \dots, y_n)}{\partial y_j}$$

3. El jacobiano de la inversa es no nulo:

$$J = \left| \det \left(\frac{\partial g_i^*(y_1, \dots, y_n)}{\partial y_j} \right)_{i,j} \right| \neq 0$$

Bajo estas condiciones, el vector aleatorio $Y = g(X)$ es de tipo continuo y su función de densidad es:

$$f_Y(y) = f_X(g^{-1}(y)) |J| \quad \forall y \in \mathbb{R}^n$$

1.4.4 Distribución del máximo y del mínimo

Un cambio de variable muy frecuente es tomar el **máximo** o el **mínimo** de un vector aleatorio:

$$Z = \min\{X_1, \dots, X_n\}, \quad W = \max\{X_1, \dots, X_n\}.$$

Estos cambios se usan para calcular las **distribuciones de orden**, muy útiles en confiabilidad, colas o estadísticos como el valor extremo.

- **Distribución del mínimo**

$$P[Z \leq x] = 1 - P[Z > x] = 1 - P[X_1 > x, \dots, X_n > x].$$

- **Distribución del máximo**

$$P[W \leq x] = P[X_1 \leq x, \dots, X_n \leq x] = F_X(x, \dots, x)$$

- **Distribución conjunta de (max, min)**

$$F_{(W,Z)}(x, y) = P[W \leq x, Z \leq y]$$

- Si $x \leq y$:

$$F_{(W,Z)}(x, y) = F_X(x, \dots, x).$$

- Si $x > y$:

$$F_{(W,Z)}(x, y) = F_X(x, \dots, x) - P[y < X_1 \leq x, \dots, y < X_n \leq x].$$

2. MODELOS DE DISTRIBUCIONES

2.1 Distribuciones discretas

2.1.1 Distribución degenerada

Sea X una variable aleatoria. Decimos que es degenerada si $X = c$, es decir, toma tan solo un valor constante.

- **Función Masa:** $f(x) = P[X = x] = \begin{cases} 1 & x = c \\ 0 & x \neq c \end{cases}$
- **Función distribución:** $F_X(x) = P[X \leq x] = \begin{cases} 0 & x < c \\ 1 & c \leq x \end{cases}$
- **Esperanza:** $E[X] = c$

2.1.2 Uniforme Discreta $\mathcal{U}(x_1, \dots, x_n)$

Una variable X cuyos posibles valores x_1, \dots, x_n son igual de probables:

- **Función Masa:** $P(X = x_i) = \frac{1}{n} \quad i = 1, \dots, n$
- **Función distribución:** $F_X(x) = \begin{cases} 0 & x < x_1 \\ \frac{i-1}{n} & x_{i-1} \leq x < x_i \quad i = 2, \dots, n \\ 1 & x \geq x_n \end{cases}$
- **Esperanza:** $E[X] = \frac{1}{n} \sum_{i=1}^n x_i$

2.1.3 Binomial $\mathcal{B}(n, p)$

Modela el número de éxitos (con probabilidad p) en n repeticiones del experimento.

- **Función Masa:** $P(X = x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$
donde x es el número de éxitos buscado.
- **Esperanza:** $E[X] = np$

Un caso concreto sería Bernouilli $\mathcal{B}(1, p)$ experimento que da lugar únicamente a dos posibles resultados: éxito (con probabilidad p) y error (con probabilidad $1 - p$).

2.1.4 Poisson $\mathcal{P}(\lambda)$

Se usa para describir eventos que ocurren de manera aleatoria pero con una tasa constante o promedio $\lambda > 0$ en el tiempo.

- **Función Masa:** $P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}$
donde x es el numero de ocurrencias en el tiempo descrito.
- **Esperanza y Varianza:** $E[X] = Var(X) = \lambda$

2.2 Distribuciones Continuas

2.2.1 Uniforme Continua $\mathcal{U}(a, b)$

Modela que para valores en dicho intervalo, la probabilidad es constante:

- **Función de densidad:** $f_X(x) = \frac{1}{b-a}$
- **Función de distribución:** $F_X(x) = \int_a^x f_X(y)dy = \frac{x-a}{b-a}$
- **Momentos no centrados:**

$$m_k = E[X^k] = \frac{b^{k+1} - a^{k+1}}{(k+1)(b-a)}$$

2.2.2 Normal $\mathcal{N}(\mu, \sigma^2)$

Se dice que X se distribuye según una normal (toma muchos valores alrededor de la media) de parámetros μ y σ^2 ($E[X]$, $Var(X)$), con:

- **Función de densidad:** $f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

Como no es fácil de recordar, en ocasiones usamos la técnica de **tipificación**, para poder usar las tablas de la $\mathcal{N}(0, 1)$:

$$X \sim \mathcal{N}(\mu, \sigma^2) \iff \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$$

- **Función generatriz de momentos:** $M_X(t) = E[e^{tX}] = e^{t\mu + \frac{t^2\sigma^2}{2}}$

Aplicaciones Importantísimas

1. Aproximación de Binomial:

Sea $X \sim \mathcal{B}(n, p)$ con $n > 30$ y $p \in (0.1, 0.9)$ entonces:

$$X \sim \mathcal{B}(n, p) \Rightarrow P[X \leq x] \cong P[Y \leq x] \Leftarrow Y \sim N(np, np(1-p))$$

2. Aproximación de Poisson:

Sea $X \sim \mathcal{P}(\lambda)$ con $\lambda > 10$ entonces:

$$X \sim \mathcal{P}(\lambda) \Rightarrow P[X \leq x] \cong P[Y \leq x] \Leftarrow Y \sim N(\lambda, \lambda)$$

Para que la aproximación de la más adecuada se utiliza la **corrección por continuidad** que consiste en **sumar y restar 0.5** a cada valor discreto de la distribución:

$$P[X = k] \Rightarrow P[k - 0.5 \leq X \leq k + 0.5]$$

$$P[X \leq k] \Rightarrow P[X \leq k + 0.5]$$

$$P[X \geq k] \Rightarrow P[k - 0.5 \leq X]$$

2.2.3 Exponencial $\exp(\lambda)$

Para esta, X se utiliza principalmente para modelar el tiempo entre eventos que ocurren de manera aleatoria y a una tasa constante en el tiempo.

- **Función de densidad:** $f_X(x) = \lambda e^{-\lambda x} \quad \forall x \geq 0$.
- **Función de distribución:** $F_X(x) = 1 - e^{-\lambda x}$
- **Función generatriz de momentos:** $M_X(t) = E[e^{tX}] = (1 - \frac{t}{\lambda})^{-1}$

3. INDEPENDENCIA DE V. ALEATORIAS

Definición 6 (Independencia de variables aleatorias). Sean X_1, \dots, X_n variables aleatorias definidas sobre el mismo espacio de probabilidad con funciones de distribución F_{X_1}, \dots, F_{X_n} . Consideramos el vector aleatorio $X = (X_1, \dots, X_n)$ con función de distribución conjunta F_X . Se dice que dichas variables son independientes si:

$$F_X(x_1, \dots, x_n) = F_{X_1}(x_1) \cdot \dots \cdot F_{X_n}(x_n), \quad \forall x_1, \dots, x_n \in \mathbb{R}$$

3.1 Caracterizaciones de independencia

Por lo general, la independencia equivale a que la función conjunta (función masa conjunta o función densidad conjunta) pueda factorizarse como producto de funciones de una sola variable (las marginales). Vemos también que pueden ser funciones arbitrarias.

3.1.1 Para variables discretas

- **Caracterización mediante funciones de masa de probabilidad** Consideramos el vector aleatorio $X = (X_1, \dots, X_n)$ discreto. Entonces, X_1, \dots, X_n son independientes si y solo si:

$$P[X_1 = x_1, \dots, X_n = x_n] = P[X_1 = x_1] \cdot \dots \cdot P[X_n = x_n]. \quad \forall x_1, \dots, x_n \in \mathbb{R}$$

- **Caracterización mediante funciones de masa de probabilidad** X_1, \dots, X_n son independientes si y solo si:

$$P[X_1 = x_1, \dots, X_n = x_n] = h_1(x_1) \cdot \dots \cdot h_n(x_n), \quad \forall x_1, \dots, x_n \in \mathbb{R}$$

siendo $h_i : \mathbb{R} \rightarrow \mathbb{R}$ funciones arbitrarias para $i = 1, \dots, n$.

3.1.2 Para variables continuas

- **Caracterización mediante funciones de densidad de probabilidad** Consideramos el vector aleatorio $X = (X_1, \dots, X_n)$ discreto. Entonces, X_1, \dots, X_n son independientes si y solo si:

$$\begin{cases} X = (X_1, \dots, X_n) \text{ es un vector aleatorio continuo} \\ \wedge \\ f_X(x_1, \dots, x_n) = f_{X_1}(x_1) \cdot \dots \cdot f_{X_n}(x_n), \quad \forall x_1, \dots, x_n \in \mathbb{R} \end{cases}$$

Esto nos garantiza que las componentes de un vector aleatorio continuo son variables aleatorias continuas. El recíproco no obstante no es cierto en general, pero hemos visto que sí lo es si las componentes son independientes.

- **Caracterización mediante factorización de la función densidad de probabilidad conjunta** X_1, \dots, X_n son independientes si y solo si:

$$\begin{cases} X = (X_1, \dots, X_n) \text{ es un vector aleatorio continuo} \\ \wedge \\ f_X(x_1, \dots, x_n) = h_1(x_1) \cdot \dots \cdot h_n(x_n), \quad \forall x_1, \dots, x_n \in \mathbb{R} \end{cases}$$

siendo $h_i : \mathbb{R} \rightarrow \mathbb{R}$ funciones arbitrarias para $i = 1, \dots, n$.

3.2 Propiedades de la independencia

1. Sea $X = c$ una variable aleatoria degenerada ($P[X = c] = 1$ $P[X = x \neq c] = 0$) $\Rightarrow X$ es independiente de cualquier otra variable aleatoria Y .
2. Las variables de cualquier subconjunto de variables independientes son independientes (X_1, \dots, X_n son independientes $\Rightarrow X_{i_1}, \dots, X_{i_k}$ son independientes para cualquier subconjunto $\{i_1, \dots, i_k\} \subset \{1, \dots, n\}$.
 - Corolario: Sean X_1, \dots, X_n variables aleatorias sobre el mismo espacio de probabilidad, son independientes \iff Distribución condicionada de X_1, \dots, X_k a X_{k+1}, \dots, X_n es igual a distribución marginal de X_1, \dots, X_k
3. Sean X_1, \dots, X_n variables aleatorias sobre el mismo espacio de probabilidad, y consideramos funciones medibles $g_i : \mathbb{R} \rightarrow \mathbb{R}$ para $i = 1, \dots, n$. Entonces:

X_1, \dots, X_n son independientes $\implies g_1(X_1), \dots, g_n(X_n)$ son independientes

3.2.1 Teorema de la multiplicación de las esperanzas

Teorema 2 (Teorema de la Multiplicación de las Esperanzas). Sean X_1, \dots, X_n variables aleatorias sobre el mismo espacio de probabilidad, donde suponemos que $\exists E[X_i]$ para todo $i = 1, \dots, n$. Entonces:

$$X_1, \dots, X_n \text{ son independientes} \implies \begin{cases} \exists E[X_1 \cdot \dots \cdot X_n] \\ \wedge \\ E[X_1 \cdot \dots \cdot X_n] = E[X_1] \cdot \dots \cdot E[X_n] \end{cases}$$

Corolario 1. Sean X, Y dos v.a. en el mismo espacio de probabilidad independientes, si $\exists E[X^2], E[Y^2]$, $\Rightarrow \text{Cov}[X, Y] = 0$

Corolario 2. Sean X_1, \dots, X_n v.a. independientes, si $\exists E[X_i^2]$ para todo $i = 1, \dots, n$, se tiene:

$$\exists \text{Var} \left[\sum_{i=1}^n X_i \right] = \sum_{i=1}^n \text{Var}[X_i]$$

Corolario 3 (Caracterización por Funciones Generatrices de Momentos). Sean X_1, X_2, \dots, X_n v.a.i. tales que existe su función generatriz de momentos. Entonces X_1, X_2, \dots, X_n son independientes \iff

$$\iff \begin{cases} \exists M_X(t_1, \dots, t_n) \quad \forall (t_1, \dots, t_n) \in I_1 \times \dots \times I_n \\ M_X(t_1, \dots, t_n) = \prod_{i=1}^n M_{X_i}(t_i) \quad \forall t_i \in I_i \end{cases}$$

3.3 Distribuciones Reproductivas

- Una **distribución reproductiva** es aquella en la que la suma de variables aleatorias independientes del mismo tipo (aunque con parámetros distintos) sigue perteneciendo a esa misma familia de distribuciones.
- Para comprobarlo, usamos la **función generatriz de momentos** (FGM). Si X_1, \dots, X_n son independientes, entonces:

$$M_{\sum_{i=1}^n X_i}(t) = \prod_{i=1}^n M_{X_i}(t).$$

Si este resultado coincide con la FGM de la misma familia, decimos que la distribución es reproductiva.

3.3.1 Distribuciones Discretas

- **Binomial:** Si $X_i \sim B(k_i, p) \Rightarrow \sum_{i=1}^n X_i \sim B(\sum_{i=1}^n k_i, p)$
- **Poisson:** Si $X_i \sim \mathcal{P}(\lambda_i) \Rightarrow \sum_{i=1}^n X_i \sim \mathcal{P}(\sum_{i=1}^n \lambda_i)$
- **Binomial Negativa:** Si $X_i \sim BN(k_i, p) \Rightarrow \sum_{i=1}^n X_i \sim BN(\sum_{i=1}^n k_i, p)$.

Distribuciones Continuas

- **Normal:** Si $X_i \sim N(\mu_i, \sigma_i^2) \Rightarrow \sum_{i=1}^n X_i \sim N(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$.
- **Erlang:** Si $X_i \sim \text{Erlang}(k_i, \lambda) \Rightarrow \sum_{i=1}^n X_i \sim \text{Erlang}(\sum_{i=1}^n k_i, \lambda)$

Las distribuciones reproductivas “se heredan” bajo sumas: la suma de independientes sigue siendo de la misma familia. Cuando no lo son (geométrica, exponencial), suelen estar ligadas a otra distribución cercana (BN y Erlang).