PROBABILIDAD

Basado en las clases y apuntes de María del Carmen Morcillo Aixelá y Julia García Galisteo

Autor:

Jorge Rodríguez Domínguez

Índice general

Capítulo 1

Introducción

Definición 1.0.1. Un conjunto $\mathcal{A}$ de subconjuntos de Ω es un σ-álgebra si verifica

(a) $\Omega \in \mathcal{A}$.

(b) Si $A \in \mathcal{A}$ entonces $A^c \in \mathcal{A}$.

(c) Si $\{A_n\}$, $A_n \in \mathcal{A}$ entonces $\cup_n A_n \in \mathcal{A}$.

Consecuencias inmediatas

- $\emptyset \in \mathcal{A}$.
- Si $A_n \in A$ entonces $\cap_n A_n \in \mathcal{A}$.
- Si $A_1, ..., A_n \in \mathcal{A}$ entonces $\cup_{i=1}^n A_i \in \mathcal{A}$ y $\cap_{i=1}^n A_i \in \mathcal{A}$.

Definición 1.0.2. Sea $\mathcal{C}$ familia de todos los intervalos. Entonces $\mathcal{C}$ define la σ-álgebra de Borel. Se denota $\mathbb{B}_1$ en $\mathbb{R}$.

Definición 1.0.3. Sea $(\Omega, \mathcal{A})$ un espacio medible. Una probabilidad es una medida $P : \mathcal{A} \longrightarrow [0, +\infty)$ tal que

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

(2) $\{A_n\}$, $A_n \in \mathcal{A}$ con $A_i \cap A_j = \emptyset$ si $i \neq j$, entonces

$$P(\cup_n A_n) = \sum_n P(A_n).$$

Consecuencias

- $P(A) = 1 - P(A^c)$, $A \in \mathcal{A}$.
- $P(\emptyset) = 0$.
- $A, B \in \mathcal{A}$ entonces $P(A \backslash B) = P(A) - P(A \cap B)$.
- $A, B \in \mathcal{A}$ y $A \subseteq B$ entonces $P(A) \leq P(B)$.
- $A, B \in \mathcal{A}$ entonces $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. En general

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k)$$
$$- ... + (-1)^{n-1} P(A_1 \cap ... \cap A_n).$$

Definición 1.0.4. Sea $\{A_n\}$, $A_n \in \mathcal{A}$. Definimos

- Límite superior de la sucesión $\{A_n\}$ como

$$\limsup_n A_n = \bigcap_{m=1}^{+\infty} \left(\bigcup_{n=m}^{+\infty} A_n \right) = A^*.$$

- Límite inferior de la sucesión $\{A_n\}$ como

$$\liminf_n A_n = \bigcup_{m=1}^{+\infty} \left(\bigcap_{n=m}^{+\infty} A_n \right) = A_*.$$

Consecuencias

- $\liminf_n A_n \subseteq \limsup_n A_n$.
- Una sucesión $\{A_n\}$, $A_n \in \mathcal{A}$ es convergente si $A^* = A_*$.

Teorema 1.0.5. *Sea $(\Omega, \mathcal{A})$ un espacio medible. La aplicación $P : \mathcal{A} \longrightarrow [0, +\infty)$ verificando*

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

(2) $A_1, ..., A_n \in \mathcal{A}$ *con* $A_i \cap A_j = \emptyset$ *si* $i \neq j$, *entonces*

$$P(\cup_{i=1}^n A_i) = \sum_{i=1}^n P(A_i).$$

(3) Si $\{A_n\} \downarrow \emptyset$, $A_n \in \mathcal{A}$, $P(\lim_n A_n) = \lim_n P(A_n) = 0$

Entonces P es una medida de probabilidad.

Definición 1.0.6. Sea $(\Omega, \mathcal{A})$ espacio de probabilidad. Sean $A, B \in \mathcal{A}$. A es independiente de B si $P(A \mid B) = P(A)$ $(P(B) > 0)$.

Recuérdese que

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad P(B) > 0.$$

Observación 1.0.7. La independencia es recírproca, es decir, A es independiente de B si y solo si B es independiente de A.

Teorema 1.0.8 (Teorema de la Probabilidad Compuesta)**.**

$$P(A_1 \cap ... \cap A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2)...P(A_n \mid A_1 \cap ... \cap A_{n-1}).$$

Teorema 1.0.9 (Teorema de la Probabilidad Total)**.** *Sea $(\Omega, \mathcal{A}, P)$ espacio de probabilidad. Sea $\{A_n\}$ partición de Ω, $P(A_n) > 0$ conocidas y sea $B \in \mathcal{A}$. Entonces*

$$P(B) = \sum_{n \in \mathbb{N}} P(B \mid A_n) P(A_n).$$

- *$P(A_n)$ se conoce como probabilidad a priori y $\sum_{n \in \mathbb{N}} P(A_n) = 1$.*
- *$P(B \mid A_n)$ se conocen como verosimilitudes.*

Teorema 1.0.10 (Teorema de Bayes)**.** *Bajos las mismas condiciones que el teorema de la probabilidad total:*

$$P(A_i \mid B) = \frac{P(B \mid A_i)P(A_i)}{\sum_{n\in\mathbb{N}} P(B \mid A_n)P(A_n)} \text{ (\textit{probabilidad a posteriori}).}$$

Y además, $\sum_{n\in\mathbb{N}} P(A_n \mid B) = 1$.

Observación 1.0.11.

- A es independiente de B si y solo si A es independiente de B^c si y solo si A^c es independiente de B^c.
- $\{A_1, ..., A_n\}$ es una familia de sucesos independientes si cualquier $\{i_1, ..., i_r\}$ verifica

$$P(A_{i_1} \cap ... \cap A_{i_r}) = \prod_{j=1}^{r} P(A_{i_j}).$$

Capítulo 2

Variables aleatorias

2.1. Variable aleatoria. Función indicadora. Variables aleatorias simples

Definición 2.1.1. Sea $(\Omega, \mathcal{A})$ espacio probabilizable. Una aplicación $X : \Omega \longrightarrow \mathbb{R}$ diremos que es variable aleatoria si $X^{-1}(B) \in \mathcal{A}$ para todo $B \in \mathbb{B}_1$.

Definición 2.1.2. Sea $(\Omega, \mathcal{A})$ espacio probabilizable. Una aplicación $X : \Omega \longrightarrow \mathbb{R}$ diremos que es variable aleatoria si $X^{-1}((-\infty, a]) \in \mathcal{A}$ para todo $a \in \mathbb{R}$.

Observación 2.1.3. Estas dos definiciones son equivalentes.

Definición 2.1.4. Sea $(\Omega, \mathcal{A})$ espacio probabilizable y sea $A \in \mathcal{A}$. La función $I_A : \Omega \longrightarrow \mathbb{R}$ definida por

$$I_A(\omega) = \begin{cases} 1 & si \quad \omega \in \mathcal{A} \\ 0 & si \quad \omega \notin \mathcal{A} \end{cases}$$

es una variable aleatoria y se denomina función indicadora o variable indicadora.

Definición 2.1.5. Sea $(\Omega, \mathcal{A})$ espacio probabilizable y sea $\{A_i\}_{i=1}^n$ una partición de Ω, $A_i \in \mathcal{A}$. Sean $x_1, ..., x_n \in \mathbb{R}$. La aplicación $X : \Omega \longrightarrow \mathbb{R}$ definida por

$$X(\omega) = \begin{cases} x_1 & si \quad \omega \in A_1 \\ x_2 & si \quad \omega \in A_2 \\ \vdots & \vdots \quad \vdots \\ x_n & si \quad \omega \in A_n \end{cases}$$

es una variable aleatoria y se denomina variable aleatoria simple. Nótese que $X(\omega) = \sum_{i=1}^n x_i I_{A_i}(\omega)$.

2.2. Propiedades y operaciones algebraicas con variables aleatorias

Teorema 2.2.1. *Sean X e Y variables aleatorias definidas sobre $(\Omega, \mathcal{A})$. Entonces*

(i) $A = \{\omega \in \Omega : X(\omega) < Y(\omega)\} \in \mathcal{A}$.

(ii) $B = \{\omega \in \Omega : X(\omega) \leq Y(\omega)\} \in \mathcal{A}$.

(iii) $C = \{\omega \in \Omega : X(\omega) = Y(\omega)\} \in \mathcal{A}$.

Demostración. Probemos (i). $A = \{\omega \in \Omega : X(\omega) < Y(\omega)\}$. Sea $\omega_0 \in A$, entonces $X(\omega_0) < Y(\omega_0)$. Ambos son números reales, por tanto, existe $r_0 \in \mathbb{Q}$ tal que $X(\omega_0) < r_0 < Y(\omega_0)$. Por tanto

$$\begin{aligned} A = \{\omega \in \Omega : X(\omega) < Y(\omega)\} &= \bigcup_{r\in\mathbb{Q}} \{\omega \in \Omega : X(\omega) < r < Y(\omega)\} \\ &= \bigcup_{r\in\mathbb{Q}} (\{\omega \in \Omega : X(\omega) < r\} \cap \{\omega \in \Omega : r < Y(\omega)\}) \\ &= \bigcup_{r\in\mathbb{Q}} \left(X^{-1}((-\infty, r)) \cap Y^{-1}((r, +\infty))\right) \end{aligned}$$

Como X e Y son variables aleatorias se tiene que $X^{-1}((-\infty, r)) \in \mathcal{A}$ e $Y^{-1}((r, +\infty)) \in \mathcal{A}$. Como A es unión numerable de elementos de $\mathcal{A}$ se tiene que $A \in \mathcal{A}$.

Para probar (ii) basta ver que

$$B = \{\omega \in \Omega : X(\omega) \leq Y(\omega)\} = (A^*)^c$$

donde $A^* = \{\omega \in \Omega : X(\omega) > Y(\omega)\}$. Como $A^* \in \mathcal{A}$ entonces $(A^*)^c = B \in \mathcal{A}$.

Para probar (iii) basta ver que

$$C = \{\omega \in \Omega : X(\omega) = Y(\omega)\} = B\backslash A = B \cap A^c \in \mathcal{A}.$$

□

Proposición 2.2.2. *Sean X e Y variables aleatorias definidas sobre $(\Omega, \mathcal{A})$. Entonces*

(1) Dado $a \in \mathbb{R}$ se tiene que $Z = aX$ es variable aleatoria.

(2) $X + Y$ es variable aleatoria.

(3) X^2 es variable aleatoria.

(4) $|X|$ es variable aleatoria.

(5) $X \cdot Y$ es variable aleatoria.

Teorema 2.2.3. *Sea X variable aleatoria definida sobre $(\Omega, \mathcal{A})$. Sea $g : \mathbb{R} \longrightarrow \mathbb{R}$ una función tal que $g^{-1}(B) \in \mathbb{B}_1$ para todo $B \in \mathbb{B}_1$. Entonces $Y = g(X)$ es también una variable aleatoria definida sobre $(\Omega, \mathcal{A})$.*

Demostración.

$$\Omega \xrightarrow{X} \mathbb{R} \xrightarrow{g} \mathbb{R}$$

Seea $B \in \mathbb{B}_1$. Entonces

$$Y^{-1}(B) = X^{-1}(g^{-1}(B)),$$

$g^{-1}(B) \in \mathbb{B}_1$ por hipótesis y como X es variable aleatoria, entonces $X^{-1}(g^{-1}(B)) = Y^{-1}(B) \in \mathbb{B}_1$. □

2.3. Probabilidad inducida en $\mathbb{R}$ por una variable aleatoria

Proposición 2.3.1. *Sea $(\Omega, \mathcal{A}), P$ un espacio de probabilidad y sea X una variable aleatoria definida en Ω, esto es, $X : \Omega \longrightarrow \mathbb{R}$. Entonces X induce sobre $(\mathbb{R}, \mathbb{B}_1)$ una medida de probabilidad $P_X : \mathbb{B}_1 \longrightarrow [0, +\infty)$ dada por*

$$P_X(B) = P[\{\omega \in \Omega : X(\omega) \in B\}] = P(X^{-1}(B))$$

para todo $B \in \mathbb{B}_1$.

Capítulo 3

Función de distribución de probabilidad

3.1. Función de distribución asociada a una medida de probabilidad

Definición 3.1.1. Una función $F : \mathbb{R} \longrightarrow [0,1]$ decimos que es una función de distribución si verifica

(i) F es creciente.

(iii) F es continua por la derecha, es decir,

$$\lim_{x \to a^*} F(x) = F(a) \text{ } para \text{ } todo \text{ } a \in \mathbb{R}$$

(iii)

$$\lim_{x \to -\infty} F(x) = 0 \quad y \quad \lim_{x \to +\infty} F(x) = 1.$$

Proposición 3.1.2. *Sea P una medida de probabilidad definida en $(\mathbb{R}, \mathbb{B}_1)$. La función $F : \mathbb{R} \longrightarrow [0,1]$ dada por $F(x) = P((-\infty, x])$ para todo $x \in \mathbb{R}$, es una distribución de probabilidad.*

Demostración. (i) F es creciente. Sean $x_1, x_2 \in \mathbb{R}$ tales que $x_1 < x_2$. Entonces

$$(-\infty, x_1] \subset (-\infty, x_2] \Longrightarrow P((-\infty, x_1]) \leq P((-\infty, x_2]) \Longrightarrow F(x1) \leq F(x2).$$

(ii) F es continua por la derecha.

$$\begin{aligned}\lim_{n \to \infty} F\left(a + \frac{1}{n}\right) &= \lim_{n \to \infty} P\left(\left(-\infty, a + \frac{1}{n}\right]\right) = P\left(\lim_{n \to \infty}\left\{\left(-\infty, a + \frac{1}{n}\right]\right\}\right) \\ &= P((-\infty, a)) = F(a)\end{aligned}$$

para todo $a \in \mathbb{R}$.

Probemos (iii).

$$\lim_{n \to \infty} F(n) = \lim_{n \to \infty} P((-\infty, n]) = P\left(\lim_{n \to \infty} \{(-\infty, n)\}\right) = P(\mathbb{R}) = 1$$

$$\lim_{n \to -\infty} F(n) = \lim_{n \to -\infty} P((-\infty, n]) = P\left(\lim_{n \to -\infty} \{(-\infty, -n)\}\right) = P(\emptyset) = 0.$$

$\square$

Observación 3.1.3. Sean $P : \mathbb{B}_1 \Longrightarrow [0,1]$ una medida de probabilidad y $F(x) = P((-\infty, x])$ función de distribución asociada a P. Sean $a, b \in \mathbb{R}$, $a < b$. Entonces

- $P((a,b]) = F(b) - F(a)$.

$$\begin{aligned}(-\infty, b] = (-\infty, a] \cup (a, b] &\Longrightarrow P((-\infty, b]) = P((-\infty, a]) + P((a, b]) \\ &\Longleftrightarrow P((a,b]) = F(b) - F(a).\end{aligned}$$

- $P(\{a\}) = F(a) - F(a^-)$.
- $P((a, +\infty)) = 1 = P((-\infty,]) = 1 - F(a)$.
- $P((a,b)) = F(b) - F(a) - P(\{b\})$.
- $P([a,b)) = F(b) - F(a) - P(\{b\}) + P(\{a\})$.
- $P([a,b]) = F(b) - F(a) + P(\{a\})$.

3.2. Aleatorización de la recta real

Sea X una variable aleatoria definida sobre $(\Omega, \mathcal{A}, P)$ y sea P_X la probabilidad inducida por X en $(\mathbb{R}, \mathbb{B}_1)$. Entonces $F : \mathbb{R} \longrightarrow [0,1]$ dada por $F(x) = P_X((-\infty, x])$ para todo $x \in \mathbb{R}$, es una función de distribución de probabilidad.

Definición 3.2.1. Sea X variable aleatoria definida sobre $(\Omega, \mathcal{A}, P)$, sea P_X la probabilidad inducida por X en $(\mathbb{R}, \mathbb{B}_1)$ y sea F_X la función de distribución de probabilidad asociada a P_X. Definimos

- Conjunto de puntos de salto de F_X como

$$D_X = \{X \in \mathbb{R} : P_X(\{x\}) > 0\}.$$

- Soporte de X (o soporte de la distribución) como

$$S_X = \{x \in \mathbb{R} : F(x+\varepsilon) - F(x-\varepsilon) > 0\}$$

  para todo $\varepsilon > 0$.

Teorema 3.2.2. *El conjunto D_X es a lo sumo infinito numerable.*

Demostración. Nótese que

$$D_X = \{X \in \mathbb{R} : P_X(\{x\}) > 0\} = \bigcup_{r \in \mathbb{N}} \left\{X \in \mathbb{R} : P_X(\{x\}) > \frac{1}{r}\right\}$$

Sea $D_r = \left\{X \in \mathbb{R} : P_X(\{x\}) > \frac{1}{r}\right\}$. Veamos que D_r es finito.

$$\begin{aligned}&D_1 = \{X \in \mathbb{R} : P_X(\{x\}) > 1\} \text{ tiene a lo sumo 1 elemento} \\ &D_2 = \left\{X \in \mathbb{R} : P_X(\{x\}) > \frac{1}{2}\right\} \text{ tiene a lo sumo 2 elementos} \\ &\quad\vdots \\ &D_r = \left\{X \in \mathbb{R} : P_X(\{x\}) > \frac{1}{r}\right\} \text{ tiene a lo sumo r elementos}\end{aligned}$$

Por tanto, D_X es a lo sumo infinito numerable. □

Teorema 3.2.3 (Teorema de Descomposición). *Dada cualquier función de distribución $F : \mathbb{R} \longrightarrow [0,1]$ siempre existe una única descomposición*

$$F = F_d + F_c$$

en la que $F_d, F_c : \mathbb{R} \longrightarrow [0,1]$ verifican

(1) $F_d(-\infty) = F_c(-\infty) = 0$.

(2) F_d *es continua por la derecha para todo $x \in \mathbb{R}$ y varía a saltos.*

(3) F_c *es continua para todo $x \in \mathbb{R}$.*

(4) F_d, F_c *son crecientes.*

(5) Las normalizaciones

$$F_1 = \frac{F_d}{F_d(+\infty)} \quad y \quad F_2 = \frac{F_c}{F_c(+\infty)}$$

son funciones de distribución tales que

$$F(x) = \alpha F_1(x) + (1-\alpha)F_2(x)$$

donde $\alpha \in (0,1)$ y $\alpha = F_d(+\infty)$ (dicho α se conoce como mixtura de la distribución).

Capítulo 4

Clasificación de variables aleatorias

4.1. Variables aleatorias discretas

Definición 4.1.1. Una variable aleatoria X tiene una distribución discreta si P_X es discreta, es decir, P_X está concentrada en un conjunto $D_X = \{x_i\}$, $x_i \in \mathbb{R}$ a lo sumo infinito numerable tal que

- $P_X(\{x_i\}) > 0$ para todo $x_i \in D_X$.
- $\sum_{x_i \in D_x} P_X(\{x_i\}) = 1$.

En tal caso

$$P_X(B) = \sum_{x_i \in D_X \cap B} P_X(\{x_i\})$$

para todo $B \in \mathbb{B}_1$ y

$$F_X(x) = \sum_{\{x_i \in D_x : x_i \leq x\}} P_X(\{x_i\}) = P_X((-\infty, x]).$$

Observación 4.1.2. Si X es una variable aleatoria discreta

(1) F_X es escalonada donde cada salto coincide con $P_X(\{x_i\})$.

(2) $P_X(D_X) = 1 = \sum_{x_i \in D_x} P_X(\{x_i\})$.

(3) $S_X = D_X$.

(4) $F_X = F_d$.

(5) Podemos establecer una partición de Ω.

$$X : \Omega \longrightarrow \mathbb{R}, \quad P_X : \mathbb{B}_1 \longrightarrow [0, 1].$$

Definimos

$$A_i = \{\omega \in \Omega : X(\omega) = x_i\}.$$

Es claro que si $x_i \neq x_j$ entonces $A_i \cap A_j = \emptyset$.

4.2. Variables aleatorias absolutamente continuas

Definición 4.2.1. Una variable aleatoria X decimos que tiene una distibución absolutamente continua si existe una función $f : \mathbb{R} \longrightarrow \mathbb{R}$ no negativa tal que

$$F_X(b) - F_X(a) = P_X((a, b]) = \int_a^b f(x)\ dx,$$

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

Proposición 4.2.2. *Una función $f : \mathbb{R} \longrightarrow \mathbb{R}$ no negativa e integrable-Riemann en cualquier intervalo $[a, b] \subset \mathbb{R}$ es función de densidad de alguna función de distribución si y solo si*

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

Observación 4.2.3. Si X es una variable aleatoria absolutamente continua

(1) $F_X(x) = \int_{-\infty}^{x} f(t)\ dt = P_X((-\infty, x])$.

(2) $F'(x) = f(x)$.

(3) $C_X = \{x \in \mathbb{R} : f(x) > 0\}$ y $D_X = \{x \in \mathbb{R} : P_X(\{x\}) > 0\} = \emptyset$.

(4) $P_X(B) = \int_B f(x)\ dx$ para todo $B \in \mathbb{B}_1$.

(5) F_X es continua.

(6) $P_X(\{a\}) = 0$ para todo $a \in \mathbb{R}$.

4.3. Distribuciones mixtas

Definición 4.3.1. Una variable aleatoria X decimos que tiene una dsitribución mixta si la función de distribución, F_X, posee a lo sumo un número infinito numerable de puntos de salto y crece de forma continua en al menos un intervalo $I \subset \mathbb{R}$.

Observación 4.3.2. Si X es una variable aleatoria mixta

(1) $D_X = \{x_n\}$ y $P_X(D_X) < 1$.

(2) $C_X^* = \{x \in \mathbb{R} : f^*(x) > 0\}$ y $P_X(C_X^*) < 1$.

(3) f^* es una distribución de pseudodensidad.

(4) $P_X(D_X) + P_X(C_X^*) = 1$.

(5) $P_X(B) = \sum_{x_n \in D_X \cap B} P_X(X_n) + \int_B f^*(x)\ dx$ para todo $B \in \mathbb{B}_1$.

Capítulo 5

Cambio de variable

5.1. Cambio de variable (variable discreta)

Proposición 5.1.1. *Sea X una variable aleatoria discreta y sea $D_X = \{x_n\}$ el conjunto de puntos de salto con $\{P_X(\{x_n\})$ función de masa. Sea $Y = g(X)$ con g función medible concentrada en $g(D_X)$. Entonces Y es una variable aleatoria y*

$$P(Y = y) = \sum_{x_n \in D_x \cap g^{-1}(y)} P_X(\{x_n\}).$$

5.2. Cambio de variable (variable absolutamente continua)

5.2.1. Variable aleatoria absolutamente continua a variable aleatoria discreta

Proposición 5.2.1. *Sea X una variable aleatoria absolutamente continua con función de densidad f_X. Sea $g : I \longrightarrow \mathbb{R}$ una función medible que toma a lo sumo un número infinito numerable de valores en $\mathbb{R}$, es decir, $g(I) = \cup_{i \in I} y_i$. Entonces $Y = g(X)$ es una variable aleatoria discreta y*

$$P(Y = y_i) = P(g(X) = y_i) = P(X \in g^{-1}(y_i)).$$

5.2.2. Variable aleatoria absolutamente continua a variable aleatoria absolutamente continua

Teorema 5.2.2 (Teorema de cambio de variable)**.** *Sea X una variable aleatoria absolutamente continua con función de densidad f_X concentrada en un intervalo $I \subseteq \mathbb{R}$. Sea $g : I \longrightarrow \mathbb{R}$ una función medible, derivable, con derivada continua y estrictamente monótona. Entonces $Y = g(x)$ es variable aleatoria absolutamente continua con función de densidad*

$$f_Y(y) = f_X(g^{-1}(y)) \cdot |(g^{-1}(y))'|.$$

Demostración. Supongamos que g es estrictamente creciente. Consideremos $Y = g(X)$. Y es una variable aleatoria absolutamente continua si existe $h : \mathbb{R} \longrightarrow \mathbb{R}$ no negativa, integrable-Riemann en [a,b], $a, b \in \mathbb{R}$ y $\int_{-\infty}^{+\infty} h(y)\ dy = 1$.

Si h es la función de densidad de Y, entonces debe verificar

$$P_Y((a, b]) = F_Y(b) - F_Y(a) = \int_a^b h(x)\ dy.$$

Construyamos h.

$$P_Y((a,b]) = P_Y(a < Y \leq b) = P(a < g(X) \leq b).$$

Como g es estrictamente creciente

$$\begin{aligned} P(a < g(X) \leq b) = P(g^{-1}(a) < X \leq g^{-1}(b)) &= \int_{g^{-1}(a)}^{g^{-1}(b)} f_X(x)\ dx \\ &= \int_{g^{-1}(a)}^{g^{-1}(b)} f_X(g^{-1}(y)) \cdot |(g^{-1}(y))'|\ dy \\ &= \int_{g^{-1}(a)}^{g^{-1}(b)} f_X(g^{-1}(y)) \cdot (g^{-1}(y))'\ dy \end{aligned}$$

Definimos $h = f_X(g^{-1}(y)) \cdot (g^{-1}(y))'$. Es claro que $h \geq 0$ y que $\int_{-\infty}^{+\infty} h(y)\ dy = 1$. Por tanto h es función de densidad de probabilidad.

El caso de g estrictamente decreciente se hace de forma análoga. □

Teorema 5.2.3 (Generalización del teorema de cambio de variable)**.** *Sea X una variable aleatoria absolutamente continua con función de densidad f_X y sea g una función medible con dominio $I = \cup_{i=1}^{n} D_i$, de forma que $g_i = g|_{D_i}$ es estrictamente monótona en cada D_i, diferenciable y con derivada no nula. Entonces $Y = g(X)$ es variable aleatoria absolutamente continua con función de densidad*

$$f_Y(y) = \sum_{y \in g_i(D_i)} f_X(g_i^{-1}(y)) \cdot |(g_i^{-1}(y))'|.$$

5.2.3. Variable aleatoria absolutamente continua a variable aleatoria mixta

Proposición 5.2.4. *Sea X una variable aleatoria absolutamente continua con función de densidad f_X y sea g una función medible constante a lo sumo en un número infinito numerable de intervalos de $\mathbb{R}$ y continua en al menos un intervalo de $\mathbb{R}$. Entonces $Y = g(X)$ es una variable aleatoria con distribución mixta.*

5.3. Variables aleatorias truncadas

Definición 5.3.1. Sea X una variable aleatoria y sea P_X la distribución de pobabilidad inducida por X en $\mathbb{R}$. Entonces $Y = X|X \in T$ donde $T \in \mathbb{B}_1$ es una variable aleatoria truncada.

Observación 5.3.2. Si X es una variable aleatoria y P_X la distribución de pobabilidad inducida por X en $\mathbb{R}$. Consideramos $Y = X|X \in T$ donde $T \in \mathbb{B}_1$. Entonces

$$P_Y(B) = P(Y \in B) = P(X \in B \mid X \in T) = \frac{P_X(B \cap T)}{P_X(T)}.$$

Proposición 5.3.3. *Sea X una variable aleatoria con función de distribución F_X. Entonces*

(i) La variable aleatoria $Y = X|X \geq x_0$, $x_0 \in \mathbb{R}$, $T = [x_0, +\infty)$ tiene como función de distribución

$$F_1(y) = \begin{cases} 0 & si \quad y < x_0 \\ \frac{F_X(y) - F_x(x_0^-)}{1 - F_X(x_0^-)} & si \quad y \geq x_0 \end{cases}$$

(ii) La variable aleatoria $Y = X|X \leq x_0$, $x_0 \in \mathbb{R}$, $T = (-\infty, x_0]$ tiene como función de distribución

$$F_2(y) = \begin{cases} \frac{F_X(y)}{F_X(x_0)} & si \quad y < x_0 \\ 1 & si \quad y \geq x_0 \end{cases}$$

(iii) *La variable aleatoria* $Y = X|x_0 \leq X < x_1$, $x_0, x_1 \in \mathbb{R}$, $T = [x_0, x_1)$ *tiene como función de distribución*

$$F_3(y) = \begin{cases} 0 & si \quad y < x_0 \\ \frac{F_X(y) - F_x(x_0^-)}{F_X(x_0^-) - F_X(x_1^-)} & si \quad x_0 \leq y < x_1 \\ 1 & si \quad y \geq x_1 \end{cases}$$

Capítulo 6

Esperanza matemática

6.1. Esperanza o valor esperado

Definición 6.1.1. Sea X una variable aleatoria con función de distribución F_X. Se define la esperanza o valor esperado de X como

$$
\begin{aligned}
E(X) &= \int_{\mathbb{R}} x \; dF_X(x) \\
&= \begin{cases} \sum_{x_n \in D_x} x_n P_X(\{x_n\}) & si \quad X \; es \; discreta \\ \int_{-\infty}^{+\infty} x f_X(x) \; dx & si \quad X \; es \; absolutamente \; conntinua \\ \sum_{x_n \in D_x} x_n P_X(\{x_n\}) + \int_{-\infty}^{+\infty} x f_X(x) \; dx & si \quad X \; es \; mixta \end{cases}
\end{aligned}
$$

siempre que la serie y/o las integrales existan, es decir, convejan absolutamente.

Proposición 6.1.2. *Sea X una variable aleatoria con función de distribución F_X. Sea g una función medible y sea $Y = g(X)$. Entonces*

$$
\begin{aligned}
E(Y) &= \int_{\mathbb{R}} g(x) \; dF_X(x) \\
&= \begin{cases} \sum_{x_n \in D_x} g(x_n) P_X(\{x_n\}) & si \quad X \; es \; discreta \\ \int_{-\infty}^{+\infty} g(x) f_X(x) \; dx & si \quad X \; es \; absolutamente \; conntinua \\ \sum_{x_n \in D_x} g(x_n) P_X(\{x_n\}) + \int_{-\infty}^{+\infty} g(x) f_X(x) \; dx & si \quad X \; es \; mixta \end{cases}
\end{aligned}
$$

siempre que la esperanza exista.

Proposición 6.1.3. *Sea X una vaiable aleatoria con función de distribución F_X y sean $g, g_1, ..., g_n$ funciones medibles Borel tales que $E[|g_i(X)|] < +\infty$. Entonces*

(1) $E[|\sum_{i=1}^n g_i(X)|] < +\infty$ *y* $E[\sum_{i=1}^n g_i(X)] = \sum_{i=1}^n E[g_i(X)]$.

(2) Si $g_1(X) \leq g_2(X)$ *en casi todo punto, entonces* $E[g_1(X)] \leq E[g_2(X)]$.

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

6.2. Momentos de una variable o de su distribución

Definición 6.2.1 (Momento de orden k respecto a c)**.**

$$M_{k,c} = E[(X-c)^k]$$

con $c \in \mathbb{R}$ y $k = 0, 1, 2, ...$

Definición 6.2.2 (Momento ordinario de orden k respecto del origen)**.**

$$m_k = \alpha_k = E[X^k]$$

- $m_0 = 1$.
- $m_1 = E[X]$.
- $m_2 = E[X^2]$.

Definición 6.2.3 (Momento central de orden k)**.**

$$\mu_k = M_k = E[(X - E(X))^k]$$

- $\mu_0 = 1$.
- $\mu_1 = 0$.
- $\mu_2 = V[X] = m_2 - m_1$.
- $\mu_k = \sum_{i=0}^{k} \binom{k}{i} (-1)^{k-i} \alpha^i \alpha_1^{k-i}$.

Definición 6.2.4 (Momento absoluto de orden k)**.**

$$\beta_k = E[|X|^k]$$

m_k existe si y solo si $\beta_k < +\infty$.

Definición 6.2.5 (Momento factorial de orden k)**.**

$$\gamma_k = E[X(X-1)(X-2)...(X-k+1)]$$

- $\gamma_1 = E[X]$.
- $V[X] = \gamma_2 + \gamma_1 - \gamma_1^2$

6.3. Otras funciones

Definición 6.3.1 (Función generatriz de probabilidad)**.** Sea X una variable aleatoria discreta con valores enteros no negativos, con distribución de probabilidad $p_0, p_1, ..., p_n$. La función genetriz de probabilidad de la variable aleatoria X es una función real de variable real dada por la expresión

$$G_X(s) = \sum_{n=0}^{\infty} s^n p_n.$$

Esta función solo está definida para aquellos valores reales que hacen convergente esta serie de potencias.

Observación 6.3.2. $G_X(s) = E(s^X)$.

Teorema 6.3.3 (Teoema de inversión para la función generatriz de probabilidad)**.** *Si X es una variable aleatoria que toma valores enteros no negativos y G_X es su función generatriz de probabilidad, entonces*

$$P(X = n) = p_n = \frac{G_X^{(n)}(0)}{n!}$$

para todo $n = 0, 1, 2, ...$

Definición 6.3.4 (Función generatriz de momentos)**.** Sea X una variable aleatoria. Se define la función generatriz de momentos asociada a X como una función $M_X : \mathbb{R} \longrightarrow \mathbb{R}$ dada por la ecuación

$$M_X(t) = E(e^{tX})$$

siempre que la esperanza exista en un entorno del cero y donde e^{tX} es la variable aleatoria transformada de X mediante la función medible $g(x) = e^{tx}$.

Observación 6.3.5. Si X es una variable aleatoria discreta con valores enteros no negativos y distribución de probabilidad $\{p_n\}_{n=1}^{\infty}$ entonces

$$M_X(t) = E(e^{tX}) = \sum_{n=0}^{\infty} e^{tn} p_n.$$

Teorema 6.3.6. *Si X tiene una función generatriz de momentos M_X que es finita en $(-t_0, t_0)$, con $t_0 > 0$, entonces la distribución de probabilidad o función de densidad asociada a dicha variable aleatoria queda determinada de forma única.*

Teorema 6.3.7. *Sea X una variable aleatoria con M_X que es finita para $|t| < t_0$, con $t_0 > 0$. Enntonces*

(a) X posee momentos ordinarios de todos los órdenes.

(b) M_X admite un desarollo en serie de Taylor en torno al cero.

$$M_X(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} M_X^{(k)}(0).$$

Además $M_X^{(k)}(0) = E(X^k)$.

Definición 6.3.8 (Función característica)**.** Sea X una variable aleatoria con función de distribución F_X. Se define la función caractrística de X como una función de $\mathbb{R}$ en $\mathbb{R}$ dada por

$$\varphi(t) = E\left[e^{itX}\right] = \int_{\mathbb{R}} e^{itx}\ dF_X(x).$$

6.4. El problema de la existencia de momentos

Teorema 6.4.1. *Sea X una variable aleatoria con función de distribución F_X. Si existe m_k entonces existe m_j para cada $j < k$.*

Demostración. Vamos a verlo para variables aleatorias absolutamente continuas. Sea $r < k$, entonces

$$\begin{aligned} E[|X|^r] &= \int_{-\infty}^{+\infty} |x|^r f(x)\ dx = \int_{-\infty}^{-1} |x|^r f(x)\ dx + \int_{-1}^{1} |x|^r f(x)\ dx + \int_{-1}^{+\infty} |x|^r f(x)\ dx \\ &\leq \int_{-\infty}^{-1} |x|^k f(x)\ dx + \int_{-1}^{1} f(x)\ dx + \int_{-1}^{+\infty} |x|^k f(x)\ dx \\ &\leq \int_{-\infty}^{+\infty} |x|^k f(x)\ dx + \int_{-\infty}^{+\infty} f(x)\ dx = E[X] + 1 < +\infty. \end{aligned}$$

Luego, existe m_r para cada $r < k$. □

Corolario 6.4.2. *(1) Si X es una variable aleatoria con $m_2 < +\infty$, para cualesquiera $a, b \in \mathbb{R}$ se verifica que*

$$E[a + bX] = a + bE[X] \quad y \quad V[a + bX] = b^2 V[X].$$

(2) Si X es una variable aleatoria, entonces

$$Y = \frac{X - E[X]}{V[X]}$$

es una variable aleatoria.

Proposición 6.4.3. *Sea X una variable aleatoria con función de distribución F_X. Si existe $c \in \mathbb{R}$ tal que $P[|X| \leq c] = 1$, entonces existen los momentos β_r, m_r y se verifica que $|E[X^r]| \leq c^r$.*

Teorema 6.4.4. *Sea X una variable aleatoria con función de distribución F_X. Si existe m_k para algún $k > 0$ entonces*

$$\lim_{n \to +\infty} n^k P[|X| > n] = 0.$$

Teorema 6.4.5. *Sea X una variable aleatoria no negativa con función de distribución F_X. Entonces $E[X]$ existe si y solo si*

$$\int_0^{+\infty} (1 - F_X(x))\ dx < +\infty.$$

Además $E[X] = \int_0^{+\infty} (1 - F_X(x))\ dx = \int_0^{+\infty} P(X > x)\ dx$.

Teorema 6.4.6. *Sea X una variable aleatoria arbitraria con función de distribución F_X. Una condición necesaria y suficiente para que exista $E[X]$ es que existan las integrales*

$$\int_0^{+\infty} (1 - F_X(x))\ dx \quad y \quad \int_{-\infty}^0 F_X(x)\ dx,$$

y en este caso

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

Teorema 6.4.7. *Sea X una variable aleatoria arbitraria con función de distribución F_X. Una condición necesaria y suficiente para que exista $E[X]$ es que*

$$\sum_{n=1}^{+\infty} P[|X| \geq n] < +\infty.$$

Capítulo 7

Características numéricas de una distribución

7.1. Medidas de posición

Definición 7.1.1 (Esperanza)**.** Sea X una variable aleatoria y F_X su función de distribución. Se define la esperanza de la distribución como

$$E[X] = \int_{\mathbb{R}} x \; dF_X(x).$$

Proposición 7.1.2. *Algunas propiedades de la esperanza.*

(1) Si X es una variable aleatoria tal que $P(X \geq 0) = 1$, entonces $E[X] \geq 0$.

(2) Si X es una variable aleatoria tal que $P(X = c) = 1$, entonces $E[X] = c$.

(3) Si X es una variable aleatoria tal que $P(a < X < b) = 1$, entonces $a < E[X] < b$.

(4) Sea X una variable aleatoria y sea $Y = a + bX$, $a, b \in \mathbb{R}$. Si $E[X] < +\infty$, entonces $E[Y] = E[a] + E[bX] = a + bE[X] < +\infty$.

(5) Sea X una variable aleatoria con función de distribución F_X y sea $Y = g(X)$ con g función medible. Entonces $E[Y] = \int_{\mathbb{R}} g(x) \; dF_X(x)$.

Teorema 7.1.3 (Teorema de König)**.** *Sea X una variable aleatoria tal que $E[X^2] < +\infty$. El valor de $a \in \mathbb{R}$ que hace mínima $E[(X-a)^2]$ es $a = E[X]$.*

Demostración.

$$\begin{aligned} E[(X-a)^2] &= E[(X - E[X] + E[X] - a)^2] \\ &= E[(X - E[X])^2 + (E[X] - a)^2 + 2(X - E[X])(E[X] - a)] \\ &= E[(X - E[X])^2] + E[(E[X] - a)^2] + 2E[(X - E[X])(E[X] - a)] \end{aligned}$$

Observamos que

(i) $E[(X - E[X])^2] = V[X]$.

(ii) $E[(E[X] - a)^2] = (E[X] - a)^2$, pues $(E[X] - a)^2$ es una constante.

(iii)

$$E[(X - E[X])(E[X] - a)] = (E[X] - a)E[X - E[X]] = (E[X] - a)(E[x] - E[x]) = 0.$$

Por tanto

$$E[(X-a)^2] = V[X] + (E[X]-a)^2.$$

Nótese que $(E[X]-a)^2 \geq 0$, luego para que sea mínimo, $(E[X]-a)^2 = 0$, es decir, $a = E[X]$. $\square$

Definición 7.1.4 (Mediana)**.** Sea X una variable aleatoria con función de distribución F_X. M es una mediana de la distribución si verifica que

$$P(X \leq M) \geq \frac{1}{2} \quad y \quad P(X \geq M) \geq \frac{1}{2}.$$

Uniendo ambas condiciones, M es una mediana de la disitribución si

$$\frac{1}{2} \leq F_X(M) \leq \frac{1}{2} + P(X = M).$$

Definición 7.1.5 (Moda)**.** Sea X una variable aleatoria con función de distribució F_X. La moda, M_0, es el valor de mayor probabilidad (variable aleatoria discreta) o donde la función de densidad alcanza el maxímo (variable aleatoria absolutamente continua).

Definición 7.1.6 (Percentiles)**.** Sea X una variable aleatoria con función de distribució F_X. P_α es el percentil de orden α si

$$P_X(X \leq P_\alpha) \geq \frac{\alpha}{100} \quad y \quad P_X(X \geq P_\alpha) \geq 1 - \frac{\alpha}{100}.$$

Definición 7.1.7 (Cuartiles)**.** Son los percetiles de orden 25, 50 y 75, es decir, $Q_1 = P_{25}$, $Q_2 = P_{50}$ y $Q_3 = P_{75}$.

7.2. Medidas de dispersión

Definición 7.2.1 (Varianza)**.** Sea X una variable aleatoria con función de distribució F_X. Se define la varianza de la distribución como

$$V[X] = E[(X - E[X])^2]$$

siempre que $E[X^2] < +\infty$.

Observación 7.2.2. $V[X] = E[X^2] - (E[X])^2$.

Proposición 7.2.3. *(i) $V[X] \geq 0$ y $V[X] = 0$ si y solo si $P(X = a) = 1$.*

(ii) $V[a + bX] = b^2 V[X]$.

Demostración. Veamos (*ii*).

$$\begin{aligned} V[a+bX] &= E[(a+bX-E[a+bX])^2] = E[(a+bX-a+bE[X])^2] \\ &= E[b^2(X-E[X])^2] = b^2E[(X-E[X])^2] \\ &= b^2V[X]. \end{aligned}$$

$\square$

Definición 7.2.4 (Desviación típica)**.** Sea X una variable aleatoria. Se define la desviación típica como $D(X) = +\sqrt{V(X)}$.

Observación 7.2.5. Si X es una variable aleatoria con $V[X]$ conocida, entonces

$$Y = \frac{X - E[X]}{D(X)}$$

es una variable aleatoria con $E[Y] = 0$ y $V[Y] = 1$.

Teorema 7.2.6 (Desigualdad de Chebyshev)**.** *Sea X una variable aleatoria con $E[X]$ y $V[X]$ conocidas. Entonces*

$$P(|X - E[X]| \leq aD(x)) \geq 1 - \frac{1}{a^2}$$

para todo $a > 0$.

7.3. Medidas de asimetría

Definición 7.3.1. Una variable aleatoria X tiene un disrtibución simétrica respecto del origen si X y $-X$ tienen la misma función de distribución (es decir, otorgan misma probabilidad a mismos conjuntos).

Definición 7.3.2. Una variable aleatoria X tiene una distribución simétrica respecto $a \in \mathbb{R}$ si $X - a$ y $a - X$ tienen la misma función de distribución.

Observación 7.3.3. Sea X variable aleatoria y sea $E[X]$. Entonces

(i) $M_1 = E[X - E[X]] = E[x] - E[X] = 0$.

(ii) $M_2 = E[(X - E[X])^2]$.

(iii) $M_3 = E[(X - E[X])^3]$. Si X es simétrica respecto $\mu = E[X]$, entonces $M_3 = 0$, es más, $M_r = 0$ para r impar.

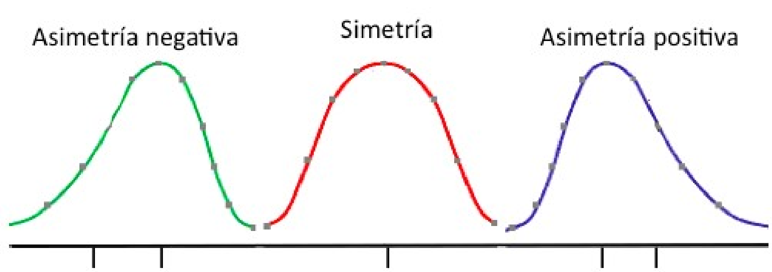
Definición 7.3.4 (Coeficiente de asimetría)**.**

$$CA = \frac{M_3}{(D(X))^2},$$

(i) Si $CA = 0$, entonces X es simértica.

(ii) Si $CA < 0$, entonces X es asimétrica a la izquierda.

(iii) Si $CA > 0$, entonces X es asimétrica a la derecha.



7.4. Coeficientes de apuntalamiento

Definición 7.4.1 (Coeficiente de apuntalamiento)**.**

$$CA_p = \frac{M_4}{(D(x))^4} - 3.$$

Observación 7.4.2. (i) $M_4 = E[(X - E[X])^4]$.

(ii) Si $X \sim N(0,1)$ entonces $CA_p = 0$.

Capítulo 8

Distribuciones discretas más usuales

8.1. Distribución degenerada

Definición 8.1.1. Se dice que una variable aleatoria $X \sim D(a)$ si su distribución de probabilidad viene dada por

$$P(X = x) = \begin{cases} 1 & si \quad x = a \\ 0 & si \quad x \neq a \end{cases}$$

Observación 8.1.2. Si $X \sim D(a)$ entonces

(i) $D_X = \{a\}$, $E[X] = a$ y $V[X] = 0$.

(ii) Su función de distribución es

$$F_X(x) = \begin{cases} 0 & si \quad x < a \\ 1 & si \quad x \geq a \end{cases}$$

(iii) $G_X(s) = E[s^X] = s^a P(X = a) = s^a$ para todo $s \in \mathbb{R}$.

(iv) $M_X(t) = e^{ta}$.

8.2. Distribución de Bernoulli

Definición 8.2.1. Se dice que una variable aleatoria $X \sim Ber(p)$ si

- $D_X = \{0, 1\}$.
- $P(X = 0) = 1 - p$ y $P(X = 1) = 1$, $0 < p < 1$.

Observación 8.2.2. Si $X \sim Ber(p)$ entonces

(i) $E[X] = p$.

(ii) $V[X] = p(1 - p)$.

(iii) $G_X(s) = E[s^X] = ps + q$ para todo $s \in \mathbb{R}$.

(iv) $M_X(t) = pe^t + q$.

Definición 8.2.3. Sean $X_1, ..., X_n$ variables aleatorias definidas sobre el mismo espacio de probabilidad. Decimos que son independientes si dados $B_1, ..., B_n \in \mathbb{B}_1$ se verifica

$$P(X_1 \in B_1, X_2 \in B_2, ..., X_n \in B_n) = \prod_{i=1}^{n} P(X_i \in B_i).$$

Observación 8.2.4. (1) Si $X_,...,X_n$ son variables aleatorias discretas, la condición de independencia se traduce como

$$P(X_1 = x_1, ..., X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i)$$
$$P(X_1 \leq x_1, ..., X_n \leq x_n) = \prod_{i=1}^{n} P(X_i \leq x_i) = \prod_{i=1}^{n} F_{X_i}(x_i).$$

(2) Si $X_1, ..., X_n$ son variables aleatorias absolutamente continuas, la condición de independencia se traduce como

$$f(x_1, ..., x_n) = \prod_{i=1}^{n} f_i(x_i)$$
$$P(X_1 \leq x_1, ..., X_n \leq x_n) = \prod_{i=1}^{n} F_{X_i}(x_i).$$

(3) Si $X_1, ..., X_n$ son variables aleatorias discretas independientes y si $T = \sum_{i=1}^{n} X_i$. Entonces T es variable aleatoria y

$$G_T(s) = \prod_{i=1}^{n} G_{X_i}(s).$$

(4) Si $X_1, ..., X_n$ son variables aleatorias discretas independientes y si $T = \sum_{i=1}^{n} X_i$. Entonces T es variable aleatoria y

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

8.3. Distribución Binomial

Definición 8.3.1. Sea X = número de éxitos al realizar n pruebas independientes de Bernoulli. Entonces $X \sim Bi(n, p)$ y

- $D_X = \{0, 1, ..., n\}$.
- $P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$, $i \in D_X$.

Observación 8.3.2. Si $X \sim Bi(n, p)$ entonces

(i) $G_X(s) = (ps + q)^n$.

(ii) $M_X(t) = (pe^t + q)^n = m(t)$.

(iii) $E[X] = m'(0) = np$.

(iv) $V[X] = np(1-p)$.

8.4. Distribución Geométrica

Definición 8.4.1. Sea X = número de fracasos hasta obtener el primer éxito al realizar pruebas independientes con $P(\text{éxito} = p)$. Entonces $X \sim Ge(p)$ y

- $D_X = \{0, 1, ...\}$.
- $P(X = n) = (1-p)^n p$, $n \in D_X$.

Observación 8.4.2. Si $X \sim Ge(p)$ entonces

(i) $= G_X(s) = \sum_{n=0}^{\infty} s^n P(X=n) = \sum_{n=0}^{\infty} s^n (1-p)^n p = p \sum_{n=0}^{\infty} (sq)^n = \frac{p}{1-sq}$ para todo $|s| \leq \frac{1}{q}$.

(ii) $M_X(t) = \frac{p}{q-e^t q}$.

(iii) $E[X] = \frac{q}{p}$.

(iv) $V[X] = \frac{q}{p^2}$.

8.5. Distribución Binomial Negativa

Definición 8.5.1. Sea X = número de fracasos hasta conseguir el r-ésimo éxito al realizar pruebas independientes con $P(\text{éxito}) = p$. Entonces $X \sim BN(r,p)$ y

- $D_X = \{0, 1, ...\}$.
- $P(X=n) = \binom{n+r-1}{n}(1-p)^n p^r$, $r \in D_X$.

Observación 8.5.2. Si $X \sim BN(r,p)$ entonces

- $X = \sum_{i=1}^{r} X_i$, donde $X_i \sim Ge(p)$ son variables aleatorias independientes.
- $G_X(s) = \left(\frac{p}{1-s(1-p)}\right)^r$.
- $m(t) = \left(\frac{p}{1-e^t(1-p)}\right)^r$.
- $E[X] = \frac{r(1-p)}{p}$.
- $V[X] = \frac{r(1-p)}{p^2}$.

8.6. Distribución de Poisson

Definición 8.6.1. $X \sim Po(\lambda)$, $\lambda > 0$, si

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

Observación 8.6.2. Si $X \sim Po(\lambda)$, $\lambda > 0$, entonces

- $G_X(s) = e^{-\lambda(1-s)}$.
- $M_X(t) = e^{-\lambda(1-e^t)}$.
- $E[X] = V[X] = \lambda$.

8.7. Distribución Hipergeométrica

Definición 8.7.1. Consideremos una población de N elementos, donde hay D elementos de la clase A y $N-D$ elementos de la clase B. Sea X = número de elementos de la clase A al tomar una muestra (sin reemplazamiento) de tamaño n. Entonces $X \sim HG(N,D,n)$ y

- $D_X = \{0, 1, ..., n\}$.
- $$P(X=i) = \frac{\binom{D}{i}\binom{N-D}{n-i}}{\binom{N}{n}}, \; i \in D_X.$$

Capítulo 9

Distribuciones continuas más usuales

9.1. Distribución Uniforme

Definición 9.1.1. Se dice que una variable aleatoria $X \sim U[a,b]$ si su función de densidad es

$$f_X(x) = \begin{cases} \frac{1}{b-a} & si \quad a \leq x \leq b \\ 0 & si \quad resto \end{cases}.$$

Observación 9.1.2. Si $X \sim U[a,b]$ entonces

- Su función de distribución es

$$F_X(x) = \begin{cases} 0 & si \quad x < a \\ \frac{x-a}{b-a} & si \quad a \leq x < b \\ 1 & si \quad x \geq b \end{cases}.$$

- $P(x \leq X \leq x + \Delta x) = \frac{\Delta x}{b-a}$.
- $E[X] = \frac{b+a}{2}$.
- $V[X] = \frac{(b-a)^2}{12}$.
- $M_X(t) = E[e^{tx}] = \int_a^b e^{tx} \frac{1}{b-a}\, dx = \frac{e^{tb}-e^{ta}}{t(b-a)}$, si $t \neq 0$. Entonces

$$M_X(t) = \begin{cases} \frac{e^{tb}-e^{ta}}{t(b-a)} & si \quad t \neq 0 \\ 0 & si \quad t = 0 \end{cases}.$$

Proposición 9.1.3. *Sea X una variable aleatoria y F_X su función de distribución con F_X estrictamente creciente. Entonces $Y = F(X)$ es una variable aleatoria e $Y \sim U(0,1)$.*

Demostración. El soporte de Y es $S_Y = [0,1]$.

- Si $y < 0$, es claro que $F_Y(y) = 0$.
- Si $0 \leq y < 1$

$$F_Y(y) = P(Y \leq y) = P(F(x) \leq y) = P(X \leq F^{-1}(y)) = F(F^{-1}(y)) = y.$$

- Si $y \geq 1$, es claro que $F_Y(y) = 1$.

Por tanto

$$F_Y(y) = \begin{cases} 0 & si & x < 0 \\ y & si & 0 \leq x < 1 \\ 1 & si & x \geq 1 \end{cases}.$$

lo que nos dice que $Y \sim U(0,1)$. □

9.2. Distribución Normal

Definición 9.2.1. Se dice que una variable aleatoria $X \sim N(\mu, \sigma)$, $\mu \in \mathbb{R}$, $\sigma > 0$, si su función de densidad es

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}},\ x \in \mathbb{R}.$$

Observación 9.2.2. Si $X \sim N(\mu, \sigma)$ entonces

- $E[X] = \mu$.
- $D(X) = \sigma$.
- Si $Z \sim N(0,1)$ entonces $E[Z] = 0$, $V[Z] = 1$ y $D(Z) = 1$.
- $Z = \frac{X-\mu}{\sigma}$ es variable aleatoria y $Z \sim N(0,1)$.
- Si $Z \sim N(0,1)$ entonces su función de densidad es

$$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2},\ z \in \mathbb{R}.$$

  Su función generatriz de momentos es

$$M_Z(t) = E[e^{tz}] = \int_{-\infty}^{+\infty} e^{tz} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\ dz = e^{t^2/2}$$

  para todo $t \in \mathbb{R}$ (para poder calcular esta integral, es útil saber que $\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\ dz = 1$).

- Haciendo uso de la función generatriz de momentos de $Z \sim N(0,1)$, calculemos la función generatriz de momentos de $X \sim N(\mu, \sigma)$. La funcuión de densidad de X es

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}},\ x \in \mathbb{R}.$$

  Nótese que, $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$, despenjando, $X = \sigma Z + \mu$. Calculemos la función generatriz de momentos de X

$$\begin{aligned} M_X(t) &= E[e^{tx}] = E[e^{t(\sigma z + \mu)}] = E[e^{t\sigma z} \cdot e^{t\mu}] = e^{t\mu} E[e^{(t\sigma)z}] \\ &= e^{t\mu} M_Z(t\sigma) \end{aligned}$$

  donde $Z \sim N(0,1)$, luego

$$M_X(t) = e^{t\mu} M_Z(t\sigma) = e^{t\mu} e^{\frac{1}{2}t^2\sigma^2}$$

  para todo $t \in \mathbb{R}$.

Observación 9.2.3. Si X e Y son varibles aleatorias independientes

(1) $E[X \cdot Y] = E[X] \cdot E[Y]$.

(2) $G_{X+Y}(s) = G_X(s) \cdot G_Y(s)$ y $M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$.

(3) $V[X+Y] = V[X] + V[Y]$ y $V[X-Y] = V[X] + V[Y]$.

Proposición 9.2.4. *(1) Si $X_1, ..., X_n$ son variables aleatorias independientes tales que $X_i \sim N(\mu_i, \sigma_i)$, entonces*

$$S = \sum_{i=1}^{n} X_i \sim N\left(\sum_{i=1}^{n} \mu_i, \sqrt{\sum_{i=1}^{n} \sigma_i^2}\right).$$

(2) Si $X_1, ..., X_n$ son variables aleatorias independientes tales que $X_i \sim N(\mu, \sigma)$, entonces

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right).$$

$\overline{X}$ se conoce como media muestral.

9.3. Distribución Exponencial

Definición 9.3.1. Se dice que una variable aleatoria $X \sim Exp(\lambda)$, $\lambda > 0$, si su función de densidad es

$$f(x) = \begin{cases} 0 & si \quad x < 0 \\ \lambda e^{-\lambda x} & si \quad x \geq 0 \end{cases}.$$

Observación 9.3.2. Si $X \sim Exp(\lambda)$ entonces

- Su función de distribución es

$$F(x) = \begin{cases} 0 & si \quad x \leq 0 \\ 1 - \lambda e^{-\lambda x} & si \quad x > 0 \end{cases}.$$

- $E[X] = \frac{1}{\lambda}$.
- $V[X] = \frac{1}{\lambda^2}$.
- $Me = \frac{\log(2)}{\lambda}$.
- Su función generatriz de momentos es

$$M_X(t) = \begin{cases} \left(1 - \frac{t}{\lambda}\right)^2 & si \quad t < \lambda \\ \text{no existe} & si \quad t \geq \lambda \end{cases}.$$

Teorema 9.3.3 (Propiedad de falta de memoria)**.** *Si $X \sim Exp(\lambda)$, entonces para todo $a > 0$ y $x > 0$ se verifica*

$$P(X > a + x \mid X \geq a) = P(X > x).$$

Relación entre las distribuciones Exponencial y Poisson

Si X = número de ocurrencias de un suceso en un intervalo de t unidades de medida, entonces $X \sim Po(\lambda t)$. Sea T = tiempo que transcurre entre dos ocurrencias sucesivas de un suceso. Entonces $T \sim Exp(\lambda)$.

9.4. Distribución Gamma

Definición 9.4.1. Se dice que una variable aleatoria $X \sim Ga(a,p)$, $a, p > 0$, si su función de densidad es

$$f(x) = \begin{cases} 0 & si \quad x \leq 0 \\ \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1} & si \quad x > 0 \end{cases},$$

donde

$$\Gamma(p) = \int_0^{+\infty} x^{p-1} e^{-x} dx.$$

Observación 9.4.2.

- $\Gamma(n) = (n-1)!$ para todo $n \in \mathbb{N}$.
- $\Gamma(p) = (p-1)\Gamma(p-1)$ para todo $p > 0$.
- $\Gamma(1) = 1$ y $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.
- $\int_0^{+\infty} x^{p-1} e^{-x} \ dx = \frac{\Gamma(p)}{a^p}$.

Observación 9.4.3. Si $X \sim Ga(a,p)$ entonces

- El momento ordinario de orden r respecto del origen

$$\begin{aligned} m_r = E[X^r] &= \int_0^{+\infty} x^r \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1} \ dx = \frac{a^p}{\Gamma(p)} \int_0^{+\infty} x^{p+r-1} e^{-ax} \ dx \\ &= \frac{\Gamma(p+r)}{\Gamma(p)} \cdot \frac{1}{a^r}. \end{aligned}$$

  En particular

  - $m_1 = \frac{p}{a}$.
  - $m_2 = \frac{p(p+1)}{a^2}$.

- $V[X] = m_2 - m_1^2 = \frac{p}{a^2}$.
- La función generatriz de momentos es

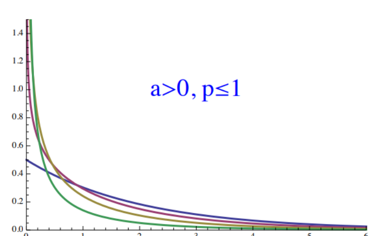
$$\begin{aligned} M_X(t) = E[e^{tx}] &= \int_0^{+\infty} e^{tx} \frac{a^p}{\Gamma(p)} e^{-ax} x^{p-1} \ dx = \frac{a^p}{\Gamma(p)} \int_0^{+\infty} e^{-x(a-t)} x^{p-1} \ dx \\ &= \frac{a^p}{\Gamma(p))} \cdot \frac{\Gamma(p)}{(a-t)^p}, \end{aligned}$$

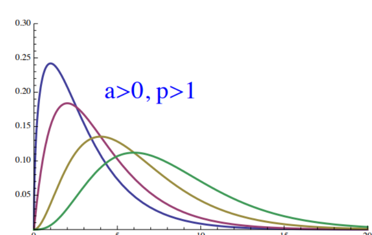
  por tanto,

$$M_X(t) = \left(\frac{a}{a-t}\right)^p$$

  para todo $a - t > 0$.

Al parámetro p se le suele llamar **parámetro de forma** y al parámetro a, **parámetro de escala**.





Proposición 9.4.4. *Si $X \sim Ga(a,p)$ entonces*

(1) Sea $Y = cX$ entonces $Y \sim Ga\left(\frac{a}{c}, p\right)$.

(2) Sea $Y = X/a$ entonces $Y \sim Ga(a^2, p)$.

(3) Sea $Y = pX$ entonces $Y \sim Ga\left(\frac{a}{p}, p\right)$.

(4) Sea $Y = X/p$ entonces $Y \sim Ga(ap, p)$.

(5) Sea $Y = 1/X$ entonces $Y \sim GaI(a,p)$ con función de densidad dada por

$$f_Y(y) = \frac{a^p}{\Gamma(p)} e^{-a\frac{1}{y}} \left(\frac{1}{y}\right)^{p+1}, \ y > 0.$$

(6) Las distribución Gamma es reproductiva respecto de p. Sean $X_1, ..., X_n$ variables aleatorias independientes tales que $X_i \sim Ga(a, p_i)$, entonces

$$S = \sum_{i=1}^{n} X_i \sim Ga\left(a, \sum_{i=1}^{p} p_i\right).$$

9.4.1. Distribución Chi-Cuadrada

Definición 9.4.5. Decimos que X se distribe según una chi cuadrada con n grados de libertad, $X \sim \chi_n^2$, si $X \sim Ga\left(\frac{1}{2}, \frac{n}{2}\right)$ con $n \in \mathbb{N}$. Su función de densidad viene dada por

$$f(x) = \begin{cases} 0 & si \quad x \leq 0 \\ \frac{\left(\frac{1}{2}\right)^{n/2}}{\Gamma(n/2)} e^{-x/2} x^{\frac{n}{2}-1} & si \quad x > 0 \end{cases},$$

Observación 9.4.6. Si $X \sim \chi_n^2$ entonces

- $E[X] = n$.
- $V[X] = 2n$.
- $M_X(t) = \left(1 - \frac{1}{2t}\right)^{n/2}$ si $t < \frac{1}{2}$.

Proposición 9.4.7. *(1) Si $X \sim U[0,1]$ y si $Y = -2\log(X)$ entonces $Y \sim \chi_2^2$.*

(2) Si $X \sim N(\mu, \sigma)$ y si $Y = \left(\frac{X-\mu}{\sigma}\right)^2$ entonces $Y \sim \chi_1^2$.

(3) Si $X_1, ..., X_n$ son variables aleatorias independientes con $X_i \sim N(0,1)$ para todo i entonces $\sum_{i=1}^{n} X_i^2 \sim \chi_n^2$.

(4) Si $X_1, ..., X_n$ son variables aleatorias independientes con $X_i \sim \chi_{n_i}^2$ entonces $\sum_{i=1}^{n} X_i^2 \sim \chi_m^2$ con $m = \sum_{i=1}^{n} n_i$.

(5) Convergencia en ley a la normal: $\chi_n^2 \xrightarrow{l} N(n, \sqrt{2n})$. Además

$$Z_n = \sqrt{2\chi_n^2} - \sqrt{2n-1} \xrightarrow{l} N(0,1).$$

9.5. Distribución de Erlang

Definición 9.5.1. Se dice que una variable aleatoria $X \sim Er(k, \lambda)$ si $X \sim Ga(a = \lambda, p = k)$, $k \in \mathbb{N}$, $\lambda > 0$.

Observación 9.5.2. Si $X_1, ..., X_n$ son variables aleatorias independientes tales que $X_i \sim Er(k_i, \lambda)$ entonces $\sum_{i=1}^{n} X_n \sim Er\left(\sum_{i=1}^{n} k_i, \lambda\right)$.

9.6. Distribución Beta

Definición 9.6.1. Se dice que una variable aleatoria $X \sim Be(\alpha, \beta)$ con $\alpha > 0$, $\beta > 0$ si su función de densidad es

$$f(x) = \begin{cases} \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} & si \quad 0 \leq x \leq 1 \\ 0 & si \quad\quad resto \end{cases}$$

donde

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}\, dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Observación 9.6.2. El momento ordinario de orden k respecto del origen

$$m_k = E[X^k] = \int_0^1 x^k \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^k x^{\alpha-1}(1-x)^{\beta-1}\, dx = \frac{\Gamma(\alpha+\beta)}{\Gamma(a)} \cdot \frac{\Gamma(\alpha+k)}{\Gamma(\alpha+\beta+k)}.$$

En particular

- $m_1 = \frac{\alpha}{\alpha+\beta}$.
- $m_2 = \frac{(\alpha+1)\alpha}{(\alpha+\beta)(\alpha+\beta+1)}$.

9.7. Aproximación de distribuciones

Definición 9.7.1. Sea $\{X_n\}$ una sucesión de variables aleatorias definidas todas ellas sobre un mismo espacio de probabilidad. Diremos que la sucesión converge en ley o distribución hacia la variable aleatoria X, definida en el mismo espacio de probabilidad, si y solo si la correspondiente sucesión de funciones de distribución $\{F_n\}$ converge hacia la función de distribución F de la variable aleatoria X en todo punto de continuidad de F.

9.7.1. Teoremas de caracterización de la convergencia en ley para variables aleatorias discretas

Teorema 9.7.2. *Si la sucesión de variables aleatorias $\{X_n\}$ toma valores enteros no negativos, para todo n, entonces la condición necesaria y suficiente para que dicha sucesión converja en ley hacia una variable aleatoria discreta X es*

$$\lim_{n\to\infty} P(X_n = k) = P(X = k) \quad \textit{para todo } k \in \mathbb{N}.$$

Teorema 9.7.3. *Si existe la función generatriz de probabilidad $G_n(s)$ para cada una de las variables aleatorias X_n y existe la función generatriz de probabilidad de la variable aleatoria X, $G_X(s)$, de la variable aleatoria límite, entonces*

$$\lim_{n\to\infty} G_n(s) = G_X(s) \quad \textit{para todo } s \textit{ donde } G_X(s) \textit{ exista.}$$

En este caso la sucesión $\{X_n\}$ converge en ley o distribución hacia la variable aleatoria X.

9.7.2. Aproximación de una distribución Binomial por una distribución de Poisson

La distribución de Poisson surgió como una convergencia de la distribución Binomial de parámetros n y p con las siguientes condiciones

- $n \to +\infty$.

- $p \to 0$.
- $np \to \lambda$, con λ constante.

En tales condiciones la distribución Binomial se puede aproximar por una distribuión de Poisson de parámetro $\lambda = np$ (Teorema de Poisson).

Obsérvese que la media de la distribución $Bi(n, p)$ es np que coincide con la media de la distribución $Po(\lambda = np)$. Sin embargo, la varianza de la distribución $Bi(n, p)$ es nqp y la de la distribución $Po(\lambda = np)$ es np. Por lo tanto esta aproximación será más exacta cuanto más próximo a 1 sea q y por lo tanto cuanto más próximo a cero sea p.

9.7.3. Teoremas de caracterización de la convergencia en ley para variables aleatorias absolutamente continuas

Teorema 9.7.4. *Sean $\{X_n\}$ y X variables aleatorias absolutamente continuas con funciones de densidad f_n y f respectivamente, tales que*

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \textit{para casi todo } x.$$

Entonces la sucesión de variables aleatorias converge en ley hacia la variable aleatoria X.

Teorema 9.7.5. *Sean $\{X_n\}$ y X variables aleatorias absolutamente continuas con funciones de densidad f_n y f respectivamente, y funciones generatrices de momentos $M_n(t)$ y $M(t)$ respectivamente, tales que*

$$\lim_{n\to\infty} M_n(t) = M(t),$$

para todos los valores de t en un intervalo alrededor del punto 0. Entonces la sucesión de variables aleatorias converge en ley hacia la variable aleatoria X.

Observación 9.7.6.

- Este último teorema es válido tanto para variables discretas como para variables continuas.
- No solo es posible aproximar distribuciones discretas por discretas, también es posible la aproximación de una distribución discreta por una distribución continua. En la mayoría de los casos se suele hacer mediante una distribución normal.

9.7.4. Teorema central del límite

Definición 9.7.7. Sea $\{X_n\}$ una sucesión de variables aleatorias, definidas todas ellas sobre el mismo espacio de probabilidad, con medias y varianzas finitas. Diremos que la sucesión $\{X_n\}$ obedece al teorema central del límite si y solo si $\{Z_n\}$ converge en ley a una distribución $N(0, 1)$, siendo

$$Z = \frac{S_n - E[S_n]}{\sigma(S_n)}$$

con $S_n = \sum_{i=1}^{n} X_i$.

Teorema 9.7.8 (Teorema de Moivre)**.** *Sea $\{X_n\}$ una sucesión de variables aleatorias independientes e idénticamente distribuidas, cada una de ellas $Bi(n = 1, p)$ (o lo que es lo mismo $Ber(p)$). Entonces la sucesión obedece al teorema central del límite.*

Corolario 9.7.9. *Si $X \sim Bi(n, p)$ y n es grande, la distribución de X se puede aproximar mediante $N(np, \sqrt{npq})$.*

Capítulo 10

Vectores aleatorios

10.1. Vectores aleatorios

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y consideremos la función $\boldsymbol{X}$ definida de la forma

$$\boldsymbol{X} : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}^n, \mathbb{B}_n, P_{\boldsymbol{X}})$$

donde $\boldsymbol{X}(\omega) = (X_1(\omega), X_2(\omega), ..., X_n(\omega))$ y $\mathbb{B}_n$ es la σ-álgebra de Borel en $\mathbb{R}^n$, es decir, la mínima σ-álgebra generada por los conjuntos

$$C_{\boldsymbol{X}} = \{\mathbf{y} \in \mathbb{R}^n : y_1 \leq x_1, ..., y_n \leq x_n\}$$

con $\boldsymbol{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$.

Definición 10.1.1. Sea $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ un vectores de funciones definido en $(\Omega, \mathcal{A}, P)$ a $\mathbb{R}^n$, donde para cada $\omega \in \Omega$

$$\boldsymbol{X}(\omega) = (X_1(\omega), X_2(\omega), ..., X_n(\omega))$$

Diremos que $\boldsymbol{X}$ es un vector aleatorio o una variable aleatoria n-dimensional si, para cada vector de números reales, $\mathbf{a} = (a_1, a_2, ..., a_n)$, la imagen inversa del intervalo n-dimensional $I = \{(x_1, ..., x_n) : x_i \leq a_i, i = 1, 2, ..., n\}$ pertenece a la σ-álgebra $\mathcal{A}$, es decir,

$$\boldsymbol{X}^{-1}(I) = \{\omega \in \Omega : X_1(\omega) \leq a_1, ..., X_n(\omega) \leq a_n\} \in \mathcal{A}.$$

Teorema 10.1.2. *Sean $X_1, X_2, ..., X_n$, n variables aleatorias definidas sobre $(\Omega, \mathcal{A}, P)$. Entonces $\boldsymbol{X} = (X_1, X_2, ..., X_n)$ es una variable aleatoria n-dimensional sobre $(\Omega, \mathcal{A}, P)$.*

Demostración. Sea $I = \{(x_1, ..., x_n) : x_i \leq a_i, a_i \in \mathbb{R}, i = 1, 2, ..., n\}$. Entonces

$$\begin{aligned}(X_1, X_2, ..., X_n)^{-1}(I) &= \{\omega \in \Omega : X_1(\omega) \leq a_1, ..., X_n(\omega) \leq a_n\} \\ &= \bigcap_{i=1}^{n} \{\omega \in \Omega : X_i(\omega) \leq a_i\} \\ &= \bigcap_{i=1}^{n} X_i^{-1}(-\infty, a_i]\end{aligned}$$

Por ser X_i variable aleatoria, se tiene que $X_i^{-1}(-\infty, a_i] \in \mathcal{A}$, por lo que $(X_1, X_2, ..., X_n)^{-1}(I) \in \mathcal{A}$. □

10.2. Distribución de probabilidad de un vector aleatorio

Dado un vector aleatorio $\boldsymbol{X}$, se denomina distribución de probabilidad inducida por $\boldsymbol{X}$ en $(\mathbb{R}^n, \mathbb{B}_n)$, a la función

$$P_{\boldsymbol{X}} : \mathbb{B}_n \longrightarrow [0,1]$$

definida para cada $B \in \mathbb{B}_n$ de la forma

$$P_{\boldsymbol{X}}(B) = P\{\boldsymbol{X}^{-1}(B)\} = P\{\boldsymbol{X} \in B\}.$$

$P_{\boldsymbol{X}}$ es una medida de probabilidad, por lo tanto, el vector aleatorio $\boldsymbol{X}$ transforma el espacio de probabilidad $(\Omega, \mathcal{A}, P)$ en un nuevo espacio de probabilidad $(\mathbb{R}^n, \mathbb{B}_n, P_{\boldsymbol{X}})$.

10.3. Función de distribucón conjunta

Dado un vector aleatorio $\boldsymbol{X}$, se define la función de distribución conjunta como la función $F_{\boldsymbol{X}} : \mathbb{R}^n \longrightarrow [0,1]$ definida de la forma

$$F_{\boldsymbol{X}}(x_1, x_2, ..., x_n) = P(X_1 \leq x_1, X_2 \leq x_2, ..., X_n \leq x_n).$$

Teorema 10.3.1. *La función de distribución de probabilidad de un vector aleatorio satisface las siguientes propiedades:*

(P1) $F(+\infty, +\infty, ..., +\infty) = 1.$

(P2) $F(x_1, ..., x_{i-1}, -\infty, x_{i+1}, ..., x_n) = 0.$

(P3) La función de distribución conjunta es creciente para cada componente.

(P4) La función de distribución conjunta es continua por la derecha para cada componente.

10.3.1. Vectores bidimensionales

Estudiaremos con más detenimiento el caso de que $n = 2$. Por lo tanto consideramos un vector aleatorio bidiensional $\boldsymbol{X} = (X, Y)$, con función de distribución $F_{(X,Y)}(x, y) := F(x, y)$. Las propiedades quedan ahora de la forma

Teorema 10.3.2. *(P1)* $F(+\infty, +\infty) = 1.$

(P2) $F(x, -\infty) = 0$ *y* $F(-\infty, y) = 0.$

(P3) La función de distribución conjunta es creciente para cada componente.

(P4) La función de distribución conjunta es continua por la derecha para cada componente.

Demostración. (P1) $F(n, n) = P(X \leq n, Y \leq n) = P_{(X,Y)}(C_{(n,n)})$ donde

$$C_{(n,n)} = \{(x, y) \in \mathbb{R}^2 : x \leq n, y \leq n\}.$$

Nótese que $\{C_{(n,n)}\} \uparrow \mathbb{R}^2$, por tanto, la probabilidad y el límite pueden conmutar y tenemos

$$\lim_{n\to\infty} F(n, n) = \lim_{n\to\infty} P_{(X,Y)}(C_{(n,n)}) = P_{(X,Y)}(\lim_{n\to\infty} C_{(n,n)}) = P_{(X,Y)}(\mathbb{R}^2) = 1.$$

(P2) $F(0, -n) = P_{(X,Y)(C_{(x,-n)})}$. Nótese que $\{C_{(x,-n)}\} \downarrow \emptyset$, por tanto, la probabilidad y el límite pueden conmutar y tenemos

$$\lim_{n\to\infty} F(x, -n) = \lim_{n\to\infty} P_{(X,Y)(C_{(x,-n)})} = P_{(X,Y)}(\lim_{n\to\infty} C_{(x,-n)}) = P_{(X,Y)}(\emptyset) = 0.$$

El argumento es análolgo para $F(-\infty, y)$.

(P3) Veamos que F es creciente en la primera componente. Sean $x_1, x_2 \in \mathbb{R}$ con $x_1 < x_2$. Nótese que $C_{(x_1,y)} \subset C_{(x_2,y)}$. Entonces

$$F(x_1, y) = P_{(X,Y)}(C_{(x_1,y)}) \leq P_{(X,Y)}(C_{(x_2,y)}) = F(x_2, y)$$

El argumento es análogo para la segunda componente.

□

Observación 10.3.3. Estas cuatro propiedades no caracterizan a la función de distribución de un vector aleatorio, en el sentido de que dada una funcón $F : \mathbb{R}^2 \longrightarrow [0,1]$ que satisfaga estas cuatro propiedades no asegura que F sea la función de distribución de un vector aleatorio. Para poder asegurarlo, tenemos que añadir una propiedad adicional.

Si $F : \mathbb{R} \longrightarrow [0,1]$ verifica $P1$, $P2$, $P3$ y $P4$, F será función de distribución si además verifica

$$P_{(X,Y)}((a,b] \times (c,d]) \in [0,1] \iff F(b,d) - F(b,c) - F(a,d) + F(a,c) \in [0,1].$$

10.3.2. Cálculo de probabilidad de conjuntos de $\mathbb{R}^2$ a partir de la función de distribución conjunta

(1) $P_{\boldsymbol{X}}((-\infty, x) \times (-\infty, y]) = P(X < x, Y \leq y) = F(x^-, y)$.

(2) $P_{\boldsymbol{X}}(\{x_1\} \times (-\infty, y]) = F(x_1, y) - F(x_1^-, y)$.

(3) $P_{\boldsymbol{X}}((-\infty, x] \times \{y\}) = F(x, y) - F(x, y^-)$.

(4) $P_{\boldsymbol{X}}((-\infty, x] \times (-\infty, y]) = F(x, y)$.

10.4. Vectores aleatorios discretos

Son vectores de la forma $\boldsymbol{X} = (X, Y)$ donde X e Y son variables aleatorias discretas. Además

- $(X, Y) = \cup_{(i,j) \in I \times J}(x_i, y_j)$.
- $P(X = x_i, Y = y_j) \in [0,1]$.
- $\sum_{i \in I} \sum_{j \in I} P(X = x_i, Y = y_j) = 1$.
- Si $D \subset \mathbb{R}^2$ entonces $P_{(X,,Y)}(D) = \sum_{(x_i,y_j) \in D} P(X = x_i, Y = y_j)$.
- $P_{ij} = P(X = x_i, Y = y_j)$.

10.5. Vectores aleatorios absolutamente continuos

Definición 10.5.1. La distribución de probabilidad de P en $(\mathbb{R}^2, \mathbb{B}_2)$ y su función de distribución $F(x,y)$ se denominan absolutamente continuas si existe una función no negativa $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ tal que para todo rectángulo $I \subset \mathbb{R}^2$

$$P(I) = \int_I f(x,y)\ dxdy.$$

No hay problema en extender esta integral a rectángulos no acotados mediante el paso al límite, ya que $F \geq 0$ por lo que $\int_I f$ crece al crecer I y además está acotada por 1. En consecuencia

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s,t)\ dsdt.$$

donde F es la función de distribución conjunta de la variable aleatoria bidimensional $P_{\boldsymbol{X}} = (X, Y)$. A esta función f se le llama *función de densidad* de la variable aleatoria bidimensional $P_{\boldsymbol{X}} = (X, Y)$. Además la función de distribución conjunta de la variable aleatoria bidimensional debe verificar que

$$F(+\infty, +\infty) = \lim_{x,y \to +\infty} \int_{-\infty}^{x} \int_{-\infty}^{y} f(s,t)\ dsdt = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(s,t)\ dsdt = 1.$$

Proposición 10.5.2. *Una función $f : \mathbb{R} \longrightarrow \mathbb{R}$ no negativa e integrable-Riemann en cualquier rectángulo de $\mathbb{R}^2$, es función de densidad de alguna distribución bidimensional si y solo si*

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(s,t)\ dsdt = 1.$$

En tal caso, f es la función de densidad de la función de distribución absolutamente continua F definida como

$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(s,t)\ dsdt.$$

Teorema 10.5.3. *Sea $\boldsymbol{X} = (X, Y)$ una variable aleatoria bidimmensional continua con función de densidad f y función de distribución F. Entonces:*

(a) *F es continua para todo $(x, y) \in \mathbb{R}^2$.*

(b) *Si f es continua en un entorno del punto (x_0, y_0), entonces F admite en dicho entorno derivada segunda respecto a (x, y) siendo para todo (x_1, y_1) de este entorno*

$$\frac{\partial F(x_1,y_1)}{\partial x \partial y} = \frac{\partial F(x_1,y_1)}{\partial y \partial x} = f(x_1, y_1).$$

(c) *El conjunto de valores de la variable aleatoria bidimensional con probabilidad positiva es el conjunto vacío.*

(d) *Para todo $B \in \mathbb{B}_2$*

$$P(B) = \int_B f(x,y)\ dxdy.$$

10.6. Distribuciones marginales

10.6.1. Caso discreto

- (X, Y) variable aleatoria bidimensional discreta.
- $(X, Y) = \cup_{i \in I, j \in J} \{(x_i, y_j)\}$.
- *Distribución marginal de X.* Sea X variable aleatoria discreta, $X = \cup_{i \in I} \{x_i\}$, entonces

$$P(X = x_i) = \sum_{j \in J} P(X = x_i, Y = y_j).$$

- *Distribución marginal de Y.* Sea Y variable aleatoria discreta, $Y = \cup_{j \in J} \{x_j\}$, entonces

$$P(Y = y_j) = \sum_{i \in I} P(X = x_i, Y = y_j).$$

10.6.2. Caso continuo

Definición 10.6.1. Llamaremos función de distribución marginal de X a una función real de variable real $F_X : \mathbb{R} \longrightarrow [0,1]$ definida por

$$F_X(x) = F(x, +\infty) = \lim_{y \to +\infty} F(x,y) = \int_{-\infty}^{x} \left(\int_{-\infty}^{+\infty} f(s,t)\ dt \right) ds.$$

A la función $f_X(x) = \int_{-\infty}^{+\infty} f(x,y)\ dy$ se le llama *densidad marginal de X* y en los puntos en los que F_X sea derivable la densidad marginal coincide con esta derivada.

Definición 10.6.2. Llamaremos función de distribución marginal de Y a una función real de variable real $F_X : \mathbb{R} \longrightarrow [0,1]$ definida por

$$F_X(y) = F(-\infty, , y) = \lim_{x \to -\infty} F(x,y) = \int_{-\infty}^{y} \left(\int_{-\infty}^{+\infty} f(s,t)\ ds \right) dt.$$

A la función $f_Y(x) = \int_{-\infty}^{+\infty} f(x,y)\ dx$ se le llama *densidad marginal de Y*.

10.7. Distribuciones condicionadas

10.7.1. Caso discreto

- (X, Y) variable aleatoria bidimensional discreta.
- $(X, Y) = \cup_{i \in I, j \in J} \{(x_i, y_j)\}$.
- $\sum_{i \in I} \sum_{j \in J} P_{ij} = 1$.
- $X|Y = y_j$.

$$P(X = x_i | y = y_j) = \frac{P(X = x_i, Y = y_j)}{P(Y = y_j)} = \frac{P_{ij}}{P(Y = y_j)}.$$

- $Y|X = x_i$.

$$P(Y = y_j, X = x_i) = \frac{P(X = x_i, Y = y_j)}{P(X = x_i)} = \frac{P_{ij}}{P(X = x_i)}$$

10.7.2. Caso continuo

Supongamos $A = (-\infty, x]$ y $B = (-\infty, y_0]$, en tal caso

$$\begin{aligned} P[X \leq x | Y \leq y_0] = \frac{P[X \leq x, Y \leq y_0]}{P[Y \leq y_0]} &= \frac{\int_{-\infty}^{x} \int_{-\infty}^{y_0} f(u,v)\ dudv}{\int_{-\infty}^{y_0} \int_{-\infty}^{+\infty} f(u,v)\ dudv} \\ &= \frac{\int_{-\infty}^{x} \int_{-\infty}^{y_0} f(u,v)\ dudv}{\int_{-\infty}^{y_0} f_Y(y)\ dy} = \frac{F(x, y_0)}{F_Y(y_0)}. \end{aligned}$$

Definición 10.7.1. Llamaremos función de distribución de la variable aleatoria X condicionada a que $Y = y_0$ al límite

$$\lim_{h \to 0^+} P[X \leq x | y_0 - h < Y \leq y_0 + h]$$

en caso de que exista y sea un función de distribución en $\mathbb{R}$ se se notará $F_{X|y_0}(x)$ o bien, $F(x|Y = y_0)$.

Definición 10.7.2. Llamaremos función de densidad de la variable aleatoria X condicionada a que $Y = y_0$ y se representará por $f_{X|y_0}(x)$ a una función de $\mathbb{R}$ en $\mathbb{R}$ no negativa, de manera que para cada $y_0 \in \mathbb{R}$ fijo

$$\int_{\mathbb{R}} f_{X|y_0}(x)\ dx = 1$$

y tal que

$$F_{X|y_0}(x) = \int_{-\infty}^{x} f_{X|y_0}(u)\ du, \text{para todo } x \in \mathbb{R}.$$

Proposición 10.7.3. *En los puntos donde f sea continua y f_Y sea positiva y continua,, la función de densidad de la variable aleatoria X condicionada a que $Y = y$, existe y puede ser expresada como*

$$f_{X|y}(x) = \frac{f(x,y)}{f_Y(y)}.$$

10.8. Independencia de variables aleatorias continuas

Sabemos que dos sucesos A y B son *estocásticamente independientes* si se verifica que

$$P(A|B) = P(A) \quad \text{o} \quad P(B|A) = P(B),$$

o de manera equivalente, si

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

Definición 10.8.1. X es independiente de Y si y solo si para cualesquiera borelianos B_1 y B_2 los sucesos $X^{-1}(B_1)$ e $Y^{-1}(B_2)$ de $\mathcal{A}$ son independientes en la probabilidad P. Es decir:

$$P[\{\omega : X(\omega) \in B_1\} \cap \{\omega : Y(\omega) \in B_2\}] = P[\{\omega : X(\omega) \in B_1\}] \cdot P[\{\omega : Y(\omega) \in B_2\}].$$

Si consideramos $x, y \in \mathbb{R}$, $B_1 = \{\omega : X(\omega) \leq x\}$ y $B_2 = \{\omega : Y(\omega) \leq y\}$, entonces X e Y son independientes se tiene que

$$P[\{\omega : X(\omega) \leq x\} \cap \{\omega : Y(\omega) \leq y\}] = P[\{\omega : X(\omega) \leq x\}] \cdot P[\{\omega : Y(\omega) \leq y\}].$$

Si F es la función de distribución conjunta y F_X y F_Y las funciones de distribución marginales de X e Y entonces

$$F(x,y) = F_X(x)F_Y(y).$$

Definición 10.8.2. Diremos que dos variables aleatorias X e Y son independientes si para todo $x, y \in \mathbb{R}$

$$F(x,y) = F_X(x)F_Y(y),$$

siendo F la función de distribución conjunta y F_X y F_Y las funciones de distribución marginales de X e Y respectivamente.

Si F es derivable respecto de x y respecto de y y F_X y F_Y son derivables, tenemos que

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial F_x(x)}{\partial x} \cdot \frac{\partial F_Y(y)}{\partial y},$$

es decir,

$$f(x,y) = f_X(x) \cdot f_Y(y)$$

donde f, f_X y f_Y son, respectivamente, la función de densidad conjunta y las funciones de densidad marginales de X e Y.

10.9. Cambio de variable

Proposición 10.9.1. *Sea* $\boldsymbol{X}$ *un vector aleatorio bidimensional absolutamente continuo con densidad conjunta* $f(x,y)$ *definida sobre* $R \subset \mathbb{R}^2$*, abierto o cerrado no degenerado donde* $f > 0$*. Si* $g : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *es una función inyectiva y diferenciable en* R*, cuyo Jacobiano* J_g *no se anula en ningún punto de* R*, entonces* $(U,V) = g(X,Y)$ *tiene una distribución absolutamente continua en* $g(R)$ *con densidad*

$$f^*(u,v) = f(g^{-1}(u,v))|J_g(g^{-1}(u,v))|^{-1}.$$

Si $h = g^{-1}$, entonces

$$f^*(u,v) = f(h(u,v))|J_h(h(u,v))|.$$

10.10. Momentos conjuntos de un vector (X,Y)

Definición 10.10.1 (Momento ordinario de orden $r, s \in \mathbb{Z}^+$)**.**

$$\begin{aligned} m_{r,s} &= E[X^r \cdot Y^s] \\ &= \sum_{i\in I}\sum_{j\in J} x_i^r y_j^s P(X = x_i, Y = y_j) \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^r y^s f(x,y)\ dxdy \end{aligned}$$

Algunos momentos ordinarios destacados son

- $m_{0,0} = 1$.
- $m_{1,1} = E[X \cdot Y]$.
- $m_{r,0} = E[X^r]$, $m_{1,0} = E[X]$, $m_{2,0} = E[X^2]$.
- $m_{0,s} = E[Y^s]$, $m_{0,1} = E[Y]$, $m_{0,2} = E[Y^2]$.

Definición 10.10.2 (Momento central de orden $r, s \in \mathbb{Z}^+$)**.**

$$\begin{aligned} M_{r,s} &= E[(X - E(X))^r \cdot (Y - E(Y))^s] \\ &= \sum_{i\in I}\sum_{j\in J} (x_i - E(X))^r \cdot (y_j - E(Y))^s P(X = x_i, Y = y_j) \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (x - E(X))^r \cdot (y - E(Y))^s f(x,y)\ dxdy \end{aligned}$$

Algunos momentos centrales destacados son

- $M_{1,1} = E[(X - E(X)) \cdot (Y - E(Y))] = Cov(X,Y)$.
- $M_{1,0} = E[X - E(X)] = 0$ y $M_{0,1} = E[Y - E(Y)] = 0$.
- $M_{2,0} = V(X)$ y $M_{0,2} = V(Y)$.

Observación 10.10.3.

- $V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$.
- $V(X-Y) = V(X) + V(Y) - 2Cov(X,Y)$.

Si X e Y son independientes

- $m_{r,s} = m_{r,0} \cdot m_{0,s}$.
- $M_{r,s} = M_{r,0} \cdot M_{0,s}$.
- $Cov(X,Y) = 0$.
- $V(X+Y) = V(X) + V(Y) = V(X-Y)$.

10.11. Distribución del máximo y del mínimo

Sea $\boldsymbol{X} = (X_1, ..., X_n)$ un vector aleatorio con función de distribución $F_{\boldsymbol{X}}$ y sean las variables aleatorias $M = \text{máx}(X_1, ..., X_n)$ y $N = \text{mín}(X_1, ..., X_n)$ definidas de la forma

$$M(\omega) = \text{máx}(X_1(\omega), ..., X_n(\omega))$$
$$N(\omega) = \text{mín}(X_1(\omega), ..., X_n(\omega))$$

Tanto M como N son variables aleatorias y

$$F_M(x) = P[M \leq x] = P[X_1 \leq x, ..., X_n \leq x] = F_X(x, ..., x)$$
$$F_N(x) = P[N \leq x] = 1 - P[N > x] = 1 - P[X_1 > x, ..., X_n > x].$$

Definición 10.11.1. Sea (X, Y) un vector aleatorio, se definen las curvas generales de regresión como las curvas

- $C_{x;y} = \{(E(X|y = y_0), y) : y_0 \in \mathbb{R}\}$.
- $C_{y;x} = \{(x_0, E(Y|x = x_0)) : x_0 \in \mathbb{R}\}$.