Tema 5. VARIABLES ALEATORIAS: CASOS DISCRETO Y CONTINUO

En este tema se va a introducir el concepto, fundamental para la Estadística, de **Variable Aleatoria** y, en particular, se van a estudiar las variables aleatorias discretas y continuas.

Los primeros pasos en la dirección de introducir la idea de **variable aleatoria** fueron dados por Poisson, en el Siglo XIX, en su libro "Sobre la Probabilidad de los Resultados Promedios de Observaciones". En él Poisson, no utilizó el término variable aleatoria, pero sí habló de "alguna cosa" que uno puede entender como un conjunto de valores con sus correspondientes probabilidades.

La palabra **variable** fue utilizada por primera vez por Chebyshev que asumió implícitamente que todas las variables aleatorias eran independientes (lo cual era un error).

1. Definición. Distribución de probabilidad. Función de distribución.

Definición de Variable Aleatoria

Sea $(\Omega, \mathcal{A}, P)$ un espacio probabilístico donde Ω es el espacio muestral. Una **variable aleatoria** $X : \Omega \to \mathbb{R}$ es una función medible, es decir, es una función que satisface:

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

siendo $X^{-1}(B) = \{\omega \in \Omega / X(\omega) \in B\} \equiv \{X \in B\}$.
Una variable aleatoria asigna a cada elemento del espacio muestral un número real.

$$\begin{aligned} X\colon \quad \Omega &\longrightarrow \mathbb{R} \\ \omega &\longrightarrow X(\omega) \end{aligned}$$

[1] $\sigma-$álgebra de Borel en $\mathbb{R}$: Es la menor clase de subconjuntos de $\mathbb{R}$ con estructura de $\sigma-$álgebra que contiene a todos los intervalos.

Ya que $\mathcal{B} = \sigma(\{(-\infty, x], x \in \mathbb{R}\})$, es decir, la $\sigma-$álgebra de Borel puede ser generada por los intervalos de la forma $(-\infty, x]$, una definición equivalente de variable aleatoria sería la siguiente:

$X : \Omega \to \mathbb{R}$ es una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$ si:

$$X^{-1}((-\infty, x]) \in \mathcal{A}, \ \forall x \in \mathbb{R}$$

siendo $X^{-1}((-\infty, x]) = \{\omega \in \Omega / X(\omega) \leq x\} \equiv \{X \leq x\}$.

Ejemplo

1. *$I_A : \Omega \to \mathbb{R}$ con $A \in \mathcal{A}$ verificando $I_A(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \omega \notin A \end{cases}$ es una variable aleatoria.*
2. *Determina la menor $\sigma-$álgebra que se debe considerar para que la función que contabiliza el número de caras obtenidas al lanzar dos veces una moneda sea una variable aleatoria.*

Operaciones con variables aleatorias:

Sean X e Y dos variables aleatorias sobre $(\Omega, \mathcal{A}, P)$. Entonces se verifica que :

- $\forall a, b \in \mathbb{R}$ $aX + b$ es una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$.

$$(aX + b)(\omega) = aX(\omega) + b, \ \forall \omega \in \Omega.$$

- $\forall a \in \mathbb{R}$ X^a (si está bien definida) es una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$.

$$X^a(\omega) = (X(\omega))^a, \ \forall \omega \in \Omega.$$

- $X + Y$ y $X - Y$ son variables aleatorias sobre $(\Omega, \mathcal{A}, P)$.

$$(X + Y)(\omega) = X(\omega) + Y(\omega), \ \forall \omega \in \Omega.$$

$$(X - Y)(\omega) = X(\omega) - Y(\omega), \ \forall \omega \in \Omega.$$

- XY y X/Y (si está bien definida) son variables aleatorias sobre $(\Omega, \mathcal{A}, P)$.

$$(XY)(\omega) = X(\omega)Y(\omega), \;\; \forall \omega \in \Omega.$$

$$(X/Y)(\omega) = X(\omega)/Y(\omega), \;\; \forall \omega \in \Omega.$$

- $\text{máx}(X, Y)$ y $\text{mín}(X, Y)$ son variables aleatorias sobre $(\Omega, \mathcal{A}, P)$.

$$\text{máx}(X, Y)(\omega) = \text{máx}(X(\omega), Y(\omega)), \;\; \forall \omega \in \Omega.$$

$$\text{mín}(X, Y)(\omega) = \text{mín}(X(\omega), Y(\omega)), \;\; \forall \omega \in \Omega.$$

- $|X|$ es una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$.

$$|X|(\omega) = |X(\omega)|, \;\; \forall \omega \in \Omega.$$

Distribución de probabilidad

Si X es una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$ su distribución de probabilidad es una función, $P_X : \mathcal{B} \to [0,1]$, definida como

$$P_X(B) = P(X \in B) = P(X^{-1}(B)) = P(\{\omega \in \Omega / \ X(\omega) \in B\}), \ \forall B \in \mathcal{B}$$

Proposición: P_X es una función de probabilidad sobre $(\mathbb{R}, \mathcal{B})$, es decir cumple los tres axiomas de Kolmogorov:

1. Si $B \in \mathcal{B}$ entonces $P_X(B) \geq 0$.
2. $P_X(\mathbb{R}) = 1$.
3. Si $\{B_n\}_{n \in \mathbb{N}} \in \mathcal{B}$, $(B_i \cap B_j = \emptyset, \forall i \neq j) \Rightarrow$

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

Por lo tanto, P_X es la probabilidad inducida por la variable aletoria en $\mathbb{R}$.

$$X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B}, P_X)$$

Ejemplo

Determina la distribución de probabilidad que induce la variable de nuestro ejemplo 2.

Función de distribución

La función de distribución de probabilidad es una función que asigna a cada valor $x \in \mathbb{R}$ la probabilidad inducida por la v.a. X asociada al intervalo $(-\infty, x]$.

Definición

Si X es una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$, se define su **función de distribución** *como una función*

$$\begin{aligned} F_X : \quad & \mathbb{R} \longrightarrow [0,1] \\ & x \longmapsto F_X(x) = P(X \in (-\infty, x]) = P(X \leq x). \end{aligned}$$

En particular, $F_X(x) = P(X \in (-\infty, x]) = P_X((-\infty, x])\ \forall x \in \mathbb{R}$.

Propiedades:

1. $\exists \lim_{x \to -\infty} F_X(x) \equiv F_X(-\infty) = 0$ y $\exists \lim_{x \to +\infty} F_X(x) \equiv F_X(+\infty) = 1.$
2. Es una función monótona no decreciente:
$$x_i < x_j \Rightarrow F_X(x_i) \leq F_X(x_j).$$
3. Es continua a la derecha:
$$\forall x_0 \in \mathbb{R},\ \exists \lim_{x \to x_0^+} F_X(x) \equiv F_X(x_0^+) = F_X(x_0).$$

Teorema (Caracterización de función de distribución de v.a.)

Toda función de distribución asociada a una v.a. verifica estas propiedades y, recíprocamente, toda función real en el intervalo $[0,1]$ *que verifique estas propiedades es la función de distribución de una variable aleatoria.*

Ejemplo

Construye la función de distribución asociada a la variable aleatoria del nuestro ejemplo 2.

Otras propiedades:

- Las posibles discontinuidades de F_X son de salto.
- $\exists \lim_{x \to x_0^-} F_X(x) \equiv F_X(x_0^-) = P(X < x_0), \ \forall x_0 \in \mathbb{R}$.
- $F_X(x_0) - F_X(x_0^-) = P(X = x_0)$.

A partir de estas propiedades se deduce que: x_0 es un punto de salto de F_X si y solo si $P(X = x_0) > 0$, y la longitud del salto es $P(X = x_0)$. Además $\forall a < b \in \mathbb{R}$:

- $P(X \leq a) = F_X(a)$.
- $P(X < a) = F_X(a^-)$.
- $P(X = a) = F_X(a) - F_X(a^-)$.
- $P(X > a) = 1 - P(X \leq a) = 1 - F_X(a)$.
- $P(X \geq a) = 1 - P(X < a) = 1 - F_X(a^-)$.
- $P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a)$.
- $P(a \leq X \leq b) = P(X \leq b) - P(X < a) = F_X(b) - F_X(a^-)$.
- $P(a < X < b) = P(X < b) - P(X \leq a) = F_X(b^-) - F_X(a)$.
- $P(a \leq X < b) = P(X < b) - P(X < a) = F_X(b^-) - F_X(a^-)$.

La variables aleatorias se clasifican en v.a. discretas o continuas:

1. Una v.a. X es **discreta** si el conjunto de valores que puede tomar la variable es contable, es decir, finito o infinito numerable.

   Ejemplo: Número de asignaturas de una matrícula universitaria, número de aprobados en un curso, número de mensajes recibidos al día, ...

2. Una v.a. X es **continua** si el conjunto de valores que puede tomar la variable es un subconjunto de los números reales, es decir, infinito no numerable. En general sus valores consisten en uno o más intervalos de la recta de números reales.

   Ejemplo: Altura de una persona, porcentaje de aprobados en una materia, temperatura ambiental, ...

2. Variables aleatorias discretas y continuas

Variables aleatorias discretas Una v.a. $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B}, P_X)$ es discreta si existe un conjunto numerable $E_X = \{x_1, x_2, ..., x_k, ...\} \subset \mathbb{R}$ tal que $P(X \in E_X) = 1$.

Definición

Se define la **función masa de probabilidad** *(f.m.p.) de la v.a. discreta* X *como una función que asigna probabilidades a los valores de la variable, es decir,*

$$\begin{aligned} p_X : \quad & E_X \longrightarrow [0,1] \\ & x_i \longmapsto p_X(x_i) = P(X = x_i), \quad i = 1, 2, \ldots, k, ... \end{aligned}$$

verificando las siguientes propiedades:

- *Es no negativa:* $P(X = x_i) \geq 0, \quad \forall i = 1, \ldots, k, ...$
- *La suma de sus valores es uno:* $\displaystyle\sum_{x_i \in E_X} P(X = x_i) = 1.$

Caracterización de f.m.p.: Toda función definida sobre un subconjunto numerable de $\mathbb{R}$, no negativa y tal que la suma de sus valores es uno, es la f.m.p. de alguna v.a. discreta con valores en dicho conjunto.

La **distribución de probabilidad** de una v.a. discreta X es:

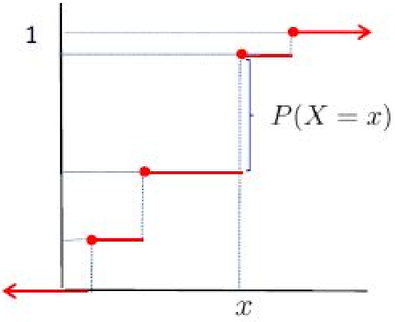
$$P_X(B) = P(X \in B) = \sum_{x_i \in B} P(X = x_i), \ \forall B \in \mathcal{B}.$$

La **función de distribución** de una v.a. discreta X es:

$$F_X(x) = P(X \leq x) = \sum_{x_i \leq x} P(X = x_i), \ \forall x \in \mathbb{R}.$$

Evidentemente, la función de distribución de una variable aleatoria discreta sigue cumpliendo las propiedades estudiadas anteriormente.

Su representación gráfica adopta una forma escalonada ya que crece únicamente a saltos, los cuales ocurren en los valores aislados que toma la variable, siendo continua por la derecha en cada uno de éstos y la longitud del salto corresponde con la probabilidad de que la variable tome dicho valor. Es más, una variable aleatoria es discreta si y solo si su función de distribución cumple estas características.



Ejemplo

Sea ahora X el número de caras al lanzar la moneda tres veces. Obtener la f.m.p. y la función de distribución asociada a la v.a. y calcular:

- $P(X < 3)$,
- $P(X \leq 3)$,
- $P(X > 1)$,
- $P(X \geq 1)$,
- $P(1 \leq X \leq 2)$,
- $P(1 \leq X < 2)$,
- $P(1 < X \leq 2)$,
- $P(1 < X < 2)$.

Variable aleatoria continua

Una v.a. $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B}, P_X)$ es continua si existe una función integrable $f_X : \mathbb{R} \to \mathbb{R}$, tal que:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt, \ \forall x \in \mathbb{R}.$$

Propiedades de f_X:

- $\lim\limits_{x \to -\infty} f_X(x) = \lim\limits_{x \to +\infty} f_X(x) = 0.$
- f_X es continua salvo, a lo sumo, en un conjunto numerable. En particular, f_X es continua en los puntos donde F_X es derivable y, en tales puntos, $f_X(x) = F'_X(x)$.
- Los valores de f_X pueden cambiarse en un conjunto numerable de puntos sin afectar a F_X.

Definición

Se define la **función de densidad** *(f.d.d.) de la v.a. continua X como la función $f_X(x) = F'_X(x)$ (donde F_X sea derivable) y satisface las siguientes condiciones:*

1. *Es no negativa: $f_X(x) \geq 0$.*
2. $\displaystyle\int_{-\infty}^{+\infty} f_X(x)dx = 1.$

Caracterización de f.d.d.:

Toda función $f : \mathbb{R} \to \mathbb{R}$ no negativa, integrable y tal que $\displaystyle\int_{-\infty}^{+\infty} f_X(x)dx = 1$, es la función de densidad de alguna variable aleatoria continua.

La **distribución de probabilidad** de una v.a. continua X es:

$$P_X(B) = P(X \in B) = \int_B f_X(x)\, dx, \ \forall B \in \mathcal{B}.$$

Así, si X es una v.a. continua, para calcular la probabilidad de que X tome valores en un determinado intervalo basta con integrar la función de densidad en dicho intervalo: $\forall a, b \in \mathbb{R},\ a < b$.

- $P_X((a,b)) = P_X((a,b]) = P_X([a,b)) = P_X([a,b]) = \int_a^b f(x)dx$.
- $P_X((-\infty, b]) = P_X((-\infty, b)) = \int_{-\infty}^b f(x)dx$.
- $P_X([a, +\infty)) = P_X((a, +\infty)) = \int_a^{+\infty} f(x)dx$.

Conviene resaltar que el valor de la función de densidad, f_X, en un punto x, no es la probabilidad de aparición de ese valor x, ya que:

$$P_X(\{x\}) = P(X \in \{x\}) = P(X = x) = \int_x^x f(t)dt = 0 \ \forall x \in \mathbb{R}.$$

Es más, la probabilidad de que una v.a. continua tome valores en un conjunto numerable sigue siendo nula:

$$\forall B \in \mathcal{B} / \ B = \{x_1, ..., x_k, ...\} \subset \mathbb{R} \Rightarrow P_X(B) = \sum_{x_i \in B} P(X = x_i) = 0.$$

La **función de distribución** de una v.a. continua X es una función continua en toda la recta real y se obtiene de la siguiente forma:

$$F_X(x) = P[X \leq x] = \int_{-\infty}^{x} f(t)dt \quad \forall x \in \mathbb{R},$$

es decir, es el área que encierra la función de densidad en el intervalo $(-\infty, x]$.

De nuevo la función de distribución así definida sigue cumpliendo las propiedades estudiadas anteriormente.

A partir de la función de distribución se pueden determinar de forma sencilla la probabilidad de que la variable tome valores en un determinado intervalo: $\forall a, b \in \mathbb{R},\ a < b$:

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

Ejemplo

Sea X una v.a. con función de densidad

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

Calcular su función de distribución y $P[0.2 < X < 0.7]$.

3. Funciones de una variable aleatoria.

A veces es necesario determinar la distribución de probabilidad de una variable aleatoria que es función de otra, en particular vamos a considerar funciones medible.

Sea X una v.a. definida sobre el espacio probabilístico $(\Omega, \mathcal{A}, P)$, es decir una función medible $X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B})$, tal que $P(X \in E_X) = 1$ y sea $g : E_X \to \mathbb{R}$ otra función medible. $Y = g(X)$ es la composición de dos funciones medibles y, por tanto, una función medible definida sobre un espacio de probabilidad, es decir una variable aleatoria verificando $P(Y \in g(E_X)) = 1$.

$$Y : \Omega \to_X E_X \to_g g(E_X)$$

Teorema (General del cambio de variable)

Si X es una variable aleatoria y g una función medible definida sobre E_X, la distribución de probabilidad de la variable $Y = g(X)$ está dada por

$$P(Y \in B) = P(g(X) \in B) = \begin{cases} \displaystyle\sum_{\{x/\ g(x) \in B\}} P(X = x), & X \text{ discreta} \\ \displaystyle\int_{\{x/\ g(x) \in B\}} f_X(x)\, dx, & X \text{ continua}, \end{cases}$$

$\forall B \in \mathcal{B}$.

Análogamente, la función de distribución de Y está dada por

$$F_Y(y) = P(Y \le y) = P(g(X) \le y) = \begin{cases} \displaystyle\sum_{\{x/\ g(x) \le y\}} P(X = x), & X \text{ discreta} \\ \displaystyle\int_{\{x/\ g(x) \le y\}} f_X(x)\, dx, & X \text{ continua}, \end{cases}$$

$\forall y \in \mathbb{R}$.

Ejemplo

Sea X *una v.a. con función de distribución* F_X*. Calcular la función de distribución de cada una de las siguientes variables en términos de* F_X*.*

1. $Y = |X|$.
2. $Y = \max(0, X)$.

Teorema (Cambio de variable discreto)

Si X *es una v.a. discreta y* $g : E_X \to \mathbb{R}$ *es una función medible arbitraria,* $Y = g(X)$ *es una v.a. discreta cuya función masa de probabilidad se obtiene a partir de la de* X *de la siguiente forma:*

$$P(Y = y) = \sum_{\{x \in E_X /\ g(x)=y\}} P(X = x)\ \forall y \in g(E_X).$$

Ejemplo

Sea X *una v.a. discreta con f.m.p.:* $P(X = -2) = \frac{1}{5}$, $P(X = -1) = \frac{1}{6}$, $P(X = 0) = \frac{1}{5}$, $P(X = 1) = \frac{1}{15}$, $P(X = 2) = \frac{11}{30}$. *Calcula la f.m.p. de las siguientes v.a. discretas:*

$$Y = X + 2; \qquad Y = X^2 + 3.$$

Teorema (Cambio de variable de continua a discreta)

Si X es una v.a. continua, con función de densidad f_X, y $g : E_X \to \mathbb{R}$ es una función medible tal que $Y = g(X)$ es una v.a. discreta, su función masa de probabilidad se obtiene a partir de f_X de la siguiente forma:

$$P(Y = y) = \int_{\{x/ \ \ g(x)=y\}} f_X(x)\, dx \ \forall y \in g(E_X).$$

Ejemplo

Sea X una v.a. continua con f.d.d.:

$$f_X(x) = \frac{1}{2}, \quad -1 < x < 1.$$

Calcula la f.m.p. de la siguiente v.a. discreta: $Y = \begin{cases} -1 & X < 0 \\ 0 & X = 0 \\ 1 & X > 0 \end{cases}.$

Teorema (Cambio de variable de continua a continua)

Si X es una v.a. continua, con función de densidad f_X, y $g : E_X \to \mathbb{R}$ es una función medible, derivable y estríctamente monótona, $Y = g(X)$ es una v.a. continua, cuya función de densidad se obtiene a partir de f_X de la siguiente forma:

$$f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{dg^{-1}(y)}{dy}\right|, \forall y \in g(E_X).$$

Ejemplo

Sea X una v.a. continua con f.d.d.:

$$f_X(x) = \lambda e^{-\lambda x}, \quad x > 0.$$

Calcula la f.d.d. de la siguiente v.a. continua: $Y = X^2$.

Teorema (Extensión del cambio de variable de continua a continua)

Si cada valor de $Y = g(X)$ *procede de un conjunto numerable de valores de* X, $\{g_n^{-1}(y)\}_{n\in\mathbb{N}}$, *y cada una de las funciones* g_n^{-1} *es derivable y estrictamente monótona,* Y *es también de tipo continuo y su función de densidad se obtiene a partir de* f_X *de la siguiente forma:*

$$f_Y(y) = \sum_{n=1}^{+\infty} f_X\left(g_n^{-1}(y)\right)\left|\frac{dg_n^{-1}(y)}{dy}\right|, \forall y \in g(E_X).$$

Ejemplo

Sea X *una v.a. continua con f.d.d.:*

$$f_X(x) = \lambda e^{-\lambda x}, \quad x > 0.$$

Calcula la f.d.d. de la siguiente v.a. continua: $Y = (X-1)^2$.