Tema 5
Variables aleatorias: distribuciones de probabilidad y características.

1. Introducción

Según se ha reflejado hasta el momento, el espacio muestral asociado a un experimento aleatorio puede ser de dos tipos:

- Cuantitativo, como el asociado al lanzamiento de un dado ($\Omega = \{1, 2, \dots, 6\}$).
- Cualitativo, como en los siguientes ejemplos:
  - Lanzamiento de dos monedas ($\Omega = \{cc, cx, xc, xx\}$).
  - Extracción de bolas de una urna.
  - Elección de un individuo de la población.

Es evidente que el tratamiento matemático de un espacio muestral de tipo cualitativo no es simple. Sin embargo, cuando un experimento da lugar a un espacio muestral de tipo cualitativo, es posible considerar una o varias características numéricas que describan las propiedades de mayor interés. Por ejemplo:

- En el lanzamiento de tres monedas: número de caras o de cruces, diferencia entre el número de caras y de cruces, etc.
- En la extracción de bolas de una urna: número de bolas de un determinado color, etc.
- En la elección de un individuo: estatura, peso, etc.

Así, cada resultado del experimento tendrá asociado un valor numérico y el espacio muestral original se transforma en un espacio cuantitativo.

Incluso en espacios muestrales cuantitativos, puede que el interés se centre no en el resultado concreto del experimento, sino en alguna característica numérica como, por ejemplo, en el lanzamiento de dos dados, la suma de los valores obtenidos.

De esta forma surge el concepto de **variable aleatoria** que, en términos generales, puede definirse como una función que asigna un valor real a cada elemento de un espacio muestral.

Al considerar una variable aleatoria sobre un espacio muestral, los conjuntos de interés estarán definidos en términos de dicha variable. Por ejemplo, conjunto de resultados elementales, tales que el valor de la variable esté comprendido entre dos números reales a y b. Para poder calcular la probabilidad de conjuntos de este tipo, es preciso exigir que tal conjunto sea un suceso. Este requerimiento implica que no toda función numérica de los resultados de un experimento es una variable aleatoria, sino que ésta debe satisfacer determinadas propiedades.

Para introducir de manera formal el concepto de variable aleatoria, comenzaremos definiendo la *σ-álgebra de Borel sobre* $\mathbb{R}$.

2. Espacio de Borel unidimensional

Sobre el conjunto de números reales $\mathbb{R}$, se define la σ-álgebra de Borel como la mínima clase de conjuntos con estructura de σ-álgebra que contiene a todos los intervalos de $\mathbb{R}$. Esto es, si $\mathcal{Y}$ denota la clase de intervalos de $\mathbb{R}$, la σ-álgebra de Borel, $\mathcal{B}$, es una clase de conjuntos de $\mathbb{R}$, ($\mathcal{B} \subset \mathcal{P}(\mathbb{R})$), tal que:

1. $\mathcal{B} \supset \mathcal{Y}$.
2. $\mathcal{B}$ es σ-álgebra.
3. Si $\mathcal{A} \subset \mathcal{P}(\mathbb{R})$ es una σ-álgebra, tal que $\mathcal{A} \supset \mathcal{Y}$, entonces $\mathcal{A} \supset \mathcal{B}$.

Al par $(\mathbb{R}, \mathcal{B})$ se le denomina *Espacio de Borel*. Los elementos de $\mathcal{B}$ se denominan *Conjuntos de Borel o Borelianos*.

- Todo intervalo y, en particular, todo número real ($\{a\} = [a, a]$), es un conjunto de Borel.
- Todo conjunto numerable y, en particular, todo conjunto finito, es un conjunto de Borel.
- Todo conjunto formado a partir de las operaciones de uniones numerables, intersecciones numerables y complementarios, realizadas a partir de intervalos de $\mathbb{R}$, es un conjunto de Borel.

Teorema: Caracterización de $\mathcal{B}$

$\mathcal{B}$ coincide con la σ-álgebra generada por los intervalos del tipo $(-\infty, x]$.
Análogamente, $\mathcal{B}$ es la σ-álgebra generada por intervalos de cualquier tipo.

3. Variables aleatorias

El concepto de variable aleatoria surge de la necesidad de calcular probabilidades de conjuntos de interés definidos en términos de dicha variable.

Así, si $(\Omega, \mathcal{A}, \mathcal{P})$ es el espacio de probabilidad asociado al experimento aleatorio en el que se pretende analizar la característica numérica de interés, ésta vendrá definida por una función

$$X : \Omega \longrightarrow \mathbb{R}.$$

Ahora bien, cada valor de X se corresponde con el subconjunto de puntos de Ω que se aplica en dicho valor esto es $\{\omega \in \Omega \ / \ X(\omega) = x\}$, que notaremos por simplicidad $\{X = x\}$. Obviamente el estudio probabilísticos de una variable aletoria conlleva el cálculo de probabilidades de dichos conjuntos así como de otros más generales como
$\{\omega \in \Omega \ / \ X(\omega) \leq x\} = \{X \leq x\}$,
$\{\omega \in \Omega \ / \ X(\omega) < x\} = \{X < x\}$,

$\{\omega \in \Omega \ / \ X(\omega) \geq x\} = \{X \geq x\}$,
$\{\omega \in \Omega \ / \ X(\omega) > x\} = \{X > x\}$,
$\{\omega \in \Omega \ / \ x_1 \leq X(\omega) \leq x_2\} = \{x_1 \leq X \leq x_2\}$,
$\ldots$
$\{\omega \in \Omega \ / \ X(\omega) \in I\} = \{X \in I\}, \ I \in \mathcal{Y}$,
$\{\omega \in \Omega \ / \ X(\omega) \in B\} = \{X \in B\}, \ B \in \mathcal{B}$.

Para poder calcular la probabilidad de dichos conjuntos es necesario que los mismos sean sucesos, esto es pertenezcan a la $\sigma-$álgebra de Borel del espacio probabilístico donde se define la función que describe la característica numérica de interés.

Así, la definición formal de variable aleatoria es

Definición

Una variable aleatoria definida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, P)$ es una función $X : \Omega \to \mathbb{R}$ que verifica $X^{-1}(\mathcal{B}) \subset \mathcal{A}$, es decir

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

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

La medibilidad proporciona el aspecto analítico de las variables aleatorias, mientras que el hecho de estar definidas sobre un espacio de probabilidad les da el carácter probabilístico.

La definición de variable aleatoria no es siempre operativa, por lo que, a la hora de probar que una función X sobre un espacio de probabilidad es una variable aleatoria, resulta conveniente hacer uso de la siguiente caracterización.

Teorema: Caracterización de variables aleatorias

$X : (\Omega, \mathcal{A}, P) \to (\mathbb{R}, \mathcal{B})$ es una variable aleatoria si y sólo si se cumple alguna de las siguientes condiciones, todas ellas equivalentes:

1. $X^{-1}((-\infty, x]) = \{\omega / X(\omega) \leq x\} = \{X \leq x\} \in \mathcal{A}, \ \forall x \in \mathbb{R}$.
2. $X^{-1}((-\infty, x)) = \{\omega / X(\omega) < x\} = \{X < x\} \in \mathcal{A}, \ \forall x \in \mathbb{R}$.
3. $X^{-1}([x, +\infty)) = \{\omega / X(\omega) \geq x\} = \{X \geq x\} \in \mathcal{A}, \ \forall x \in \mathbb{R}$.
4. $X^{-1}((x, +\infty)) = \{\omega / X(\omega) > x\} = \{X > x\} \in \mathcal{A}, \ \forall x \in \mathbb{R}$.
5. $X^{-1}((a, b]) = \{\omega / a < X(\omega) \leq b\} = \{a < X \leq b\} \in \mathcal{A}, \ \forall a, b \in \mathbb{R}$.
6. $X^{-1}([a, b)) = \{\omega / a \leq X(\omega) < b\} = \{a \leq X < b\} \in \mathcal{A}, \ \forall a, b \in \mathbb{R}$.
7. $X^{-1}((a, b)) = \{\omega / a < X(\omega) < b\} = \{a < X < b\} \in \mathcal{A}, \ \forall a, b \in \mathbb{R}$.
8. $X^{-1}([a, b]) = \{\omega / a \leq X(\omega) \leq b\} = \{a \leq X \leq b\} \in \mathcal{A}, \ \forall a, b \in \mathbb{R}$.

Ejemplos de variables aleatorias

Si se trabaja en un espacio probabilístico $(\Omega, \mathcal{P}(\Omega), P)$, en el que todo subconjunto de Ω es un suceso, entonces cualquier función $X : \Omega \to \mathbb{R}$ es una variable aleatoria. Esto es lo usual si el espacio muestral Ω es finito.

Ejemplo 1: Funciones indicadoras de conjuntos medibles

Una función de la forma $I_A : \Omega \to \mathbb{R}$, definida como

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

donde $A \subseteq \Omega$, se denomina **función indicadora del conjunto** A.

Si $(\Omega, \mathcal{A}, P)$ es un espacio de probabilidad, y $A \in \mathcal{A}$, entonces I_A es una variable aleatoria. En efecto:

$$I_A^{-1}(B) = \begin{cases} \emptyset & 1 \notin B, 0 \notin B \\ A & 1 \in B, 0 \notin B \\ A^C & 1 \notin B, 0 \in B \\ \Omega & 1 \in B, 0 \in B \end{cases} \Longrightarrow I_A^{-1}(B) \in \mathcal{A} \ \ \forall B \in \mathcal{B}.$$

Ejemplo 2: Funciones simples

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad, y $\{A_i\}_{i=1}^n$ una partición de Ω, tal que $A_i \in \mathcal{A}$ (partición medible). Sean $x_1, \ldots, x_n \in \mathbb{R}$.

La función $X : \Omega \to \mathbb{R}$ definida como $X(\omega) = x_i$ si $w \in A_i$ se denomina **función simple**, y es una variable aleatoria, pues

$$X^{-1}(B) = \bigcup_{i=1/x_i \in B}^{n} A_i \in \mathcal{A}.$$

O sea, toda función definida en un espacio de probabilidad que tome un número finito de valores, cada uno sobre un conjunto medible, es una variable aleatoria. Esto es obviamente extensible a particiones no finitas numerables, ya que, en tal caso, la anti-imagen de B sería una unión, no necesariamente finita, pero numerable.

Ejemplo 3: En el lanzamiento de un dado, se asigna el valor 1 a los resultados pares, y el resultado 0 a los impares.
El espacio de probabilidad de la variable aleatoria es $(\Omega, \mathcal{A}, P)$, donde

$$\Omega = \{1, 2, 3, 4, 5, 6\}; \quad \mathcal{A} = \mathcal{P}(\Omega); \quad P \equiv \text{uniforme}.$$

Al considerar $\mathcal{A} = \mathcal{P}(\Omega)$, cualquier función real definida en $(\Omega, \mathcal{A}, P)$ sería una variable aleatoria.
Veamos cómo se comprobaría que X es una variable aleatoria en el caso de considerar otra σ-álgebra:
Por una parte

$$X : \Omega \to \mathbb{R} \qquad X(\omega) = \begin{cases} 1 & \omega \in \{2,4,6\} \\ 0 & \omega \in \{1,3,5\} \end{cases}$$

$X = I_{\{2,4,6\}} \Longrightarrow X$ es variable aleatoria si $\{2,4,6\} \in \mathcal{A}$. Para ello bastaría considerar una σ-álgebra que contenga a dicho conjunto. Sin embargo, si se considera por ejemplo la σ-álgebra $\mathcal{A} = \{\Omega, \emptyset, \{1\}, \{2,3,4,5,6\}\}$, entonces X no sería una variable aleatoria.

También se podría haber razonado usando la definición de variable aleatoria. Así

$$X^{-1}(B) = \begin{cases} \emptyset & 0,1 \notin B \\ \{1,3,5\} & 0 \in B, 1 \notin B \\ \{2,4,6\} & 0 \notin B, 1 \in B \\ \Omega & 0,1 \in B \end{cases}$$

y de nuevo se llega a la conclusión de que basta exigir que $\{2,4,6\} \in \mathcal{A}$ (o, equivalentemente $\{1,3,5\} \in \mathcal{A}$).
Veamos cómo, aún en este caso tan simple, se puede simplificar el razonamiento usando la caracterización de variable aleatoria. Así

$$X^{-1}((-\infty, x]) = \begin{cases} \emptyset & x < 0 \\ \{1,3,5\} & 0 \leq x < 1 \\ \Omega & x \geq 1 \end{cases}$$

llegando a la misma conclusión.

Ejemplo 4:
Dos jugadores, A y B, lanzan cada uno un dado. Describir el espacio de probabilidad asociado a este experimento y cada una de las siguientes variables:

1. Puntuación obtenida por el jugador A.

2. Mayor puntuación.

3. Ganancia del jugador A si se conviene que el que obtenga mayor puntuación recibe una moneda del contrario.

4. Ganancia del jugador A si se conviene que el que obtenga menor puntuación paga al contrario la diferencia de puntuación.

Notamos cada resultado elemental por (i,j), donde i es la puntuación de A y j la de B.

$$\Omega = \{(i,j)/i,j = 1,\ldots,6\}; \quad \mathcal{A} = \mathcal{P}(\Omega); \quad P((i,j)) = 1/36$$

1. $X(i,j) = i; \quad X : 1,2,3,4,5,6.$

2. $Y(i,j) = \begin{cases} i & i > j \\ j & i \leq j \end{cases} \qquad Y : 1,2,3,4,5,6.$

3. $Z(i,j) = \begin{cases} 1 & i > j \\ 0 & i = j \\ -1 & i < j \end{cases}$

4. $U(i,j) = i - j \quad U = -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5.$

Ejemplo 5:
En el ejemplo anterior, especificar los sucesos:

- La puntuación obtenida por A es 2: $\{X = 2\} = \{(2,1), (2,2), (2,3), (2,4), (2,5), (2,6)\}$.
- La mayor puntuación es ≤ 2: $\{Y \leq 2\} = \{(1,1), (1,2), (2,1), (2,2)\}$.
- Se se ha convenido que el que obtenga menor puntuación paga al contrario la diferencia:
  - A gana al menos 4 monedas: $\{U \geq 4\} = \{(5,1), (6,1), (6,2)\}$.
  - A pierde más de 2 monedas: $\{U < -2\} = \{(1,4), (1,5), (1,6), (2,5), (2,6), (3,6)\}$.
  - No pierde ni A ni B: $\{U = 0\} = \{(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)\}$.

4. Operaciones algebraicas con variables aleatorias

La clase de variables aleatorias definidas sobre un mismo espacio de probabilidad es cerrada para las operaciones algebraicas usuales.
En efecto:

- **Teorema 1:** Si X es una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$, y $a, b \in \mathbb{R}$, entonces la aplicación
$$aX + b : \Omega \to \mathbb{R}, \text{ dada por } (aX + b)(\omega) = a(X(\omega)) + b$$
es una variable aleatoria sobre el mismo espacio.

- **Teorema 2:** Si X e Y son variables aleatorias sobre $(\Omega, \mathcal{A}, P)$, y $a, b \in \mathbb{R}$, entonces la aplicación
$$X + Y : \Omega \to \mathbb{R}, \text{ dada por } (X + Y)(\omega) = X(\omega) + Y(\omega)$$
es una variable aleatoria sobre el mismo espacio.

  - **Corolario 1:** Si $X_1, X_2, \ldots, X_n$ son variables aleatorias definidas sobre $(\Omega, \mathcal{A}, P)$, entonces $X_1 + X_2 + \cdots + X_n$ es una variable aleatoria definida sobre $(\Omega, \mathcal{A}, P)$.
  - **Corolario 2:** Cualquier combinación lineal de variables aleatorias sobre $(\Omega, \mathcal{A}, P)$ es una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$.
  - **Corolario 3:** Si X e Y son variables aleatorias sobre $(\Omega, \mathcal{A}, P)$, entonces $\{\omega \ / \ X(\omega) \leq Y(\omega)\} \in \mathcal{A}$ y $\{\omega \ / \ X(\omega) = Y(\omega)\} \in \mathcal{A}$.

- **Teorema 3:** Si X es una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$, y $a, b \in \mathbb{R}$, entonces la aplicación
$$X^2 : \Omega \to \mathbb{R}, \text{ dada por } (X^2)(\omega) = (X(\omega))^2$$
es una variable aleatoria sobre el mismo espacio.

- **Teorema 4:** Si X e Y son variables aleatorias sobre $(\Omega, \mathcal{A}, P)$, entonces la aplicación
$$X \cdot Y : \Omega \to \mathbb{R}, \text{ dada por } (X \cdot Y)(\omega) = X(\omega) \cdot Y(\omega)$$
es una variable aleatoria sobre el mismo espacio.

- **Teorema 5:** Si X e Y son variables aleatorias sobre $(\Omega, \mathcal{A}, P)$, e $Y \neq 0$, entonces la aplicación
$$X \cdot Y : \Omega \to \mathbb{R}, \text{ dada por } \left(\frac{X}{Y}\right)(\omega) = \frac{X(\omega)}{Y(\omega)}$$
es una variable aleatoria sobre el mismo espacio.

- **Teorema 6:** Si X e Y son variables aleatorias sobre $(\Omega, \mathcal{A}, P)$, entonces $\max(X, Y)$ y $\min(X, Y)$ son variables aleatorias sobre el mismo espacio.

  Corolario: Si X es una variable aleatoria sobre $(\Omega, \mathcal{A}, P)$, entonces $X^+ = \max(0, X)$, $X^- = -\min(0, X)$ y $|X| = X^+ + X^-$ son variables aleatorias sobre $(\Omega, \mathcal{A}, P)$.

  El recíproco no es cierto: Si $|X|$ es variable aleatoria, X no tiene por qué serlo.

  Ejemplo

  $\Omega = \{1, 2, 3, 4\}$, $\mathcal{A} = \{\emptyset, \Omega, \{1, 2, 3\}, \{4\}\}$

$$\begin{array}{rcrr} X: & \Omega & \longrightarrow & \mathbb{R} \\ & 1 & \mapsto & 1 \\ & 2 & \mapsto & -1 \\ & 3 & \mapsto & 1 \\ & 4 & \mapsto & -1 \end{array}$$

$$\begin{array}{rcrr} |X|: & \Omega & \longrightarrow & \mathbb{R} \\ & \omega & \mapsto & 1 \end{array}$$

  $|X| \equiv 1$ es una variable aleatoria, sin embargo, $X^{-1}(\{1\}) = \{1, 3\} \notin \mathcal{A}$ y, por tanto, X no lo es.

5. Distribución de probabilidad de una variable aleatoria

Cuando se considera una variable aleatoria X sobre un espacio de probabilidad $(\Omega, \mathcal{A}, P)$, los únicos sucesos de interés son los que se expresan en términos de esta variable, esto es, los sucesos de la forma
$$\{\omega \in \Omega \ / \ X(\omega) \in B\} = X^{-1}(B) = \{X \in B\} \in \mathcal{A}$$

Las probabilidades de estos sucesos describen completamente el comportamiento de la variable X y dan lugar a lo que se denomina la **distribución de probabilidad** de X.

Definición

Dada una variable aleatoria X sobre $(\Omega, \mathcal{A}, P)$, se denomina *distribución de probabilidad de* X o probabilidad inducida por X a la función de conjunto

$$P_X = P \circ X^{-1}: \quad \mathcal{B} \longrightarrow [0,1]$$

$$B \longmapsto P_X(B) = P(X^{-1}(B)) = P(\{X \in B\})$$

Teorema
P_X es una medida de probabilidad sobre $(\mathbb{R}, \mathcal{B})$.

Demostración.-
AI: $P_X(B) \geq 0$, $\forall B \in \mathcal{B}$
AII: $P_X(\mathbb{R}) = P(X^{-1}(\mathbb{R})) = P(\Omega) = 1$
AIII: $\{B_n\} \in \mathcal{B}$ mutuamente excluyentes $\Rightarrow$ $\{X^{-1}(B_n)\}$ mutuamente excluyentes, entonces

$$P_X\left(\bigcup_n B_n\right) = P\left(X^{-1}\left(\bigcup_n B_n\right)\right) = P\left(\bigcup_n X^{-1}(B_n)\right) = \sum_n P_X(B_n)$$

■

Por lo tanto, la variable aleatoria X transforma el espacio probabilístico original en un nuevo espacio probabilístico $(\mathbb{R}, \mathcal{B}, P_X)$

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

y el interés se centra exclusivamente en el estudio de este nuevo espacio, esto es, en el estudio de P_X.

Esta es la característica esencial de las variables aleatorias, que transforman un espacio probabilístico arbitrario en un espacio de probabilidad numérico.

Ejemplo En el ejemplo 3 de la sección anterior asociado al lanzamiento de un dado se había considerado la variable aleatoria

$$X : \Omega \to \mathbb{R} \qquad X(\omega) = \begin{cases} 1 & \omega \in \{2,4,6\} \\ 0 & \omega \in \{1,3,5\} \end{cases}$$

y se había obtenido que

$$X^{-1}(B) = \begin{cases} \emptyset & 0,1 \notin B \\ \{1,3,5\} & 0 \in B, 1 \notin B \\ \{2,4,6\} & 0 \notin B, 1 \in B \\ \Omega & 0,1 \in B \end{cases}$$

Por tanto la distribución de probabilidad de X es

$$P_X(B) = P(X^{-1}(B)) = \begin{cases} P(\emptyset) = 0 & 0,1 \notin B \\ P(\{1,3,5\}) = \dfrac{1}{2} & 0 \in B, 1 \notin B \\ P(\{2,4,6\}) = \dfrac{1}{2} & 0 \notin B, 1 \in B \\ P(\Omega) = 1 & 0,1 \in B \end{cases}$$

6. Función de distribución de una variable aleatoria

Como ya hemos indicado, el estudio de una variable aleatoria se reduce al de su distribución de probabilidad, que es una función de conjunto, definida sobre la σ-álgebra de Borel $\mathcal{B}$.

Si bien para ciertos tipos de variables aleatorias el manejo de estas funciones de conjunto puede ser simple (como es el caso de variables con un número finito de valores), en general, trabajar con este tipo de funciones puede ser complicado.

Este inconveniente se resuelve asignando a cada distribución de probabilidad P_X una función de punto que la describe completamente y se denomina *función de distribución* de la variable aleatoria X.

Definición

Dada una variable aleatoria X definida sobre un espacio de probabilidad $(\Omega, \mathcal{A}, P)$ con distribución de probabilidad P_X, se denomina función de distribución de la variable a

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

Teorema

La función de distribución de una variable aleatoria X satisface

1) Es monótona no decreciente

2) Es continua a la derecha

3) $\lim\limits_{x \to +\infty} F(x) = 1$ y $\lim\limits_{x \to -\infty} F(x) = 0$

Demostración

1) $x_1 < x_2 \Longrightarrow (-\infty, x_1] \subset (-\infty, x_2] \Longrightarrow$ (usando la monotonía de P_X) $F_X(x_1) = P_X((-\infty, x_1]) \leq P_X((-\infty, x_2]) = F_X(x_2)$.

2) La demostración rigurosa de esta propiedad exige trabajar con sucesiones de conjuntos y usaremos (aunque no se ha probado) la continuidad de una medida de probabilidad, es decir, que si $\{A_n\}_{n \in \mathbb{N}}$ es una sucesión de conjuntos tal que $\exists lim A_n = A$, entonces $\exists lim P(A_n) = P(A)$.

Puesto que F_X es monótona, $\forall x_0 \in \mathbb{R}$, $\exists \lim\limits_{x \to x_0^+} F_X(x)$ y este límite puede obtenerse por sucesiones monótonas decrecientes a x_0. Por tanto, hemos de probar que

$$\lim_{x_n \downarrow x_0} F_X(x_n) = F_X(x_0)$$

En efecto

$$\{(-\infty, x_n]\} \downarrow \bigcap_n (-\infty, x_n] = (-\infty, x_0]$$

y, por tanto,

$$\lim_n P_X((-\infty, x_n]) = P_X((-\infty, x_0])$$

3) Análogamente, considerando $x_n \uparrow +\infty \Longrightarrow (-\infty, x_n] \uparrow \mathbb{R}$ y $x_n \downarrow -\infty \Longrightarrow (-\infty, x_n] \downarrow \emptyset$ ■

Notemos que la demostración de estas propiedades se basa exclusivamente en el hecho de que P_X es una medida de probabilidad.

Por tanto, cualquier medida de probabilidad P sobre $(\mathbb{R}, \mathcal{B})$ define una función de punto $F_P : \mathbb{R} \longrightarrow [0,1]$ no decreciente, continua a la derecha y tal que $\lim\limits_{x \to +\infty} F_P(x) = 1$ y $\lim\limits_{x \to -\infty} F_P(x) = 0$. Dicha función se define por $F_P(x) = P((-\infty, x])$.

Sin embargo, lo realmente importante en Cálculo de Probabilidades es que el recíproco de este resultado es también cierto. Esto es, toda función $F : \mathbb{R} \longrightarrow \mathbb{R}$ no decreciente, continua a la derecha y tal que $\lim\limits_{x \to +\infty} F(x) = 1$ y $\lim\limits_{x \to -\infty} F(x) = 0$, determina una única medida de probabilidad P_F sobre $(\mathbb{R}, \mathcal{B})$ tal que $P_F((-\infty, x]) = F(x)$.

Teorema de Correspondencia

- Si P es una medida de probabilidad sobre $(\mathbb{R}, \mathcal{B})$, $F_P : \mathbb{R} \longrightarrow \mathbb{R}$ definida como $F_P(x) = P((-\infty, x])$ es no decreciente, continua a la derecha y verifica $\lim\limits_{x \to +\infty} F_P(x) = 1$ y $\lim\limits_{x \to -\infty} F_P(x) = 0$.

- Si $F : \mathbb{R} \longrightarrow \mathbb{R}$ es no decreciente, continua a la derecha y verifica $\lim\limits_{x \to +\infty} F(x) = 1$ y $\lim\limits_{x \to -\infty} F(x) = 0$, entonces existe una única medida de probabilidad P_F sobre $(\mathbb{R}, \mathcal{B})$ que satisface $P_F((-\infty, x]) = F(x), \ \forall x \in \mathbb{R}$.

Existe, por tanto, una correspondencia biunívoca entre las medidas de probabilidad en $(\mathbb{R}, \mathcal{B})$ y las funciones de punto sobre $\mathbb{R}$ verificando tales propiedades. Según esta correspondencia, a la distribución de probabilidad de una variable aleatoria X, P_X, le corresponde su función de distribución, esto es, $F_{P_X} = F_X$ y a la función de distribución F_X le corresponde P_X, ya que $P_X((-\infty, x])) = F(x)$.

Por tanto, la función de distribución de una variable aleatoria determina completamente su distribución de probabilidad.

Teorema
Toda función $F : \mathbb{R} \longrightarrow \mathbb{R}$ verificando 1), 2) y 3) es la función de distribución de alguna variable aleatoria definida sobre algun espacio de probabilidad.

Demostración
En efecto, por el Teorema de Correspondencia, existe una única medida de probabilidad P_F sobre $(\mathbb{R}, \mathcal{B})$ que satisface $P_F((-\infty, x]) = F(x),\ \forall x \in \mathbb{R}$.
Asi, si definimos $X : (\mathbb{R}, \mathcal{B}, P_F) \longrightarrow (\mathbb{R}, \mathcal{B})$ como $X(x) = x$, la distribución de probabilidad P_X de dicha variable aleatoria coincide con P_F y, por tanto, su función de distribución $F_X(x) = P_X((-\infty, x]) = P_F((-\infty, x]) = F(x), \forall x \in \mathbb{R}$.

Otras propiedades de la función de distribución

1. $\forall x \in \mathbb{R},\ \exists \lim\limits_{y \to x^-} F_X(y) = F_X(x^-) = P(X < x)$ y

   $$\exists \lim_{y \to x^+} F_X(y) = F_X(x) = P(X \leq x)$$

   Demostración:

   La existencia de los límites está garantizada por ser F_X monótona y por la continuidad a la derecha, es claro que $\lim\limits_{y \to x^+} F_X(y) = F_X(x)$.

   Veamos ahora que $\lim\limits_{y \to x^-} F_X(y) = P(\{X < x\})$. Ya que el límite existe, puede tomarse por sucesiones crecientes

   $$\lim_{y \to x^-} F_X(y) = \lim_{x_n \uparrow x} F_X(x_n)$$
   $$F_X(x_n) = P(X \leq x_n) = P_X((-\infty, x_n]) \uparrow P_X((-\infty, x)) = P(X < x).$$

2. Los únicos puntos de discontinuidad de F_X son de salto y la longitud del salto en cualquier punto $x \in \mathbb{R}$ es
   $$P(X = x) = F_X(x) - F_X(x^-)$$

   Demostración:

   Esto es debido a la continuidad a la derecha, no decrecimiento y existencia de límite a la izquierda. Además, el salto es

   $$\lim_{y \to x^+} F_X(y) - \lim_{y \to x^-} F_X(y) = P(X \leq x) - P(X < x) = P(X = x)$$

3. x es un punto de continuidad de $F_X \iff P(X = x) = 0$.

4. El conjunto de puntos de discontinuidad de F_X es numerable.

   Demostración:

$$D = \{x \in \mathbb{R} \ / \ F_X(x) > F_X(x^-)\}$$

   $x \in D \iff F_X(x) - F_X(x^-) > 0 \iff \exists n \ / \ F_X(x) - F_X(x^-) \geq 1/n.$

   Entonces, si $E_n = \{x \in D \ / \ F_X(x) - F_X(x^-) \geq 1/n\}$ es claro que $D = \bigcup_n E_n$. Ya que E_n contiene a lo más n puntos (en caso contrario, la suma de los saltos sería mayor que uno), D es numerable.

Cálculo de probabilidades mediante funciones de distribución

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

Ejemplo: En el ejemplo 3 de la sección anterior asociado al lanzamiento de un dado se había considerado la variable aleatoria

$$X : \Omega \to \mathbb{R} \qquad X(\omega) = \begin{cases} 1 & \omega \in \{2,4,6\} \\ 0 & \omega \in \{1,3,5\} \end{cases}$$

y se había obtenido que

$$X^{-1}((-\infty, x]) = \begin{cases} \emptyset & x < 0 \\ \{1,3,5\} & 0 \leq x < 1 \\ \Omega & x \geq 1 \end{cases}$$

Por tanto la función de distribución de X es

$$F_X(x) = \begin{cases} P(\emptyset) = 0 & x < 0 \\ P(\{1,3,5\}) = 1/2 & 0 \leq x < 1 \\ P(\Omega) = 1 & x \geq 1 \end{cases}$$

7. Variables aleatorias discretas

Defnición Una variable aleatoria $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_X)$ es discreta si toma valores en un conjunto numerable, es decir, $\exists E_X \subset \mathbb{R}$ tal que

$$P(X \in E_X) = 1$$

Nota: E_X es el conjunto de valores de la variable y es claro que

$$x \notin E_X, \quad P(X = x) = 0$$

Función masa de probabilidad
Si X es una variable aletoria discreta con valores en el conjunto $E_X = \{x_n,\ n \in \mathbb{N}\}$, se define su función masa de probabilidad como

$$\begin{array}{rccl} p_X: & E_X & \longrightarrow & [0,1] \\ & x_n & \longmapsto & P_X(x_n) = P(X = x_n) = P_X(\{x_n\}) \end{array}$$

Propiedades de la función masa de probabilidad

- Es no negativa, $p_X(x) \geq 0,\ \forall x \in E_X$.
- La suma de sus valores es la unidad, $\displaystyle\sum_{x \in E_X} p_X(x) = 1$
- La función masa de probabilidad de una variable aleatoria discreta determina la distribución de probabilidad de la variable

$$\forall B \in \mathcal{B},\ P_X(B) = P(X \in B) = \sum_{x \in B \cap E_X} P(X = x)$$

y, por tanto, su función de distribución

$$\forall x \in \mathbb{R},\ F_X(x) = \sum_{\substack{x_n \in E_X \\ x_n \leq x}} P(X = x_n) = \sum_{n=1}^{\infty} P(X = x_n)\epsilon(x - x_n)$$

con

$$\epsilon(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Por tanto, la función de distribución de una variable aleatoria discreta es una función de salto cuyos puntos de salto son aquellos números reales x_n tales que $P(\{X = x_n\}) > 0$ y dicha probabilidad es la longitud del salto.

El crecimiento a saltos de la función de distribución caracteriza a las variables discretas:

Teorema: Caracterización de variables discretas

$X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_X)$ es discreta si y sólamente si F_X crece sólo a saltos.

Demostración:

$\Longleftarrow$) Si F_X sólo crece a saltos, ya que $\lim_{x \to +\infty} F_X(x) = 1$ y $\lim_{x \to -\infty} F_X(x) = 0$, la suma de las longitudes de todos los saltos debe ser uno.

Sea D el conjunto de puntos de salto. Puesto que D es numerable

$$P(X \in D) = \sum_{x \in D} P(X = x) = 1.$$

Entonces, X toma valores en D, numerable, y es, por tanto, discreta.

$\Longrightarrow$) El recíproco es inmediato, como se ha visto anteriormente.

$\blacksquare$

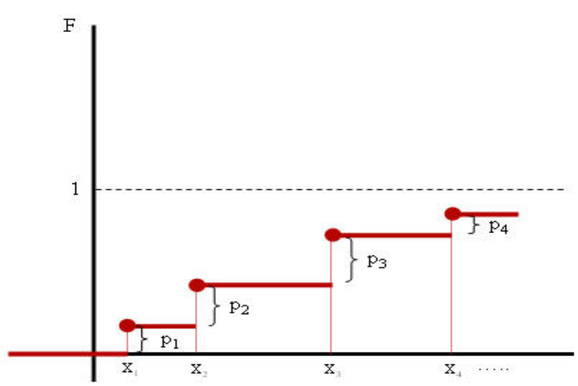
Teorema

Toda colección numerable de números no negativos, de suma la unidad constituyen los valores de la función masa de probabilidad de alguna variable aleatoria de tipo discreto

Demostración

Dada $\{p_n;\ n \geq 1\}$ con $p_n \geq 0,\ \forall n \geq 1$ y $\sum p_n = 1$, se define

$$F(x) = \begin{cases} 0 & x < x_1 \\ \displaystyle\sum_{n \leq i} p_n & x_i \leq x < x_{i+1} \end{cases}, \text{ o bien } F(x) = \begin{cases} 0 & x < 0 \\ \displaystyle\sum_{n \leq i} p_n & i \leq x < i+1 \end{cases}$$



Esta función es, evidentemente, no decreciente, continua a la derecha y con límites cero y uno en ∞ y $+\infty$, respectivamente. Por tanto, por la aplicación del Teorema de Correspondencia, existe una variable aleatoria con tal función como función de distribución y, además, por las características mostradas por ella, dicha variable es de tipo discreto.

■

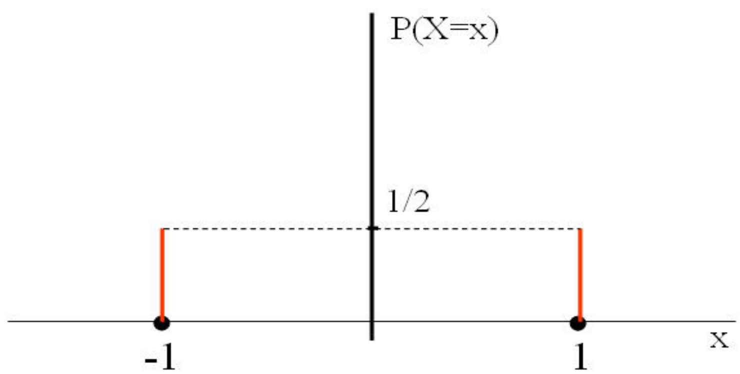
EJEMPLOS

Ejemplo 1

Un jugador recibe una moneda si en el lanzamiento de un dado sale un número par; en caso contrario, pierde una moneda. Calcular la función masa de probabilidad y la función de distribución de la variable aleatoria que describe la ganancia.

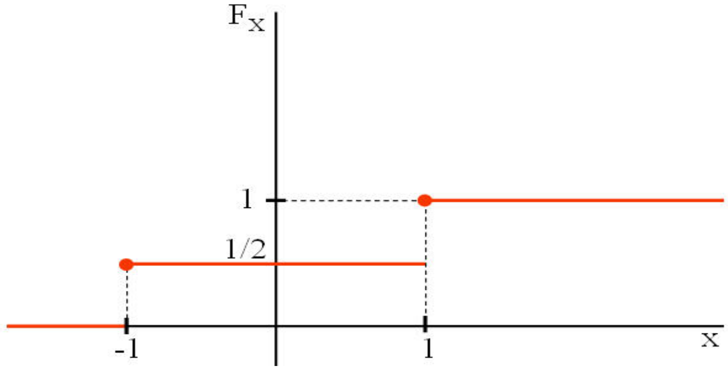
Función masa de probabilidad

$P(X = 1) = P_X(1) = P(X^{-1}(\{1\})) = P(\{2, 4, 6\}) = 1/2$
$P(X = -1) = 1/2$



Función de distribución

$$\begin{array}{rccl} F_X : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & P(\{X \leq x\}) = \begin{cases} 0 & x < -1 \\ 1/2 & -1 \leq x < 1 \\ 1 & x \geq 1 \end{cases} \end{array}$$



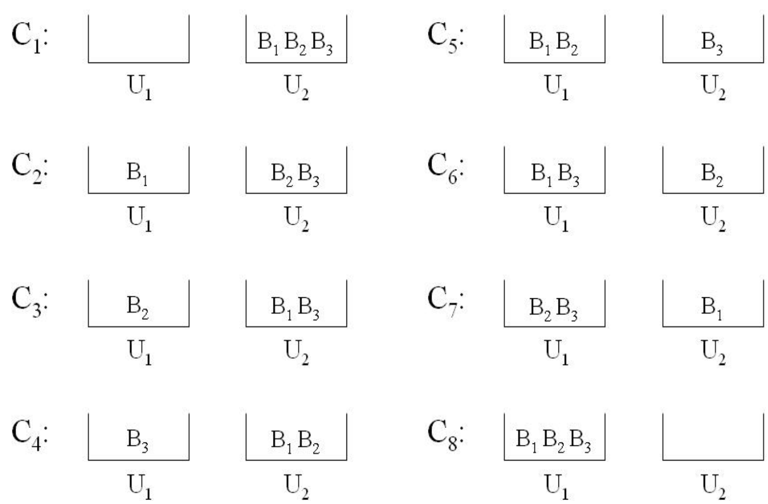
En general, si B es un conjunto de Borel arbitrario

$$P(X \in B) = P_X(B) = P(X^{-1}(B)) = \begin{cases} P(\emptyset) = 0 & -1 \notin B,\ 1 \notin B \\ P(\{1,3,5\}) = 1/2 & -1 \in B,\ 1 \notin B \\ P(\{2,4,6\}) = 1/2 & -1 \notin B,\ 1 \in B \\ P(\Omega) = 1 & -1 \in B,\ 1 \in B \end{cases}$$

Ejemplo 2

Se introducen al azar tres bolas en dos urnas U_1 y U_2. Sea N el número de bolas en la primera urna, especificar su función masa de probabilidad y su función de distribución.

Sean B_1, B_2 y B_3 las tres bolas; las distintas colocaciones, todas equiprobables son:



Cada disposición puede identificarse con una ordenación de las dos urnas

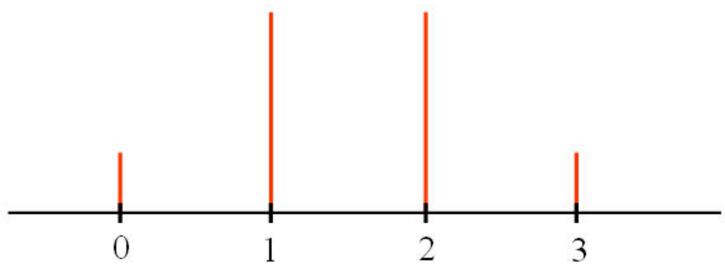
$$\Omega = \{U_1U_1U_1, U_1U_1U_2, U_1U_2U_1, U_2U_1U_1, U_1U_2U_2, U_2U_1U_2, U_2U_2U_1, U_2U_2U_2\}$$

$$\begin{array}{rlcl} N: & \Omega & \longrightarrow & \mathbb{R} \\ & U_1U_1U_1 & \longmapsto & 3 \\ & U_1U_1U_2 & \longmapsto & 2 \\ & U_1U_2U_1 & \longmapsto & 2 \\ & U_2U_1U_1 & \longmapsto & 2 \\ & U_1U_2U_2 & \longmapsto & 1 \\ & U_2U_1U_2 & \longmapsto & 1 \\ & U_2U_2U_1 & \longmapsto & 1 \\ & U_2U_2U_2 & \longmapsto & 0 \end{array}$$

Función masa de probabilidad

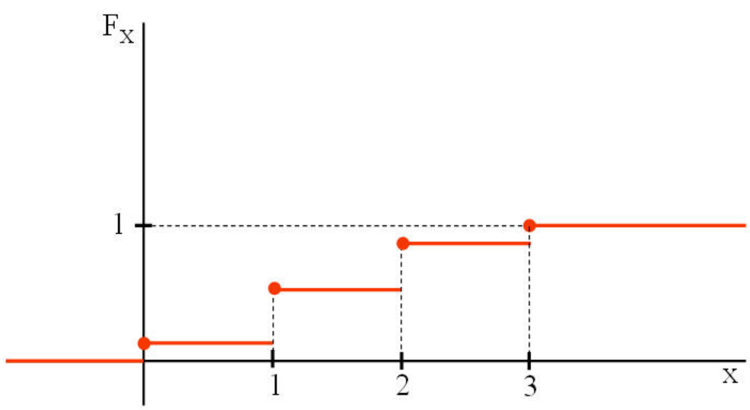
$P(N = 0) = 1/8$

$P(N = 1) = 3/8$
$P(N = 2) = 3/8$
$P(N = 3) = 1/8$



Función de distribución

$$F_N(x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \leq x < 1 \\ 4/8 & 1 \leq x < 2 \\ 7/8 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$



Distribución de probabilidad

$$P_N(B) = P(N \in B) = \sum_{\substack{i=0 \\ i \in B}}^{3} P(N = i)$$

Ejemplo 3

Un juego consiste en lanzar una modeda cuatro veces y contar el número de rachas de resultados iguales. El jugador recibe una moneda por cada racha pero tiene que pagar una moneda por participar en el juego. Dar la distribución de probabilidad del beneficio del jugador.

RESULTADOS POSIBLES Ω		BENEFICIO X
$CCCC$	$\longmapsto$	0
$CCCX$	$\longmapsto$	1
$CCXC$	$\longmapsto$	2
$CXCC$	$\longmapsto$	2
$XCCC$	$\longmapsto$	1
$CCXX$	$\longmapsto$	1
$CXCX$	$\longmapsto$	3
$XCCX$	$\longmapsto$	2
$CXXC$	$\longmapsto$	2
$XCXC$	$\longmapsto$	3
$XXCC$	$\longmapsto$	1
$XXXC$	$\longmapsto$	1
$XXCX$	$\longmapsto$	2
$XCXX$	$\longmapsto$	2
$CXXX$	$\longmapsto$	1
$XXXX$	$\longmapsto$	0

Función masa de probabilidad

$P(X = 0) = 1/8$
$P(X = 1) = 3/8$
$P(X = 2) = 3/8$
$P(X = 3) = 1/8$

Función de distribución

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1/8 & 0 \leq x < 1 \\ 4/8 & 1 \leq x < 2 \\ 7/8 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Distribución de probabilidad

$$P_X(B) = P(X \in B) = \sum_{\substack{i=0 \\ i \in B}}^{3} P(X = i)$$

Observemos que $P_X \equiv P_N$ (del ejemplo anterior). Por tanto, las dos variables aleatorias, aunque referidas a espacios de probabilidad distintos, tienen la misma distribución.

Si, en el ejemplo anterior, un jugador recibe una moneda por cada bola que haya en la primera urna, N es su beneficio. Por tanto, los dos juegos serían equivalentes, en el sentido de que el jugador puede ganar las mismas cantidades con las mismas probabilidades.

8. Variables aleatorias continuas

Existen muchas situaciones prácticas en las que las variables de interés toman valores en un conjunto no numerable de puntos:

- Tiempo de vida de un determinado producto.
- Suma de dos puntos elegidos al azar de un intervalo.

En tales casos, no es posible asignar una probabilidad positiva a cada uno de los valores de la variable de forma que la suma de esas probabilidades sea la unidad como ocurre en el caso discreto. Así, el tratamiento probabilístico de este tipo de variables es diferente al de las variables discretas.

Aunque la clase de variables de este tipo es más amplia, vamos a estudiar ahora un tipo particular, las denominadas **variables absolutamente continuas**, a las que es usual llamar simplemente **de tipo continuo o continuas**.

Definición
Una variable aleatoria $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_X)$ es continua si su función de distribución es continua en todo $\mathbb{R}$ y derivable, y con derivada continua, salvo en a lo sumo un conjunto numerable de puntos. Esto es equivalente a que $\exists f_X : \mathbb{R} \longrightarrow \mathbb{R}$ no negativa e integrable tal que

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

La función f_X se denomina **función de densidad** de la variable aleatoria X.

Propiedades

1. La función de densidad de una variable aleatoria X es una función no negativa, integrable y $\displaystyle\int_{-\infty}^{+\infty} f_X(t)dt = 1$.

   La no negatividad e integrabilidad forman parte de la definición y

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

2. Si x_0 es un punto de continuidad de f_X, F_X es derivable en x_0 y $F_X'(x_0) = f_X(x_0)$.

3. La función de densidad de un variable aleatoria continua determina a su función de distribución F_X y a su distribución de probabilidad P_X.

   En efecto, por la definición se tiene

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

y además se puede probar que

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

Para intervalos, es evidente

- $P(X \leq b) = P_X((-\infty, b]) = F_X(b) = \int_{-\infty}^{b} f_X(x)\,dx$
  Además, como F_X es continua y, por tanto,
  $$\forall x \in \mathbb{R}, \quad P_X(\{x\}) = P(X = x) = F_X(x^+) - F_X(x^-) = 0,$$
  $P(X < b) = P(X \leq b) = F_X(b) = \int_{-\infty}^{b} f_X(x)\,dx$
- $P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(x)\,dx$
  De nuevo esa probabilidad coincide con $P(a \leq X < b)$, $P(a < X < b)$ y $P(a \leq X \leq b)$.

Para conjuntos de Borel más generales esta integral debe entenderse en el sentido de Lebesgue (más general que la integral de Riemann y que escapa a los contenidos de este curso)

Finalmente, como consecuencia del Teorema de Correspondencia, se tiene el siguiente resultado.

Teorema
Si $f : \mathbb{R} \longrightarrow \mathbb{R}$ es no negativa, integrable y tal que $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = 1$, f es la función de densidad de alguna variable aleatoria de tipo continuo.

Demostración
Definimos $F : \mathbb{R} \longrightarrow \mathbb{R}$ como

$$F(x) = \int_{-\infty}^{x} f(t)dt$$

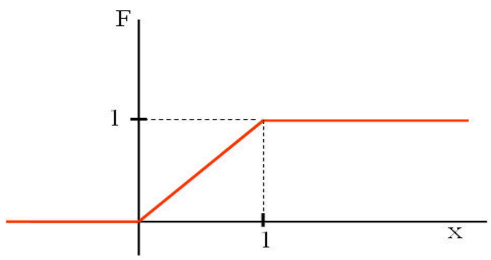
F es no decreciente, continua (y, por tanto, continua a la derecha) y $F(-\infty) = 0$ y $F(+\infty) = 1$. Entonces F es la función de distribución de alguna variable aleatoria, la cual, evidentemente, es de tipo continuo con función de densidad f.

■

EJEMPLOS

Ejemplo 1
Sea X una variable aleatoria con función de distribución

$$F(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$



F es derivable en $\mathbb{R} - \{0, 1\}$ y

$$F^{'}(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 < x < 1 \\ 0 & x > 1 \end{cases}$$

Entonces, podemos definir

$$f(x) = \begin{cases} 0 & x < 0 \\ 1 & 0 \leq x \leq 1 \\ 0 & x > 1 \end{cases}$$

que es continua salvo en $x = 0, 1$ y es evidente que

$$F(x) = \int_{-\infty}^{x} f(t)dt, \quad \forall x \in \mathbb{R}$$

Por tanto, f es la función de densidad de X (notemos que f puede cambiarse arbitrariamente en 0 y 1 y no es única; tomamos la versión más usual)

$$P(\{0.4 < X \leq 0.6\}) = F(0.6) - F(0.4) = 0.2 = \int_{0.4}^{0.6} f(t)dt = \int_{0.4}^{0.6} 1\, dt$$

Ejemplo 2

Comprobar que la siguiente función es función de densidad. Calcular su función de distribución y, si X es una variable aleatoria con dicha distribución, calcular $P(\{0.3 < X \leq 1.5\})$.

$$f(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 < x \leq 1 \\ 2 - x & 1 \leq x < 2 \\ 0 & x \geq 2 \end{cases}$$

Dicha función es no negativa y además,

$$\int_{-\infty}^{+\infty} f(t)\, dt = \int_{0}^{1} t\, dt + \int_{1}^{2} (2 - t)\, dt = \left.\frac{t^2}{2}\right|_0^1 + \left(2t - \left.\frac{t^2}{2}\right|_1^2 = \frac{1}{2} + 4 - 2 - 2 + \frac{1}{2} = 1.$$

$$F(x) = \begin{cases} 0 & x < 0 \\ \displaystyle\int_0^x t\,dt = \frac{x^2}{2} & 0 \le x < 1 \\ \displaystyle\int_0^1 t\,dt + \int_1^x (2-t)\,dt = 2x - \frac{x^2}{2} - 1 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

$$P[0{,}3 < X \le 1{,}5] = F(1{,}5) - F(0{,}3) = 0{,}83$$

En general, si $f = 0$ salvo en (a, b), entonces $F(x) = 0$ si $x < a$ y $F(x) = 1$ si $x \ge b$. Si $f = 0$ salvo en $(a, +\infty)$, entonces $F(x) = 0$ si $x < a$. Si $f = 0$ salvo en $(-\infty, b)$, entonces $F(x) = 1$ si $x \ge b$.

9. Variables aleatorias mixtas

Aunque la mayoría de las distribuciones que aparecen en problemas prácticos son discretas o continuas, a veces es necesario considerar otro tipo de distribuciones que son "mezcla" de ambos tipos. Son las denominadas **distribuciones mixtas**.

Definición

Una variable aleatoria $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_X)$ se dice que es mixta si $\exists D_X \subset \mathbb{R}$ numerable y $\exists h_X : \mathbb{R} \longrightarrow \mathbb{R}$ no negativa e integrable tal que

$$\forall x \in \mathbb{R}, \quad F_X(x) = P(\{X \le x\}) = \sum_{\substack{x_j \in D_X \\ x_j \le x}} P(\{X = x_j\}) + \int_{-\infty}^{x} h_X(t)\,dt.$$

Notemos que en este tipo de variables, h_X no es una función de densidad ya que $\displaystyle\int_{-\infty}^{+\infty} h_X(t)\,dt \neq 1$.

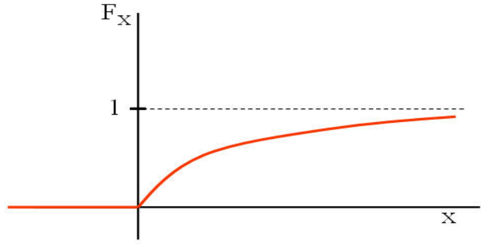
Ejemplo

Supóngase que en un sistema eléctrico, el voltaje X es una variable aleatoria continua con función de densidad

$$f_X(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{1}{(1+x)^2} & x \ge 0 \end{cases}$$

Así,

$$F_X(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{x}{1+x} & x \ge 0 \end{cases}$$

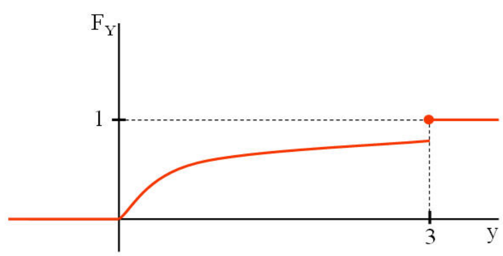


El voltaje de este sistema se mide con un voltímetro que registra el valor real de X si $X \leq 3$, pero si $X > 3$ el voltímetro registra siempre el valor 3. Se define Y como el valor registrado por el voltímetro. Entonces la distribución de Y es mixta:

- $P(\{Y = 3\}) = P(\{X \geq 3\}) = \displaystyle\int_3^{+\infty} \frac{1}{(1+x)^2}\,dx = -\frac{1}{1+x}\Big|_3^{+\infty} = \frac{1}{4} > 0$

- De lo anterior se deduce que $P(\{0 < Y < 3\}) = 3/4$

$$F_Y(y) = \begin{cases} 0 & y \leq 0 \\ F_X(y) = \dfrac{y}{1+y} & 0 \leq y < 3 \\ 1 & y \geq 3 \end{cases}$$



En este caso, $D_X = \{3\}$ con $P(Y = 3) = F_Y(3) - F_Y(3^-) = 1 - 3/4 = 1/4$ y

$$h_Y(y) = f_X(y) = \frac{1}{(1+y)^2}, \ 0 < y < 3.$$

10. Funciones de variables aleatorias. Cambio de variable

Teorema
Si $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_X)$ es una variable aleatoria y $g : (\mathbb{R}, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B})$ es una función medible (esto es, $g^{-1}(\mathcal{B}) \subset \mathcal{B}$), entonces $Y = g(X) : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B})$ es una variable aleatoria.

Demostración
$\forall B \in \mathcal{B},\ Y^{-1}(B) = (g(X))^{-1}(B) = X^{-1}(\underbrace{g^{-1}(B)}_{\in \mathcal{B}}) \in \mathcal{A}$

Por tanto, $Y^{-1}(\mathcal{B}) \subset \mathcal{A}$ y, puesto que está definida sobre un espacio de probabilidad es una variable aleatoria.

Nos planteamos el problema de obtener la distribución de probabilidad de Y a partir de la de X. En teoría, el problema se resuelve de forma inmediata mediante el siguiente teorema general.

Teorema general de Cambio de Variable
Sea $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_X)$ una variable aleatoria y $g : (\mathbb{R}, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B})$ una función medible. Sea $Y = g(X)$, entonces

- $\forall B \in \mathcal{B},\ P_Y(B) = P_X(g^{-1}(B))$
- $\forall y \in \mathbb{R},\ F_Y(y) = P_X(g^{-1}((-\infty, y]))$

Demostración

- $P_Y(B) = P[Y^{-1}(B)] = P[X^{-1}(g^{-1}(B))] = P_X(g^{-1}(B))$
- $F_Y(y) = P_Y((-\infty, y]) = P_X(g^{-1}((-\infty, y]))$

■

Pero en la práctica trabajaremos con variables discretas o continuas, o sea, con funciones masa de probabilidad o funciones de densidad. Nos interesa, por tanto, especificar las fórmulas de cambio de variable para tales casos.

Cambio de variable discreto

Teorema general de Cambio de Variable
Si $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_X)$ una variable aleatoria discreta, $P(\{X \in E_X\}) = 1$ y $g : (\mathbb{R}, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B})$ una función medible, entonces $Y = g(X)$ es discreta con valores en $g(E_X)$ y

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

Demostración

- $g(E_X)$ es numerable y $P\{Y \in g(E_X)\} = P\{g(X) \in g(E_X)\} \geq P\{X \in E_X\} = 1$
- $\forall y \in g(E_X),\ P\{Y = y\} = P\{g(X) = y\} = P\{X \in g^{-1}(y)\} = \displaystyle\sum_{x \in g^{-1}(y) \cap E_X} P\{X = x\}$

■

EJEMPLOS

Ejemplo 1
Sea X una variable aleatoria discreta con función masa de probabilidad

$$P\{X = x\} = \frac{\mathrm{e}^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2 \ldots \quad (\lambda > 0)$$

(colección numerable de números no negativos de suma la unidad).
Calcular la distribución (función masa de probabilidad) de $Y = X^2$.

$$E_X = \{0, 1, 2, \ldots\} \overset{g(x)=x^2}{\longrightarrow} g(E_X) = \{0, 1, 4, \ldots\}$$

$$\forall y \in g(E_X) \;\; P\{Y = y\} = P\{X^2 = y\} = P\{X = \sqrt{y}\} = \frac{\mathrm{e}^{-\lambda}\lambda^{\sqrt{y}}}{\sqrt{y}!}$$

Ejemplo 2
Sea X una variable aleatoria discreta con función masa de probabilidad $P\{X = -2\} = 1/5$, $P\{X = -1\} = 1/6$, $P\{X = 0\} = 1/5$, $P\{X = 1\} = 1/15$, $P\{X = 2\} = 11/30$. Calcular la función masa de probabilidad de $Y = X^2 + 3$.

$$E_X = \{-2, -1, 0, 1, 2\} \overset{g(x)=x^2+3}{\longrightarrow} g(E_X) = \{7, 4, 3\}$$

$$P\{Y = 3\} = P\{X^2 + 3 = 3\} = P\{X^2 = 0\} = P\{X = 0\} = 1/5$$
$$P\{Y = 4\} = P\{X^2 + 3 = 4\} = P\{X^2 = 1\} = P\{X = 1\} + P\{X = -1\} = 7/30$$
$$P\{Y = 7\} = P\{X^2 + 3 = 7\} = P\{X^2 = 4\} = P\{X = 2\} + P\{X = -2\} = 17/30$$

Cambio de variable continuo

Este caso no es tan simple como el discreto, ya que una variable aleatoria de tipo continuo no tiene por qué transformarse en otra de tipo continuo. Por ejemplo, sea X una variable aleatoria continua con función de densidad

$$f(x) = \begin{cases} 1/2 & x \in (-1, 1) \\ 0 & x \notin (-1, 1) \end{cases}$$

y $g : (\mathbb{R}, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B})$ una función medible definida como

$$g(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

$Y = g(X)$ toma valores en el conjunto $\{0, 1\}$ y es, por tanto, una variable aleatoria discreta con $P\{Y = 0\} = P\{X < 0\} = 1/2$ y $P\{Y = 1\} = P\{X \geq 0\} = 1/2$.

Sin embargo si se define g como

$$g(x) = \begin{cases} x & x < 0 \\ \\ 1 & x \geq 0 \end{cases}$$

$Y = g(X)$ toma valores en el conjunto $(-1, 0) \cup \{1\}$ y es una variable de tipo mixto con $P\{Y = 1\} = P\{X \geq 0\} = 1/2$ y

$$h_y(y) = \begin{cases} 1/2 & y \in (-1, 0) \\ \\ 0 & y \notin (-1, 0) \end{cases}$$

Aquí analizaremos el caso de transformaciones que convierten una variable continua en discreta y el caso de transformaciones biunívocas de variables continuas (que transforman variables continuas en continuas).

Teorema (CV continua a discreta): Sea $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_X)$ una variable aleatoria continua con función de densidad f_X y sea $Y = g(X)$ una variable aleatoria discreta con valores en un conjunto numerable E_Y. Entonces,

$$\forall y \in E_Y, \quad P\{Y = y\} = \int_{g^{-1}(y)} f_X(x)\, dx$$

Demostración
$P\{Y = y\} = P\{g(X) = y\} = P\{X \in g^{-1}(y)\} = \int_{g^{-1}(y)} f_X(x)\, dx$

■

Ejemplo
Sea X una variable aleatoria continua con función de densidad

$$f(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2}, \quad x \in \mathbb{R}$$

y sea Y una variable aleatoria definida como

$$Y = \begin{cases} 1 & X > 0 \\ 1/2 & X = 0 \\ -1 & X < 0 \end{cases}$$

Evidentemente, Y es una variable discreta y, además $P\{Y = 1/2\} = 0$. Por tanto, el conjunto de valores de Y es $\{-1, 1\}$ y su función masa de probabilidad es

$$P\{Y = 1\} = P\{X > 0\} = \int_0^{+\infty} \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} dx$$

$$P\{Y = -1\} = P\{X < 0\} = \int_{-\infty}^0 \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-x^2/2} dx = [y = -x,\ dy = -dx] = P\{Y = 1\}$$

Así, $P\{Y = 1\} = P\{Y = -1\} = 1/2$.

Teorema (CV continua a continua): Sea $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_X)$ una variable aleatoria continua con valores en el intervalo (a, b) ($f_X > 0 \;\Leftrightarrow x \in (a, b)$) y sea $g : (a, b) \longrightarrow \mathbb{R}$ una función derivable y estríctamente monótona (creciente o decreciente). Entonces $Y = g(X)$ es una variable aleatoria continua con función de densidad

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| & y \in g((a, b)) \\ \\ 0 & y \notin g((a, b)) \end{cases}$$

EJEMPLOS

Ejemplo 1

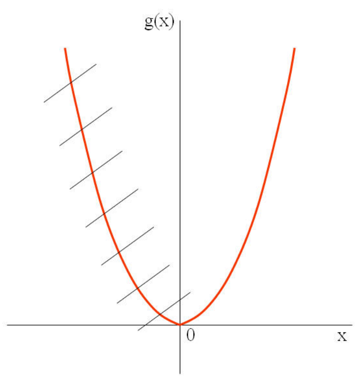
Sea X una variable aleatoria continua con función de densidad

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

Calcular la función de densidad de $Y = X^2$.

Consideramos la función

$$\begin{array}{rccc} g: & \mathbb{R}^+ & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x^2 \end{array}$$



Observemos que esta transformación es creciente estrictamente para $x \geq 0$ que es el recinto donde toma valores nuestra variable.

Entonces, $Y = g(X) = X^2 \geq 0,\; (Y \in (0, +\infty))$ y

$$f_Y(y) = \begin{cases} 0 & y < 0 \\ f(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| & y \geq 0 \end{cases}$$

$y = g(x) = x^2 \;\longrightarrow\; x = g^{-1}(y) = \sqrt{y}, \quad \frac{dg^{-1}(y)}{dy} = \frac{1}{2\sqrt{y}}$ y, por tanto,

$$f_Y(y) = \begin{cases} 0 & y < 0 \\ \lambda \mathrm{e}^{-\lambda\sqrt{y}} \frac{1}{2\sqrt{y}} & y \geq 0 \end{cases}$$

Otra opción, sería calcular directamente la función de distribución y luego derivar.

$$F_Y(y) = P[X^2 \leq y] = \begin{cases} 0 & y < 0 \\ P[X \leq \sqrt{y}] & y \geq 0 \end{cases}$$

donde

$$P[X \leq \sqrt{y}] = \int_0^{\sqrt{y}} \lambda \mathrm{e}^{-\lambda x} = \left[-\mathrm{e}^{-\lambda x}\right]_0^{\sqrt{y}} = 1 - \mathrm{e}^{-\lambda\sqrt{y}}$$

De aquí se deduce también que

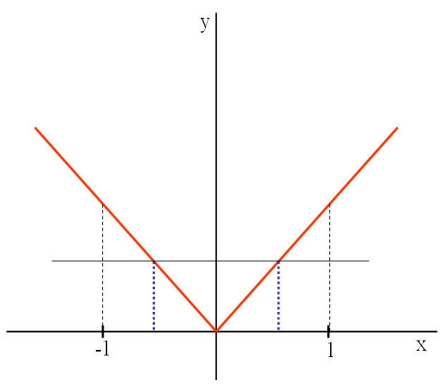
$$f_Y(y) = \begin{cases} 0 & y < 0 \\ \lambda \mathrm{e}^{-\lambda\sqrt{y}} \frac{1}{2\sqrt{y}} & y \geq 0 \end{cases}$$

Ejemplo 2
Sea X una variable aleatoria continua con función de densidad

$$f(x) = \begin{cases} 1/2 & x \in (-1,1) \\ 0 & x \notin (-1,1) \end{cases}$$

Calcular la distribución de $Y = |X|$.

$X \in (-1,1)$ pero $g : (-1,1) \longrightarrow \mathbb{R}$ con $g(x) = |x|$ no está en las condiciones del Teorema de cambio de variable continuo.



Observamos que $Y \in [0,1)$, pero cada valor de Y tiene dos antiimágenes $X = \pm Y$. Calculemos la función de distribución de Y

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ P[|X| \leq y] & 0 \leq y < 1 \\ 1 & y \geq 1 \end{cases}$$

con $P[|X| \leq y] = P[-y \leq X \leq y] = F_X(y) - F_X(-y) = \displaystyle\int_{-y}^{y} \frac{1}{2}\, dx = y$

$$[-y, y] \subset [-1, 1] \uparrow.$$

Entonces

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

de donde

$$f_Y(y) = \begin{cases} 0 & y \notin (0,1) \\ 1 & y \in (0,1) \end{cases}$$

Observemos en este ejemplo lo siguiente:

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ \\ F_X(y) - F_X(-y) & 0 \leq y < 1 \\ \\ 1 & y \geq 1 \end{cases} \longrightarrow \quad f_Y(y) = \begin{cases} 0 & y \notin (0,1) \\ \\ f_X(y) + f_X(-y) & y \in (0,1) \end{cases}$$

Entonces, en realidad, lo que se hace es aplicar la fórmula de Cambio de Variable a cada una de las antiimágenes y después, sumar.

1) $Y = |X| \quad X = Y \quad \dfrac{dx}{dy} = 1 > 0$ (estamos considerando sólo las antiimágenes positivas)

$$f_Y^1(y) = \begin{cases} 0 & y \notin (0,1) \\ f_X(y)|1| & y \in (0,1) \end{cases}$$

2) $Y = |X| \quad X = -Y \quad \dfrac{dx}{dy} = -1 < 0$ (estamos considerando sólo las antiimágenes negativas)

$$f_Y^2(y) = \begin{cases} 0 & y \notin (0,1) \\ f_X(-y)| - 1| & y \in (0,1) \end{cases}$$

Si ahora sumamos f_Y^1 y f_Y^2 obtenemos f_Y.

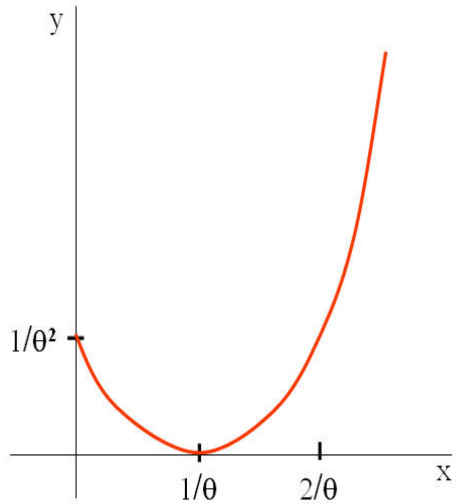
Este resultado es cierto en general y constituye una extensión del Teorema de Cambio de Variable que hemos estudiado para el caso continuo.

Extensión del Teorema de C. V. continuo

Sea $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_X)$ y sea $Y = g(X)$ una nueva variable aleatoria tal que para cada valor de Y, existe un número finito (o incluso numerable) de antiimágenes, cada una de ellas derivable y estrictamente monótona. Entonces, Y es una variable aleatoria de tipo continuo y su función de densidad se obtiene aplicando la fórmula de Cambio de Variable a cada antiimagen y sumando las funciones resultantes.

Ejemplo
Sea X una variable aleatoria con función de densidad $f_\theta(x) = \theta e^{-\theta x},\ x > 0\ (\theta > 0)$. Calcular la función de densidad de $Y = (X - 1/\theta)^2$.

$y = g(x) = (x - 1/\theta)^2$



La variable aleatoria Y toma valores en $(0, +\infty)$.

- Si $0 < y < 1/\theta^2$, la función g tiene dos imágenes inversas, cada una de ellas derivable y estrictamente monótona:

  $g_1^{-1}(y) = \dfrac{1}{\theta} + \sqrt{y} \quad \text{con} \quad \dfrac{dg_1^{-1}(y)}{dy} = \dfrac{1}{2\sqrt{y}} > 0$

  $g_2^{-1}(y) = \dfrac{1}{\theta} - \sqrt{y} \quad \text{con} \quad \dfrac{dg_2^{-1}(y)}{dy} = -\dfrac{1}{2\sqrt{y}} < 0$

  Así

$$f_Y(y) = f_X\left(\frac{1}{\theta} + \sqrt{y}\right)\left|\frac{1}{2\sqrt{y}}\right| + f_X\left(\frac{1}{\theta} - \sqrt{y}\right)\left|-\frac{1}{2\sqrt{y}}\right| =$$

$$\left[f_X\left(\frac{1}{\theta} + \sqrt{y}\right) + f_X\left(\frac{1}{\theta} - \sqrt{y}\right)\right]\frac{1}{2\sqrt{y}}, \quad 0 < y < \frac{1}{\theta^2}$$

- Si $y > 1/\theta^2$, la función g tiene una imagen inversa, derivable y estrictamente monótona

  $g^{-1}(y) = \dfrac{1}{\theta} + \sqrt{y} \quad \text{con} \quad \dfrac{dg_1^{-1}(y)}{dy} = \dfrac{1}{2\sqrt{y}} > 0$

  Así

$$f_Y(y) = f_X\left(\frac{1}{\theta}+\sqrt{y}\right)\left|\frac{1}{2\sqrt{y}}\right| = f_X\left(\frac{1}{\theta}+\sqrt{y}\right)\frac{1}{2\sqrt{y}}, \quad y > \frac{1}{\theta^2}$$

Por tanto,

$$f_Y(y) = \begin{cases} 0 & y < 0 \\ \left[f_X\left(\frac{1}{\theta}+\sqrt{y}\right)+f_X\left(\frac{1}{\theta}-\sqrt{y}\right)\right]\frac{1}{2\sqrt{y}} & 0 < y < \frac{1}{\theta^2} \quad (\leftarrow \text{ dos antiimágenes}) \\ f_X\left(\frac{1}{\theta}+\sqrt{y}\right)\frac{1}{2\sqrt{y}} & y > \frac{1}{\theta^2} \quad (\leftarrow \text{ una antiimagen}) \end{cases}$$

$$f_Y(y) = \begin{cases} 0 & y < 0 \\ \frac{\theta}{2\sqrt{y}}\left[\mathrm{e}^{-1-\theta\sqrt{y}}+\mathrm{e}^{-1+\theta\sqrt{y}}\right] & 0 < y < \frac{1}{\theta^2} \\ \frac{\theta}{2\sqrt{y}}\mathrm{e}^{-1-\theta\sqrt{y}} & y > \frac{1}{\theta^2} \end{cases}$$

Se podría haber calculado directamente la función de distribución

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ P[(X-1/\theta)^2 \le y] & y \ge 0 \end{cases}$$

con $P[(X-1/\theta)^2 \le y] = P[-\sqrt{y}+1/\theta \le X \le \sqrt{y}+1/\theta] = F_X(1/\theta+\sqrt{y}) - F_X(1/\theta-\sqrt{y})$. Observemos que si $y > 1/\theta^2$ entonces $F_X(1/\theta-\sqrt{y}) = 0$.

11. Esperanza matemática de una variable aleatoria

Existe una definición global de esperanza matemática aplicable a cualquier tipo de variable aleatoria, que se expresa en términos de integrales de Lebesgue. Aquí daremos la particularización de esta definición para las variables discretas, continuas y mixtas.

Esperanza de una variable aleatoria discreta

Sea $X : (\Omega, \mathcal{A}, P) \longrightarrow (\mathbb{R}, \mathcal{B}, P_X)$ un variable aleatoria discreta con valores en E_X. Se define la esperanza de X como

$$\mathrm{E}X = \sum_{x\in E_X} xP(X=x)$$

siempre que dicha suma sea absolutamente convergente, es decir

$$\sum_{x\in E_X} |x|P(X=x) < \infty.$$

Notemos que si la variable toma un número finito de valores, la suma que define la esperanza tiene un número finito de sumandos y es siempre absolutamente convergente. Por tanto, la esperanza de una variable aleatoria con un número finito de valores siempre existe.

Por el contrario, si E_X es infinito, es necesario imponer la condición de convergencia absoluta de la serie ya que, en caso contrario, distintas reordenaciones de los sumandos podrían dar distintos valores de la esperanza.

Ejemplo: Sea X una variable aleatoria con función masa de probabilidad

$$\mathrm{P}\left[X = (-1)^{j+1}\frac{3^j}{j}\right] = \frac{2}{3^j}, \quad j = 1, 2, \ldots$$

$$\sum_{j=1}^{\infty}\left|(-1)^{j+1}\frac{3^j}{j}\right|\frac{2}{3^j} = \sum_{j=1}^{\infty}\frac{2}{j} = \infty$$

la serie no es absolutamente convergente y, por tanto, X no tiene esperanza.

Sin embargo, notemos que

$$\sum_{j=1}^{\infty}(-1)^{j+1}\frac{3^j}{j}\frac{2}{3^j} = \sum_{j=1}^{\infty}(-1)^{j+1}\frac{2}{j} = 2\left[\sum_{k=0}^{\infty}\frac{(-1)^{2k+2}}{2k+1} + \sum_{k=0}^{\infty}\frac{(-1)^{2k+3}}{2k+2}\right] =$$

$$= 2\left[\sum_{k=0}^{\infty}\frac{1}{(2k+1)(2k+2)}\right] < \infty.$$

Esperanza de una variable aleatoria continua

Sea X un variable aleatoria continua con función de densidad f_X. Se define la esperanza de X como

$$\mathrm{E}X = \int_{\mathbb{R}} x f_X(x)dx$$

siempre que la función $xf_X(x)$ sea absolutamente convergente, es decir

$$\int_{\mathbb{R}} |x| f_X(x)dx < \infty$$

Esperanza de una variable aleatoria mixta

Sea X un variable aleatoria mixta con función de distribución

$$F_X(x) = \sum_{\substack{x_j \le x \\ x_j \in D_X}} \mathrm{P}[X = x_j] + \int_{-\infty}^{x} h(t)dt$$

Se define la esperanza de X como

$$\mathrm{E}X = \sum_{x_j \in D_X} x_j \mathrm{P}[X = x_j] + \int_{\mathbb{R}} x h(x) dx$$

siempre que la suma sea absolutamente convergente y $\int_{\mathbb{R}} |x| h(x) dx < \infty$.

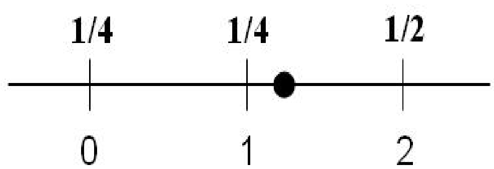
El número EX se denomina también **valor esperado o media** de la variable aleatoria X o de la distribución de X (está determinado por la distribución de la variable).

EJEMPLOS

Ejemplo 1.- Sea X una variable aleatoria con valores en $\{0, 1, 2\}$ y función masa de probabilidad

$$\begin{array}{c} \mathrm{P}[X = 0] = 1/4 \\ \mathrm{P}[X = 1] = 1/4 \\ \mathrm{P}[X = 2] = 1/2 \end{array}$$

$$\mathrm{E}X = 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 2 \cdot \frac{1}{2} = \frac{5}{4}$$



Obsérvese que la media no es el valor central, que sería 1, sino un valor comprendido entre los valores extremos, que no tiene por qué ser un valor de la variable (centro de gravedad).

Ejemplo 2.- Sea X una variable aleatoria con función masa de probabilidad $\mathrm{P}[X = n] = \dfrac{1}{n(n+1)}$, $n = 1, 2, \ldots$

$$\sum_{n=1}^{\infty} |n| \frac{1}{n(n+1)} = \sum_{n=1}^{\infty} n \frac{1}{n(n+1)} = \infty,$$

luego no existe la esperanza de X.

Ejemplo 3.- Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} 1/4 & x \in [1, 5] \\ 0 & \text{en otro caso} \end{cases}$$

$$\int_1^5 |x|f(x)dx = \int_1^5 xf(x)dx = \int_1^5 \frac{x}{4}dx = 3$$

Por tanto, $\exists \mathrm{E}X = 3$.

Ejemplo 4.- Sea X una variable aleatoria con función de densidad

$$f(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x} & x > 0 \\ 0 & x \leq 0 \end{cases} \qquad (\lambda > 0)$$

$$\int_0^\infty |x|\lambda \mathrm{e}^{-\lambda x} = \int_0^\infty x\lambda \mathrm{e}^{-\lambda x} = \lambda \int_0^\infty x\mathrm{e}^{-\lambda x} = [u = x, dv = \mathrm{e}^{-\lambda x}dx] = \lambda \left[-\frac{x\mathrm{e}^{-\lambda x}}{\lambda}\Big|_0^\infty + \int_0^\infty \frac{\mathrm{e}^{-\lambda x}}{\lambda} \right]$$

$$= \lambda \left[0 - \frac{\mathrm{e}^{-\lambda x}}{\lambda^2}\Big|_0^\infty \right] = \lambda \frac{1}{\lambda^2} = \frac{1}{\lambda}$$

Por tanto, $\exists \mathrm{E}X = \dfrac{1}{\lambda}$.

Ejemplo 5.- Sea X una variable aleatoria con función de densidad

$$f(x) = \frac{1}{\pi(1+x^2)} \qquad x \in \mathbb{R}$$

$$\int_{-\infty}^{+\infty} |x|f(x)dx = \frac{1}{\pi}\int_{-\infty}^{+\infty} \frac{|x|}{1+x^2}dx = \frac{2}{\pi}\int_0^{+\infty} \frac{x}{1+x^2}dx = \frac{2}{\pi}\,\frac{\ln(1+x^2)}{2}\Big|_0^\infty = \infty$$

luego $\nexists \mathrm{E}X$.

Ejemplo 6.- Sea X una variable aleatoria con función de distribución

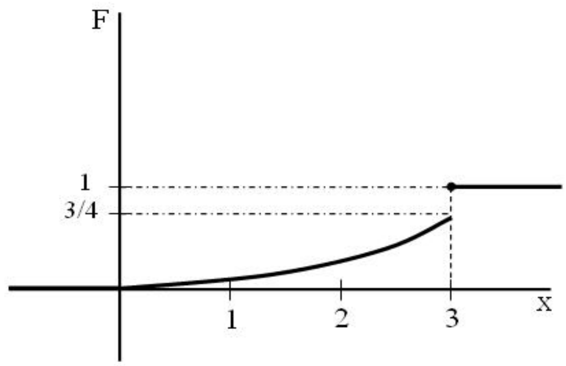
$$F_X(x) = \begin{cases} 0 & x < 0 \\ \dfrac{x}{1+x} & 0 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

Por una parte,

$$D_X = \{3\} \qquad\qquad \mathrm{P}[X = 3] = 1/4$$

y, por tanto, la suma en la expresión de la media consta de un único término. Por otra parte,

$$h_X(x) = \begin{cases} \dfrac{1}{(1+x)^2} & x \in (0,3) \\ 0 & \text{en otro caso} \end{cases}$$



$$\int_0^3 |x|\frac{1}{(1+x)^2} = \int_0^3 x\frac{1}{(1+x)^2} = \frac{1}{2}\int_0^3 \left(\frac{2(1+x)}{(1+x)^2} - \frac{2}{(1+x)^2}\right)dx =$$

$$= \frac{1}{2}\left[\ln(1+x)^2\Big|_0^3\right] - \int_0^3 \frac{1}{(1+x)^2}dx = \frac{1}{2}\ln 16 + \left.\frac{1}{1+x}\right|_0^3 = \ln 4 + \frac{1}{4} - 1 = \ln 4 - 3/4 < \infty$$

Así,

$$\exists \mathrm{E}X = 3\cdot\frac{1}{4} + \int_0^3 x\frac{1}{(1+x)^2} = \ln 4.$$

Esperanza de una función de una variable aleatoria

Sea X una variable aleatoria y $g : (\mathbb{R}, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B})$ una función medible. Entonces, $Y = g(X)$ es una variable aleatoria cuya esperanza, en caso de existir, viene dada por

$$\mathrm{E}Y = \begin{cases} \sum_y y\mathrm{P}[Y=y] & \text{si es discreta} \\ \\ \int y f_Y(y)dy & \text{si es continua} \end{cases}$$

Sin embargo, no será necesario calcular la función de densidad o función masa de probabilidad de la nueva variable para determinar si existe su esperanza y cuanto vale, sino que esto puede hacerse a partir de la distribución de la variable original X, como probamos a continuación en dos casos particulares: X discreta, o bien X continua y g en las condiciones del Teorema de Cambio de Variable (derivable y estrictamente monótona).

Teorema

Sea X una variable aleatoria y $g : (\mathbb{R}, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B})$ una función medible.

a) Si X es discreta

$$\exists \mathrm{E}[g(X)] \iff \sum_{x\in E_X} |g(x)|\mathrm{P}[X=x] < \infty$$

$$\mathrm{E}[g(X)] = \sum_{x \in E_X} g(x)\mathrm{P}[X = x]$$

b) Si X es continua y g en las condiciones del T. C. V. (derivable y estrictamente monótona)

$$\exists \mathrm{E}[g(X)] \iff \int_{\mathbb{R}} |g(x)| f_X(x)dx < \infty$$

$$\mathrm{E}[g(X)] = \int_{\mathbb{R}} g(x) f_X(x)dx$$

Ejemplo.- Sea X el valor obtenido al lanzar un dado. Calcular $\mathrm{E}X$ y $\mathrm{E}X^2$.

$$\mathrm{E}X = \frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2}$$

$$\mathrm{E}X^2 = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6}$$

Observemos que si se quiere calcular la esperanza de muchas funciones de una variable aleatoria ($\mathrm{E}X^3, \mathrm{E}X^4$, ...), no hay que hacer un cambio de variable cada vez. Además, notemos que, en general, $\mathrm{E}X^2 \neq (\mathrm{E}X)^2$

Corolario

$\exists \mathrm{E}X \iff \exists \mathrm{E}|X|$

Propiedades de la esperanza matemática

1. $X = c \implies \exists \mathrm{E}X = c$.

2. $X \geq 0 \implies$ Si $\exists \mathrm{E}X,\ \mathrm{E}X \geq 0$

- DISCRETA $\mathrm{E}X = \sum_{x \in E_X} x\mathrm{P}[X = x] \geq 0$, ya que $x \in E_X \Rightarrow x \geq 0$

- CONTINUA $\mathrm{E}X = \int_{-\infty}^{+\infty} x f_X(x)dx = [f_X(x) = 0, x < 0] = \int_0^{+\infty} x f_X(x)dx \geq 0$

3. $X \geq 0$ y $\mathrm{E}X = 0 \iff P(X = 0) = 1$

$\Leftarrow$) Es evidente.

$\Rightarrow$) Si X es continua con función de densidad f_X, se debe verificar que $f_X(x) = 0,\ \forall x < 0$ y, por tanto

$$\mathrm{E}[X] = \int_0^{+\infty} x f_X(x)\, dx.$$

Dado que el integrando es una función no negativa

$$\mathrm{E}[X] = 0 \;\Leftrightarrow\; x f_X(x) = 0,\ \forall x \geq 0 \;\Leftrightarrow\; f_X(x) = 0,\ \forall x > 0.$$

Así $X \geq 0$ y $\mathrm{E}X = 0 \;\Rightarrow\; f_X(x) = 0,\ \forall x \neq 0$, que está en contradicción con que f_X sea una función de densidad. Por tanto X es discreta. Dado que $\mathrm{E}[X]$ es una suma de cantidades no negativas, el hecho de ser igual a cero implica que todos los términos son nulos, es decir, si $x > 0 \;\Rightarrow\; P[X = x] = 0$. De esto se deduce que sólo el 0 puede tener probabilidad positiva.

4. Si $\exists \mathrm{E}X \implies |\mathrm{E}X| \leq \mathrm{E}|X|$

- DISCRETA $|\mathrm{E}X| = \left|\sum_{x\in E_X} x\mathrm{P}[X = x]\right| \leq \sum_{x\in E_X} |x|\mathrm{P}[X = x] = \mathrm{E}|X|$
- CONTINUA $|\mathrm{E}X| = \left|\int x f_X(x)dx\right| \leq \int |x| f_X(x)dx = \mathrm{E}|X|$

5. Si X está acotada ($\exists M \in \mathbb{R}$ / $\mathrm{P}(|X| \leq M) = 1$) $\implies \exists \mathrm{E}X$ y $|\mathrm{E}X| \leq M$

- DISCRETA $x \in E_X \;\Rightarrow\; |x| \leq M$

$$\sum_{x\in E_X} |x|\mathrm{P}[X = x] \leq M \sum_{x\in E_X} \mathrm{P}[X = x] = M < \infty$$

$$|\mathrm{E}X| \leq \mathrm{E}|X| \leq M$$

- CONTINUA: $|x| > M \;\Rightarrow\; f_X(x) = 0$ y, por tanto $\int_{-M}^{M} f_X(x)\,dx = 1$

$$\int_{-\infty}^{\infty} |x| f_X(x)dx = \int_{-M}^{M} |x| f_X(x)\,dx \leq M \int_{-\infty}^{\infty} f_X(x)dx = M < \infty$$

$$|\mathrm{E}X| \leq \mathrm{E}|X| \leq M$$

6. Linealidad: Si $\exists \mathrm{E}X \implies \forall a, b \in \mathbb{R},\ \exists \mathrm{E}[aX + b] = a\mathrm{E}X + b$

- DISCRETA $\sum_{x\in E_X} |ax + b|\mathrm{P}[X = x] \leq |a| \sum_{x\in E_X} |x|\mathrm{P}[X = x] + |b| < \infty$

$$\sum_{x\in E_X} (ax + b)\mathrm{P}[X = x] = a \sum_{x\in E_X} x\mathrm{P}[X = x] + b = a\mathrm{E}X + b$$

- CONTINUA $\int |ax + b| f_X(x)dx \leq |a| \int |x| f_X(x)dx + |b| < \infty$

$$\int (ax + b) f_X(x)dx = a \int x f_X(x)dx + b = a\mathrm{E}X + b$$

7. Linealidad para funciones de una misma variable: Sean $h_1(X), \dots, h_n(X)$ v. a. tales que $\exists \mathrm{E}[h_i(X)],\ \forall i = 1, \dots, n$. Entonces

$$\exists \mathrm{E}[a_1 h_1(X) + \cdots + a_n h_n(X)] = a_1 \mathrm{E}[h_1(X)] + \cdots + a_n \mathrm{E}[h_n(X)], \forall a_1, \dots, a_n \in \mathbb{R}$$

- DISCRETA

$$\sum_{x \in E_X} |a_1 h_1(x) + \cdots + a_n h_n(x)| \mathrm{P}[X = x] \leq \sum_{x \in E_X} |a_1||h_1(x)| \mathrm{P}[X = x] + \cdots + \sum_{x \in E_X} |a_n||h_n(x)| \mathrm{P}[X = x] < \infty$$

$$\sum_{x \in E_X} (a_1 h_1(x) + \cdots + a_n h_n(x)) \mathrm{P}[X = x] = a_1 \sum_{x \in E_X} h_1(x) \mathrm{P}[X = x] + \cdots + a_n \sum_{x \in E_X} h_n(x) \mathrm{P}[X = x] =$$

$$= a_1 \mathrm{E}[h_1(X)] + \cdots + a_n \mathrm{E}[h_n(X)]$$

- CONTINUA: Análoga

8. Conservación del orden: Sea X una v. a. y h y g funciones medibles tales que $h(X) \leq g(X)$, entonces si $\exists \mathrm{E}[h(X)]$ y $\mathrm{E}[g(X)]$, entonces $\mathrm{E}[h(X)] \leq \mathrm{E}[g(X)$.

Dem.- $(g - h)(X) \geq 0 \implies \exists \mathrm{E}[(g - h)(X)] \geq 0$

12. Momentos de una variable aleatoria

En muchas situaciones, no es fácil determinar completamente la distribución de probabilidad de una variable aleatoria. Sin embargo, sí pueden determinarse diversas características de la distribución que, si bien no proporcionan una descripción completa de la distribución, pueden dar una visión general de la misma. Destacamos en primer lugar los momentos de una distribución, que pueden ser de distintos tipos como vemos a continuación.

Momentos no centrados (centrados respecto al origen u ordinarios)

Sea X una variable aleatoria y k un entero no negativo. Se define el momento no centrado de orden k de X como $\mathrm{E}X^k$, siempre que exista.

$$m_k = \mathrm{E}X^k = \begin{cases} \displaystyle\sum_x x^k \mathrm{P}[X = x] & \text{caso discreto} \\ \\ \displaystyle\int x^k f_X(x) dx & \text{caso continuo} \end{cases}$$

En particular,

$$m_1 = \mathrm{E}X, \qquad m_2 = \mathrm{E}X^2, \qquad m_3 = \mathrm{E}X^3, \dots$$

Momentos centrados (respecto a la media)

Sea X una variable aleatoria, tal que $\exists \mathrm{E}X$ y sea k un entero no negativo. Se define el momento centrado de orden k de X como $\mathrm{E}[(X-\mathrm{E}X)^k]$, siempre que exista.

$$\mu_k = \mathrm{E}[(X-\mathrm{E}X)^k] = \begin{cases} \displaystyle\sum_x (x-\mathrm{E}X)^k \mathrm{P}[X=x] & \text{caso discreto} \\ \\ \displaystyle\int (x-\mathrm{E}X)^k f_X(x)dx & \text{caso continuo} \end{cases}$$

En particular, $\mu_1 = 0$ y el momento centrado de orden 2, si existe, se denomina la **varianza** de la variable

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

y a su raíz cuadrada, la **desviación típica**. Posteriormente, estudiaremos con detenimiento algunas de sus propiedades y su significado.

Momentos absolutos (respecto al origen)

Sea X una variable aleatoria y $\alpha \in \mathbb{R}^+$. Se define el momento absoluto de orden α de X como $\mathrm{E}|X|^\alpha$, siempre que exista.
Observemos que $\exists \mathrm{E}X^k \iff \exists \mathrm{E}|X|^k$

Teorema de existencia de momentos

Si $\exists \mathrm{E}|X|^t$ $(t \in \mathbb{R}^+) \implies \exists \mathrm{E}|X|^s\ \forall s \leq t$
Demostración:

- DISCRETA $\exists \mathrm{E}|X|^s \iff \displaystyle\sum_x |x|^s \mathrm{P}[X=x] < \infty$

$$\begin{aligned} \sum_x |x|^s \mathrm{P}[X=x] &= \sum_{x/|x|\leq 1} |x|^s \mathrm{P}[X=x] + \sum_{x/|x|>1} |x|^s \mathrm{P}[X=x] \leq \\ &\leq \sum_{x/|x|\leq 1} \mathrm{P}[X=x] + \sum_{x/|x|>1} |x|^t \mathrm{P}[X=x] \leq \\ &\leq \mathrm{P}[|X|\leq 1] + \mathrm{E}|X|^t < \infty \end{aligned}$$

- CONTINUA $\exists \mathrm{E}|X|^s \iff \displaystyle\int |x|^s f_X(x)dx < \infty$

$$\begin{aligned} \int_x |x|^s f_X(x)dx &= \int_{x/|x|\leq 1} |x|^s f_X(x)dx + \int_{x/|x|>1} |x|^s f_X(x)dx \leq \\ &\leq \int_{x/|x|\leq 1} f_X(x)dx + \int_{x/|x|>1} |x|^t f_X(x)dx \leq \\ &\leq \mathrm{P}[|X|\leq 1] + \mathrm{E}|X|^t < \infty \end{aligned}$$

Corolario

- $\exists m_k = \mathrm{E}X^k \implies \exists m_j = \mathrm{E}X^j \ \forall j \leq k$
- $\exists \mu_k = \mathrm{E}[(X - m_1)^k] \implies \exists \mu_j = \mathrm{E}[(X - m_1)^j] \ \forall j \leq k$

Demostración:

- $\exists m_k \iff \exists \mathrm{E}|X|^k \implies \exists \mathrm{E}|X|^j \ \ \forall j \leq k \ (\iff \exists m_j \ \forall j \leq k)$
- $\exists \mu_k \iff \exists \mathrm{E}|X - m_1|^k \implies \exists \mathrm{E}|X - m_1|^j \ \ \forall j \leq k \ (\iff \exists \mu_j \ \forall j \leq k)$

Relación entre momentos centrados y no centrados

1) Si $\exists m_k \implies \exists \mu_k = \displaystyle\sum_{n=0}^{k} (-1)^n \binom{k}{n} m_{k-n} m_1^n$

2) Si $\exists \mu_k \implies \exists m_k = \displaystyle\sum_{n=0}^{k} \binom{k}{n} \mu_{k-n} m_1^n$

Demostración:

1) $\mu_k = \mathrm{E}[(X - m_1)^k] = \mathrm{E}\left[\displaystyle\sum_{n=0}^{k} (-1)^n \binom{k}{n} X^{k-n} m_1^n\right]$

   Dado que existe la esperanza de cada sumando (por el corolario), todos funciones de X, existen la esperanza de la suma y se tiene, por la linealidad, la expresión para μ_k.

2) $m_k = \mathrm{E}X^k = \mathrm{E}[((X - m_1) + m_1)^k] = \mathrm{E}\left[\displaystyle\sum_{n=0}^{k} \binom{k}{n} (X - m_1)^{k-n} m_1^n\right]$

   y se tiene la propiedad mediante un razonamiento similar a 1.

Caso particular: $VarX = E[X^2] - (E[X])^2$

13. Desigualdades relacionadas con momentos de una variable aleatoria: Teorema de Markov y Desigualdad de Chebychev

Como ya hemos indicado, en ciertas ocasiones, la distribución de probabilidad de una variable aleatoria no es conocida pero se conocen ciertas características de ella como pueden ser los momentos. En tal caso, como vemos a continuación, podemos calcular cotas para ciertas probabilidades relativas a la variable bajo estudio.

Teorema de Markov (Desigualdad básica)

Si X es una variable aleatoria arbitraria y $h : (\mathbb{R}, \mathcal{B}) \longrightarrow (\mathbb{R}, \mathcal{B})$ una función medible no negativa tal que $\exists \mathrm{E}[h(X)]$, entonces

$$\forall \epsilon > 0 \;\; \mathrm{P}[h(X) \geq \epsilon] \leq \frac{\mathrm{E}[h(X)]}{\epsilon}$$

Demostración:

- DISCRETA

$$\begin{aligned}\mathrm{E}[h(X)] = &\sum_x h(x)\mathrm{P}[X = x] = \sum_{x/h(x)\geq\epsilon} h(x)\mathrm{P}[X = x] + \sum_{x/h(x)<\epsilon} h(x)\mathrm{P}[X = x] \geq \\ &\geq \sum_{x/h(x)\geq\epsilon} h(x)\mathrm{P}[X = x] \geq \epsilon \sum_{x/h(x)\geq\epsilon} \mathrm{P}[X = x] = \epsilon\mathrm{P}[h(X) \geq \epsilon]\end{aligned}$$

- CONTINUA

$$\begin{aligned}\mathrm{E}[h(X)] = &\int h(x)f_X(x)dx = \int_{\{x/h(x)\geq\epsilon\}} h(x)f_X(x)dx + \int_{\{x/h(x)<\epsilon\}} h(x)f_X(x)dx \geq \\ &\geq \int_{\{x/h(x)\geq\epsilon\}} h(x)f_X(x)dx \geq \epsilon \int_{\{x/h(x)\geq\epsilon\}} f_X(x)dx = \epsilon\mathrm{P}[h(X) \geq \epsilon]\end{aligned}$$

Desigualdad de Chebychev

Si $\exists \mathrm{E}X^2 \implies \forall k > 0 \;\; \mathrm{P}[|X - \mathrm{E}X| \geq k] \leq \dfrac{\mathrm{Var}X}{k^2}$

Demostración: Se obtiene de la desigualdad básica aplicada a $(X - \mathrm{E}X)^2$ con $\epsilon = k^2$.

Expresiones alternativas de la desigualdad de Chebychev

- $\mathrm{P}[|X - \mathrm{E}X| < k] \geq 1 - \dfrac{\mathrm{Var}X}{k^2}$
- $\mathrm{P}[|X - \mathrm{E}X| \geq k\sqrt{\mathrm{Var}X}] \leq \dfrac{1}{k^2}$
- $\mathrm{P}[|X - \mathrm{E}X| < k\sqrt{\mathrm{Var}X}] \geq 1 - \dfrac{1}{k^2}$

La desigualdad de Chebychev permite acotar la probabilidad de que una variable (con momento de segundo orden) esté dentro de un intervalo centrado en la media y de longitud $2k$, en términos de la varianza

$$\mathrm{P}[\mathrm{E}X - k < X < \mathrm{E}X + k] \geq 1 - \frac{\mathrm{Var}X}{k^2},$$

o, dentro de un intervalo centrado en la media y de longitud $2k\sqrt{\mathrm{Var}X}$

$$\mathrm{P}[\mathrm{E}X - k\sqrt{\mathrm{Var}X} < X < \mathrm{E}X + k\sqrt{\mathrm{Var}X}] \geq 1 - \frac{1}{k^2}.$$

Así, para $k = 4$, tenemos que la probabilidad de que una variable aleatoria esté comprendida en un intervalo centrado en la media de la variable y de longitud 8 veces la desviación típica es, al menos, 0.9375.
Para $k = 5$, la probabilidad de que una variable esté dentro de un intervalo centrado en su media y de longitud 10 veces la desviación típica es, al menos, 0.96.

NOTA: Observemos que para mayores valores de k, las cotas son más altas (precisas) pero, sin embargo, los intervalos son mayores.

Ejemplo.- Si X es una variable aleatoria con media cero y varianza uno, encontrar un intervalo en el que esté contenida la variable con probabilidad mayor o igual que
1) 0.99
2) 1/3

$$\mathrm{P}[|X| < k] \geq 1 - \frac{1}{k^2}$$

1) $1 - 1/k^2 = 0{,}99 \longrightarrow k = 10 \longrightarrow \mathrm{P}[-10 < X < 10] \geq 0{,}99$
2) $1 - 1/k^2 = 1/3 \longrightarrow k = \sqrt{3/2} \longrightarrow \mathrm{P}[-\sqrt{3/2} < X < \sqrt{3/2}] \geq 1/3$
Entonces, toda variable con media cero y varianza uno toma valores entre -10 y 10 con probabilidad mayor o igual que 0.99, sea cual sea su distribución.

14. Función generatriz de momentos

Los momentos de una variable aleatoria pueden determinarse, en general, sin más que aplicar su definición; sin embargo, en ciertos casos, los cálculos no son simples. Un método alternativo para el cálculo de los momentos es a través de la denominada función generatriz de momentos de una variable aleatoria cuya importancia, además de su relación con los momentos, es que esta función, si existe, determina la distribución de la variable.

Definición
Sea X una variable aleatoria tal que $\exists \mathrm{E}[\mathrm{e}^{tX}]\ \forall t \in (-t_0, t_1),\ (t_0, t_1 \in \mathbb{R}^+)$. Se define su función generatriz de momentos como $M_X : (-t_0, t_1) \longrightarrow \mathbb{R}^+$ dada por

$$M_X(t) = \mathrm{E}\left[\mathrm{e}^{tX}\right], \quad t \in (-t_0, t_1).$$

Como se observa de la definición, la f.g.m. de una variable puede no existir.

Ejemplo 1.- Sea X una variable aleatoria con función masa de probabilidad

$$\mathrm{P}[X=x]=\frac{6}{\pi^2x^2},\qquad x=1,2,\dots$$

$$\mathrm{E}\left[\mathrm{e}^{tX}\right]=\sum_{x=1}^{\infty}\mathrm{e}^{tx}\mathrm{P}[X=x]=\sum_{x=1}^{\infty}\mathrm{e}^{tx}\frac{6}{\pi^2x^2}=\frac{6}{\pi^2}\sum_{x=1}^{\infty}\frac{\mathrm{e}^{tx}}{x^2}$$

Notemos que $\forall t>0$, $\dfrac{\mathrm{e}^{tx}}{x^2}\underset{x\to\infty}{\longrightarrow}\infty$ y, por tanto, la serie diverge, $\forall t>0$ y $\nexists$ la f.g.m..

Ejemplo 2.- Sea X una variable aleatoria con función masa de probabilidad

$$\mathrm{P}[X=x]=\frac{\mathrm{e}^{-\lambda}\lambda^x}{x!},\quad x=0,1,2,\dots\quad(\lambda>0)$$

$$\mathrm{E}\left[\mathrm{e}^{tX}\right]=\sum_{x=0}^{\infty}\mathrm{e}^{tx}\frac{\mathrm{e}^{-\lambda}\lambda^x}{x!}=\mathrm{e}^{-\lambda}\sum_{x=0}^{\infty}\frac{(\lambda\mathrm{e}^t)^x}{x!}=\mathrm{e}^{-\lambda}\mathrm{e}^{\lambda\mathrm{e}^t}=\mathrm{e}^{\lambda(\mathrm{e}^t-1)},\quad\forall t\in\mathbb{R}$$

$$\begin{array}{rccc} M_X: & \mathbb{R} & \longrightarrow & \mathbb{R}^+\\ & t & \longmapsto & \mathrm{e}^{\lambda(\mathrm{e}^t-1)}\end{array}$$

Ejemplo 3.- Sea X una variable aleatoria con función de densidad

$$f_X(x)=\frac{1}{2}\mathrm{e}^{-x/2},\quad x>0$$

$$\begin{aligned}\mathrm{E}\left[\mathrm{e}^{tX}\right]&=\int_0^{\infty}\mathrm{e}^{tx}\frac{1}{2}\mathrm{e}^{-x/2}dx=\frac{1}{2}\int_0^{\infty}\mathrm{e}^{x(t-1/2)}dx=\\&=\frac{1}{2}\left[\frac{\mathrm{e}^{x(t-1/2)}}{t-1/2}\right]_0^{\infty}=\begin{cases}+\infty & t>1/2\ (\nexists\mathrm{E}[\mathrm{e}^{tX}])\\ \frac{1}{2}\frac{1}{1/2-t}=\frac{1}{1-2t} & t<1/2\end{cases}\end{aligned}$$

Entonces

$$\begin{array}{rccc} M_X: & (-\infty,1/2) & \longrightarrow & \mathbb{R}^+\\ & t & \longmapsto & \dfrac{1}{1-2t}\end{array}$$

Propiedades de la función generatriz de momentos

- **Teorema de Unicidad**

  La f. g. m. de una variable aleatoria, si existe, determina de forma única la distribución de la variable.

- **Relación con los momentos**

  Sea X una variable aleatoria con f. g. m. $M_X(t),\ t \in (-t_0, t_1)$. Entonces

  1) $\exists \mathrm{E}\left[X^k\right], \quad \forall k \in \mathbb{N}$

  2) $M_X(t) = \displaystyle\sum_{k=0}^{\infty} \frac{t^k}{k!} \mathrm{E}\left[X^k\right], \quad t \in (-t_0, t_1)$

  3) $\mathrm{E}\left[X^k\right] = M_X^{(k)}(t)\Big|_{t=0} = \dfrac{d^k M_X(t)}{dt^k}\Bigg|_{t=0}$

Ejemplo 4. Sea X una variable aleatoria con función de densidad

$$f_X(x) = \frac{1}{2} \mathrm{e}^{-x/2}, \quad x > 0$$

Calcular la media y la varianza de X a partir de su f. g. m.

Según se ha visto en el Ejemplo 3 anterior

$$M_X(t) = \frac{1}{1-2t} \qquad t < 1/2$$

$$\frac{dM_X(t)}{dt} = \frac{2}{(1-2t)^2} \longrightarrow \frac{dM_X(t)}{dt}\bigg|_{t=0} = \mathrm{E}X = 2$$

$$\frac{d^2M_X(t)}{dt^2} = \frac{8(1-2t)}{(1-2t)^4} = \frac{8}{(1-2t)^3} \longrightarrow \frac{d^2M_X(t)}{dt^2}\bigg|_{t=0} = \mathrm{E}X^2 = 8 \longrightarrow \mathrm{Var}X = 4$$

El recíproco del resultado anterior no es cierto; esto es, puede que existan todos los momentos de una variable aleatoria y no exista la f.g.m. como probamos en el siguiente ejemplo.

Ejemplo 5 (Existencia de momentos no implica la existencia de f. g. m.)

Sea X una variable aleatoria con función de densidad

$$f_X(x) = \frac{1}{\alpha\Gamma(\alpha)} \mathrm{e}^{-x^{1/\alpha}}, \quad x > 0 \quad (\alpha > 1)$$

$$\exists \mathrm{E}X^k \iff \int_0^{\infty} \frac{1}{\alpha\Gamma(\alpha)} x^k \mathrm{e}^{-x^{1/\alpha}} dx < \infty$$

Dicha integral, haciendo el cambio $y = x^{1/\alpha} \to x = y^{\alpha} \to dx = \alpha y^{\alpha-1} dy$, queda

$$\int_0^{\infty} \frac{1}{\alpha\Gamma(\alpha)} y^{k\alpha} \mathrm{e}^{-y} \alpha y^{\alpha-1} dy = \frac{1}{\Gamma(\alpha)} \int_0^{\infty} e^{-y} y^{k\alpha+\alpha-1} dy = \frac{\Gamma(\alpha+k\alpha)}{\Gamma(\alpha)} < \infty \quad \forall k \in \mathbb{N}$$

Por otra parte

$$\exists M_X(t) \iff \int_0^{\infty} \mathrm{e}^{tx} f_X(x) dx < \infty \quad \forall t \in (-t_0, t_1)$$

$$\int_0^{\infty} \mathrm{e}^{tx} f_X(x) dx = \frac{1}{\alpha\Gamma(\alpha)} \int_0^{\infty} \mathrm{e}^{x(t - x^{1/\alpha - 1})} dx$$

Puesto que

$$\frac{1}{\alpha} - 1 < 0 \implies x^{1/\alpha - 1} \underset{x \to +\infty}{\longrightarrow} 0 \implies t - x^{1/\alpha-1} \underset{x \to +\infty}{\longrightarrow} t$$

Entonces, si $t > 0$ y $s \in (0, t)$, $\exists x_0$ tal que $x > x_0 \longrightarrow t - x^{1/\alpha-1} > s$ y, por tanto, $\mathrm{e}^{x(t-x^{1/\alpha-1})} > \mathrm{e}^{sx}$ y

$$\int_0^{\infty} \mathrm{e}^{x(t-x^{1/\alpha-1})} dx \geq \int_{x_0}^{\infty} \mathrm{e}^{x(t-x^{1/\alpha-1})} dx > \int_{x_0}^{\infty} \mathrm{e}^{sx} dx = \infty$$

15. Características numéricas asociadas a una distribución

Además de los momentos, existen otras características numéricas asociadas a la distribución de una variable aleatoria, que proporcionan una descripción general de la misma y permiten, en ciertos casos, establecer comparaciones entre distintas distribuciones. Aquí estudiaremos algunas de estas características clasificadas en

- Medidas de posición.
- Medidas de dispersión.
- Medidas de forma.

Medidas de posición o localización
Indicadores de la posición o localización de la distribución de probabilidad en $\mathbb{R}$.

MEDIA $\mathrm{E}X$

Es la medida de posición central por excelencia aunque tiene el inconveniente de que puede no existir y es muy sensible a cambios en valores de la variable aunque éstos sean de probabilidad muy pequeña.

$$\mathrm{P}[X=0]=1/4, \quad \mathrm{P}[X=1]=1/2, \quad \mathrm{P}[X=2]=1/5, \quad \mathrm{P}[X=3]=1/20, \quad \mathrm{E}X=1{,}05$$

$$\mathrm{P}[Y=0]=1/4, \quad \mathrm{P}[Y=1]=1/2, \quad \mathrm{P}[Y=2]=1/5, \quad \mathrm{P}[Y=1000]=1/20, \quad \mathrm{E}Y=50{,}9$$

Así, el valor de la media es indicativo de la posición en poblaciones homogéneas.

<u>MEDIANA</u> Me_X

Me_X es un valor real tal que $\mathrm{P}[X \leq \mathrm{Me}_X] \geq 1/2$ y $\mathrm{P}[X \geq \mathrm{Me}_X] \geq 1/2$ o, equivalentemente,

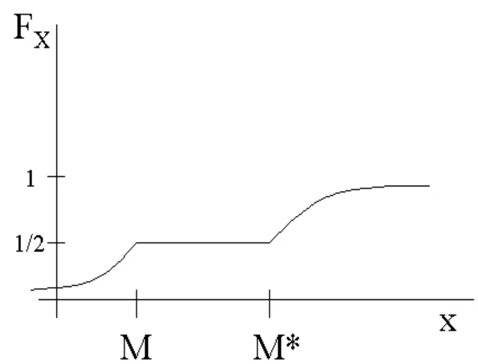
$$F_X(Me_X^-) \leq 1/2 \leq F_X(Me_X).$$

La mediana divide a la distribución en dos partes en el sentido de que la probabilidad a la izquierda y a la derecha es $\geq 1/2$.

Nótese que si X es una variable aleatoria de tipo continuo, la mediana es un valor que satisface

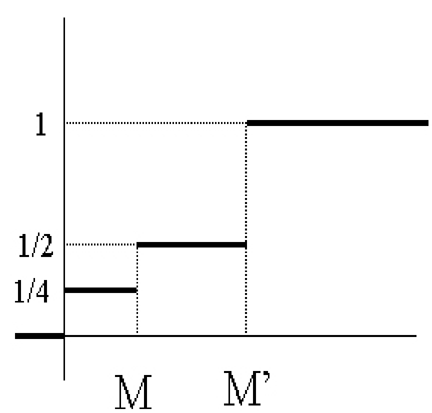
$$F_X(\mathrm{Me}_X) = 1/2$$

• Así, la mediana de una distribución no tiene por qué ser única



En este caso todos los puntos de $[M, M^*]$ son medianas.

También, en distribuciones discretas puede no haber unicidad de la mediana.

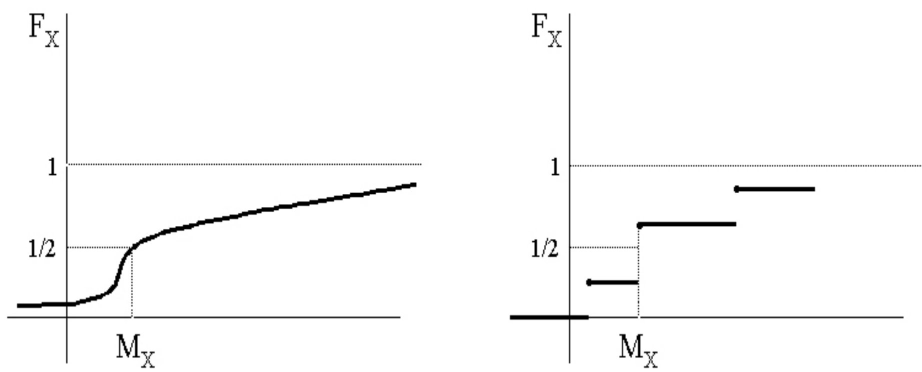


También en este caso todos los puntos de $[M, M']$ son medianas.

Y, en general, si M y M^* son medianas, todos los puntos de $[M, M^*]$ lo son

$$m \in (M, M^*) \implies \begin{cases} \mathrm{P}[X \leq m] \geq \mathrm{P}[X \leq M] \geq 1/2 \\ \\ \mathrm{P}[X \geq m] \geq \mathrm{P}[X \geq M^*] \geq 1/2 \end{cases}$$

• La mediana es única en casos como los siguientes o similares



• Como principal ventaja respecto de la media, además de que siempre existe, debemos indicar que la mediana es menos sensible a cambios en valores de la variable con masa de probabilidad pequeña

$\mathrm{P}[X = 0] = 1/4, \quad \mathrm{P}[X = 1] = 1/2, \quad \mathrm{P}[X = 2] = 1/5, \quad \mathrm{P}[X = 3] = 1/20, \quad \mathrm{Me}_X = 1$

$\mathrm{P}[Y = 0] = 1/4, \quad \mathrm{P}[Y = 1] = 1/2, \quad \mathrm{P}[Y = 2] = 1/5, \quad \mathrm{P}[Y = 1000] = 1/20, \quad \mathrm{Me}_Y = 1$

MODA

Es un número real que maximiza la función masa de probabilidad (valor más probable) en el caso de variables discretas, o la función de densidad, para variables continuas.

La moda puede no ser única. Por ejemplo

- $\mathrm{P}[X = 0] = 1/3$, $\mathrm{P}[X = 1] = 1/3$, $\mathrm{P}[X = 2] = 1/6$, $\mathrm{P}[X = 3] = 1/6$ y son modas el 0 y el 1.
- $f_X(x) = \begin{cases} 1/3 & 0 < x < 1 \\ 2/3 & 1 < x < 2 \end{cases}$ y son modas todos los valores del intervalo $(1, 2)$

PERCENTILES x_q

Son medidas de posición que generalizan a la mediana.

Se define el cuantil o percentil de orden q $(0 \leq q \leq 1)$ como un valor x_q tal que $\mathrm{P}[X \leq x_q] \geq q$ y $\mathrm{P}[X \geq x_q] \geq 1 - q$ o, equivalentemente,

$$F_X(x_q^-) \leq q \leq F_X(x_q).$$

Así, un cuantil de orden q divide a la distribución en dos partes, quedando a la izquierda una masa de probabilidad $\geq q$ y a la derecha una masa de probabilidad $\geq 1-q$.

Para distribuciones continuas, se cumple

$$F_X(x_q) = q$$

- Si $q = 1/2$, entonces $x_q = Me_X$
- Siempre existe un cuantil de cualquier orden, aunque no tiene por qué ser único (igual que la mediana)

- $x_{0,25} = Q_1 \quad \longrightarrow \quad 1^{\underline{er}}$ Cuartil
- $x_{0,5} = Q_2 \quad \longrightarrow \quad 2^{\underline{o}}$ Cuartil (Mediana)
- $x_{0,75} = Q_3 \quad \longrightarrow \quad 3^{\underline{er}}$ Cuartil

Medidas de dispersión

Las medidas de posición por sí solas no son suficiente para dar una descripción adecuada de la distribución de una variable aleatoria ya que no reflejan la dispersión de los valores de la variable respecto de ellas.

$\mathrm{P}[X=-1] = 1/4, \quad \mathrm{P}[X=0] = 1/2, \quad \mathrm{P}[X=1] = 1/4 \quad \longrightarrow \quad \mathrm{E}X = 0, \quad Me_X = 0$

$\mathrm{P}[Y=-100] = 1/4, \quad \mathrm{P}[Y=0] = 1/2, \quad \mathrm{P}[Y=100] = 1/4 \quad \longrightarrow \quad \mathrm{E}Y = 0, \quad Me_Y = 0$

Para dar un indicador de la representatividad de las medidas de posición (entendiéndose por ello la menor o mayor desviación de la distribución respecto de ellas) se definen las medidas de dispersión.

VARIANZA

Es el momento centrado de orden dos

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

Mide la desviación cuadrática media de la variable respecto de su media. La varianza proporciona una medida de la variación o dispersión de una distribución alrededor de su media. Un valor pequeño de la varianza indica que la distribución de probabilidad está muy concentrada alrededor de la media y un valor grande de la varianza generalmente indica que la distribución de probabilidad tiene una dispersión amplia alrededor de la media. Veamos a continuación algunas propiedades que justifican su uso como medida de dispersión.

Propiedades de la varianza

1) $\mathrm{Var}X \geq 0$ y $\mathrm{Var}X = 0 \iff \exists c \in \mathbb{R} \;/\; \mathrm{P}[X=c] = 1$

    - $\mathrm{Var}X \geq 0$ por ser la esperanza de una variable aleatoria no negativa.

- $\Longleftarrow]$ $\mathrm{P}[X=c]=1 \implies \mathrm{P}[(X-c)^2=0]=1 \implies \mathrm{E}X=c$ y $\mathrm{E}[(X-c)^2]=\mathrm{Var}X=0$
- $\Longrightarrow]$ $\mathrm{Var}X=\mathrm{E}[(X-\mathrm{E}X)^2]=0$ y $\mathrm{P}[(X-\mathrm{E}X)^2\geq 0]=1 \implies \mathrm{P}[(X-\mathrm{E}X)^2=0]=1 \implies \mathrm{P}[X=\mathrm{E}X]=1$

2) $\mathrm{Var}(aX+b)=a^2\mathrm{Var}X, \quad \forall a,b\in\mathbb{R}$

$$\mathrm{Var}(aX+b)=\mathrm{E}[(aX+b-a\mathrm{E}X-b)^2]=a^2\mathrm{Var}X$$

2.1. $\mathrm{Var}(X+b)=\mathrm{Var}X$ es invariante frente a cambios de origen en las unidades de medida (desplazamiento de la distribución).

Este resultado es intuitivo puesto que desplazando la distribución de X una distancia de b unidades a lo largo de la recta real, la media de la distribución variará en b unidades, pero el desplazamiento no afectará a la dispersión de la distribución alrededor de su media.

2.2. $\mathrm{Var}X=\mathrm{Var}(-X)$

3) $\mathrm{Var}X<\mathrm{E}[(X-c)^2], \quad \forall c\neq \mathrm{E}X$. Esto indica que si se mide la dispersión de una variable por la desviación cuadrática media alrededor de un punto c arbitario, ésta es mínima si $c=\mathrm{E}X$.

$$\mathrm{E}[(X-c)^2]=\mathrm{E}[(X-\mathrm{E}X+\mathrm{E}X-c)^2=\mathrm{Var}X+(\mathrm{E}X-c)^2+2\,\underbrace{\mathrm{E}[(X-\mathrm{E}X)]}_{=0}\,[\mathrm{E}X-c]$$

Luego,

$\mathrm{E}[(X-c)^2]=\mathrm{Var}X+(\mathrm{E}X-c)^2, \quad \forall c\in\mathbb{R}$

y queda probado el resultado.

Nota.- Para $c=0$, se obtiene $\mathrm{Var}X=\mathrm{E}X^2-(\mathrm{E}X)^2$ (expresión ya obtenida a partir de la relación entre los momentos)

• Hay que indicar que la varianza de una variable se mide en unidades cuadradas, por lo que es común usar como medida de dispersión su raíz cuadrada o DESVIACIÓN TÍPICA, expresada en la misma unidad de medida que la variable.

• La varianza, y por tanto la desviación típica, es muy influenciable por valores extremos. Tal y como ocurría para la esperanza, la varianza de cualquier distribución se puede hacer arbitrariamente grande colocando una masa de probabilidad positiva, aunque sea pequeña, suficientemente lejos del "centro"de los datos.

TIPIFICACIÓN: Dada una variable aleatoria X siempre se puede obtener otra, Z, a partir de ella, con media cero y varianza uno,

$$Z=\frac{X-E[X]}{\sqrt{VarX}}$$

A este proceso se le llama **tipificación** de la variable aleatoria X y a la nueva variable aleatoria Z, **variable tificada**.

DESVIACIÓN ABSOLUTA MEDIA RESPECTO DE LA MEDIANA

$$\mathrm{E}|X - Me_X|$$

Su uso queda justificado por la siguiente propiedad

$$\mathrm{E}[|X - Me_X|] < \mathrm{E}[|X - c|] \quad \forall c \neq Me_X$$

RANGO O RECORRIDO: Diferencia entre el mayor y el menor valor de la variable aleatoria.

RANGO O RECORRIDO INTERCUARTÍLICO $Q_3 - Q_1$

Es la medida de dispersión usual asociada a los cuartiles. La ventaja de esta frente a la anterior es que disminuye los efectos distorsionantes que pueden provocar en el rango los valores extremos.

Todas la anteriores medidas de dispersión son absolutas en el sentido de que dependen de la unidad de medida de la variable y no son apropiadas, por tanto, para comparar distribuciones medidas en escalas diferentes. Además, una dispersión medida, por ejemplo, por una desviación típica igual a 10 puede tener significado distinto (es pequeña o grande) en función del valor de la media (es mucho menor una desviación típica de 10 en una distribución de media 1000 que en una de media 20).

Por ello, para comparar dos distribuciones, debe usarse una medida de dispersión relativa, que tiene como ventaja el no verse afectada por cambios en las unidades de medida

COEFICIENTE DE VARIACIÓN (DE PEARSON)

Se define para variables aleatorias con $E[X] \neq 0$ como

$$CV_X = \frac{\sigma_X}{\mathrm{E}X}$$

- Es adimensional
- $|CV_{aX}| = |CV_X|$

$$CV_{aX} = \frac{\sigma_{aX}}{|\mathrm{E}[aX]|} = \frac{\sqrt{\mathrm{Var}[aX]}}{|a\mathrm{E}X|} = \frac{|a|\sqrt{\mathrm{Var}X}}{|a||\mathrm{E}X}| = \pm CV_X$$

Ejemplo

$\mathrm{E}X = 100$ y $\sigma_X^2 = 25 \quad \longrightarrow \quad CV_X = \frac{5}{100} = 0{,}05$
$\mathrm{E}Y = 30$ y $\sigma_Y^2 = 16 \quad \longrightarrow \quad CV_Y = \frac{4}{30} = 0{,}13$

$$\sigma_X^2 > \sigma_Y^2 \;\text{ y, sin embargo, }\; CV_X < CV_Y$$

O sea, la distribución de X tiene una menor dispersión relativa que la de X aunque la dispersión absoluta sea mayor.

Medidas de forma

Para describir una distribución son también interesantes ciertas medidas de forma, que proporcionan indicadores de la simetría de la función masa de probabilidad o de la función de densidad de una variable aleatoria y del apuntamiento (aplastamiento) de dichas funciones.

SIMETRÍA

Definición.- Una variable aleatoria (distribución) es **simétrica** alrededor de un punto $\alpha \in \mathbb{R}$ si

$$\forall x \in \mathbb{R} \qquad \mathrm{P}[X \leq \alpha - x] = \mathrm{P}[X \geq \alpha + x]$$

o equivalentemente, las variables aleatorias $X - \alpha$ y $\alpha - X$ tienen la misma distribución.

• Si X es simétrica alrededor del origen, se dirá simplemente <u>simétrica</u>.

Definición para variables discretas. X discreta, es simétrica alrededor de $\alpha \in \mathbb{R}$ si y sólo si

$$\mathrm{P}[X = \alpha - x] = \mathrm{P}[X = \alpha + x] \;\; \forall x \in \mathbb{R}.$$

Demostración: Inmediata de la definición.

Definición para variables continuas. X continua, es simétrica alrededor de $\alpha \in \mathbb{R}$ si y sólo si

$$f_X(\alpha - x) = f_X(\alpha + x) \qquad \forall x \in \mathbb{R}.$$

Demostración
$\Longrightarrow$] En el caso continuo, la definición de simetría equivale a $F_X(\alpha+x) = 1-F_X(\alpha-x), \forall x \in \mathbb{R}$ y derivando se obtiene la condición anterior.

$\Longleftarrow$] $f_X(\alpha - x) = f_X(\alpha + x) \quad \forall x \in \mathbb{R} \implies \forall t \in \mathbb{R} \quad \int_{-\infty}^{t} f_X(\alpha - x)dx = \int_{-\infty}^{t} f_X(\alpha + x)dx$

• $\displaystyle\int_{-\infty}^{t} f_X(\alpha - x)dx = \left[\begin{array}{c} y = \alpha - x \\ dy = -dx \end{array} \right] = -\int_{+\infty}^{\alpha - t} f_X(y)dy = \int_{\alpha - t}^{+\infty} f_X(y)dy = 1 - F_X(\alpha - t)$

• $\displaystyle\int_{-\infty}^{t} f_X(\alpha + x)dx = \left[\begin{array}{c} y = \alpha + x \\ dy = dx \end{array} \right] = \int_{-\infty}^{\alpha + t} f_X(y)dy = F_X(\alpha + t)$

Propiedades de las distribuciones simétricas

1) Si X es simétrica alrededor de $\alpha \in \mathbb{R}$, todos los momentos centrados en α de orden impar, si existen, son nulos.

   Dem.- Por la definición de simetría $\mathrm{E}[(X-\alpha)^{2k+1}] = \mathrm{E}[(\alpha - X)^{2k+1}]$ y ya que $\mathrm{E}[(\alpha - X)^{2k+1}] = -\mathrm{E}[(X-\alpha)^{2k+1}]$ se tiene la propiedad.

2) Si X es simétrica alrededor de un punto $\alpha \in \mathbb{R}$ y $\exists \mathrm{E}X$, entonces $\mathrm{E}X = \alpha$.

   Dem.- Consecuencia inmediata de 1).

3) Si X es simétrica alrededor de un punto $\alpha \in \mathbb{R}$, α es Mediana.

   Dem.- Haciendo $x = 0$ en la condición de simetría: $\mathrm{P}[X \leq \alpha] = \mathrm{P}[X \geq \alpha]$. Entonces, esa probabilidad no puede ser menor de 1/2, porque, en tal caso,

$$\mathrm{P}[X \leq \alpha] + \mathrm{P}[X \geq \alpha] < 1 \quad \text{absurdo}$$

4) Si X es simétrica alrededor de un punto $\alpha \in \mathbb{R}$ y es unimodal, Moda$= \alpha$.

A continuación definimos un coeficiente que proporciona un indicador de la asimetría de una distribución.

Coeficiente de asimetría de Fisher

$$\frac{\mu_3}{\sigma_X^3}$$

El uso de este coeficiente se basa en el hecho de que en distribuciones simétricas todos los momentos de orden impar y, en particular, μ_3, son cero. Entonces, una distribución será más asimétrica cuanto mayor sea $|\mu_3|$, esto es, cuanto menos se compensen las desviaciones $(X - \mathrm{E}X)^3$ (se usa μ_3 porque μ_1 siempre vale cero).

Con objeto de obtener una medida adimensional, se divide por σ_X^3. Con ello se consigue, además, que el coeficiente sea invariante (realmente, su valor absoluto) frente a cambios de escala y origen en las unidades de medida de la variable:

$$X \longrightarrow \frac{\mu_3}{\sigma_X^3}$$

$$aX+b \longrightarrow \frac{\mathrm{E}[(aX+b-(a\mathrm{E}X+b))^3]}{\left(\sqrt{\mathrm{Var(aX+b)}}\right)^3} = \frac{a^3\mu_3}{\left(\sqrt{a^2\mathrm{Var}X}\right)^3} = \frac{a^3\mu_3}{|a|^3\sigma_X^3} = \pm\frac{\mu_3}{\sigma_X^3}$$

Ejemplo. $\mathrm{P}[X=-1]=1/4, \quad \mathrm{P}[X=0]=1/2, \ \ \mathrm{P}[X=3]=1/4$

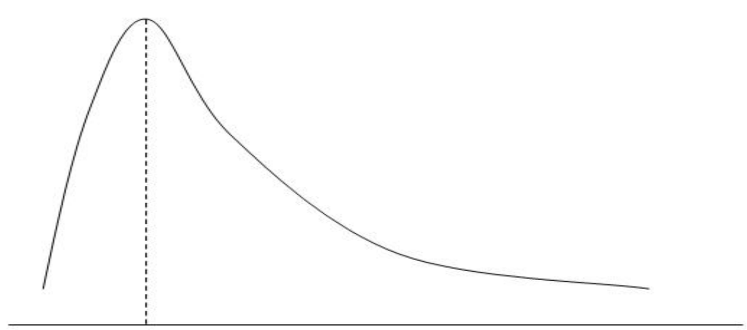
$\mathrm{E}X=1/2$
$\mu_3=\mathrm{E}[(X-1/2)^3]=3$
$\mathrm{Var}X=\mathrm{E}[(X-1/2)^2]=17/8 \quad \longrightarrow \quad \sigma_x=\sqrt{17/8}$

$$\frac{\mu_3}{\sigma_X^3}=\frac{3}{(\sqrt{17/8})^3}>0$$

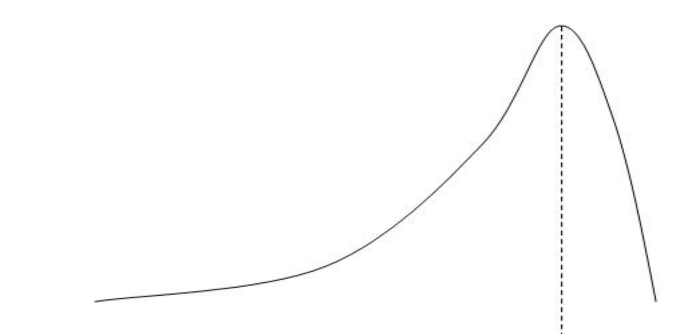
Observamos que la función masa de probabilidad es asimétrica por la derecha: las desviaciones a la derecha de $\mathrm{E}X$ están más cargadas que las desviaciones a la izquierda: $(3-\mathrm{E}X)$ más carga que $(\mathrm{E}X-0)$ y $(\mathrm{E}X+1)$, desviaciones a la izquierda.

En general:

- Distribuciones simétricas $\quad \longrightarrow \quad \dfrac{\mu_3}{\sigma_X^3}=0$
- Distribuciones asimétricas a la derecha $\quad \longrightarrow \quad \dfrac{\mu_3}{\sigma_X^3}>0$



- Distribuciones asimétricas a la izquierda $\quad \longrightarrow \quad \dfrac{\mu_3}{\sigma_X^3}<0$



CURTOSIS (APUNTAMIENTO)

Para distribuciones unimodales, moderadamente asimétricas y campaniformes, Fisher dio una medida del apuntamiento de la distribución que indica mayor o menor concentración de la variable respecto de su media.

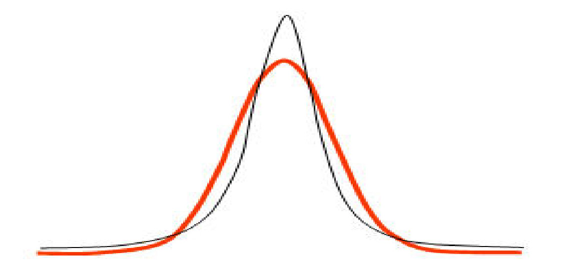
Esta medida se define en relación a una distribución patrón, la distribución normal (campana de Gauss), mediante el siguiente coeficiente

• **Coeficiente de curtosis de Fisher**

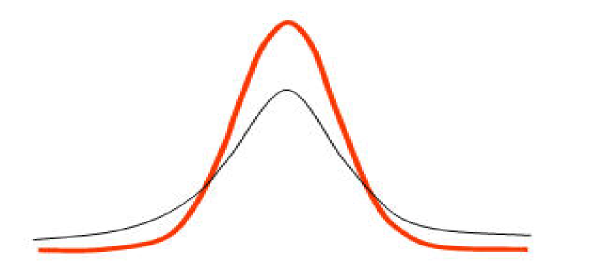
$$\frac{\mu_4}{\sigma_X^4} - 3$$

que vale cero para la distribución patrón.

- $\frac{\mu_4}{\sigma_X^4} - 3 = 0 \quad \longrightarrow$ DISTRIBUCIÓN MESOCÚRTICA
- $\frac{\mu_4}{\sigma_X^4} - 3 > 0 \quad \longrightarrow$ DISTRIBUCIÓN LEPTOCÚRTICA (Más apuntada (menos aplastada) que la normal $\rightarrow$ Más concentración)



$\frac{\mu_4}{\sigma_X^4} - 3 < 0 \quad \longrightarrow$ DISTRIBUCIÓN PLATICÚRTICA (Menos apuntada (más aplastada) que la normal $\rightarrow$ Menos concentración)



• No está afectado por cambios de escala y origen

$$X \longrightarrow \frac{\mu_4}{\sigma_X^4} - 3$$

$$aX+b \longrightarrow \frac{\mathrm{E}[((aX+b)-(a\mathrm{E}X+b))^4]}{(\mathrm{Var}(aX+b))^2} - 3 = \frac{a^4\mu_4}{a^4\mathrm{Var}X} - 3$$

Ejemplo

- $\mathrm{P}[X=-1]=1/4, \quad \mathrm{P}[X=0]=1/2, \quad \mathrm{P}[X=1]=1/4$

$$\begin{array}{l} \mathrm{E}X = 0 \\ \mathrm{E}X^2 = 1/2 \ \rightarrow \ \mathrm{Var}X = 1/2 \\ \mu_4 = 1/2 \end{array}$$

$$\frac{\mu_4}{\sigma_X^4} - 3 = \frac{1/2}{(1/2)^2} - 3 = -1 < 0 \quad \longrightarrow \quad \text{Platicúrtica (menos apuntada que la normal)}$$

- $\mathrm{P}[X=-1]=1/5, \quad \mathrm{P}[X=0]=3/5, \quad \mathrm{P}[X=1]=1/5$

$$\begin{array}{l} \mathrm{E}X = 0 \\ \mathrm{E}X^2 = 2/5 \ \rightarrow \ \mathrm{Var}X = 2/5 \\ \mu_4 = 2/5 \end{array}$$

$$\frac{\mu_4}{\sigma_X^4} - 3 = \frac{2/5}{(2/5)^2} - 3 = 5/2 - 3 = -0.5 < 0 \quad \longrightarrow \quad \text{Platicúrtica, pero menos que la anterior.}$$