2

Características de una variable aleatoria

En este capítulo se estudian algunas características numéricas de las variables aleatorias. Dichas características o parámetros de la distribución juegan un papel importante en la descripción de una variable aleatoria, además de ser fundamentales en el desarrollo de la inferencia estadística.

Las características de una variable aleatoria se agrupan en tres tipos de medidas: medidas de centralización, dispersión y forma. Además, en este Capítulo se introducen los momentos de una variable aleatoria y la desigualdad de Tchebychev que será de gran utilidad para la obtención de cotas de probabilidad.

2.1. Medidas de centralización

A continuación, se presentan algunas medidas de centralización como son la esperanza matemática, los cuantiles, la mediana (como un caso particular de los cuantiles) y la moda.

La definición de las medidas de centralización es diferente dependiendo del tipo de v.a. que se considere: v.a. discreta con función de masa y v.a. continua con función de densidad. Por tanto, cada medida se define de distinta forma dependiendo de si la v.a. es discreta o continua.

2.1.1. Esperanza matemática

La *esperanza matemática* se interpreta como el valor medio de la distribución teórica de probabilidades del fenómeno estudiado, es decir, el valor hacia el que tiende la media aritmética, estudiada en estadística descriptiva, si se tiene un número suficientemente grande de observaciones.

Por tanto, la esperanza matemática esta acotada entre el mínimo y el máximo de los valores que toma la v.a. y representa el punto de equilibrio o centro de gravedad de una distribución de probabilidad.

La esperanza matemática se denota como $E[X] = \mu$.

La definición de la esperanza matemática depende del tipo de v.a. que se considere:

- *Variable aleatoria discreta*

  Sea X una v.a. discreta con puntos de masa $\{x_i\}$ y función de masa $p_i = P(X = x_i)$ donde $i \in \mathbb{N}$. Se define la esperanza matemática de la v.a. X como

$$E[X] = \sum_{i=1}^{\infty} x_i P(X = x_i) = \sum_{i=1}^{\infty} x_i p_i$$

  Si el dominio de la v.a. X está formado por un conjunto infinito numerable, la esperanza matemática es una serie infinita que puede ser convergente o divergente. En este caso, la esperanza matemática de la v.a. X existe, si la serie que se presenta a continuación converge absolutamente

$$\sum_{i=1}^{\infty} |x_i| \, p_i < \infty$$

  Si el domino de la v.a. X es finito, la esperanza matemática existe siempre, ya que se obtiene como la suma de un número finito de valores.

  Para el caso discreto, la esperanza matemática no tiene porqué ser uno de los puntos de masa de la v.a. X, ya que representa un valor medio de la distribución de probabilidad de X.

  Cuando todos los puntos de masa $\{x_1, .., x_N\}$ de una v.a. discreta X tienen la misma probabilidad, $P(X = x_i) = \frac{1}{N}$, la esperanza matemática de X

coincide con la media aritmética, tal que

$$E[X] = \sum_{i=1}^{N} x_i P(X = x_i) = \frac{1}{N} \sum_{i=1}^{N} x_i = \bar{x}$$

Ejemplo 2.1
Sea X una v.a. discreta con puntos de masa $\{-2, 0, 1, 3\}$ y función de masa $P(X = -2) = 0.2, P(X = 0) = 0.5, P(X = 1) = 0.1$ y $P(X = 3) = 0.2$. Calcular la esperanza matemática de la v.a. X.

$E[X] = \sum_{i=1}^{4} x_i P(X = x_i) = -2 \cdot 0.2 + 0 \cdot 0.5 + 1 \cdot 0.1 + 3 \cdot 0.2 = 0.3$

- *Variable aleatoria continua*

Sea X una v.a. continua con función de densidad $f(x)$. Se define la esperanza matemática de X como

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx$$

Para que exista la esperanza matemática de una v.a. continua se debe de cumplir que la integral que se muestra a continuación sea convergente absolutamente, es decir

$$\int_{-\infty}^{\infty} |x| f(x) dx < \infty$$

Si X es una v.a. acotada, de forma que $P(a \leq X \leq b) = 1$, la esperanza matemática de X existe siempre.

Sea X una v.a. con función de densidad simétrica respecto a un cierto valor $c \in \mathbb{R}$, entonces si existe $E[X]$, se verifica que $E[X] = c$.

Ejemplo 2.2
Sea X una v.a. continua con función de densidad

$$f(x) = \begin{cases} \dfrac{8}{7x^2} & si \quad 1 \leq x \leq 8 \\ 0 & en \quad otro\ caso \end{cases}$$

Calcular la esperanza matemática de la v.a. X.

$E[X] = \int_1^8 x f(x) dx = \int_1^8 x \frac{8}{7x^2} dx = \frac{8}{7} \int_1^8 \frac{1}{x} dx = \frac{8}{7} [\ln x]_1^8 = 2.376$

Propiedades de la esperanza matemática

1. Sea $Y = h(X)$ una transformación de la v.a. X. Se define la esperanza matemática de dicha transformación como

   - *Variable aleatoria discreta*

   $$E[Y] = E[h(X)] = \sum_{i=1}^{\infty} h(x_i) P(X = x_i) = \sum_{i=1}^{\infty} h\left(x_i\right) p_i$$

   - *Variable aleatoria continua*

   $$E[Y] = E[h(X)] = \int_{-\infty}^{\infty} h(x) f(x) dx$$

2. Para la transformación lineal de la v.a. X, dada por $Y = aX + b$, con $a, b \in \mathbb{R}$, la esperanza matemática de Y se define como

   $$E[Y] = E[aX + b] = aE[X] + b$$

   Obsérvese que la esperanza de una constante es la propia constante, es decir, $E[b] = b$, con $b \in \mathbb{R}$.

3. Esta propiedad surge como generalización de la propiedad anterior.

   Sea $Y = \sum_{i=0}^{n} a_i x_i$ donde $x_0 = 1$. La esperanza matemática de Y viene dada por

   $$E[Y] = E\left[\sum_{i=0}^{n} a_i x_i\right] = \sum_{i=0}^{n} a_i E[X_i] = a_0 + \sum_{i=1}^{n} a_i E[X_i]$$

4. Si la v.a. X está acotada entre los valores a y b con $a, b \in \mathbb{R}$, tal que $P(a \leq X \leq b) = 1$, la esperanza matemática de la v.a. X también está acotada entre dichos valores, $a \leq E[X] \leq b$.

5. Sean $X_1, ..., X_n$ v.a. independientes tales que $\forall i = 1, .., n\ \exists E[X_i]$, entonces

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

2.1.2. Cuantiles

Sea X una v.a. con función de distribución $F(x)$. El *cuantil de orden* p, denotado como C_p, es el valor de la variable aleatoria tal que la función de distribución en dicho punto es igual a p.

Como casos particulares de los cuantiles se pueden encontrar los *cuartiles*, los *deciles* y los *percentiles*, cuya definición depende de las partes en las que se divide la función de distribución; así, para el caso de los cuartiles, $F(x)$ se divide en cuatro partes siendo el primer cuartil $Q_1 = C_{\frac{1}{4}}$, el segundo cuartil, también denominado mediana, $Q_2 = C_{\frac{1}{2}}$ y el tercer cuartil $Q_3 = C_{\frac{3}{4}}$. Para los deciles, la función de distribución se divide en diez partes iguales, tales que $D_1 = C_{\frac{1}{10}}$, $D_2 = C_{\frac{1}{5}}, ..., D_9 = C_{\frac{9}{10}}$. Y para los percentiles, la $F(x)$ está dividida en cien partes iguales, siendo $P_1 = C_{\frac{1}{100}}$, $P_2 = C_{\frac{1}{50}}, ..., P_{99} = C_{\frac{99}{100}}$.

La definición depende del tipo de v.a. que se considere:

- *Variable aleatoria discreta*

  Sea X una v.a. discreta con función de distribución $F(x)$. Se denomina cuantil de orden p de la variable aleatoria X, con $p \in (0, 1)$, al valor C_p que satisface:

  $$F(C_p) = P(X \leq C_p) \geq p \qquad \text{y} \qquad P(X \geq C_p) \geq 1 - p$$

  Representando gráficamente $F(x)$ se observa como dependiendo del valor de p el cuantil vendrá determinado de un modo u otro. Esto se puede comprobar en las Figuras 2.1 y 2.2.

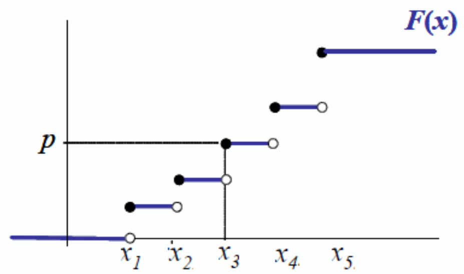

Figura 2.1. Cuantil de orden p dado por $[x_3, x_4]$

  Si el valor de p coincide con un escalón de la función de distribución, entonces el cuantil es todo el intervalo cerrado $[x_3, x_4]$.

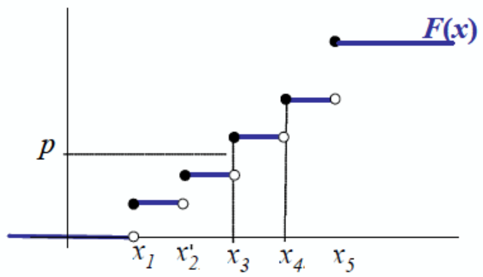

Figura 2.2. Cuantil de orden p dado por x_3

Si el valor de p no coincide con un escalón de la función de distribución, entonces el cuantil es el punto de masa del escalón siguiente (x_3).

Ejemplo 2.3

Sea X una v.a. discreta con función de distribución $F(x)$.

$$F(x) = \begin{cases} 0 & si \quad x < -2 \\ \frac{1}{4} & si \quad -2 \leq x < 0 \\ \frac{1}{2} & si \quad 0 \leq x < 1 \\ \frac{5}{6} & si \quad 1 \leq x < 2 \\ 1 & si \quad x \geq 2 \end{cases}$$

Calcular los cuantiles de orden $p = \frac{1}{5}$ y $p = \frac{5}{6}$.

Para encontrar los cuantiles se ha de tener en cuenta que:

$$F(C_p) = P(X \leq C_p) \geq p \qquad y \qquad P(X \geq C_p) \geq 1 - p$$

Por tanto, $C_{\frac{1}{5}} = -2$ y $C_{\frac{5}{6}} = [1, 2]$

- *Variable aleatoria continua*

Sea X una v.a. continua con función de distribución $F(x)$. Se define el cuantil de orden p de la v.a. X, con $p \in (0, 1)$, como el valor C_p que satisface

$$F(C_p) = P(X \leq C_p) = p$$

Representando la función de distribución $F(x)$ de una v.a. X, el cuantil de orden p viene dado por el valor de la v.a. para el que se acumula la probabilidad p.

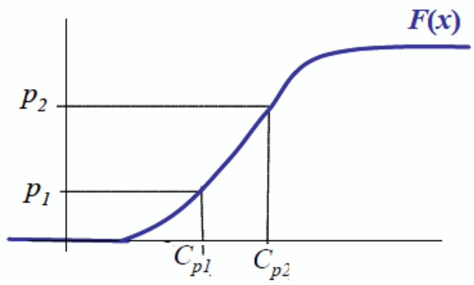

Figura 2.3. Cuantiles de una v.a. X continua

Ejemplo 2.4

Sea X una v.a. continua con función de densidad tal que

$$f(x) = \begin{cases} \frac{8}{7x^2} & si \quad 1 \leq x \leq 8 \\ 0 & en \quad otro\ caso \end{cases}$$

Calcular el primer y el tercer cuartil, Q_1 y Q_3, el decil 7, D_7, y el percentil 85, P_{85}.

Para el cálculo de los cuantiles especificados se ha de verificar la siguiente expresión:

$$F(C_p) = P(X \leq C_p) = p$$

Por tanto,

$F(Q_1) = \frac{1}{4} \Longrightarrow F(Q_1) = \int_{-\infty}^{Q_1} f(x)dx = \int_1^{Q_1} \frac{8}{7x^2}dx = \left[\frac{-8}{7x}\right]_1^{Q_1} =$

$= \frac{-8}{7}\left[\frac{1}{Q_1} - 1\right] = \frac{1}{4} \Longrightarrow Q_1 = 1.28$

$F(Q_3) = \frac{3}{4} \Longrightarrow F(Q_3) = \int_{-\infty}^{Q_3} f(x)dx = \int_1^{Q_3} \frac{8}{7x^2}dx = \left[\frac{-8}{7x}\right]_1^{Q_3} =$

$= \frac{-8}{7}\left[\frac{1}{Q_3} - 1\right] = \frac{3}{4} \Longrightarrow Q_3 = 2.91$

$F(D_7) = \frac{7}{10} \Longrightarrow F(D_7) = \int_{-\infty}^{D_7} f(x)dx = \int_1^{D_7} \frac{8}{7x^2}dx = \left[\frac{-8}{7x}\right]_1^{D_7} =$

$= \frac{-8}{7}\left[\frac{1}{D_7} - 1\right] = \frac{7}{10} \Longrightarrow D_7 = 2.58$

$F(P_{85}) = \frac{85}{100} \Longrightarrow F(P_{85}) = \int_{-\infty}^{P_{85}} f(x)dx = \int_1^{P_{85}} \frac{8}{7x^2}dx = \left[\frac{-8}{7x}\right]_1^{P_{85}} =$

$= \frac{-8}{7}\left[\frac{1}{P_{85}} - 1\right] = \frac{85}{100} \Longrightarrow P_{85} = 3.9$

Ejemplo 2.5

Sea X una v.a. continua con función de densidad dada por

$$f(x) = \begin{cases} \frac{2}{x^2} - \frac{1}{x^3} & si \quad \frac{1}{2} \leq x \leq 1 \\ \frac{1}{x^3} & si \quad x > 1 \\ 0 & en \quad otro\ caso \end{cases}$$

Calcular el tercer cuartil, Q_3.

Para calcular el tercer cuartil se ha de verificar $F(Q_3) = P(X \leq Q_3) = \frac{3}{4}$. Teniendo en cuenta como está definida la función de densidad hay que estudiar en que región de dicha función se acumula la probabilidad de $\frac{3}{4}$. Para ello, lo primero que hay que hacer es integrar la primera parte de $f(x)$, definida cuando $\frac{1}{2} \leq x \leq 1$, para ver que probabilidad se acumula en dicha parte. Es decir,

$$\int_{\frac{1}{2}}^{1} f(x)dx = \int_{\frac{1}{2}}^{1} \left(\frac{2}{x^2} - \frac{1}{x^3}\right) dx = \left[\frac{-2}{x} + \frac{1}{2x^2}\right]_{\frac{1}{2}}^{1} = \frac{1}{2}$$

Por tanto, en la primera parte de la función de densidad $\left(\frac{1}{2} \leq x \leq 1\right)$ la probabilidad acumulada es $\frac{1}{2}$, por lo que el tercer cuartil no se encuentra en esta primera parte de $f(x)$.

Por la definición de esta función de densidad el tercer cuartil tiene que estar definido cuando $x > 1$. Entonces:

$F(Q_3) = \int_{\frac{1}{2}}^{1} \left(\frac{2}{x^2} - \frac{1}{x^3}\right) dx + \int_{1}^{Q_3} \frac{1}{x^3}dx = \frac{3}{4} \Longrightarrow \int_{1}^{Q_3} \frac{1}{x^3}dx = \frac{3}{4} - \frac{1}{2} = \frac{1}{4}$

$\int_{1}^{Q_3} \frac{1}{x^3}dx = \frac{1}{4} \Longrightarrow \int_{1}^{Q_3} \frac{1}{x^3}dx = \left[\frac{-1}{2x^2}\right]_{1}^{Q_3} = \frac{-1}{2Q_3^2} + \frac{1}{2} = \frac{1}{4} \Longrightarrow Q_3 = +\sqrt{2}$

2.1.3. Mediana

Sea X una v.a. con función de distribución $F(x)$. Se define la *mediana* de la v.a. X como el valor de la variable aleatoria tal que la probabilidad a la izquierda de ese valor coincide con la probabilidad a la derecha de dicho valor, y es igual a $\frac{1}{2}$. Por tanto, la mediana es el cuantil de orden $\frac{1}{2}$ y se denota como Me, tal que

$$Me = C_{\frac{1}{2}} = Q_2 = D_5 = P_{50}$$

Además, la mediana es una medida de centralización que existe siempre y es muy importante, sobre todo si la esperanza matemática no existe o no es representativa.

El cálculo de la mediana también es diferente dependiendo del tipo de v.a. que se considera

- *Variable aleatoria discreta*

  Sea X una v.a. con función de distribución $F(x)$. Se define la mediana de la v.a. X como el valor que cumple

  $$F(Me) = P(X \leq Me) \geq \frac{1}{2} \qquad \text{y} \qquad P(X \geq Me) \geq \frac{1}{2}$$

  Recuérdese que para el caso discreto la mediana no tiene porqué ser única.

  Ejemplo 2.6
  Sea X una v.a. discreta con función de masa $P(X = 0) = 0.1$; $P(X = 1) = 0.3$; $P(X = 2) = 0.4$ y $P(X = 3) = 0.2$. Calcular la mediana.

  La mediana ha de verificar:

  $$F(Me) = P(X \leq Me) \geq \frac{1}{2} \quad y \quad P(X \geq Me) \geq \frac{1}{2}$$

  Por tanto, la mediana es 2, $Me = 2$.

- *Variable aleatoria continua*

  Sea X una v.a. continua con función de distribución $F(x)$. Se define la mediana de la v.a. X como el valor Me para el que la función de distribución acumula exactamente la mitad de la probabilidad, tal que

  $$F(Me) = P(X \leq Me) = \frac{1}{2}$$

  Ejemplo 2.7
  Considérese la v.a. descrita en el **Ejemplo 4.4.** *Calcular la mediana de dicha v.a.*

  $F(Me) = \int_{-\infty}^{Me} f(x)dx = \int_{1}^{Me} \frac{8}{7x^2}dx = \left[\frac{-8}{7x}\right]_1^{Me} = \frac{-8}{7}\left[\frac{1}{Me} - 1\right] = \frac{1}{2}$

  $\Longrightarrow Me = 1.78$

2.1.4. Moda

Se define la *moda* de una v.a. X como el valor o valores más frecuentes de dicha v.a. X, es decir, el valor de la v.a. X con mayor probabilidad de ocurrir. La moda se denota como Mo.

El cálculo de la moda depende del tipo de v.a.:

- *Variable aleatoria discreta*

  Sea X una v.a. discreta con puntos de masa $\{x_i\}$ $\forall i \in \mathbb{N}$. Se define la moda Mo como el valor de la v.a. X que maximiza la función de masa, tal que
  $$P(X = Mo) \geq P(X = x_i) \qquad \forall i \in \mathbb{N}$$

  Ejemplo 2.8
  Considérese la v.a. discreta presentada en el **Ejemplo 4.6.** *Calcular la moda de dicha v.a.*

  El valor más frecuente de la v.a. X es el correspondiente a $X = 2$, con probabilidad 0.4*, $P(X = 2) = 0.4$, que maximiza la función de masa.*

  Entonces, $Mo = 2$.

- *Variable aleatoria continua*

  Sea X una v.a. continua con función de densidad $f(x)$. Se define la moda Mo de la v.a. X como cualquier valor de la v.a. X que se corresponda con un máximo relativo de la función de densidad $f(x)$.

A continuación, se presenta la función de densidad de una v.a. X con dos máximos relativos. Por tanto, dicha v.a. X es bimodal, siendo las modas Mo_1 y Mo_2.

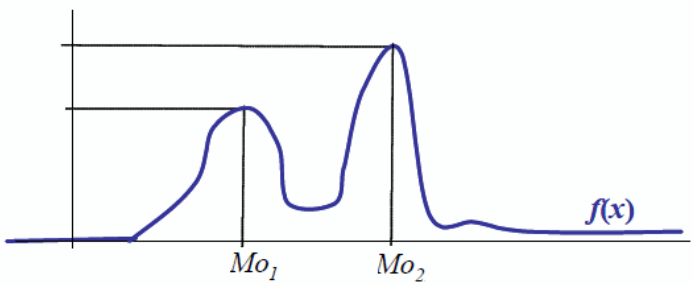

Figura 2.4. Densidad de una variable aleatoria X bimodal

Ejemplo 2.9

Considérese la v.a. continua descrita en el **Ejemplo 4.4.** *Calcular la moda de dicha v.a.*

La moda se obtiene calculando el máximo de la función de densidad de la v.a. X *tal que*

$f'(x) = \frac{-112x}{49x^4} = -\frac{16}{7x^3} < 0$

Por lo que la función es decreciente, de forma que $f(1) \geq f(x) \geq f(8)$.

Entonces, $Mo = 1$

2.2. Medidas de dispersión

Las medidas de dispersión indican la mayor o menor variabilidad de los valores de una variable aleatoria. En algunas de las medidas que se introducen en este apartado, como la varianza o la desviación típica, se mide la variabilidad de los valores de la v.a. respecto de sus valores centrales.

Además de las medidas anteriormente citadas, en este apartado también se define el recorrido y el recorrido intercuartílico.

2.2.1. Varianza

Se define la *varianza* de una v.a. X, como una medida cuadrática de la dispersión de los valores de la v.a. X respecto a su esperanza $E[X]$. Es decir, la varianza es el valor esperado de la diferencia al cuadrado entre cada uno de los valores de la v.a. X y su valor medio. La varianza se denotada como $V(X) = \sigma^2$.

$$V(X) = E\left[(X - E[X])^2\right] = E\left[(X - \mu)^2\right]$$

De nuevo, la definición de varianza depende del tipo de v.a. que se considere:

- *Variable aleatoria discreta*

  Sea X una v.a. discreta con puntos de masa $\{x_i\}$ $\forall i \in \mathbb{N}$ y función de masa

$P(X = x_i)$ $\forall i \in \mathbb{N}$. Se define la varianza de X como

$$V(X) = \sum_{i=1}^{\infty}(x_i - \mu)^2 P(X = x_i) = \sum_{i=1}^{\infty}(x_i - \mu)^2 p_i$$

Si el conjunto de valores de la v.a. X es finito, entonces la varianza se obtiene como una suma finita; si por el contrario, el conjunto es infinito numerable, se necesita la convergencia de la serie para que la varianza esté definida.

Ejemplo 2.10
Sea X una v.a. discreta con puntos de masa $\{-2, 0, 1, 3, 4\}$ equiprobables. Calcular la varianza de la v.a. X.

Para obtener la varianza es necesario el cálculo de la esperanza matemática de la v.a. X.

$E[X] = \sum_{i=1}^{5} x_i P(X = x_i) = p\sum_{i=1}^{5} x_i = \frac{6}{5} = \mu$

$V(X) = \sum_{i=1}^{5}(x_i - \mu)^2 P(X = x_i) = p\sum_{i=1}^{5}(x_i - \mu)^2 =$

$= \frac{1}{5}\left[\left(-2 - \frac{6}{5}\right)^2 + ... + \left(4 - \frac{6}{5}\right)^2\right] = \frac{114}{25} = 4.56.$

- *Variable aleatoria continua*

  Sea X una v.a. continua con función de densidad $f(x)$. Se define la varianza de la v.a. X como

$$V(X) = \int_{-\infty}^{\infty}(x - \mu)^2 f(x)dx$$

  En el caso continuo, es necesaria la convergencia de la integral impropia para que la varianza exista.

Si los valores de la v.a. X están concentrados alrededor de la esperanza de dicha v.a., entonces la varianza es pequeña, como se puede observar en la Figura 2.5.

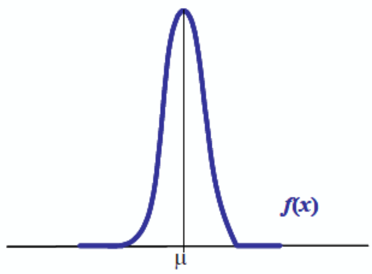

Figura 2.5. Varianza de una v.a. continua, donde σ^2es pequeña

Por el contrario, cuando los valores de la v.a. X están dispersos respecto a la esperanza, entonces la varianza es grande, tal y como se observa en la Figura 2.6.

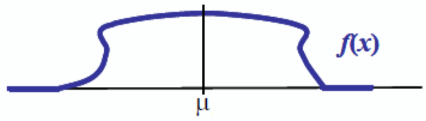

Figura 2.6. Varianza de una v.a. continua, donde σ^2es grande

Ejemplo 2.11

Sea X una v.a. continua con función de densidad tal que

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

Calcular la varianza de la v.a. X.

De igual forma que en el ejemplo anterior, es necesario el cálculo de la esperanza matemática de X para la obtención de la varianza.

$E\left[X\right] = \int_0^1 xf(x)dx = \int_0^1 2x^2dx = \frac{2}{3} = \mu$

$V(X) = \int_0^1 (x-\mu)^2 f(x)dx = \int_0^1 \left(x - \frac{2}{3}\right)^2 2xdx = \frac{1}{18}$

Propiedades de la varianza

1. Sea X una variable aleatoria, su varianza también puede expresarse como

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

2. La varianza de una v.a. X es nula si y solo si X es constante, siendo $P(X = c) = 1$
$$V(X) = 0 \Leftrightarrow X = c \qquad c \in \mathbb{R}$$

3. Sea Y una transformación lineal de la v.a. X, tal que $Y = aX + b$, con $a, b \in \mathbb{R}$. Entonces la varianza de Y es tal que
$$V(Y) = V(aX + b) = a^2 V(X)$$
Como consecuencia de esta propiedad se tiene que la varianza es invariante frente a los cambios de origen y además $V(-X) = V(X)$ siendo $a = -1$

4. Sean $X_1, ..., X_n$ v.a. independientes $\forall i = 1, .., n$. Entonces
$$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i)$$

2.2.2. Desviación típica

Dado que la varianza es una suma de cuadrados, la unidad de medida de la misma será el cuadrado de la unidad de medida en la que aparece expresada la variable aleatoria. Por tanto, es importante introducir otra medida de dispersión que se exprese en las mismas unidades que la v.a. en estudio.

Sea X una v.a., se define la *desviación típica* de X como la raíz positiva de la varianza, siendo su unidad de medida la misma que la de la v.a. X.
$$DT(X) = +\sqrt{V(X)} = \sigma$$

En ocasiones, es necesario trabajar con una v.a. centrada en el cero y con desviación típica 1. Esta nueva v.a. Z se denomina *variable aleatoria tipificada* y se define tal que
$$Z = \frac{X - \mu}{\sigma}$$
donde X es una v.a. con esperanza μ y desviación típica σ.

En este caso, Z verifica que $E[Z] = 0$ y $DT(Z) = 1$ $(V(Z) = 1)$.

2.2.3. Recorrido

Se define el *recorrido* de una v.a. X como la diferencia entre el mayor y el menor valor que toma dicha v.a., esto es

$$R = max(X) - min(X)$$

Esta medida presenta una gran inestabilidad ya que solo depende de los valores extremos de la v.a. X.

2.2.4. Recorrido intercuartílico

Para evitar la inestabilidad que presenta el recorrido, se define el *recorrido intercuartílico* de una v.a. X, como la diferencia entre el tercer y el primer cuartil, tal que

$$RI = Q_3 - Q_1 = C_{\frac{3}{4}} - C_{\frac{1}{4}}$$

Esta medida muestra el intervalo del 50% de los valores con más probabilidad de ocurrencia de la v.a. X.

Ejemplo 2.12

Sea X una v.a. discreta con función de distribución $F(x)$.

Calcular el recorrido y el recorrido intercuartílico de la v.a. X, siendo $F(x)$ tal que

$$F(x) = \begin{cases} 0 & si \quad x < -2 \\ \frac{1}{5} & si \quad -2 \leq x < 0 \\ \frac{1}{2} & si \quad 0 \leq x < 1 \\ \frac{5}{6} & si \quad 1 \leq x < 2 \\ 1 & si \quad x \geq 2 \end{cases}$$

Los puntos de masa de la v.a. X son $\{-2, 0, 1, 2\}$, por tanto, el recorrido que se obtiene como la diferencia entre el mayor y el menor valor que toma la v.a. X es el siguiente

$$R = 2 - (-2) = 4$$

Ejemplo 2.13

Para calcular el recorrido intercuartílico es necesario obtener previamente el primer y el tercer cuartil de X.

$$Q_1 = 0 \quad y \quad Q_3 = 1 \Longrightarrow RI = Q_3 - Q_1 = 1$$

2.3. Momentos

Algunas de las características de las variables aleatorias que se han visto anteriormente, como pueden ser la esperanza matemática y la varianza, son casos particulares de los momentos. Fundamentalmente, se distinguen:

- Momentos respecto al origen
- Momentos respecto a la media

2.3.1. Momentos respecto al origen

Dada una v.a. X, se define el *momento respecto al origen de orden* r, $\forall r \in \mathbb{N}$, como

$$\alpha_r = E[X^r]$$

Los momentos respecto al origen también se denominan momentos no centrados.

Distinguiendo en función del tipo de v.a. se tiene que:

- *Variable aleatoria discreta*

$$\alpha_r = E[X^r] = \sum_{i=1}^{\infty} x_i^r P(X = x_i) = \sum_{i=1}^{\infty} x_i^r p_i$$

- *Variable aleatoria continua*

$$\alpha_r = E[X^r] = \int_{-\infty}^{\infty} x^r f(x) dx$$

Como caso particular de los momentos respecto al origen, se observa que cuando $r = 1$ el momento respecto al origen de orden 1 es la esperanza matemática de la v.a. X, tal que $\alpha_1 = E[X]$.

Si la v.a. X está acotada, de forma que $\exists\ a, b \in \mathbb{R}\ /\ P(a \leq X \leq b) = 1$, entonces existen todos los momentos respecto al origen de la v.a. X. Sin embargo, si la v.a. X no está acotada, los momentos respecto al origen de orden r existen sii $E[|X|^r] < +\infty$.

Además, si existe el momento respecto al origen de orden r, también existen todos los momentos respecto al origen de orden k, con $k \leq r$.

Ejemplo 2.14

Sea X una v.a. continua con función de densidad dada por

$$f(x) = \begin{cases} \frac{2}{x^3} & si \quad x \geq 1 \\ 0 & en \quad otro\ caso \end{cases}$$

Calcular todos los posibles momentos respecto al origen.

Se comienza calculando el momento respecto al origen de orden 1, es decir, la esperanza matemática de la v.a. X:

$\alpha_1 = E[X] = \int_1^\infty x f(x)dx = \int_1^\infty \frac{2}{x^2}dx = \left[\frac{-2}{x}\right]_1^\infty = 2$

El momento respecto al origen de orden 2, es tal que

$\alpha_2 = E[X^2] = \int_1^\infty x^2 f(x)dx = \int_1^\infty \frac{2}{x}dx = [2\ln(x)]_1^\infty = \infty$

Por tanto, existe el momento respecto al origen de orden 1 y no existe ningún otro momento respecto al origen de orden k, $\forall k > 1$.

2.3.2. Momentos respecto a la media

Dada una v.a. X, se define el *momento respecto a la media de orden* r, $\forall r \in \mathbb{N}$, como

$$\mu_r = E[(X - \mu)^r]$$

siendo μ la esperanza matemática de la v.a. X.

Los momentos respecto a la media de orden r también se denominan momentos centrados.

Distinguiendo en función del tipo de v.a., se tiene que:

- *Variable aleatoria discreta*

$$\mu_r = E[(X-\mu)^r] = \sum_{i=1}^{\infty} (x_i - \mu)^r P(X = x_i) = \sum_{i=1}^{\infty} (x_i - \mu)^r p_i$$

- *Variable aleatoria continua*

$$\mu_r = E[(X-\mu)^r] = \int_{-\infty}^{\infty} (x-\mu)^r f(x)dx$$

Como caso particular de los momentos respecto a la media de orden r, se observa que:

- Cuando $r = 1$, el momento respecto de la media de orden 1 es cero, ya que $\mu_1 = E[(X-\mu)] = E[X] - \mu = \mu - \mu = 0$
- Cuando $r = 2$, el momento respecto de la media coincide con la varianza de la v.a. X, tal que $\mu_2 = E[(X-\mu)^2] = \sigma^2$

Además, cuando la distribución (función de masa y función de densidad) de la v.a. X es simétrica respecto a su media y existe el momento respecto a la media de orden r, siendo r un valor natural impar, entonces dicho momento es igual a cero, es decir, $\mu_r = E[(X-\mu)^r] = 0 \ \ \forall r \in \mathbb{N}$ *impar*.

2.3.3. Relaciones entre los momentos

Los momentos respecto al origen se pueden obtener mediante los momentos respecto a la media y viceversa. A continuación, se presenta la relación que existe entre los momentos respecto al origen y los momentos respecto a la media.

- $\alpha_r = E[X^r] = \sum\limits_{i=0}^{r} \begin{pmatrix} r \\ i \end{pmatrix} \mu_{r-i}\mu^i$

  Para obtener esta expresión se aplica el binomio de Newton, tal que

  $$\alpha_r = E[X^r] = E[((X-\mu)+\mu)^r] = E\left[\sum_{i=0}^{r} \begin{pmatrix} r \\ i \end{pmatrix} (X-\mu)^{r-i} \mu^i\right] =$$

  $$= \sum_{i=0}^{r} \begin{pmatrix} r \\ i \end{pmatrix} E\left[(X-\mu)^{r-i}\right] \mu^i = \sum_{i=0}^{r} \begin{pmatrix} r \\ i \end{pmatrix} \mu_{r-i}\mu^i$$

- $\mu_r = E[(X-\mu)^r] = \sum_{i=0}^{r} \begin{pmatrix} r \\ i \end{pmatrix} (-1)^i \alpha_{r-i}\mu^i$

  Aplicando el binomio de Newton

  $\mu_r = E[(X-\mu)^r] = E\left[\sum_{i=0}^{r} \begin{pmatrix} r \\ i \end{pmatrix} (-1)^i \, (X)^{r-i} \, \mu^i\right] =$

  $= \sum_{i=0}^{r} \begin{pmatrix} r \\ i \end{pmatrix} (-1)^i E\left[X^{r-i}\right] \mu^i = \sum_{i=0}^{r} \begin{pmatrix} r \\ i \end{pmatrix} (-1)^i \alpha_{r-i}\mu^i$

2.3.4. Función generatriz de momentos

Los momentos se pueden determinar directamente aplicando la definición de los mismos, aunque algunos momentos de órdenes elevados pueden resultar muy laboriosos. Es por ello, que en este apartado se define la *función generatriz de momentos* para calcular de forma sencilla cualquier momento de interés.

Se define la función generatriz de momentos (*f.g.m.*) de una v.a. X como la función $M_X : \mathbb{R} \to \mathbb{R}$ dada por

$$\forall t \in \mathbb{R} \qquad M_X(t) = E[e^{tX}]$$

El cálculo de la f.g.m. depende del tipo de v.a. que se considere:

- *Variable aleatoria discreta*

$$M_X(t) = E[e^{tX}] = \sum_{i=1}^{\infty} e^{tx_i} P\,(X = x_i) = \sum_{i=1}^{\infty} e^{tx_i} p_i$$

- *Variable aleatoria continua*

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

La f.g.m. $M_X(t)$, puede ser finita o infinita ya que e^{tX} es una función positiva para todo $t \in \mathbb{R}$.

Cálculo de los momentos con la f.g.m.

- *Momentos respecto al origen*

  Los momentos respecto al origen de orden r se obtienen mediante la derivada $r-$ésima de la f.g.m. evaluada en el punto $t = 0$, tal que

  $$\alpha_r = M_X^{(r)}(0) = \frac{\partial^r}{\partial t^r}\left(M_X(t)\right)|_{t=0}$$

- *Momentos respecto a la media*

  Para obtener los momentos respecto a la media de orden r, al igual que para los momentos respecto al origen de orden r, es necesario derivar la f.g.m. y evaluarla en el punto $t = 0$, además de multiplicar por $e^{-t\mu}$.

  $$\mu_r = e^{-t\mu}M_X^{(r)}(0) = \frac{\partial^r}{\partial t^r}\left(e^{-t\mu}M_X(t)\right)|_{t=0}$$

Propiedades de la f.g.m.

1. Si $t = 0 \Rightarrow M_X(t = 0) = E[e^0] = 1$. Por tanto, la función generatriz de momentos existe siempre que $t = 0$.

2. La f.g.m., en caso de existir, es siempre única y determina univocamente a la distribución de probabilidad de la variable en estudio. Como consecuencia de esta propiedad se enuncia la siguiente propiedad.

3. Sean X e Y dos variables aleatorias con la misma función generatriz de momentos, $M_X(t) = M_Y(t)$. Entonces, la distribución de X e Y también es la misma.

4. Sea X una v.a. y sea Y una transformación lineal de la v.a. X dada por, $Y = aX + b$, con $a, b \in \mathbb{R}$, entonces la f.g.m. de la v.a. Y es tal que

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

5. Sean $X_1, ..., X_n$ variables aleatorias independientes y sea la v.a. $S = \sum_{i=1}^{n} X_i$.

   La función generatriz de momentos de la v.a. S viene dada por

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

$$M_S(t) = E[e^{tS}] = E[e^{t(X_1+...+X_n)}] = E[e^{tX_1}e^{tX_2}...e^{tX_n}] =$$

por ser v.a. independientes la esperanza del producto es el producto de las esperanzas

$$= E[e^{tX_1}]E[e^{tX_2}]...E[e^{tX_n}] = \prod_{i=1}^{n} M_{X_i}(t)$$

Ejemplo 2.15

Sea una v.a. X con función de densidad $f(x) = \begin{cases} e^{-x} & si \quad x > 0 \\ 0 & en \quad otro\ caso \end{cases}$

a) *Calcular la f.g.m.*

b) *Calcular la esperanza y la varianza de la v.a. X a partir de la f.g.m.*

a) $M_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x)dx = \int_0^{\infty} e^{tx}e^{-x}dx = \int_0^{\infty} e^{x(t-1)}dx =$

$$= \left[\frac{1}{t-1}e^{x(t-1)}\right]_0^{\infty} = \frac{1}{1-t} \ si \ t < 1$$

b) $\alpha_1 = M_X^{(1)}(0) = \frac{\partial}{\partial t}\left(M_X(t)\right)|_{t=0} = \dfrac{1}{(1-t)^2}|_{t=0} = 1$

Por tanto, la $E[X] = 1$

La varianza se puede obtener de dos formas:

a través de la f.g.m., donde $\mu_r = e^{-t\mu}M_X^{(r)}(0)$, *o bien a través de los momentos respecto del origen, tal que* $V(X) = E[X^2] - E^2[X] = \alpha_2 - \alpha_1^2$.

Para el cálculo de la varianza en este ejemplo, se utilizan los momentos respecto al origen, siendo

$$\alpha_2 = M_X^{(2)}(0) = \frac{\partial^2}{\partial t^2}\left(M_X(t)\right)|_{t=0} = \frac{2(1-t)}{(1-t)^4}|_{t=0} = 2$$

Entonces, $V(X) = \alpha_2 - \alpha_1^2 = 2 - 1 = 1$

2.4. Medidas de forma

Las medidas de forma indican, según la tipología de la distribución de la v.a. de interés y de acuerdo a su representación gráfica, la forma de dicha variable. Estas medidas hacen referencia principalmente a la asimetría y el aplastamiento o kurtosis de la variable aleatoria. Por tanto, las medidas de forma se pueden clasificar en dos grupos: las medidas de asimetría y las medidas de aplastamiento o kurtosis.

2.4.1. Medida de asimetría

Una v.a. X se dice que es simétrica cuando su función de masa o función de densidad lo es respecto del eje $X = E[X]$.

Sea X una v.a., se define el *índice de asimetría* de X como el cociente entre el momento respecto a la media de orden 3 y la desviación típica elevada al cubo, es decir

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{E[(X-\mu)^3]}{\sigma^3}$$

En función del signo del índice de simetría, γ_1, se tiene que

- Si $\gamma_1 = 0 \Rightarrow$ La función de densidad de la v.a. X es *simétrica* respecto a su esperanza μ.

- Si $\gamma_1 < 0 \Rightarrow$ La función de densidad de la v.a. X es *asimétrica negativa.* Por tanto, la probabilidad de que la v.a. X tome valores a la izquierda de la media es mayor que la probabilidad de que los tome a la derecha de la misma.

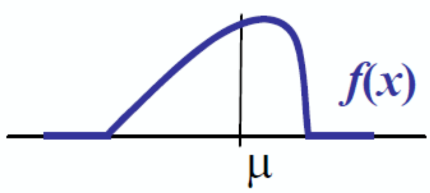

Figura 2.7. Variable aleatoria asimétrica negativa

- Si $\gamma_1 > 0 \Rightarrow$ La función de densidad de la v.a. X es *asimétrica positiva.* Por tanto, la probabilidad de que la v.a. tome valores a la derecha de la media es mayor que la probabilidad de que los tome a la izquierda de dicha media.

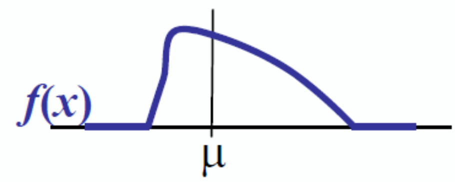

Figura 2.8. Variable aleatoria asimétrica positiva

2.4.2. Medida de aplastamiento o kurtosis

Esta medida refleja el grado de aplastamiento de una v.a. X comparándola con la curva normal.

Se define el *índice de aplastamiento o kurtosis* (ε_k) de una v.a. X como

$$\varepsilon_k = \frac{\mu_4}{\sigma^4} - 3 = \frac{E[(X-\mu)^4]}{\sigma^4} - 3$$

En función del signo de ε_k se tiene que:

- Si $\varepsilon_k = 0$ $\Rightarrow$La función de densidad de la v.a. X muestra un grado de aplastamiento similar al de una v.a. con distribución normal y se denomina *mesocúrtica*.

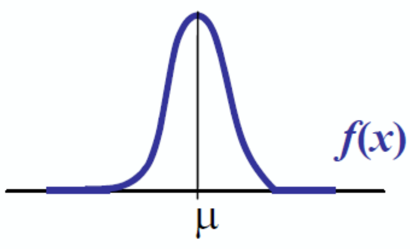

Figura 2.9. Variable aleatoria con distribución Normal (Mesocúrtica)

- Si $\varepsilon_k < 0 \Rightarrow$ La función de densidad de la v.a. X es más aplastada que la de una v.a. con distribución normal y se denomina *platicúrtica*.

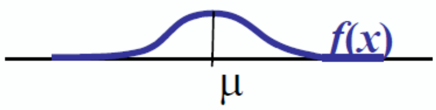

Figura 2.10. Variable aleatoria con índice de aplastamiento negativo (Platicúrtica)

- Si $\varepsilon_k > 0 \Rightarrow$ La función de densidad de la v.a. X es más apuntada que la de la normal y se denomina *leptocúrtica*.

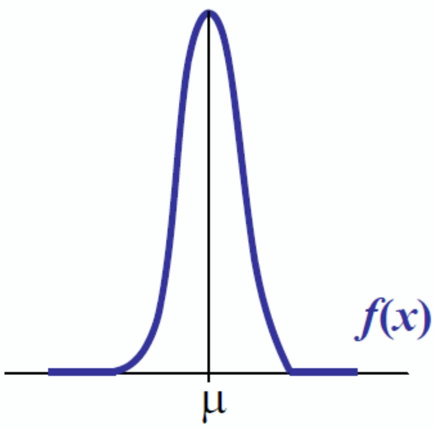

Figura 2.11. Variable aleatoria con índice de aplastamiento positivo (Leptocúrtica)

2.5. Desigualdad de Markov y Tchebychev

2.5.1. Desigualdad de Markov

Sea X una variable aleatoria y $g(X)$ una transformación de la v.a. X, tal que $g(X) \geq 0$, entonces

$$\forall t > 0 \text{ se verifica que } P\left(g(X) \geq t\right) \leq \frac{E[g(X)]}{t}$$

Demostración 2.1

$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx = \int_{g(X)\geq t} g(x)f(x)dx + \int_{g(X)<t} g(x)f(x)dx \geq$

$\geq \int_{g(X)\geq t} g(x)f(x)dx \geq \int_{g(X)\geq t} tf(x)dx = t\int_{g(X)\geq t} f(x)dx = tP\left(g(X) \geq t\right)$

Por tanto,

$$P\left(g(X) \geq t\right) \leq \frac{E[g(X)]}{t}$$

Como caso particular, se tiene que cuando una una v.a. X está acotada inferiormente por el 0, es decir $P\left(X \geq 0\right) = 1$, por la desigualdad de Markov se tiene que

$$P\left(X \geq t\right) \leq \frac{E[X]}{t} \qquad \forall t > 0$$

2.5.2. Desigualdad de Tchebychev

Como consecuencia de la desigualdad de Markov, se obtiene la *desigualdad de Tchebychev* que se utiliza para calcular cotas de las probabilidades de una v.a. X cuya distribución es desconocida, pero su esperanza y varianza son conocidas. Por tanto, la desigualdad de Tchebychev se enuncia como:

Sea X una v.a. con media μ y desviación típica σ, entonces

$$\forall k > 0 \text{ se verifica que } P\left(|X-\mu| < k\sigma\right) \geq 1 - \frac{1}{k^2} \text{ o también}$$

$$P\left(|X-\mu| \geq k\sigma\right) \leq \frac{1}{k^2}$$

Demostración 2.2

Aplicando la desigualdad de Markov y siendo $g(X) = (X-\mu)^2 \geq 0$, se obtiene que:

$$P\left((X-\mu)^2 \geq t\right) \leq \frac{E[(X-\mu)^2]}{t} = \frac{\sigma^2}{t} \text{ de igual forma, } P\left(|X-\mu| \geq \sqrt{t}\right) \leq \frac{\sigma^2}{t}$$

Considerando $t = k^2\sigma^2$, la desigualdad queda tal que

$$P\left(|X-\mu| \geq k\sigma\right) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$$

En consecuencia, la cota inferior para la probabilidad del suceso contrario queda definida como

$$P\left(|X-\mu| < k\sigma\right) \geq 1 - \frac{1}{k^2}$$

Si se considera $k\sigma = t$, la desigualdad de Tchebychev es tal que

$$\forall t > 0 \text{ se verifica que } P\left(|X-\mu| < t\right) \geq 1 - \frac{\sigma^2}{t^2} \text{ o también}$$

$$P\left(|X-\mu| \geq t\right) \leq \frac{\sigma^2}{t^2}$$

Ejemplo 2.16

El número de estudiantes del campus universitario de Madrid que utilizan las líneas de autobuses, es en término medio de 8550 estudiantes con una varianza de 900. Calcular una cota a la probabilidad de que un día concreto el número de estudiantes que usan los autobuses del campus esté comprendido entre 8450 y 8650.

Sea $X =$"Número de estudiantes que utilizan las líneas de autobuses del campus".

Como sólo son conocidas la media y la varianza de X, para calcular esta cota es necesario utilizar la desigualdad de Tchebychev.

$$P\left(8450 < X < 8650\right) = P\left(8450 - 8550 < X - \mu < 8650 - 8550\right) =$$

$$= P\left(-100 < X - \mu < 100\right) = P\left(|X-\mu| < 100\right) \geq 1 - \frac{900}{100^2} = 0.91$$

La probabilidad de que el número de estudiantes que utilizan los autobuses del campus esté comprendida entre 8450 y 8650 es al menos 0.91.