IV. Variables Aleatorias Continuas

M.A. Fernando Jesús Martínez Eissa

02H2019

Variables Aleatorias

- Se dice que X es una Variable Aleatoria si $X:\Omega \to \mathbb{R}$, es decir una función cuyo dominio es el espacio muestral y su contradominio la recta real.

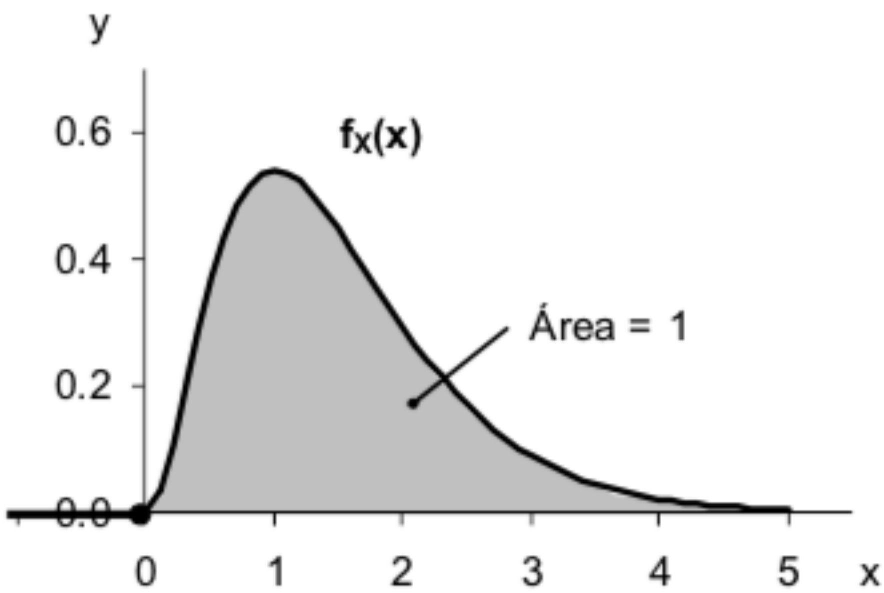
- A la imagen o rango de la función X se le denomina soporte de la variable aleatoria X.
- Con base en su soporte las v.a. se clasifican en:
  a) Discretas: Se asocian a un espacio muestral discreto, su soporte es finito o infinito numerable y provienen de un proceso de conteo.
  b) Continuas: Se asocian a un espacio muestral continuo, su soporte es infinito no numerable y provienen de un proceso de medición.

Función de densidad de probabilidad

Si X es una v.a. continua, se dice que la función $f_X(\cdot), f_X: \mathbb{R} \to [0, \infty)$, es la función de densidad de probabilidad de X si:

i. $f_X(x) \geq 0 \; para \; toda \; x \in \mathbb{R}; y$

ii. $\int_{-\infty}^{\infty} f_X(x)dx = 1$



Función de densidad de probabilidad

La principal diferencia entre las variables aleatorias discretas y continuas es que en el caso discreto se consideran sumas de masas de probabilidad y en el caso continuo se consideran áreas bajo la curva, ya que el soporte de las variables aleatorias continuas es denso (propiedad de los números reales).

Importante. Para el caso continuo el contradominio de la $f.d.p.$ es $[0, \infty)$ (y no [0, 1] como en el caso discreto) ya que $f_X(\cdot)$ no es una probabilidad (de hecho ni siquiera es relevante).

En el caso continuo lo relevante para el cálculo de probabilidades no es el valor que toma la $f.d.p.$ sino el área bajo la curva.

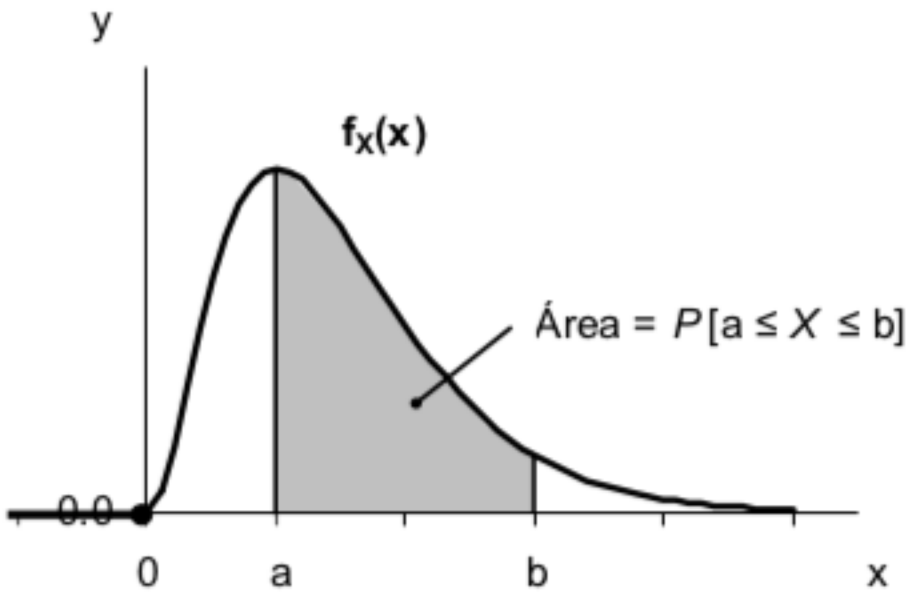
Función de densidad de probabilidad

Proposición

Si X es una v.a. continua y $A \subset \mathbb{R}$., entonces $P[x \in A] = \int_A f_X(x)dx$

Corolario

Si X es una v.a. continua $P[a \leq x \leq b] = \int_a^b f_X(x)dx$



Función de densidad de probabilidad

Proposición

Si X es una v.a. continua entonces:

i. $P[X = a] = 0$

ii. $P[X \leq a] = P[X < a]$

iii. $P[a \leq X \leq b] = P[a < X < b] = P[a \leq X < b] = P[a < X \leq b]$

Función de densidad de probabilidad

Ejemplo 1

Considere que X es una v.a. continua para la cual:

$$f_X(x) = \begin{cases} 0.2, & si\ 0 < x \leq 1 \\ 0.3, & si\ 1 < x \leq 2 \\ -0.5 + 0.4x, & si\ 2 < x \leq 3 \\ 0 & e.o.c. \end{cases}$$

a) Grafique $f_X(x)$ y exprésela en términos de funciones indicadoras
b) Verifique que $f_X(x)$ es función de densidad de probabilidad
c) Calcule $P[0.5 \leq X < 1]$ y $P[0.5 < X \leq 1]$¿Existe alguna diferencia?
d) Determine $P[X = 2]$Comente su resultado
e) Calcule $P[X > 2], P[X < 0.5]$ y $P[X \geq 0.5]$
f) Calcule $P[X < 1 | X \geq 0.5]$

Función de densidad de probabilidad

Ejemplo 2

La Distribución Exponencial con parámetro $\theta, \theta > 0$, se caracteriza por la siguiente función de densidad de probabilidad:

$$f_x(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}} I_{(0,\infty)}(x)$$

a) Verifique que $f_X(x)$ es una función de densidad de probabilidad

b) Calcule $P[a<X<b]$ para $0<a<b$

Función de densidad de probabilidad

Ejemplo 3

Suponga que $f_1(x)\ y\ f_2(x)$ son funciones de densidad de probabilidad. Demuestre que la función

$$f(x) = \alpha f_1(x) + (1-\alpha) f_2(x)$$

también es función de densidad de probabilidad.

Función de Acumulación Continua

Se define para una variable aleatoria continua X, la función acumulación $F(x)$ como:

$$F(x) = P(X \leq x) \; para \; -\infty < x < \infty$$

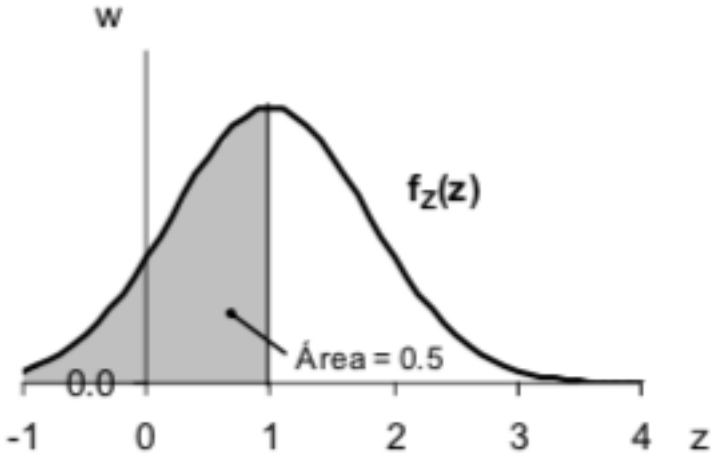
Esta función especifica, para todos lo reales, la probabilidad de que la variable aleatoria sea menor o igual a x.

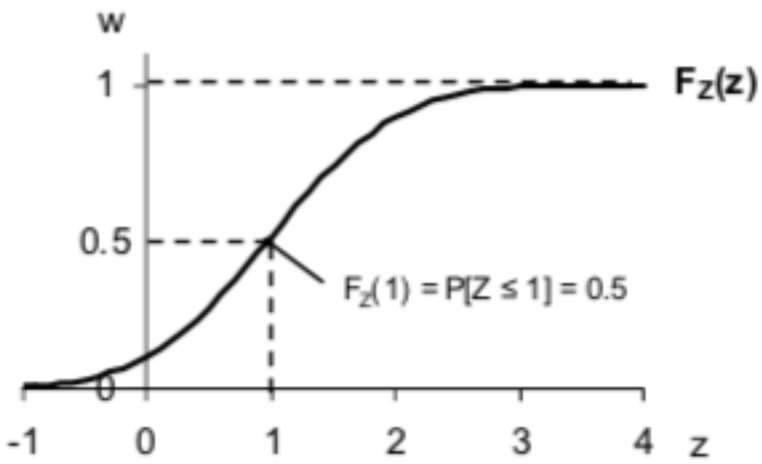
Considere ahora que $a \leq b$. Entonces se sigue de $[X \leq a] \epsilon [X \leq b]$, por lo tanto se sigue que $F(a) \leq F(b)$.

Nota: Tiene la misma definición para variables discretas y continuas.

Función de Acumulación Continua

Caso continuo





Función de Acumulación Continua

Proposición

Si Z es v.a. continua, entonces

$$F_Z(\mathrm{z}) = P[Z \leq z] = \int_{-\infty}^{z} f_z\,(t)dt \;\forall z \in \mathbb{R}$$

Nota: El cálculo de la f.d.a. requiere un cambio de notación en la variable del integrando para evitar ambigüedad. Este cambio es válido ya que en cualquier integral definida la variable del integrando es una variable muda.

Función de Acumulación Continua

1. F_X es una función no decreciente, es decir si $a < b$, entonces $F_X(a) \leq F_X\,(b)$

2. $\lim_{x \longrightarrow +\infty} F_X(x) = 1$

3. $\lim_{x \longrightarrow -\infty} F_X(x) = 0$

4. F_X es una función continua por la derecha. Es decir,

$$\lim_{h \longrightarrow 0} F_X(x+h) = F_X(x)\; para\; h > 0$$

Función de Acumulación Continua

Proposición

Si Z es v.a. continua con $a < b$, entonces

$$P[a \leq Z \leq b] = F_Z(b) - F_Z(a)$$

Si Z es v.a. continua, entonces

$$f_Z(z) = \frac{d}{dz} F_Z(z)$$

Función de Acumulación Continua

Ejemplo 4

El período de funcionamiento de cierto transmisor hasta su primera falla (en cientos de horas) es una variable aleatoria T con función de distribución acumulada:

$$F_T(t) = \begin{cases} 0, & si\ t < 0 \\ 1 - e^{-t^2}, & si\ t \geq 0 \end{cases}$$

a) Calcule la probabilidad de que el transmisor trabaje por lo menos durante 200 horas hasta tener su primera falla.

b) Determine la función de densidad de probabilidad de la variable aleatoria T.

Función de supervivencia

Se dice que la función $S_X(\cdot), S_X: [0, \infty) \longrightarrow [0,1]$ es la función de supervivencia de la v.a. X si $S_X(x) = P[X > x] \forall x \geq 0$

Las funciones de supervivencia son útiles para modelar el tiempo futuro de vida. *vgr.*, el tiempo a la muerte de algunas especies (seres humanos, animales, bacterias, virus y algunas otras especies biológicas), o la vida útil de algunos objetos (maquinaria, mobiliario, equipo de cómputo, artículos electrodomésticos). A la rama de la Estadística encargada del estudio de esta clase de fenómenos se le conoce como Análisis de Supervivencia.

Proposición

$$S_X(x) = 1 - F_X(x)\ para\ x \geq 0$$

Función de Acumulación Continua

Ejemplo 5

La distribución Weibull con parámetros α y $\beta, \alpha > 0$ y $\beta > 0$ es utilizada frecuentemente para modelar el tiempo de vida (X) de algunas máquinas. Si la función de supervivencia de X está dada por:

$$S_X(x) = e^{-\left(\frac{x}{\alpha}\right)^{\beta}} para\ x > 0$$

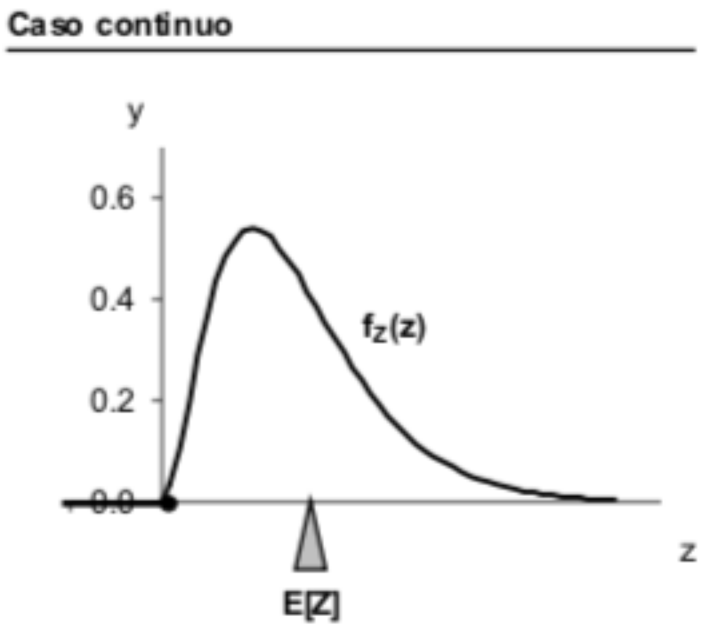
Calcule la función de densidad de X.

Esperanza de una v.a. contínua

Si X es una variable aleatoria continua con función de probabilidad $f(x)$, entonces su esperanza o valor esperado $E[X]$ es:

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

Geométricamente representa el centro de masa de la función, es decir en donde se equilibran.



Esperanza de una v.a. contínua

Ejemplo 6

El tiempo en horas Z requerido por los estudiantes de un curso para resolver un examen final es una v.a. con función de densidad :

$$f_Z(Z) = \begin{cases} \dfrac{z(4-z)}{9}, & 0 \leq z \leq 3 \\ 0, & e.o.c. \end{cases}$$

a) Realice un bosquejo de la gráfica de $f_Z(z)$.

b) Calcule e interprete la esperanza matemática de la v.a. Z

Esperanza de una v.a. contínua

Ejemplo 7

La v.a. Y tiene función de densidad :

$$f_Y(y) = \begin{cases} a + by, & 0 < y < 1 \\ 0, & e.o.c. \end{cases}$$

Encuentre los valores de a y b tales que la media de Y sea $\frac{2}{3}$

Esperanza de una v.a. continua

Si $E[X]$ existe y $a \leq X \leq b \Longrightarrow a \leq E[X] \leq b$.

Cuando $E[X]$ es de la forma ∞ ó $\infty - \infty$, se dice que la esperanza no existe. Únicamente existe si $E[X] < \infty$, es decir cuando converge a un valor.

Proposición

Si X es una v.a. continua, entonces

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

Esperanza de una v.a. continua

Ejemplo 8

Sea X la v.a. Con función de densidad $f_X(x) = x^{-2} I_{[1,\infty)}(x)$.

a) Grafique y verifique que $f_X(x)$ es función de densidad de probabilidad
b) Demuestre que $E[X]$ no existe.

Valor Esperado (*continuo*)

El concepto de esperanza se puede generalizar a través del valor esperado.

Considere que se tiene una v.a. X y su correspondiente función de densidad $g(x)$ y que se busca calcular su valor esperado.

Dado que $g(x)$ es en sí una variable aleatoria continua, entonces buscamos:

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

Ejemplos de transformaciones pueden ser: $g(x) = x^2; g(x) = ln(x); g(x) = (x-c)^2$. El valor esperado es la generalización de la esperanza ya que $Y = g(X)$ es una nueva v.a.

Valor Esperado (*continuo*)

Al igual que con el caso de la esperanza cuando $E[g(X)]$ es de la forma ∞ ó $\infty - \infty$, se dice que la esperanza no existe. Únicamente existe si $E[g(X)] < \infty$, es decir cuando converge a un valor.

Aunque X es la v.a. para la cual se conoce $f_X(x)$, la definición de valor esperado permite calcular $E[Y] = E[g(X)]$ sin necesidad de conocer $f_Y(y)$.

Valor Esperado (*continuo*)

Ejemplo 9

Una casa de bolsa ofrece a sus clientes un fondo de renta variable con una tasa de rendimiento anual con la siguiente función de densidad:

$$f_X(x) = \begin{cases} \dfrac{2}{9}(x+1), & -1 \leq x \leq 2 \\ 0, & e.o.c. \end{cases}$$

Un inversionista puede ganar o perder en este fondo dependiendo si la tasa X es positiva

o negativa, respectivamente.

a) Calcular la tasa de rendimiento anual esperada..

b) Si un cliente decide invertir 50 mil pesos en este fondo, entonces la comisión del ejecutivo de cuenta será de $C = 5\sqrt{X+1}$ (en miles de pesos). Calcule el valor esperado de la comisión del ejecutivo de cuenta.

Valor Esperado (*continuo*)

Ejemplo 9 (*Continuación*)

c) La tasa de rendimiento continuo del fondo está dada por $\delta = ln(1+X)$. Calcule la probabilidad de que la tasa de rendimiento continuo sea mayor a 100%.

d) Si otro cliente de la casa de bolsa decide invertir 90 mil pesos en este fondo, entonces el saldo de la inversión al final de 2 años será $S = 90(1+X)^2$(en miles de pesos). Calcule la probabilidad de que al final de los 2 años el saldo del inversionista sea menor al saldo final promedio de esta inversión.

Momentos

A la cantidad $E[X^n], n\epsilon\ \mathbb{R}$ se le llama el *n-ésimo* momento de X

$$\mu'_n = E[X^n] = \int_A x^n f(x)dx$$

y su *n-ésimo* momento central se define como:

$$\mu_n = E[(X - E[X])^n]$$

El cálculo de los momentos y de los momentos centrales se realiza aplicando la ley del estadístico inconsciente considerando las transformaciones $g(X) = X^n$ y $g(X) = (X - E[X])^n$

Nótese que cuando n=1 $\mu'_1 = E[X^1] = \int_{x,f(x)>0} x^1 f(x)dx = E[X]$, es decir el primer momento de una variable aleatoria es su esperanza.

Momentos

Corolario

Si X es una v.a. y $E[X] = \mu$, entonces:

I. $\mu_2 = \mu_2' - (\mu_1')^2, i.e. E[(X-\mu)^2] = E[X^2] - (E[X])^2$

II. $\mu_3 = \mu_3' - 3\mu_1'\mu_2' + 2(\mu_1')^3, i.e.$
$E[(X-\mu)^3] = E[X^3] - 3E[X]E[X^2] + 2(E[X])^3$

III. $\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6(\mu_1')^2\mu_2' - 3(\mu_1')^4 i.e.$
$E[(X-\mu)^4] = E[X^4] - 4E[X]E[X^3] + 6(E[X])^2E[X^2] - 3(E[X])^4$

Ejemplo 10

Se sabe que si X tiene una Distribución Uniforme Continua sobre el intervalo $[\theta_1, \theta_2]$ con $\theta_1, \theta_2 \in \mathbb{R}$ y $\theta_1 < \theta_2$, entonces $f_X(x) = \frac{1}{\theta_2 - \theta_1} I_{[\theta_1, \theta_2]}(x)$, es decir, la distribución asigna la misma probabilidad a intervalos de igual longitud en $[\theta_1, \theta_2]$.

a) Calcular el *k-ésimo* momento de la variable aleatoria X, k = 1, 2,...
b) Graficar e interpretar el valor del primer momento (k = 1).
c) Calcule el *k-ésimo* momento central de la variable aleatoria X, k = 1, 2,....

Varianza (*continuo*)

Si X es una v.a. Con media μ, entonces su varianza denotada por Var(X), se define como:

$$Var[X] = E[(X-\mu)^2]$$

Entonces

$$Var[X] = E[X^2 - 2X\mu + \mu^2]$$

$$Var[X] = E[X^2] - 2\mu E[X] + \mu^2$$

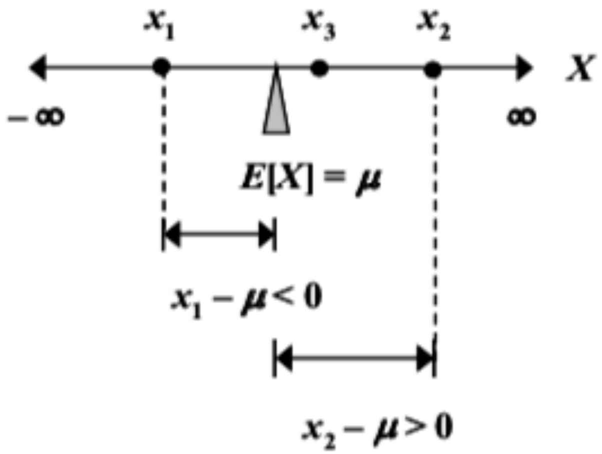
$$Var[X] = E[X^2] - 2E[X]E[X] + E^2[X]$$

$$Var[X] = E[X^2] - E^2[X]$$

Nota: observe que la varianza corresponde al segundo momento central.

Varianza (*continuoo*)

La varianza es una medida de dispersión de la distribución de la v.a. X respecto a su esperanza como se muestra a continuación:



Observe

$X - E[X]$ es una distancia aleatoria con signo (+ o -)

$g(X) = (X - E[X])^2 \geq 0$ es una distancia aleatoria positiva o cero, pero con unidades al cuadrado.

Proposición $Var[X] \geq 0$

Varianza (*continuo*)

Ejemplo 11

Se sabe que X tiene una Distribución Exponencial con parámetro β, $\beta > 0$, entonces $f_X(x) = \frac{1}{\beta} e^{-\frac{x}{\beta}} I_{[0,\infty)}(x)$.

a) Calcular la varianza y la desviación estándar de X.

Varianza (*continuo*)

Ejemplo 12

La producción mensual de jarabe para la tos de un laboratorio (X, en miles de litros) es una variable aleatoria con función de densidad $f_X(x) = \frac{6-x}{8} I_{[2,6]}(x)$. Se sabe que para la elaboración de este jarabe, el laboratorio enfrenta *costos fijos* por \$145,000 y una función de *costos variables* (en miles de pesos) dada por $C(x) = \frac{105(x-2)^2}{4}$.

a) Calcular la esperanza y la varianza de la producción mensual de jarabe.
b) Calcule el *costo total esperado* que debe pagar mensualmente el laboratorio
c) Calcule el *costo medio esperado* que debe pagar mensualmente el laboratorio.

Medidas Poblacionales

Medidas poblacionales

Medidas que caracterizan y cuantifican aspectos relevantes de una distribución

Grupo	Descripción	Medidas
Tendencia central	• Ubican la parte central de una distribución de probabilidades	• Media • Mediana • Moda
Localización	• Segmentan el soporte de una distribución mediante valores que cuantifican la probabilidad acumulada	• Percentiles • Cuartiles • Cuantiles
Dispersión	• Miden el grado de variabilidad de la distribución, es decir, su grado de concentración o dispersión	• Varianza • Desviación estándar • Coeficiente de variación • Amplitud intercuartílica • Desviación media
Simetría y forma	• Permiten inferir el perfil de la distribución de probabilidades	• Coeficiente de asimetría • Coeficiente de curtosis

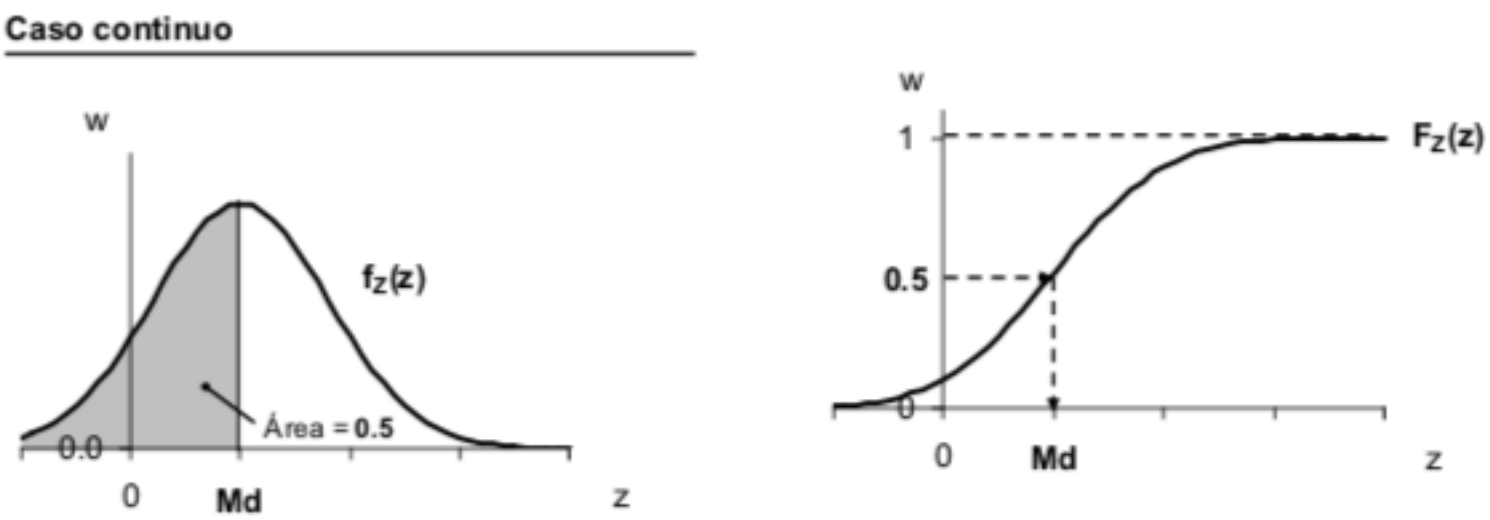
v.a.'s Discretas | v.a.'s Continuas

Medidas estudiadas previamente

Mediana (*Continuo*)

Mediana (Med): Si X es una v.a. continua con $F_X(x)$, entonces su mediana, corresponde al soporte de X tal que:,

$$F_X(Med) = \frac{1}{2}$$



Mediana (*Continuo*)

Ventajas de la mediana sobre la media:

i. La mediana siempre existe
ii. La mediana es mejor medida de tendencia central en distribuciones no simétricas

Distribución Simétrica

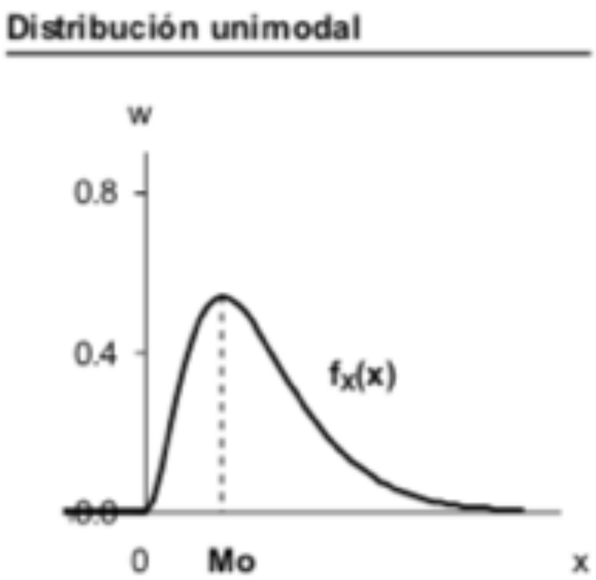
Para el caso continuo, Si X es una v.a. con $f.d.p$ y $c \in \mathbb{R}$, se dice que X tiene una distribución simétrica respecto a c si:

$$f_X(c - x) = f_X(c + x)$$

Moda (Mo): Corresponde al valor del soporte de X que maximiza a $f_X(x)$, es decir, su función de densidad.

La Moda puede no ser única o incluso no existir:

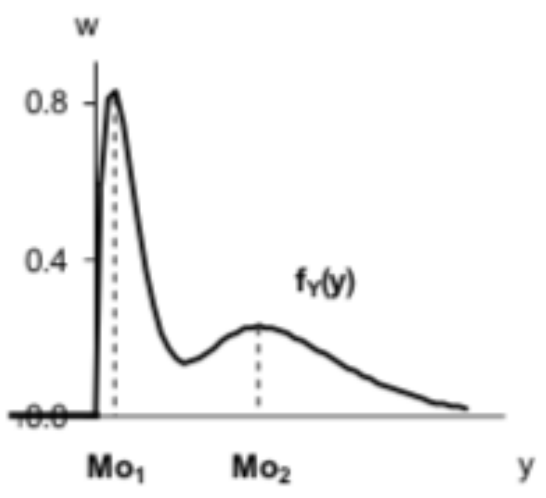
- Si $f_X(x)$ alcanza su máximo en un solo punto se, dice que X tiene una distribución unimodal.



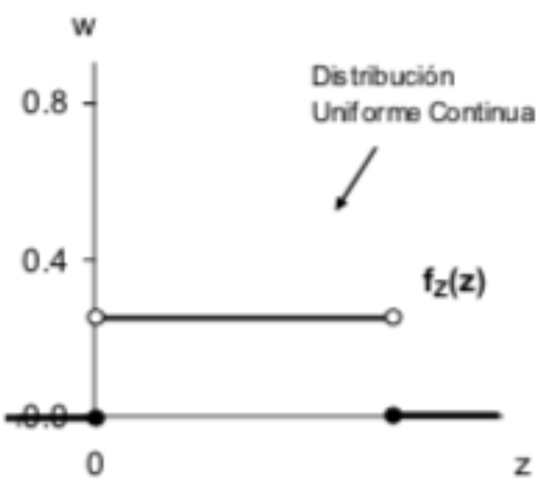
- Si $f_X(x)$ alcanza su máximo en más de un punto, se dice que X tiene una distribución multimodal. En estricto sentido, la moda es el valor real para el cual $f_X(x)$ es un máximo global, sin embargo, es común considerar como moda a los valores en los que $f_X(x)$ alcanza un máximo local.

- Si $f_X(x)$ toma el mismo valor en todos los puntos del soporte, se dice que X tiene una distribución amodal.

Distribución multimodal



Distribución amodal



Esperanza de una v.a. contínua

Ejemplo 13

Retomado el ejemplo 6, el tiempo en horas Z requerido por los estudiantes de un curso para resolver un examen final es una v.a. con función de densidad :

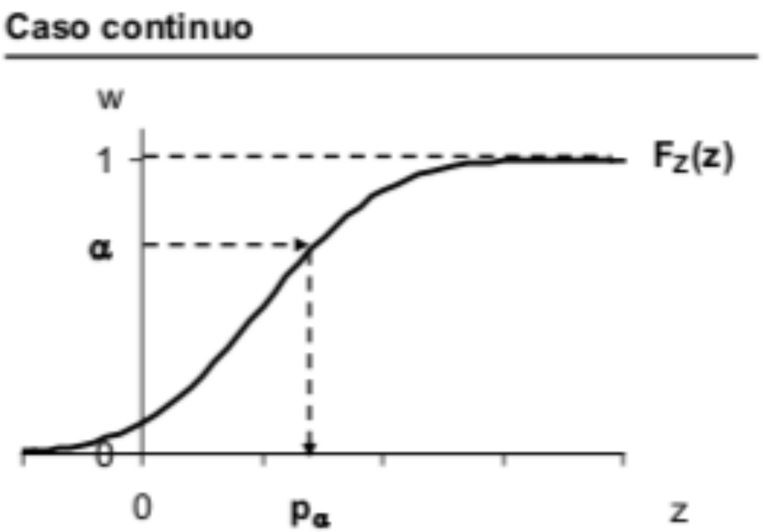
$$f_Z(Z) = \begin{cases} \dfrac{z(4-z)}{9}, & 0 \leq z \leq 3 \\ 0, & e.o.c. \end{cases}$$

a) Calcular la moda y la mediana de la v.a. Z.

Percentiles

Si X es una v.a. continua con $F_X(x)$, entonces el percentil α, 0<α<1, denotado por p_α es el valor del soporte de X tal que

$$F_X(p_\alpha) = \alpha$$



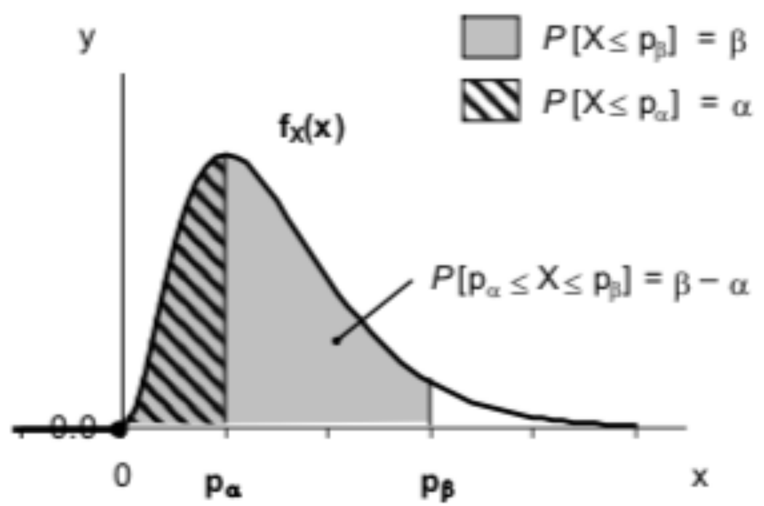
El percentil α es el valor de X por debajo del cual se acumula el $100\alpha\%$ de la distribución de probabilidades. A los percentiles 0.1, 0.2,…, 0.9 se les denomina deciles.

Percentiles

Si X es una v.a. continua y $0 < \alpha < \beta < 1$, entonces:

$$P[p_\alpha \le X \le p_\beta] = F_X(p_\beta) - F_X(p_\alpha) = \beta - \alpha$$

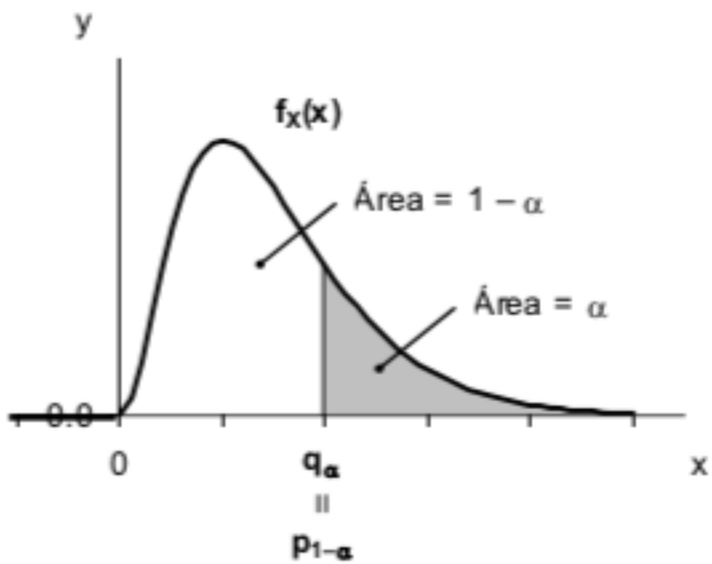
Figura 2.13



El percentil α de una distribución involucra un área a, $0 < \alpha < 1$, en la cola izquierda de la distribución de la v.a. X.

Cuantiles

El valor del soporte de X que deja un área α en la cola derecha de la distribución, se le denomina cuantil α y se denota por $q\alpha$.



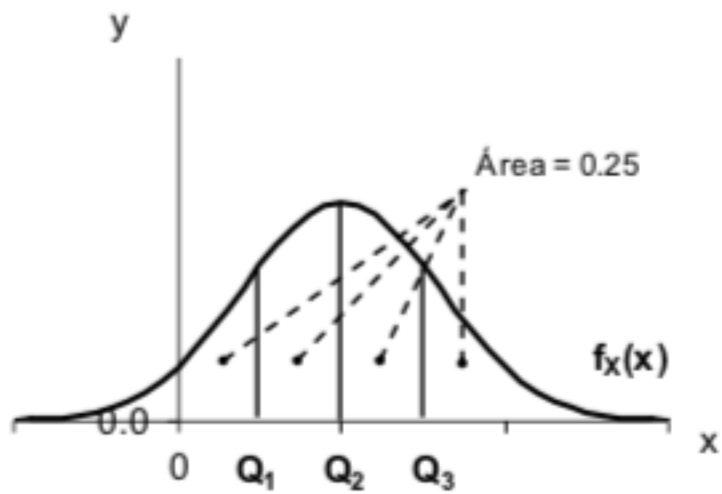
Proposición

$$q_\alpha = p_{1-\alpha}$$

Cuartiles

Si X es v.a., entonces:

a. El primer cuartil o cuartil inferior se define por $Q_1 = p_{0.25}$
b. El segundo cuartil o mediana se define por $Q_2 = p_{0.5}$
c. El tercer cuartil o cuartil superior se define por $Q_3 = p_{0.75}$



Los cuartiles son valores del soporte de X que dividen la distribución de probabilidades en 4 regiones equiprobables.

Ejemplo 14

El tiempo (en minutos) que pasa una persona en un verificentro es una variable aleatoria con función de distribución acumulada:

$$F_T(t) = \begin{cases} 0, & si\ t < 0 \\ 1 - e^{-\frac{1}{20}t}, & si\ t \geq 0 \end{cases}$$

a) Calcule la función de densidad de probabilidad de la variable aleatoria T.

b) Determine el valor de la mediana de la distribución de T.

c) Calcule el percentil 0.10 de la distribución de probabilidad de la variable aleatoria T. Interprete dicho valor.

d) Determine el valor de p_α(el percentil α) para 0 < α < 1.

e) ¿Qué tiempo de espera máximo debiera comunicar el verificentro a sus clientes para asegurar que será correcto en el 95% de los casos?

Ejemplo 14 (*continuación*)

El tiempo (en minutos) que pasa una persona en un verificentro es una variable aleatoria con función de distribución acumulada:

$$F_T(t) = \begin{cases} 0, & si\ t < 0 \\ 1 - e^{-\frac{1}{20}t}, & si\ t \geq 0 \end{cases}$$

f) Considere el intervalo $[k,\ p_{0.95}]$. ¿Qué valor debe tomar k para afirmar que el tiempo de espera se ubica en este intervalo con una probabilidad de 90%?

g) Calcule los cuartiles de la distribución de la variable aleatoria T. Ubique estos valores en la gráfica de la función de densidad de T.

Medidas de dispersión

Coeficiente de Variación (CV)

Si X es variable aleatoria, su esperanza existe y es diferente de cero, entonces el coeficiente de variación se define por:

$$CV[X] = \frac{\sigma_X[X]}{E[X]}$$

Este coeficiente es adimensional ya que la desviación estándar y la esperanza están expresadas en las mismas unidades y en ocasiones se expresa como %. De hecho es una medida de dispersión relativa.

Amplitud intercuantílica (AI)

Mide la variabilidad de la distribución de la v.a. X considerando la distancia entre el primer y tercer cuartiles.

$$AI = Q_3 - Q_1$$

Medidas de dispersión

Desviación Media respecto a la Media

Si X es v.a. con $\mu = E[X] < \infty$ su desviación media respecto a la media se define por:

$$DM(\mu) = E[|X - \mu|]$$

La transformación $g(X) = |X - \mu|$ es la variación absoluta de la v.a. X respecto de su media $\mu = E[X]$. La desviación media respecto a la media mide la variación absoluta promedio de una v.a. respecto de su media.

El concepto de desviación media se puede aplicar respecto de otras medidas de tendencia central como la moda o la mediana:

- Desviación media respecto a la moda: $DM(Mo) = E[|X - Mo|]$, siempre que X tenga distribución unimodal.
- Desviación media respecto a la mediana: $DM(Med) = E[|X - Med|]$

Medidas de dispersión

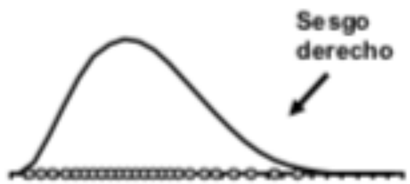
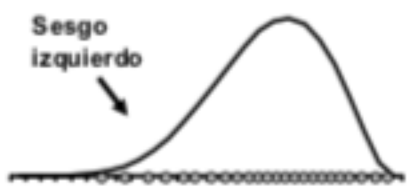
Desviación Estándar

La desviación estándar es preferida sobre la desviación media respecto a la media por que, en general, el cálculo de σ_X es el más simple que el de $DM(\mu)$ y por las propiedades matemáticas de la transformación $g(X) = (X - \mu)^2$ (función cuadrática) frente a las de $g(X) = |X - \mu|$ (función valor absoluto).

En general la desviación estándar sobredimensiona la variabilidad en comparación con la desviación media respecto a la media.

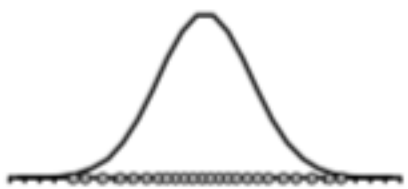
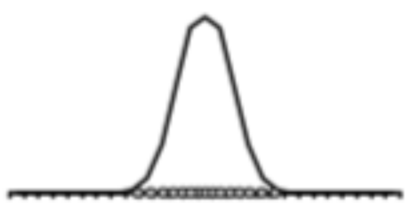
Medidas de simetría y forma

Las distribuciones de probabilidad (o curvas poblacionales) se clasifican de acuerdo con su grado de simetría en simétricas o asimétricas (con sesgo derecho - positivo o izquierdo - negativo)

Tipo de curva	Gráfica	Caracteristicas
Simétrica		• Existe un valor central alrededor del cual se concentran los datos • Son igualmente probables valores superiores o inferiores
Asimétrica con sesgo derecho		• Su extremidad o cola derecha está más extendida • Presenta un corte en el lado derecho de su figura • Hay mayor concentración de datos en los valores pequeños
Asimétrica con sesgo izquierdo		• Su extremidad o cola izquierda está más extendida • Presenta un corte en el lado izquierdo de su figura • Hay mayor concentración de datos en los valores grandes

Medidas de simetría y forma

De acuerdo con su picudez (o planura), las distribuciones de probabilidad se clasifican en mesocúticas (curva media), platicúrticas (curva plana) o leptocúrticas (curva pronunciada)

Tipo de curva	Gráfica	Características
Mesocúrtica		• Presenta un pico mediano • Los datos presentan un grado de concentración moderado
Platicúrtica		• Presenta un pico poco pronunciado • Los datos presentan poco grado de concentración y gran dispersión
Leptocúrtica		• Presenta un pico muy pronunciado • Los datos presentan alto grado de concentración y poca dispersión

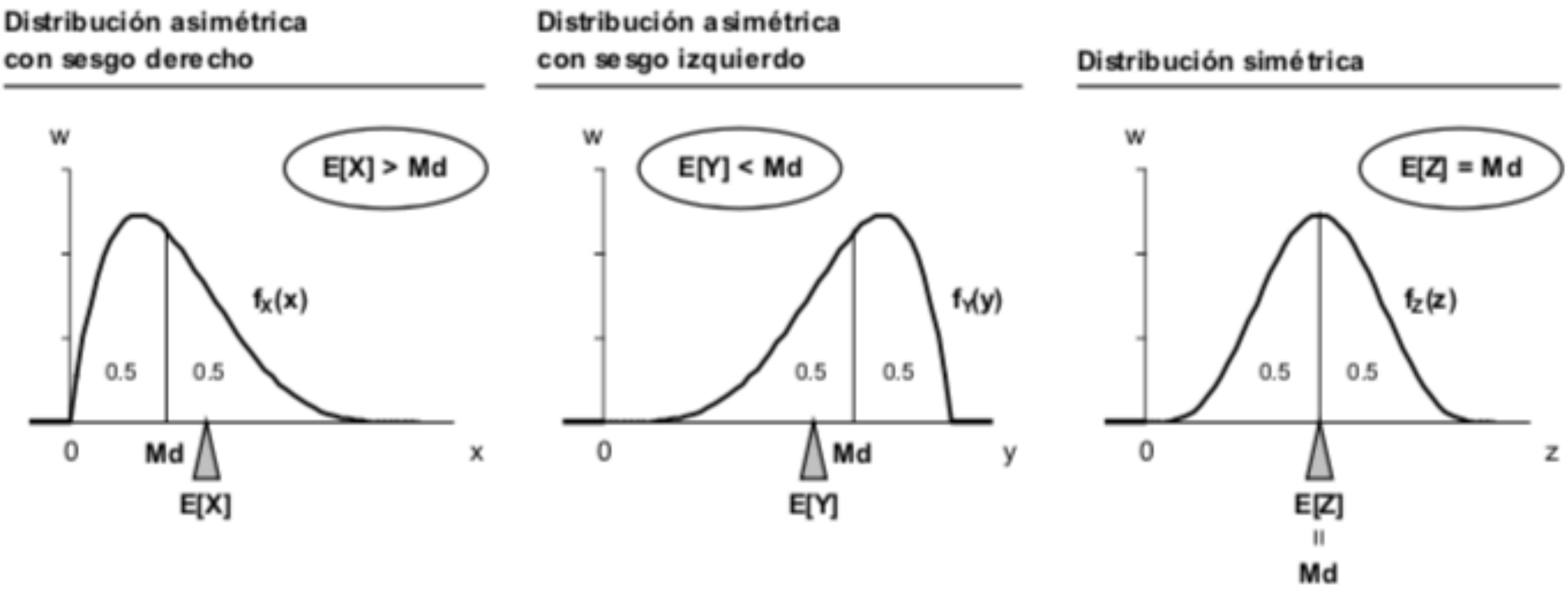
Medidas de simetría y forma

Si X tiene distribución unimodal y $E[X] < \infty$, entonces:

i) $E[X] > Mod \implies X$ tiene distribución asimétrica con sesgo derecho

ii) $E[X] < Mod \implies X$ tiene distribución asimétrica con sesgo izquierdo

iii) $E[X] = Mod \implies X$ tiene distribución simétrica.



Coeficientes de asimetría y curtosis

Si X es una v.a. con σ_X y $\mu_K = E[(X - E[X])^k], k \in \mathbb{Z}^+$, entonces:

i) $CA[X] = \alpha_3 = \frac{\mu_3}{\sigma_X^3}$ define su coeficiente de asimetría

ii) $CC[X] = \alpha_4 = \frac{\mu_4}{\sigma_X^4}$ define su coeficiente de curtosis.

Importante: El cálculo de los coeficientes de asimetría y curtosis involucra el cálculo de momentos centrales. Estos últimos pueden calcularse via momentos no centrales.

El coeficiente de asimetría permite clasificar una distribución de probabilidad de acuerdo con su grado de simetría.

Coeficientes de asimetría y curtosis

Interpretación del coeficiente de asimetría

i) $CA[X] > 0 \Rightarrow X$ tiene una distribución asimétrica con sesgo derecho o positivo.
ii) $CA[X] < 0 \Rightarrow X$ tiene una distribución asimétrica con sesgo izquierdo o negativo.
iii) $CA[X] < 0 \Rightarrow X$ tiene una distribución simétrica (sin sesgo).

Interpretación del coeficiente de curtosis

i) $CC[X] > 3 \Rightarrow X$ tiene una distribución leptocútica.
ii) $CC[X] < 3 \Rightarrow X$ tiene una distribución platicúrtica.
iii) $CC[X] < 3 \Rightarrow X$ tiene una distribución mesocútica.

Coeficientes de asimetría y curtosis

Ejemplo 15

Se sabe que si X tiene una Distribución Uniforme Continua sobre el intervalo $[\theta_1, \theta_2]$ con $\theta_1, \theta_2 \in \mathbb{R}$ y $\theta_1 < \theta_2$, entonces $f_X(x) = \frac{1}{\theta_2 - \theta_1} I_{[\theta_1, \theta_2]}(x)$. Además en el ejemplo 10 se demostró que

$$\mu_k = \frac{1}{(k+1)(\theta_2 - \theta_1)} \left[\left(\frac{\theta_2 - \theta_1}{2} \right)^{k+1} - \left(-\frac{\theta_2 - \theta_1}{2} \right)^{k+1} \right].$$

a) Calcular el *coeficiente de asimetría de X*. Interprete.

b) Calcular el *coeficiente de curtosis de X*. Interprete

Bibliografía

1. Wackerly, D.D., Mendenhall, W., and Scheaffer, R.L. (2008), *Mathematical Statistics with Applications* 7th Edition, Duxbury, Thompson, Brooks/Cole.

2. Wackerly, D.D., Mendenhall, W., and Scheaffer, R.L. (2010), *Estadística Matemática con aplicaciiones* 7ma Edición, CENGAGE Learning.

3. Pitman, J. (1993), *Probability*. Springer 6ª. Ed.

4. Ross, S. (1993), *A First Course in Probability*. Pearson 9th. Ed.

5. Canavos, G.C. (1987), *Probabilidad y Estadística*, McGraw Hill.

6. Mittelhammer (2013),*Mathematical Statistics for Economics and Business*, 2nd Ed. Springer.