TEMA 4. CARACTERISTICAS DE LAS VARIABLES ALEATORIAS

La función de distribución y las de probabilidad y densidad contienen toda la información relevante sobre las propiedades de la v.a. En muchos casos tan sólo necesitamos conocer alguna propiedad de estas variables.
Dentro de estas propiedades o características destacan dos muy importantes: las medidas de centralización o localización (media, mediana y moda) y las medidas de dispersión (varianza y desviación típica), así como otras que ya veremos más adelante.

Con las medidas de centralización pretendemos obtener un valor significativo que pueda sustituir al conjunto de valores que puede tomar la v.a. Las medidas de dispersión nos indican precisamente si el valor tomado como medida de centralización representa bien o no a los valores de la v.a.

Esperanza matemática o valor esperado

a) Variable aleatoria discreta

Sea X una v.a. que puede tomar los valores (x_i), $i=1,2,\ldots$
Se define la media, valor esperado o esperanza de X como:

$$E[X] = \sum_i x_i \cdot P[X=x_i]$$

La esperanza de una v.a X, $E[X]$, no tiene porqué existir ya que la serie $\sum_i x_i \cdot P[X=x_i]$ no tiene porqué ser convergente.

Veamos un ejemplo (ejercicio 1): Sea X el número de llantas defectuosas en un lote de 4 llantas. Su distribución de probabilidad es:

x_i	0	1	2	3	4
$P[X=x_i]$	0'6561	0'2916	0'0486	0'0036	0'0001

Determínese el nº esperado de llantas defectuosas por lote

$E[X] = \sum_i x_i \cdot P[X=x_i] = 0 \cdot 0'6561 + 1 \cdot 0'2916 + 2 \cdot 0'0486 + 3 \cdot 0'0036 +$
$+ 4 \cdot 0'0001 = 0'04 \Rightarrow E[X] = 0'04$

b) Variable aleatoria continua

Sea X una v.a. continua con función de densidad $f(x)$. Se define la media, valor esperado o esperanza de X como:

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

No está garantizada la existencia de la esperanza de X, pues la integral:

$$\int_{-\infty}^{+\infty} x \cdot f(x)\, dx$$ no tiene porqué ser convergente.

Veamos un ejemplo (ejercicio 5): El rendimiento de metal puro de un mineral, expresado en tanto por ciento, es una v.a. cuya función de densidad viene dado por:

$$f(x) = \begin{cases} \frac{x^2}{9} & \text{si } 0 < x < 3 \\ 0 & \text{en el resto} \end{cases}$$

Calcule el rendimiento medio de cada piedra.

$$E[X] = \int_{-\infty}^{+\infty} x\, f(x)\, dx = \int_0^3 x \cdot \frac{x^2}{9} = \frac{1}{9} \int_0^3 x^3\, dx = \frac{1}{9} \left[\frac{x^4}{4} \right]_0^3$$

$$E[X] = \frac{1}{9} \cdot \frac{3^4}{4} = \frac{81}{36} = \frac{9}{4}$$

El rendimiento medio es $E[X] = \frac{9}{4}$

Observación: función de una variable aleatoria

Si X es una variable aleatoria, una función de la misma es $g[X]$ donde $g: \mathbb{R} \longrightarrow \mathbb{R}$. Si X toma el valor x, la función de la v.a. $g[X]$ toma el valor $g(x)$.

c) Esperanza de una función de una v.a.

- Caso discreto

Sea X una v.a. discreta que puede tomar los valores (x_i) $i=1,2,\dots$ Sea $g[X]$ una función de la v.a. X. Se define la esperanza de $g[X]$ como:

$$E[g[X]] = \sum_i g(x_i) \cdot P[X = x_i] \qquad i = 1,2,\dots$$

De nuevo tenemos que observar que $E[g[X]]$ no tiene porqué existir, todo depende de que la serie $\sum_i g(x_i) \cdot P[X=x_i]$ sea convergente.

Veamos un ejemplo: sea X una v.a. discreta con función de probabilidad:

x_i	0	1	2	3	4
$P[X=x_i)$	0'6561	0'2916	0'0486	0'0036	0'0001

Determínese la esperanza de $g[X] = X^3$.

$E[g[X]] = \sum_i x_i^3 \cdot P[X=x_i] = 0^3 \cdot 0'6561 + 1^3 \cdot 0'2916 + 2^3 \cdot 0'0486 + 3^3 \cdot 0'0036 +$
$+ 4^3 \cdot 0'0001 = 0'784$; luego $E[X^3] = 0'784$

Observemos que se verifica que $E[X^3] \neq (E[X])^3$

ya que $0'784 \neq 0'04^3$

- Caso continuo

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

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

De nuevo insistimos en que no tiene porqué existir la esperanza $E[g[X]]$, todo depende de la convergencia de la integral $\int_{-\infty}^{\infty} g(x) \cdot f(x)\, dx$

Veamos un ejemplo: Sea X una v.a. cuya función de densidad viene dado por:

$$f(x) = \begin{cases} \dfrac{x^2}{9} & \text{si } 0 < x < 3 \\ 0 & \text{en el resto} \end{cases}$$

Calcule la esperanza de la función $g[X] = 2X + X^2$

$$E[g[X]] = E[2X+X^2] = \int_{-\infty}^{\infty} (2x+x^2)\cdot f(x)\,dx = \int_0^3 (2x+x^2)\cdot\frac{x^2}{9}\,dx$$

$$E[2X+X^2] = \frac{2}{9}\int_0^3 x^3\,dx + \frac{1}{9}\int_0^3 x^4\,dx = \frac{2}{9}\left[\frac{x^4}{4}\right]_0^3 + \frac{1}{9}\left[\frac{x^5}{5}\right]_0^3$$

$$E[2X+X^2] = \frac{162}{36} + \frac{243}{45} = \frac{99}{100}$$

Vamos a enunciar y a demostrar una por una las propiedades de la esperanza. En los apuntes del grupo piloto tan sólo se exige el enunciado, como las demostraciones son tan fáciles merece la pena que las mires para que tengas un recurso más pero no te las aprendas las demostraciones de memoria.

Propiedades de la esperanza

1ª $E[C] = C \quad \forall C \in \mathbb{R}$, C = constante.

Demostración caso discreto

La variable aleatoria X toma valores (x_i) $i = 1, 2, \dots$ y la función de la v.a. es $g[X] = C$. Entonces:

$$E[g[X]] = E[C] = \sum_i g(x_i)\cdot P[X = x_i] = \sum_i C\cdot P[X = x_i]$$

$$E[C] = C\cdot\sum_i P[X = x_i] \Rightarrow E[C] = C\cdot 1 = C.$$

Demostración caso continuo

La v.a. X tiene como función de densidad $f(x)$

La función de la v.a. X es $g[X] = c$ (constante)
Entonces:

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

$$E[c] = c \cdot \int_{-\infty}^{\infty} f(x)\,dx = c \cdot 1 = c.$$

2ª Sea X v.a y a, b constantes. Entonces:

$$E[a+bX] = a + bE[X]$$

Demostración caso discreto

X v.a. discreta que toma valores (x_i) $i = 1, 2, \ldots$
$g[X] = a + bX$. Entonces:

$$E[g[x]] = E[a+bx] = \sum_i (a + b x_i) \cdot P[X = x_i]$$

$$E[a+bX] = \sum_i a \cdot P[X = x_i] + b \cdot \sum_i x_i \cdot P[X = x_i]$$

$$E[a+bX] = a \cdot \sum_i P[X = x_i] + b \cdot E[X] = a \cdot 1 + b \cdot E[X].$$

Luego $E[a+bX] = a + b \cdot E[X]$

Demostración caso continuo

X v.a. con función de densidad $f(x)$. Sea la función $g[X] = a + bX$, a, b constantes. Entonces

$$E[g[x]] = E[a+bX] = \int_{-\infty}^{\infty} (a+bx) \cdot f(x)\,dx = \int_{-\infty}^{\infty} [a\,f(x) + b\,x\,f(x)]\,dx$$

$$E[a+bX] = a \cdot \int_{-\infty}^{\infty} f(x)\,dx + b \int_{-\infty}^{\infty} x \cdot f(x)\,dx = a \cdot 1 + b \int_{-\infty}^{\infty} x\,f(x)\,dx$$

Luego $E[a+bX] = a + b \cdot E[X]$.

3ª Sea X una v.a., a y b constantes y $g[X]$, $h[X]$ dos funciones de X. Entonces:

$$E[a \cdot g[X] + b \cdot h[X]] = a \cdot E[g[X]] + b \cdot E[h[X]]$$

Demostración caso discreto

Sea X v.a. que toma valores (x_i) $i=1,2,\ldots$ con función de probabilidad $P[X=x_i]$. Entonces:

$$E[a\,g[X]+b\,h[X]] = \sum_i [a\,g(x_i)+b\,h(x_i)]\cdot P[X=x_i] =$$

$$= \sum_i a\cdot g(x_i)\cdot P[X=x_i] + b\cdot h(x_i)\cdot P[X=x_i].$$ Luego:

$$E[a\cdot g[X]+b\,h[X]] = a\cdot\sum_i g(x_i)\cdot P[X=x_i] + b\cdot\sum_i h(x_i)\cdot P[X=x_i]$$

$$E[a\cdot g[X]+b\,h[X]] = a\cdot E[g[X]] + b\cdot E[h[X]]$$

Demostración caso continuo

Sea X v.a. con función de densidad $f(x)$. Entonces:

$$E[a\cdot g[X]+b\cdot h[X]] = \int_{-\infty}^{\infty}[a\cdot g(x)+b\cdot h(x)]\,f(x)\,dx = a\cdot\int_{-\infty}^{\infty} g(x)f(x)dx +$$

$$+b\int_{-\infty}^{\infty} h(x)\cdot f(x)\,dx = a\cdot E[g[X]] + b\cdot E[h(x)].$$

4º Si X es v.a entonces $E[X-E[X]]=0$

Demostración para los dos casos (continuo y discreto)

Al ser $E[X]$ una constante

$$E[X-E[X]] = E[X] - E[E[X]] = E[X]-E[X] = 0$$

5º Sea X una v.a., sean $g[X]$ y $h[X]$ dos funciones de X tales que $g[X]\leq h[X]$. Entonces se verifica que:

$$E[g[X]] \leq E[h[X]]$$

Demostración caso discreto

Sea X v.a. que toma valores (x_i) $i=1,2,\ldots$

Como $g[X]\leq h[X] \Rightarrow g(x_i)\leq h(x_i)$ $\forall i$. Por tanto:

$$E[g[X]] = \sum_i g(x_i)\cdot P[X=x_i] \leq \sum_i h(x_i)\cdot P[X=x_i] = E[h[X]]$$

Demostración caso continuo

Sea X v.a. con función de densidad $f(x)$.
Si $g[X] \leq h[X] \Rightarrow g(x) \leq h(x) \quad \forall x \in \mathbb{R}$. Entonces:

$$E[g[X]] = \int_{-\infty}^{\infty} g(x) \cdot f(x)\, dx \leq \int_{-\infty}^{\infty} h(x) \cdot f(x)\, dx = E[h[X]]$$

6ª Si $P[X \geq 0] = 1$ y existe $E[X] \Rightarrow E[X] \geq 0$

Demostración:

Si $P[X \geq 0] = 1 \Rightarrow P[X < 0] = 1 - P[X \geq 0] = 1 - 1 = 0 \Rightarrow$

$\Rightarrow X < 0$ es el suceso imposible $\Rightarrow X$ solo toma valores positivos. Por tanto:

$E[X] = \sum_i x_i \cdot P[X = x_i] \geq 0$ ya que $x_i \geq 0$ y $P[X = x_i] \geq 0$

MEDIDAS DE DISPERSIÓN

a) Varianza

Sea X una v.a. con media $\mu = E[X]$. La varianza de X la representamos por σ^2, $Var[X]$, $V[X]$ y se define como:

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

Observación respecto a la notación

Hemos de llevar mucho cuidado con la notación que empleamos. En realidad sería mejor, si hay posibilidad de confusión, que pusiéramos:

$\sigma_X^2 = Var[X] = E[(X - \mu_X)^2]$, $\mu_X = E[X]$. Con esta notación, si por ejemplo $Y = \alpha X + \beta$, α y β constantes, tendríamos que:

$\sigma_Y^2 = Var[Y] = E[(Y - \mu_Y)^2]$. Ahora bien $\mu_Y = E[\alpha X + \beta] \Rightarrow$

$\Rightarrow \mu_Y = \alpha \cdot E[X] + \beta \Rightarrow \mu_Y = \alpha \mu_X + \beta$

Por tanto:

$$\sigma_y^2 = E[(Y-\mu_y)^2] = E[(\alpha X+\beta-\alpha\mu_x-\beta)^2] = E[\alpha^2(X-\mu_x)^2]$$

$\sigma_y^2 = E[\alpha^2(X-\mu_x)^2] = \alpha^2 \cdot E[(X-\mu_x)^2] \Rightarrow \sigma_y^2 = \alpha^2 \cdot \sigma_x^2$. Estas cuestiones las vamos a ver a continuación.

Vamos a enunciar 6 propiedades de la varianza. Haremos una por una las demostraciones que no te has de estudiar de memoria pero sí entenderlas. Son bastante fáciles.

PROPIEDADES DE LA VARIANZA

1ª $Var[X] = E[(X-\mu)^2] \geq 0$

Demostración:

Hemos visto una propiedad de la esperanza matemática que decía que si $g[X] \geq h[X]$ entonces $E[g[X]] \geq E[h[X]]$

Como $(X-\mu)^2 \geq 0 \Rightarrow E[(X-\mu)^2] \geq E[0] = 0$ (la esperanza de una constante es ella misma)

2ª Si C es constante entonces $Var[C] = 0$

Demostración:

$Var[C] = E[(C-\mu_C)^2]$ donde $\mu_C = E[C] = C$. Luego:

$Var[C] = E[(C-C)^2] = E[0] = 0$

3ª El valor de c para el cual la expresión $E[(X-c)^2]$ alcanza un valor mínimo es $c = \mu = E[X]$.

Demostración:

$$E[(X-c)^2] = E[X^2 - 2cX + c^2] = E[X^2] - 2c\,\overset{\mu}{\overset{\shortparallel}{E[X]}} + c^2$$

Luego:

$E[(X-c)^2] = E[X^2] - 2\mu c + c^2$. Si queremos que se alcance un valor mínimo variando c:

$$\frac{d\,E[(X-c)^2]}{dc} = 0 \Rightarrow -2\mu + 2c = 0 \Rightarrow 2c = 2\mu \Rightarrow c = \mu.$$

4ª $\sigma^2 = E[x^2] - E[x]^2$

Demostración:

$\sigma^2 = Var[x] = E[(x-\mu)^2] = E[x^2 - 2\mu x + \mu^2]$ siendo $\mu = E[x]$

Por tanto:

$\sigma^2 = E[x^2] - 2\mu E[x] + \mu^2 \Rightarrow \sigma^2 = E[x^2] - 2\mu^2 + \mu^2 = E[x^2] - \mu^2$

Como $\mu = E[x]$ nos queda que:

$\sigma^2 = E[x^2] - E[x]^2$

5ª $Var[x \pm a] = Var[x] \quad \forall a \in \mathbb{R}$

Demostración:

Sea $Y = x + a$ entonces $\mu_y = E[x+a] = E[x] + a = \mu_x + a$

Por lo tanto:

$\sigma_y^2 = Var[Y] = Var[x+a] = E[(Y-\mu_y)^2] = E[(x + \cancel{a} - \mu_x - \cancel{a})^2] \Rightarrow$

$\Rightarrow Var[x+a] = E[(x-\mu_x)^2] = Var[x]$

Del mismo modo se demostraría que $Var[x-a] = Var[x]$ tomando $Y = x - a$.

6ª $Var[bx] = b^2 Var[x] \quad \forall b \in \mathbb{R}$

Demostración:

Sea $Y = bX \Rightarrow \mu_y = E[Y] = E[bx] = b \cdot E[x] = b \cdot \mu_x$

Se tiene que:

$\sigma_y^2 = Var[Y] = Var[bx] = E[(Y-\mu_y)^2] = E[(bx - b\mu_x)^2] \Rightarrow$

$Var[bx] = E[b^2(x-\mu_x)^2] = b^2 \cdot E[(x-\mu_x)^2]$. luego:

$Var[bx] = b^2 \cdot Var[x]$.

Corolario (consecuencia): $Var[a+bx] = b^2 \cdot Var[x] \quad \forall a, b \in \mathbb{R}$

Demostración:

Según la propiedad 5ª $Var[a+bx] = Var[bx]$.

Según la propiedad 6ª:

$Var[a+bx] = Var[bX] = b^2 \cdot Var[x]$.

La varianza se puede usar para conocer la proximidad de los valores de la variable aleatoria X a su media. A grandes valores de la varianza corresponden grandes dispersiones de X sobre μ y pequeños valores de σ^2 corresponden a pequeñas dispersiones de la variable respecto a μ. La varianza no se mide en las mismas unidades que X por eso se introduce el concepto de desviación típica que sí se mide en las mismas unidades que X. Por lo demás la desviación típica juega el mismo papel que la varianza: medir la dispersión de la variable aleatoria X respecto a $\mu = E[X]$

Desviación típica

La desviación típica es la raíz cuadrada de la varianza. Se representa por σ

$$\sigma = \sqrt{V[X]} = \sqrt{Var[X]} = \sqrt{E[(X-\mu)^2]} \quad \text{siendo } \mu = E[X]$$

Propiedades de la desviación típica

1ª $\sigma \geq 0$

Demostración:

$\sigma = \sqrt{E[(X-\mu)^2]} \geq 0$ por definición de raíz cuadrada

2ª $\sigma(c) = 0$

Demostración:

Sabemos que $\sigma^2(c) = 0 = Var[c]$ (2ª propiedad de la varianza) $\Rightarrow \sigma(c) = \sqrt{0} = 0$

3ª $\sigma = \sqrt{E[X^2] - E[X]^2}$

Demostración:

Como $\sigma^2 = Var[X] = E[X^2] - E[X]^2 \Rightarrow \sigma = \sqrt{E[X^2] - E[X]^2}$

4ª $\sigma_{X+a} = \sigma_X$

Demostración

$\sigma_{X+a}^2 = Var[X+a] = Var[X]$ (propiedad 5 de la varianza) $= \sigma_X^2$

Como $\sigma_{x+a}^2 = \sigma_x^2 \Rightarrow \sigma_{x+a} = \sigma_x$

5ª $\sigma_{bx} = |b| \cdot \sigma_x$

Demostración:

$\sigma_{bx}^2 = Var[bx] = b^2 Var[x] = b^2 \sigma_x^2$ (propiedad 6ª de la varianza)

Luego $\sqrt{\sigma_{bx}^2} = \sqrt{b^2 \sigma_x^2} \Rightarrow \sigma_{bx} = |b| \cdot \sigma_x$

6ª $\sigma_{a+bx} = |b| \cdot \sigma_x$

Demostración:

$\sigma_{a+bx} = \sigma_{bx} = |b| \cdot \sigma_x$

Veamos un ejemplo (Ejercicio 1):
Sea X el número de llantas defectuosas en un lote de 4 llantas moldeadas simultáneamente. Supongamos que la distribución de probabilidad de X es:

x_i	0	1	2	3	4
$P[X=x_i]$	0'6561	0'2916	0'0486	0'0036	0'0001

a) Calcule la varianza y la desviación típica de X
b) Si una llanta defectuosa supone una penalización de 1000 u.m., calcule la penalización esperada por lote, así como la varianza de dicha penalización

a) $Var[X] = E[(X-\mu)^2] = \sum_i (x_i - \mu)^2 \cdot P[X = x_i]$

$\mu = E[X] = \sum x_i \cdot P[X = x_i] = 0 \cdot 0'6561 + 1 \cdot 0'2916 + 2 \cdot 0'0486 + 3 \cdot 0'0036 +$
$+ 4 \cdot 0'0001 \Rightarrow \mu = 0'4$. Entonces

$Var[X] = \sum_{i=1}^{4} (x_i - 0'4)^2 \cdot P[X = x_i] = (0-0'4)^2 \cdot 0'6561 + (1-0'4)^2 \cdot$
$\cdot 0'2916 + (2-0'4)^2 \cdot 0'0486 + (3-0'4)^2 \cdot 0'0036 + (4-0'4)^2 \cdot 0'0001 = 0'36$

La varianza es $\sigma^2 = 0'36$. Por tanto la desviación típica es $\sigma = \sqrt{0'36} = 0'6 \Rightarrow \sigma = 0'6$

b) Si X es el número de llantas defectuosas por lote, la penalización por lote será $Y = 1000X$

La penalización esperada por lote es:

$E[Y] = E[1000\,X] = 1000\,E[X] = 1000 \cdot 0'4 = 400$

La varianza de la penalización es:

$Var[Y] = Var[1000\,X] = 1000^2 \cdot Var(x) = 10^6 \cdot 0'36 = 36 \cdot 10^4 = 360.000 \text{ um}^2$

Medidas de posición.

a) Moda

La moda es el valor más probable de la variable aleatoria. Será pues el valor para el cual la función de probabilidad (si es variable discreta) o la función de densidad (si es variable continua) se hace máxima.

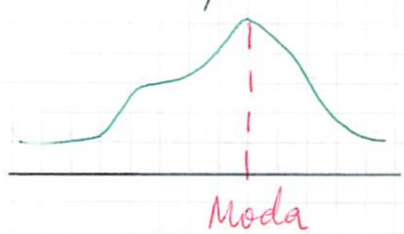
En variable discreta la moda es el valor $X = Mo$ tal que

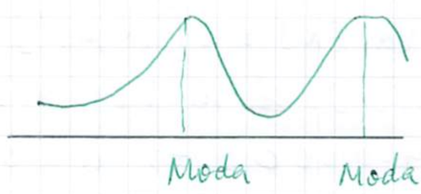
$$P[X = Mo] \geq P[X = x_i] \quad \forall i$$

En variable continua

$$f(Mo) \geq f(x) \quad \forall x$$

Una función puede tener más de un máximo por lo que puede haber más de una moda. Así una distribución puede ser unimodal (una sola moda), bimodal (dos modas) o plurimodal (más de dos modas). Gráficamente podríamos representar esta situación como sigue:





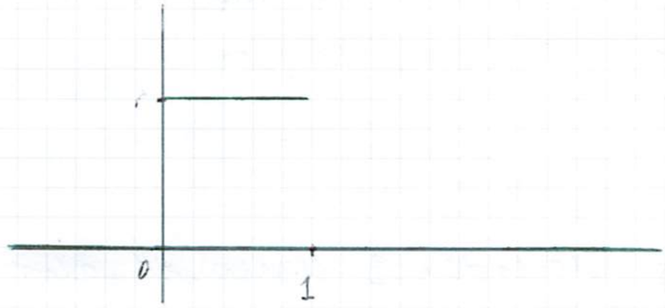
Veamos ejemplos.

Dado la función de densidad

$$f(x) = \begin{cases} 1 & \text{si } 0 < x < 1 \\ 0 & \text{en el resto} \end{cases}$$

Hallar el valor de la moda

Hacemos una representación gráfica



La moda es cualquier valor comprendido entre 0 y 1

Dada la función de densidad

$$f(x) = \begin{cases} e^{-x} & si \ x \in (0, +\infty) \\ 0 & resto \end{cases}$$

Hallar el valor de la moda

$f'(x) = -e^{-x} \neq 0 \quad \forall x \Rightarrow$ no hay máximo $\Rightarrow$ no existe la moda.

Dada la función de densidad

$f(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \quad \forall x \in \mathbb{R}$, hallar la moda.

$f'(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}} \cdot (-x) = 0 \Rightarrow x = 0$. luego la moda es $Mo = 0$.

b) Mediana

La mediana Me es un valor de la variable aleatoria que verifica:

$F(Me) = P[X \leq Me] = \frac{1}{2}$ o lo que es lo mismo:

$P[X > Me] = 1 - P[X \leq Me] = \frac{1}{2}$

- Si X es v.a. continua, siempre existe Me

$$F(Me) = \int_{-\infty}^{Me} f(x)\,dx = \frac{1}{2} = \int_{Me}^{+\infty} f(x)\,dx$$

• Si X es v.a. discreta no siempre existe este valor, por ello para estas variables la mediana Me es el valor más pequeño de la v.a. que verifica:

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

Cuantiles

Se define el cuantil C_q de orden q como el valor más pequeño de la v.a. X que verifica:

$F(C_q) = P[X \leq C_q] = q \quad 0 < q < 1$ si X es v.a. continua

$F(C_q) = P[X \leq C_q] \geq q \quad 0 < q < 1$ si X es v.a. discreta

Los casos particulares más importantes de cuantiles son:

- Cuartiles: $Q_i \quad i = 1, 2, 3 \quad q_i = \frac{1}{4}, \frac{2}{4}, \frac{3}{4}$

$F(Q_i) = P[X \leq Q_i] = \frac{i}{4}$ (v.a. continua) $i = 1 \cdots 4$

$F(Q_i) = P[X \leq Q_i] \geq \frac{i}{4}$ (v.a. discreta) $i = 1 \cdots 4$

- Deciles: Hay nueve deciles $D_i = \frac{i}{10}$, $i = 1, 2, \cdots 9$

$F(D_i) = P[X \leq D_i] = \frac{i}{10}$ (v.a. continua) $i = 1 \cdots 9$

$F(D_i] = P[X \leq D_i] \geq \frac{i}{10}$ (v.a. discreta) $i = 1 \cdots 9$

- Percentiles: hay 99 percentiles P_i, $i = 1 \cdots 99$

$P_i = \frac{i}{100} \quad i = 1 \cdots 99$

$F(P_i) = P[X \leq P_i] = \frac{i}{100} \quad i = 1 \cdots 99$ (v.a. continua)

$F(P_i) = P[X \leq P_i] \geq \frac{i}{100} \quad i = 1, 2, \cdots 99$

Veamos algunos ejemplos de los conceptos que hemos definido.

Ejemplo (ejercicio 15)
Con objeto de establecer un plan de producción, una empresa ha estimado que la demanda de sus clientes se comporta semanalmente con arreglo a la ley de probabilidad dado por la función de densidad

$$f(x) = \begin{cases} \frac{1}{5} & \text{si } 4 \le x \le 9 \\ 0 & \text{en el resto} \end{cases}$$

, si x se representa en miles de unidades.
¿Qué cantidad debe tener dispuesta a la venta, al comienzo de cada semana, para poder satisfacer toda la demanda con probabilidad al menos de 0'5?

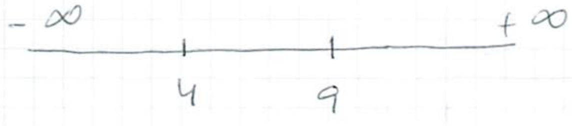
Hay que buscar un valor a de la v.a. X tal que

$P[X \le a] \geq 0'5 \Rightarrow a = Me$

$$P[X \le Me] = 0'5 = \int_0^{Me} f(x)dx = \int_4^{Me} \frac{1}{5}\, dx = \left[\frac{1}{5}x\right]_4^{Me} = 0'5$$

$$\Rightarrow \frac{Me}{5} - \frac{4}{5} = 0'5 \Rightarrow Me - 4 = 2'5 \Rightarrow \boxed{Me = 6'5}$$

Calculemos la función de distribución. La partición de la recta real para localizar los valores de X es



Si $X < 4 \Rightarrow F(x) = 0$

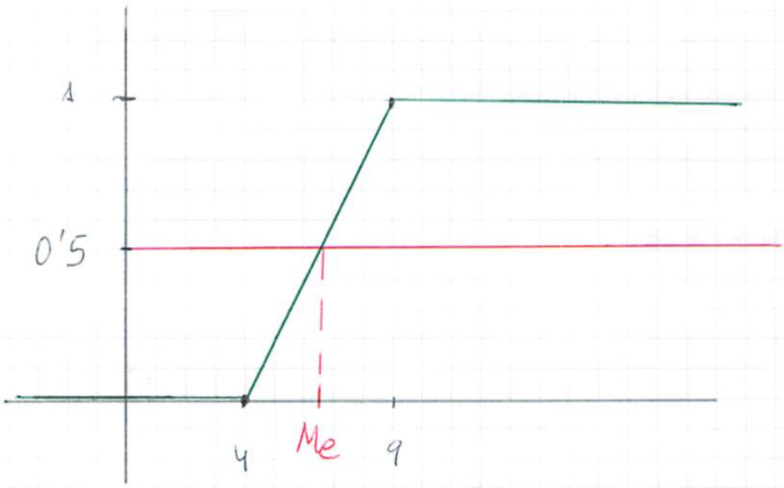
Si $4 \le x \le 9$ $\quad F(x) = \int_{-\infty}^{x} f(x)dx = \int_4^x \frac{1}{5}\, dx = \frac{1}{5}x\Big]_4^x = \frac{x-4}{5}$

Si $x \geq 9 \Rightarrow F(x) = \int_{-\infty}^{x} f(x)dx = \int_4^9 \frac{1}{5}\, dx = \frac{1}{5}x\Big]_4^9 = \frac{9}{5} - \frac{4}{5} = 1$

Luego la función de distribución es:

$$F(x) = \begin{cases} 0 & \text{si } x < 4 \\ \frac{x-4}{5} & \text{si } 4 \le x < 9 \\ 1 & \text{si } x \geq 9 \end{cases}$$

La representación gráfica de $F(x)$ es:



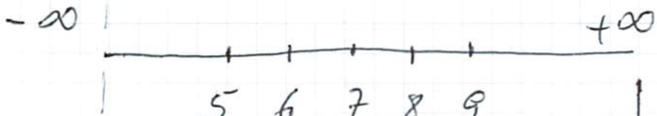
Ejemplo (ejercicio 16): Repita el ejercicio anterior suponiendo que la demanda se comporta con arreglo a la ley de probabilidad que se da en la siguiente tabla:

x_i	5	6	7	8	9
$P[X=x_i]$	0'2	0'2	0'2	0'2	0'2

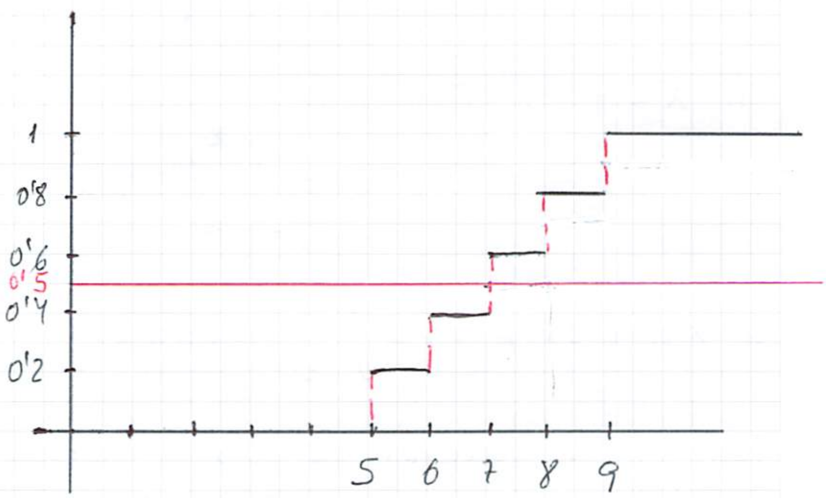
Para hallar la mediana tenemos que encontrar un valor Me de la v.a tal que:

$$P[X \leq Me] \geq \frac{1}{2} \Rightarrow F(Me) \geq \frac{1}{2}$$

El campo de variación de X viene dado por la siguiente partición:



$$F(x) = \begin{cases} 0 & \text{si } x < 5 \\ 0'2 & \text{si } 5 \leq x < 6 \\ 0'4 & \text{si } 6 \leq x < 7 \\ 0'6 & \text{si } 7 \leq x < 8 \\ 0'8 & \text{si } 8 \leq x < 9 \\ 1 & \text{si } x \geq 9 \end{cases}$$



Luego $Me = 7$

Ejemplo (ejercicio 17)

Sea X una v.a. con función de densidad

$$f(x) = \begin{cases} \frac{1}{\lambda} \cdot e^{-\frac{x}{\lambda}} & \text{si } x > 0 \\ 0 & \text{en el resto.} \end{cases}$$

Hallar la mediana y el percentil 90

La mediana Me es el valor de la v.a. que verifica

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

$$F(Me) = \int_{-\infty}^{Me} f(x)\,dx = \int_0^{Me} \frac{1}{\lambda} \cdot e^{-\frac{x}{\lambda}}\,dx = \frac{1}{2}$$

$$\int \frac{1}{\lambda} e^{-\frac{x}{\lambda}}\,dx \begin{cases} t = -\frac{x}{\lambda} \\ dt = -\frac{1}{\lambda}\,dx \Rightarrow dx = -\lambda\,dt \end{cases} = \int \frac{1}{\lambda} \cdot e^t \cdot (-\lambda\,dt) =$$

$= -e^t = -e^{-\frac{x}{\lambda}}$. Luego:

$$\int_0^{Me} \frac{1}{\lambda} e^{-\frac{x}{\lambda}}\,dx = \left[-e^{-\frac{x}{\lambda}}\right]_0^{Me} = \frac{1}{2} \Rightarrow -e^{-\frac{Me}{\lambda}} + 1 = \frac{1}{2} \Rightarrow$$

$$\Rightarrow e^{-\frac{Me}{\lambda}} = \frac{1}{2} \Rightarrow \ln e^{-\frac{Me}{\lambda}} = \ln\frac{1}{2} = \ln 1 - \ln 2$$

$$-\frac{Me}{\lambda} = -\ln 2 \Rightarrow \boxed{Me = \lambda \ln 2}$$

El percentil 90 es el valor de la v.a. que verifica que

$P[X \leq P_{90}] = F(P_{90}) = \frac{90}{100} = \frac{9}{10}$. Luego:

$$\int_{-\infty}^{P_{90}} f(x)\,dx = \frac{9}{10} \Rightarrow \int_0^{P_{90}} \frac{1}{\lambda} \cdot e^{-\frac{P_{90}}{\lambda}} = \frac{9}{10} \Rightarrow \left[-e^{-\frac{x}{\lambda}}\right]_0^{P_{90}} = \frac{9}{10} \Rightarrow$$

$$1 - e^{-\frac{P_{90}}{\lambda}} = \frac{9}{10} \Rightarrow e^{-\frac{P_{90}}{\lambda}} = \frac{1}{10} \Rightarrow \ln e^{-\frac{P_{90}}{\lambda}} = \ln\frac{1}{10} \Rightarrow$$

$$-\frac{P_{90}}{\lambda} = -\ln 10 \Rightarrow \boxed{P_{90} = \lambda \ln 10}$$

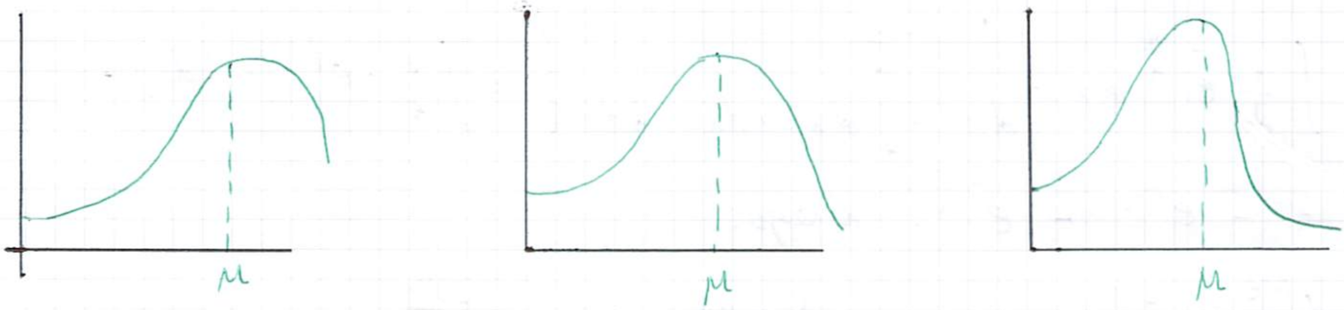
MEDIDAS DE FORMA: ASIMETRIA Y CURTOSIS

Las medidas de forma informan sobre la representación gráfica de la función de probabilidad (variable discreta) o de la función de densidad (variable continua).

Asimetría

Para medir la asimetría de una distribución respecto a un eje perpendicular al de abcisas que pasa por μ se utiliza el coeficiente de asimetría:

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



$\gamma_1 < 0$ asimetría negativa o asimétrica a la izquierda

$\gamma_1 = 0$, simétrica

$\gamma_1 > 0$ asimétrica positiva o asimétrica a la derecha

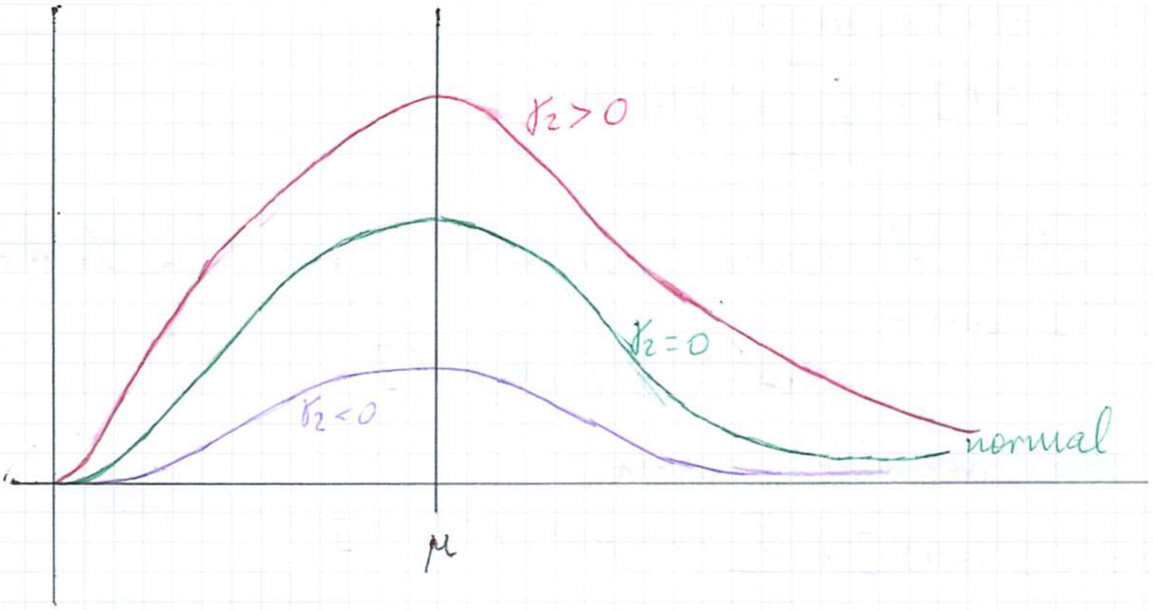
Curtosis

La curtosis o apuntamiento hace referencia al mayor o menor apuntamiento de una distribución cuando se compara con la distribución normal que es una distribución continua con función de densidad:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty$$

cuya representación gráfica se denomina campana de Gauss.

Se define el coeficiente de curtosis como $\gamma_2 = \frac{E[(x-\mu)^4]}{\sigma^4} - 3$



$\gamma_2<0$: platicúrtica, menos apuntada que la normal

$\gamma_2=0$: mesocúrtica, apuntamiento de la normal

$\gamma_2>0$: leptocúrtica, más apuntada que la normal

Para evaluar la curtosis es preciso que la distribución analizada tenga un perfil campaniforme.

MOMENTOS DE UNA VARIABLE ALEATORIA

El concepto de momento de una v.a. es una generalización del concepto de valor esperado. Los momentos se pueden definir respecto a cualquier valor de la variable aleatoria. Sea c un valor de la variable aleatoria X. Se define el momento de orden r de X respecto a c como:

$$E[(X-c)^r]$$

Los momentos más usuales son respecto al origen o con respecto a la media.

a) Se define el momento de orden r respecto al origen como:

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

• Si X es variable discreta

$$\mu_r = E[X^r] = \sum_i x_i^r \cdot P[X=x_i]$$

• Si X es variable continua

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

b) Se define el momento de orden r respecto a la media $\mu = E[X]$ como:

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

• Si X es variable discreta

$$m_r = E[(X-\mu)^r] = \sum_i (x_i-\mu)^r \cdot P[X=x_i]$$

• Si X es variable continua

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

La existencia de los momentos está sujeta a la convergencia de la serie o integral que los define.

- El momento de orden $r=1$ respecto al origen es la esperanza:

$$\mu_1 = \mu = E[X]$$

- El momento de orden $r=2$ respecto a la media es la varianza:

$$m_2 = E[(X-\mu)^2] = Var[X] = \sigma^2$$

Se puede observar que, en efecto, el concepto de momento es una generalización del concepto de esperanza ya que dando valores a r podemos obtener media, varianza, desviación típica, podemos calcular coeficientes de asimetría y curtosis, etc. Pero todo esto es un lío.

¿Qué relación existe entre el momento de orden r respecto al origen y respecto a la media?

$m_r = E[(X-\mu)^r]$, según el desarrollo del binomio de Newton:

$$(X-\mu)^r = \sum_{i=0}^{r} \binom{r}{i} (-1)^i \cdot X^{r-i} \mu^i$$

Si tenemos en cuenta las propiedades de la esperanza:

$$E[(X-\mu)^r] = E\left[\sum_{i=0}^{r} \binom{r}{i}(-1)^i \cdot X^{r-i} \cdot \mu^i\right] = \sum_{i=0}^{r} \binom{r}{i}(-1)^i \cdot \mu^i \cdot E[X^{r-i}]$$

Es decir:

$$E[(X-\mu)^r] = \sum_{i=0}^{r} \binom{r}{i}(-1)^i \cdot \mu^i \cdot E[X^{r-i}]$$

Por ejemplo si $r=2$:

$$Var[X] = E[(X-\mu)^2] = \binom{2}{0}(-1)^0 \cdot \mu^0 \cdot E[X^2] + \binom{2}{1} \cdot (-1)^1 \cdot \mu^1 \cdot E[X]$$

$$+ \binom{2}{2} \cdot (-1)^2 \cdot \mu^2 \cdot E[X^0] = E[X^2] - 2\mu \cdot \mu + \mu^2 = E[X^2] - \mu^2 \Rightarrow$$

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

$$\begin{array}{ccccc} & & 1 & & \\ & 1 & & 1 & \\ 1 & & 2 & & 1 \\ \downarrow & & \downarrow & & \downarrow \\ \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} \end{array}$$

Como hemos podido observar con bastante frecuencia la serie o integral que permite calcular el momento de orden r suele desembocar en cálculos complicados, lo que puede simplificarse bastante mediante procedimientos alternativos. Uno de estos procedimientos consiste en definir una función tal que a partir de ella y mediante derivadas sucesivas se obtengan los momentos. A esta se le llama función generatriz de momentos.

Función generatriz de momentos

La función generatriz de momentos de una v.a X es el valor esperado de e^{tX}, se representa por $\phi_X(t)$, es decir que:

$$\phi_X(t) = E[e^{tX}]$$

- Si X es v.a. discreta

$$\phi_X(t) = E[e^{tX}] = \sum_i e^{tx_i} \cdot P[X = x_i]$$

• Si X es v.a. continua

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

Veamos ejemplos de estos conceptos.

Ejemplo (ejercicio 23)

Sea X una v.a. con función de probabilidad

$P[X=K] = p \cdot q^K \quad K = 0, 1, 2, \dots \quad p+q=1, \quad 0<p<1.$

Obtenga la función generatriz de momentos.

Al ser X una variable discreta se tiene que

$$\phi_x(t) = E[e^{tx}] = \sum_i e^{tx_i} \cdot P[X=x_i] = \sum_i e^{ti} \cdot P[X=i] =$$

$$= \sum_i e^{ti} \cdot p \cdot q^i = p \cdot \sum_i e^{ti} \cdot q^i = p\left[1 + e^t q + e^{2t} \cdot q^2 + e^{3t} q^3 + \dots\right]$$

La expresión $1 + e^t \cdot q + e^{2t} q^2 + e^{3t} \cdot q^3 + \dots$ corresponde a la suma de los infinitos términos de una progresión geométrica de razón $r = e^t q$. Para que esto se pueda sumar es necesario que $|e^t q| < 1 \Rightarrow e^t \cdot q < 1 \Rightarrow e^t < \frac{1}{q} \Rightarrow$

$\Rightarrow \ln e^t < \ln \frac{1}{q} \Rightarrow \boxed{t < -\ln q}$. En este caso la suma de los infinitos términos vale $\frac{a_1}{1-r}$ es decir:

$$\phi_x(t) = \frac{1}{1-q \cdot e^t} \cdot p \quad \text{si } t < -\ln q$$

Luego la función generatriz de momentos existe si se verifica que $t < -\ln q$ y vale:

$$\phi_x(t) = \frac{p}{1-q \cdot e^t}$$

Ejemplo (ejercicio 25)

Sea X una v. a. continua con función de densidad:

$$f(x) = \begin{cases} \frac{1}{2} e^{x} & \text{si } x \leq 0 \\ \frac{1}{2} e^{-x} & \text{si } x \geq 0 \end{cases}$$

Obtenga la función generatriz de momentos

Al ser una variable continua la función generatriz de momentos es:

$$\phi_X(t) = E[e^{tX}] = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx = \int_{-\infty}^{0} e^{tx} \cdot \frac{1}{2} e^{x} dx + \int_{0}^{+\infty} e^{tx} \cdot \frac{1}{2} e^{-x} dx$$

- $$\int e^{tx} \cdot \frac{1}{2} e^{x} dx = \frac{1}{2} \int e^{tx+x} dx = \frac{1}{2} \int e^{(1+t)x} dx \quad \left\{ \begin{array}{l} w = (1+t)x \\ dw = (1+t)\, dx \\ dx = \frac{1}{1+t} dw \end{array} \right\}$$

$$= \frac{1}{2} \int e^{w} \cdot \frac{1}{1+t} dw = \frac{1}{2(1+t)} e^{w} = \frac{1}{2(1+t)} e^{(1+t)x}$$

$$\int_{-\infty}^{0} e^{tx} \cdot \frac{1}{2} e^{x} dx = G(0) - G(-\infty) = \left[\frac{1}{2(1+t)} e^{(1+t)x} \right]_{-\infty}^{0} = \frac{1}{2(1+t)} \quad \text{si } \boxed{t+1>0}$$

- $$\int e^{tx} \cdot \frac{1}{2} \cdot e^{-x} dx = \frac{1}{2} \int e^{tx-x} dx = \frac{1}{2} \int e^{x(t-1)} dx \quad \left\{ \begin{array}{l} w = x(t-1) \\ dw = (t-1)\, dx \\ dx = \frac{1}{t-1} dw \end{array} \right\}$$

$$= \frac{1}{2} \int e^{w} \cdot \frac{1}{t-1} dw = \frac{1}{2(t-1)} e^{w} = \frac{1}{2(t-1)} e^{(t-1)x}$$

$$\int_{0}^{\infty} e^{tx} \cdot \frac{1}{2} \cdot e^{-x} dx = G(\infty) - G(0) = \left[\frac{1}{2(t-1)} e^{(t-1)x} \right]_{0}^{\infty} = \frac{1}{2(t-1)} \quad \boxed{\text{si } t-1<0}$$

Para que exista la función generatriz de momentos ha de verificarse:

$t+1>0 \Rightarrow t>-1$
$t-1<0 \Rightarrow t<1$
$\Rightarrow -1<t<1$. En este caso la función generatriz de momentos es $\phi_X(t) = \frac{1}{2(t+1)} + \frac{1}{2(t-1)}$

$$\phi_X(t) = \frac{1}{2}\left[\frac{1}{t+1} + \frac{1}{t-1}\right] \quad \text{si} \quad -1 < t < 1$$

$$\phi_X(t) = \frac{1}{2} \cdot \frac{t \cancel{-1} + t \cancel{+1}}{(t+1)(t-1)} \quad \text{si} \quad -1 < t < 1$$

$$\phi_X(t) = \frac{1}{t^2-1} \quad \text{si} \quad -1 < t < 1$$

Cálculo de momentos ordinarios (respecto al origen) a partir de la función generatriz de momentos.

Sea X una v.a. con función generatriz de momentos $\phi_X(t)$ $\forall t \in (-t_0, t_0)$. Entonces la derivada r-ésima de dicha función en el punto $t=0$ coincide con el momento r-ésimo centrado en el origen, es decir:

$$E[X^r] = \left.\frac{\partial^r \phi_X(t)}{dt^r}\right|_{t=0}$$

Veamos ejemplos de esto.

Ejemplo (ejercicio 23)

Sea X una v.a. con función de probabilidad $P[X=k] = p \cdot q^k$ $k = 0, 1, 2, \ldots$ $0 < p < 1$ $p + q = 1$. Obtenga la media y la varianza.

Anteriormente hemos obtenido la función generatriz de momentos para esta v.a X, que es:

$$\phi_X(t) = \frac{p}{1 - q \cdot e^t} \quad \text{si} \quad t < -\ln q$$

La media es el momento de orden 1

$$E[X] = E[X^1] = \left.\frac{\partial \phi_X(t)}{\partial t}\right|_{t=0} = \left.\frac{p \cdot q \cdot e^t}{(1 - q \cdot e^t)^2}\right|_{t=0} = \frac{p \cdot q}{(1-q)^2} =$$

$$E[X] = \frac{p \cdot q}{p^2} = \frac{q}{p}$$

Sabemos que la varianza es $Var[X] = E[(X-\mu)^2] = E[X^2] - E[X]^2$

Necesitamos calcular $E[X^2]$.

Se tiene que:

$$E[x^2] = \frac{\partial^2 \phi_x(t)}{\partial t^2}\Big|_{t=0} = \frac{\partial}{\partial t}\left(\frac{\partial \phi_x(t)}{\partial t}\right)\Big|_{t=0} =$$

$$E[x^2] = \frac{\partial}{\partial t}\left[\frac{pqe^t}{(1-qe^t)^2}\right]\Big|_{t=0} = \frac{pqe^t(1-qe^t)^2 + 2pq^2e^{2t}(1-qe^t)}{(1-qe^t)^4}\Big|_{t=0} =$$

$$\Rightarrow E[x^2] = \frac{p.q.e^t(1-qe^t) + 2pq^2e^{2t}}{(1-qe^t)^3}\Big|_{t=0} = p.q.e^t\,\frac{1-qe^t+2qe^t}{(1-qe^t)^3}\Big|_{t=0}$$

$$E[x^2] = p.q.e^t\,\frac{1+q.e^t}{(1-qe^t)^3}\Big|_{t=0} \Rightarrow E[x^2] = \frac{pq(1+q)}{(1-q)^3}$$

$$E[x^2] = \frac{pq(1+q)}{p^3} \Rightarrow E[x^2] = \frac{q(1+q)}{p^2}$$

Luego $Var[x] = \frac{q(1+q)}{p^2} - \left(\frac{q}{p}\right)^2 = \frac{q+q^2-q^2}{p^2} \Rightarrow Var[x] = \frac{q}{p^2}$

Ejemplo (ejercicio 25)

Sea X la v. a. con función de densidad

$$f(x) = \begin{cases} \frac{1}{2}e^{x} & \text{si } x \leq 0 \\ \frac{1}{2}e^{-x} & \text{si } x \geq 0 \end{cases}$$

Obtenga la media y la varianza.

Anteriormente hemos obtenido la función generatriz de momentos $\phi_x(t)$ que valía

$$\phi_x(t) = \frac{1}{1-t^2} \quad \text{si} \quad -1 < t < 1.$$

$$E[x^1] = \frac{\partial \phi_x(t)}{\partial t}\Big|_{t=0} = \frac{2t}{(1-t^2)^2}\Big|_{t=0} = 0$$

Luego $\mu = E[x] = 0$.

Sabemos que $Var[X] = E[(X-\mu)^2] = E[X^2] - \overset{0}{\overset{\shortparallel}{E[X]}}^2 = E[X^2]$

Como:

$$E[X^2] = \frac{\partial^2 \phi_X(t)}{\partial t^2}\bigg|_{t=0} = \frac{\partial}{\partial t}\left[\frac{\partial \phi_X(t)}{\partial t}\right]\bigg|_{t=0} = \frac{\partial}{\partial t}\left[\frac{2t}{(1-t^2)^2}\right]\bigg|_{t=0}$$

$$E[X^2] = \frac{2(1-t^2)^2 - 2t\cdot 2\cdot(1-t^2)(-2t)}{(1-t^2)^4}\bigg|_{t=0} = \frac{2(1-t^2)+8t^2}{(1-t^2)^4}\bigg|_{t=0} =$$

$$E[X^2] = \frac{6t^2+2}{(1-t^2)^4}\bigg|_{t=0} \Rightarrow E[X^2] = 2$$

Luego $Var[X] = 2$.

Función generatriz de momentos de una transformación lineal

Sea X una v.a. con función generatriz de momentos $\phi_X(t) = E[e^{Xt}]$. Sea Y la transformación lineal $Y = a+bX$, $a, b \in \mathbb{R}$. Entonces la función generatriz de momentos de Y es:

$\phi_Y(t) = E[e^{tY}] = E[e^{at+btX}] = E[e^{at}\cdot e^{tbX}]$ en esta expresión todo lo que no es X es constante. Luego:

$$\phi_Y(t) = e^{at}\cdot E[e^{tbX}] \Rightarrow \phi_Y(t) = e^{at}\cdot \phi_X(tb)$$

Veamos un ejemplo

Ejemplo (ejercicio 27)

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

$$f(x) = \begin{cases} e^{-x} & \text{si } x \geq 0 \\ 0 & \text{en el resto} \end{cases}$$

con función generatriz de momentos $\phi_X(t) = \frac{1}{1-t}$ si $t < 1$, y media $\mu = 1$.

Demuéstrese que la función generatriz de momentos de la v.a. $X - E[X]$ viene dada por $\frac{e^{-t}}{1-t}$ si $t < 1$

Sea $Y = X - E[X] = X - \mu = X - 1$. Entonces

$$\phi_Y(t) = E[e^{tY}] = E[e^{tx-t}] = E[e^{-t} \cdot e^{tx}] = e^{-t} \cdot E[e^{tx}]$$

$$\phi_Y(t) = e^{-t} \cdot \phi_X(t) = \frac{e^{-t}}{1-t} \quad \text{si } t < 1.$$

Unicidad de la función generatriz de momentos (muy importante)

Sean X e Y dos v.a. con funciones generatrices de momentos $\phi_X(t)$ y $\phi_Y(t)$ respectivamente. Entonces:

X e Y tienen la misma distribución de probabilidad $\Leftrightarrow \phi_X(t) = \phi_Y(t)$

Teorema de Markov

Sea X una v.a, $g[X]$ una función de la misma tal que $g(X) \geq 0$. Sea K una constante positiva. Entonces se verifica que:

$$P[g(x) \geq K] \leq \frac{E[g(x)]}{K}$$

Demostración:

$$E[g(x)] = \sum_i g(x_i) \cdot P[X = x_i] = \sum_{g(x) \geq K} g(x_i) \cdot P[X = x_i] + \sum_{g(x_i) < K} g(x_i) \cdot P[X = x_i] \geq \sum_{g(x_i) \geq K} g(x_i) \cdot P[X = x_i] \geq \sum_{g(x_i) \geq K} K \cdot P[X = x_i]$$

Luego $E[g(x)] \geq K \cdot \sum_{g(x_i) \geq K} P[X = x_i] = K \cdot P[g(x) \geq K] \Rightarrow$

$$\Rightarrow K \cdot P[g(x) \geq K] \leq E[g(x)] \Rightarrow P[g(x) \geq K] \leq \frac{E[g(x)]}{K}$$

El teorema de Markov permite calcular una cota superior de la probabilidad del suceso de que una variable $g(x) \geq 0$ tome un valor no inferior a una cantidad $K > 0$ fijada de antemano. Para conocer esa cota necesitamos saber $E[g(x)]$.

Teorema de Chebychev

Sea X una v.a. con media $\mu = E[X]$ y varianza $\sigma^2 = E[(X-\mu)^2]$. Entonces si $K>0$ es constante:

$$P[(x-\mu)^2 > K] \leq \frac{\sigma^2}{K} \quad \text{o bien:}$$

$$P[|x-\mu| > \sqrt{K}] \leq \frac{\sigma^2}{K}$$

Demostración:

El teorema de Chebychev es un caso particular del teorema de Markov, basta con tomar $g(X) = (x-\mu)^2 \Rightarrow E[g(x)] = E[(x-\mu)^2] = \sigma^2$. Si sustituimos estos valores en el teorema de Markov:

$$P[(x-\mu)^2 \geq K] \leq \frac{E[(x-\mu)^2]}{K} = \frac{\sigma^2}{K}$$ y esto es equivalente a poner $P[|x-\mu| \geq \sqrt{K}] \leq \frac{\sigma^2}{K}$

Observaciones sobre el Teorema de Tchebychev

Si en vez del valor $\sqrt{K}$ tomamos $K\sigma$ entonces:

$P[|x-\mu| \geq K\sigma] \leq \frac{\sigma^2}{K^2\sigma^2} \Rightarrow P[|x-\mu| \geq K\sigma] \leq \frac{1}{K^2}$. La probabilidad del suceso contrario será:

$$P[|x-\mu| < K\sigma] = 1 - P[|x-\mu| \geq K\sigma] \geq 1 - \frac{1}{K^2}$$

Como $|x-\mu| < K\sigma \Leftrightarrow \mu - K\sigma < x < \mu + K\sigma$ podemos poner:

$$P[\mu - K\sigma < X < \mu + K\sigma] \geq 1 - \frac{1}{K^2}$$

- La probabilidad de que la variable X tome valores dentro del intervalo $(\mu - K\sigma, \mu + K\sigma)$ es mayor o igual que $1 - \frac{1}{K^2}$
- La probabilidad de que la variable X tome valores fuera del intervalo es $\frac{1}{K^2}$

Conforme K crece la probabilidad de que la v.a. X tome valores fuera del intervalo $(\mu - K\sigma, \mu + K\sigma)$ es cada vez menor la utilidad del teorema de Tchebychev se centra en aquellos casos en los que no conocemos la distribución de probabilidad pero suponemos conocidas la media μ y la desviación típica σ. Con esto podemos calcular cotas sobre la probabilidad.

Una última cuestión antes de ver algunos ejemplos. Hemos demostrado el teorema de Markov para variable discreta, para variable continua la demostración sería así:

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) f(x)\,dx = \int_{g(x)\geq K} g(x)\cdot f(x)\,dx + \int_{g(x)<K} g(x)\cdot f(x)\,dx \Rightarrow$$

$$\Rightarrow E[g(x)] \geq \int_{g(x)\geq K} g(x)\cdot f(x)\,dx \geq \int_{g(x)\geq K} K\cdot f(x)\,dx = K \int_{g(x)\geq K} f(x)\,dx = K\cdot P[g(x)\geq K]$$

Luego $K\cdot P[g(x)\geq K] \leq E[g(x)] \Rightarrow P[g(x)\geq K] \leq \dfrac{E[g(x)]}{K}$

Vamos a los ejemplos

Ejemplo (ejercicio 21)

Se sabe que $E[X] = 10$ y $V(X) = 25$. Determine los intervalos alrededor de la media de esta v.a. que contengan al menos el 50% y el 90% de la probabilidad.

Según la observación hecha en torno al teorema de Tchebychev:

$$P[\mu - K\sigma < X < \mu + K\sigma] \geq 1 - \frac{1}{K^2}$$

Que contengan al menos el 50%: $1 - \dfrac{1}{K^2} = \dfrac{1}{2} \Rightarrow \dfrac{1}{K^2} = \dfrac{1}{2}$

$K^2 = 2 \Rightarrow K = \sqrt{2} \Rightarrow K = \sqrt{2} > 0$ (K ha de ser positivo)

El intervalo es $(\mu - K\sigma, \mu + K\sigma) = (10 - \sqrt{2}\cdot\sqrt{25}, 10 + \sqrt{2}\cdot\sqrt{25}) =$

$= (10 - 5\sqrt{2}, 10 + 5\sqrt{2}) = (2'929, 17'071)$

- Que contengan el 90%: $1 - \frac{1}{k^2} = \frac{}{} \Rightarrow \frac{1}{k^2} = \frac{1}{10}$

$k = \sqrt{10}$

El intervalo es $(\mu - k\sigma, \mu + k\sigma) = (10 - \sqrt{10} \cdot \sqrt{25}, 10 + \sqrt{10} \cdot \sqrt{25})$

$= (-5'811, 25'811)$

Ejemplo (ejercicio 22)

El tiempo necesario para transportar un cargamento por vía marítima entre dos puertos sigue una distribución de probabilidad de media 90h. y desviación típica 3h. El capitán del barco pretende llegar al otro puerto entre 80 y 100 horas después de haber abandonado el primero. Obtenga un límite inferior para la probabilidad de que se cumpla lo que dice el capitán.

Según el teorema de Tchebychev:

$P[\,|x - \mu| < k\sigma] \geq 1 - \frac{1}{k^2}$ y esto es equivalente a poner:

$P[\mu - k\sigma < x < \mu + k\sigma] \geq 1 - \frac{1}{k^2}$

$\mu - k\sigma = 80$ $\Rightarrow$ $90 - 3k = 80 \Rightarrow k = \frac{10}{3}$

$\mu + k\sigma = 100$ $\quad$ $90 + 3k = 100 \Rightarrow k = \frac{10}{3}$

La cota inferior de la probabilidad es $1 - \frac{1}{k^2} = 1 - \frac{1}{\frac{100}{9}} =$

$= 1 - \frac{9}{100} = \frac{91}{100} = 0'91$

Lo que dice el capitán se cumplirá con una probabilidad al menos de un 91%.