VALOR ESPERADO Y VARIANZA DE UNA VARIABLE ALEATORIA

DEF: (Esperanza) Sea X una variable aleatoria definida sobre el espacio de probabilidad $(\Omega, \mathcal{F}, P)$. La esperanza (valor esperado, media o promedio) de X es definida como

$$E(X) = \int_{\mathbb{X}} x \, dF(x)$$

que adopta la forma

$$E(X) = \sum_{x \in \mathbb{X}} x \, P(X = x) = \sum_{x \in \mathbb{X}} x \, p_X(x)$$

para X v.a. discreta, o bien

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

para X v.a continua. (para $\mathbb{X} = X(\Omega)$)

siempre que la integral (caso continuo) o la suma (caso discreto) exista, esto es siempre que

$$\sum_{x \in \mathbb{R}} |x| \cdot P(X=x) < \infty,$$

o bien

$$\int_{-\infty}^{\infty} |x| f_X(x) < \infty.$$

OBS: Si $E(|X|)=\infty$ entonces decimos que $E(X)$ no existe.

Además, si X es una variable aleatoria que sólo toma valores en un número finito de valores, entonces $E(X)$ siempre existe

DEF: Si $E(X)$ es finita, decimos que X es integrable

EJEMPLO Suponga que X es v.a. con densidad

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

Entonces

$$E(X) = \int_0^{\infty} x f_X(x)\,dx = \int_0^{\infty} x \lambda e^{-\lambda x}\,dx.$$

- Tomando,

$$u = x, \quad dv = \lambda e^{-\lambda x}\,dx, \quad \text{así}$$
$$du = dx, \quad v = -e^{-\lambda x}$$

- integrando por partes, es decir

- $$\int_0^{\infty} x \lambda e^{-\lambda x}\,dx = \int_0^{\infty} u\,dv = uv\Big|_0^{\infty} - \int_0^{\infty} v\,du$$

$$= -xe^{-\lambda x}\Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x}\,dx$$

$$= \int_0^{\infty} e^{-\lambda x}\,dx = -\frac{1}{\lambda} e^{-\lambda x}\Big|_0^{\infty}$$

$$= \frac{1}{\lambda} \qquad \left(\int e^{ax}\,dx = \frac{1}{a} e^{ax} \right).$$

Ejemplo Considere

$$f_X(x) = \frac{1}{b-a} I_{[a,b]}(x)$$

Entonces

$$E(X) = \int_a^b \frac{1}{b-a} x \, dx = \frac{1}{b-a} \int_a^b x \, dx$$

$$= \frac{1}{b-a} \left. \frac{x^2}{2} \right|_a^b = \frac{1}{b-a} \left(\frac{b^2}{2} - \frac{a^2}{2} \right).$$

$$= \frac{1}{b-a} \left(\frac{b^2-a^2}{2} \right) = \frac{(b-a)(a+b)}{2(b-a)}.$$

$$= \frac{a+b}{2}.$$

EJEMPLO Suponga X con función de probabilidad

$$P(X=x) = \frac{e^{-\mu}\mu^{x}}{x!}, \quad \mu > 0$$

y $x \in \{0,1,2,\dots\}$. Luego.

- $$E(X) = \sum_{x=0}^{\infty} x\,P(X=x) = \sum_{x=0}^{\infty} x\,\frac{e^{-\mu}\mu^{x}}{x!}$$

$$= e^{-\mu}\sum_{x=0}^{\infty}\frac{x\,\mu^{x}}{x!} = e^{-\mu}\sum_{x=1}^{\infty}\frac{x\,\mu^{x}}{x(x-1)!}$$

$$= e^{-\mu}\sum_{x=1}^{\infty}\frac{\mu\cdot\mu^{x-1}}{(x-1)!} = \mu\, e^{-\mu}\sum_{x=1}^{\infty}\frac{\mu^{x-1}}{(x-1)!}$$

- Sea $z = x-1$, de ahí que

$$E(X) = \mu\, e^{-\mu}\cdot\sum_{z=0}^{\infty}\frac{\mu^{z}}{z!} = \mu\, e^{-\mu} e^{\mu}$$

$$= \mu.$$

PROPIEDAD: Sea Y una variable aleatoria continua con densidad f. Si $E(Y)$ existe, entonces

$$E(Y) = \int_0^{\infty} \{1 - F_Y(y)\} dy - \int_0^{\infty} F_Y(-y) dy$$

RESULTADO: Sea X una variable aleatoria con densidad f y considere $g: \mathbb{R} \to \mathbb{R}$ una función tal que $Y = g(X)$ es una variable aleatoria. Entonces

$$E(g(X)) = \begin{cases} \sum_x g(x) f(x), & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{\infty} g(x) f(x) dx, & \text{si } X \text{ es continua} \end{cases}$$

siempre que la suma o la integral converja dem: (a) Suponga que X es discreta con valores $x_1, x_2, \ldots$ En este caso.

$$p(x) = \begin{cases} P(X=x), & \text{si } x = x_1, x_2, \ldots \\ 0, & \text{en otro caso.} \end{cases}$$

La variable aleatoria $Y = g(X)$ toma valores $g(x_1), g(x_2), \ldots$ (además algunos de estos valores pueden ser iguales). Suponga que y_j con $j \geq 1$ representa los diferentes valores de $g(x_i)$. Luego, agrupando todos los $g(x_i)$ con el mismo valor sigue que

$$\sum_i g(x_i)\, p(x_i) = \sum_j \sum_{i : g(x_i) = y_j} g(x_i)\, p(x_i)$$

$$= \sum_j y_j \sum_{i : g(x_i) = y_j} p(x_i) = \sum_j y_j\, P(g(X) = y_j)$$

$$= \sum_j y_j\, P(Y = y_j) = E(Y)$$

$$= E(g(X)).$$

(b) Suponga que g es función no negativa

Entonces

$$E(g(x)) = \int_0^\infty P(g(x) > y)\,dy - \int_0^\infty P(g(x) \leq -y)\,dy$$

$$= \int_0^\infty P(g(x) > y)\,dy$$

$$= \int_0^\infty \left(\int_B f(x)\,dx \right) dy$$

donde $B = \{x : g(x) > y\}$. De este modo

$$E(g(x)) = \int_0^\infty \int_0^{g(x)} f(x)\,dy\,dx$$

$$= \int_0^\infty g(x) f(x)\,dx.$$

Escaneado con CamScanner

Ejemplo: Sea $A \subseteq \Omega$ y considere

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

Note que

- $$E(I_A(x)) = \int_\Omega I_A(x) f_x(x)\,dx$$

$$= \int_{\{x \in A\}} f_x(x)\,dx$$

$$= P(x \in A).$$

- Obs: Note que $E(1/x) \neq 1/E(x)$. En efecto,

$$E\left(\frac{1}{x}\right) = \int_\Omega \frac{1}{x} f(x)\,dx \neq \frac{1}{\int_\Omega x f(x)\,dx} = \frac{1}{E(x)}.$$

Ejemplo: Sea

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

Entonces

$$E(x^3) = \int_{-\infty}^{\infty} x^3 f(x)\,dx = \int_0^1 2x^4\,dx$$

$$= 2\int_0^1 x^4\,dx = 2\,\frac{x^5}{5}\Big|_0^1$$

$$= \frac{2}{5}.$$

RESULTADO. Sea X una v.a. real

(a) Si $P(X \geq 0) = 1$ y $E(X)$ existe entonces $E(X) \geq 0$.

(b) $E(\alpha) = \alpha$ para α una constante.

(c) Si X es acotada. Es decir, existe una constante $M > 0$ tal que $P(|X| \leq M) = 1$ entonces $E(X)$ existe.

dem: para notar (a) considere el caso continuo

$$E(X) = \int_0^\infty [1-F(x)]\,dx - \int_0^\infty F(-x)\,dx \quad \Big\{ P(X < -x)$$

$$= \int_0^\infty [1-F(x)]\,dx = \int_0^\infty P(X > x)\,dx$$

$$\geq 0$$

- Ahora, (b) sigue desde

$$E(\alpha) = \alpha P(X=\alpha) = \alpha$$

o bien

$$E(\alpha) = \int_{-\infty}^\infty \alpha f(x)\,dx = \alpha \int_{-\infty}^\infty f(x)\,dx$$

- $$= \alpha$$

(c) Suponga X una v.a discreta tomando valores $x_1, x_2, \dots$ entonces como $P(|X| > M) = 0$ entonces podemos considerar $\{x_1, x_2, \dots\} \subseteq [-M, M]$

Escaneado con CamScanner

(d) Sean α y β constantes y $g(x)$, $h(x)$ funciones de variables aleatorias. Entonces el valor esperado

$$E\{\alpha g(x) + \beta h(x)\} = \alpha E(g(x)) + \beta E(h(x)).$$

(es decir la esperanza es lineal).

(e) Si g y h son funciones tales que $g(x) \leq h(x)$. $\forall x$ y sus esperanzas existen. Entonces

$$E(g(x)) \leq E(h(x)).$$

En particular

$$|E(x)| \leq E(|x|).$$

Escaneado con CamScanner

De ahí que

$$\sum_i |x_i| P(X=x_i) \leq \sum_i M\, P(X=x_i)$$

$$= M \sum_i P(X=x_i) = M < \infty$$

Para X continua y como $P(|X|>M)=0$ de ahí que $f(x)=0 \quad \forall\, x \notin [-M,M]$. Luego

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

$$\leq \int_{-M}^{M} M f(x)\,dx = M \int_{-M}^{M} f(x)\,dx$$

$$= M\, P(|X| \leq M) = M < \infty$$

(d) Solo consideraremos el caso continuo

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

$$= \int_{-\infty}^{\infty} \alpha g(x) f(x)\,dx + \int_{-\infty}^{\infty} \beta h(x) f(x)\,dx$$

$$= \alpha \int_{-\infty}^{\infty} g(x) f(x)\,dx + \beta \int_{-\infty}^{\infty} h(x) f(x)\,dx$$

(e) Note que

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

$$= E(h(x))$$

Además, como

$$-|x| \leq x \leq |x|$$

sigue que

$$E(-|x|) \leq E(x) \leq E(|x|)$$

$$\Rightarrow \quad -E(|x|) \leq E(x) \leq E(|x|)$$

y de ahí que

$$|E(x)| \leq E(|x|).$$

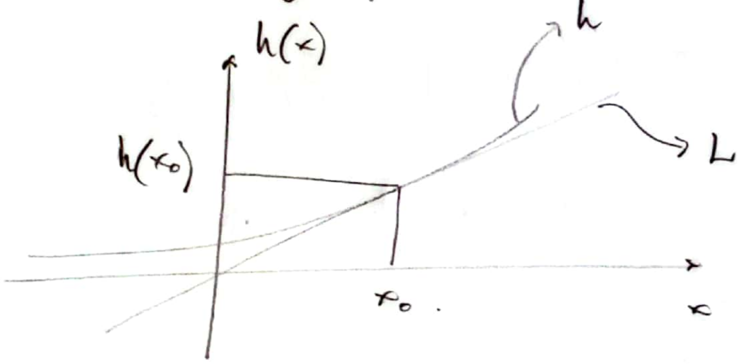
RESULTADO (Desigualdad de Jensen)

Sea h una funcion convexa definida en $\mathbb{R}$.
Si X es integrable, entonces

$$E(h(X)) \geq h(E(X))$$

$$\left\{ \begin{array}{l} h(tx + (1-t)y) \leq t h(x) + (1-t) h(y) \\ \text{con } x, y \in \mathbb{R} \quad y \quad t \in [0,1] \end{array} \right.$$

dem. Considere el gráfico



Dado x_0 y el punto $h(x_0)$ del gráfico de h por la convexidad existe una recta L que pasa por $h(x_0)$ y deja la curva h arriba de ella.

Sea $\lambda \in \mathbb{R}$

$$y - h(x_0) = \lambda (x - x_0)$$

es la ecuación de la recta L. Así

$$h(x) \geq L(x) = h(x_0) + \lambda (x - x_0).$$

De este modo,

$$E(h(x)) \geq E(L(x))$$

$$= E\{h(x_0) + \lambda (x - x_0)\}$$

$$= h(x_0) + \lambda (E(x) - x_0).$$

tomando $x_0 = E(x)$ * sigue que

$$E(h(x)) \geq h(E(x)) + \lambda (E(x) - E(x)).$$

(* pues $L(E(x)) = h(E(x))$).

OBS: Si h es concava entonces

$$E\{h(x)\} \leq h(E(x))$$

PROPIEDAD. Considere que deseamos minimizar.

$$\min_{b} E\{(x-b)^2\}$$

cuya solución es $\hat{b} = E(x)$. Para notar lo anterior, escriba

$$E\{(x-b)^2\} = E\{(x - E(x) + E(x) - b)^2\}$$

$$= E[(x-E(x))^2 + 2(x-E(x))(E(x)-b)$$

$$+ (E(x)-b)^2]$$

Ahora

$$E[(x-E(x))(E(x)-b)]$$

$$= (E(x)-b) \cdot E[(x-E(x))]$$

$$= (E(x)-b)[E(x)-E(x)].$$

$$= 0$$

De este modo

$$E[(X-b)^2] = E[(X-E(X))^2] + (E(X)-b)^2.$$

el primer término no depende de b mientras que el 2° siempre es mayor o igual a cero. Además este es cero (su valor mínimo) si $b = E(X)$. De ahí que

$$\min_b E[(X-b)^2] = E[(X-E(X))^2]$$

(Ejercicio : Resolver usando cálculo).

Escaneado con CamScanner

Ejemplo:

(a) Sea $h(x) = |x|$. Entonces

$$E(|x|) \geq |E(x)|$$

(b) Sea $h(x) = x^2$. Entonces

$$E(x^2) \geq (E(x))^2$$

(c) Sea $h(x) = |x|^p$, $p \geq 1$. Entonces

$$E(|x|^p) \geq |E(x)|^p$$

Obs: Para la validad de la desigualdad de Jensen basta que la función h sea convexa en un intervalo (a,b) tal que $P(a < x < b) = 1$

Por ejemplo si $P(x > 0) = 1$. entonces podemos considerar $h(x) = 1/x$ con $(a,b) = (0,\infty)$. En este caso

$$E\left(\frac{1}{x}\right) \geq \frac{1}{E(x)}$$

Bajo la misma condición $P(x>0) = 1$ podemos usar $h(x) = \log x$ (concava) y $E(\log x) \leq \log(E(x))$.

DEF. Sea X una variable aleatoria, el r-ésimo momento de X es definido como.

$$\mu_r' = E(X^r)$$

siempre que la esperanza exista.

OBS: Note que

$$\mu_1' = E(X) = \mu$$

es la media de X

DEF Sea X una variable aleatoria con esperanza finita. Entonces el r-ésimo momento central (en torno de $\mu = E(X)$) es dado por

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

$$= E[(X - E(X))^r].$$

siempre que la esperanza exista.

Escaneado con CamScanner

Obs: Tambien podemos definir el r-ésimo momento en torno de a, como.

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

Evidentemente $\mu_r(a) = \mu_r$ para $a = \mu_1$.

DEF: la varianza de una variable aleatoria X es su 2° momento central, y anotamos

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

Obs: La raíz cuadrada positiva de $Var(X)$ es la desviación estándar de X

$$Sd(X) = \sqrt{Var(X)}.$$

RESULTADO Sea X una v.a cuyo valor esperado existe y $\alpha, \beta \in \mathbb{R}$ son constantes. Entonces

(a) $Var(X) \geq 0$.

(b) $Var(\alpha) = 0$

(c) $Var(\alpha X) = \alpha^2 Var(X)$

(d) $Var(X + \beta) = Var(X)$

(e) $Var(X) = 0$ si y solo si $P(X = E(X)) = 1$

dem: (a) es evidente desde la definición.
Para notar (b)

$$
\begin{aligned}
Var(\alpha) &= E[(\alpha - E(\alpha))^2] = E[(\alpha - \alpha)^2] \\
&= E(0) = 0
\end{aligned}
$$

(c) Note que

$$
\begin{aligned}
\alpha X - E(\alpha X) &= \alpha X - \alpha E(X) \\
&= \alpha (X - E(X))
\end{aligned}
$$

Escaneado con CamScanner

De ahí que

$$
\begin{aligned}
\operatorname{Var}(\alpha X) &= E[(\alpha X - E(\alpha X))^2] \\
&= E[\alpha^2 (X - E(X))^2] \\
&= \alpha^2 E[(X - E(X))^2] = \alpha^2 \operatorname{Var}(X).
\end{aligned}
$$

(d) Sigue de notar que

$$
\begin{aligned}
\operatorname{Var}(X+\beta) &= E[(X+\beta - E(X+\beta))^2] \\
&= E[(X+\beta - E(X) - \beta)^2] \\
&= E[(X - E(X))^2] = \operatorname{Var}(X).
\end{aligned}
$$

(e) Primero, considere $X = E(X)$ con probabilidad 1. Entonces es claro que $\operatorname{Var}(X) = 0$

Ahora, suponga que $\operatorname{Var}(X) = 0$ y sea $a := E(X)$. Si $P(X=a) < 1$, entonces existe $c > 0$ tal que

$$P((X-a)^2 > c) > 0.$$

Ahora, como (por qué?)

$$(x-a)^2 \geq c\, I_{\{(x-a)^2 > c\}}$$

Entonces

$$E[(x-a)^2] \geq E[c\, I_{\{(x-a)^2>c\}}]$$

$$\Rightarrow \quad \operatorname{var}(x) \geq c\, E[I_{\{(x-a)^2>c\}}]$$

$$\Rightarrow \quad \operatorname{var}(x) \geq c\, P((x-a)^2 > c) > 0.$$

lo cual es una contradicción. De ahí que $P(x = E(x)) = 1$.

RESULTADO Sea x una v.a. tal que $E(x^2)$ existe. Entonces

$$\operatorname{var}(x) = E(x^2) - (E(x))^2$$

Escaneado con CamScanner

dem:

$$\begin{aligned} Var(X) &= E[(X - E(X))^2] \\ &= E[X^2 - 2X E(X) + (E(X))^2] \\ &= E(X^2) - 2E(X)E(X) + E[(E(X))^2] \\ &= E(X^2) - 2(E(X))^2 + (E(X))^2 \\ &= E(X^2) - (E(X))^2 \end{aligned}$$

DEF Para una variable aleatoria X con 2° momento finito, sea $\mu = E(X)$ y $\sigma^2 = Var(X)$ El 3r y 4° momentos centrales estandarizados son definidos como

$$\gamma = E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right], \quad \kappa = E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right].$$

y se denominan sesgo y curtosis. Además $\kappa - 3$ es conocido como exceso de curtosis.

OBS los coeficientes de sesgo y curtosis arrojan información sobre la simetría y grado de agudeza (o tamaño de la colas) de la distribución de X.

DEF El cuantil α-ésimo de una variable aleatoria X es el valor más pequeño $q \in \mathbb{R}$ satisfaciendo $F_X(q) \geq \alpha$ donde α es alguna constante $\alpha \in (0,1)$. En ocasiones anotaremos q_α para resaltar la conexión con α.

Para X variable aleatoria continua, entonces $F_X(q) = \alpha$. Si además F es estrictamente creciente entonces podemos calcular el cuantil usando la función inversa

$$q = F_X^{-1}(\alpha)$$

(**OBS:** En general, para F no decreciente, la inversa (generalizada) de F, denotada por F^{-} está definida por

$$F_X^{-}(u) = \inf\{x : F(x) \geq u\}.$$
)

● **DEF:** Sea X una variable aleatoria continua con densidad $f(x)$. La mediana de X es un punto $c \in \mathbb{R}$ tal que

$$\int_{-\infty}^{c} f(x)\,dx = \int_{c}^{\infty} f(x)\,dx \quad \left(= \frac{1}{2}\right).$$

● **RESULTADO** La mediana de una v.a. X es el punto c tal que

$$F_X(c) = \frac{1}{2}.$$

dem: Dado que

$$F_X(x) = \int_{-\infty}^{x} f_X(u)\, du$$

Tenemos que

$$F_X(c) = \int_{-\infty}^{c} f_X(x)\, dx$$

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

es decir

$$F_X(c) = 1 - F_X(c)$$

$$\Rightarrow \quad 2F_X(c) = 1 \quad \Rightarrow \quad F_X(c) = \frac{1}{2}.$$

obs: En ocasiones anotamos

$$x_{0.5} = \text{med}(X).$$

Ejemplo: Sea X una variable aleatoria con

$$f_X(x) = \frac{1}{b-a}, \quad a \leq x \leq b.$$

De ahí que

$$F_X(x) = \frac{x-a}{b-a}, \quad a \leq x \leq b.$$

Así, la mediana de X es el número $c \in \mathbb{R}$ tal que

$$F_X(c) = \frac{1}{2}.$$

es decir

$$F_X(c) = \frac{c-a}{b-a} = \frac{1}{2}$$

De ahí que

$$c-a = \frac{b-a}{2} \rightarrow c = \frac{b-a}{2} + a$$

$$\Rightarrow \quad c = \frac{a+b}{2}.$$

Ejemplo: Sea X v.a con $f_X(x) = \lambda e^{-\lambda x}$, para $x \geq 0$. Entonces

$$F_X(x) = 1 - e^{-\lambda x}, \quad x \geq 0$$

La mediana de X es el punto c tal que

$$F_X(c) = \frac{1}{2} \Rightarrow 1 - e^{-\lambda c} = \frac{1}{2}.$$

$$\Rightarrow e^{-\lambda c} = \frac{1}{2} \Rightarrow -\lambda c = \log(1/2)$$

$$\Rightarrow -\lambda c = -\log 2 \Rightarrow c = \frac{\log 2}{\lambda}.$$

DEF: Sea X una v.a continua. La moda es el punto c tal que $f_X(x)$ alcanza su máximo

$$c = \arg\max_x f_X(x)$$

$$= \arg\max_x \frac{d}{dx} F_X(x).$$

Ejemplo Sea X v.a. con densidad

$$f_X(x) = 6x(1-x), \quad 0 \leq x \leq 1.$$

para obtener el máximo, tomamos la derivada de f_X. Note que

- $$0 = \frac{d}{dx} f_X(x) = \frac{d}{dx} 6x(1-x).$$
  $$= 6(1-2x).$$

resolviendo, obtenemos

$$6(1-2x) = 0 \Rightarrow x = \frac{1}{2}.$$

- Además

$$\frac{d^2}{dx^2} f_X(x) = \frac{d}{dx} 6(1-2x) = -12 < 0.$$

luego $x = \frac{1}{2}$ es máximo global (único).