4 Esperança matemàtica. Variància.

4.1 Esperança matemàtica. Propietats.

L'esperança matemàtica és una generalització del concepte de mitjana aritmètica i pot interpretar-se com:

1. Valor mig teòric de tots els valors que pot prendre la variable. Representa una mesura de centralització.
2. Centre de gravetat dels punts que corresponen als valors de la variable, assignant-los una quantitat de massa proporcional a la funció de probabilitat en cada punt.
3. Si la variable aleatòria és el guany o pèrdua en un determinat joc d'atzar, l'esperança representa el guany mig per jugada. Un joc és equitatiu o just si la seva esperança és zero.

L'**esperança matemàtica** d'una variable aleatòria discreta X amb valors sobre S i funció de probabilitat f es denota per $E(X)$ i es defineix com

$$E(X) = \sum_{x \in S} x\, f(x),$$

sempre que aquesta sèrie sigui absolutament convergent. Si S és un conjunt finit, $S = \{x_1, \ldots, x_n\}$, aleshores,

$$E(X) = \sum_{i=1}^{n} x_i\, f(x_i),$$

on $f(x_i) = P(X = x_i)$, $i = 1, \ldots, n$.

L'**esperança matemàtica** d'una variable aleatòria contínua X amb funció de densitat $f(x)$, $x \in \mathbb{R}$ és:

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

sempre que aquesta integral sigui absolutament convergent.

Propietats de l'esperança.

1. si $X = c \in \mathbb{R}$, $E(X) = c$,
2. linealitat de l'esperança: $a, b \in \mathbb{R}$, $E(a\,X + b) = a\; E(X) + b$,
3. si $h : \mathbb{R} \to \mathbb{R}$ és una funció,

$$E(h(X)) = \begin{cases} \sum_{x \in S} h(x)\, f(x), & \text{cas discret,} \\ \int_{-\infty}^{+\infty} h(x)\, f(x)\, dx, & \text{cas continu.} \end{cases}$$

Demostració: Demostrem aquestes propietats en el cas discret. Considerem X una v.a. que pren valors sobre el conjunt $S = \{x_1, x_2, \ldots, x_n\}$ finit, amb funció de probabilitat $f(x)$.

1. $X = c \in \mathbb{R}$,

$$E(X) = \sum_{i=1}^{n} x_i\, f(x_i) = \sum_{i=1}^{n} c\, f(x_i) = c \sum_{i=1}^{n} f(x_i) = c.$$

3. Considerem el cas particular que h és una transformació injectiva amb $h(x_i) = y_i$, per $i = 1, 2, \ldots, n$. aleshores:

$$E(h(X)) = E(Y) = \sum_{i=1}^{n} y_i\, P(Y = y_i) = \sum_{i=1}^{n} h(x_i)\, P(X = x_i) = \sum_{i=1}^{n} h(x_i)\, f(x_i).$$

2. Linealitat de l'esperança: és un cas particular de l'anterior utilitzant la funció $h(x_i) = a\, x_i + b$, amb $a, b \in \mathbb{R}$,

$$E(a\, X + b) = \sum_{i=1}^{n} (a\, x_i + b)\, f(x_i) = \sum_{i=1}^{n} a\, x_i\, f(x_i) + \sum_{i=1}^{n} b\, f(x_i)$$

$$= a \sum_{i=1}^{n} x_i\, f(x_i) + b \sum_{i=1}^{n} f(x_i) = a\ E(X) + b.$$

$\square$

5 Variància. Propietats. Desigualtat de Txebixev.

La **variància** d'una variable aleatòria discreta X amb valors en S i funció de probabilitat f es denota per $var(X)$ i es defineix com

$$var(X) = E((X - E(X))^2) = \sum_{x \in S} \left(x - \sum_{x \in S} x\, f(x) \right)^2 f(x),$$

sempre que aquesta sèrie sigui absolutament convergent.

La **variància** d'una variable aleatòria contínua X amb funció de densitat $f(x)$, $x \in \mathbb{R}$, és:

$$var(X) = E((X - E(X))^2) = \int_{-\infty}^{+\infty} \left(x - \int_{-\infty}^{+\infty} x\, f(x)\, dx \right)^2 f(x)\, dx.$$

L'arrel quadrada de la variància s'anomena **desviació típica** de X.

Propietats de la variància.

1. $var(X) = E(X^2) - (E(X))^2$.
   Si X pren valors en un conjunt finit $S = \{x_1, \ldots, x_n\}$, aleshores:
   $var(X) = \sum_{i=1}^n x_i^2\, f(x_i) - (\sum_{i=1}^n x_i\, f(x_i))^2$.
   Si X és una variable aleatòria contínua, aleshores:
   $var(X) = \int_{-\infty}^{+\infty} x^2 f(x)\, dx - \left(\int_{-\infty}^{+\infty} x\, f(x)\, dx\right)^2$
2. Per a qualsevol variable X, la variància és no negativa, és a dir, $var(X) \geq 0$, i $var(X) = 0$ si i només si X és constant.
3. $\forall a, b \in \mathbb{R}$, $var(a\,X + b) = a^2\ var(X)$.
4. Si X pren valors en un conjunt finit S i $g : S \subseteq \mathbb{R} \to \mathbb{R}$ és una funció, $var(g(X)) = \sum_{x \in S}(g(x))^2\, f(x) - (E(g(X)))^2$.
   Si X és una variable aleatòria contínua, aleshores:
   $var(g(X)) = \int_{-\infty}^{+\infty}(g(x))^2 f(x)\, dx - (E(g(X)))^2$.
5. **Desigualtat de Txebixev** Si X és una variable aleatòria amb esperança m i variància $\sigma^2 < \infty$ i k és una constant positiva, aleshores

   $$P(|X - m| \geq k) \leq \sigma^2/k^2.$$

   Passant al complementari es té la desigualtat contrària:

   $$P(|X - m| < k) \geq 1 - \sigma^2/k^2.$$

   Aquest és un resultat molt important que s'utilitza per a estimar una probabilitat en el cas que no es conegui la distribució de la variable aleatòria.

Demostració: Demostrem algunes propietats de la variància:
1. $var(X) = E(X^2) - E(X)^2$.

$$var(X) = E((X - E(X))^2) = E(X^2 + E(X)^2 - 2X\,E(X)) = E(X^2) - E(X)^2.$$

2. $var(X) = 0 \Leftrightarrow X$ és constant.

$$var(X) = E((X - E(X))^2) = 0 \Leftrightarrow X = E(X).$$

3. $a, b \in \mathbb{R}$, $var(a\,X + b) = a^2\ var(X)$.

$$var(a\,X + b) = E((a\,X + b - E(a\,X + b))^2) = E((a\,X - a\,E(X))^2)$$

$$= E(a^2(X - E(X))^2) = a^2\, var(X).$$

□

Exemple 13 (continuació) *Calculem l'esperança i variància de X:*

$$E(X) = (-3)\frac{1}{27} + (-1)\frac{2}{9} + 1\frac{4}{9} + 3\frac{8}{27} = 1.$$

$$E(X^2) = 9\frac{1}{27} + 1\frac{2}{9} + 1\frac{4}{9} + 9\frac{8}{27} = \frac{99}{27},$$

$$\mathrm{var}(X) = E(X^2) - (E(X))^2 = \frac{99}{27} - 1^2 = \frac{72}{27}.$$

Exemple 16 *Una variable aleatòria X té la següent funció de densitat de probabilitat*

$$f(x) = \begin{cases} (1+x^2)/12, & si\ x \in (0,3), \\ 0, & si\ x \notin (0,3). \end{cases}$$

Calculeu:

a) la funció de distribució de X,

b) les probabilitats $P(1 < X < 2)$ i $P(X < 1)$,

c) l'esperança i variància de X,

d) la probabilitat $P(|X - E(X)| \geq 1)$ i compareu-la amb la fita que s'obtindria mitjançant la desigualtat de Txebixev.

a) Calculem la funció de distribució de X:

$$F(x) = P(X \leq x) = P(X \in (-\infty, x]) = \int_{-\infty}^{x} f(t)\,dt,$$

és a dir,

$$F(x) = \begin{cases} 0, & si\ x < 0, \\ (x + x^3/3)/12, & si\ 0 \leq x < 3, \\ 1, & si\ x \geq 3. \end{cases}$$

b) Calculem les probabilitats $P(1 < X < 2)$ i $P(X < 1)$:

$$P(1 < X < 2) = \int_1^2 f(x)\,dx = \int_1^2 (1+x^2)/12\,dx = 0.278,$$

$$P(X < 1) = \int_{-\infty}^{1} f(x)\,dx = \int_{-\infty}^{0} 0\,dx + \int_0^1 (1+x^2)/12\,dx = 0.111.$$

c) Calculem l'esperança i variància de X:

$$E(X) = \int_{-\infty}^{+\infty} x\,f(x)\,dx = \int_{-\infty}^{0} x\,0\,dx + \int_0^3 x\,(1+x^2)/12\,dx + \int_3^{+\infty} x\,0\,dx = 2.0625.$$

$$E(X^2) = \int_{-\infty}^{+\infty} x^2\,f(x)\,dx = \int_{-\infty}^{0} x^2\,0\,dx + \int_0^3 x^2\,(1+x^2)/12\,dx + \int_3^{+\infty} x^2\,0\,dx = 4.8,$$

i per tant, $\mathrm{var}(X) = E(X^2) - (E(X))^2 = 4.8 - 2.0625^2 = 0.546.$

d) Primer calculem exactament la probabilitat $P(|X - E(X)| \geq 1)$, o bé utilitzant la funció de densitat de X o bé la seva funció de distribució.

$$\begin{aligned} P(|X - E(X)| \geq 1) &= 1 - P(|X - E(X)| < 1) = 1 - P(-1 < X - E(X) < 1) \\ &= 1 - P(1.0625 < X < 3.0625) = 1 - P(1.0625 < X < 3) \\ &= 1 - [F(3) - F(1.0625)] = F(1.0625) = 0.1219, \end{aligned}$$

on hem tingut en compte que $E(X) = 2.0625$, i que X és una variable aleatòria contínua amb funció de densitat diferent de zero en l'interval $(0, 3)$. En cavi, mitjançant la desigualtat de Txebixev obtenim:

$$P(|X - E(X)| \geq 1) \leq \frac{\text{var}(X)}{1^2} = 0.546,$$

que no és fals, però tampoc és molt precís. Recordeu que aquesta desigualtat s'utilitza com una aproximació de la probabilitat quan no es disposa de la llei de probabilitat de la variable aleatòria.

Exemple 17 *S'ha observat que el nombre mig de clients per dia que entra en una determinada botiga és de 20, amb una desviació típica de 2 clients. Què es pot dir de la probabilitat que un determinat dia el nombre mig de clients per dia es trobi entre 17 i 23?*
Considerem la variable aleatòria $X =$"nombre de clients per dia", que té esperança $m = E(X) = 20$ i desviació típica $\sigma = 2$. Ens demanen calcular $P(17 < X < 23)$, per tant

$$P(17 < X < 23) = P(17 - 20 < X - 20 < 23 - 20) = P(-3 < X - 20 < 3)$$

$$= P(|X - 20| < 3) = 1 - P(|X - 20| \geq 3).$$

Aplicant la desigualtat de Txebixev,

$$P(|X - 20| \geq 3) \leq 4/9,$$

per tant,

$$P(|X - 20| < 3) \geq 1 - 4/9 = 5/9.$$

5.1 Moments d'una variable aleatòria.

A més de l'esperança matemàtica i la variància d'una variable aleatòria hi ha altres mesures que serveixen per a caracteritzar la distribució de la variable aleatòria: els **moments**.
Es defineix el **moment central d'ordre** k com

$$\mu_k = E[(X - E(X))^k] = E[(X - m)^k], \tag{5}$$

on $m = E(X)$, sempre que existeixi (és a dir, si la v.a. és discreta caldrà que la sèrie corresponent sigui convergent, i si la v.a. és contínua, caldrà que la integral corresponent sigui convergent). si X és una v.a. discreta que pren valors sobre un conjunt S amb funció de probabilitat f, aleshores la fórmula (5) s'escriu

$$\mu_k = \sum_{x \in S} (x-m)^k\, f(x) = \sum_{x \in S} (x-m)^k\, P(X = x).$$

Si X és una v.a. contínua amb funció de densitat $f(x)$, $x \in \mathbb{R}$, aleshores la fórmula (5) s'escriu

$$\mu_k = \int_{\mathbb{R}} (x-m)^k\, f(x)\, dx.$$

Casos particulars:

1. $\mu_0 = E[(X-m)^0] = 1$, el moment central d'orde 0 sempre val 1.
2. $\mu_1 = E[(X-m)^1] = E(X) - m = 0$, el moment central d'ordre 1, sempre val zero.
3. $\mu_2 = E[(X-m)^2] = var(X)$, el moment central d'ordre 2 és la variància de X.

El moment central d'ordre 3, μ_3, informa de la simetria de la v.a. Es defineix el **coeficient d'asimetria** de X com

$$\alpha_3 = \frac{\mu_3}{\mu_2^{3/2}}.$$

Si la distribució de X és simètrica respecte de m, aleshores $\alpha_3 = 0$. Si la probabilitat que X prengui valors superiors a m és més gran que la probabilitat que X prengui valors inferiors a m, aleshores $\alpha_3 > 0$ i, es diu que X és asimètrica cap a la dreta (de m). En cas contrari, és a dir, si la probabilitat que X prengui valors superiors a m és menor que la probabilitat que X prengui valors inferiors a m, aleshores $\alpha_3 < 0$ i, es diu que X és asimètrica cap a l'esquerra (de m).
El moment central d'ordre 4, μ_4, informa de l'apuntament de la distribució de X. Es defineix el **coeficient de curtosi** de X com

$$\alpha_4 = \frac{\mu_4}{\mu_2^2}.$$

Aquest coeficient val $\alpha_4 = 3$ per a la llei normal (que és la llei de referència). Si $\alpha_4 > 3$ es diu que la distribució de X és més apuntada que la llei normal, en canvi, si $\alpha_4 < 3$ es diu que la distribució de X és més plana que la llei normal.

Proposició 5.1 *Sigui X una v.a. amb esperança m i variància σ^2. Considerem la v.a. $Y = (X-m)/\sigma$, aleshores:*

1. *$E(Y) = 0$,*

2. $\mathrm{var}(Y) = 1$,

3. $\mu_r(Y) = \mu_r(X)/(\mu_2(X))^{r/2}$,

on μ_r és el moment central d'ordre r.

Demostració:

$$E(Y) = E\left(\frac{X-m}{\sigma}\right) = \frac{1}{\sigma}\,E(X-m) = \frac{1}{\sigma}(E(X)-m) = 0.$$

$$\mathrm{var}(Y) = \mathrm{var}\left(\frac{X-m}{\sigma}\right) = \frac{1}{\sigma^2}\,\mathrm{var}(X-m) = \frac{1}{\sigma^2}\;\mathrm{var}(X) = 1.$$

$$\begin{aligned}\mu_r(Y) = E(Y^r) = E\left(\left(\frac{X-m}{\sigma}\right)^r\right) &= E\left(\frac{(X-m)^r}{\sigma^r}\right)\\ = \frac{1}{\sigma^r}\;E((X-m)^r) = \frac{1}{\sigma^r}\,\mu_r(X) &= \frac{\mu_r(X)}{(\mu_2(X))^{r/2}},\end{aligned}$$

ja que $\sigma^r = (\mathrm{var}(X))^{r/2} = (\mu_2(X))^{r/2}$. $\square$

5.2 Altres mesures informatives: La mediana.

L'esperança matemàtica de X pot interpretar-se com el *centre* de la distribució de X. Existeix un altre punt en la recta real que també podria ser considerat com el centre de la distribució de X: és el punt que divideix la probabilitat total en dues parts iguals, és a dir, és el punt m_0 tal que la probabilitat a l'esquerra de m_0 és $1/2$ i la probabilitat a la dreta de m_0 també és $1/2$. Aquest punt s'anomena **mediana**. La definició formal de la mediana és:

- Si X és una v.a. contínua,
  $m_0 \in \mathbb{R}$ tal que $P(X \leq m_0) = F_X(m_0) = 1/2$.
- Si X és una v.a. discreta,
  $m_0 \in \mathbb{R}$ tal que $P(X \leq m_0) \geq 1/2$ i $P(X \geq m_0) \geq 1/2$, o equivalentment, $m_0 \in \mathbb{R}$ tal que $P(X < m_0) \leq 1/2$ i $P(X \geq m_0) \geq 1/2$.

Exemple 18 *Si X és una v.a. que només pren valors sobre el conjunt $S = \{1,2,3,4\}$ amb probabilitats $P(X=1) = 0.1$, $P(X=2) = 0.2$, $P(X=3) = 0.3$ i $P(X=4) = 0.4$, la mediana de X és $m_0 = 3$, ja que $P(X \leq 3) = 0.1+0.2+0.3 = 0.6 \geq 1/2$ i $P(X \geq 3) = 0.3+0.4 = 0.7 \geq 1/2$.*

Exemple 19 *Si X és una v.a. que només pren valors sobre el conjunt $S = \{1,2,3,4\}$ amb probabilitats $P(X=1) = 0.1$, $P(X=2) = 0.4$, $P(X=3) = 0.3$ i $P(X=4) = 0.2$, aleshores X té dues medianes: $m_0 = 2$ i $m_0 = 3$, ja que $P(X \leq 2) = 0.5$, $P(X \geq 2) = 0.9$ i $P(X \leq 3) = 0.8$, $P(X \geq 3) = 0.5$.*

Exemple 20 *Si X és una v.a. contínua amb funció de densitat*

$$f(x) = \begin{cases} 4\,x^3, & x \in (0,1), \\ 0, & x \notin (0,1), \end{cases}$$

la mediana de X és

$$m_0 = P(X \leq m_0) = \int_0^{m_0} 4\,x^3\,dx = m_0^4 = \frac{1}{2},$$

per tant, $m_0 = (1/2)^{1/4}$.