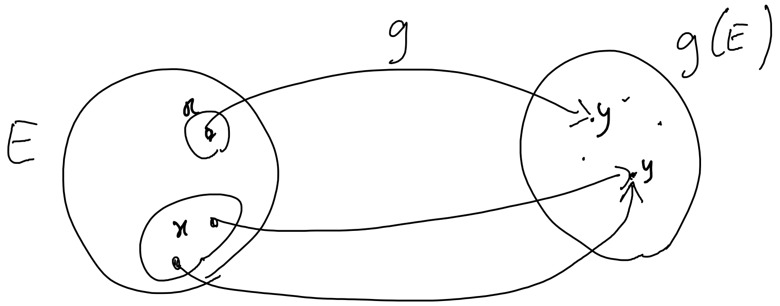
VARIANZA

Se abbiamo una variabile aleatoria X a valori in E e $g: E \to \mathbb{R}$, allora $g(X)$ è una v. al. reale.

Come si calcola $E[g(X)]$?

1° caso: se X è discreta, allora anche $g(X)$ è discreta, e

$$E[g(X)] = (\text{supponendo } E \text{ finito}) \sum_{y \in g(E)} y \, \mathbb{P}\{g(X) = y\}$$



$$= \sum_{y \in g(E)} y \, \mathbb{P}\Big(\bigcup_{\substack{x \text{ t.c.} \\ g(x)=y}} \{X = x\}\Big) = \sum_{y \in g(E)} y \sum_{\substack{x \text{ t.c.} \\ g(x)=y}} \mathbb{P}\{X = x\} =$$

$$= \sum_{y \in g(E)} \sum_{\substack{x \text{ t.c.} \\ y=g(x)}} g(x) \, \mathbb{P}\{X = x\} = \sum_{x \in E} g(x) \, \mathbb{P}\{X = x\}$$

<u>formula della speranza di una funzione composta</u>

2° caso: X continua di densità f, allora

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

ogni volta che l'integrale è ben definito

Se prendiamo $g(x) = (x - \mu)^2$, con $\mu = E[X]$, allora definiamo

$$\mathrm{Var}[X] = E[g(X)] = E[(X-\mu)^2] = E[(X - E[X])^2]$$

<u>varianza di X</u>

<u>esempio</u>: $X \sim Be(p)$, $E[X] = p$

$$\mathrm{Var}[X] = (1-p)^2 \cdot \mathbb{P}\{X = 1\} + (0-p)^2 \cdot \mathbb{P}\{X = 0\} =$$

$$= (1-p)^2 \cdot p + p^2(1-p) =$$

$$= p(1-p)(1-p+p) = p(1-p)$$

<u>esempio</u>: $X \sim N(0,1)$, $E[X] = 0$

$$\mathrm{Var}[X] = \int_{-\infty}^{+\infty} (x-0)^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx = \text{(int. per parti)}$$

$$= \left[-x \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} 1 \times \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx =$$

$$= 0 - 0 + 1 = 1$$

<u>esempio</u>: $X \sim N(\mu, \sigma^2)$, $E[X] = \mu$

$X = \sigma Y + \mu$, con $Y \sim N(0,1)$

$$\mathrm{Var}[X] = E[(X-\mu)^2] = E[(\sigma Y \cancel{+\mu} \cancel{-\mu})^2] =$$

$$= E[\sigma^2 Y^2] = \sigma^2 E[Y^2] = \sigma^2 \int^{+\infty} x^2 \frac{1}{\cdots} e^{-\frac{1}{2}x^2} dx =$$

$$= E[\sigma^2 Y] = \sigma^2 E[Y] = \sigma^2 \int_{-\infty} \frac{}{\sqrt{2\pi}} dx =$$

$$= \sigma^2 Var[Y] = \sigma^2 \times 1 = \sigma^2$$

Proprietà della varianza

1) se $X \equiv b \in \mathbb{R}$, allora $Var[X] = 0$, e anche

$Var[X] = 0 \Rightarrow X =$ costante

2) dopo

3) $Var[X] \geq 0$, e $Var[X] = 0 \Leftrightarrow X$ costante

4) se X e Y i.d., allora $Var[X] = Var[Y]$

La 2) si divide in:

2a) se $a \in \mathbb{R}$, allora

$$Var[aX] = a^2 Var[X]:$$

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

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

2b) se X, Y v.al., allora

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

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

$$= E\Big[\big(\underbrace{X - E[X]} + \underbrace{Y - E[Y]}\big)^2\Big] =$$

$$= E\Big[\big(X - E[X]\big)^2 + \big(Y - E[Y]\big)^2 + 2\big(X - E[X]\big)\big(Y - E[Y]\big)\Big] =$$

$$= Var[X] + Var[Y] + 2\,Cov(X, Y)$$

con $Cov(X, Y) = E\Big[\big(X - E[X]\big)\big(Y - E[Y]\big)\Big]$

covarianza tra X e Y: misura la dipendenza (lineare) tra X e Y

In particolare:

- se $Cov(X, Y) > 0$, X e Y positivamente correlate
- se $Cov(X, Y) < 0$, X e Y negativamente correlate
- se $Cov(X, Y) = 0$, X e Y scorrelate o non correlate

nota: $Cov(X, Y) = E\Big[\big(X - E[X]\big)\big(Y - E[Y]\big)\Big] =$

$$= E\Big[XY - XE[Y] - YE[X] + E[X]E[Y]\Big] =$$

$$= E[XY] - E\big[XE[Y]\big] - E\big[YE[X]\big] + E\big[E[X]E[Y]\big] =$$

$$= E[XY] - \cancel{E[Y]E[X]} - E[X]E[Y] + \cancel{E[X]E[Y]} =$$

$$= E[XY] - E[X]E[Y]$$

nota: se X e Y ind. e hanno speranza finita, allora

che $E[XY] = E[X]E[Y]$; allora $Cov(X, Y) = 0$

se poi X e Y hanno varianza finita, allora

$$\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$$

nota: $\mathrm{Cov}(X,X) = E\big[(X - E[X])(X - E[X])\big] = \mathrm{Var}[X]$

e $\mathrm{Var}[X] = E[X^2] - E[X]^2$

esempio: $X \sim \mathcal{B}(n,p)$, $E[X] = np$

$$\mathrm{Var}[X] = E\big[(X-np)^2\big] = \sum_{k=0}^{n} (K-np)^2 \binom{n}{K} p^K (1-p)^{n-K} = ?$$

usiamo le proprietà: X ha la stessa legge di $S_n = \sum_{i=1}^{n} Y_i$

con $(Y_i)_i$ i.i.d. $\sim Be(p)$; allora

$$\mathrm{Var}[X] \overset{(4)}{=} \mathrm{Var}[S_n] = \mathrm{Var}\Big[\sum_{i=1}^{n} Y_i\Big] \overset{(2b)}{=} \sum_{i=1}^{n} \mathrm{Var}[Y_i] =$$

$$= \sum_{i=1}^{n} p(1-p) = np(1-p)$$

esempio: $X \sim Po(\lambda)$ $\qquad E[X] = \lambda$

$$\mathrm{Var}[X] = E[X^2] - E[X]^2 = E[X^2] - \lambda^2$$

$$E[X^2] = \sum_{n=0}^{\infty} n^2 \cdot \mathbb{P}\{X=n\} = \sum_{n=\cancel{0}\,1}^{\infty} n^2 e^{-\lambda} \frac{\lambda^n}{n!} = \sum_{n=1}^{\infty} n\, e^{-\lambda} \frac{\lambda^n}{(n-1)!} =$$

$K = n-1$

$$= \sum_{K=0}^{\infty} (K+1) e^{-\lambda} \frac{\lambda^{K+1}}{K!} =$$

$$= \sum_{K=0}^{\infty} K e^{-\lambda} \frac{\lambda^{K+1}}{K!} + \sum_{K=0}^{\infty} 1 \cdot e^{-\lambda} \frac{\lambda^{K+1}}{K!} =$$

$$= \lambda\, e^{-\lambda} \sum_{K} K \frac{\lambda^K}{\ldots} + \lambda \sum_{K} e^{-\lambda} \lambda^K$$

$$= \lambda e^{-\lambda} \sum_{0}^{\infty} k\, K \frac{\lambda^K}{K!} + \lambda \sum_{0}^{\infty} k\, e^{-\lambda} \frac{\lambda^K}{K!} =$$

$$= \lambda E[X] + \lambda \times 1 = \lambda^2 + \lambda$$

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