Varianza di una v.a. discreta

Sia X una v.a. discreta, integrabile, e sia E l'insieme dei suoi valori. Si dice che X è centrata se $\mathbb{E}[X]=0$.
Se X è una v.a. discreta integrabile e $\mu \in \mathbb{R}$, l'unica scelta di μ che rende centrata la v.a. $X-\mu$ è $\mu = \mathbb{E}[X]$; la differenza $X - \mathbb{E}[X]$ si chiama la v.a. centrata associata a X.

Definizione Sia X una v.a. discreta tale che X^2 sia integrabile; si chiama varianza di X il numero

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

La varianza di X è una misura della dispersione dei valori di X attorno alla media $\mathbb{E}[X]$. Essa è nulla se e solo se $X - \mathbb{E}[X]$ è quasi certamente nulla, ossia X è equivalente ad una costante secondo P.
Il numero $\sqrt{\mathrm{Var}[X]}$ si chiama scarto quadratico medio, o deviazione standard, di X.

Si noti che

$$\mathrm{Var}[X+c] = \mathrm{Var}[X], \quad \mathrm{Var}[cX] = c^2\,\mathrm{Var}[X] \quad \forall c \in \mathbb{R}.$$

Risulta infine

$$\mathrm{Var}[X] = \sum_{k \in E} (k-\mathbb{E}[X])^2\, P(X=k) .$$

Siano adesso X, Y v.a. discrete con X^2, Y^2 integrabili, e con insiemi di valori E_X, E_Y rispettivamente.

<u>Definizione</u> la <u>covarianza</u> di X e Y è il numero

$$\mathrm{Cov}(X,Y) = \mathbb{E}\left[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])\right] = \mathbb{E}[XY]-\mathbb{E}[X]\,\mathbb{E}[Y].$$

Dunque la covarianza è la differenza fra la speranza del prodotto e il prodotto delle speranze. Si noti che

$$\mathbb{E}[XY] = \sum_{k\in E_{XY}} k\, P(XY=k) = \sum_{k\in E_{XY}} k \sum_{h\in E_X} P(XY=k, X=h) =$$

$$= \sum_{h\in E_X} \sum_{k\in E_{XY}} k\, P\left(Y=\frac{k}{h}, X=h\right) = \left[j=\frac{k}{h}\right]$$

$$= \sum_{h\in E_X} \sum_{j\in E_Y} jh\, P(Y=j, X=h).$$

Quindi, se le v.a X, Y sono indipendenti, allora

$$P(Y=j, X=h) = P(Y=j)\, P(X=h)$$

e dunque

$$\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y],$$

ed in tal caso

$$\mathrm{Cov}(X,Y) = 0.$$

<u>Teorema</u> Siano X, Y v.a. discrete con varianza finita. Allora

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

<u>dim.</u> Si ha

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

$$= \mathbb{E}[X^2+Y^2+2XY] - (\mathbb{E}[X]+\mathbb{E}[Y])^2 =$$

$$= \mathbb{E}[X^2] + \mathbb{E}[Y^2] + 2\,\mathbb{E}[XY] -$$

$$- (\mathbb{E}[X])^2 - (\mathbb{E}[Y])^2 - 2\,\mathbb{E}[X]\,\mathbb{E}[Y] =$$

$$= \mathrm{Var}[X] + \mathrm{Var}[Y] + 2\,\mathrm{Cov}(X,Y). \quad \square$$

<u>Esempi</u> (1) Sia X una v.a. bernoulliana di parametro p. Allora

$$\mathbb{E}[X] = P(X=1) = p,$$

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - p^2 = p^2 - p = pq.$$

(2) Sia X una v.a. dotata di legge binomiale $B(n,p)$. Allora, essendo $k\binom{n}{k} = n\binom{n-1}{k-1}$, si ha

$$\mathbb{E}[X] = \sum_{k=1}^{n} k\binom{n}{k} p^k q^{n-k} = n\sum_{h=0}^{n-1}\binom{n-1}{h} p^{h+1} q^{n-1-h} = np(p+q)^{n-1} = np.$$

mentre

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - n^2p^2 = \sum_{k=1}^{n} k^2 \binom{n}{k} p^k q^{n-k} - n^2p^2 =$$

$$= \sum_{k=2}^{n} (k^2-k)\binom{n}{k} p^k q^{n-k} + np - n^2p^2 =$$

$$= \sum_{k=2}^{n} n(n-1)\binom{n-2}{k-2} p^k q^{n-k} + np - n^2p^2 =$$

$$= n(n-1)p^2 + np - n^2p^2 = -np^2 + np = npq.$$

(3) Sia X una v.a. con legge di Poisson di parametro λ. Allora

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \sum_{h=0}^{\infty} e^{-\lambda} \frac{\lambda^{h+1}}{h!} = \lambda e^{-\lambda} \sum_{h=0}^{\infty} \frac{\lambda^h}{h!} = \lambda,$$

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \lambda^2 = \sum_{k=1}^{\infty} k^2 e^{-\lambda} \frac{\lambda^k}{k!} - \lambda^2 =$$

$$= \sum_{k=2}^{\infty} (k^2-k) e^{-\lambda} \frac{\lambda^k}{k!} + \lambda - \lambda^2 = e^{-\lambda}\lambda^2 \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} + \lambda - \lambda^2 = \lambda.$$

(4) Sia X una v.a. con legge geometrica di parametro p. Allora

$$\mathbb{E}[X] = \sum_{n=1}^{\infty} npq^{n-1} = p \sum_{n=1}^{\infty} \frac{d}{dq} q^n = p \frac{d}{dq} \sum_{n=0}^{\infty} q^n = p \frac{d}{dq} \frac{1}{1-q} =$$

$$= \frac{p}{(1-q)^2} = \frac{1}{p},$$

$$\mathrm{Var}[X] = \sum_{n=1}^{\infty} n^2 p q^{n-1} - \frac{1}{p^2} = \sum_{n=2}^{\infty} (n^2-n) p q^{n-1} + \frac{1}{p} - \frac{1}{p^2} =$$

$$= pq \sum_{n=2}^{\infty} \frac{d^2}{dq^2} q^n + \frac{1}{p} - \frac{1}{p^2} = pq \frac{d^2}{dq^2} \frac{1}{1-q} + \frac{1}{p} - \frac{1}{p^2} =$$

$$= \frac{2pq}{(1-q)^3} + \frac{1}{p} - \frac{1}{p^2} = \frac{2q}{p^2} + \frac{1}{p} - \frac{1}{p^2} = \frac{1}{p^2} - \frac{1}{p} = \frac{q}{p^2}.$$