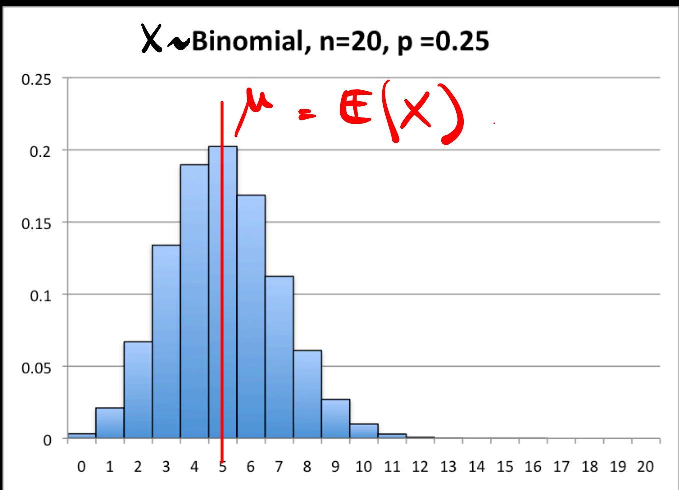
Richiamo: Consideriamo un esperimento con risultati

$$X \in \{0,1,2,\dots m\}$$



fdm $\mathbb{P}(X=k)$

$$\Longrightarrow \mathbb{P}(X\in B) = \sum_{k\in B} \mathbb{P}(X=k)$$

- $\mathbb{E}(X) = \sum_{k \in \mathrm{range}(X)} k\,\mathbb{P}(X=k)$
- $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$
- $\mathbb{E}(aX) = a\,\mathbb{E}(X)$
- $I_A = \begin{cases} 1 & \text{se } A \text{ avviene} \\ 0 & \text{altrimenti} \end{cases}$

$$\mathbb{P}(I_A = 1) = p$$

$$\mathbb{E}(I_A) = \mathbb{P}(I_A = 1) = p$$

$$X = \sum_{j=1}^{n} I_{A_j} \Longrightarrow \mathbb{E}(X) = \mathbb{E}\left(\sum_{j=1}^{n} I_{A_j}\right) =$$

$$= \sum_{j=1}^{n} \mathbb{E}(I_{A_j}) = \sum_{j=1}^{n} \mathbb{P}(I_{A_j} = 1)$$

<u>Es</u>: $X \sim \mathrm{Bin}(n,p)$ $\quad \mathbb{E}(X) = ?$

$$X = I_{A_1} + I_{A_2} + \dots + I_{A_n} \qquad A_j = \{\text{successo alla prova } j\}.$$

$$\mathbb{P}(I_{A_j} = 1) = p$$

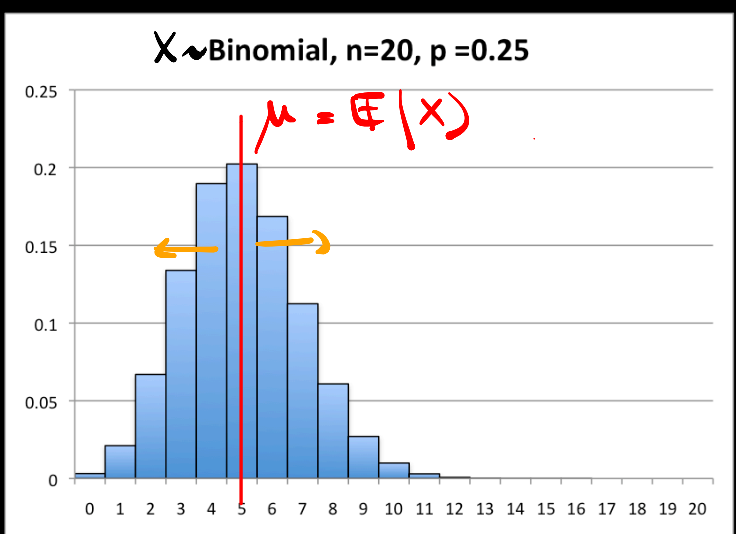
$$\mathbb{E}(X) = \mathbb{E}\Big(\sum_{j=1}^{n} I_{A_j}\Big) = \sum_{j=1}^{n} \mathbb{E}(I_{A_j})$$

$$= \sum_{j=1}^{n} \mathbb{P}(I_{A_j} = 1) = \sum_{j=1}^{n} p = n \cdot p.$$

Se avessi applicato la definizione di valore atteso

$$\mathbb{E}(X) = \sum_{k=0}^{n} k \cdot \mathbb{P}(X=k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \dots = n \cdot p$$



$$\mathrm{Var}(X) = \underbrace{\sum_{k \in \text{range}} (k-\mu)^2 \, \mathbb{P}(X=k)}_{\mathbb{E}((X-\mu)^2)}$$

$$\mu = \mathbb{E}(X)$$

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \geq 0$$

$$SD(X) = \sqrt{\mathrm{Var}(X)}$$

Somme di VA indipendenti

Richiamo: $\mathbb{E}(X+Y)=\mathbb{E}(X)+\mathbb{E}(Y)$

Cosa succede nel caso della varianza?

Richiamo
- Se $X \perp\!\!\!\perp Y$ $\mathbb{E}(XY)=\mathbb{E}(X)\mathbb{E}(Y)$
- $\mathrm{Var}(X)=\mathbb{E}((X-\mu)^2)$

Fatto: Se X, Y sono indipendenti

$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)$$

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

$$=\mathbb{E}\Big(\big(\underbrace{(X-\mathbb{E}(X))}_{\tilde{X}}+\underbrace{(Y-\mathbb{E}(Y))}_{\tilde{Y}}\big)^2\Big)=\mathbb{E}((\tilde{X}+\tilde{Y})^2)$$

$$=\mathbb{E}(\tilde{X}^2+2\tilde{X}\tilde{Y}+\tilde{Y}^2)=\mathbb{E}(\tilde{X}^2)+\mathbb{E}(2\tilde{X}\tilde{Y})+\mathbb{E}(\tilde{Y}^2)$$

$$=\underbrace{\mathbb{E}((X-\mathbb{E}(X))^2)}_{\mathrm{Var}(X)}+2\,\cancel{\underbrace{\mathbb{E}(X-\mathbb{E}(X))}_{\mathbb{E}(X)-\mathbb{E}(X)=0}}^{\,0}\,\mathbb{E}(Y-\mathbb{E}(Y))+\underbrace{\mathbb{E}((Y-\mathbb{E}(Y))^2)}_{\mathrm{Var}(Y)}.$$

Es: lancio due volte una moneta $I_1, I_2 \in \{0,1\}$

$$\mathrm{Var}(I_1)=\frac{1}{2}\left(1-\frac{1}{2}\right)=\frac{1}{4} \Rightarrow \mathrm{Var}(I_1+I_2)=\mathrm{Var}(I_1)+\mathrm{Var}(I_2)$$

$$=\frac{1}{4}+\frac{1}{4}=\frac{1}{2}$$

In generale Siano $X_1, \dots X_n$ indipendenti allora

$$\boxed{\mathrm{Var}(X_1+X_2+\dots+X_n)=\mathrm{Var}(X_1)+\dots+\mathrm{Var}(X_n)}$$

Se $X_1, \dots X_n$ hanno la stessa distribuzione e indipend.
(Indipendenti e Identicamente Distribuite IID)

$\Longrightarrow \mathbb{E}(X_i) = \mu$, $Var(X_i) = \sigma^2$ per ogni i

Definiamo $S_n = X_1 + \dots + X_n$

- $\mathbb{E}(S_n) = \mathbb{E}(X_1 + \dots + X_n) = \sum_{i=1}^{n} \mathbb{E}(X_i) = n\,\mathbb{E}(X_1)$

$\Longrightarrow$
- $Var(S_n) = Var(X_1 + \dots X_n) = \sum Var(X_i) = n\,Var(X_1)$
- $SD(S_n) = \sqrt{Var(S_n)} = \sqrt{n} \cdot \sqrt{Var(X_1)} = \sqrt{n}\,SD(X_1)$

<u>Esempio</u>: lancio 50 dadi con risultato X_j $\quad \mathbb{E}(X_j) = 3.5$

S_{50} = somma dei risultati $= X_1 + \dots + X_{50}$ $\quad Var(X_j) = 2.91$

$\mathbb{E}(S_{50}) = 50 \cdot 3.5 = 175$ $\qquad Var(S_n) = 50\,Var(X_1) = 50 \cdot 2.91$
$= 145$

Traslazioni e scaling

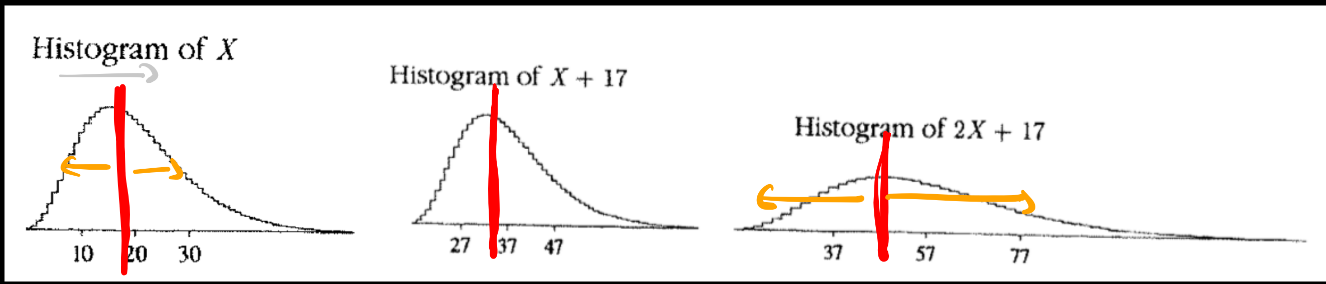
Consideriamo $Y = \underbrace{a}_{\text{scaling}} \cdot X + \underbrace{b}_{\text{traslazione}}$ per a, b fissi

Allora $\mathbb{E}(Y) = \mathbb{E}(aX + b) = \mathbb{E}(aX) + \mathbb{E}(b) = a\,\mathbb{E}(X) + b$

$Var(Y) = \mathbb{E}((Y - \mathbb{E}(Y))^2) = \mathbb{E}((aX \cancel{+b}) - \underbrace{\mathbb{E}(aX+b)}_{a\mathbb{E}(X) \cancel{+b}})^2)$

$= \mathbb{E}(a^2(X - \mathbb{E}(X))^2) = a^2\,\mathbb{E}((X - \mathbb{E}(X))^2) =$

$= a^2\,Var(X)$.



traslare la VA non influenza la forma della sua fdm, e quindi varianza, ma solo il valore medio (centro della distribuzione) invece scalare la VA influenza anche la varianza.

Ex: $\mathrm{Var}(X_j) = 7$ $\quad \mathrm{Var}(n X_j) = n^2 \mathrm{Var}(X_j)$

$= n^2 7$

⚠ questo non è uguale a

$$\mathrm{Var}(S_n) = \mathrm{Var}(X_1 + \dots + X_n) = n \,\mathrm{Var}(X_j) = n7$$

Consideriamo: $\bar{X}_n = \frac{1}{n}(X_1 + X_2 + \dots X_n)$ (media empirica)

- $$\mathbb{E}(\bar{X}_n) = \mathbb{E}\left(\frac{1}{n}(X_1 + \dots + X_n)\right) = \frac{1}{n}\mathbb{E}(X_1 + \dots + X_n) = \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}(X_j)$$
$$= \frac{1}{n} \cdot n\,\mathbb{E}(X_1) = \mathbb{E}(X_1)$$

- $$\mathrm{Var}(\bar{X}_n) = \mathrm{Var}\left(\frac{1}{n}(X_1 + \dots + X_n)\right) = \frac{1}{n^2}\cdot \mathrm{Var}(X_1 + \dots X_n) = \frac{1}{n^2}\sum \mathrm{Var}(X_j)$$
$$= \frac{1}{n^{\cancel{2}}} \cdot \cancel{n}\,\mathrm{Var}(X_1) = \frac{1}{n}\mathrm{Var}(X_1)$$

Esempio: $X \sim Bin(n, p)$ $\quad X = I_{A_1} + I_{A_2} + \dots + I_{A_n}$

$$\mathbb{E}(X) = n \cdot p.$$

$$Var(X) = Var(I_{A_1} + \dots + I_{A_n}) = n \cdot Var(I_{A_1}) = np(1-p)$$

$$\overline{X}_n = \frac{1}{n}(I_{A_1} + \dots + I_{A_n}) = \frac{1}{n} X$$

$$\mathbb{E}(\overline{X}_n) = \mathbb{E}(I_{A_1}) = \mathbb{P}(I_{A_1} = 1) = p$$

$$Var(\overline{X}_n) = \frac{1}{n} Var(I_{A_1}) = \frac{1}{n} p(1-p)$$

Esempio dell' esempio $\quad p = \frac{1}{2}$

$$\mathbb{E}(\overline{X}_n) = \frac{1}{2}$$

$$Var(\overline{X}_n) = \frac{1}{n}\frac{1}{4} \xrightarrow{n \to \infty} 0$$

$\longrightarrow$ dispersione $\to 0$

$$\mathbb{P}\left(\left|\overline{X}_n - \frac{1}{2}\right| > \varepsilon\right) \xrightarrow{n \to \infty} 0 \qquad \forall \varepsilon > 0$$