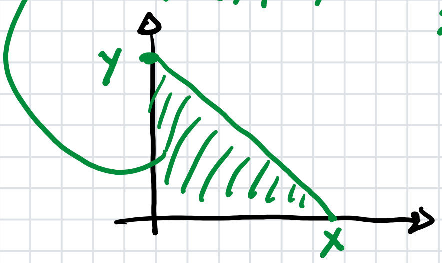
- Ricordiamo dalla lezione precedente:
  Esercizio: Calcolare l'area media del triangolo definito dai vertici (0,0), (X,0), (0,Y), dove (X,Y) è un vettore di variabili casuali uniforme discreto con X > Y, per X in {2,3,4} e Y in {1,2,3}

$A = A(X,Y) = \frac{XY}{2}$



DALLA SCORSA LEZIONE:

$\mathbb{E}[A] = \frac{35}{12}$

ABBIAMO VISTO CHE:

$$\mathbb{E}[A] \neq \frac{\mathbb{E}[X]\,\mathbb{E}[Y]}{2} = A(\mathbb{E}[X], \mathbb{E}[Y])$$

PERCHÈ X E Y NON SONO INDIP. E QUINDI $\mathbb{E}[XY] \neq \mathbb{E}[X]\,\mathbb{E}[Y]$

DOMANDA:

$$\mathbb{E}[A \mid Y=2] = \sum_{x=3}^{4} A(x,2)\,\frac{\mathbb{P}[X=x, Y=2]}{\mathbb{P}[Y=2]}$$

$$= \sum_{x=3}^{4} \frac{x\cdot 2}{2}\,\frac{1/6}{1/3} = \frac{1}{2}(3+4) = \frac{7}{2}$$

IN GENERALE POSSIAMO CALCOLARE:

$$\mathbb{E}[A \mid Y=y] = \sum_{x=y+1}^{4} \frac{xy}{2}\,\frac{\mathbb{P}[X=x, Y=y]}{\mathbb{P}[Y=y]}$$

$$= \sum_{x=1}^{4} \frac{xy}{2}\,\mathbb{P}[X=x \mid Y=y]$$

ANCORA PIÙ IN GENERALE, PER QUALSIASI FUNZIONE $g(x,y)$:

$$\mathbb{E}[g(X,Y) \mid Y=y] = \sum_{x} g(x,y)\,\mathbb{P}[X=x \mid Y=y]$$

$$= \sum_{x} g(x,y)\, p_{X|Y}(x|y)$$

• SE $X \perp Y \Rightarrow p_{X|Y}(x|y) = p_X(x)$

$$\underbrace{\mathbb{E}[g(X,Y) \mid Y=y]}_{\text{IL VALORE ATTESO CONDIZIONATO È UNA FUNZIONE DEL EVENTO CONDIZIONANTE}} = \sum_x g(x,y)\, p_X(x) \to \text{DIPENDE DAL VALORE } y$$

IL VALORE ATTESO CONDIZIONATO È UNA FUNZIONE DEL EVENTO CONDIZIONANTE → SE QUESTO EVENTO DIPENDE DAL VALORE DI UN'ALTRA VAR, ALLORA SARÀ FUNZIONE DI QUEL VALORE

IN GENERALE, ANCHE SE X E Y NON SONO INDIP:

$$\mathbb{E}[g(X,Y) \mid Y=y] = h(y) \to \forall\, y \text{ NEL SUPPORTO DI } Y$$

POTREMMO CHIEDERCI

$$\mathbb{E}\big[\mathbb{E}[g(X,Y) \mid \underline{Y}]\big] = \mathbb{E}[h(\underline{Y})] = \sum_y h(y)\, p_Y(y)$$

TUTTI I POSSIBILI VALORI DI Y

$$= \sum_y \underbrace{\mathbb{E}[g(X,Y) \mid Y=y]}\, p_Y(y)$$

$$= \sum_y \Big(\sum_x g(x,y)\, \underline{p_{X|Y}(x|y)} \Big)\, p_Y(y)$$

$$= \sum_y \sum_x g(x,y)\, \frac{p_{X,Y}(x,y)}{\cancel{p_Y(y)}}\, \cancel{p_Y(y)}$$

CI FERMIAMO QUI SE LA g DIPENDE DA X e Y →

$$= \sum_x \sum_y g(x,y)\, p_{X,Y}(x,y)$$

MA SE $g(X,Y) = g(X)$ DIPENDE SOLO DA X

$$\mathbb{E}\left[\mathbb{E}[g(X) \mid Y]\right] = \sum_x \sum_y g(x)\, p_{X,Y}(x,y)$$

$$= \sum_x g(x) \underbrace{\sum_y p_{XY}(x,y)}$$

$$= \sum_x g(x)\, p_X(x) = \mathbb{E}[g(X)]$$

QUINDI:

$$\mathbb{E}\left[\mathbb{E}[X \mid Y]\right] = \mathbb{E}[X] \rightarrow g(X)=X$$

$$\mathbb{E}\left[\mathbb{E}[g(X) \mid Y]\right] = \mathbb{E}[g(X)] \rightarrow \text{per qualsiasi } g(X)$$

$\Rightarrow$ QUELLO CHE ABBIAMO FATTO È SCRIVERE

$$p_X(x) = \sum_y p_{XY}(x,y)$$

$$= \sum_y p_{X|Y}(x|y)\, p_Y(y)$$

$\Rightarrow$ USANDO LA LEGGE DELLE PROBABILITÀ TOTALI.

$Y \sim \text{Pois}(\lambda = 15)$

Ricordiamo dalla lezione 10:

Esempio: Il numero di errori in un sistema si può modellare con una distribuzione di Poisson di parametro 15. Ogni errore è fatale con probabilità 0.2. Qual è la distribuzione congiunta del numero di errori e il numero di errori fatali? Qual è la distribuzione marginale del numero di errori fatali?

Qual è il valore numero atteso di errori fatali nel sistema?

$X \mid Y \sim \text{Bin}(n = Y, p = 0.2)$

SENZA DOVER CONOSCERE $p_X(x)$ POSSIAMO CALCOLARE $E[X]$

1) $E[X \mid Y] = \underbrace{0.2\, Y}_{\text{DIPENDE DAL NUMERO TOT. DI ERRORI}}$

IL VALORE ATTESO DELLA BINOMIALE È np

2) $E[X] = E\left[E[X \mid Y]\right] = E[0.2\, Y] = 0.2\, \underbrace{E[Y]}_{\lambda = 15}$

$= 0.2 \cdot 15 = \dfrac{2 \cdot 15}{10} = 3$

SE $g(X) = (X - \underbrace{E[X]}_{\text{UN NUMERO NON CASUALE}})^2$

$E[g(X)] = E\left[(X - E[X])^2\right] \equiv \text{Var}[X]$

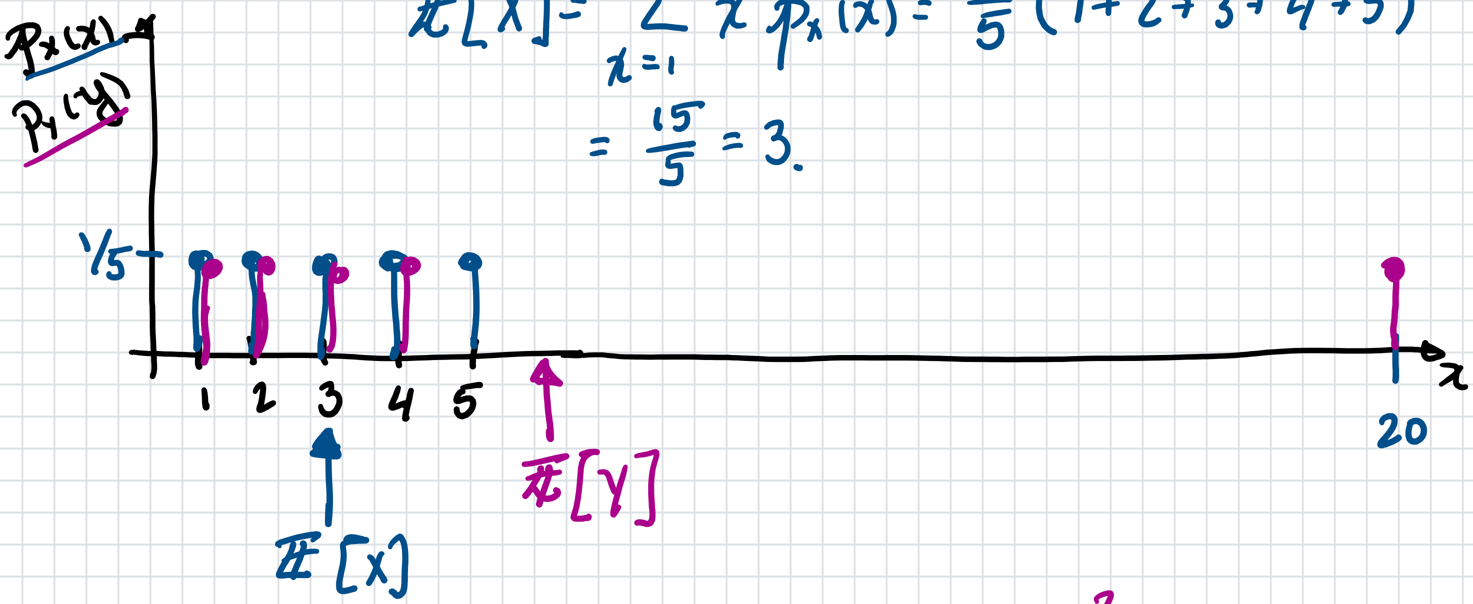
DEFINIZIONE DI VARIANZA.

TORNIAMO PROSSIMA LEZIONE

ESEMPIO: $X \sim UD\{1, 2, 3, 4, 5\}$

$$p_X(x) = \begin{cases} 1/5 & \text{se } x \in \{1, \dots, 5\} \\ 0 & \text{altrimenti} \end{cases}$$

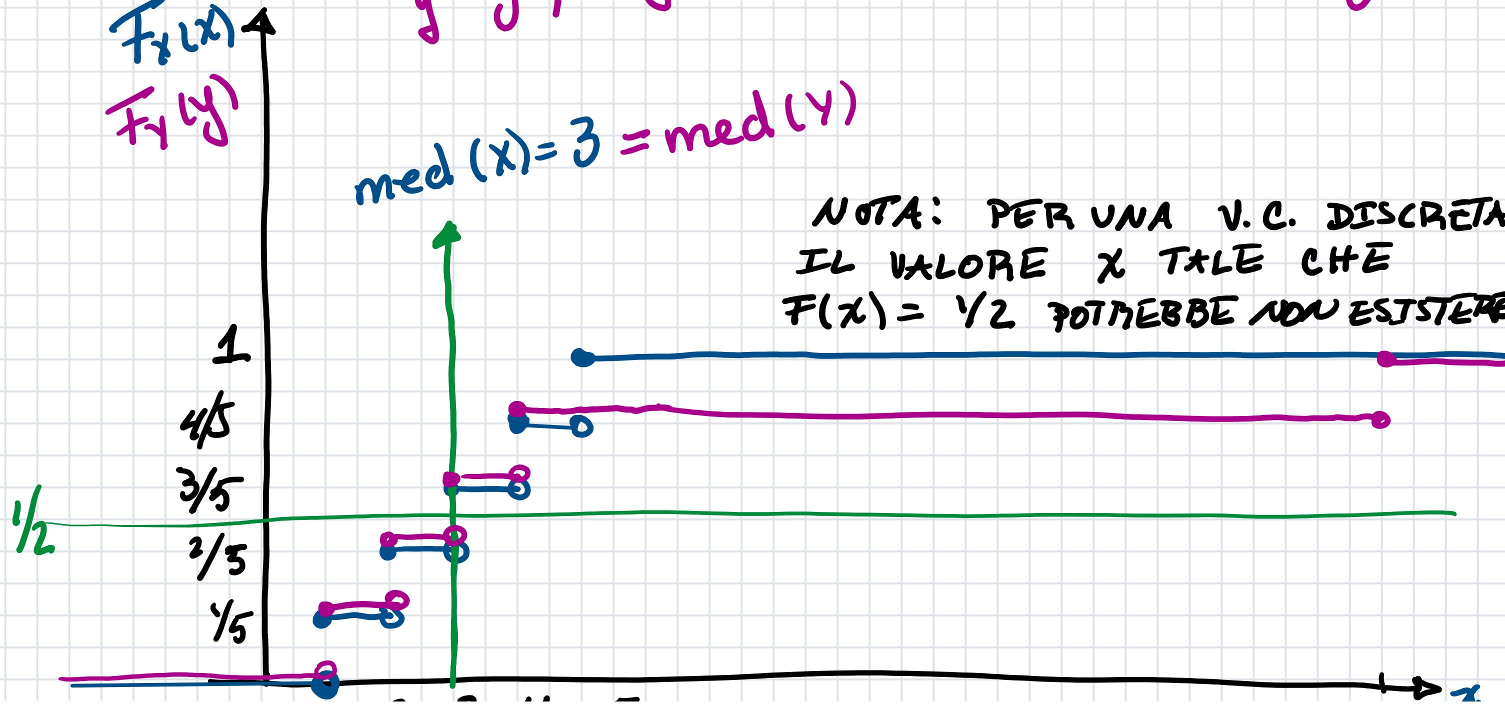
$$\mathbb{E}[X] = \sum_{x=1}^{5} x\, p_X(x) = \frac{1}{5}(1+2+3+4+5) = \frac{15}{5} = 3.$$



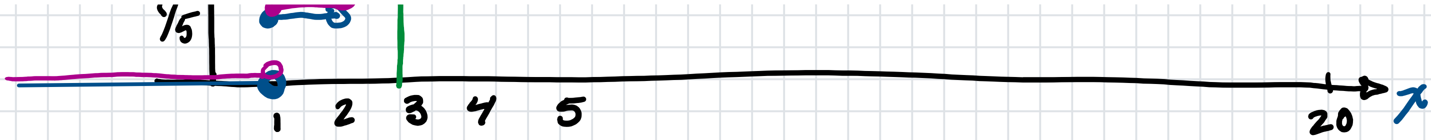
$Y \sim UD\{1, 2, 3, 4, 20\}$

$$p_Y(y) = \begin{cases} 1/5 & \text{se } y \in \{1,2,3,4\} \cup \{20\} \\ 0 & \text{altrimenti} \end{cases}$$

$$\mathbb{E}[Y] = \sum_{y} y\, p_Y(y) = \frac{1}{5}(1+2+3+4+20) = \frac{30}{5} = 6$$



NOTA: PER UNA V.C. DISCRETA IL VALORE x TALE CHE $F(x) = 1/2$ POTREBBE NON ESISTERE



4/5
1
2
3
4
5
20
x

NOTA: $\mathbb{E}[X]$ minimiza l'errore quadratico medio:

$$\min_{a \in \mathbb{R}} \mathbb{E}[(X-a)^2] = \mathbb{E}[(X-\mathbb{E}[X])^2] = Var[X]$$

med(X) minimiza l'errore assoluto (medio)

$$\min_{a \in \mathbb{R}} \mathbb{E}[|X-a|] = \mathbb{E}[|X-med(X)|]$$

SE LA DISTRIBUZIONE È SIMMETRICA

$$\mathbb{E}[X] = med(X)$$

MEDIA E MEDIANA SONO MISSURE DI TENDENZA CENTRALE (POSIZIONE).
LA VARIANZA È UNA MISSURA DI DISPERSIONE.

★ Per la prossima lezione ricordatevi di leggere, dal capitolo 5 del libro di testo (Variance, Higher Moments and Random Sums) le sezioni 5.2 a 5.6 e/o le pagine 14-23 del quarto set di slides (4 CostantiCaratteristiche).