Lezione 13

Ieri abbiamo visto la prima distribuzione di probabilità (discreta): la distribuzione binomiale o di Bernoulli.

$X := \#$ teste su n lanci $\rightsquigarrow \Omega = \{0, 1, \dots, n\}$ [n fissato] [p fissato = prob. che esca testa]

$$\mathbb{P}(X=k) \doteq \binom{n}{k} p^k (1-p)^{n-k}$$

$$\| $$

$$\mathbb{P}\Big(\bigcup \underbrace{TTTCCTC\dots T}_{k \text{ teste}} \Big)$$

unione di tutte le stringhe di testa con k lettere T e $n-k$ lettere C.

$$\|$$

$$\sum_{\substack{\text{stringhe} \\ \|}} \mathbb{P}(\text{stringa come sopra}) = \mathbb{P}(T \cap T \cap T \cap C \cap C \dots)$$

$\| \leftarrow$ ogni lancio è indipendente dagli altri

$$\sum_{\text{stringhe}} \underbrace{\mathbb{P}(T)}_{=p}\underbrace{\mathbb{P}(T)}_{=p}\underbrace{\mathbb{P}(T)}_{=p}\underbrace{\mathbb{P}(C)}_{=1-p}\cdot\ldots\cdot\underbrace{\mathbb{P}(T)}_{=p}$$

$$\| $$

$$\sum_{\text{stringhe}} p^k(1-p)^{n-k} = p^k(1-p)^{n-k}\left(\sum_{\text{stringhe}} 1\right) = p^k(1-p)^{n-k}\underbrace{(\#\text{stringhe})}_{\text{modi di distribuire } k \text{ teste su } n \text{ oggetti}}$$

$$\|$$

$$p^k(1-p)^{n-k}\binom{n}{k}$$

§ III.2. DISTRIBUZIONE DI POISSON

Modello: arrivo di
- clienti alle poste
- utenti in un server
- decadimenti radioattivi
- virus che infettano cellule

in un certo intervallo di tempo

$X :=$ # eventi accaduti nell'intervallo di tempo fissato.

Come si fa a calcolare $\mathbb{P}(X=k)$?

$\Omega = \mathbb{N} = \{0, 1, 2, \dots\}$

suddivido il nostro intervallo di tempo in n intervallini.

Def. una nuova variabile aleatoria:

$Y_n :=$ # intervallini in cui accade almeno un evento [esempio: arriva almeno un cliente]

[penso n grande.

Se prendo n grande abbastanza:

$$\underbrace{\mathbb{P}(Y_n = k)}_{\text{questo lo sappiamo fare invece}} = \underbrace{\mathbb{P}(X = k)}_{\text{calcolare questo era troppo difficile}}$$

$$\mathbb{P}(Y_n = k) = \binom{n}{k} p^k (1-p)^{n-k} =$$

Come prima sia $p = \mathbb{P}(\text{evento})$ che arrivi un cliente

[$np = \mathbb{E}[Y_n] = \mu$ è il valore atteso come ieri]

Adesso vogliamo prende il limite con n grande perché il problema iniziale non dipende da n!

esercizio: semplificate il binomiale

$$= \ldots = \boxed{\frac{\mu^k}{k!}} \underbrace{\frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdot \ldots \cdot \left(1-\frac{k-1}{n}\right)}{\left(1-\frac{\mu}{n}\right)^k}} \underbrace{\left(1-\frac{\mu}{n}\right)^n}$$

$\lim_{n\to\infty}$

$$\frac{\mu^k}{k!} \cdot 1 \cdot e^{-\mu}$$

$$\Rightarrow \mathbb{P}(X=k) = \underbrace{\frac{\mu^k}{k!} e^{-\mu}}_{\geq 0}$$

→ è sempre in $[0,1]$?

→ è una probabilità? ← basta far vedere questa ⊕ verificare che sono non negative, per rispondere anche alla prima domanda.

$$\sum_{k=0}^{\infty} \mathbb{P}(X=k) \overset{?}{=} 1$$

$$\sum_{k=0}^{\infty} = \sum_{k \in \Omega}$$

$$e^x = \frac{1}{0!} + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

<u>Conto</u>: $\displaystyle\sum_{k=0}^{\infty} \frac{\mu^k}{k!} e^{-\mu} = e^{-\mu} \boxed{\underbrace{\left(\sum_{k=0}^{\infty} \frac{\mu^k}{k!} \right)}_{= e^{\mu}}} = e^{-\mu} \cdot e^{\mu} = 1$ ✓

Domande che ci rimangono?

$Var[X]$, σ_X, $\mathbb{E}[X]$ ← vogliamo verificare che sia veramente μ.

$$\mathbb{E}[X] = \sum_{x_i \in \underbrace{X(\Omega)}_{\text{è } \mathbb{N}}} x_i \cdot \mathbb{P}(X=x_i) = \sum_{k=\cancel{0}\,1}^{\infty} k \cdot \mathbb{P}(X=k) = \boxed{\sum_{k=\cancel{0}\,1}^{\infty} \cancel{k} \cdot \frac{\mu^{k-1}}{\cancel{k!}\,(k-1)!} e^{-\mu}}\,\mu$$

$$= \mu \cdot e^{-\mu} \cdot \left(\sum_{k=\cancel{0}\,1}^{\infty} \frac{\mu^{k-1}}{(k-1)!} \right) = \mu \cdot e^{-\mu} \cdot \underbrace{\left(\sum_{t=0}^{\infty} \frac{\mu^t}{t!} \right)}_{=e^{\mu}} = \mu \cdot \cancel{e^{-\mu}} \cdot \cancel{e^{\mu}} = \mu$$

cambio di indice: $\boxed{t = k-1}$

$$\mathrm{Var}[X] = \sum_{x_i \in X(\Omega)} \mathbb{P}(X=x_i)\left(x_i - \mathbb{E}[X]\right)^2 = \sum_{k=0}^{\infty} \frac{\mu^k}{k!} e^{-\mu} \underbrace{(k-\mu)^2}_{= k^2 + \mu^2 - 2k\mu}$$

$$= \underbrace{e^{-\mu} \sum_{k=0}^{\infty} \frac{\mu^k}{k!} k^2}_{= \mu^2 + \mu} - \underbrace{2\mu\, e^{-\mu} \sum_{k=0}^{\infty} \frac{\mu^k}{k!} \cdot k}_{= -2\mu^2} + \underbrace{\mu^2 e^{-\mu} \underbrace{\left(\sum_{k=0}^{\infty} \frac{\mu^k}{k!} \right)}_{e^{\mu}}}_{= \mu^2}$$

$$= \cancel{\mu^2} + \mu - \cancel{2\mu^2} + \cancel{\mu^2} = \mu$$

$$\sigma_X = \sqrt{\mu}.$$

Cosa abbiamo fatto?

Introdotto X di Poisson, di valore atteso $= \mu$.

$$\mathbb{P}(X=k) = \frac{\mu^k}{k!} e^{-\mu} \qquad \Omega = \{0, 1, \dots, \dots\} = \mathbb{N}.$$

$$\mathbb{E}[X] = \mu. \qquad \boxed{\mathrm{Var}[X] = \mu} \qquad \sigma_X = \sqrt{\mu}$$

Abbiamo che per la distribuzione di Poisson.

Varianza = Valore atteso!

Distribuzioni che abbiamo visto:

- Binomiale o di Bernoulli (n,p) $\mathbb{P}$ Ω $\mathbb{E}$ Var σ — Ω finito
- Poisson (μ) " " " " " — Ω infinito ma discreto

E se Ω non è discreto ?!?

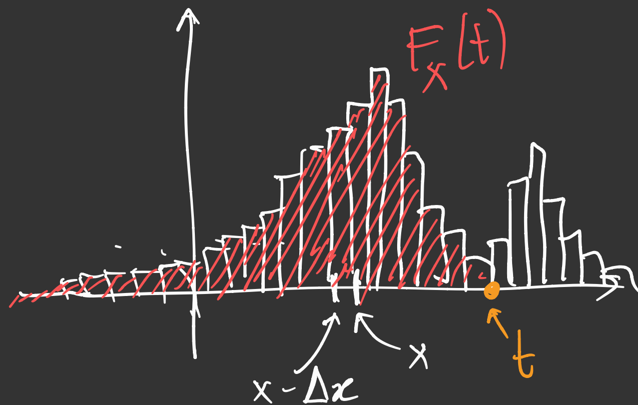
Sostanzialmente: $\boxed{\sum \longmapsto \int}$

DEF (VARIABILE ALEATORIA CONTINUA) (a valori reali)

$X : \Omega \longrightarrow \mathbb{R}$ è una v.a.c. $\overset{DEF}{\Longleftrightarrow}$

$$\forall x \in \mathbb{R} \quad \exists \lim_{\Delta x \to 0^+} \frac{\mathbb{P}(x - \Delta x < X \leq x)}{\Delta x} =: f_X(x)$$

e se f_X è integrabile su ogni intervallo limitato di $\mathbb{R}$.



$\lim_{\Delta x \to 0}$

DA NON CONFONDERE:

$$\begin{cases} X = \text{VARIABILE ALEATORIA} \\ x \in \mathbb{R} \end{cases}$$

$\boxed{\text{DEF}}$ (FUNZIONE di DISTRIBUZIONE)

$$F_X : \mathbb{R} \longrightarrow [0,1]$$

$$F_X(t) := \mathbb{P}(X \leq t) = \int_{-\infty}^{t} f_X(x)\,dx$$

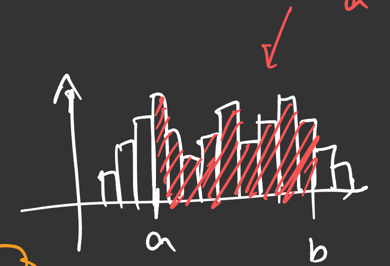
$$f_X = \frac{dF_X}{dx}$$

$\boxed{\text{OSS}}$ $\boxed{\int_{-\infty}^{\infty} f_X(x)\,dx = 1}$, $\lim_{t \to -\infty} F_X(t) = 0$, $\boxed{\lim_{t \to +\infty} F_X(t) = 1}$

$\Rightarrow$ = AREA TOTALE DEVE ESSERE 1.

OSS $\mathbb{P}(a \le X \le b) = \int_a^b f_X(x)\,dx$

$= F_X(b) - F_X(a)$



F_X è la PRIMITIVA di f_X ← DENSITÀ DI PROBABILITÀ

f_X è la DERIVATA di F_X ← FUNZIONE DI DISTRIBUZIONE

OSS $\mathbb{P}(X > t) \overset{?}{=} 1 - \mathbb{P}(X \le t) = 1 - F_X(t)$.

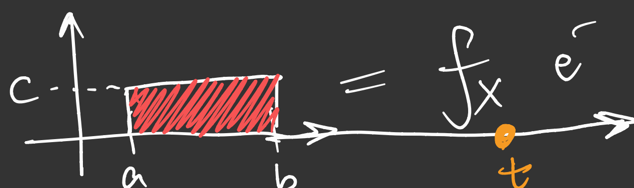
OSS $\underline{\mathbb{P}(a \le X \le b)} = \underline{\mathbb{P}(a < X \le b)} = \underline{\mathbb{P}(a \le X < b)}$

$= \underline{\mathbb{P}(a < X < b)}$

Perché $\boxed{\mathbb{P}(X=k)=0}$ per le variabili continue, perché $\{k\}$ in tutto $\mathbb{R}$ ha "misura nulla", e la sua probabilità è zero.

Questo è uno dei passaggi chiave dal DISCRETO al CONTINUO.

§ III.3. DISTRIBUZIONE UNIFORME (CONTINUA) sull'intervallo $[a,b]$.



$= f_X$ è la funzione costante su $[a,b]$.

$c = ?$ La normalizzazione ci dice

Area totale $= 1$

$\parallel$

$$1 = \int_{-\infty}^{\infty} f_X(x)\,dx = c \cdot \left(\int_a^b 1 \right) = (b-a) \cdot c$$

$$f_X(x) = \begin{cases} c & \text{su } [a,b] \\ 0 & \text{altrove} \end{cases}$$

Possiamo concludere

$$c = \frac{1}{(b-a)}$$

$$F_X(t) = \begin{cases} 0 & \text{se } t \leq a \\ \frac{t-a}{b-a} & \text{se } t \in [a,b] \\ 1 & \text{se } t \geq b \end{cases}$$

(da verificare!)

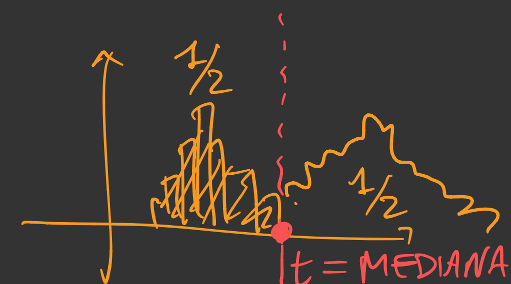
Cosa sono allora $\mathbb{E}[X]$, $Var[X]$ e σ_X per variabili continue?

- $$\mathbb{E}[X] := \underbrace{\int_{-\infty}^{\infty} x \cdot f_X(x)\, dx}_{\text{prima era } \sum_{x \in X(\Omega)} x \cdot \mathbb{P}(X=x)}$$

- $$Var[X] := \underbrace{\int_{-\infty}^{\infty} f_X(x)\,(x - \mathbb{E}[X])^2\, dx}_{\text{prima era } \sum_{x \in X(\Omega)} (x-\mathbb{E}[X])^2\, \mathbb{P}(X=x)}$$

- $$\sigma_X := \sqrt{Var[X]}$$

- $$\text{Mediana} := F_X^{-1}\left(\tfrac{1}{2}\right)$$



Proprietà: (dim. per esercizio)

numero

- $\mathbb{E}[X_1 + X_2] = \mathbb{E}[X_1] + \mathbb{E}[X_2]$, $\quad \mathbb{E}[\lambda X] = \lambda \cdot \mathbb{E}[X]$.
- $\mathrm{Var}[\underbrace{X_1 + X_2}_{}] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2]$ $\quad \mathrm{Var}[\lambda \cdot X] = \lambda^2 \cdot \mathrm{Var}[X]$

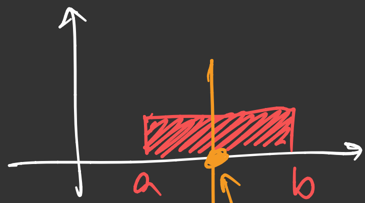
nuova variabile ottenuta come somme di due variabili

VERO SE X_1, X_2 SONO INDIPENDENTI.

Calcoliamo $\mathbb{E}[X]$, $\mathrm{Var}[X]$ per la distribuzione uniforme:

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} x f_X(x)\,dx = \int_a^b x \cdot \left(\frac{1}{b-a}\right) dx = \frac{1}{b-a}\left[\frac{x^2}{2}\right]_a^b$$

$$= \frac{b^2 - a^2}{(b-a)\,2} = \frac{\cancel{(b-a)}(b+a)}{\cancel{(b-a)}\;2} = \frac{b+a}{2}$$



$$\frac{a+b}{2} = \mathbb{E}[X]$$

$$\mathrm{Var}[X] = \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 f_X(x)\, dx$$

$$= \int_a^b \left(x - \frac{a+b}{2}\right)^2 \cdot \frac{dx}{(b-a)}$$

$$= \frac{1}{(b-a)} \int_a^b x^2 + \frac{(a+b)^2}{4} - (a+b)x$$

$$\boxed{\begin{aligned} \int x^2 &= \frac{x^3}{3} \\ \int x &= \frac{x^2}{2} \\ \int 1 &= x \end{aligned}}$$

$$= \frac{1}{(b-a)} \left[\frac{b^3 - a^3}{3} + \frac{(b+a)^2(b-a)}{4} - \frac{(b+a)(b^2-a^2)}{2} \right]$$

$$= \frac{(b-a)^2}{12} \quad \Rightarrow \quad \sigma_X = \frac{b-a}{2\cdot\sqrt{3}} \quad ✓$$

§ III.4: DISTRIBUZIONE ESPONENZIALE.

Studiamo il tempo che intercorre fra due eventi consecutivi nella distribuzione di Poisson.

[Binomiale $\rightsquigarrow$ Poisson $\rightsquigarrow$ Esponenziale]

$\boxed{\text{OSS}}$ Ogni evento (arrivo di un cliente) è indipendente. Quindi possiamo cominciare a misurare nel momento in cui accade un evento (arriva qualcuno) e attendere il prossimo.

Sia $\lambda := \#$ medio di arrivi nell'intervallo unitario

Sia $Y(t) := \#$ eventi in $[0,t] = \text{Poisson}(\mu = t \cdot \lambda)$

è una variabile discreta $\forall t$, quindi $\mathbb{P}(Y(t) = k) \neq 0$.

sempre.

$$\mathbb{P}(Y(t)=k) = \frac{(\lambda t)^k}{k!} \cdot e^{-\lambda t}$$

$X :=$ tempo al prossimo arrivo [X è v.a.c.]

$$\underline{\mathbb{P}(X>t)} = \underline{\mathbb{P}(Y(t)=0)} = \frac{(\lambda t)^0}{0!} e^{-\lambda t} = e^{-\lambda t}$$

$0! = 1$.

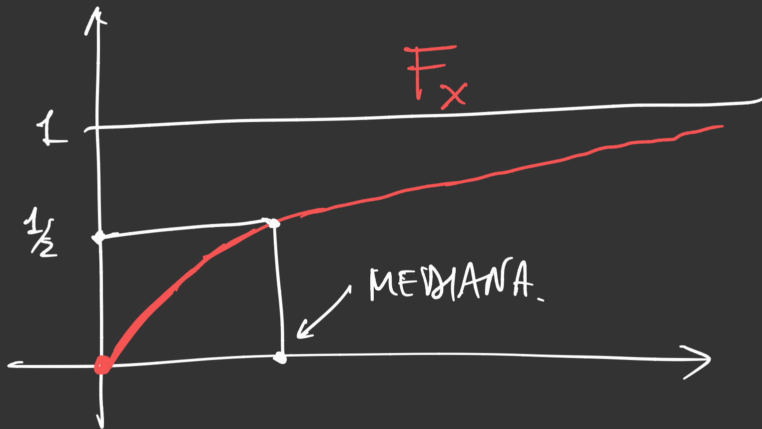
non è arrivato nessuno fino al tempo t

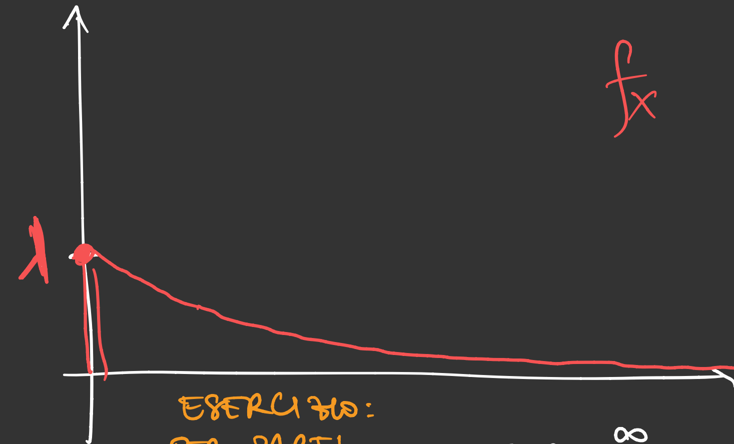
$$F_X(t) = \mathbb{P}(X \le t) = 1 - \mathbb{P}(X>t)$$

$$\Rightarrow F_X(t) = \begin{cases} 1 - e^{-\lambda t} & t>0 \\ 0 & t \le 0 \end{cases}$$

prendo la derivata in t

$$\Rightarrow f_X = \begin{cases} \lambda e^{-\lambda x} & x>0 \\ 0 & x \le 0 \end{cases}$$





ESERCIZIO: PER PARTI

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x \cdot f_X(x)\, dx = \int_0^{\infty} x \lambda e^{-\lambda x} \overset{\downarrow}{=} \underbrace{\left[-x e^{-\lambda x} \right]_0^{+\infty}}_{= 0 - 0 = 0} - \underbrace{\int_0^{\infty} e^{-x\lambda} dx}_{= 1/\lambda}$$

$$= \frac{1}{\lambda}$$

Cosa significa che il valor medio è $1/\lambda$?

Facile: se arrivano λ clienti nell'intervallo di tempo, mediamente, ne arriverà uno ogni $\frac{1}{\lambda}$.

$$Var[X] = \int_{-\infty}^{\infty} (x - \mathbb{E}[X])^2 f_X(x)\,dx = \int_0^{\infty} \underbrace{\left(x - \frac{1}{\lambda}\right)^2}_{= x^2 + \frac{1}{\lambda^2} - \frac{2x}{\lambda}} \lambda e^{-\lambda x}\,dx$$

ogni integrale è stato integrato per parti

$$= \underbrace{\int_0^{\infty} x^2 \lambda e^{-\lambda x}\,dx}_{= 0 - 0} - \underbrace{\frac{2}{\lambda}\int_0^{\infty} \lambda x e^{-\lambda x}\,dx}_{= \frac{1}{\lambda}} \underbrace{- \frac{2}{\lambda^2} + \frac{1}{\lambda^2}}_{= -\frac{1}{\lambda^2}}$$

$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \implies \sigma_X = \frac{1}{\lambda}\ !$$

Nella distribuzione esponenziale

$$\boxed{\sigma_X = \mathbb{E}[X]}\ !$$

Invece nella Poisson $Var[X] = \mathbb{E}[X]$.