ESAME 06/06/2025

1) a) Data X v.a. discreta con $P(X=2)=P(X=-1)=\frac{1}{2}$, allora $E[X^2]=\frac{5}{2}$

Vero, $E[X^2]=\sum a_k^2 \cdot p(X=a_k) = 2^2\cdot\frac{1}{2}+(-1)^2\cdot\frac{1}{2}=2+\frac{1}{2}=\frac{5}{2}$

b) La mediana dei dati $(x_1,\dots,x_n)$ è il valore più frequente nei dati

Falso, la moda è il valore più frequente, la mediana è il dato x_i t.c. metà degli altri valori è $\leq x_i$ e l'altra metà è $\geq x_i$

c) Se X è una variabile con densità f (p.es. v.a. espon. o gauss.) allora $f(x)=P(X=x)$ $\forall x\in\mathbb{R}$

Falso, se X ha densità allora $P(X=x)=0$ $\forall x$ (è continua)

d) Dati A, A', B eventi con $P(B)>0$, allora

$P(A\cup A'|B)=P(A|B)+P(A'|B)-P(A\cap A'|B)$

Vero, perché $P(A\cup B)=P(A)+P(B)-P(A\cap B)$ e qua sta venendo applicata la medesima proprietà

e) Dato un campione $X_1, X_2, \dots,$ uno stimatore corretto di un parametro θ è anche consistente (per la precisione, una famiglia di stimatori corretti $(U_n)_n$ di θ è anche consistente)

Falso, per essere corretto $\hat{\theta}=\theta$, mentre per essere consistente $\hat{\theta}\xrightarrow[n\to\infty]{P}\theta$. Quindi, per affermare che sono consistenti, non è suffic. sapere che sono corretti.

f) Siano (x_i,y_i), $i=1,\dots,n$, i dati relativi a peso x_i in kg e altezza y_i in m (metri) di n individui. Sia $\tilde{y}_i=100\cdot y_i$ le altezze espresse in cm (centimetri). La retta di regressione $\tilde{y}=a+bx$ dei dati $(x_i,\tilde{y}_i)$ è la stessa di $y=a+bx$ dei dati (x_i,y_i)?

Falso, ad esempio il coefficiente angolare $\tilde{b}$ dei dati $(x_i,\tilde{y}_i)$ e quello b dei dati (x_i,y_i) sono legati dalla relazione:

$$\tilde{b}=\frac{cov(x,\tilde{y}_i)}{var(x)}=\frac{100\cdot cov(x,y)}{var(x)}=100\cdot b$$

2) Il numero di domande processate da un server in un minuto di un giorno feriale segue una distribuzione di Poisson di parametro 1. Supponiamo che il num di domande processate in un un dato minuto sia indipendente dal num di domande processate in altri momenti.

a) Calcolare la probabilità che, in 2 dati minuti consecutivi (di un gg feriale), il server processi almeno 2 domande.

$X =$ "# domande processate in 1 min" $\sim Poiss(\lambda)$

$Y = X_1 + X_2 = 2X \sim Poiss(1+1) = Poiss(2)$

$$P\{Y \geq 2\} = 1 - P\{Y=0\} - P\{Y=1\} = 1 - \frac{2^0 \cdot e^{-2}}{0!} - \frac{2^1 \cdot e^{-2}}{1!} =$$

$$= 1 - e^{-2} - 2e^{-2} = 1 - 3e^{-2} \approx 0,594$$

b) Calcolare approssimativamente la prob. che in un dato gg (cioè 1440 min) feriale, il server processi almeno 1400 domande

$Y = X_1 + \dots + X_{1440} = 1440X \sim Poiss(1440)$

Si può applicare TCL, in quanto n è relativamente grande

$E[Y] = 1440 = m \qquad Var(x) = 1440 = \sigma^2$

$$P\{Y \geq 1400\} = P\left\{\frac{Y-1440}{\sqrt{1440}} \geq \frac{1400-1440}{\sqrt{1440}}\right\} \approx P\left\{Z \geq \frac{-40}{37,947}\right\}$$

$$= P\{Z \geq -1.05\} = P\{Z \leq 1.05\} = \Phi(1.05) = 0,8531$$

In un gg festivo il num di domande processate dallo stesso server segue una distribuzione di Poisson di parametro 2.

c) Individuare approssimativamente la distribuzione del num complessivo di domande tra domenica (festivo) e lunedì (feriale)

$X =$ "# domande processate in un min in un gg feriale" $\sim P(1)$

$Y =$ "# " " " " festivo" $\sim P(2)$

$A = X_1 + \dots + X_{1440} = 1440X \sim P(1 \cdot 1440) = P(1440)$

$B = Y_1 + \dots + Y_{1440} = 1440Y \sim P(2 \cdot 1440) = P(2880)$

Quindi, essendo X e Y indipendenti lo è anche la loro somma

e si ha $Poiss(1440+2880) = Poiss(4320) \approx N(4320, 4320)$

3) Sia $X_1, \dots, X_n$ un campione i.i.d. di densità

$$f_\theta(x) = e^{\theta - x} \mathbb{1}_{[\theta, +\infty)}(x) = \begin{cases} e^\theta \cdot e^{-x} & \text{se } x \geq \theta \\ 0 & \text{se } x < \theta \end{cases} \quad \text{con } \theta > 0$$

a) verificare che la funzione di ripartizione F_θ delle X_i è

$$F_\theta(x) = \begin{cases} 1 - e^{\theta - x} & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

$$\int_{-\infty}^{x} f_\theta(t)dt = \int_{-\infty}^{\theta} f_\theta(t)dt + \int_{\theta}^{x} f_\theta(t)dt = \int_{-\infty}^{\theta} 0\, dt + \int_{\theta}^{x} e^{\theta - t} dt$$

$$= \int_{\theta}^{x} e^{\theta} \cdot e^{-t} dt = \underbrace{e^{\theta}}_{\text{costante rispetto a } t} \int_{\theta}^{x} e^{-t} dt = e^{\theta} \left[-e^{-t} \right]_{\theta}^{x} = e^{\theta} (-e^{-x} + e^{-\theta})$$

$$= e^{\theta}(-e^{-x} + e^{-\theta}) = e^{\theta - \theta} - e^{\theta - x} = 1 - e^{\theta - x} \quad \text{ok!}$$

b) Calcolare $E[e^{aX_1}]$ al variare di $a > 0$; per quali valori di $a > 0$, $E[e^{aX_1}]$ è finito?

$$E[e^{aX_1}] = \int_{-\infty}^{+\infty} e^{ax} \cdot f_\theta(x)dx = \int_{\theta}^{+\infty} e^{ax} \cdot e^{\theta - x} dx = \int_{\theta}^{+\infty} e^{ax} \cdot e^{\theta} \cdot e^{-x} dx$$

$$= e^{\theta} \int_{\theta}^{+\infty} e^{ax} \cdot e^{-x} dx = e^{\theta} \int_{\theta}^{+\infty} e^{x(a-1)} dx = e^{\theta} \left[\frac{1}{(a-1)} \cdot e^{x(a-1)} \right]_{\theta}^{+\infty}$$

$$= e^{\theta} \left(\frac{1}{(a-1)} \cdot e^{+\infty(a-1)} - \left(\frac{1}{(a-1)} \cdot e^{\theta(\theta - 1)} \right) \right)$$

È finito se e solo se $a < 1$, altrimenti per $e^{+\infty(a-1)}$ per $a \geq 1$ risulterebbe $+\infty$

c) Calcolare, se esiste, lo stimatore di massima verosimiglianza di θ

La funzione $f_\theta(x) = e^{\theta - x} \mathbb{1}_{(\theta, +\infty)}(x)$ è positiva $\forall x \in (\theta, +\infty)$ in quanto è noto dal testo che $\theta > 0$. Il massimo della funzione si ha quindi quando $x = \min_i X_i$, valore oltre il quale la funzione decresce. Quindi $\hat{\theta} = \min_{i=1,\dots,n} X_i$

ESAME 15/05/2025

1) a) Dato un campione i.i.d. $X_1,\dots,X_n$ (con momento primo finito), un intervallo di fiducia per il valore atteso delle X_i, di livello 95% è un intervallo I tale che la media campionaria $\bar{X}_n$ delle X_i cade in I con prob. almeno 95%

Falso, non significa che il parametro (media campionaria) cade nell'intervallo con probabilità del 95% ma che se ripetessimo l'esperimento molte volte, nel 95% dei casi il parametro cadrebbe nell'intervallo di fiduco.

b) Se X e Y sono due v.a. Bernoulli di parametro $p\in(0,1)$ allora $X+Y$ è v.a. binomiale di parametri $2,p$

Falso, ponendo che $X=Y$, risulterebbe che $X+Y=2X$, quindi se fosse una variabile di Bernoulli si potrebbe ottenere come risultato solamente 0 o 2, mentre col binomiale anche 1

c) Se due v.a. (discrete o con densità) hanno la stessa legge, allora $E[\varphi(X)]=E[\varphi(Y)]$ per ogni funzione $\varphi:\mathbb{R}\to\mathbb{R}$ limitata

Vero, nel caso discreto $E[\varphi(X)]=\sum_x \varphi(x)p(x)$ e la funzione di massa p dipende solo dalla legge. Nel caso con densità è analogo $E[\varphi(X)]=\int_{\mathbb{R}}\varphi(x)f(x)\,dx$.

d) Se $E[X]=0$, allora X è nulla

Falso, ad esempio nel caso continuo $E[X]=\int_{\mathbb{R}} x\cdot f(x)\,dx$, potrebbe essere nulla la $f(x)$

e) Dato un campione i.i.d. $X_1,\dots,X_n$ (con momento primo finito), $\frac{X_1+X_2}{2}$ è uno stimatore corretto di $E[X_i]$

Vero, per la linearità si ha che $E[(X_1+X_2)/2]=(E[X_1]+E[X_2])/2=E[X_1]$

f) Se $(x_1,y_1),\dots,(x_n,y_n)$ sono dati bivariati, il loro coefficiente di correlazione "r" vale ±1 se e solo se i dati (x_i,y_i) sono allineati (cioè disposti sulla stessa retta)

Vero, il coefficiente di correlazione misura quanto i punti sono vicini a una retta. Per essere $r=1$, i dati devono necessariamente trovarsi perfettamente allineati sulla retta.

2) Aldo e Bruno giocano a carte. In ciascun turno, Aldo estrae due carte distinte da un mazzo di carte napoletano (4 semi, 10 carte/seme). Aldo vince se almeno una carta è del seme "denari".

a) Calcolare la probabilità di vittoria di Aldo in un singolo turno

Sia S l'insieme delle carte

$$\Rightarrow \Omega = \{a, b \in S \mid a \neq b\} \qquad \#\Omega = 40 \cdot 39$$

$$P(\geq 1 \text{ denari}) = 1 - P(\text{nessun denari}) = 1 - \frac{30 \cdot 29}{40 \cdot 39} = 0{,}442$$

*non si può usare da subito il binomiale perché è senza reinserimento

b) Calcolare la probabilità che, su 5 turni, Aldo vinca almeno in due turni

$$P(\text{Aldo vince}) = P(\geq 1 \text{ denari}) = 0{,}442 = p$$

$$X = \text{"\# partite vinte da Aldo"} \sim B(5, 0.442)$$

$$P\{X \geq 2\} = 1 - P\{X = 0\} - P\{X = 1\} = 1 - \binom{5}{0} p^0 (1-p)^{5-0} - \binom{5}{1} p^1 (1-p)^{5-1}$$

$$= 1 - (1-p)^5 - 5p(1-p)^4 = 1 - (0{,}558)^5 - 5(0{,}442)(0{,}558)^4 \simeq 0{,}732$$

Il mazzo è stato scelto a caso tra 4 mazzi. Aldo e Bruno vengono però a sapere che uno di questi 4 mazzi è truccato, in modo che tutte le carte di quel mazzo siano di denari.

c) Se Aldo vince in almeno 2 turni su 5, prob. che il mazzo scelto sia truccato?

X = "Aldo vince due turni" $\qquad A$ = "mazzo truccato"

$$P(A) = \frac{1}{4} \qquad P(A \mid X) = ?$$

$$P(X \mid A^c) = 0{,}732 \qquad P(X \mid A) = 100\% \quad \left(\text{se truccato vince sicuramente}\right)$$

$$P(A \mid X) = \frac{P(X \mid A) \cdot P(A)}{P(X)} = \frac{P(X \mid A) \cdot P(A)}{P(X \mid A^c) \cdot P(A^c) + P(X \mid A) \cdot P(A)} =$$

$$= \frac{1 \cdot 1/4}{0{,}732 \cdot \frac{3}{4} + 1 \cdot \frac{1}{4}} \simeq 0{,}313$$

3) Sia $X_1, \dots, X_n$ un campione i.i.d. di densità

$$f_\theta(x) = \theta x^{\theta-1} \mathbb{1}_{(0,1)}(x) = \begin{cases} \theta x^{\theta-1} & \text{se } x \in (0,1) \\ 0 & \text{altrimenti} \end{cases} \quad \text{con } \theta > 0$$

a) Calcolare in funzione di θ il valore atteso $E[X_i]$ e $Var(X_i)$

$$E[X] = \int_{-\infty}^{+\infty} x \cdot f_\theta(x)\,dx = \int_0^1 x \cdot \theta x^{\theta-1}\,dx = \theta \int_0^1 x^\theta\,dx = \theta \left[\frac{x^{\theta+1}}{\theta+1} \right]_0^1$$

$$= \theta \left(\frac{1^{\theta+1}}{\theta+1} - \frac{0^{\theta+1}}{\theta+1} \right) = \theta \cdot \frac{1}{\theta+1} = \frac{\theta}{\theta+1}$$

$$E[X^2] = \int_{-\infty}^{+\infty} x^2 \cdot f_\theta(x)\,dx = \dots = \theta \left[\frac{x^{\theta+2}}{\theta+2} \right] = \frac{\theta}{\theta+2}$$

$$Var(X) = E[X^2] - E[X]^2 = \frac{\theta}{\theta+2} - \frac{\theta^2}{(\theta+1)^2}$$

b) Calcolare in funzione di θ la densità $Y_i = -\log X_i$, se esiste

$$h(x) = -\log x \quad \text{con } h: (0, +\infty) \to \mathbb{R}$$

$$y = -\log x \iff -y = \log x \iff e^{-y} = x \iff x = e^{-y}$$

$\log_b A = C \iff b^C = A$

$$h^{-1}(y) = e^{-y} \qquad \frac{d h^{-1}(y)}{dy} = \frac{d(e^{-y})}{dy} = e^{-y} \cdot (-1) = -e^{-y}$$

composta

$$f_Y(y) = f_X(h^{-1}(y)) \cdot \left| \frac{d h^{-1}}{dy} \right| \mathbb{1}_{(0,+\infty)}(y) = \theta (e^{-y})^{\theta-1} \cdot |-e^{-y}| \mathbb{1}_{(0,+\infty)}(y)$$

$$= \theta e^{-y\theta+y} \cdot e^{-y} \mathbb{1}_{(0,+\infty)}(y) = \theta e^{-y\theta+y-y} \mathbb{1}_{(0,+\infty)}(y) = \theta e^{-y\theta} \mathbb{1}_{(0,+\infty)}(y)$$

c) Calcolare, se esiste, lo stimatore di mas. ver. (MLE) di θ

$$L(\theta; x_1, \dots, x_n) = \underbrace{P_\theta(x_1) \cdot \dots \cdot P_\theta(x_n)}_{n} = \theta x_1^{\theta-1} \cdot \dots \cdot \theta x_n^{\theta-1} = \theta^n \cdot \prod_{i=0} x_i^{\theta-1}$$

$$\text{Log}(L) = \log(\theta^n \cdot \prod x_i^{\theta-1}) = \log(\theta^n) + \log(\prod x_i^{\theta-1})$$

$$= n \log\theta + \sum_{i=0}^{n} \log(x_i^{\theta-1}) = n\log\theta + \sum^{n} (\theta-1)\log x_i$$

$$\frac{d\,\text{Log}(L)}{d\theta} = \frac{d(n\log\theta + \sum(\theta-1)\log x_i)}{d\theta} = \frac{n}{\theta} + \frac{d((\theta-1)\sum \log x_i)}{d\theta} = \frac{n}{\theta} + \sum \log x_i$$

costante

$$\frac{n}{\theta} + \sum \log x_i = 0 \iff \frac{n}{\theta} = -\sum \log x_i \iff \theta = -\frac{n}{\sum \log x_i} = \hat{\theta}$$

5.2) Overbooking: 180 posti, venduti 195, su 10 passeggeri mediamente se ne presentano 9

a) Probabilità di Overbooking?

X = "# passeggeri che si presentano" $\sim B(195, 0.9)$

$P\{\text{Overbooking}\} = P\{X \geq 181\}$

Si può applicare il TCL perché n è grande

$E[X] = n \cdot p = 175.5 = m$ $\quad Var(X) = 175{,}5 \cdot 0{,}10 = 17{,}55 = \sigma^2$

$$P\{X \geq 181\} = \left\{\frac{X - 175.5}{\sqrt{17{,}55}} \geq \frac{181 - 175.5}{\sqrt{17{,}55}}\right\} \approx \left\{Z \geq \frac{5{,}5}{4{,}189}\right\} = \{Z \geq 1{,}31\}$$

$$= 1 - \Phi(1.31) = 1 - 0.9049 = 0.0951$$

b) Numero di biglietti da vendere t.c. $P\{\text{overbooking}\} \leq 0{,}01$

$$P\{X \geq 181\} = \left\{\frac{X - n \cdot 0{,}9}{\sqrt{n \cdot 0{,}9 \cdot 0{,}1}} \geq \frac{181 - n \cdot 0{,}9}{\sqrt{n \cdot 0{,}9 \cdot 0{,}1}}\right\} \approx P\left\{Z \geq \frac{181 - n \cdot 0{,}9}{\sqrt{n \cdot 0{,}09}}\right\}$$

$$= 1 - \Phi\left(\frac{181 - n \cdot 0{,}9}{\sqrt{n \cdot 0{,}09}}\right) = 0.01$$

$$\Phi\left(\frac{181 - n \cdot 0{,}9}{\sqrt{n \cdot 0{,}09}}\right) = 1 - 0{,}01 = 0{,}99 \implies \frac{181 - n \cdot 0{,}9}{\sqrt{n \cdot 0{,}99}} = q_{0.99} = 2.33$$

$$\frac{181 - n \cdot 0{,}9}{\sqrt{n \cdot 0{,}09}} = 2.33 \iff 181 - n \cdot 0{,}9 = 2{,}33 \cdot \sqrt{n \cdot 0{,}09}$$

$$\iff (181 - n \cdot 0{,}9)^2 = (2{,}33)^2 \cdot 0{,}09n$$

$$\iff (181)^2 + (0{,}9)^2 n^2 - 2 \cdot 0{,}9(181)n = 5{,}429 \cdot 0{,}09n$$

$$\iff n^2(0{,}9)^2 + n(-0{,}9 \cdot 181 - 5{,}429 \cdot 0{,}09) + (181)^2$$

$$\iff n^2(0{,}81) + n(-163{,}39) + 32761$$

$$n = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \begin{cases} 223{,}84 \implies \text{non accettabile, è maggiore del valore al punto a} \\ 180{,}69 \quad \text{ok!} \end{cases}$$

$n = 180$ (arrotondo per difetto perché $\leq$)

BINOMIALE

5.3) Altezza $\sim N(177.8,\ 10.6^2)$. Probabilità di media di altezza di due estratti > 180?

X_i = "altezza i-esimo individuo estratto" $\sim N(177.8,\ 10.6^2)$

$X_i \sim N(m_i, \sigma_i^2)$ indip. $\Longrightarrow X_1 + X_2 \sim N(m_1 + m_2, \sigma_1^2 + \sigma_2^2)$

$\bar{X}_2 = \frac{1}{2}(X_1 + X_2) \sim N(177.8,\ \frac{1}{2}\cdot(10.6)^2) \longrightarrow E[\bar{X}_n] = E[X_i],\ Var(\bar{X}_n) = \frac{1}{n}\cdot Var(X_i)$

$P\{\bar{X}_2 > 180\} = P\left\{\frac{X - 177.8}{\sqrt{\frac{1}{2}}\cdot 10{,}6} \geq \frac{180 - 177.8}{\sqrt{1/2}\cdot 10{,}6}\right\} = P\{Z > 0{,}29\} = 1 - \Phi(0{,}29)$

$= 1 - 0{,}6141 = 3859$

$\to$ > 180 e non ≥ 181 perché ≥ 181 non comprenderebbe 180.1, 180.2, ...
era comunque specificato nel testo "> 180"

Con 100 estratti sarebbe $\bar{X}_{100} \sim N(177.8,\ \frac{1}{100}\cdot(10.6)^2)$

5.4) Server riceve email di dimensioni: medie 4,73MB e deviazione standard 0,53 MB.

a) Su 70 email, prob. di dim. totale > 340 MB?

X_i = "dimensione della i-esima email" $i = 1, \dots, 70 \sim N(4.73MB, (0.53MB)^2)$

Y = "dim. totale email" $= \sum_{i=0}^{70} X_i = 70\cdot\bar{X}_{70} \sim N(4{,}73MB\cdot 70, (0.53MB)^2\cdot 70) = (331.1,\ 19.663)$

$n = 70$ è abbastanza grande, uso TCL

$P\{Y > 340\} = \left\{\frac{Y - 331.1}{\sqrt{19.663}} > \frac{340 - 331.1}{\sqrt{19.663}}\right\} \approx \{Z > 2.01\} = 1 - \Phi(2.01)$

$= 1 - 0.9778 = 0.0222$

b) Y (in MB) t.c. dim. totale $\leq y$ con prob. 95%?

$P\{Y \leq y\} \approx \left\{Z \leq \frac{y - 331.1}{\sqrt{19.663}}\right\} = \Phi\left(\frac{y - 331.1}{\sqrt{19.663}}\right) = 0.95$

$\Longrightarrow \frac{y - 331.1}{\sqrt{19.663}} = q_{0.95} = 1.64$

$y - 331.1 = 1.64\cdot\sqrt{19.663} \iff y = 331 + 7.27 \iff y = 338.37$

GAUSSIANA

5.5) Max dei valori di due lanci di dado equilibrato: legge?

X_i = "esito dell'i-esimo lancio" $P\{X_i = k\} = \frac{1}{6} \ \forall k \in \{1,2,\dots,6\} \ \forall i$

$X = \max\{X_1, X_2\}$

$P\{X \leq k\} = P\{X_1 \leq k, X_2 \leq k\} = P\{X_1 \leq k\} \cdot P\{X_2 \leq k\} = \frac{1}{6}k \cdot \frac{1}{6}k = \frac{1}{36}k^2$

$P\{X = k\} = P\{X \leq k\} - P\{X \leq k-1\} = \frac{1}{36}(k^2 - (k-1)^2) = \frac{1}{36}(2k-1)$

$E[X] = \sum_{k=1}^{6} (2k-1)\frac{1}{36} \cdot k = \frac{161}{36}$

5.6) 42 Studenti di cui 13 mancini, divisi in gruppo (a) e (b), ciascuno di 21 elementi, con sorteggio

a) Funzione di massa di X e Y?

X = "# mancini in (a)" Y = "numero mancini in b"

$X = Y \sim IP(42, 13, 21)$

$P\{X = k\} = \frac{\binom{13}{k}\binom{29}{21-k}}{\binom{42}{21}} \implies$ Y ha la stessa f. di massa

b) X e Y indipendenti?

Non possono essere indipendenti perché $X+Y=13$, perciò $P\{X=1, Y=1\} = 0$ ma $P\{X=1\} \cdot P\{Y=1\} > 0$, non indipendenti

c) Calcolare E[X] e E[Y] senza usare le leggi di X e Y

Poiché X e Y sono equidistribuite, $E[X] = E[Y]$ e $X+Y=13$

Quindi $E[X] + E[Y] = E[X+Y] = E[13] = 13$

Quindi $2E[X] = 13$ e perciò $E[X] = \frac{13}{2} = 6{,}5$

d) Calcolare E[X] usando le leggi ma con pochi conti

$X_i = \begin{cases} 1 & \text{se i-esimo individuo estratto di (a) è mancino} \\ 0 & \text{altrimenti} \end{cases}$

$X = \sum_{i=0}^{21} X_i$

$E[X_i] = P\{\text{estrarre mancino}\} = \frac{13}{42}$

$E[X] = \sum_{i=0}^{21} E[X_i] = 21 \cdot \frac{13}{42} = \frac{13}{2}$

5.8) N° di clienti a uno sportello in un gg $\sim P(3)$, indipendente dal numero di clienti in altri gg. Stimare, con disuguaglianza di Chebyshev, la prob. di almeno 20 clienti in una settimana.

$X =$ "# clienti in un gg" $\sim Poiss(3)$

Settimana lavorativa = 5gg, perciò $Y = X_1 + \dots + X_5 \sim P(3 \cdot 5) = P(15)$

$E[Y] = \lambda = 15 \quad Var(Y) = \lambda = 15$

$P\{Y \geq 20\} = P\{X - 15 \geq 20 - 15\} = P\{X - 15 \geq 5\}$

$P\{|X - E[X]| \geq d\} \leq \frac{1}{d^2} \cdot Var(X) \Longrightarrow P\{|X - 15| \geq 5\} \leq \frac{1}{(20-15)^2} \cdot 15$

$\Longleftrightarrow P\{|X - 15| \geq 5\} \leq \frac{1}{25} \cdot 15 = \frac{15}{25} \approx 0,6$

5.7) Tempo di vita di un componente $\sim Exp(0.5)$

a) prob. di guasto in un anno?

$X =$ "tempo di vita" $\sim Exp(\frac{1}{2})$

$P\{X \leq 1\} = \int_{-\infty}^{1} f(x)dx = \int_0^1 \frac{1}{2} e^{-\frac{1}{2}x} dx = \frac{1}{2} \int_0^1 e^{-\frac{1}{2}x} dx = \frac{1}{2} \left[-2e^{-\frac{1}{2}x} \right]_0^1 \approx 0,393$

Se un componente si guasta, viene sostituito da un altro.

b) Media e varianza del tempo totale di vita di 2 componenti?

$Y = X =$ "tempo di vita" $\sim Exp(\frac{1}{2})$

$E[X] = E[Y] = \frac{1}{1/2} = 2 \qquad Var(X) = Var(Y) = \frac{1}{(1/2)^2} = 4$

X e Y sono indipendenti $\Longrightarrow V = X + Y =$ "tempo di vita totale"

$E[V] = E[X] + E[Y] = 2 + 2 = 4 \qquad Var(V) = Var(X) + Var(Y) = 4 + 4 = 8$

c) Prob. che il secondo componente si guasti entro 4 anni dall'installazione del 1° componente?

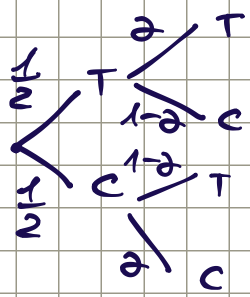
V ha densità

$f_V(v) = \int_{-\infty}^{+\infty} f_X(x) \cdot f_Y(v-x) dx = \int_{-\infty}^{+\infty} \frac{1}{2} e^{-\frac{x}{2}} \mathbb{1}_{(0,+\infty)}(x) \cdot \frac{1}{2} e^{-\frac{(v-x)}{2}} \mathbb{1}_{(0,+\infty)}(v-x) dx$

- $v \leq 0 \quad f_V(v) = 0$
- $v > 0 \quad f_V(v) = \int_0^v \frac{1}{4} e^{-\frac{v}{2}} dx = \frac{1}{4} v e^{-\frac{v}{2}}$

$P\{V \leq 4\} = \int_{-\infty}^{4} f_V(v) dv = \int_0^4 \frac{1}{4} v e^{-v/2} dv = 1 - 3e^{-2} \approx 0,594$

5.9) Lancio di una moneta equilibrata. Se esce testa, lanciamo una moneta con prob. a di testa. Se esce croce, lanciamo una moneta con prob. $1-a$ di testa.



a) legge congiunta delle X_i?

$$P\{X_1=T, X_2=T\} = P\{X_1=C, X_2=C\} = \frac{1}{2}a$$

$$P\{X_1=T, X_2=C\} = P\{X_1=C, X_2=T\} = \frac{1}{2}(1-a)$$

b) leggi marginali

$$P\{X_1=T\} = P\{X_1=C\} = \frac{1}{2}$$

$$P\{X_2=T\} = P\{X_2=C\} = \frac{1}{2}a + \frac{1}{2}(1-a) = \frac{1}{2}a + \frac{1}{2} - \frac{1}{2}a = \frac{1}{2}$$

c) Sono correlate?

X_i a valori finiti $\Longrightarrow$ ammettono momenti di ogni ordine

X_1 e X_2 sono correlati se $\rho = \frac{\mathrm{cov}(x,y)}{\sqrt{\mathrm{var}(x)\mathrm{var}(y)}} = 0$

$$\mathrm{Cov}(X_1, X_2) = E[X_1X_2] - E[X_1]E[X_2]$$

$$E[X_1] = \frac{1}{2} \quad E[X_2] = \frac{1}{2}$$

$$E[X_1X_2] = (\underset{C}{0}\cdot 0)P(X_1=0, X_2=0) + (0\cdot 1)P(X_1=0, X_2=1) + (1\cdot 0)P(X_1=0, X_2=1) + (1\cdot 1)P(X_1=1, X_2=1) = \frac{1}{2}\overset{T}{a}$$

$$\mathrm{Cov}(X_1,X_2) = \frac{1}{2}a - \frac{1}{2}\cdot\frac{1}{2} = \frac{a}{2} - \frac{1}{4} = \frac{2a-1}{4}$$

$$\mathrm{var}(X_1,X_2) = E[X^2] - E[X]^2 = \frac{1}{2} - \frac{1}{4} = \frac{2-1}{4} = \frac{1}{4}$$

$$\rho(X_1, X_2) = \frac{\frac{2a-1}{4}}{\frac{1}{4}} = 2a-1$$

d) Per quali a, X_1 e X_2 sono indipendenti?

Per essere indipendenti, devono essere anche scorrelate, perciò $\rho = 0$

$$2a-1=0 \iff 2a=1 \iff a=\frac{1}{2}$$

6.1) X v.a. con densità $f_X(x) = \theta x^{-\theta-1} \mathbb{1}_{(1,+\infty)}(x)$, $\theta > 0$

a) Verificare che f è una densità

Per essere una densità
- $f \geq 0$: ok $\forall x \in (1, +\infty)$
- $\int_{-\infty}^{+\infty} f(x)dx = 1$

$$\int_{-\infty}^{+\infty} f(x)dx = \int_1^{+\infty} \theta x^{-\theta-1} dx = \theta \int_1^{+\infty} x^{-\theta-1} dx = \theta \left[-\frac{x^{-\theta}}{\theta} \right]_1^{+\infty}$$

$$= \theta \left(-\frac{1}{+\infty^\theta} + \frac{1^{-\theta}}{\theta} \right) \underset{\theta>0}{=} \theta \left(0 + \frac{1}{\theta} \right) = 1 \quad \text{ok!}$$

b) Stimatore MLE di θ

$$L(\theta; x_1, \dots, x_n) = P_\theta(x_1) \cdot \dots \cdot P_\theta(x_n) = \theta x_1^{-\theta-1} \cdot \dots \cdot \theta x_n^{-\theta-1} = \theta^n \prod_{i=0}^{n} X_i^{-\theta-1}$$

$$\mathrm{Log}(L) = \log(\theta^n) + \log\left(\prod X_i^{-\theta-1}\right) = n \log\theta + \sum \log(x_i^{-\theta-1})$$

$$= n\log\theta + \sum(-\theta-1)\log X_i = n\log\theta - (\theta+1)\sum \log X_i$$

$$\frac{d\,\mathrm{Log}(L)}{d\theta} = \frac{d(n\log\theta - (\theta+1)\underbrace{\sum\log X_i}_{c})}{d\theta} = \frac{n}{\theta} - \sum\log X_i$$

$$\frac{n}{\theta} - \sum\log X_i = 0 \iff \frac{n}{\theta} = \sum\log X_i \iff \theta = \frac{n}{\sum\log X_i} = \hat{\theta}$$

c) Stimatore MLE corretto per $n=1$?

$\hat\theta$ è corretto se $\hat\theta = \theta$. Per $n=1$, $\hat\theta = \frac{1}{\log X_1}$

$$E\left[\frac{1}{\log X}\right] = \int_1^{+\infty} \frac{1}{\log x} \cdot \theta x^{-\theta-1} dx \qquad \boxed{\begin{array}{ll} t = \log x & dx = e^t dt \\ \log(1) = 0 & \log(+\infty) = +\infty \end{array}}$$

$$\int_0^{+\infty} \frac{1}{t} \cdot e^{-t\theta-t} \cdot e^t dt = \int_0^{+\infty} \frac{1}{t} \cdot e^{-\theta t} dt = +\infty \quad \underline{\text{non}} \text{ corretto}$$

d) Stimatore MLE consistente?

Per LGN, $\frac{1}{n}\sum\log X_i \longrightarrow E[\log X]$, quindi $\hat\theta \longrightarrow \frac{1}{E[\log X]}$

$$E[\log X] = \int_1^{+\infty} \log x \cdot \theta x^{-\theta-1} = \int_0^{+\infty} t \cdot \theta e^{-t(\theta+1)} \cdot e^t dt = \int_0^{+\infty} t \cdot \theta e^{-t\theta} dt$$

$$= \left[t \cdot \frac{e^{-t\theta}}{\theta} \right]_0^{+\infty} - \int_0^{+\infty} \theta e^{-t\theta} = \left[t \cdot \frac{e^{-t\theta}}{\theta} - \frac{\theta e^{-t\theta}}{\theta} \right]_0^{+\infty} = \left(0 - 0 - \left(0 - \frac{1}{\theta}\right)\right) = \frac{1}{\theta}$$