Calcolo delle Probabilità e Statistica, Ingegneria Civile e A&T e Informatica
I prova finale a.a. 2022/23

Durata della prova 2.5 h

Punteggi: **1:** $4+3+4$; **2 :** $4+4+4$; **3:** $3+4$.

1. Si consideri il vettore aleatorio (X,Y) a valori in $\{-2,0,1\} \times \{-1,0,1\}$ con densità congiunta:

$$p_{XY}(-2,-1) = 0; \;\; p_{XY}(-2,0) = \frac{2}{36}; \;\; p_{XY}(-2,1) = 0;$$
$$p_{XY}(0,-1) = \frac{4}{36}; \;\; p_{XY}(0,0) = \frac{2}{36}; \;\; p_{XY}(0,1) = 0;$$
$$p_{XY}(1,-1) = \frac{1}{36}; \;\; p_{XY}(1,0) = \frac{26}{36}; \;\; p_{XY}(1,1) = \frac{1}{36}.$$

(i) Si determinino le densità marginali p_X, p_Y e si calcoli la $cov(X,Y)$. Le variabili aleatorie X ed Y sono indipendenti?
(ii) Si calcoli la probabilità dell'evento $\{X - |Y| = -1\}$.
(iii)Si calcoli la densità della variabile aleatoria $Z = X + Y$.

2. Si consideri il vettore aleatorio (X,Y) con densità congiunta:

$$f(x,y) = \begin{cases} 2e^{-(x+2y)} & \text{se } x \geq 0 \text{ e } y \geq 0 \\ 0 & \text{altrimenti} \end{cases}$$

(i) Trovare le densità marginali di X e Y, e calcolare $cov(X,Y)$. Dire se X e Y sono indipendenti.
(ii) Calcolare $P(X < Y)$ e $P(X + Y < 2 | X > 1)$.
(iii) Trovare la densità congiunta del vettore aleatorio $(U,V) = (X+Y, X-Y)$.
Le v.a. U e V sono indipendenti? In caso contrario, calcolare $cov(U,V)$.

3. (i) Un campione di 500 luci al led viene estratto da una grossa fornitura ed esaminato per rilevare eventuali difetti. Supponiamo che 350 pezzi superino il controllo.
Calcolare un intervallo di confidenza a livello $1 - \alpha = 0.95$ per la percentuale p di luci al led della fornitura accettabili.
(ii) Sia data una successione $(X_n)_{n\geq 1}$ di variabili aleatorie indipendenti identicamente distribuite con legge di Bernoulli di parametro $p = 0.7$ e sia $\overline{X}_n = \frac{X_1 + \ldots + X_n}{n}$. Stimare quanto deve essere grande n affinché

$$P(|\overline{X}_n - p| \leq 0.1) \geq 0.95.$$

Soluzioni della I prova finale a.a. 2022/23

(i) Si ha:

$$P(X=i)=\sum_{j\in\{-1,0,1\}} P(X=i,Y=j),$$

per cui:

$$P(X=-2)=2/36,\ P(X=0)=4/36+2/36=6/36,$$

$$P(X=1)=1/36+26/36+1/36=28/36.$$

Analogamente:

$$P(Y=j)=\sum_{i\in\{-2,0,1\}} P(X=i,Y=j),$$

per cui:

$$P(Y=-1)=4/36+1/36=5/36,\ P(Y=0)=2/36+2/36+26/36=30/36,$$

$$P(Y=1)=1/36.$$

Siccome $0=p_{XY}(-2,-1)=\neq P(X=-2)P(Y=-1)=\frac{6}{36}\cdot\frac{30}{36}$, le v.a. X e Y non sono indipendenti. Si ha inoltre:

$$E(X)=\sum_{i\in\{-2,0,1\}} i\cdot P(X=i)=(-2)\cdot\frac{2}{36}+0\cdot\frac{6}{36}+1\frac{28}{36}=\frac{24}{36}=\frac{2}{3}\ .$$

$$E(Y)=\sum_{j\in\{-1,0,1\}} j\cdot P(Y=j)=-1\cdot\frac{5}{36}+0\cdot\frac{30}{36}+1\cdot\frac{1}{36}=-\frac{4}{36}=-\frac{1}{9}$$

$$E(XY)=\sum_{(i,j)\in\{-2,0,1\}\times\{-1,0,1\}} ij\cdot p_{XY}(i,j)=(-1)\cdot\frac{1}{36}+1\cdot\frac{1}{36}=0\ ,$$

da cui

$$Cov(X,Y)=E(X)E(Y)=\frac{2}{3}\cdot-\frac{1}{9}=-\frac{2}{27}\ .$$

(ii) Scriviamo l'evento $\{X-|Y|=-1\}$ come unione di eventi elementari.

$$\{X-|Y|=-1\}=\{X=-2,Y=-1\}\cup\{X=-2,Y=1\}.$$

e quindi

$$P(X-|Y|=-1)=P(X=-2,Y=-1)+P(X=-2,Y=1)=0.$$

(iii) Osserviamo che $Z\in\{-3,-2,-1,0,1,2\}$. Si ha:

$$P(Z=-3)=P(X=-2,Y=-1)=0,\ P(Z=-2)=P(X=-2,Y=0)=2/36,$$

$$P(Z=-1) = P(X=-2, Y=1) + P(X=0, Y=-1) = 0 + 4/36 = 1/9,$$

$$P(Z=0) = P(X=0, Y=0) + P(X=1, Y=-1) = 2/36 + 1/36 = 3/36,$$

$$P(Z=1) = P(X=0, Y=1) + P(X=1, Y=0) = 26/36, \; P(Z=2) = P(X=1, Y=1) = 1/36. \blacksquare$$

2. (i) Si ha, per $x \geq 0$:

$$f_X(x) = \int_0^{+\infty} 2e^{-x}e^{-2y}dy = e^{-x},$$

mentre $f_X(x) = 0$ per $x < 0$. Pertanto, X ha densità esponenziale di parametro 1.
Per $y \geq 0$:

$$f_Y(y) = \int_0^{+\infty} 2e^{-x}e^{-2y}dx = 2e^{-2y},$$

mentre $f_Y(y) = 0$ per $y < 0$. Pertanto, Y ha densità esponenziale di parametro 2.
Risulta, ovviamente, $f(x,y) = f_X(x)f_Y(y)$, ovvero X e Y sono indipendenti e quindi $cov(X,Y) = 0$.
(ii) Si ha:

$$P(X < Y) = \int_0^{+\infty} dx \int_x^{+\infty} dy f(x,y) = \int_0^{+\infty} e^{-x}dx \int_x^{+\infty} 2e^{-2y}dy$$

$$= \int_0^{+\infty} e^{-x}e^{-2x}dx = \int_0^{+\infty} e^{-3x}dx = \frac{1}{3} \; .$$

Si ha:

$$P(X+Y < 2 | X > 1) = \frac{P(X > 1, X+Y < 2)}{P(X > 1)}.$$

Calcoliamo prima $P(X > 1, X+Y < 2)$; si ottiene:

$$P(X > 1, X+Y < 2) = \int_1^2 e^{-x}dx \int_0^{2-x} 2e^{-2y}dy = \int_1^2 e^{-x}(1 - e^{-4+2x})dx$$

$$= \int_1^2 (e^{-x} - e^{-4}e^{x})dx = (-e^{-x} - e^{-4}e^{x})\big|_1^2 = e^{-1} + e^{-3} - 2e^{-2}.$$

Allora

$$P(X+Y < 2 | X > 1) = \frac{e^{-1} + e^{-3} - 2e^{-2}}{e^{-1}} = 1 + e^{-2} - 2e^{-1}.$$

(iii) Si ha:

$$\begin{cases} X = (U+V)/2 \\ Y = (U-V)/2 \end{cases}$$

Naturalmente, tenendo conto che $X \geq 0,\ Y \geq 0$, che $U = X + Y,\ V = X - Y$, deve essere $U \geq 0,\ \ U + V \geq 0$ e $V \leq U$, ovvero deve essere $U \geq 0$ e $-U \leq V \leq U$. La matrice Jacobiana della trasformazione inversa è:

$$J = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}$$

il cui determinante vale $-1/2$. Pertanto, la densità congiunta del vettore (U, V) è:

$$f_{(U,V)}(u,v) = \frac{1}{2} f\left(\frac{u+v}{2}, \frac{u-v}{2}\right) \mathbf{1}_{\{u+v \geq 0\}} \cdot \mathbf{1}_{\{u-v \geq 0\}}.$$

Le funzioni indicatrici valgono entrambe uno quando $u \geq 0$ e $-u \leq v \leq u$; pertanto

$$f_{(U,V)}(u,v) = \frac{1}{2} \cdot 2e^{-(u+v)/2} e^{-2(u-v)/2} \mathbf{1}_{\{u \geq 0,\ -u \leq v \leq u\}}$$

$$= e^{-3u/2} e^{v/2} \mathbf{1}_{\{u \geq 0,\ -u \leq v \leq u\}}.$$

Come si vede facilmente, $f_{(U,V)}(u,v)$ non si può scrivere come prodotto di una funzione della sola variabile u e di una funzione della sola variabile v, quindi U e V non sono indipendenti (d'altra parte, $X + Y$ è dipendente da $X - Y$). Si ha, inoltre:

$$cov(U,V) = cov(X+Y, X-Y) = E[(X+Y)(X-Y)] - E(X+Y)E(X-Y)$$

$$= E(X^2 - Y^2) - (1 + 1/2)(1 - 1/2) = 2 - 1/2 - 3/4 = 3/4 \ .$$

3.

(i) Per $n = 500$, sia $X_i,\ i = 1, .., n$, la v.a. che vale 1 se l'i–esima luce al led è priva di difetti, e 0 altrimenti. Le v.a. X_i sono indipendenti e Bernoulliane di parametro p incognito (quindi anche la media $\mu = p$ e la varianza $\sigma^2 = p(1-p)$ sono incognite); $X = X_1 + X_2 + .. + X_n$ rappresenta il numero totale di luci al led prive di difetti. Il problema fornisce per la media campionaria di luci al led non difettose il valore $\bar{x} = (x_1 + x_2 + .. + x_n)/n = \frac{350}{500} = 0.7$. Un intervallo I di confidenza a livello $1 - \alpha$ per la media incognita di una distribuzione avente varianza σ^2, è:

$$I = \left[\bar{x} - \frac{\sigma}{\sqrt{n}} \phi_{1-\alpha/2}, \bar{x} + \frac{\sigma}{\sqrt{n}} \phi_{1-\alpha/2}\right] \qquad (*)$$

dove $\bar{x}$ è la media campionaria, e ϕ_β è il quantile della Gaussiana standard, tale che $\Phi(\phi_\beta) = \beta$. Nel caso in esame, si ha $n = 500,\ \bar{x} = 0.7$ e σ è incognito. Da $1 - \alpha = 0.95$ segue $1 - \alpha/2 = 0.975$, e quindi dalla tavola dei valori di Φ si ricava $\phi_{1-\frac{\alpha}{2}} = 1.96$. Sostituendo in $(*)$, si ottiene l'intervallo:

$$I(\sigma) = \left[0.7 - \frac{\sigma}{10\sqrt{5}} \cdot 1.96,\ 0.7 + \frac{\sigma}{10\sqrt{5}} \cdot 1.96\right].$$

Poiché $\sigma = \sqrt{p(1-p)} \leq 1/2,\ \forall p \in [0,1]$, l'intervallo $I(\sigma)$ è certamente contenuto in

$$I = \left[0.7 - \frac{0.5}{10\sqrt{5}} \cdot 1.96,\ 0.7 + \frac{0.5}{10\sqrt{5}} \cdot 1.96\right] = [0.6562, 0.7438]\ ,$$

che è l'intervallo di confidenza cercato per la percentuale di luci al led accettabili. Si noti che l'intervallo trovato ha un'ampiezza $\geq$ di quella che si sarebbe trovata se σ fosse stata nota, avendo dovuto fare una maggiorazione.

(ii) Si ha:

$$\begin{aligned} P(|\overline{X}_n - p| \leq 0.1) &= P(|X_1 + X_2 + \ldots X_n - np| \leq n \cdot 0.1) \\ &= P\left(\frac{-n \cdot 0.1}{\sigma\sqrt{n}} \leq \frac{X_1 + X_2 + \ldots X_n - np}{\sigma\sqrt{n}} \leq \frac{n \cdot 0.1}{\sigma\sqrt{n}}\right) \\ &\approx 2\Phi\left(\frac{0.1 \cdot \sqrt{n}}{\sqrt{0.21}}\right) - 1. \end{aligned}$$

Imponendo che tale quantità sia ≥ 0.95 si ottiene

$$\Phi\left(\frac{0.1 \cdot \sqrt{n}}{\sqrt{0.21}}\right) \geq 0.975 = \Phi(1.96).$$

Quindi, deve essere

$$\sqrt{n} \cdot 0.1/\sqrt{0.21} \geq 1.96$$

e, risolvendo, si ricava $\sqrt{n} \geq 8.98$ ovvero $n \geq 81$.