Calcolo delle Probabilità e Statistica, Ingegneria Civile e A&T e Informatica
II prova finale a.a. 2017/18

Punteggi: **1:** $4 + 5$; **2 :** $4 + 4 + 5$; **3:** $4 + 4$.

1. Sia X una v.a. con distribuzione esponenziale di parametro $\lambda = 3$.
(i) Posto $Y = [X]$, trovare la legge di Y e calcolare $E(Y)$ e $E(Y^2)$;
($[x] = \max\{n \in \mathbf{Z} : n \leq x\}$ denota la parte intera di x).
(ii) Si effettuano lanci successivi di un dado equo: sia T il minimo numero di lanci affinché si ottenga la faccia 4; supponiamo che T e la v.a. Y del punto precedente siano indipendenti.
(a) Calcolare $P(T - 1 > Y)$.
(b) Calcolare $P(T - 1 = 3Y)$.

2. Si consideri la v.a. bidimensionale (X, Y) con densità congiunta:

$$f(x, y) = \begin{cases} 2(1 - x - y + 2xy) & \text{se } (x, y) \in (0, 1)^2 \\ 0 & \text{altrimenti} \end{cases}$$

(i) Trovare le densità marginali di X e Y e calcolare $E(X)$, $E(Y)$, $Var(X)$, $Var(Y)$ e $cov(X, Y)$.
(ii) Trovare la densità di X condizionata a $\{Y = 1/3\}$ e calcolare $E(X|Y = 1/3)$.
(iii) Posto $U = X + Y$ e $V = X - Y$, calcolare la densità congiunta del vettore aleatorio (U, V) e calcolare $E(U) + 7E(V)$.

3. Si effettua un test su un campione di 500 tastiere per computer, per rilevare eventuali difetti di produzione. Supponiamo che 350 pezzi superino il controllo.
(i) Trovare un intervallo di confidenza al livello $1 - \alpha = 0.95$ per la percentuale p di tastiere della fornitura accettabili.
(ii) Siano $X_1, X_2, \ldots, X_n$ v.a. di Bernoulli di parametro $\theta = 0.9$ e poniamo
$\bar{X}_n = (X_1 + X_2 + \ldots + X_n)/n$. Stimare quanto deve essere grande n affinché risulti
$P(|\bar{X}_n - \theta| \geq 0.01) \leq 0.01$.

Soluzioni della II prova finale a.a. 2017/18

1. (i) $Y = [X]$ risulta una v.a. discreta a valori nell' insieme dei numeri naturali, compreso lo zero. Se $k \geq 0$ è un intero, si ha

$$P(Y = k) = P(X \in [k, k+1)) = \int_k^{k+1} 3e^{-3x} dx = (1 - e^{-3})e^{-3k},$$

per cui Y ha distribuzione Geometrica di parametro $p = 1 - e^{-3}$. Si ha poi:

$$E(Y) = 1/p - 1 = 1/(e^3 - 1); \; Var(Y) = (1-p)/p^2;$$

quindi

$$E(Y^2) = E^2(Y) + Var(Y) = 1/(e^3 - 1)^2 + e^{-3}/(1 - e^{-3})^2 = (1 + e^{-3})/(1 - e^{-3})^2.$$

(ii) T ha distribuzione geometrica modificata di parametro $1/6$.
(a) Si ha:

$$P(T > Y + 1) = \sum_{k=0}^{\infty} P(T > k+1, Y = k)$$

e, per l' indipendenza di Y e T questa probabilità vale

$$\sum_{k=0}^{\infty} P(T > k+1) P(Y = k) = \sum_{k=0}^{\infty} \left(\frac{5}{6}\right)^{k+1} p(1-p)^k$$

$$= \frac{5p}{6} \sum_{k=0}^{\infty} \left(\frac{5}{6}(1-p)\right)^k = \frac{5p}{6} \cdot \frac{1}{1 - 5(1-p)/6} = \frac{5 - 5e^{-3}}{6 - 5e^{-3}} \; .$$

(b) Poiché $S = T - 1$ ha distribuzione geometrica di parametro $1/6$, si ha:

$$P(T - 1 = 3Y) = P(S = 3Y) = \sum_{k=0}^{\infty} P(S = 3k) P(Y = k)$$

$$= \sum_{k=0}^{\infty} \frac{1}{6} \left(\frac{5}{6}\right)^{3k} p(1-p)^k = \frac{p}{6} \sum_{k=0}^{\infty} \left((\frac{5}{6})^3 (1-p)\right)^k$$

$$= \frac{p}{6} \cdot \frac{1}{1 - (5/6)^3(1-p)} \; .$$

2. (i) Per $x \in (0, 1)$ si ha:

$$f_X(x) = \int_0^1 2(1 - x - y + 2xy) dy = 2\left(1 - x + \Big[-y^2/2 + 2xy^2/2\Big]_0^1\right) = 1.$$

Analogamente per $y \in (0,1)$ (l' espressione della densità congiunta è simmetrica in x e y):

$$f_Y(y) = 1.$$

Quindi, le marginali X e Y sono entrambe uniformemente distribuite nell' intervallo (0.1). Calcoliamo ora la covarianza. Intanto

$$E(XY) = \int_0^1 dx \int_0^1 dyxyf(x,y) = 2\int_0^1 xdx \int_0^1 y(1-x-y+2xy)dy = \ldots = \frac{5}{18}\ .$$

Allora

$$cov(X,Y) = E(XY) - E(X)E(Y) = \frac{5}{18} - \frac{1}{4} = -\frac{1}{36}\ .$$

Le v.a. X e Y non sono indipendenti.
(ii) Si ha:

$$f_{X|Y}(x|1/3) = f(x,1/3)/f_Y(1/3) = 2(1-x-1/3+2x/3)/1 = \frac{2}{3}(2-x),\ x \in (0,1)$$

Inoltre

$$E(X|Y=1/3) = \int_0^1 \frac{2}{3}(2-x)xdx = \frac{2}{3}\Big[x^2 - x^3/3\Big]_0^1 = \frac{4}{9}\ .$$

(iii) Effettuando la trasformazione $(X,Y) \longrightarrow (U,V) = (X+Y, X-Y)$, si trova che $U \in (0,2)$ e $V \in (-1,1)$. La matrice Jacobiana della trasformazione inversa $(X,Y) = ((U+V)/2, (U-V)/2)$ vale $-1/4$, pertanto si trova:

$$f_{(U,V)}(u,v) = f_{(X,Y)}((u+v)/2,(u-v)/2)\cdot\frac{1}{4} = \frac{1}{4}(2-2u+u^2-v^2)\mathbf{1}_{(0,2)\times(-1,1)}(u,v).$$

Si ha, senza dover trovare le densità di U e V :

$$E(U + 7E(V)) = E(X) + E(Y) + 7(E(X) - E(Y)) = \frac{1}{2} + \frac{1}{2} = 1,$$

visto che $E(V) = E(X) - E(Y) = 0$, e X e Y sono uniformemente distribuite in $(0,1)$ e quindi hanno media $1/2$.

3. (i) Per $n = 500$, sia $X_i,\ i = 1,..,n$, la v.a. che vale 1 se l'i–esima tastiera 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 tastiere prive di difetti. Il problema fornisce per la media campionaria di tastiere 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 tastiere 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:

$$P\left(|\bar{X}_n - \theta| \geq 0.01\right) = P\left(|X_1 + \ldots X_n - n\theta| \geq 0.01 \cdot n\right) =$$

$$= P\left(\left|\frac{X_1 + \ldots X_n - n\theta}{\sigma\sqrt{n}}\right| \geq \frac{0.01 \cdot n}{\sigma\sqrt{n}}\right)$$

e, siccome $\sigma^2 = 0.9 \cdot 0.1 = 0.09$, ovvero $\sigma = 0.3$, per l'approssimazione normale questa probabilità vale circa

$$P\left(|W| \geq \frac{0.01 \cdot \sqrt{n}}{0.3}\right) = 1 - P\left(|W| \leq \sqrt{n}/30\right) = 2(1 - \Phi\left(\sqrt{n}/30\right)),$$

dove abbiamo indicato con W una v.a. Gaussiana standard. Affinché l'ultima quantità ottenuta sia ≤ 0.01, deve aversi $\Phi\left(\sqrt{n}/30\right) \geq 0.995 \approx \Phi(2.575)$. Dunque deve essere $\sqrt{n}/30 \geq 2.575$, ovvero $n \geq 5968$.