CPS II Prova scritta del 9/07/2021.

1. Si lanciano contemporaneamente e ripetutamente una moneta ed un dado equilibrati. Sia X il numero minimo di lanci della moneta affinché si ottenga Testa e sia Y il numero minimo di lanci del dado per ottenere un punto ≤ 4.
(i) Qual è la densità discreta di X? e di Y? Trovare $E(X)$ e $E(Y)$.
(ii) Trovare la legge di $Z = \min(X, Y)$ e calcolare $E(Z)$.
(iii) Calcolare $P(X + Y = 5)$.

2. Per $k > 0$, si consideri la funzione:

$$f(x,y) = \begin{cases} k(x+y) & \text{se } (x,y) \in [0,3]^2 \\ 0 & \text{altrimenti} \end{cases}$$

(i) Dopo aver determinato la costante k, in modo che f sia la densità di una v.a. bidimensionale assolutamente continua (X, Y), trovare le densità marginali di X e Y e dire se X e Y sono stocasticamente indipendenti.
(ii) Calcolare media e varianza e covarianza di X, Y.
(iii) Posto $Z = X - Y$, trovare la densità di Z e calcolare $P(Z \leq 0)$.
(iv) Fornire un esempio di densità congiunta $g(u, v)$ di un vettore aleatorio (U, V), in modo che U abbia la stessa densità di X e inoltre U e V siano indipendenti.

3. Un campione di 100 lampade a led viene estratto da una grossa fornitura ed esaminato per rilevare eventuali difetti. Si trova che 20 pezzi non superano il controllo.
(i) Calcolare un intervallo di confidenza a livello $1 - \alpha = 0.95$ per la percentuale p di lampade della fornitura accettabili.
(ii) Siano $X_1, X_2, \ldots$ v.a. di Bernoulli di parametro $p = 0.8$, e sia $\bar{X}_n = \frac{1}{n}(X_1 + X_2 + \ldots + X_n)$. Stimare quanto deve essere grande n affinché sia $P(|\bar{X}_n - p| \leq 0.01) \geq 0.99$).

CPS   Soluzioni Prova scritta del 9/07/2021.

1. (i) X è l'istante di primo successo in una serie di prove indipendenti e Bernoulliane, in ciascuna delle quali la probabilità del successo (cioé che esca Testa) è $p = \frac{1}{2}$; Y è l'istante di primo successo in una serie di prove indipendenti e Bernoulliane, in ciascuna delle quali la probabilità del successo (cioé che esca un punteggio ≤ 4) è $p' = \frac{4}{6} = \frac{2}{3}$. Dunque, per $k = 1, 2, \dots$, si ha:

$$P(X = k) = \frac{1}{2}(\frac{1}{2})^{k-1} = (\frac{1}{2})^k; \quad P(Y = k) = \frac{2}{3}(\frac{1}{3})^{k-1}$$

Inoltre $E(X) = \frac{1}{p} = 2$, $E(Y) = \frac{1}{p'} = \frac{3}{2}$.
(ii) Siccome X e Y sono indipendenti, per $k = 1, 2, \dots$:

$$P(Z > k) = P(X > k)P(Y > k).$$

Ricordando che per una v.a. Geometrica modificata T di parametro θ risulta $P(T > n) = (1 - \theta)^n$,si ha:

$$P(Z > k) = (1 - 1/2)^k(1 - 2/3)^k = (1/6)^k.$$

Quindi:

$$P(Z = k) = P(z > k - 1) - P(Z > k) = (\frac{1}{6})^{k-1} - (\frac{1}{6})^k = (\frac{1}{6})^{k-1}(1 - \frac{1}{6}) = \frac{5}{6}(\frac{1}{6})^{k-1}$$

Pertanto Z è un istante di primo successo, relativo ad un esperimento aleatorio con $p'' = \frac{5}{6}$, e quindi $E(Z) = \frac{1}{p''} = \frac{6}{5}$.
(iii) Si ha:

$$P(X + Y = 5) = P(X = 2, Y = 3) + P(X = 3, Y = 2) + P(X = 1, Y = 4) + P(X = 4, Y = 1) =$$

$$= P(X = 2)P(Y = 3) + P(X = 3)P(Y = 2) + P(X = 1)P(Y = 4) + P(X = 4)P(Y = 1) = \dots$$

2. Si ha:

$$\int_0^3 dx \int_0^3 dy(x + y) = 27,$$

per cui $k = \frac{1}{27}$.
(i) Si ha:

$$f_X(x) = \int_0^3 dy \frac{1}{27}(x + y) = \frac{x}{9} + \frac{1}{6} \ , x \in [0, 3].$$

Analogamente, vista la simmetria, si ottiene $f_Y(y) = f_X(y) = \frac{y}{9} + \frac{1}{6}$, $y \in [0, 3]$. Ovviamente, X e Y non sono indipendenti, in quanto $f(x, y) \neq f_X(x)f_Y(y)$.
(ii) Si ha:

$$E(X) = E(Y) = \int_0^3 x\left(\frac{x}{9} + \frac{1}{6}\right) dx = \frac{7}{4}$$

$$E(X^2) = E(Y^2) = \int_0^3 x^2\left(\frac{x}{9} + \frac{1}{6}\right) dx = \frac{15}{4}$$

per cui

$$Var(X) = Var(Y) = \frac{15}{4} - \left(\frac{7}{4}\right)^2 = \frac{11}{16}$$

(iii) Si ha:

$$E(XY) = \int_0^3 dx \int_0^3 dy \frac{1}{27}(x^2y + xy^2) = 3,$$

Pertanto $cov(X,Y) = E(XY) - E(X)E(Y) = 3 - \frac{49}{16} = -\frac{1}{16}$.
(iii) Come è facile vedere, $Z = X - Y \in [-3,3]$.
Per $t \in [-3,0]$ si ha:

$$P(Z \leq t) = \int\int_{[0,3]^2 \cap \{y \geq x-t\}} \frac{1}{27}(x+y)dxdy = \int_{-t}^{3} dy \int_0^{y+t} \frac{1}{27}(x+y)dx =$$

$$\frac{1}{27}\int_{-t}^{3}(\frac{3}{2}y^2 + 2ty + \frac{t^2}{2})dy = \frac{1}{27}(\frac{y^3}{2} + ty^2 + \frac{t^2}{2}y)\|_{-t}^{3} = (\frac{1}{18}t^2 + \frac{t}{3} + \frac{1}{2}).$$

Per $t \in [0,3]$ si ha:

$$P(Z \leq t) = 1 - \int_t^3 dx \int_0^{x-t} \frac{1}{27}(x+y)dy = 1 - \frac{1}{27}\int_t^3 (\frac{3}{2}x^2 - 2tx + \frac{t^2}{2})dx =$$

$$= 1 - \frac{1}{27}(\frac{x^3}{2} - tx^2 + \frac{t^2 x}{2})\|_t^3 = (1 - \frac{t^2}{18} + \frac{t}{3} - \frac{1}{2}) = -\frac{t^2}{18} + \frac{t}{3} + \frac{1}{2}.$$

Derivando, si ottiene la densità di Z:

$$f_Z(z) = I_{[-3,0]}(z)(\frac{z}{9} + \frac{1}{3}) + I_{[0,3]}(z)(-\frac{z}{9} + \frac{1}{3})$$

Inoltre $P(Z \leq 0) = \frac{1}{2}$.
(iv) Basta prendere $g(u,v) = f_X(u)f_Y(v) = \left(\frac{u}{9} + \frac{1}{6}\right)\left(\frac{v}{9} + \frac{1}{6}\right)$, $u, v \in [0,3]$.

3. (i) Per $n = 100$, sia X_i, $i = 1,..,n$ la v.a. che vale 1 se l'i–esima lampada 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 lampade prive di difetti. Il problema fornisce per la media campionaria di lampade non difettose il valore $\bar{x} = (x_1 + x_2 + .. + x_n)/n = 0.8$, ottenuta dalle n lampade esaminate. 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 $n = 100$, $\bar{x} = 0.8$ 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:

$$I(\sigma) = \left[0.8 - \frac{\sigma}{10}\cdot 1.96,\ 0.8 + \frac{\sigma}{10}\cdot 1.96\right].$$

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

$$I = \left[0.8 - \frac{0.5}{10}\cdot 1.96,\ 0.8 + \frac{0.5}{10}\cdot 1.96\right] = [0.702, 0.898]\ ,$$

che è l'intervallo di confidenza cercato per la percentuale di dischetti accettabili. Si noti che l'intervallo trovato ha un'ampiezza eventualmente maggiore di quella che si sarebbe trovata se σ fosse stata nota, avendo dovuto fare una maggiorazione.
(ii) Si ha:

$$P(|\bar{X}_n - p| \leq 0.01) = P(|X_1 + \ldots X_n - np| \leq 0.01\cdot n) = P(|\frac{X_1 + \ldots X_n - np}{\sigma\sqrt{n}}| \leq \frac{0.01\cdot n}{\sigma\sqrt{n}})$$

e, siccome $\sigma^2 = 0.8\cdot 0.2 = 0.16$, per l'approssimazione normale questa probabilità vale circa

$$P(|W| \leq \frac{0.01\cdot\sqrt{n}}{0.4}) = P(|W| \leq \sqrt{n}/40) = 2\Phi(\sqrt{n}/40) - 1,$$

dove abbiamo indicato con W una v.a. Gaussiana standard. Affinché l'ultima quantitá ottenuta sia ≥ 0.99, deve aversi $\Phi(\sqrt{n}/40) \geq 0.99 \approx \Phi(2.33)$. Dunque deve essere $\sqrt{n}/40 \geq 2.33$, ovvero $n > 8687$.

Calcolo delle Probabilità e Statistica, Ingegneria Civile e A&T e Informatica
I prova finale a.a. 2016/17

Punteggi: **1:** $3+6$; **2 :** $4+5+2+2$; **3:** $4+4$.

1. Una scatola contiene 100 monete; 80 di queste sono equilibrate, mentre le altre 20 danno testa con probabilità 3/4 e croce con probabilità 1/4.
(i) Una moneta viene scelta a caso e lanciata $n=10$ volte. Qual è la probabilità di ottenere 7 volte testa? Qual è la probabilità che la moneta sia una di quelle equilibrate sapendo che in 10 lanci si è ottenuto 7 volte testa?
(ii) Selezioniamo ora una moneta equilibrata, M_1, ed una moneta non equilibrata, M_2, e lanciamo più volte e indipendentemente queste due monete; sia T il minimo numero di lanci necessario ad ottenere testa per la prima volta, lanciando la moneta M_1 ed S il minimo numero di lanci necessario ad ottenere testa per la prima volta, lanciando la moneta M_2. Calcolare:
a) $P(S \geq T^2 | S \leq 2)$; b) $P(S \leq T)$; c) $P(\frac{5}{2}S = 10 - \frac{5}{2}T)$.

2. La v.a. Y ha legge $\Gamma(3/2, 1/2)$, mentre la densità condizionale di X dato $Y=y>0$ è normale di media 0 e varianza $1/y$.
(i) Calcolare $E(\sqrt{Y})$ e $Var(\sqrt{Y})$.
(ii) Calcolare la densità congiunta del vettore aleatorio (X,Y) e la densità marginale di X; le v.a. X e Y sono indipendenti?
(iii) La v.a. X ha speranza matematica finita? Nel caso, quanto vale $E(X)$?
(iv) Calcolare $P(XY \leq 0)$.

3. (i) I semiconduttori prodotti da una fabbrica risultano privi di difetti con probabilità $p=0.973$. Dare un'approssimazione della probabilità che in un lotto di 1000 semiconduttori ve ne siano meno ($\leq$) di 970 senza difetti.
ii) Una nuova linea di produzione di semiconduttori viene messa in servizio e di essa si vuole valutare la proporzione q di quelli privi di difetti. In un lotto di 1000 pezzi prodotti se ne trovano 980 senza difetti. Qual è un intervallo di fiducia di livello $1-\alpha=0.95$ per la proporzione q?

Soluzioni della I prova finale a.a. 2016/17

1. (i) Consideriamo gli eventi "si è scelta una moneta equilibrata" (chiamiamolo E) e "si è scelta una moneta non equilibrata" (chiamiamolo N). Indichiamo poi con A l'evento "in 10 lanci si sono ottenute 7 teste". I dati del problema implicano che $P(E) = \frac{80}{100} = \frac{4}{5}, P(N) = \frac{20}{100} = \frac{1}{5}$. Se la moneta prescelta è equilibrata, allora la probabilità di ottenere 7 teste in 10 lanci è data dalla legge binomiale $B(10, 1/2)$; quindi la probabilità di avere 7 teste in 10 lanci per la moneta equilibrata è

$$\binom{10}{7}\left(\frac{1}{2}\right)^{10}.$$

Dunque

$$P(A|E) = \binom{10}{7}\left(\frac{1}{2}\right)^{10}.$$

Se invece la moneta non è equilibrata, la probabilità di avere 7 teste in 10 lanci è data dalla legge binomiale $B(10, 3/4)$; pertanto, la probabilità di avere 7 teste in 10 lanci per la moneta non equilibrata è

$$\binom{10}{7}\left(\frac{3}{4}\right)^{7}\left(\frac{1}{4}\right)^{3}.$$

Dunque

$$P(A|N) = \binom{10}{7}\left(\frac{3}{4}\right)^{7}\left(\frac{1}{4}\right)^{3}.$$

(ii) Per la formula delle probabilità totali la probabilità di avere 7 teste in 10 lanci risulta essere

$$P(A) = P(E)P(A|E) + P(N)P(A|N) = \frac{4}{5}\binom{10}{7}\left(\frac{1}{2}\right)^{10} + \frac{1}{5}\binom{10}{7}\left(\frac{3}{4}\right)^{7}\left(\frac{1}{4}\right)^{3}$$

$$= \frac{4}{5}\cdot 0.117 + \frac{1}{5}\cdot 0.25 = 0.094 + 0.05 = 0.144$$

Ora occorre calcolare $P(E|A)$; per la formula di Bayes:

$$P(E|A) = \frac{P(A|E)P(E)}{P(A)} = \frac{\binom{10}{7}\left(\frac{1}{2}\right)^{10}\cdot\frac{4}{5}}{0.144}$$

$$= \frac{0.94}{0.144} = 0.65\ .$$

(iii) T ha distribuzione geometrica modificata di parametro $1/2$, mentre T ha distribuzione geometrica modificata di parametro $3/4$. Inoltre S e T sono indipendenti.

a) Si ha:

$$P(S \geq T^2, S \leq 2) = P(\{T = 1, S = 1\} \cup \{T = 1, S = 2\})$$

$$= P(T=1)P(S=1) + P(T=1)P(S=2) = \frac{1}{2}\cdot\frac{3}{4} + \frac{1}{2}\cdot\frac{3}{16} = \frac{15}{32} \ .$$

Quindi:

$$P(S \geq T^2 | S \leq 2) = P(S \geq T^2, S \leq 2)/P(S \leq 2) = (15/32)/(15/16) = 1/2.$$

b)
$$P(S > T) = \sum_{k=1}^{\infty} P(S > k, T = k) = \sum_{k=1}^{\infty} P(S > k)P(T = k)$$

$$= \sum_{k=1}^{\infty} (1 - 3/4)^k \cdot \frac{1}{2}\left(\frac{1}{2}\right)^{k-1}$$

$$= \sum_{k=1}^{\infty} \left(\frac{1}{4}\right)^k \left(\frac{1}{2}\right)^k = \sum_{k=1}^{\infty} \left(\frac{1}{8}\right)^k = \frac{1}{1-1/8} - 1 = \frac{1}{7} \ .$$

Quindi $P(S \leq T) = 1 - P(S > T) = 6/7$.
c) L'evento $\{\frac{5}{2}S = 10 - \frac{5}{2}T\}$ non è altro che l'evento $\{S + T = 4\}$; pertanto la probabilità cercata è:

$$P(S+T=4) = P(S=1, T=3) + P(S=3, T=1) + P(S=2, T=2)$$

$$= \frac{3}{4}\cdot\frac{1}{2}\left(\frac{1}{2}\right)^2 + \frac{3}{4}\cdot\left(\frac{1}{4}\right)^2\cdot\frac{1}{2} + \frac{3}{4}\cdot\frac{1}{4}\left(\frac{1}{2}\right)^2$$

$$= \frac{3}{32} + \frac{3}{8\cdot 16} + \frac{3}{16\cdot 4} = \frac{21}{16\cdot 8} = 0.1640625 \ .$$

2. (i) La densità di Y è, per $y > 0$:

$$f_Y(y) = \frac{(1/2)^{3/2}}{\Gamma(3/2)} y^{1/2} e^{-y/2} = \frac{1}{\sqrt{2\pi}} y^{1/2} e^{-y/2}$$

(ricordare che $\Gamma(1/2) = \sqrt{\pi}$ e $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, per cui $\Gamma(3/2) = \sqrt{\pi}/2$), mentre è zero per $y \leq 0$. Dunque

$$E(\sqrt{Y}) = \int_0^{+\infty} \sqrt{y} f_Y(y) dy = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} y e^{-y/2} dy$$

$$= \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} y \frac{1}{2} e^{-y/2} dy = \sqrt{\frac{2}{\pi}} \int_0^{+\infty} y \frac{1}{2} e^{-y/2} dy$$

$$= \sqrt{\frac{2}{\pi}} \cdot 2,$$

essendo l'integrale uguale alla media di una v.a esponenziale di parametro 1/2, che vale 2. Il momento del second'ordine è ancora più facile da calcolare, ricordando il valore della media delle leggi Gamma:

$$E((\sqrt{Y})^2) = E(Y) = 3.$$

Pertanto, $Var(\sqrt{Y}) = E((\sqrt{Y})^2) - E^2(\sqrt{Y}) = 3 - 8/\pi$.
(ii) Siccome

$$f_{X|Y}(x|y) = \frac{1}{\sqrt{2\pi}}\sqrt{y}e^{-x^2y/2},$$

la densità congiunta di (X, Y) è per $x \in (-\infty, +\infty)$ e $y > 0$:

$$f(x, y) = f_{X|Y}(x|y) \cdot f_Y(y) = \frac{1}{2\pi}ye^{-y(1+x^2)/2}.$$

La densità marginale di X è:

$$f_X(x) = \int_0^{+\infty} f(x, y)dy = \frac{1}{2\pi}\int_0^{+\infty} ye^{-y(1+x^2)/2}dy$$

$$= \frac{1}{2\pi}\frac{\Gamma(2)}{((1+x^2)/2)^2}\int_0^{+\infty}\frac{((1+x^2)/2)^2}{\Gamma(2)}y^{2-1}e^{-y(1+x^2)/2}dy$$

$$= \frac{2}{\pi}\cdot\frac{1}{(1+x^2)^2},$$

poiché la funzione integranda è una densità Gamma di parametri 2 e $(1 + x^2)/2$ e quindi l'integrale vale 1.
Ovviamente, le v.a. X e Y non sono indipendenti.
(iii) Siccome $|x|f_X(x) \sim 1/x^3$ per $x \to \infty$, risulta che $E(X)$ è finita; anzi, visto che $f_X(x)$ è una funzione pari, si ha $E(X) = 0$.
(iv) Siccome $Y > 0$, si ha $P(XY \leq 0) = P(X \leq 0) = 1/2$, sempre per la parità di $f_X(x)$.

3.
(i) Il numero, X, di semiconduttori senza difetti nel lotto di 1000 pezzi segue una legge binomiale $B(1000, 0.973)$. Applicando l'approssimazione normale con la correzione di continuità, otteniamo:

$$P(X \leq 970) = P(X \leq 970.5)$$

$$\approx \Phi\left(\frac{970.5 - 1000 \cdot 0.973}{\sqrt{0.973 \cdot 0.027}\sqrt{1000}}\right) = \Phi(-0.487)$$

$$= 1 - \Phi(0.487) = 1 - 0.687 = 0.313.$$

Il risultato esatto, disponendo della funzione di ripartizione della binomiale $B(1000, 0.973)$ è 0.305.

(ii) Un intervallo I di confidenza a livello $1-\alpha$ per la media incognita di una distribuzione avente varianza σ^2, è:

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

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 $n = 1000$, $\bar{x} = 0.98$, mentre σ è incognita; considerando che il campione proviene da una sequenza di v.a. di Bernoulli, la cui varianza σ^2 è $\leq 1/4$, possiamo maggiorare σ con $1/2$. 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 $(*)$ $\bar{x} = 0.98$, $\phi_{1-\frac{\alpha}{2}} = 1.96$ e $\sigma = 1/2$ si ottiene l'intervallo $I = [0.95,\ 1.01]$.

CPS I Prova scritta del 22/06/2021.

1. Si consideri la v.a. bidimensionale discreta (X, Y) con densità:

$$p(x, y) = [2^{x+y}(5/2)^y]/[e^7 x!\ y!], \quad x, y \in \mathbb{N} \cup \{0\}$$

(i) Trovare le densità marginali di X e Y. Si tratta di densità note? X e Y sono stocasticamente indipendenti?
(ii) Calcolare $cov(X + 1, Y + 1)$, $E(X - 2Y)$ e $Var(X - 2Y)$.
(iii) Trovare la densità di $Z = X + Y$ e calcolare $P(Z > 3)$.
(iv) Per $h, k \in \mathbb{N} \cup \{0\}$ con $h \geq k$, calcolare $P(X = k|Z = h)$; in particolare calcolare $P(X = 1|Z = 2)$. Condizionatamente a $Z = h$, che cosa si può dire riguardo alla distribuzione di X?

2. Per $\alpha > 0$, si consideri la funzione:

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

(i) Dopo aver determinato la costante α, in modo che f sia la densità di una v.a. bidimensionale assolutamente continua (X, Y), trovare le densità marginali di X e Y e dire se X e Y sono stocasticamente indipendenti.
(ii) Calcolare media e varianza di X e Y e $Cov(X, Y)$.
(iii) Calcolare $P(X \leq Y)$.
(iv) Calcolare $P(Y > 1/2|X < 1/2)$.

3. Si vuole verificare il carico di rottura medio di alcuni cavi. A tale scopo, si eseguono 20 prove di rottura su altrettanti cavi identici, ottenendo carichi di rottura $x_1, \ldots, x_{20}$ con media campionaria pari a $\bar{x}_n = 10.49\ t$ (t = tonnellate). Da considerazioni teoriche, si può assumere che il campione $x_1, \ldots, x_{20}$ abbia deviazione standard pari a $\sigma = 1.2\ t$.
(i) Trovare un intervallo di confidenza a livello $1 - \alpha = 0.99$ per il carico di rottura medio μ.
(ii) Determinare la dimensione minima del campione affinché l'ampiezza dell'intervallo di confidenza a livello 0.99 per la media μ non superi $0.514\ t$.

CPS Soluzioni Prova scritta del 22/06/2021.

1. (i) Osserviamo che $p(x,y)$ può essere scritta come

$$p(x,y) = \frac{2^x 5^y}{x!\ y!}e^{-7}, \quad x,y \in \mathbb{N} \cup \{0\}$$

Pertanto, si ha, per $x = 0,1,2,\dots$:

$$p_X(x) = P(X = x) = \sum_{y=0}^{\infty} \frac{2^x 5^y}{x!\ y!}e^{-7} = e^{-7}\frac{2^x}{x!}\sum_{y=0}^{\infty}\frac{5^y}{y!}$$

$$= e^{-7}\frac{2^x}{x!}e^5 = e^{-2}\frac{2^x}{x!}$$

e quindi $X \sim$ Poisson(2). Analogamente, per $y = 0,1,2,\dots$:

$$p_Y(x) = P(Y = y) = \sum_{x=0}^{\infty} \frac{2^x 5^y}{x!\ y!}e^{-7} = e^{-7}\frac{5^y}{y!}\sum_{x=0}^{\infty}\frac{2^x}{x!}$$

$$= e^{-7}\frac{5^y}{y!}e^2 = e^{-5}\frac{5^y}{y!}$$

e quindi $X \sim$ Poisson(5). Le v.a. X e Y sono stocasticamente indipendenti, essendo $p(x,y) = p_X(x)p_Y(y)$, e quindi $cov(X,Y) = 0$.
(ii) Poiché per ogni a,b, si ha $cov(X+a, Y+b) = E[(X+a-E(X+a))(Y+a-E(Y+b))] = E(X-E(X))(Y-E(Y)] = cov(X,Y)$, allora $cov(X+1,Y+1) = cov(X,Y) = 0$. Inoltre, ricordando che, sia la media che la varianza di una v.a. con distribuzione di Poisson di parametro λ sono uguali a λ, si ottiene $E(X-2Y) = E(X) - 2E(Y) = 2 - 2\cdot 5 = -8$; inoltre, siccome X e Y sono indipendenti (e quindi anche X e $-2Y$ lo sono), si ha $Var(X-2Y) = Var(X) + Var(-2Y) = Var(X) + (-2)^2 Var(Y) = 2 + 4\cdot 5 = 22$.
(iii) Ricordando che la somma di v.a. indipendenti di Poisson, di parametri λ e μ, rispettivamente, ha distribuzione di Poisson di parametro $\lambda + \mu$, si ottiene che $Z \sim$ Poisson(7) e quindi

$$P(Z > 3) = 1 - P(Z=0) - P(Z=1) - P(Z=2) - P(Z=3)$$

$$= 1 - e^{-7}(1 + 7 + 7^2/2! + 7^3/3!) = 0.918$$

(iii) Per $h \geq k$, si ha

$$P(X = k|Z = h) = \frac{P(X = k, X+Y = h)}{P(X+Y = h)} = \frac{P(X = k, Y = h-k)}{P(Z = h)} =$$

(essendo X e Y indipendenti)

$$\left[e^{-2}\frac{2^k}{k!}\cdot e^{-5}\frac{5^{h-k}}{(h-k)!}\right] \Big/ \left[e^{-7}\frac{7^h}{h!}\right]$$

$$= \frac{2^k 5^{h-k}}{7^h}\cdot\frac{h!}{k!(h-k)!} = \binom{h}{k}\frac{2^k}{7^k}\cdot\frac{5^{h-k}}{7^{h-k}} = \binom{h}{k}\left(\frac{2}{7}\right)^k\left(\frac{5}{7}\right)^{h-k}$$

Pertanto, condizionatamente a $Z = h$, la v.a. X ha distribuzione binomiale di parametri h e $2/7$.
Per $k = 1$ e $h = 2$, si ottiene

$$P(X = 1|Z = 2) = \binom{2}{1}\frac{2}{7}\cdot\frac{5}{7} = 2\cdot\frac{10}{49} = \frac{20}{49} = 0.4081$$

2. Si ha:

$$\int_0^1 dx \int_0^2 dy(x^2 + xy/2) = 7/6,$$

per cui $k = \frac{6}{7}$.
(i) Per $x \in (0,1)$ si ha:

$$f_X(x) = \frac{6}{7}\int_0^2 (x^2 + xy/2)dy = \frac{6}{7}(2x^2 + x),$$

e 0 altrimenti. Analogamente, per $y \in (0,2)$, si ha:

$$f_Y(y) = \frac{6}{7}\int_0^1 (x^2 + xy/2)dx = \frac{6}{7}(1/3 + y/4),$$

e 0 altrimenti. Ovviamente, X e Y non sono indipendenti, in quanto $f(x,y) \neq f_X(x)f_Y(y)$.
(ii) Si ha:

$$E(X) = \frac{6}{7}\int_0^1 x(2x^2 + x)dx = 5/7,$$

$$E(Y) = \frac{6}{7}\int_0^2 (y/3 + y^2/4)dy = 8/7,$$

mentre

$$E(XY) = \frac{6}{7}\int_0^1 dx \int_0^2 dy\ xy(x^2 + xy/2) = \frac{17}{21}\ .$$

Pertanto $cov(X,Y) = E(XY) - E(X)E(Y) = \frac{17}{21} - \frac{5}{7}\frac{8}{7} = -\frac{1}{147}$.
(iii) Si ha:

$$P(X > Y) = \frac{6}{7}\int_0^1 dx \int_0^x (x^2 + xy/2)dy = \frac{15}{56},$$

e quindi $P(X < Y) = 1 - \frac{15}{56} = \frac{41}{56}$.
(iv) Si ha:

$$P(Y > 1/2|X < 1/2) = [P(Y > 1/2, X < 1/2)]/P(X < 1/2);$$

il numeratore vale

$$\int_0^{1/2} dx \int_{1/2}^2 dy \frac{6}{7}(x^2 + xy/2) = \frac{69}{448},$$

ed il denominatore è

$$\int_0^{1/2} dx \frac{6}{7}(2x^2 + x) = \frac{5}{28};$$

Quindi, effettuando i calcoli:

$$P(Y > 1/2|X < 1/2) = 69/80 = 0.8625$$

3. (i) 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 $n = 20$, $\bar{x} = 10.49$ e $\sigma = 1.2$ Da $1 - \alpha = 0.99$ segue $1 - \alpha/2 = 0.995$, e quindi dalla tavola dei valori di Φ si ricava $\phi_{1-\frac{\alpha}{2}} = 2.57$. Sostituendo in $(*)$, si ottiene l'intervallo:

$$I = \left(10.49 - \frac{1.2 \cdot 2.57}{\sqrt{20}},\ 10.49 + \frac{1.2 \cdot 2.57}{\sqrt{20}}\right) = (9.79, 11.18).$$

(ii) L'ampiezza dell'intervallo di confidenza è $a = \frac{2 \cdot 1.2}{\sqrt{n}} \cdot 2.57$; imponendo che $a \leq 0.514$ si ottiene una disequazione nell'incognita n che, risolta, fornisce $n \geq 144$.

FACOLTÀ DI INGEGNERIA
CALCOLO DELLE PROBABILITÀ E STATISTICA
INGEGNERIA CIVILE E A&T E INFORMATICA

PRIMA PROVA SCRITTA
A.A. 2018-2019

Durata della prova 2 h
Punteggi: 1) $3+3+3$; **2)** $2+5+4+4$; **3)** $3+3$.

Totale = 30.

Esercizio 1 Sia Z una variabile aleatoria a valori interi positivi tale che, per $n \geq 1$, risulti $P(Z \leq n) = 1-(2/3)^n$, e W un'altra v.a., indipendente da Z e con la stessa distribuzione di Z.

(i) Calcolare $P(Z^2 < W^2)$.

(ii) Calcolare $E[(Z-1)^2 - \frac{1}{2}(1-W)^2]$.

(iii) Siano X e Y v.a. indipendenti con distribuzione di Poisson e tali che $E(X^2) = E(Y^2) = 12$. Calcolare

$$P\left(2 - XY \leq \frac{X^2+Y^2}{2} < 8 - XY\right).$$

Esercizio 2 Per $\alpha > 0$ si consideri la funzione $f : \mathbf{R}^2 \longrightarrow \mathbf{R}$ definita da:

$$f_\alpha(x,y) = \begin{cases} \alpha\left[e^{-(x+y)} + e^{-2(x+y)}\right] & x>0,\ y>0 \\ 0 & \text{altrimenti} \end{cases}$$

$$= \alpha\left[e^{-(x+y)} + e^{-2(x+y)}\right] \mathbf{1}_{(0,+\infty)}(x)\mathbf{1}_{(0,+\infty)}(y).$$

(i) Trovare il valore $\bar{\alpha}$ di α in modo che $f_{\bar{\alpha}}(x,y)$ sia la densità congiunta di un vettore aleatorio bidimensionale.

(ii) Si consideri la v.a. bidimensionale (X,Y) che ha per densità $f_{\bar{\alpha}}(x,y)$; si calcolino le densità marginali di X e Y, $E(X)$, $E(Y)$ e $cov(X,Y)$. Le v.a. X e Y sono stocasticamente indipendenti?

(iii) Trovare la densità di $Z = X + 2Y$ e calcolare $P(Y \leq 2X+1)$.

(iv) Calcolare $P(Y \leq X | X > 1)$.

Esercizio 3 Alcuni studenti effettuano delle misurazioni per determinare il punto di fusione dello stagno. Dai dati ottenuti da 100 diverse misurazioni effettuate, si trova una media campionaria di 231.8°C.

(i) Supponendo che il punto di fusione cercato sia distribuito secondo una v.a. con deviazione standard $\sigma = 15.4$°C, trovare un intervallo di confidenza al livello $1-\alpha = 0.95$ per la temperatura media di fusione dello stagno.

(ii) Supponiamo ora che i dati $X_i,\ i = 1,\dots,100$, ottenuti dalle misurazioni siano estratti da una stessa distribuzione con media 231.8 e varianza σ^2, dove σ è quella del precedente punto (i). Utilizzando l'approssimazione normale, calcolare $P(\overline{X}_{100} \leq 232)$, dove $\overline{X}_{100} = \frac{X_1+X_2+\cdots+X_{100}}{100}$. Si tratta di una probabilità minore o maggiore di 1/2 ?

Calcolo delle Probabilità e Statistica, a.a. 2018-19

Soluzioni della prima prova scritta

Esercizio 1 Dal fatto che $P(Z > n) = (2/3)^n$, si desume che Z è geometrica modificata di parametro $p = 1/3$, quindi $P(Z = k) = \frac{1}{3}\left(\frac{2}{3}\right)^{k-1}$, $k = 1, 2, \dots$ e $E(Z) = 1/p = 3$.

(i) Si ha:

$$P(Z^2 < W^2) = P(|Z|) < |W|) = P(Z < W)$$

$$= \sum_{k=1}^{\infty} P(Z < k)P(W = k) = \sum_{k=1}^{\infty} \frac{1}{3}\left(\frac{2}{3}\right)^{k-1}\left[1 - \left(\frac{2}{3}\right)^{k-1}\right],$$

visto che $P(Z < k) = 1 - P(Z \geq k) = 1 - P(Z > k-1) = 1 - (\frac{2}{3})^{k-1}$. Continuando il calcolo:

$$P(Z^2 < W^2) = \frac{1}{3}\left[\sum_{k=1}^{\infty}\left(\frac{2}{3}\right)^{k-1} - \sum_{k=1}^{\infty}\left(\frac{2}{3}\right)^{2k-2}\right]$$

$$= \sum_{k=1}^{\infty}\frac{1}{3}\left(\frac{2}{3}\right)^{k-1} - \frac{1}{3}\cdot\frac{9}{4}\sum_{k=1}^{\infty}\left(\frac{4}{9}\right)^{k}$$

$$= 1 - \frac{3}{4}\left(\frac{1}{1-4/9} - 1\right) = 1 - \frac{3}{4}(9/5 - 1) = 1 - \frac{3}{4}\cdot\frac{4}{5} = \frac{2}{5}$$

(ii) Si ha:

$$E[(Z-1)^2 - \frac{1}{2}(1-W)^2] = E[Z^2 - 2Z + 1/2 + Z - W^2/2].$$

Visto che

$$E(Z^2) = E(W^2) = Var(Z) + E^2(Z) = \frac{1-p}{p^2} + \frac{1}{p^2} = \frac{2-p}{p^2},$$

la media cercata vale

$$\frac{2-p}{p^2} - 2/p + 1/2 + 1/p - \frac{1}{2}\cdot\frac{2-p}{p^2} =$$

$$= \frac{1}{2}\cdot\frac{2-p}{p^2} - 1/p + 1/2 = \cdots = 5,$$

dove si è sostituito $p = 1/3$.

(iii) X e Y hanno distribuzione di Poisson di parametro λ, da trovare. Siccome media e varianza di una v.a. di Poisson sono uguali a λ, si ha $E(X^2) = Var(X) + E^2(X) = \lambda + \lambda^2$; imponendo che $\lambda + \lambda^2 = 12$, si ottiene una equazione di secondo grado che, risolta, fornisce $\lambda = -4$, che si scarta, perchè negativa, e $\lambda = 3$, che è il valore cercato. Dunque, X e Y hanno distribuzione di Poisson di parametro 3. Si ha:

$$P\left(2 - XY \leq \frac{X^2+Y^2}{2} < 8 - XY\right) = P\left(4 \leq X^2 + Y^2 + 2XY < 16\right)$$

$$= P(2 \leq X + Y < 4) = P(X + Y \in \{2,\ 3\}).$$

Siccome $X + Y \sim Poisson(3+3)$, la probabilità cercata vale

$$e^{-6}\left(6^2/2 + 6^3/6\right) = 36\cdot\frac{3}{2}e^{-6} = 0.1338.$$

Esercizio 2 (i) Per $\alpha > 0$ la funzione $f_\alpha(x,y)$ è continua e positiva per $x,y > 0$. Calcolando l'integrale doppio di $f_\alpha(x,y)$ esteso a ${\mathbf{R}_+}^2$, si ottiene $\frac{5}{4}\alpha$; imponendo che tale valore sia uguale a 1, si ottiene $\bar{\alpha} = \frac{4}{5}$. Dunque, si ottiene la densità:

$$f_{\bar{\alpha}}(x,y) = \frac{4}{5}\left[e^{-(x+y)} + e^{-2(x+y)}\right]\mathbf{1}_{(0,+\infty)}(x)\mathbf{1}_{(0,+\infty)}(y).$$

(ii) La densità marginale di X è:

$$\begin{aligned} f_X(x) &= \frac{4}{5}\left[e^{-x}\int_0^{+\infty} e^{-y}dy + \frac{e^{-2x}}{2}\int_0^{+\infty} 2e^{-2y}dy\right]\mathbf{1}_{(0,+\infty)}(x) \\ &= \begin{cases} \frac{4}{5}\left[e^{-x} + e^{-2x}/2\right] = \frac{4}{5}e^{-x}(1+e^{-x}/2), & x > 0 \\ 0 & \text{altrimenti} \end{cases}. \end{aligned}$$

Visto che l'espressione di $f_{\bar{\alpha}}(x,y)$ è simmetrica rispetto a x e y, si ottiene analogamente

$$f_Y(y) = \begin{cases} \frac{4}{5}e^{-y}(1+e^{-y}/2), & y > 0 \\ 0 & \text{altrimenti} \end{cases}.$$

Si ha:

$$E(X) = \int_0^{+\infty} x f_X(x)dx = \frac{4}{5}\int_0^{+\infty}(xe^{-x} + \frac{x}{2}e^{-2x})dx = \frac{4}{5}\left(1 + \frac{1}{4}\cdot\frac{1}{2}\right) = \frac{4}{5}\cdot\frac{9}{8} = \frac{9}{10}$$

(il calcolo è stato effettuato agevolmente, ricordando la formula per la media di una v.a. esponenziale di parametro λ). Naturalmente, si ottiene anche $E(Y) = E(X) = \frac{9}{10}$, visto che X e Y hanno la stessa distribuzione.

Si ha:

$$\begin{aligned} E(XY) &= \frac{4}{5}\left[\int_0^{+\infty} dx \int_0^{+\infty} dy\ xy\left(e^{-(x+y)} + e^{-2(x+y)}\right)\right] \\ &= \frac{4}{5}\left[\int_0^{+\infty} dx\ xe^{-x}\int_0^{+\infty} dy\ ye^{-y} + \int_0^{+\infty} dx\ xe^{-2x}\int_0^{+\infty} dy\ ye^{-2y}\right] \\ &= \frac{4}{5}\left(1\cdot 1 + \frac{1}{2}\cdot\frac{1}{2}\right) = 1 \end{aligned}$$

(anche qui, abbiamo sfruttato la formula per la media di una v.a. esponenziale di parametro λ).

Quindi, otteniamo:

$$cov(X,Y) = E(XY) - E(X)E(Y) = 1 - \left(\frac{9}{10}\right)^2 = \frac{19}{100} = 0.19 \neq 0,$$

per cui le v.a. X e Y non sono stocasticamente indipendenti; d'altra parte, ciò segue anche dal fatto che la densità congiunta di (X,Y) non è uguale al prodotto delle densità marginali.

(iii) La v.a. $Z = X + 2Y$ assume valori positivi; utilizziamo il metodo di cambiamento di variabili. Consideriamo la trasformazione

$$\psi : \begin{cases} x = x \\ z = x + 2y \end{cases} \qquad \psi^{-1} : \begin{cases} x = x \\ y = (z-x)/2 \end{cases}$$

La matrice Jacobiana della trasormazione inversa è :

$$J_{\psi^{-1}}(x,z) = \begin{pmatrix} 1 & 0 \\ -1/2 & 1/2 \end{pmatrix},$$

che ha determinante uguale a 1/2. Quindi, si ottiene per la densità congiunta di (X,Z) :

$$f_{(X,Z)}(x,z) = f_{(X,Y)}(x,(z-x)/2)\ \frac{1}{2}$$

$$= \frac{4}{5}\cdot\frac{1}{2}\left[e^{-(x+(z-x)/2)} + e^{-2(x+(z-x)/2)}\right]\mathbf{1}_{(0,+\infty)}(x)\mathbf{1}_{\{x<z\}}(z)$$

$$= \frac{2}{5}\left[e^{-(x+z/2)} + e^{(x+z)}\right]\mathbf{1}_{(0,+\infty)}(x)\mathbf{1}_{\{x<z\}}(z).$$

Quindi, per $z>0$ la densità di Z è:

$$f_Z(z) = \int_{-\infty}^{\infty} dx\frac{2}{5}\left[e^{-(x+z/2)} + e^{(x+z)}\right]\mathbf{1}_{(0,+\infty)}(x)\mathbf{1}_{\{x<z\}}(z)$$

$$= \frac{2}{5}\int_0^z dx\left(e^{-z/2}e^{-x/2} + e^{-z}e^{-x}\right) = \frac{2}{5}\left[e^{-z/2}\cdot 2(1-e^{-z/2}) + e^{-z}(1-e^{-z})\right]$$

$$= \frac{2}{5}\left[2e^{-z/2} - e^{-z} - e^{-2z}\right],\ z>0.$$

Inoltre:

$$P(Y\le 2X+1) = \int_0^{+\infty} dx\int_0^{2x+1} dy f_{\bar{\alpha}}(x,y)$$

$$= \frac{4}{5}\left[\int_0^{+\infty} dxe^{-x}\int_0^{2x+1} e^{-y}dy + \int_0^{+\infty} dxe^{-2x}\cdot\frac{1}{2}\int_0^{2x+1} 2e^{-2y}dy\right]$$

$$= \frac{4}{5}\left[\int_0^{+\infty} dxe^{-x}\left(1-e^{-2x-1}\right) + \frac{1}{2}\int_0^{+\infty} dxe^{-2x}\left(1-e^{-4x-2}\right)\right]$$

$$= \frac{4}{5}\left[\int_0^{+\infty} dx\left(e^{-x} - e^{-3x}e^{-1}\right) + \frac{1}{2}\int_0^{+\infty} dx\left(e^{-2x} - e^{-6x}e^{-2}\right)\right]$$

$$= \frac{4}{5}\left[1 - \frac{1}{3}e^{-1} + \frac{1}{2}\left(\frac{1}{2} - \frac{1}{6}e^{-2}\right)\right] = \frac{1}{15}(15 - 4e^{-1} - e^{-2}) \approx 0.8928$$

(iv) Si ha:

$$P(Y\le X|X>1) = \frac{P(Y\le X,\ X>1)}{P(X>1)} = \frac{\int\int_A f(x,y)dxdy}{\int_1^{+\infty} f_X(x)dx},$$

dove $f(x,y)$ è la densità di (X,Y), $f_X(x)$ è la densità di X e
$A = \{(x,y)\in\mathbf{R}^2 : x>1,\ 0<y<x\}$.

Il numeratore vale:

$$Num = \frac{4}{5}\left[\int_1^{+\infty} dx\int_0^x dy\left(e^{-(x+y)} + e^{-2(x+y)}\right)\right]$$

$$= \frac{4}{5}\left[\int_1^{+\infty} dxe^{-x}\int_0^x dye^{-y} + \int_1^{+\infty} dxe^{-2x}\cdot\frac{1}{2}\int_0^x dy2e^{-2y}\right]$$

$$= \frac{4}{5}\left[\int_1^{+\infty} dxe^{-x}(1-e^{-x}) + \frac{1}{2}\int_1^{+\infty} dxe^{-2x}(1-e^{-2x})\right]$$

$$= \frac{4}{5}\left[\int_1^{+\infty} dx(e^{-x}-e^{-2x}) + \frac{1}{2}\int_1^{+\infty} dx(e^{-2x}-e^{-4x})\right]$$

$$= \frac{e^{-1}}{10}(8-2e^{-1}-e^{-2}) \approx 0.2654$$

Il denominatore vale:

$$Denom = \frac{4}{5}\left[\int_1^{+\infty}(e^{-x}+\frac{1}{2}e^{-2x})dx\right] = \frac{e^{-1}}{5}(4+e^{-1}).$$

Dunque, la probabilità cercata è:

$$\frac{Num}{Denom} = \frac{\frac{e^{-1}}{10}(8-2e^{-1}-e^{-2})}{\frac{e^{-1}}{5}(4+e^{-1})} = \frac{1}{2}\cdot\frac{8-2e^{-1}-e^{-3}}{4+e^{-1}} \approx 0.8258\ .$$

Esercizio 3 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 $n = 100$, la media campionaria è $\bar{x} = 231.8$, e $\sigma = 15.4$;

(i) se $1-\alpha = 0.95$, allora $1-\alpha/2 = 0.975$, e quindi dalla tavola dei valori di Φ si ricava $\phi_{1-\frac{\alpha}{2}} = 1.96$. Sostituendo in $(*)$, si ottiene che un intervallo di confidenza per la temperatura media di fusione dello stagno, al livello 0.95 è:

$$I = \left[231.8 - \frac{15.4}{10}\cdot 1.96,\ 231.8 + \frac{15.4}{10}\cdot 1.96\right] = [228.7816,\ 234.8184]\ .$$

(ii) Si ha:

$$P\left(\overline{X}_{100} \le t\right) = P\left(\frac{X_1+X_2+\cdots+X_{100}}{100} \le t\right) = P(X_1+X_2+\cdots+X_{100} \le 100t)$$

$$= P\left(\frac{X_1+X_2+\cdots+X_{100} - 100\cdot 231.8}{15.4\sqrt{100}} \le \frac{100t - 100\cdot 231.8}{15.4\sqrt{100}}\right)$$

che, per l'approssimazione normale vale circa

$$\Phi\left(\frac{100(t-231.8)}{15.4\sqrt{100}}\right) = \Phi\left(\frac{10(t-231.8)}{15.4}\right).$$

Quindi, per $t = 232$, si ottiene:

$$P(\overline{X}_{100} \le 232) \approx \Phi\left(\frac{10(232-231.8)}{15.4}\right) = \Phi(0.1298) \approx 0.59\ .$$

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 $n \geq 5968$.