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:

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

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

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

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

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

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 $P(Z \le 0) = \frac{1}{2}$.
(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 $\sigma = \sqrt{p(1-p)} \le 1/2$, $\forall p \in [0,1]$, 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:

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

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

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

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