NOME E COGNOME:

MATRICOLA:

PROVA SCRITTA DI MDP, 13/06/2023

Esercizio 1 Il numero N di fratelli/sorelle di una persona scelta a caso è una variabile geometrica $N \sim G(p)$ di media 1.5. Lanciamo un dado e scegliamo un numero di persone a caso corrispondente al numero uscito. Denotiamo con X il valore uscito dal dado e Y la variabile "numero di persone scelte che hanno esattamente 1 fratello/sorella".

a. Mostrare che $p = 2/5$.
b. Determinare la probabilità che una persona abbia esattamente un fratello/sorella.
c. Stabilire se X e Y sono indipendenti.
d. Determinare $P(Y = 1|X = 2)$ e $P(X = 2|Y = 1)$.
e. Determinare $E[Y]$.

Soluzione. a) Sappiamo che il valore atteso di una variabile geometrica modificata di parametro p è $1/p$. D'altra parte una variabile geometrica e una geometrica modificata differiscono di una traslazione di 1: o meglio, nel nostro caso, $N + 1 \sim \tilde{G}(p)$ è una variabile geometrica modificata di parametro p. Pertanto

$$E(N) = E(\tilde{G}(p)) - 1 = \tfrac{1}{p} - 1$$

Nel nostro caso, segue che il parametro della variabile geometrica che definisce il numero di fratelli/sorelle è definito dall'equazione $\frac{1}{p} - 1 = \frac{3}{2}$, quindi $\frac{1}{p} = \frac{5}{2}$ e otteniamo $p = \frac{2}{5} = 0.4$.

b) N è una variabile geometrica di parametro $p = 0.4$, dunque $P(N = 1) = (1 - p)p = \frac{6}{25} = 0.24$.

c) Le due variabili sono evidentemente dipendenti: se $X = k$ allora $Y \leq k$.

d) $P(Y = 1|X = 2)$ è la probabilità che su due persone scelte a caso ve ne sia esattamente una che abbia esattamente un fratello/sorella. Dunque

$$P(Y = 1|X = 2) = 2P(N = 1)P(N \neq 1) = 2 \cdot \tfrac{6}{25} \cdot \tfrac{19}{25} = \tfrac{228}{625} = 2 \cdot 0.24 \cdot 0.76 = 0.3648.$$

Osserviamo più in generale che se $n \in \{1, 2, 3, 4, 5, 6\}$ allora

$$P(Y = 1|X = n) = nP(N = 1)P(N \neq 1)^{n-1} = n \cdot 0.24 \cdot (0.76)^{n-1}$$

Pertanto abbiamo

$$P(Y = 1) = \sum_{n=1}^{6} P(Y = 1, X = n) = \sum_{n=1}^{6} P(Y = 1|X = n)P(X = n) = \sum_{n=1}^{6} \frac{n}{6} \cdot 0.24 \cdot (0.76)^{n-1} = 0.3679$$

e con la formula di Bayes possiamo calcolare

$$P(X = 2|Y = 1) = P(X = 2) \cdot \frac{P(Y = 1|X = 2)}{P(Y = 1)} = \frac{1}{6} \cdot \frac{0.3648}{0.3679} = 0.1653$$

e) La variabile aleatoria Y assume valori in $\{0, 1, 2, 3, 4, 5, 6\}$. Abbiamo

$$P(Y = k) = \sum_{n=1}^{6} P(Y = k, X = n) = \sum_{n=1}^{6} P(Y = k|X = n)P(X = n) = \frac{1}{6} \sum_{n=1}^{6} \binom{n}{k} \cdot (0.24)^k \cdot (0.76)^{n-k}$$

da cui ricaviamo

$$P(Y = 1) = 0.36793 \qquad P(Y = 2) = 0.15494 \qquad P(Y = 3) = 0.04284$$
$$P(Y = 4) = 0.00745 \qquad P(Y = 5) = 0.00074 \qquad P(Y = 6) = 0.00003$$

Pertanto otteniamo il valore atteso

$$E(Y) = \sum_{k=0}^{6} kP(Y = k) = 0.84$$

In alternativa, facendo meno conti, possiamo arrivare al risultato utilizzando le proprietà del valore atteso. Pensiamo ogni risultato del fenomeno aleatorio in questione come un elenco ordinato $(n; p_1, \ldots, p_6)$, dove

$n \leq 6$ è il risultato del lancio del dado e dove $p_1, \dots, p_6$ sono 6 persone prese come segue: le prime n persone $p_1, \dots, p_n$ sono scelte a caso e tutte diverse, e se $n < 6$ allora le rimanenti persone nell'elenco $p_{n+1}, \dots, p_6$ sono tutte prese uguali a p_n. Se $\omega = (n; p_1, \dots, p_6)$ è un risultato, indichiamo $X(\omega) = n$ il risultato del dado e con $p_i(\omega) = p_i$ la i-esima persona nell'elenco. Se $i \leq 6$, poniamo infine

$$X_i(\omega) = \begin{cases} 1 & \text{se } i \leq X(\omega) \\ 0 & \text{altrimenti} \end{cases} \qquad Y_i(\omega) = \begin{cases} 1 & \text{se } p_i(\omega) \text{ ha esattamente un fratello/sorella} \\ 0 & \text{altrimenti} \end{cases}$$

Allora $X_i \sim B(1, \frac{i}{6})$, mentre $Y_i \sim B(1, 0.24)$. Inoltre X_i e Y_i sono variabili aleatorie indipendenti. Osserviamo che

$$Y = \sum_{i=1}^{6} X_i Y_i$$

Per le proprietà del valore atteso troviamo quindi

$$E(Y) = \sum_{i=1}^{6} E(X_i Y_i) = \sum_{i=1}^{6} E(X_i) E(Y_i) = \sum_{i=1}^{6} \tfrac{i}{6} \cdot 0.24 = 0.24 \cdot \sum_{i=1}^{6} \frac{i}{6} = 0.24 \cdot E(X) = 0.24 \cdot 3.5 = 0.84.$$

Esercizio 2 Dato un parametro reale $a > 0$, sia

$$f(s) = \begin{cases} \frac{2}{3}(s+1) & \text{se } s \in [-1, 0] \\ a - s^2 & \text{se } s \in [0, \sqrt{a}] \\ 0 & \text{se } s \notin [-1, \sqrt{a}] \end{cases}$$

a. Disegnare un grafico approssimato di $f(s)$.
b. Mostrare che l'unico valore di $a > 0$ per cui f è la densità di una variabile aleatoria continua X è $a = 1$.
c. Calcolare $P(X > 0)$.
d. Calcolare la funzione di ripartizione di X.
e. Calcolare il valore atteso di X.

Soluzione. b) Abbiamo

$$\int_{-1}^{\sqrt{a}} f(s)ds = \int_{-1}^{0} \tfrac{2}{3}(s+1)ds + \int_{0}^{\sqrt{a}} (a-s^2)ds = \tfrac{1}{3} + [as - \tfrac{1}{3}s^3]_0^{\sqrt{a}} = \tfrac{1}{3} + \tfrac{2}{3}a\sqrt{a}$$

Affinché f sia la densità di una variabile aleatoria deve essere $\frac{1}{3} + \frac{2}{3}a\sqrt{a} = 1$. Pertanto $\sqrt{a}^3 = 1$, da cui $a = 1$.
c) $P(X > 0) = 1 - P(X < 0) = 1 - \frac{1}{3} = \frac{2}{3}$.
d) Abbiamo $a = 1$. Se $t < 0$ allora vale $F_X(t) = 0$, mentre se $t > 1$ vale $F_X(t) = 1$.
Se $t \in [-1, 0]$ abbiamo

$$F_X(t) = \int_{-1}^{t} \tfrac{2}{3}(s+1)ds = \tfrac{1}{3}(t+1)^2$$

Se $t \in [0, 1]$ infine abbiamo

$$F_X(t) = \tfrac{1}{3} + \int_{0}^{t} (1 - s^2)ds = \tfrac{1}{3} + t - \tfrac{1}{3}t^3$$

e) Abbiamo

$$E(X) = \int_{-1}^{1} sf(s)ds = \int_{-1}^{0} \tfrac{2}{3}(s^2 + s)ds + \int_{0}^{1} (s - s^3)ds = (\tfrac{2}{9} - \tfrac{1}{3}) + (\tfrac{1}{2} - \tfrac{1}{4}) = \frac{5}{36}$$

Esercizio 3 Il peso di una pigna ha densità normale di media 130 gr e scarto quadratico medio 20 gr. Sappiamo che 30 kg di pigne producono mediamente 1 kg di pinoli.

a. Quanti grammi di pinoli producono mediamente 8 kg di pigne?
b. Qual'è la probabilità che una pigna pesi più di 120 gr?
c. Quante pigne bisogna raccogliere per avere una probabilità di almeno il 90% che il loro peso totale sia di almeno 30 kg?

Soluzione. a) Sia X la variabile aleatoria che indica i grammi di pinoli contenuti in un dato chilogrammo di pigne, sia Y la variabile aleatoria che indica i grammi di pinoli contenuti in dati 30 chilogrammi di pigne, e sia Z la variabile aleatoria che indica i grammi di pinoli contenuti in dati 8 chilogrammi di pigne. Allora possiamo scrivere

$$Y = Y_1 + \ldots + Y_8, \qquad Z = Z_1 + \ldots + Z_{30}$$

con $Y_i \sim Z_j \sim X$ per tutti gli indici i, j. Pertanto abbiamo $E(Y) = 8E(X)$ e $E(Z) = 30E(X)$. D'altra parte sappiamo che $E(Z) = 1000$, pertanto troviamo $E(X) = \frac{1000}{30}$ e $E(Y) = \frac{8000}{30} = 266.66$. Dunque deduciamo che mediamente 8 kg di pigne producono 266.66 gr di pinoli.

b) Sia X il peso di una data pigna, allora

$$P(X > 120) = P(\zeta_0 > \frac{120 - 130}{20}) = P(\zeta_0 > -0.5) = 0.6915$$

c) Sia n il numero di pigne raccolte, sia X_i il peso della pigna e sia $S_n = X_1 + \ldots + X_n$. Allora $S_n \sim N(130n, 400n)$. Dunque

$$P(S_n > 30000) = P(\zeta_0 > \frac{30000 - 130n}{20\sqrt{n}}).$$

Abbiamo $P(\zeta_0 > \lambda) \geq 0.9$ se e solo se $1 - \Phi(\lambda) > 0.9$ se e solo se $\Phi(\lambda) \leq 0.1$ se e solo se $\Phi(-\lambda) \geq 0.9$ se e solo se e solo se $-\lambda \leq 1.29$. Pertanto affinché $P(S_n > 30000)$ deve essere

$$\frac{30000 - 130n}{20\sqrt{n}} \leq -1.29,$$

vale a dire

$$130n - 25.8\sqrt{n} - 30000 \geq 0$$

L'unica radice positiva della precedente equazione in $\sqrt{n}$ è 15.29. Pertanto deve essere $n \geq 234$.