ESERCIZI DI PROBABILITA' E DISTRIBUZIONI

Indice

1 Esercizi parametrici

1.1 Caso Discreto

Esercizio

Si consideri la seguente funzione

$$p(x) = \begin{cases} k/3 & x = 1, 2, 3 \\ k & x = 4 \\ k/3 & x = 5, 6, 7 \end{cases} \tag{1}$$

a) si determini il valore di k affinché $p(x)$ possa essere funzione di probabilità per una v.c. X;

b) si determini la funzione di ripartizione $F(x)$ della v.c. X e se ne disegni il grafico;

c) si determinino i quartili della v.c. X;

d) si calcoli il momento terzo centrale di X e si commenti.

Soluzione

a) La condizione $p(x_i) \geq 0$ è verificata per ogni $i = 1, \ldots, 7$. Pertanto, affinché la funzione (1) sia funzione di probabilità, si deve verificare la relazione:

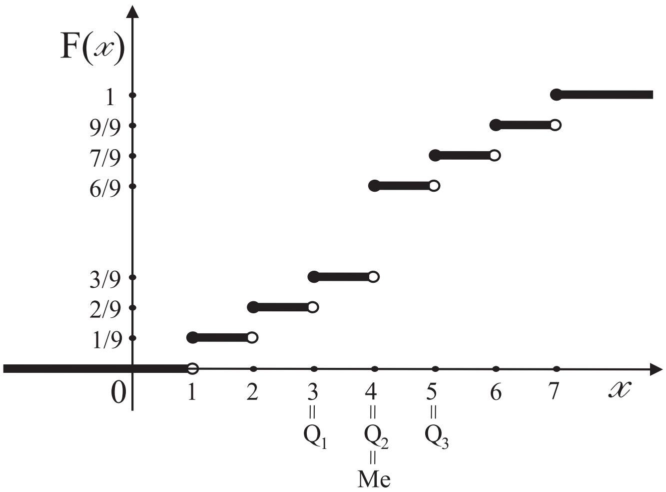
$$\sum_{i=1}^{7} p(x_i) = 1 \; . \tag{2}$$

Dunque

$$\begin{aligned} \frac{k}{3} + \frac{k}{3} + \frac{k}{3} + k + \frac{k}{3} + \frac{k}{3} + \frac{k}{3} &= 1 \\ \frac{3k + 3k + 3k}{3} &= 1 \\ \frac{9k}{3} &= 1 \\ 3k &= 1 \qquad \Rightarrow \qquad k = \frac{1}{3} \end{aligned}$$

b)

$$F(x) = \begin{cases} 0 & x < 1 \\ 1/9 & 1 \leq x < 2 \\ 2/9 & 2 \leq x < 3 \\ 3/9 & 3 \leq x < 4 \\ 6/9 & 4 \leq x < 5 \\ 7/9 & 5 \leq x < 6 \\ 8/9 & 6 \leq x < 7 \\ 1 & x \geq 7 \end{cases} \tag{3}$$



c) I quartili relativi alla v.c. X sono:

- $Q_1 = F^{-1}(0,25) = 3$;
- $Q_2 = Me = F^{-1}(0,5) = 4$;
- $Q_3 = F^{-1}(0,75) = 5$;

d)

$$\begin{aligned}
E(X) &= \sum_{i=1}^{7} x_i \cdot p(x_i) = 4 \\
E\left\{[X-E(X)]^3\right\} &= \sum_{i=1}^{7} [x_i - E(X)]^3 \cdot p(x_i) \\
&= (1-4)^3\frac{1}{9} + (2-4)^3\frac{1}{9} + (3-4)^3\frac{1}{9} + (4-4)^3\frac{3}{9} + \\
&\quad +(5-4)^3\frac{1}{9} + (6-4)^3\frac{1}{9} + (7-4)^3\frac{1}{9} \\
&= 0
\end{aligned}$$

Essendo $E\left\{[X-E(X)]^3\right\} = 0$ si può concludere che la distribuzione di X simmetrica. In generale si ha che se X è una v.c. è simmetrica allora ogni suo momento centrale di ordine r dispari è uguale a 0. Naturalmente non vale il contrario.

Esercizio

Sia X una v.c. discreta avente la seguente funzione di probabilità:

$$p(x) = \begin{cases} \dfrac{x^2}{k} & x = 1, 2, 3 \\ \\ 0 & \text{altrove} \end{cases} \tag{4}$$

a) si determini il valore di k e si tracci il grafico di $p(x)$;

b) si ricavi la funzione di ripartizione e la si rappresenti graficamente;

c) si determinino i quartili della v.c. X;

d) Si fornisca l'espressione della funzione generatrice dei momenti e da essa si ricavi $E(X)$.

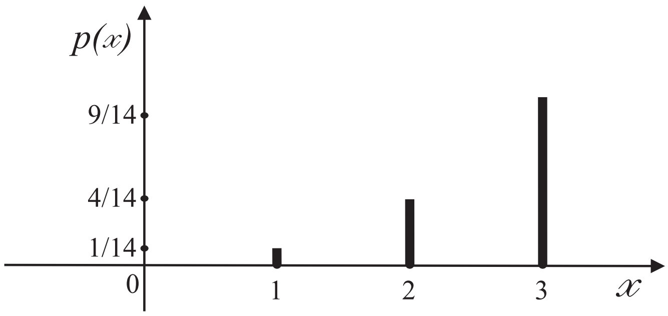
Soluzione

a) La condizione $p(x) \geq 0$ è verificata per $x = 1, 2, 3$. Pertanto la costante k si può ricavare attraverso l'uguaglianza:

$$\sum_{x=1}^{3} p(x) = 1\ . \tag{5}$$

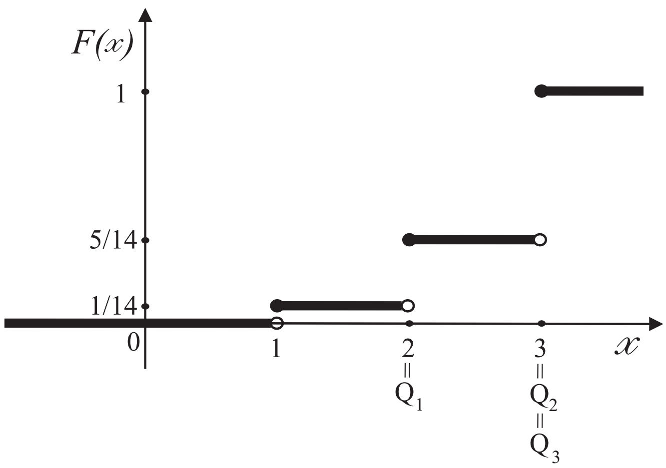
$$\begin{aligned}\frac{1^2}{k}+\frac{2^2}{k}+\frac{3^2}{k} &= 1\\ \frac{1}{k}+\frac{4}{k}+\frac{9}{k} &= 1\\ \frac{14}{k} &= 1 \Rightarrow \qquad k = 14\,.\end{aligned}$$

x	$p(x)$
1	1/14
2	4/14
3	9/14



b)

$$F(x) = \begin{cases} 0 & x < 1 \\ 1/14 & 1 \le x < 2 \\ 5/14 & 2 \le x < 3 \\ 1 & x \ge 3 \end{cases}$$



c) I quartili della v.c. X sono:

- $Q_1 = F^{-1}(0,25) = 2$;
- $Q_2 = Q_3 = F^{-1}(0,5) = F^{-1}(0,75) = 3$.

d)

$$\begin{aligned} m_X(t) &= \sum_{x=1}^{3} e^{tx} \cdot p(x) \\ &= e^{t\cdot 1} \cdot \frac{1}{14} + e^{t\cdot 2} \cdot \frac{4}{14} + e^{t\cdot 3} \cdot \frac{9}{14} \ . \end{aligned}$$

$$\begin{aligned} E(X) &= m'_X(t)\big|_{t=0} \\ &= e^{t\cdot 1} \cdot \frac{1}{14} + e^{t\cdot 2} \cdot \frac{8}{14} + e^{t\cdot 3} \cdot \frac{27}{14}\Big|_{t=0} \\ &= \frac{1}{14} + \frac{8}{14} + \frac{27}{14} = \frac{36}{14} = 2,571 \ . \end{aligned}$$

1.2 Caso Continuo

Esercizio

Sia data la funzione

$$f(x) = \begin{cases} |1-x| & 0 \leq x \leq a \\ 0 & \text{altrove} \ , \end{cases} \tag{6}$$

dove a è un parametro incognito.

a) Si determini il valore di a che rende $f(x)$ la funzione di densità di una v.c. X.

b) Si calcoli il valore atteso di X.

c) Si ricavi l'espressione della funzione di ripartizione di X e se ne determini il terzo quartile.

Soluzione

a) La funzione $|1-x| \geq 0$ sempre quindi l'incognita a si ricava dalla relazione

$$\int_0^a f(x)dx = 1 \ . \tag{7}$$

L'integrale

$$\int_0^1 |1-x|\, dx = \frac{1}{2}$$

quindi deve essere $a > 1$. Pertanto

$$\begin{aligned}
\int_0^a |1-x|\,dx &= 1 \\
\int_0^1 (1-x)dx + \int_1^a (x-1)dx &= 1 \\
\left\{x - \frac{x^2}{2}\right\}_0^1 + \left\{\frac{x^2}{2} - x\right\}_1^a &= 1 \\
1 - \frac{1}{2} + \frac{a^2}{2} - a - \frac{1}{2} + 1 &= 1 \\
\frac{a^2}{2} - a &= 0 \\
a(a-2) &= 0 \qquad \Rightarrow \qquad \text{soluzioni} \begin{cases} a = 0 \\ a = 2\ . \end{cases}
\end{aligned}$$

La soluzione $a = 0$ rende nullo l'integrale (7) quindi l'unica soluzione accettabile è $a = 2$.

b)

$$\begin{aligned}
E(X) &= \int_0^2 x\,|1-x|\,dx \\
&= \int_0^1 x(1-x)dx + \int_1^2 x(x-1)dx \\
&= \int_0^1 (x-x^2)dx + \int_1^2 (x^2-x)dx \\
&= \left\{\frac{x^2}{2} - \frac{x^3}{3}\right\}_0^1 + \left\{\frac{x^3}{3} - \frac{x^2}{2}\right\}_1^2 \\
&= \frac{1}{2} - \frac{1}{3} + \frac{8}{3} - \frac{4}{2} - \frac{1}{3} + \frac{1}{2} = 1\ .
\end{aligned}$$

c) Per $0 \le x < 1$ risulta:

$$F(X) = \int_0^x (1-t)\,dt = \left\{t - \frac{t^2}{2}\right\}_0^x = x - \frac{x^2}{2}$$

Per $1 \leq x < 2$ risulta:

$$\begin{aligned}
F(X) &= \int_0^1 (1-t)\,dt + \int_1^x (t-1)\,dt \\
&= 1 - \frac{1}{2} + \left\{ \frac{t^2}{2} - t \right\}_1^x \\
&= \frac{1}{2} + \frac{x^2}{2} - x - \frac{1}{2} + 1 \\
&= 1 - x + \frac{x^2}{2}
\end{aligned}$$

Pertanto la funzione di ripartizione risulta:

$$F(x) = \begin{cases} 0 & x < 0 \\ x - \dfrac{x^2}{2} & 0 \leq x < 1 \\ 1 - x + \dfrac{x^2}{2} & 1 \leq x < 2 \\ 1 & x \geq 2 \end{cases}$$

Il terzo quartile, cioè il valore di x tale che $F(x) = 0,75$ si trova nell'intervallo $x \in [1, 2)$. Tale considerazione deriva dal fatto che $F(1) = 0,5$. Essendo la funzione di ripartizione una funzione non decrescente ne consegue che il terzo quartile deve essere maggiore o uguale a 1. Ma per $x \geq 2$ la $F(x)$ assume valore 1 pertanto il terzo quartile deve essere $1 \leq x < 2$. Si prende in considerazione allora la forma funzionale della $F(x)$ nell'intervallo $[1, 2)$ e si ricava

$$\begin{aligned}
1 - x + \frac{x^2}{2} &= 0,75 \\
x^2 - 2x + 0,5 &= 0 \\
x_{1,2} &= \frac{2 \pm \sqrt{4-2}}{2}
\end{aligned}$$

$$\text{soluzioni} = \begin{cases} x_1 = \frac{2-\sqrt{2}}{2} = 0,29 & \text{fuori dall'intervallo [1,2);} \\ x_2 = \frac{2+\sqrt{2}}{2} = 1,71 = Q_3 & \text{compresa nell'intervallo [1,2)} \\ & \text{quindi soluzione accettabile.} \end{cases}$$

Esercizio

Sia X una v.c. continua avente la seguente funzione di densità

$$f(x) = \begin{cases} kx & -1 < x < 0 \\ -kx & 0 \leq x < 3 \\ 0 & \text{altrove} . \end{cases} \tag{8}$$

a) Si determini il valore della costante k che rende $f(x)$ la funzione di densità della v.c. X e si tracci il grafico di $f(x)$.

b) Si ricavi la funzione di ripartizione e il primo quartile di X.

c) Si calcoli il valore atteso di X.

d) Si calcoli il secondo momento standardizzato della v.c. X e si commenti.

Soluzione

a) Affinché la funzione $f(x)$ sia maggiore o uguale a zero per ogni x deve risultare $kx \geq 0$ per $-1 < x < 0$ e deve essere $-kx \geq 0$ per $0 \leq x < 3$. Ciò accade solamente se la costante k è minore o uguale a zero. In realtà se fosse $k = 0$ la funzione $f(x)$ non sarebbe funzione di densità in quanto risulterebbe costantemente pari a 0. Si cerca dunque $k < 0$ tale che sia soddisfatta la condizione

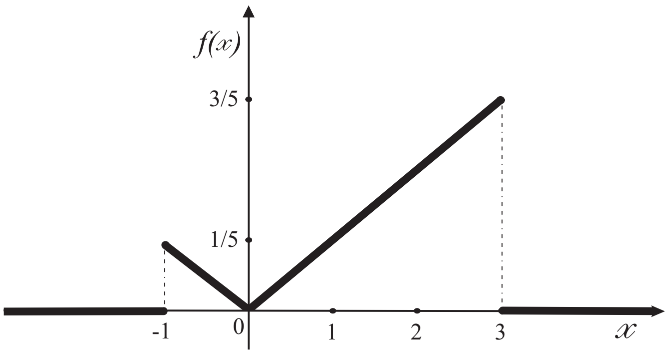
$$\int_{-1}^{3} f(x)dx = 1 \ . \tag{9}$$

Pertanto

$$\begin{aligned} \int_{-1}^{3} f(x)dx &= 1 \\ \int_{-1}^{0} kx\ dx + \int_{0}^{3} (-kx)dx &= 1 \\ k\left\{\frac{x^2}{2}\right\}_{-1}^{0} - k\left\{\frac{x^2}{2}\right\}_{0}^{3} &= 1 \\ -\frac{k}{2} - k\frac{9}{2} &= 1 \\ -5k &= 1 \qquad \Rightarrow \qquad k = -\frac{1}{5} \end{aligned}$$

Risulta che la funzione di densità è la seguente

$$f(x) = \begin{cases} -\dfrac{1}{5}x & -1 < x < 0 \\ \dfrac{1}{5}x & 0 \leq x < 3 \\ 0 & \text{altrove .} \end{cases}$$



b) Per $-1 \leq x < 0$ risulta:

$$F(X) = \int_{-1}^{x} \left(-\frac{1}{5}t\right) dt = -\frac{1}{5}\left\{\frac{t^2}{2}\right\}_{-1}^{x} = -\frac{x^2}{10} + \frac{1}{10}$$

Per $0 \leq x < 3$ risulta:

$$\begin{aligned} F(X) &= \underbrace{\int_{-1}^{0} \left(-\frac{1}{5}t\right) dt}_{F(0)} + \int_{0}^{x} \frac{1}{5}t\, dt \\ &= \frac{1}{10} + \frac{1}{5} \cdot \left\{\frac{t^2}{2}\right\}_{0}^{x} \\ &= \frac{x^2}{10} + \frac{1}{10} \end{aligned}$$

Pertanto la funzione di ripartizione risulta:

$$F(x) = \begin{cases} 0 & x < -1 \\ \\ -\dfrac{x^2}{10} + \dfrac{1}{10} & -1 \leq x < 0 \\ \\ \dfrac{x^2}{10} + \dfrac{1}{10} & 0 \leq x < 3 \\ \\ 1 & x \geq 3 \end{cases}$$

$F(0) = 0,1$ pertanto deve essere $Q_1 \in [0,3)$. Il primo quartile $Q_1 = x_{0,25}$ è quel valore di x tale per cui $F(x) = 0,25$ pertanto

$$\begin{aligned} \frac{x_{0,25}^2}{10} + \frac{1}{10} &= 0,25 \\ x_{0,25}^2 &= 1,5 \\ x_{0,25} &= \pm\sqrt{1,5}\ . \end{aligned}$$

Dunque $Q_1 = \sqrt{1,5}$.

c)

$$\begin{aligned} E(X) &= \int_{-1}^{3} x f(x) dx \\ &= \int_{-1}^{0} x\left(-\frac{1}{5}x\right) dx + \int_{0}^{3} x\left(\frac{1}{5}x\right) dx \\ &= -\frac{1}{5}\left\{\frac{x^3}{3}\right\}_{-1}^{0} + \frac{1}{5}\left\{\frac{x^3}{3}\right\}_{0}^{3} \\ &= \frac{1}{15} + \frac{9}{5} = \frac{28}{15}\ . \end{aligned}$$

d)

$$E\left[\left(\frac{X-\mu}{\sigma}\right)^2\right] = \frac{1}{\sigma^2} E\left[(X-\mu)^2\right] = \frac{1}{\sigma^2}\sigma^2 = 1\ ,$$

dove $\mu = E(X)$ e $\sigma^2 = Var(X)$. Il secondo momento standardizzato è uguale ad 1 per qualsiasi variabile casuale.

2 Probabilità totali e Bayes

Esercizio

Alle primarie per il voto del leader di uno schieramento politico si presentano due candidati: Rossi e Bianchi. Da precedenti elezioni dello stesso tipo è emerso che il 53% dei votanti è di sesso maschile. Inoltre recenti sondaggi prevedono che il 54% della popolazione femminile voterà Bianchi mentre il 52% della popolazione maschile voterà Rossi.

a) Con le informazioni a disposizione, calcolare la probabilità che Rossi ha di vincere le elezioni.

b) Rossi vince le elezioni e si riscontra che: il 58% dei votanti è risultato essere di sesso maschile, il 57% della popolazione femminile votante ha votato per Rossi ed il 41% della popolazione maschile votante ha votato per Bianchi (non ci sono state schede nulle). Avendo estratto a caso una scheda dall'urna che contiene tutti i voti e avendo riscontrato che tale voto è a favore di Bianchi, si calcoli la probabilità che la scheda sia stata compilata da un maschio.

c) Estratto a caso un elettore dei votanti, qual'è la probabilità che sia di sesso maschile ed abbia votato per Rossi.

Soluzione

a) Con le informazioni a disposizione si può ottenere la seguente tabella

	Maschio	Femmina	
Rossi	$0,2756$	$0,2162$	$0,4918$
Bianchi	$0,2544$	$0,2538$	$0,5082$
	$0,53$	$0,47$	1

Il punto di partenza è la probabilità $P(\text{Maschio}) = 0,53$ da cui si ricava $P(\text{Femmina}) = 0,47$. Da quest'ultima probabilità si ottiene

$$\begin{aligned} P(\text{Femmina} \cap \text{Bianchi}) &= P(\text{Bianchi}|\text{Femmina}) \cdot P(\text{Femmina}) \\ &= 0,54 \cdot 0,47 = 0,2538 \,. \end{aligned}$$

Inoltre

$$\begin{aligned} P(\text{Maschio} \cap \text{Rossi}) &= P(\text{Rossi}|\text{Maschio}) \cdot P(\text{Maschio}) \\ &= 0,52 \cdot 0,53 = 0,2756 \,. \end{aligned}$$

Tali probabilità, insieme alle informazioni iniziali sulle marginali, permettono di ricavare la tabella sopra riportata.

La probabilità che vinca Rossi è dunque

$$\begin{aligned} P(\text{Rossi}) &= P(\text{Maschio} \cap \text{Rossi}) + P(\text{Femmina} \cap \text{Rossi}) \\ &= 0,2756 + 0,2162 = 0,4918 \ . \end{aligned}$$

che è lo stesso valore riportato in tabella.

b) Dallo spoglio dei voti è possibile costruire la seguente tabella di probabilità congiunte e marginali

	Maschi	Femmine	
Rossi	$0,3422$	$0,2394$	$0,5816$
Bianchi	$0,2378$	$0,1806$	$0,4184$
	$0,58$	$0,42$	1

È possibile sfruttare la formula di Bayes o formula delle probabilità a priori

$$\begin{aligned} P(\text{Maschio}|\text{Bianchi}) &= \frac{P(\text{Maschio} \cap \text{Bianchi})}{P(\text{Bianchi})} \\ &= \frac{P(\text{Bianchi}|\text{Maschio}) \cdot P(\text{Maschio})}{P(\text{Bianchi})} \\ &= \frac{0,41 \cdot 0,58}{0,4184} = 0,568 \ . \end{aligned}$$

c) Dalla tabella ricavata al punto b) è facile ricavare che

$$P(\text{Maschio} \cap \text{Rossi}) = 0,3422 \ .$$

Esercizio

Il 26% degli iscritti alla Facoltà di Sociologia di una università italiana è costituito da maschi. Il 45% di questi è fumatore mentre il 25% delle femmine è fumatore.

a) Avendo estratto casualmente un iscritto e avendo verificato che è fumatore, si determini la probabilità che sia maschio.

b) Si determini la probabilità che, estraendo un iscritto, risulti maschio o fumatore.

c) Estratti con riposizione 120 iscritti, si determini la probabilità che almeno 40 individui selezionati nel campione siano fumatori.

Soluzione

a) Si adotti la seguente notazione: F = femmina, M=maschio, FU=fumatore, NF=non fumatore. Le informazioni a disposizione sono le seguenti:

$$\begin{aligned} P(\text{M}) &= 0,26 \quad \text{quindi} \quad P(\text{F}) = 0,74 \\ P(\text{FU}|\text{M}) &= 0,45 \\ P(\text{FU}|\text{F}) &= 0,25\ . \end{aligned}$$

Pertanto

$$\begin{aligned} P(\text{M}|\text{FU}) &= \frac{P(\text{M} \cap \text{FU})}{P(\text{FU})} = \frac{P(\text{M}) \cdot P(\text{FU}|\text{M})}{P(\text{M}) \cdot P(\text{FU}|\text{M}) + P(\text{F}) \cdot P(\text{FU}|\text{F})} \\ &= \frac{0,26 \cdot 0,45}{0,26 \cdot 0,45 + 0,74 \cdot 0,25} = \frac{0,117}{0,302} = 0,387\ . \end{aligned}$$

b) Si ricava dai dati forniti la seguente tabella

	FU	NF	
M	$0,117$	$0,143$	$0,26$
F	$0,185$	$0,556$	$0,74$
	$0,302$	$0,698$	1

da cui si ottiene:

$$\begin{aligned} P(\text{M} \cup \text{FU}) &= P(\text{M}) + P(\text{FU}) - P(\text{FU} \cap \text{M}) \\ &= 0,26 + 0,302 - 0,117 = 0,445 \end{aligned}$$

c) Si consideri la v.c. X che conta fra i 120 iscritti selezionati quelli che sono fumatori. Tale variabile si distribuisce con legge di probabilità binomiale di parametri $n = 120$ e $p = P(\text{FU}) = 0,302$. Il numero delle estrazioni risulta maggiore di 30 quindi è ragionevole approssimare X con una variabile casuale Normale Y di aspettativa $E(Y) = 120 \cdot 0,302 = 36,24$ e di varianza $Var(Y) = 120 \cdot 0,302 \cdot (1 - 0,302) = 25,296$. Puntualizzato ciò la probabilità che X sia maggiore o uguale di 40 si può calcolare come segue

$$
\begin{aligned}
P(X \geq 40) &= 1 - P(X < 40) \\
&\approx 1 - P(Y < 40) \\
&\approx 1 - \Phi\left(\frac{40 - 36,24}{\sqrt{25,296}}\right) \\
&\approx 1 - \Phi(0,74) \\
&\approx 1 - 0,7703 = 0,2297 \ .
\end{aligned}
$$

3 Funzione generatrice dei momenti

Esercizio

Sia data la funzione generatrice dei momenti

$$
m(t) = \frac{1}{1 - 2t} \qquad t < 0,5 \ .
$$

a) Sia X una v.c. dotata della f.g.m di cui sopra. Si calcolino il valore atteso e la varianza di X.

b) Sia X una v.c. dotata della f.g.m. di cui sopra. Si determini la f.g.m. della v.c. $Y = X/2$.

c) Siano X_1 ed X_2 due v.c. indipendenti ciascuna dotata di f.g.m. di cui sopra. Si determini la f.g.m. della v.c. $Y = X_1 + X_2$.

d) Siano X_1 ed X_2 due v.c. indipendenti ciascuna dotata di f.g.m. di cui sopra. Si determini la f.g.m. della v.c. $Z = X_1 + 3X_2$. Si calcolino inoltre il valore atteso e la varianza di Z.

Soluzione

a) Il calcolo dei momenti si ottiene derivando opportunamente la funzione generatrice dei momenti $m(t)$ e ponendo t uguale a 0:

$$
\begin{aligned}
E(X) &= \left[\frac{d}{dt} m(t)\right]_{t=0} \\
&= \left[\frac{2}{(1 - 2t)^2}\right]_{t=0} = 2 \ ,
\end{aligned}
$$

$$
\begin{aligned}
E(X^2) &= \left[\frac{d^2}{dt^2}m(t)\right]_{t=0} \\
&= \left[\frac{8}{(1-2t)^3}\right]_{t=0} = 8\ ,
\end{aligned}
$$

$$
\begin{aligned}
Var(X) &= E(X^2) - [E(X)]^2 \\
&= 8 - 2^2 = 4\ .
\end{aligned}
$$

b) Se $Y = a + bX$ allora

$$m_Y(t) = e^{at} m_X(bt)\ .$$

In questo caso dato che $Y = \frac{1}{2}X$ risulta $a = 0$ e $b = \frac{1}{2}$, pertanto

$$
\begin{aligned}
m_Y(t) &= e^{0t} m_X\left(\frac{1}{2}t\right) \\
&= (1-t)^{-1}\ .
\end{aligned}
$$

c) Se X_1 è indipendente da X_2 ed $Y = X_1 + X_2$ risulta che

$$
\begin{aligned}
m_Y(t) &= m_{X_1}(t) \cdot m_{X_2}(t) \\
&= (1-2t)^{-1}(1-2t)^{-1} \\
&= (1-2t)^{-2}\ .
\end{aligned}
$$

d) Posto $W = 3X_2$ si ha che X_1 è indipendente da W. Ciò è dovuto al fatto che se due v.c. X ed Y sono indipendenti allora risultano indipendenti anche le v.c. $g(X)$ ed $h(Y)$ qualsiasi siano le funzioni $g(\cdot)$ ed $h(\cdot)$.

$$
\begin{aligned}
m_W(t) &= e^{0t} m_{X_2}(3t) \\
&= (1-6t)^{-1}\ .
\end{aligned}
$$

$$
\begin{aligned}
m_Z(t) &= m_{X_1}(t) \cdot m_W(t) \\
&= (1-2t)^{-1}(1-6t)^{-1}\ .
\end{aligned}
$$

$$
\begin{aligned}
E(Z) &= E(X_1) + E(W) = E(X_1) + 3E(X_2) \\
&= 2 + 3 \cdot 2 = 8 \\
Var(Z) &= Var(X_1) + Var(W) = Var(X_1) + 9Var(X_2) \\
&= 4 + 9 \cdot 4 = 40\ .
\end{aligned}
$$

4 Variabili casuali discrete

4.1 Binomiale

Esercizio

Presso uno sportello bancomat del centro di Milano 4 persone su 5 fanno operazione di prelievo.

a) Supponendo di estrarre a caso 10 persone (con riposizione) che si sono recate allo sportello, calcolare la probabilità che il numero totale di persone che ha prelevato sia maggiore di 9.

b) Posto che in 6 estrazioni (con riposizione) almeno 3 persone hanno prelevato, calcolare la probabilità che il numero totale di persone che ha prelevato sia maggiore di 4.

c) Supponendo di estrarre a caso 300 persone (con riposizione) che si sono recate allo sportello, calcolare la probabilità che il numero totale di persone che ha prelevato sia compreso tra 245 e 255.

d) Calcolare il numero medio di estrazioni (con riposizione) che si devono effettuare per ottenere una persona che non ha prelevato.

e) Si supponga di estrarre (con riposizione) in maniera sequenziale le persone che si sono presentate allo sportello. Calcolare la probabilità che la seconda persona che ha prelevato sia estratta alla terza estrazione.

Soluzione

a) La probabilità di estrarre una persona che ha prelevato è pari a $4/5 = 0,8$. Estraendo 10 persone con riposizione si ha che la variabile casuale X che conta il numero di persone che hanno prelevato si distribuisce con legge Binomiale di parametri $p = 0,8$ ed $n = 10$:

$$X \backsim \text{Bin}(n = 10; p = 0,8) \ .$$

$$\begin{aligned} P(X \geq 9) &= P(X = 9) + P(X = 10) \\ &= \binom{10}{9} 0,8^9 0,2^1 + \binom{10}{10} 0,8^{10} \\ &= 0,268 + 0,1073 = 0,3753 \ . \end{aligned}$$

b) Si vuole calcolare

$$P(X > 4|X \geq 3) \quad = \quad \frac{P(X > 4)}{P(X \geq 3)} \ ,$$

dove in questo caso

$$X \backsim \text{Bin}(n = 6; p = 0,8) \ .$$

Calcolando

$$\begin{aligned}
P(X \geq 3) &= 1 - P(X < 3) \\
&= 1 - P(X = 0) - P(X = 1) - P(X = 2) \\
&= 1 - \binom{6}{0}0,2^6 - \binom{6}{1}0,8^1 0,2^5 - \binom{6}{2}0,8^2 0,2^4 \\
&= 1 - 0,000064 - 0,00153 - 0,01536 = 0,9830 \ ;
\end{aligned}$$

e

$$\begin{aligned}
P(X > 4) &= P(X = 5) + P(X = 6) \\
&= \binom{6}{5}0,8^5 0,2^1 + \binom{6}{6}0,8^6 0,2^0 \\
&= 0,393 + 0,262 = 0,655 \ ,
\end{aligned}$$

si ottiene

$$\begin{aligned}
P(X > 4|X \geq 3) &= \frac{P(X > 4)}{P(X \geq 3)} \\
&= \frac{0,655}{0,9830} = 0,666 \ ;
\end{aligned}$$

c) Si consideri la v.c. X che conta fra i 300 estratti quelli che hanno prelevato. Tale variabile si distribuisce con legge di probabilità Binomiale di parametri $n = 300$ e $p = 0,8$, Essendo il numero delle estrazioni molto elevato (> 30) è possibile approssimare X con una variabile casuale Normale Y di aspettativa $E(Y) = 300 \cdot 0,8 = 240$ e di varianza $Var(Y) = 300 \cdot 0,8 \cdot (1 - 0,8) = 48$. La probabilità che $245 < X < 255$ è

$$\begin{aligned}
P(245 < X < 255) &= P(X = 246) + P(X = 247) + \cdots \\
&\quad \cdots + P(X = 253) + P(X = 254) \\
&= P(X < 255) - P(X \leq 245) \\
&\approx \Phi\left(\frac{255 - 240}{\sqrt{48}}\right) - \Phi\left(\frac{245 - 240}{\sqrt{48}}\right) \\
&\approx \Phi\left(2,16\right) - \Phi\left(0,72\right) \\
&\approx 0,9846 - 0,7642 = 0,2204 \ .
\end{aligned}$$

d) Se Y è la variabile casuale che conta il numero di estrazioni necessarie per ottenere una persona che non ha prelevato allora

$$Y \backsim \text{Pascal}(p = 0,2) \ ,$$

e dunque

$$E(Y) = \frac{1}{p} = \frac{1}{0,2} = 5 \ .$$

Quindi il numero medio di estrazioni (sequenziali) per ottenere una persona che non ha prelevato è 5.

e) Se Z è la variabile casuale che conta il numero di estrazioni (sequenziali) necessarie per ottenere 2 persone che hanno prelevato allora

$$Z \backsim \text{Binomiale Negativa}(r = 2; p = 0,8) \ .$$

Quindi la probabilità che la seconda persona che ha prelevato sia estratta alla terza prova è uguale alla probabilità che $Z = 3$:

$$P(Z = 3) = \binom{2}{1} 0,8^2 0,2 = 0,256 \ .$$

Esercizio

Da un mazzo di 52 carte (13 di picche, 13 di cuori, 13 di fiori e 13 di quadri) ne vengono estratte cinque con reinserimento. Si è interessati alla variabile casuale X che descrive il numero di carte di cuori ottenute nelle estrazioni. Determinare:

a) il valore atteso e la varianza della variabile X;

b) la probabilità di estrarre tre carte di cuori;

c) la probabilità di estrarre almeno tre carte di cuori;

d) la probabilità di estrarre al più tre carte di cuori;

Soluzione

a) Dato che l'estrazione delle carte dal mazzo di 52 avviene con riposizione, la probabilità di ottenere una carta di cuori rimane costante da estrazione ad estrazione. Le singole estrazioni sono inoltre indipendenti in quanto fisicamente separate. Alla luce di ciò, la variabile casuale X risulta essere una variabile casuale binomiale di parametri $n = 5$ e $p = \frac{13}{52} = \frac{1}{4}$. Si ha dunque che:

$$E(X) = n \cdot p = 5 \cdot 0.25 = 1.25 \ \ ;$$

$$Var(X) = n \cdot p \cdot (1 - p) = 5 \cdot 0.25 \cdot 0.75 = 0.9375 \ \ .$$

b) Per quanto appena osservato, la funzione di probabilità della variabile casuale X è data da:

$$P(X=x) = \begin{cases} \binom{5}{x}(0.25)^x(1-0.25)^{5-x} & x = 0,1,2,3,4,5 \\ \\ 0 & altrove \end{cases}$$

Si ha dunque che:

$$\begin{aligned} P(X=3) &= \binom{5}{3}(0.25)^3(0.75)^{5-3} \\ &= \frac{5!}{3!\cdot 2!}\cdot(0.0156)\cdot(0.5625) \\ &= 10\cdot(0.0156)\cdot(0.5625) = 0.0879 \quad . \end{aligned}$$

c) La probabilità di estrarre almeno tre carte di cuori coincide con la probabilità che la variabile casuale X assuma valori maggiori o uguali a 3.
Si ha quindi:

$$\begin{aligned} P(X\geq 3) &= P(X=3)+P(X=4)+P(X=5) \\ &= \binom{5}{3}(0.25)^3(0.75)^{5-3}+\binom{5}{4}(0.25)^4(0.75)^{5-4}+\binom{5}{5}(0.25)^5(0.75)^{5-5} \\ &= 0.0879+5\cdot(0.25)^4\cdot(0.75)+(0.25)^5 \\ &= 0.0879+0.0146+0.0010 = 0.1035 \quad . \end{aligned}$$

In alternativa, il medesimo valore si sarebbe potuto ottenere seguendo il seguente procedimento:

$$\begin{aligned} P(X\geq 3) &= 1-P(X<3) = 1-(P(X=0)+P(X=1)+P(X=2)) \\ &= 1-\left(\binom{5}{0}(0.25)^0(0.75)^5+\binom{5}{1}(0.25)(0.75)^4+\binom{5}{2}(0.25)^2(0.75)^3\right) \\ &= 1-(0.75)^5-5\cdot(0.25)\cdot(0.75)^4-10\cdot(0.25)^2\cdot(0.75)^3 \\ &= 1-0.2373-0.3955-0.2637 = 1-0.8965 = 0.1035 \quad . \end{aligned}$$

d) La probabilità di estrarre al più tre carte di cuori coincide con la probabilità che la variabile casuale X assuma valori minori o uguali a 3:

$$\begin{aligned} P(X\leq 3) &= P(X=0)+P(X=1)+P(X=2)+P(X=3) \\ &= 0.2373+0.3955+0.2637+0.0879 = 0.9844 \\ &= 1-[P(X=4)+P(X=5)] \quad . \end{aligned}$$

Esercizio

Il 19% dei nuclei familiari di una collettività possiede almeno due televisori. Si estrae un campione casuale (con reimmissione) di 400 nuclei familiari. Considerata la variabile casuale X=*'numero dei nuclei familiari estratti che possiedono almeno due televisori'*:

a) si calcolino aspettativa e varianza di X;

b) si calcoli $P(66 \leq X \leq 91)$.

Soluzione

a) L'esito dell'estrazione di un nucleo familiare dalla collettività in considerazione, può essere interpretato utilizzando una variabile casuale indicatore I di parametro $p = 0.19$. In particolare, tale variabile casuale assume valore 1 nel caso in cui la famiglia estratta possiede almeno due televisori, assume valore 0 se la famiglia estratta non possiede almeno due televisori. Si ha inoltre che:

$$P(I = 1) = 0.19 \qquad\qquad P(I = 0) = 1 - 0.19 = 0.81 \quad .$$

Si indichi con I_j la variabile indicatore che interpreta l'esito della j-esima estrazione ($j = 1, 2, ..., 400$). Dato che le estrazioni avvengono con riposizione si ha che le variabili casuali $I_1, I_2, ..., I_{400}$ hanno la stessa distribuzione di probabilità (sono tutte v.c. indicatori di parametro $p = 0.19$) e sono indipendenti in quanto le estrazioni sono fisicamente separate. In questa prospettiva, la v.c.

X = "numero dei nuclei familiari estratti che possiedono almeno due televisori"

è data dalla somma delle variabili indicatori $I_1, I_2, ..., I_{400}$:

$$X = \sum_{j=1}^{400} I_j \quad .$$

Dato che le v.c. $I_1, I_2, ..., I_{400}$ sono identicamente ed indipendentemente distribuite, si ha che X risulta essere una binomiale di parametri $n = 400$ e $p = 0.19$. Si ha dunque che:

$$E(X) = n \cdot p = 400 \cdot 0.19 = 76$$

$$Var(X) = n \cdot p \cdot (1 - p) = 400 \cdot 0.19 \cdot 0.81 = 61.56 \quad .$$

b) Come si è osservato durante lo svolgimento del punto precedente, la variabile casuale X può essere vista come la somma di 400 v.c. che godono delle seguenti caratteristiche:

- $I_1, I_2, ..., I_{400}$ hanno la medesima distribuzione di probabilità;
- $I_1, I_2, ..., I_{400}$ sono indipendenti;
- $I_1, I_2, ..., I_{400}$ hanno media e varianza finite e rispettivamente pari a 0.19 e $0.19 \cdot 0.81 = 0.1538$.

Tali risultati ed il fatto che $n = 400 > 30$ ci permettono di dire, in forza del teorema del limite centrale, che

$$P\left(\frac{\sum_{j=1}^{400} I_j - (400 \cdot 0.19)}{\sqrt{400 \cdot (0.19) \cdot (0.81)}} \leq z\right) = P\left(\frac{X - 76}{\sqrt{61.56}} \leq z\right) \cong P\left(Z \leq z\right) = \Phi(z)$$

dove Z indica una variabile casuale normale standardizzata. Sfruttando tale approssimazione, si ha che:

$$\begin{aligned} P(66 \leq X \leq 91) &= P\left(\frac{66-76}{\sqrt{61.56}} \leq \frac{X-76}{\sqrt{61.56}} \leq \frac{91-76}{\sqrt{61.56}}\right) \\ &\cong P\left(\frac{66-76}{\sqrt{61.56}} \leq Z \leq \frac{91-76}{\sqrt{61.56}}\right) \\ &= \Phi\left(\frac{91-76}{\sqrt{61.56}}\right) - \Phi\left(\frac{66-76}{\sqrt{61.56}}\right) \\ &= \Phi(1.9118) - \Phi(-1.2745) \end{aligned}$$

Dalle tavole della normale si ottiene che:

$$\Phi(1.91) = 0.9719 \quad .$$

In forza della simmetria della distribuzione normale si ha inoltre che:

$$\Phi(-1.27) = 1 - \Phi(1.27) = 1 - 0.8980 = 0.1020 \quad .$$

Concludendo si ha:

$$P(66 \leq X \leq 91) \cong 0.9719 - 0.1020 = 0.8699 \quad .$$

Esercizio

Un macchina produce pezzi difettosi con una probabilità pari a 0.4. Calcolare:

a) la probabilità che su 6 pezzi prodotti il numero di pezzi difettosi sia un numero dispari;

b) il valore atteso e la varianza della variabile che conta il numero di pezzi difettosi su 100 prodotti e di quella che li conta su 1000 pezzi prodotti;

c) la probabilità che su 100 pezzi prodotti, i pezzi difettosi siano in numero compreso tra 15 e 30 (estremi inclusi).

Soluzione

a) Dal contesto descritto dal testo dell'esercizio si ha che:

- la probabilità che la macchina produca pezzi difettosi rimane costante nel tempo;
- la macchina produce pezzi difettosi in maniera casuale.

Le caratteristiche (difettosità$\backslash$non difettosità) di ciascun pezzo prodotto possono essere interpretate utilizzando una variabile casuale indicatore di parametro $p = 0.4$. Si considerino ora le variabili casuali indicatori associate a 6 pezzi prodotti. Le osservazioni sopra riportate implicano che queste 6 variabili casuali sono identicamente ed indipendentemente distribuite e, di conseguenza, la variabile casuale $X =$"*numero di pezzi difettosi su 6 prodotti*", è una binomiale di parametri $n = 6$ e $p = 0.4$. Si ha quindi che:

$$\begin{aligned}
& P(\text{"su 6 pezzi prodotti il numero di prodotti difettosi è dispari"}) \\
= \; & P(X = 1) + P(X = 3) + P(X = 5) \\
= \; & \binom{6}{1} 0.4^1 (1 - 0.4)^5 + \binom{6}{3} 0.4^3 (1 - 0.4)^3 + \binom{6}{5} 0.4^5 (1 - 0.4)^1 \\
= \; & 6 \cdot 0.4 \cdot 0.6^5 + 20 \cdot 0.4^3 \cdot 0.6^3 + 6 \cdot 0.4^5 \cdot 0.6^1 \\
= \; & 0.1866 + 0.2765 + 0.0369 = 0.5 \quad .
\end{aligned}$$

b) La variabile casuale $Y =$ "n° di pezzi difettosi su 100 pezzi prodotti" è una binomiale di parametri $n = 100$ e $p = 0.4$. Si ha dunque che:

$$E(Y) = n \cdot p = 100 \cdot 0.4 = 40 \quad ;$$

$$Var(Y) = n \cdot p \cdot q = 100 \cdot 0.4 \cdot 0.6 = 24 \quad .$$

La variabile casuale $K =$ "n° di pezzi difettosi su 1000 pezzi prodotti" è una binomiale di parametri $n = 1000$ e $p = 0.4$. Si ha dunque che:

$$E(K) = n \cdot p = 1000 \cdot 0.4 = 400 \quad ;$$

$$Var(K) = n \cdot p \cdot q = 1000 \cdot 0.4 \cdot 0.6 = 240 \quad .$$

c) Dato che $n = 100$ può ritenersi sufficientemente elevato, in forza del teorema del limite centrale si ha che la distribuzione della variabile Y è approssimabile con la distribuzione normale di media 40 e varianza 24. Si ha dunque che:

$$\begin{aligned} P(15 \leq Y \leq 30) &= P\left(\frac{15-40}{\sqrt{24}} \leq \frac{Y-40}{\sqrt{24}} \leq \frac{30-40}{\sqrt{24}}\right) \\ &\cong P\left(\frac{15-40}{\sqrt{24}} \leq Z \leq \frac{30-40}{\sqrt{24}}\right) \\ &= P\left(-5.1031 \leq Z \leq -2.0412\right) = \Phi(-2.0412) - \Phi(-5.1031) \ . \end{aligned}$$

In forza della simmetria della distribuzione normale, si ha che:

$$\Phi(-2.04) = 1 - \Phi(2.04) = 1 - 0.9793 = 0.0207 \ ;$$

$$\Phi(-5.10) = 1 - \Phi(5.10) = 1 - 1 = 0 \ .$$

Si ha dunque che:

$$P(15 \leq Y \leq 30) \cong \Phi\left(-2.04\right) - \Phi\left(-5.10\right) = 0.0207 \ .$$

Esercizio

Sia X la variabile casuale che descrive il numero di teste ottenute in 100 lanci di una moneta regolare. Determinare

$$\Pr\{|X - \mu| \leq 8\}$$

dove μ è il valore atteso di X.

Soluzione

Ogni lancio della moneta viene ripetuto sotto le stesse condizioni in modo tale che: le probabilità di ottenere testa o croce non variano da lancio a lancio e i lanci della moneta sono indipendenti.

La variabile casuale X è di conseguenza una binomiale di parametri $n = 100$ e $p = 0.5$. Alla luce di ciò si ha che:

$$\mu = E(X) = n \cdot p = 100 \cdot 0.5 = 50 \ ;$$

$$Var(X) = n \cdot p \cdot (1-p) = 100 \cdot 0.5 \cdot 0.5 = 25 \ .$$

Dato che $n = 100$ può ritenersi sufficientemente elevato, in forza del teorema del limite centrale, si ha che la distribuzione della variabile X è approssimabile con la distribuzione normale di media 50 e varianza 25. Si ha dunque che:

$$
\begin{aligned}
P(|X-\mu| \leq 8) &= P(-8 \leq (X-\mu) \leq 8) \\
&= P(-8 \leq (X-50) \leq 8) = P\left(\frac{-8}{\sqrt{25}} \leq \frac{X-50}{\sqrt{25}} \leq \frac{8}{\sqrt{25}}\right) \\
&\cong P\left(-1.6 \leq Z \leq 1.6\right) \\
&= \Phi(1.6) - \Phi(-1.6) = \Phi(1.6) - (1-\Phi(1.6)) \\
&= 2 \cdot \Phi(1.6) - 1 = 2 \cdot 0.9452 - 1 = 1.8904 - 1 = 0.8904 \quad .
\end{aligned}
$$

Esercizio

Sia *X* una variabile casuale indicatore di parametro $p = 0.6$ e sia *Y* una variabile casuale binomiale di parametri $n = 2$ e $p = 0.4$.

a) si esplicitino le distribuzioni di *X* e *Y*;

b) si determini la distribuzione di probabilità congiunta di *X* e *Y* nell'ipotesi di indipendenza;

c) si calcolino aspettativa e varianza della variabile casuale $Z = X + Y$.

Soluzione

a) La variabile casuale X è un indicatore di parametro $p = 0.6$ e di conseguenza:

$$
P(X = x) = \begin{cases} 0.6 & x = 1 \\ 0.4 & x = 0 \end{cases} \quad .
$$

La variabile casuale Y è una binomiale di parametri $n = 2$ e $p = 0.4$ e di conseguenza:

$$P(Y=y) = \begin{cases} \binom{2}{0} \cdot 0.4^0 \cdot (1-0.4)^2 & y=0 \\ \\ \binom{2}{1} \cdot 0.4^1 \cdot (1-0.4)^1 & y=1 \\ \\ \binom{2}{2} \cdot 0.4^2 \cdot (1-0.4)^0 & y=2 \\ \\ 0 & \text{altrove} \end{cases}$$

$$= \begin{cases} (1-0.4)^2 & y=0 \\ \\ 2 \cdot 0.4 \cdot (1-0.4) & y=1 \\ \\ 0.4^2 & y=2 \\ \\ 0 & \text{altrove} \end{cases}$$

$$= \begin{cases} 0.36 & y=0 \\ \\ 0.48 & y=1 \\ \\ 0.16 & y=2 \\ \\ 0 & \text{altrove} \end{cases}$$

b) Se le variabili casuali X e Y sono indipendenti allora:

$$P(X=x, Y=y) = P(X=x) \cdot P(Y=y) \qquad x=0,1\ ; \quad y=0,1,2\ .$$

Si ha dunque che:

- $P(X=0, Y=0) = P(X=0) \cdot P(Y=0) = 0.6 \cdot 0.36 = 0.216$
- $P(X=0, Y=1) = P(X=0) \cdot P(Y=1) = 0.6 \cdot 0.48 = 0.288$
- $P(X=0, Y=2) = P(X=0) \cdot P(Y=2) = 0.6 \cdot 0.16 = 0.096$
- $P(X=1, Y=0) = P(X=1) \cdot P(Y=0) = 0.4 \cdot 0.36 = 0.144$
- $P(X=1, Y=1) = P(X=1) \cdot P(Y=1) = 0.4 \cdot 0.48 = 0.192$
- $P(X=1, Y=2) = P(X=1) \cdot P(Y=2) = 0.4 \cdot 0.16 = 0.064$

Nella seguente tabella viene riportata la distribuzione congiunta di X e Y in ipotesi di indipendenza:

$X \backslash Y$	0	1	2	tot
0	0.216	0.288	0.096	0.6
1	0.144	0.192	0.064	0.4
tot	0.36	0.48	0.16	1

c) Si ha che:

- $E(X) = 0.6$;
- $Var(X) = 0.6 \cdot 0.4 = 0.24$;
- $E(Y) = 2 \cdot 0.4 = 0.8$;
- $Var(Y) = 2 \cdot 0.4 \cdot 0.6 = 0.48$;

Per quanto riguarda il calcolo di $E(Z)$ si ha:

$$E(Z) = E(X+Y) = E(X) + E(Y) = 0.6 + 0.8 = 1.4 \quad .$$

Per quanto riguarda il calcolo della varianza di Z si ha:

$$Var(Z) = Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y) \quad .$$

Come si osserva dalla formula appena riportata, non è possibile calcolare la varianza della variabile Z senza conoscere la distribuzione congiunta di X e Y. Supponendo che X e Y siano indipendenti, e dunque $Cov(X,Y) = 0$, si ha che:

$$Var(Z) = 0.24 + 0.48 = 0.72 \quad .$$

4.2 Ipergeometrica

Esercizio

Un commerciante acquista una partita di diamanti confezionata in tre scatole ognuna da un venditore diverso. Tutte e tre le scatole contengono 10 diamanti ciascuna. Nella scatola confezionata dal venditore A c'è una pietra falsa, in quella del venditore B ci sono 4 pietre false, mentre nella scatola del venditore C ci sono 2 pietre false. Il commerciante sceglie a caso una delle tre scatole ed estrae in blocco 3 diamanti:

a) si calcoli la probabilità che nessuno dei diamanti estratti sia falso.

b) Avendo estratto un campione in blocco di 3 diamanti non falsi, si calcoli la probabilità che esso provenga dalla scatola confezionata dal venditore B.

c) Si supponga che il commerciante abbia scelto la scatola del venditore B. Se le estrazioni dei 3 diamanti avvenissero con riposizione, quale sarebbe la probabilità che il terzo diamante estratto fosse il primo a risultare falso?

Soluzione

a) Si indichino con A, B e C rispettivamente gli eventi "scegliere la scatola A", "scegliere la scatola B" e "scegliere la scatola C". Si indichino con X_A, X_B ed X_C le variabili casuali che contano il numero di pietre false estratte rispettivamente dalla scatola A, B e C. Sia inoltre X la variabile casuale che conta il numero di pietre false estratte con lo schema di campionamento descritto nell'esercizio. Allora

$$\begin{aligned} P(X=0) &= P(X_A=0|A)\cdot P(A)+P(X_B=0|B)\cdot P(B)+ \\ &\quad +P(X_C=0|C)\cdot P(C)\ . \end{aligned} \tag{10}$$

Scegliendo a caso una delle tre scatole si avrà $P(A)=P(B)=P(C)=1/3$. Inoltre, essendo le estrazioni senza reinserimento, si ha che

$$\begin{aligned} X_A &\backsim \text{Ipergeometrica}(N=10; r=1; n=3)\ , \\ X_B &\backsim \text{Ipergeometrica}(N=10; r=4; n=3)\ , \\ X_C &\backsim \text{Ipergeometrica}(N=10; r=2; n=3)\ . \end{aligned}$$

Precisato ciò è possibile ricavare la probabilità in (10) nel seguente modo:

$$\begin{aligned} P(X=0) &= \frac{\binom{9}{3}\binom{1}{0}}{\binom{10}{3}}\cdot\frac{1}{3}+\frac{\binom{6}{3}\binom{4}{0}}{\binom{10}{3}}\cdot\frac{1}{3}+\frac{\binom{8}{3}\binom{2}{0}}{\binom{10}{3}}\cdot\frac{1}{3} \\ &= \frac{7}{30}+\frac{1}{18}+\frac{7}{45}=0,444 \end{aligned}$$

b)

$$P(B|X=0) = \frac{P[B\cap(X=0)]}{P(X=0)}=\frac{\frac{1}{18}}{0,444}=0,125$$

c) Indicando con Y la v.c. che conta il numero di estrazioni (con riposizione) necessarie per ottenere una pietra falsa allora

$$Y \backsim \text{Pascal}(p=\frac{4}{10}=0,4)\ ,$$

la probabilità di ottenere la prima pietra falsa alla terza estrazione è pari a

$$P(Y=3)=(1-0,4)^{3-1}0,4=0,144\ .$$

4.3 Poisson

Esercizio

Il numero di meteoriti che colpisce un satellite durante ogni sua orbita si distribuisce come una variabile casuale di Poisson di parametro λ. Nel compiere la sua orbita il satellite impiega 1 giorno, ed è mediamente colpito da 3 meteoriti.

a) Si calcoli la probabilità che nel percorrere 5 orbite il numero di meteoriti che colpiscono il satellite sia minore o uguale a 3.

b) Supponendo invece $\lambda = 0,5$ si fornisca il valore atteso e la varianza della v.c. Y che misura il tempo (in giorni) che intercorre fra l'impatto di un meteorite e quello successivo.

Soluzione

a) Indicando con X la variabile casuale che conta il numero di meteoriti che colpiscono il satellite in un giorno si ha che $E(X) = \lambda_X = 3$ pertanto

$$X \backsim \text{Poisson}(\lambda_X = 3) \; .$$

Indicando con W la v.c. che conta il numero di meteoriti che colpiscono il satellite in 5 giorni si ha che

$$W \backsim \text{Poisson}(\lambda_W = 5 \cdot 3 = 15) \; ,$$

perciò

$$\begin{aligned} P(W \leq 3) &= P(W = 0) + P(W = 1) + P(W = 2) + P(W = 3) \\ &= \frac{e^{-15}15^0}{0!} + \frac{e^{-15}15^1}{1!} + \frac{e^{-15}15^2}{2!} + \frac{e^{-15}15^3}{3!} \\ &= e^{-15}\left(1 + 15 + \frac{15^2}{2} + \frac{15^3}{6}\right) = 0,0002 \; . \end{aligned}$$

b) Se Y è la v.c. che misura il tempo (in giorni) tra 2 impatti successivi allora

$$X \backsim \text{Esponenziale}(\theta) \; ,$$

dove $\lambda = \theta \cdot t$ in cui $\lambda = 0,5$ è il parametro di $X \backsim \text{Poisson}(\lambda = 0,5)$ v.c. che conta il numero di meteoriti in un giorno. Pertanto

$$\lambda = \theta \cdot \underbrace{t}_{\text{1 giorno}} \rightarrow \lambda = \theta = 0,5 \; ,$$

da cui deriva

$$E(Y) = \frac{1}{\theta} = \frac{1}{0,5} = 2 \; , \qquad Var(Y) = \frac{1}{\theta^2} = \frac{1}{0,5^2} = 4 \; .$$

4.4 Pascal & Binomiale Negativa

Esercizio

Un negozio di videonoleggio mette in vendita 20 DVD usati di un certo film, 4 dei quali presentano problemi di audio.

a) Estratti casualmente 3 DVD senza riposizione, qual è la probabilità che soltanto il terzo estratto presenti problemi di audio.

b) Considerando ora estrazioni sequenziali con riposizione, si determini la probabilità che:

- il primo DVD con problemi di audio sia il quinto estratto;
- il secondo DVD con problemi di audio sia il settimo estratto.

Soluzione

a) Indicando con A l'evento "il DVD presenta problemi di audio" risulta che la probabilità che solo il terzo presenti la caratteristica A è la seguente

$$P(\bar{\mathrm{A}} \cap \bar{\mathrm{A}} \cap \mathrm{A}) = \frac{16}{20} \cdot \frac{15}{19} \cdot \frac{4}{18} = 0,1404$$

b) Sia

$$Y \backsim \text{Pascal}(p = \frac{4}{20} = 0,2) \ ,$$

pertanto

$$P(Y = 5) = 0,8^4 \cdot 0,2 = 0,08192 \ .$$

Inoltre sia

$$W \backsim \text{Binomiale Negativa}(r = 2; p = 0,2) \ ,$$

dunque

$$P(W = 7) = \binom{6}{1} 0,2^2 \cdot 0,8^5 = 0,0786 \ .$$

4.5 Trinomiale

Esercizio

L'ufficio ricerche di mercato chiede ad un campione di 12 persone se preferiscono un nuovo prodotto al vecchio. Si suppone che il 30% della popolazione preferisce il vecchio tipo di prodotto, il 50% preferisce il nuovo prodotto ed il 20% è indifferente:

a) si determini la probabilità che nel campione ci siano 8 clienti che preferiscono il nuovo prodotto e 2 che preferiscono il vecchio.

b) Sapendo che 3 clienti preferiscono il vecchio tipo, si determini la probabilità che 5 clienti preferiscano il nuovo tipo.

c) Si calcoli il valore atteso e la varianza del numero di clienti che sono indifferenti e la covarianza fra il numero di clienti che preferiscono il vecchio tipo e quelli che preferiscono il nuovo tipo.

Soluzione

a) Siano X_1 ed X_2 le v.c. che descrivono il numero di persone (nel campione di 12 unità statistiche) che preferiscono rispettivamente il nuovo ed il vecchio prodotto. Dunque si ha che

$$(X_1, X_2) \backsim \text{Trinomiale}(n = 12; p_1 = 0,5; p_2 = 0,3) \ ,$$

pertanto

$$\begin{aligned} P\left(X_1 = 8, X_2 = 2\right) &= \frac{12!}{8! \cdot 2! \cdot 2!} 0,5^8 \cdot 0,3^2 \cdot 0,2^2 \ , \\ &= \frac{12 \cdot 11 \cdot 10 \cdot 9 \cdot 8!}{8! \cdot 2! \cdot 2!}(0,000014062) = 0,0418 \ . \end{aligned}$$

b) Fissato che si conosca il numero di persone nel campione che preferiscono il vecchio tipo, la v.c. che conta il numero di persone che preferiscono il nuovo tipo si distribuisce con legge binomiale:

$$(X_1 | X_2 = n_2) \backsim \text{Binomiale}\left(n - n_2; \frac{p_1}{1 - p_2}\right) \ ,$$

dove

$$\frac{p_1}{1 - p_2} = \frac{0,5}{1 - 0,3} = 0,7143 \ .$$

Si ha quindi

$$(X_1 | X_2 = 3) \backsim \text{Binomiale}\left(n = 9; p = 0,7143\right) \ .$$

La probabilità che $X_1 = 5$ dato $X_2 = 3$ risulta allora

$$\begin{aligned} P\left(X_1 = 5 | X_2 = 3\right) &= \binom{9}{5} 0,7143^5 \cdot (1 - 0,7143)^4 \\ &= \frac{9!}{5! \cdot 4!}(0,0012389) = 0,1561 \ . \end{aligned}$$

c) Sia X_3 la v.c. che descrive il numero di persone (nel campione di 12 unità statistiche) che sono indifferenti ai due prodotti. Naturalmente

$$X_3 \backsim \text{Binomiale}\,(n = 12; p = 0,2)\ ,$$

pertanto

$$\begin{aligned} E(X_3) &= n \cdot p \\ &= 12 \cdot 0,2 = 2,4\ , \end{aligned}$$

e

$$\begin{aligned} Var(X_3) &= n \cdot p \cdot q \\ &= 12 \cdot 0,2 \cdot 0,8 = 1,92\ . \end{aligned}$$

La covarianza tra X_1 ed X_2 è

$$\begin{aligned} Cov(X_1, X_2) &= n \cdot (n-1) \cdot p_1 \cdot p_2 - n \cdot p_1 \cdot n \cdot p_2 \\ &= -n \cdot p_1 \cdot p_2 \\ &= -12 \cdot 0,5 \cdot 0,3 = -1,8\ . \end{aligned}$$

Da osservare che la covarianza è negativa tra le due variabili poiché esiste una dipendenza lineare negativa fra le stesse. Più sono alti (bassi) i valori che assume X_1 e più sono bassi (alti) i valori che assume X_2. Questa relazione è evidente dato che, essendo il numero delle estrazioni fissate (12 in questo caso), più sono le persone che preferiscono il prodotto nuovo meno sono le persone che preferiscono il prodotto vecchio (e viceversa naturalmente).

5 Variabili casuali continue

5.1 Pareto

Esercizio

Il fatturato delle maggiori imprese meccaniche italiane (espresso in migliaia di Euro) può essere descritto da una v.c. X avente distribuzione di Pareto $\theta = 1,2$. È noto inoltre che il 20% di queste imprese ha fatturato superiore a 2 milioni di Euro.

a) Si specifichi la funzione di densità di X dopo avere determinato il parametro x_0 (arrotondare il valore di x_0 all'unità più vicina).

b) Si determinino il valore atteso e la varianza di X, commentando opportunamente i risultati ottenuti.

c) Si calcolino la probabilità che un'impresa abbia fatturato compreso fra $1,5$ e 2 milioni di Euro.

Soluzione

a) L'unico parametro incognito è x_0 cioè

$$X \backsim \text{Pareto}(x_0 = ?; \theta = 1,2) \ .$$

Il parametro incognito si ricava dalla seguente relazione

$$P(X > 2000) = 0,2 \ ,$$

dunque si sfrutta la funzione di ripartizione di X:

$$\left(\frac{x_0}{2000}\right)^{1,2} = 0,2 \rightarrow x_0 = 0,2^{\frac{1}{1,2}} \cdot 2000 \cong 533 \ ,$$

da cui deriva

$$f_X(x; x_0, \theta) = \begin{cases} 2194,75 \cdot x^{-2,2} & x \geq 523 \\ 0 & \text{altrove} \ . \end{cases}$$

b) Il valore atteso (momento primo) di X risulta

$$E(X) = x_0 \cdot \frac{\theta}{\theta - 1} = 523 \cdot \frac{1,2}{0,2} = 3138 \ ,$$

mentre il momento secondo $E(X^2)$ non esiste dal momento che $2 > \theta$. Perciò non esiste nemmeno la varianza che sarebbe uguale al momento secondo meno il momento primo al quadrato:

$$Var(X) = E(X^2) - [E(X)]^2 \ .$$

c) La probabilità si può ottenere sfruttando la funzione di ripartizione di X:

$$\begin{aligned} P(1500 \leq X \leq 2000) &= F_X(2000) - F_X(1500) \\ &= 1 - \left(\frac{523}{2000}\right)^{1,2} - \left[1 - \left(\frac{523}{1500}\right)^{1,2}\right] \\ &= -0,2 + 0,2824 = 0,0824 \ , \end{aligned}$$

oppure calcolando direttamente l'integrale

$$\begin{aligned} P(1500 \leq X \leq 2000) &= \int_{1500}^{2000} f(x) dx \\ &= \int_{1500}^{2000} 2194,75 \cdot x^{-2,2} dx \\ &= \ldots \end{aligned}$$

5.2 Esponenziale

Esercizio

La direzione di un supermercato deve acquistare un nuovo freezer per l'esposizione dei surgelati in vendita. La durata di funzionamento in mesi di un primo modello di freezer per supermercato è descritta da una v.c. X con distribuzione esponenziale di parametro $\theta = 0,01$.

a) Si valutino la durata attesa e quella mediana del freezer.

b) Si determini la probabilità che la durata del freezer sia superiore a 200 mesi.

c) Nell'ipotesi che il freezer sia già funzionante da 80 mesi, si calcoli la probabilità che questo funzioni per altri 200 mesi e la si confronti con la probabilità ottenuta al punto precedente.

d) Se la direzione del supermercato deve scegliere fra il freezer di cui sopra ed un secondo modello di freezer che ha durata in mesi che si può descrivere con una v.c. di Pareto di parametri $x_0 = 50$ e $\theta = 2$, quale modello deve scegliere per avere il maggiore valore atteso della durata?

Soluzione

a) Il valore atteso della v.c. X risulta

$$E(X) = \frac{1}{\theta} = \frac{1}{0,01} = 100 \ ,$$

pertanto il freezer dura mediamente (senza rompersi) 100 mesi. La mediana relativa alla v.c. X risulta

$$Me(X) = \frac{1}{\theta} \cdot \ln 2 = E(X) \cdot 0,6931 = 69,31 \ ,$$

dunque il valore x tale per cui $P(X \leq x) = 0,5$ è pari a $69,31$. Si ricordi che la mediana è quel valore x tale per cui la probabilità che X assuma un valore inferiore ad x è la stessa della probabilità che X assuma un valore maggiore di x.

b) Semplicemente

$$
\begin{aligned}
P(X > 200) &= \int_{200}^{+\infty} 0,01 e^{-0,01\cdot x} dx \\
&= 1 - \int_{0}^{200} 0,01 e^{-0,01\cdot x} dx \\
&= 1 - \underbrace{P(X \leq 200)}_{F_X(200)} \\
&= 1 - \left[1 - e^{-0,01\cdot 200}\right] = e^{-2} = 0,135 \,.
\end{aligned}
$$

c) La probabilità condizionata risulta

$$
\begin{aligned}
P(X > 280 | X > 80) &= \frac{P(X > 280 \cap X > 80)}{P(X > 80)} \\
&= \frac{P(X > 280)}{P(X > 80)} \\
&= \frac{e^{-0,01\cdot 280}}{e^{-0,01\cdot 80}} \\
&= e^{-0,01\cdot 200} \\
&= P(X > 200) = 0,135 \,.
\end{aligned}
$$

Le due probabilità $P(X > 280|X > 80)$ e $P(X > 200)$ risultano uguali. In altre parole la probabilità che il freezer non si rompa per altri 200 mesi dato che ha funzionato per 80 mesi è la stessa che esso funzioni continuativamente per 200 mesi dalla prima accensione. L'"assenza di memoria" è la proprietà della v.c. esponenziale tale per cui risulta

$$
P(X > x_2 | X > x_1) = P(X > x_2 - x_1) \qquad \text{con } x_2 > x_1 \,.
$$

d) Come calcolato al punto a) risulta $E(X) = 100$. Se Y è la durata del secondo freezer che si distribuisce come una Pareto di parametri $x_0 = 50$ e $\theta = 2$ allora

$$
E(Y) = x_0 \frac{\theta}{\theta - 1} = 50 \cdot \frac{2}{2-1} = 100 \,.
$$

In termini di valore atteso di durata, è indifferente scegliere il primo piuttosto che il secondo freezer dato che $E(X) = E(Y)$.

5.3 Normale

Esercizio

Sia X una variabile casuale normale che descrive la portata del fiume Adige nel mese di Giugno nella località di Boara Pisani. È noto che in tale località la portata media del fiume nel mese di giugno è di 243 metri cubici al secondo e che $P(X < 400) = 0,9$.

a) Si determini la varianza di X.

b) Supponendo che $\sigma^2 = 100$, si determini la probabilità che in giugno il fiume Adige abbia una portata compresa fra 230 e 260 metri cubici al secondo.

c) Si supponga che $Y = 10 + X$ sia la v.c. che descrive la portata del fiume Adige nella località di Boscochiaro. Supponendo che $\sigma_X^2 = 100$, si calcoli $P(240 < Y < 270)$ e si confronti il valore ottenuto con quello ricavato al punto b).

Soluzione

a) La varianza va ricavata dalla seguente relazione

$$P(X < 400) = \Phi\left(\frac{400 - 243}{\sigma}\right) = 0,9$$

da cui si evince che

$$\frac{400 - 243}{\sigma} = 1,285$$

dal momento che il nono decile relativo alla v.c. normale standard è $z_{0,9} = 1,285$. Si ha dunque che lo scarto quadratico medio è $\sigma = 122,17$ da cui si ottiene la varianza di X che risulta pari a $Var(X) = \sigma^2 = 14927$.

b) Supponendo che $\sigma^2 = 100$ e dunque $\sigma = 10$, la probabilità richiesta si calcola come segue

$$\begin{aligned}
P(230 \leq X \leq 260) &= \int_{230}^{260} f(x)dx \\
&= F_X(260) - F_X(230) \\
&= \Phi\left(\frac{260 - 243}{10}\right) - \Phi\left(\frac{230 - 243}{10}\right) \\
&= \Phi(1,7) - \Phi(-1,3) \\
&= \Phi(1,7) - [1 - \Phi(1,3)] \\
&= 0,9554 - 1 + 0,9032 = 0,8586\ .
\end{aligned}$$

c) Essendo X una v.c. normale ed essendo Y una trasformazione lineare di X allora risulta che anche Y è una v.c. normale. In particolare risulta che

$$E(Y) = E(10 + X) = 10 + E(X) = 10 + 243 = 253$$

e

$$Var(Y) = Var(10 + X) = Var(X) = 100 \; .$$

Con questi elementi è possibile calcolare la probabilità $P(240 \leq Y \leq 270)$ (come fatto al punto b) per la variabile casuale X) sfruttando il valore atteso e la varianza di Y appena calcolati. In alternativa si può osservare che

$$\begin{aligned} P(240 \leq Y \leq 270) &= P(240 \leq 10 + X \leq 270) \\ &= P(230 \leq X \leq 260) \; , \end{aligned}$$

e quindi è possibile concludere che $P(240 \leq Y \leq 270) = 0,8586$.

5.4 Log-Normale

Esercizio

Avendo selezionato dalla popolazione di una regione italiana un gruppo di uomini adulti coniugati si è verificato che la distribuzione del reddito mensile segue una legge lognormale con parametri $\gamma = 6,109$ e $\delta = 0,699$.

a) Si calcoli la probabilità che, estraendo un individuo dal gruppo, questi abbia un reddito mensile compreso fra 1300 e 2600 Euro.

b) Si calcoli il valore atteso e la varianza del reddito.

c) Nella medesima regione, si è osservato che il logaritmo del reddito mensile delle mogli ha distribuzione normale con valore atteso pari a $5,3$ e varianza pari a $0,9$. Si calcoli il valore atteso e la varianza del logaritmo del reddito complessivo del gruppo di coniugi supponendo che la correlazione fra il logaritmo del reddito dei mariti e delle mogli del gruppo sia nulla.

Soluzione

a) Posta X la v.c. che descrive il reddito mensile degli uomini adulti coniugati si ha che

$$\begin{aligned} P(1300 \leq X \leq 2600) &= \int_{1300}^{2600} f(x)dx \\ &= F_X(2600) - F_X(1300) \\ &= \Phi\left(\frac{\ln 2600 - 6,109}{0,699}\right) - \Phi\left(\frac{\ln 1300 - 6,109}{0,699}\right) \\ &= \Phi\left(2,510\right) - \Phi\left(1,518\right) \\ &= 0,99396 - 0,93574 = 0,058 \ . \end{aligned}$$

b) Il valore atteso e la varianza di X risultano rispettivamente

$$E(X) = e^{\gamma + \frac{1}{2}\delta^2} = e^{6,109 + \frac{0,699^2}{2}} = 574,385 \ ,$$

$$Var(X) = e^{2\gamma + \delta^2}\left(e^{\delta^2} - 1\right) = 207859,954 \ .$$

c) Si indichi con Y la v.c. che descrive il logaritmo del reddito mensile delle mogli degli adulti coniugati del gruppo in analisi. Si indichi inoltre con Z la v.c. che descrive il logaritmo del reddito mensile complessivo del gruppo di coniugi (mariti e mogli). Allora si ha che

$$Z = X + Y \ ,$$

da cui si ricava

$$E(Z) = E(X+Y) = E(X) + E(Y) = 6,109 + 5,3 = 11,409 \ .$$

e

$$\begin{aligned} Var(Z) &= Var(X+Y) = Var(X) + Var(Y) + \underbrace{\cancelto{0}{2Cov(X,Y)}}_{\text{nulla per ipotesi}} \\ &= 0,699^2 + 0,9^2 = 1,389 \ . \end{aligned}$$