STATISTICA: esercizi svolti sulle

VARIABILI CASUALI

1 VARIABILI CASUALI

1.1 Variabili casuali generiche

1. Si supponga che un dado truccato, formato da sei facce contrassegnate dai numeri da 1 a 6, sia costruito in modo tale che la probabilità di ottenere "6" è doppia rispetto a quella degli altri punteggi. Indicata con X la variabile casuale "punteggio ottenuto in un lancio del dado", si determini:

   a) la legge di probabilità della variabile casuale X;

   b) il valore atteso e la varianza della variabile casuale X.

   Svolgimento
   Svolgimento punto a) La variabile casuale X può assumere i valori $1, 2, 3, 4, 5, 6$. Per quanto specificato nel testo dell'esercizio si ha che la funzione di probabilità della variabile casuale X ha la seguente forma:

   $$P(X = x) = \begin{cases} p & x = 1, 2, 3, 4, 5 \\ 2p & x = 6 \end{cases}$$

   Affinchè $P(X = x)$ sia effettivamente una funzione di probabilità per la variabile casuale X, il valore di p deve essere tale da assicurare che:

   (a) $P(X = x) \geq 0 \qquad \forall\, x = 1, 2, 3, 4, 5, 6;$

   (b) $\sum_{x=1}^{6} P(X = x) = 1$

   La prima condizione impone che $p \geq 0$.
   Osservando che

   $$\sum_{x=1}^{6} P(X = x) = 7p\ ,$$

   la seconda condizione prescrive che $7p = 1$. Quest'ultima equazione è soddisfatta solo se $p = \frac{1}{7} \geq 0$.
   La funzione di probabilità della variabile casuale X è di conseguenza data da:

   $$P(X = x) = \begin{cases} \dfrac{1}{7} & x = 1, 2, 3, 4, 5 \\ \\ \dfrac{2}{7} & x = 6 \end{cases}$$

   Svolgimento punto b) Il valore atteso $E(X)$ della variabile casuale X è dato da:

   $$E(X) \ = \ \sum_{x=1}^{6} x \cdot P(X = x)$$

$$
\begin{aligned}
&= \frac{1}{7}\sum_{x=1}^{5} x + 6 \cdot \frac{2}{7} = \frac{15}{7} + \frac{12}{7} \\
&= \frac{27}{7} = 3.8571 \quad .
\end{aligned}
$$

La varianza $Var(X)$ della variabile casuale X è data da:

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

Al fine del calcolo di $Var(X)$, si ricava il valore atteso $E(X^2)$.

$$
\begin{aligned}
E(X^2) &= \sum_{x=1}^{6} x^2 \cdot P(X = x) \\
&= \frac{1}{7}\sum_{x=1}^{5} x^2 + 6^2 \cdot \frac{2}{7} = \frac{55}{7} + \frac{72}{7} \\
&= \frac{127}{7} \quad .
\end{aligned}
$$

Si ha dunque:

$$
\begin{aligned}
Var(X) &= \frac{127}{7} - \left(\frac{27}{7}\right)^2 \\
&= \frac{127}{7} - \frac{729}{49} = \frac{889 - 729}{49} \\
&= \frac{160}{49} = 3.2653 \quad .
\end{aligned}
$$

2. Un gioco consiste nel lanciare un dado ed una moneta non truccati. Come risultato del lancio del dado si considera il numero riportato sulla faccia superiore, mentre per il lancio della moneta si considera il punteggio 0 se si presenta testa, punteggio 1 se si presenta croce. Determinare il valore atteso e la varianza della variabile casuale che descrive la somma dei punteggi riportati nel lancio del dado e della moneta.

   Svolgimento
   Si indichi con X la variabile casuale che interpreta l'esito del lancio del dado e con Y la variabile casuale che interpreta il punteggio associato all'esito del lancio della moneta. La variabile casuale X può assumere i valori appartenenti all'insieme $\{1, 2, 3, 4, 5, 6\}$ mentre la variabile casule Y può assumere i valori 0 o 1. Si ha inoltre che:

$$
P(X = x) = \frac{1}{6} \qquad x = 1, 2, 3, 4, 5, 6 \quad ;
$$

$$
P(Y = y) = \frac{1}{2} \qquad y = 0, 1 \quad .
$$

La variabile casuale che descrive la somma dei punteggi ottenuti lanciando il dado e la moneta è $Z = X + Y$. Si ha che:

$$\begin{aligned} E(Z) &= E(X+Y) = E(X) + E(Y) \\ &= \sum_{x=1}^{6} x \cdot P(X=x) + \sum_{y=0}^{1} y \cdot P(Y=y) \\ &= \frac{1}{6}\sum_{x=1}^{6} x + \frac{1}{2} \\ &= \frac{21}{6} + \frac{1}{2} = 3.5 + 0.5 = 4 \end{aligned}$$

Le variabili casuali X e Y sono indipendenti in quanto interpreti di esperimenti fisicamente separati. In questo caso si ha:

$$P[(X=x) \cap (Y=y)] = P(X=x) \cdot P(Y=y) \qquad x = 1, ..., 6; \quad y = 0, 1 \ ;$$

$$Var(Z) = Var(X+Y) = Var(X) + Var(Y) \ .$$

Per comodità, si ricavano dapprima le varianze $Var(X)$ e $Var(Y)$.

$$\begin{aligned} Var(X) &= E(X^2) - [E(X)^2] \\ &= \sum_{x=1}^{6} x^2 \cdot P(X=x) - (3.5)^2 \\ &= \frac{1}{6}\sum_{x=1}^{6} x^2 - (3.5)^2 \\ &= \frac{91}{6} - 12,25 = 15,1667 - 12,25 = 2.9167 \ ; \end{aligned}$$

$$\begin{aligned} Var(Y) &= E(Y^2) - [E(Y)^2] \\ &= \frac{1}{2} - \frac{1}{4} = \frac{1}{4} = 0.25 \ . \end{aligned}$$

Concludendo si ha:

$$Var(Z) = 2.9167 + 0.25 = 3.1667 \ .$$

3. Un'urna contiene tre palline contrassegnate dai numeri 1, 2, 3. Si estrae con riposizione un campione di ampiezza due. Sia Y la variabile casuale che esprime la media aritmetica dei numeri riportati sulle palline estratte. Calcolare aspettativa e varianza di Y.

   Svolgimento
   Si ricava dapprima lo spazio campionario Ω dell'esperimento che consiste nell'estrazione con riposizione di due palline dall'urna specificata dal testo dell'esercizio.

$$\begin{aligned} \Omega &= \{ (1,1); (1,2); (1,3); \\ &\quad (2,1); (2,2); (2,3); \\ &\quad (3,1); (3,2); (3,3)\} \\ &= \{(x_1, x_2) : x_1 = 1,2,3 \ \text{ e } \ x_2 = 1,2,3\}. \end{aligned}$$

Si osservi che gli eventi elementari in Ω sono equiprobabili e

$$P[(x_1, x_2)] = \frac{1}{9} \qquad \forall\ (x_1, x_2) \in \Omega \ \ .$$

La variabile casuale Y è una funzione così definita:

$$\begin{aligned} Y &: \Omega \to \mathbb{R} \\ (x_1, x_2) &\mapsto \frac{x_1 + x_2}{2} \ \ . \end{aligned}$$

Il seguente prospetto è utile a ricavare l'insieme dei valori assumibili dalla v.c. Y e la sua funzione di probabilità.

$(x_1; x_2)$	$p[(x_1, x_2)]$	$y = (x_1 + x_2)/2$
$(1, 1)$	$\frac{1}{9}$	1
$(1, 2)$	$\frac{1}{9}$	1.5
$(1, 3)$	$\frac{1}{9}$	2
$(2, 1)$	$\frac{1}{9}$	1.5
$(2, 2)$	$\frac{1}{9}$	2
$(2, 3)$	$\frac{1}{9}$	2.5
$(3, 1)$	$\frac{1}{9}$	2
$(3, 2)$	$\frac{1}{9}$	2.5
$(3, 3)$	$\frac{1}{9}$	3

Come si osserva dal prospetto appena riportato, la variabile Y può assumere i valori appartenenti al seguente insieme:

$$S = \{1, 1.5, 2, 2.5, 3\} \ \ .$$

Nel dettaglio:

- la v.c. Y assume valore 1 in corrispondenza della sola realizzazione campionaria $(1, 1)$. Tale realizzazione campionaria ha probabilità $\frac{1}{9}$ e di conseguenza

$$P(Y = 1) = \frac{1}{9} \ \ ;$$

- la v.c. Y assume valore 1.5 in corrispondenza di due realizzazioni campionarie, ciascuna con probabilità $\frac{1}{9}$. Di conseguenza:

$$P(Y = 1.5) = \frac{1}{9} + \frac{1}{9} = \frac{2}{9} \ \ ;$$

- la v.c. Y assume valore 2 in corrispondenza di tre realizzazioni campionarie, ciascuna con probabilità $\frac{1}{9}$. Di conseguenza:

$$P(Y = 2) = \frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{3}{9} \ \ ;$$

- la v.c. Y assume valore 2.5 in corrispondenza di due realizzazioni campionarie, ciascuna con probabilità $\frac{1}{9}$. Di conseguenza:

$$P(Y = 2.5) = \frac{1}{9} + \frac{1}{9} = \frac{2}{9} \ ;$$

- la v.c. Y assume valore 3 in corrispondenza della sola realizzazione campionaria $(3,3)$. Tale realizzazione campionaria ha probabilità $\frac{1}{9}$ e di conseguenza

$$P(Y = 3) = \frac{1}{9} \ .$$

In sintesi si ha:

$$Y = \begin{cases} 1 & 1.5 & 2 & 2.5 & 3 \\ \\ \dfrac{1}{9} & \dfrac{2}{9} & \dfrac{1}{3} & \dfrac{2}{9} & \dfrac{1}{9} \end{cases} \ .$$

Alla luce di ciò si ha che:

$$\begin{aligned} E(Y) &= \sum_{i=1}^{5} y_i \cdot P(Y = y_i) \\ &= 1 \cdot \frac{1}{9} + 1.5 \cdot \frac{2}{9} + 2 \cdot \frac{3}{9} + 2.5 \cdot \frac{2}{9} + 3 \cdot \frac{1}{9} \\ &= \frac{18}{9} = 2 \ . \end{aligned}$$

$$\begin{aligned} E(Y^2) &= \sum_{i=1}^{5} y_i^2 \cdot P(Y = y_i) \\ &= 1^2 \cdot \frac{1}{9} + (1.5)^2 \cdot \frac{2}{9} + 2^2 \cdot \frac{2}{9} + (2.5)^2 \cdot \frac{2}{9} + 3^2 \cdot \frac{1}{9} \\ &= \left(1 \cdot \frac{1}{9}\right) + \left(\frac{9}{4} \cdot \frac{2}{9}\right) + \left(4 \cdot \frac{3}{9}\right) + \left(\frac{25}{4} \cdot \frac{2}{9}\right) + \left(9 \cdot \frac{1}{9}\right) \\ &= \frac{156}{36} \ . \end{aligned}$$

$$\begin{aligned} Var(Y) &= E(Y^2) - [E(Y)]^2 \\ &= \frac{156}{36} - 4 = \frac{156 - 144}{36} = \frac{1}{3} \ . \end{aligned}$$

4. Data la distribuzione di probabilità congiunta delle variabili casuali X e Y:

X / Y	0	1	2
1	0,25	0	0,25
2	0,15	0	0,15
3	0	0,2	0

   a) stabilire se le variabili X e Y sono indipendenti;
   b) calcolare $\Pr\{Y \geq 2, X > 0\}$ e $\Pr\{Y \geq 2 | X > 0\}$;
   c) calcolare il valore atteso e la varianza della variabile casuale $Z = X - 2Y$;

Svolgimento

Svolgimento punto a) Di seguito si riportano la distribuzione congiunta e le distribuzioni marginali delle v.c. X e Y:

X / Y	0	1	2	tot
1	0.25	0	0.25	0.5
2	0.15	0	0.15	0.3
3	0	0.2	0	0.2
tot	0.4	0.2	0.4	1

Se le variabili casuali X e Y fossero indipendenti si avrebbe che:

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y) \qquad y = 1, 2, 3 \quad x = 0, 1, 2 \quad .$$

Si osservi che, secondo la distribuzione congiunta fornita dal testo dell'esercizio, si ha:

$$P(X = 1, Y = 1) = 0 \neq 0.2 \cdot 0.5 = P(X = 1) \cdot P(Y = 1) \quad .$$

Di conseguenza le variabili casuali X e Y non sono indipendenti.

Svolgimento punto b)

$$\begin{aligned} P(X > 0, Y \geq 2) &= P(X = 1, Y = 2) + P(X = 2, Y = 2) + \\ &\quad + P(X = 1, Y = 3) + P(X = 2, Y = 3) \\ &= 0 + 0.15 + 0.2 + 0 = 0.35 \end{aligned}$$

$$\begin{aligned} P(Y \geq 2 | X > 0) &= \frac{P(X > 0, Y \geq 2)}{P(X > 0)} \\ &= \frac{0.35}{P(X = 1) + P(X = 2)} = \frac{0.35}{0.2 + 0.4} \\ &= \frac{0.35}{0.6} = 0.5833 \quad . \end{aligned}$$

Svolgimento punto c) Dalle proprietà del valore atteso si ha che:

$$E(Z) = E(X - 2Y) = E(X) - 2E(Y).$$

$$\begin{aligned}
E(X) &= \sum_{x=0}^{2} x \cdot P(X=x) \\
&= (0 \cdot 0.4) + (1 \cdot 0.2) + (2 \cdot 0.4) = 1
\end{aligned}$$

$$\begin{aligned}
E(Y) &= \sum_{y=1}^{3} y \cdot P(Y=y) \\
&= (1 \cdot 0.5) + (2 \cdot 0.3) + (3 \cdot 0.2) = 1.7 \ \ .
\end{aligned}$$

Si ha dunque che:

$$E(Z) = E(X) - 2E(Y) = 1 - 3.4 = -2.4 \ \ .$$

Svolgimento punto d) Dalle proprietà della varianza e della covarianza si ha che:

$$Var(Z) = Var(X - 2Y) = Var(X) + 2^2 Var(Y) - (2 \cdot 2) Cov(X, Y) \ \ .$$

$$\begin{aligned}
Var(X) &= E(X^2) - [E(X)]^2 \\
&= \sum_{x=0}^{2} x^2 \cdot P(X=x) - 1^2 \\
&= (0 \cdot 0.4) + (1 \cdot 0.2) + (4 \cdot 0.4) - 1 = 1.8 - 1 = 0.8 \ \ ;
\end{aligned}$$

$$\begin{aligned}
Var(Y) &= E(Y^2) - [E(Y)]^2 \\
&= \sum_{y=1}^{3} y^2 \cdot P(Y=y) - (1.7^2) \\
&= (1 \cdot 0.5) + (4 \cdot 0.3) + (9 \cdot 0.2) - 2.89 = 3.5 - 2.89 = 0.61 \ \ ;
\end{aligned}$$

$$\begin{aligned}
Cov(X, Y) &= E(X \cdot Y) - E(X) \cdot E(Y) \\
&= \sum_{x=0}^{2} \sum_{y=1}^{3} xy \cdot P(X=x, Y=y) - (1)(1.7) \\
&= [(0 \cdot 1) \cdot 0.25] + [(0 \cdot 2) \cdot 0.15] + [(0 \cdot 3) \cdot 0] + \\
&\quad + [(1 \cdot 1) \cdot 0] + [(1 \cdot 2) \cdot 0] + [(1 \cdot 3) \cdot 0.2] + \\
&\quad + [(2 \cdot 1) \cdot 0.25] + [(2 \cdot 2) \cdot 0.15] + [(2 \cdot 3) \cdot 0] - 1.7 \\
&= 0.6 + 0.5 + 0.6 - 1.7 = 0 \ \ .
\end{aligned}$$

Questo ultimo risultato informa che le variabili casuali X e Y, pur non essendo indipendenti, risultano essere incorrelate. Alla luce di tali risultati si ha che:

$$\begin{aligned}
Var(Z) &= Var(X) + 4Var(Y) - 4Cov(X, Y) \\
&= 0.8 + 4 \cdot 0.61 - 0 = 3.24 \ \ .
\end{aligned}$$

1.2 Variabili casuali notevoli

5. 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;

   Svolgimento
   Svolgimento punto 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:

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

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

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

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

   Si ha dunque che:

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

   Svolgimento punto 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:

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

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

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

Svolgimento punto 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:

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

6. 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 $P(66 \leq X \leq 91)$.

   Svolgimento
   Svolgimento punto 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.

$$X = \text{“numero dei nuclei familiari estratti che possiedono almeno due televisori”}$$

   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:

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

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

Svolgimento punto 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:

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

7. 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).

   Svolgimento
   Svolgimento punto 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\ 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:

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

   Svolgimento punto 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:

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

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

   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:

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

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

   Svolgimento punto 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:

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

8. 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.

Svolgimento
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;
- 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:

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

9. 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 $Z = X + Y$.

   Svolgimento

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

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

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

$$\begin{aligned} P(Y = y) &= \begin{cases} \binom{2}{0} \cdot 0.4^0 \cdot (1 - 0.4)^2 & y = 0 \\ \binom{2}{1} \cdot 0.4^1 \cdot (1 - 0.4)^1 & y = 1 \\ \binom{2}{2} \cdot 0.4^2 \cdot (1 - 0.4)^0 & y = 2 \end{cases} \\ &= \begin{cases} (1 - 0.4)^2 & y = 0 \\ 2 \cdot 0.4 \cdot (1 - 0.4) & y = 1 \\ 0.4^2 & y = 2 \end{cases} \\ &= \begin{cases} 0.36 & y = 0 \\ 0.48 & y = 1 \\ 0.16 & y = 2 \end{cases} \end{aligned}$$

   Svolgimento punto 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:

Y / X	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

Svolgimento punto 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:

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