Esercitazione 1 del corso di Statistica 2

Prof. Domenico Vistocco

Dott.ssa Paola Costantini

28 Aprile 2008

Esercizio n. 1

Si consideri un campione di 200 studenti di cui si conoscono le seguenti probabilità distinti secondo il sesso (M=maschio, F=femmina) e l'età (G = di età 15-16, A = di età 17-18). Sapendo che:P(M) = 0,55, P(G) = 0,6 e che P(G | M) = 0,71 ricavare la distribuzione di frequenza doppia:

Sesso / Età	M	F	Tot.
G			
A			
Tot.			200

$$P(M) = \frac{x}{200} = 0{,}55 \Rightarrow x = 0{,}55 * 200 = 110$$

$$P(G) = \frac{x}{200} = 0{,}6 \Rightarrow x = 0{,}6 * 200 = 120$$

$$P(G|M) = \frac{x/200}{110/200} = \frac{x}{110} = 0{,}71 \Rightarrow x = 0{,}71 * 110 = 78$$

Esercizio n. 2

Supponiamo di avere un'urna contenente 12 palline, di cui 5 verdi, 3 blu e 4 rosse. Si vuole determinare la probabilità che, estraendo due palline con e senza reimmissione, esse siano:

a. dello stesso colore;
b. al più una pallina verde;
c. due palline di cui almeno una blu;
d. due palline di cui solo la prima rossa;
e. due palline di cui una verde e l'altra non rossa;

Soluzione

Estrazione Bernoulliana (o con reimmissione)

a. ("due palline dello stesso colore")

$$P[(R_1 \cap R_2) \cup (V_1 \cap V_2) \cup (B_1 \cap B_2)] = P(R_1 \cap R_2) + P(V_1 \cap V_2) + P(B_1 \cap B_2) =$$

$$= P(R_1) \cdot P(R_2) + P(V_1) \cdot P(V_2) + P(B_1) \cdot P(B_2) =$$

$$= \left(\frac{4}{12} \cdot \frac{4}{12}\right) + \left(\frac{5}{12} \cdot \frac{5}{12}\right) + \left(\frac{3}{12} \cdot \frac{3}{12}\right) = \frac{16}{144} + \frac{25}{144} + \frac{9}{144} = \frac{50}{144} = 0{,}347$$

b. ("al più una pallina verde")

$$P\left[\left(V_1 \cap \overline{V_2}\right) \cup \left(\overline{V}_1 \cap V_2\right) \cup \left(\overline{V_1} \cap \overline{V_2}\right)\right] =$$

$$= P\left[\left(V_1 \cap \overline{V_2}\right) + \left(\overline{V}_1 \cap V_2\right) + \left(\overline{V_1} \cap \overline{V_2}\right)\right] = P(V_1) \cdot P(\overline{V_2}) + P(\overline{V_1}) \cdot P(V_2) + P(\overline{V_1}) \cdot P(\overline{V_2}) =$$

$$= \frac{5}{12} \cdot \frac{7}{12} + \frac{7}{12} \cdot \frac{5}{12} + \frac{7}{12} \cdot \frac{7}{12} = \frac{35}{144} + \frac{35}{144} + \frac{49}{144} = \frac{119}{144} = 0{,}82$$

c. ("due palline di cui almeno una blu")=

$$P\left[\left(B_1 \cap \overline{B_2}\right) \cup \left(\overline{B_1} \cap B_2\right) \cup (B_1 \cap B_2)\right] = P(B_1) \cdot P(\overline{B_2}) + P(\overline{B_1}) \cdot P(B_2) + P(B_1) \cdot P(B_2) =$$

$$= \frac{3}{12} \cdot \frac{9}{12} + \frac{9}{12} \cdot \frac{3}{12} + \frac{3}{12} \cdot \frac{3}{12} = 0{,}4375$$

d. ("due palline di cui solo la prima rossa")=

$$P(R_1 \cap \overline{R_2}) = P(R_1) \cdot P(\overline{R_2}) = \frac{4}{12} \cdot \frac{8}{12} = \frac{32}{144} = 0{,}22$$

e. ("due palline di cui una verde e l'altra non rossa")

$P\left[\left(V_1 \cap \overline{R_2}\right) \cup \left(\overline{R_1} \cap V_2\right)\right] =$ posto che

$$P(V_1 \cap V_2) \cup (V_1 \cap B_2) \cup (B_1 \cap V_2) = P(V_1) \cdot P(V_2) + P(V_1) \cdot P(B_2) + P(B_1) \cdot P(V_2) =$$

$$= \frac{5}{12} \cdot \frac{5}{12} + \frac{5}{12} \cdot \frac{3}{12} + \frac{3}{12} \cdot \frac{5}{12} = \frac{25}{144} + \frac{15}{144} + \frac{15}{144} = \frac{55}{144} = 0{,}38$$

Estrazione in blocco (o senza reimmissione):

a. ("due palline dello stesso colore")

$$P[(R_1 \cap R_2) \cup (V_1 \cap V_2) \cup (B_1 \cap B_2)] = P(R_1 \cap R_2) + P(V_1 \cap V_2) + P(B_1 \cap B_2) =$$

$$= P(R_1) \cdot P(R_2 | R_1) + P(V_1) \cdot P(V_2 | V_1) + P(B_1) \cdot P(B_2 | B_1) =$$

$$= \left(\frac{4}{12} \cdot \frac{3}{11}\right) + \left(\frac{5}{12} \cdot \frac{4}{11}\right) + \left(\frac{3}{12} \cdot \frac{2}{11}\right) = \frac{12}{132} + \frac{20}{132} + \frac{6}{132} = \frac{38}{132} = 0{,}287$$

b. ("al più una pallina verde")

$$P\left[\left(V_1 \cap \overline{V_2}\right) \cup \left(\overline{V}_1 \cap V_2\right) \cup \left(\overline{V_1} \cap \overline{V_2}\right)\right] =$$

$$= P\left[\left(V_1 \cap \overline{V_2}\right) + \left(\overline{V}_1 \cap V_2\right) + \left(\overline{V_1} \cap \overline{V_2}\right)\right] = P(V_1) \cdot P(\overline{V_2} | V_1) + P(\overline{V_1}) \cdot P(V_2 | V_1) + P(\overline{V_1}) \cdot P(\overline{V_2} | \overline{V_1}) =$$

$$= \frac{5}{12} \cdot \frac{7}{11} + \frac{7}{12} \cdot \frac{5}{11} + \frac{7}{12} \cdot \frac{6}{11} = \frac{119}{132} = 0{,}84$$

c. ("due palline di cui almeno una blu")=

$$P\left[\left(B_1 \cap \overline{B_2}\right) \cup \left(\overline{B_1} \cap B_2\right) \cup (B_1 \cap B_2)\right] = P(B_1) \cdot P(\overline{B_2} | B_1) + P(\overline{B_1}) \cdot P(B_2 | \overline{B_1}) + P(B_1) \cdot P(B_2 | B_1) =$$

$$= \frac{3}{12} \cdot \frac{9}{11} + \frac{9}{12} \cdot \frac{3}{11} + \frac{3}{12} \cdot \frac{2}{11} = 0{,}4545$$

d. ("due palline di cui solo la prima rossa")=

$$P(R_1 \cap \overline{R_2}) = P(R_1) \cdot P(\overline{R_2} | R_1) = \frac{4}{12} \cdot \frac{8}{11} = \frac{32}{132} = 0{,}24$$

e. ("due palline di cui una verde e l'altra non rossa")

$P\left[\left(V_1 \cap \overline{R_2}\right) \cup \left(\overline{R_1} \cap V_2\right)\right] =$ posto che

$$P(V_1 \cap V_2) \cup (V_1 \cap B_2) \cup (B_1 \cap V_2) = P(V_1) \cdot P(V_2 | V_1) + P(V_1) \cdot P(B_2 | V_1) + P(B_1) \cdot P(V_2 | B_1) =$$

$$= \frac{5}{12} \cdot \frac{4}{11} + \frac{5}{12} \cdot \frac{3}{11} + \frac{3}{12} \cdot \frac{5}{11} = \frac{20}{132} + \frac{15}{132} + \frac{15}{132} = \frac{50}{132} = 0{,}378$$

Esercizio n. 3

Una fabbrica produce ganci per appendere quadri e li vende in confezioni da 20. Il proprietario della fabbrica ha stimato che la probabilità che in una scatola non vi siano ganci difettosi è 0,85, la probabilità che vi sia un gancio difettoso è 0,12 e la probabilità che vi siano due ganci difettosi è 0,03.
Sia *X* la variabile casuale che descrive il numero di ganci difettosi; la distribuzione di X è riportata nella tabella 2.

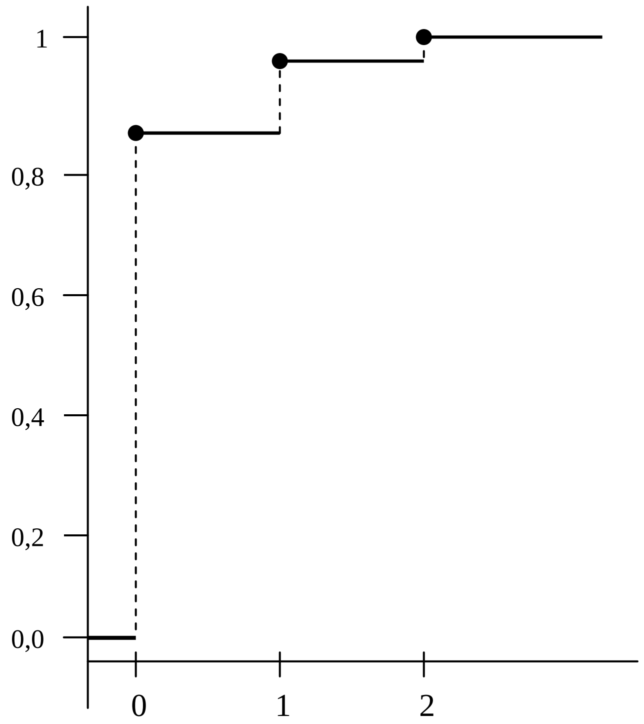
a. Calcolare e rappresentare graficamente la funzione di ripartizione di X.
b. Calcolare il valore atteso e la varianza di X.

Soluzione

Tab. 2

Ganci difettosi	0	1	2
Probabilità	0,85	0,12	0,03

F(x)=0, per x < 0
F(x)=0,85 per 0 ≤ x ≤ 1
F(x)=0,97 per 1 ≤ x ≤ 2
F(x)=1 per x ≥ 2



b. Nel caso di variabili casuali discrete il valore atteso è dato da

$$\mu = E[X] = \sum_i x_i p_i$$

Nel nostro caso abbiamo:

$$\mu = E[X] = \sum_{x=0}^{2} xP(X = x) = 0 \times 0{,}85 + 1 \times 0{,}12 + 2 \times 0{,}03 = 0{,}18$$

Tale risultato può essere interpretato nel modo seguente: se si considera un elevato numero di scatole il numero medio di ganci difettosi per scatola è 0,18.

Per le variabili casuali discrete la varianza è data da:

$$\sigma^2 = \sum_i (xi - \mu)^2 p_i$$

Nel nostro caso la varianza è data da:

$$\sigma^2 = \sum_{x=0}^{2}(x-\mu)^2 P(X=x) = (0-0{,}18)^2 \times 0{,}85 + (1-0{,}18)^2 \times 0{,}12 + (2-0{,}18)^2 \times 0{,}03 = 0{,}21$$

La varianza può anche essere calcolata come differenza tra il valore atteso del quadrato di X e il quadrato della media.

$$\sigma^2 = Var(X) = E[X^2] - \mu^2$$
$$E[X^2] = 0 \times 0{,}85 + 1 \times 0{,}12 + 4 \times 0{,}03 = 0{,}24$$
$$\mu^2 = 0{,}18^2 = 0{,}03$$
$$\sigma^2 = 0{,}24 - 0{,}03 = 0{,}21$$
$$\sigma = \sqrt{\sigma_2} = 0{,}458$$

Esercizio n. 4

Lanciando tre monete, si consideri la v.c. descritta dal numero delle teste che si presentano in una prova.

Spazio degli eventi	x_i	p_i	P_i
CCC	0	1/8	1/8
CCC	1	3/8	4/8
CTC			
TCC			
TTC	2	3/8	7/8
TCT			
CTT			
TTT	3	1/8	1

Calcolare il valore atteso e la varianza di X.

Applicando la formula del valore atteso: $\mu = E[X] = \sum_i x_i p_i$ abbiamo:

$E(X) = 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = 1{,}5$ per cui in una prova il numero atteso di "teste" è pari a 1,5.

Var(X) $= \sigma^2 = \sum_i (x_i - \mu)^2 p_i =$

$$(0-1{,}5)^2 \cdot 0{,}125 + (1-1{,}5)^2 \cdot 0{,}375 + (2-1{,}5)^2 \cdot 0{,}375 + (3-1{,}5)^2 \cdot 0{,}125 = 0{,}75$$

Solitamente σ risulta un indice di variabilità più conveniente da utilizzare perché è espresso nella stessa unità di X.

$\sigma = \sqrt{\sigma_2} = \sqrt{0,75} = 0,866$

Il numero di "teste" che si presenta in ciascun lancio è diverso da quello atteso (1,5) mediamente per 0,866 teste.

Esercizio n. 5

Da un'indagine statistica emerge che gli italiani hanno in media il livello di colesterolo a 225 unità, con una varianza pari a 75 unità.

a. Calcolare la probabilità che il livello di colesterolo sia compreso tra 150 e 300 unità.

$$P(|X - \mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$$

$$\Delta = k\sigma \Rightarrow k = \frac{\Delta}{\sigma} = \frac{75}{8,66} = 8,66$$

$$P(|X - 225| < 75) = P(150 \leq X \leq 300) \geq 1 - \frac{1}{8,66^2} = 0,987$$

b. Calcolare la probabilità che il livello di colesterolo sia compreso tra 200 e 250 unità.

$$P(|X - \mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$$

$k = \frac{\Delta}{\sigma} = \frac{25}{8,66} = 2,9$ per k > 0 la disuguaglianza è valida.

$$P(|X - 225| < 25) = P(200 \leq X \leq 250) \geq 1 - \frac{1}{2,9^2} = 0,88$$

c. Calcolare la probabilità che il livello di colesterolo sia compreso tra 215 e 235 unità.

$$P(|X - \mu| \leq k\sigma) \geq 1 - \frac{1}{k^2}$$

$k = \frac{\Delta}{\sigma} = \frac{10}{8,66} = 1,15$ per k > 0 la disuguaglianza è valida.

$$P(|X - 225| < 10) = P(215 \leq X \leq 235) \geq 1 - \frac{1}{1,15^2} = 0,2438$$

c. Calcolare la probabilità che il livello di colesterolo sia compreso tra 220 e 230 unità.

$$P(|X-\mu| \le k\sigma) \ge 1-\frac{1}{k^2}$$

$k = \frac{\Delta}{\sigma} = \frac{5}{8{,}66} = 0{,}57$ per k < 0 la disuguaglianza è banale in termini di contenuto informativo

$$P(|X-225| < 5) = P(220 \le X \le 230) \ge 1-\frac{1}{0{,}57^2} = -2$$

Di conseguenza, quando l'ampiezza dell'intervallo aumenta X assume con maggiore probabilità valori distanti da μ.