Corso di Laurea in Ingegneria Informatica
Anno Accademico 2014/2015
Calcolo delle Probabilità e Statistica Matematica

Nome ....

N. Matricola .... Ancona, 6 novembre 2015

1. In un'urna ci sono 5 palline rosse e 3 palline bianche. Se ne estraggono due a caso, in successione, senza rimpiazzo e senza osservare il risultato della prima estrazione prima di aver affettuato la seconda. Siano X_1 e X_2 le v.a. tali che

$$X_1 = \begin{cases} 1 & \text{se la prima estratta è rossa} \\ 0 & \text{se la prima estratta è bianca} \end{cases}$$

e

$$X_2 = \begin{cases} 1 & \text{se la seconda estratta è rossa} \\ 0 & \text{se la seconda estratta è bianca} \end{cases}$$

Calcolare il coefficiente di correlazione di X_1 e X_2.

Soluzione. Dobbiamo determinare la distribuzione congiunta di X_1 e X_2 e le marginali. Per la marginale di X_1 possiamo procedere per via elementare, in quanto $P(X_1 = 1) = 5/8$ e $P(X_1 = 0) = 3/8$. Anche la marginale di X_2 può essere calcolata senza far ricorso alla densità congiunta, facendo uso del teorema delle probabilità totali. Abbiamo:

$$\begin{aligned} &P(X_2 = 1) = P(X_2 = 1|X_1 = 1)\, P(X_1 = 1) + P(X_2 = 1|X_1 = 0)\, P(X_1 = 0) = \\ &\quad = \frac{4}{7}\,\frac{5}{8} + \frac{5}{7}\,\frac{3}{8} = \frac{35}{56} \\ &P(X_2 = 0) = 1 - P(X_2 = 1) = \frac{19}{56} \end{aligned}$$

Da qui:

$$\begin{aligned} &E[X_1] = \frac{5}{8} \\ &E[X_2] = \frac{35}{56} \\ &Var(X_1) = E[X_1^2] - E[X_1]^2 = \frac{5}{8} - \left(\frac{5}{8}\right)^2 = \frac{15}{64} \\ &Var(X_2) = E[X_2^2] - E[X_2]^2 = \frac{35}{56} - \left(\frac{35}{56}\right)^2 = \frac{15}{64}. \end{aligned}$$

Per il calcolo della densità congiunta, valutiamo preventivamente tutti i casi possibili dell'estrazione delle due palline, che indichiamo con $\#S$. Si tratta di tutte le possibili coppie che si possono formare con le 8 palline totali, ma contando anche l'ordine; abbiamo pertanto:

$$\#S = D_{8,2} = C_{8,2}\, P_2 = 2\binom{8}{2} = 2\,\frac{8!}{2!\,6!} = 56.$$

I valori della densità congiunta sono dunque:

$$
\begin{aligned}
P(X_1 = 1, X_2 = 1) &= \frac{D_{5,2}}{56} = \frac{2\binom{5}{2}}{56} = \frac{20}{56} \\
P(X_1 = 1, X_2 = 0) &= \frac{5 \times 3}{56} = \frac{15}{56} \\
P(X_1 = 0, X_2 = 1) &= \frac{3 \times 5}{56} = \frac{15}{56} \\
P(X_1 = 0, X_2 = 0) &= \frac{D_{3,2}}{56} = \frac{2\binom{3}{2}}{56} = \frac{6}{56},
\end{aligned}
$$

che possiamo organizzare nella matrice

$$
\begin{pmatrix} 20/56 & 15/56 \\ \\ 15/56 & 6/56 \end{pmatrix},
$$

dove le righe si riferiscono a X_1 e le colonne a X_2. Le somme sulle righe e sulle colonne sono consistenti con quanto già trovato per le marginali di X_1 e X_2. Possiamo ora calcolare

$$
E[X_1\, X_2] = 1 \times \frac{20}{56} + 0 \times \frac{36}{56} = \frac{20}{56},
$$

che dà per la covarianza:

$$
Cov(X_1, X_2) = E[X_1\, X_2] - E[X_1]\, E[X_2] = \frac{20}{56} - \frac{5}{8}\,\frac{35}{56} = -\frac{15}{448}
$$

e il coefficiente di correlazione è

$$
\rho = \frac{Cov(X_1, X_2)}{\sqrt{Var(X_1)\, Var(X_2)}} = -\frac{15}{448}\,\frac{64}{15} = -\frac{1}{7} \approx 0.14
$$

2. Si vuole stimare il reddito familiare medio annuo della popolazione di un certo quartiere. Un campione di 10 famiglie dà i seguenti risultati, in migliaia di euro:

$$
60, 45, 72, 84, 55, 91, 62, 63, 72, 80.
$$

Determinare gli intervalli di confidenza bilaterali per il reddito medio al 90%, 95% e 99%.

Soluzione. Abbiamo $n = 10$, $\overline{X}_n = 68.4$ e $\sigma^2 = Var(X) = 198.0$. Dalle tavole:

$$
t_{\alpha/2}(9) = 1.833, \qquad 2.262, \qquad 3.250
$$

per i livelli di confidenza rispettivamente del 90%, 95% e 99%. Gli intervalli $\overline{X}_n \pm \sigma\, t_{\alpha/2}(9)/\sqrt{n}$ sono pertanto:

$$
\begin{aligned}
68.4 &\pm \sqrt{\frac{198.0}{10}}\, 1.833 = [60.24, 76.56] \\
68.4 &\pm \sqrt{\frac{198.0}{10}}\, 2.262 = [58.33, 78.47] \\
68.4 &\pm \sqrt{\frac{198.0}{10}}\, 3.250 = [53.94, 82.86]
\end{aligned}
$$

3. Una fiera vede un'affluenza media giornaliera di 1000 visitatori, distribuiti secondo una legge di Poisson. Per incrementare il numero di visite, gli organizzatori si rivolgono a una agenzia pubblicitaria specializzata che, dall'esperienza passata, riesce ad aumentare la media di un fattore 1.5 nell' 80% dei casi (mentre è ininfluente nel restante 20 %). Nella giornata successiva alla fine della campagna pubblicitaria, gli organizzatori rilevano una presenza di 1230 visitatori. Qual è la probabilità che l'aumento sia dovuto alla campagna pubblicitaria?

Soluzione. Sia X la variabile casuale che rappresenta il numero di visitatori in una particolare giornata. Tale variabile possiede una distribuzione di Poisson di parametro $\lambda_N = 1000$ in assenza della campagna pubblicitaria (o con campagna ininfluente) e $\lambda_E = 1500$ con una campagna efficace. Indichiamo con E e N gli eventi

$$E = \text{“la campagna pubblicitaria è stata efficace”}$$
$$N = \text{“la campagna pubblicitaria non è stata efficace”}.$$

Abbiamo allora $P(E) = 0.8$ e $P(N) = 0.2$. La domanda del problema si traduce nel calcolo della probabilità condizionata $P(E|X = 1230)$, che eseguiamo con l'ausilio della formula di Bayes e del teorema delle probabilità totali.

$$P(E|X = 1230) = \frac{P(X = 1230|E)\,P(E)}{P(X = 1230)}$$

Calcoliamo separatamente i vari pezzi, tenendo presente che $k = 1230$.

$$\begin{aligned}
&P(X = k|E) = e^{-\lambda_E}\,\frac{\lambda_E^k}{k!}\\
&P(X = k) = P(X = k|E)\,P(E) + P(X = k|N)\,P(N) =\\
&\quad = e^{-\lambda_E}\,\frac{\lambda_E^k}{k!} \times 0.8 + e^{-\lambda_N}\,\frac{\lambda_N^k}{k!} \times 0.2 = \frac{\lambda_E\,e^{-\lambda_E}}{k!}\left[0.8 + 0.2 \times \frac{\lambda_N^k}{\lambda_E^k}\,e^{\lambda_E - \lambda_N}\right]
\end{aligned}$$

E quindi

$$P(E|X = 1230) = \frac{0.8}{0.8 + 0.2 \times \frac{\lambda_N^k}{\lambda_E^k}\,e^{\lambda_E - \lambda_N}} \approx 0.53$$