Corso di Laurea in Ingegneria Informatica
Anno Accademico 2017/2018
Probabilità e Statistica - Prova pratica

Nome ....

N. Matricola .... Ancona, 18 gennaio 2018

1. La media dei voti di una classe in una verifica di latino è stata di 8.5 (nella scala da 1 a 10) e la deviazione standard campionaria di 2.34. Se il 90 % degli studenti ha ottenuto una votazione inferiore a 9, nell'ipotesi di una distribuzione normale dei voti, quanti sono gli studenti della classe?

   Svolgimento. Indichiamo con X la variabile casuale che indica il voto di uno studente preso a caso. Possiamo assumere che $\sim N(\mu, \sigma^2)$ con $\mu = 8.5$ e $\sigma = 2.34$. Se la classe ha n studenti, la variabile standardizzata è
$$Z = \frac{X - \mu}{\sigma/\sqrt{n}}$$
   Abbiamo pertanto che dal dato $P(X < 9) = 0.9$ otteniamo
$$0.9 = P(X < 9) = P\left(Z < \frac{9 - 8.5}{2.34/\sqrt{n}}\right) = P\left(Z < 0.21\sqrt{n}\right) = 1 - P\left(Z > 0.21\sqrt{n}\right)$$
   e quindi $P\left(Z > 0.21\sqrt{n}\right) = 0.1$ da cui $0.21\sqrt{n} = z_{0.1}$ dove $z_{0.1} = 1.282$ è il quantile della normale al 10 %. Otteniamo pertanto
$$\sqrt{n} = \frac{1.282}{0.21} \approx 6.1$$
   che fornisce $n \approx 36$ (o 37).

2. Si estraggono 2 palline da un'urna che ne contiene 15, di cui 10 rosse e 5 blu. L'estrazione di una pallina rossa comporta una vincita di 50 euro, l'estrazione di una blu una perdita di 100 euro. Quanto vale la vincita media in questo gioco?

   Svolgimento. Introduciamo gli eventi

- RR="entrambe le palline estratte sono rosse";
- RB="una pallina estratta è rossa, l'altra blu";
- BB="entrambe le palline estratte sono blu".

Indichiamo con X la vincita, che può assumere i valori $+100$ (due palline rosse), -50 (una rossa e una blu) e -200 (due palline blu). Abbiamo:

$$E[X] = 100\,P(RR) - 50\,P(RB) - 200\,P(BB)$$

Le probabilità sono date da

$$P(RR) = \frac{\binom{10}{2}}{\binom{15}{2}} = \frac{3}{7}$$
$$P(BB) = \frac{\binom{5}{2}}{\binom{15}{2}} = \frac{2}{21}$$
$$P(RB) = \frac{50}{\binom{15}{2}} = \frac{10}{21}$$

Sostituendo:

$$E[X] = 100\,\frac{3}{7} - 50\,\frac{10}{21} - 200\,\frac{2}{21} = \frac{900-500-400}{21} = 0$$

3. Ad un incrocio pericoloso si verificano in media 2 incidenti ogni settimana, distribuiti secondo la legge di Poisson. Quant' è la probabilità che

   (i) ci siano più di 3 incidenti in esattamente 2 settimane in un mese (con 1 mese=4 settimane)?

   (ii) non ci sia nessun incidente in un mese?

   (iii) in un dato mese, esista una settimana senza incidenti?

Svolgimento.

(i) La probabilità p che ci siano più di 3 incidenti in una settimana è data da

$$p = \sum_{k=4}^{\infty} e^{-2}\,\frac{2^k}{k!} = 1 - \sum_{k=0}^{3} e^{-2}\,\frac{2^k}{k!} = 1 - e^{-2}\left(1 + 2 + \frac{4}{2} + \frac{8}{6}\right) = e^{-2}\,\frac{19}{3} \approx 0.857$$

quindi la probabilità che ci siano esattamente due settimane su quattro in cui ciò avviene segue la legge binomiale $B(4, p)$. Abbiamo quindi per tale probabilità:

$$\binom{4}{2} p^2 (1-p)^2 \approx 0.09$$

(ii) La probabilità p che non ci sia nessun incidente in una settimana è e^{-2}; quindi la probabilità che non ci sia nessun incidente in quattro settimane è

$$(e^{-2})^4 = e^{-8} \approx 3.3\ 10^{-4}$$

(iii) Seguendo il calcolo del punto precedente, la probabilità che ci sia una settimana senza incidenti è

$$\binom{4}{1} (e^{-2})^1 (1-e^{-2})^3 \approx 0.35$$

4. La seguente tabella di contingenza mostra i dati relativi alla durata delle automobili prodotte da tre diverse case automobilistiche, A, B e C:

		Casa	
Km (migliaia)	A	B	C
< 100	120	150	130
$[100, 150]$	100	140	130
$[150, 200]$	90	100	90
> 200	80	80	80

Mediante un test del χ^2 di indipendenza, stabilire se, al 5% di significatività, la vita delle macchine esaminate sia indipendente dalla casa costruttrice.

Svolgimento. Dobbiamo innanzitutto stimare le probabilità marginali. Il rango del campione è $n = 1290$.Indichiamo con

$$\widehat{p}_i = \sum_{j=1}^{3} \frac{N_{ij}}{n} \qquad \widehat{q}_j = \sum_{i=1}^{4} \frac{N_{ij}}{n}$$

$i = 1, 2, 3, 4,\ j = 1, 2, 3$, gli stimatori delle densità marginali relative rispettivamente alla durata e alla casa costruttrice. Si ha, aggiungendo una riga ed una colonna alla tabella (arrotondando a due cifre): La statistica da confrontare con il valore tabulare del χ^2 è

	Casa			
Km (migliaia)	A	B	C	p_i
< 100	120	150	130	0.31
$[100, 150]$	100	140	130	0.29
$[150, 200]$	90	100	90	0.22
> 200	80	80	80	0.19
q_j	0.30	0.36	0.33	

$$T = \sum_{i=1}^{4} \sum_{j=1}^{3} \frac{(N_{ij} - n\,\widehat{p}_i\,\widehat{q}_j)^2}{n\,\widehat{p}_i\,\widehat{q}_j} = 3.93.$$

Il numero di gradi di libertà è $3 \times 2 = 6$, quindi il valore sulle tavole è $\chi^2_{0.05}(6) = 12.592$ quindi non si può rigettare l'ipotesi nulla di indipendenza della durata rispetto alla casa costruttrice.