Corso di Laurea Specialistica in Ingegneria Elettronica
Curriculum di Biomedica
Anno Accademico 2012/2013
Statistica Medica

Nome ....

N. Matricola .... Ancona, 15 novembre 2013

1. In un'elezione comunale si presentano 4 liste. Siano n_1, n_2, n_3 ed n_4 i voti riportati da ciascuna lista. Si intervista quindi un elettore a caso e si denota con X il numero totale di voti riportati dalla lista per cui ha dichiarato di aver dato il voto. Si sceglie, poi, indipendentemente e a caso, una delle quattro liste e si denota con Y il numero totale di voti riportati da quella lista.

   (i) Qualè la distribuzione di probabilità di X e di Y?

   (ii) Qualè il valore di aspettazione di X e di Y?

   (iii) Rispondere alle domande precedenti nel caso in cui gli elettori in totale siano 1480 e le quattro liste abbiano riportato rispettivamente 400, 330, 250 e 500 voti.

Soluzione.

(i) Le due variabili casuali X e Y sono discrete e possono assumere gli stessi valori, $\{n_1, n_2, n_3, n_4\}$, con probabilità $p_X(n_i) = n_i/N$ e $p_Y(n_i) = 1/4$, $i = 1, 2, 3, 4$.

(ii) $E[X] = \sum_{i=1}^4 n_i\, p_X(n_i) = \sum_{i=1}^4 (n_i^2)/N$; $E[Y] = \sum_{i=1}^4 n_i\, p_Y(n_i) = \sum_{i=1}^4 (n_i)/4 = N/4$.

(iii) $p_X(n_1) = 400/1480 = 0.27$, $p_X(n_2) = 330/1480 = 0.22$, $p_X(n_3) = 250/1480 = 0.17$, $p_X(n_4) = 500/1480 = 0.34$; dunque $E[X] = 392.8$ e $E[Y] = 370$.

2. II numero di pazienti in arrivo al pronto soccorso di un ospedale segue una distribuzione di Poisson con media di 3 all'ora.

   (i) Qual'è la probabilità che vi siano almeno 2 pazienti dalle 10.30 alle 11.00?

   (ii) Assumendo ancora che l'evento del punto *(i)* si verifichi, qual'è la probabilità che dalle 10.30 alle 11.15 si presentino almeno 4 pazienti?

Soluzione. Sia X il numero di pazienti in un'ora. Allora X è distribuita secondo

$$\frac{\lambda^k}{k!}\, e^{-\lambda}$$

con $\lambda = 3$.

(i) Il numero di pazienti in mezz'ora, che indichiamo con Y, avrà ancora una distribuzione di Poisson con $\lambda_Y = 3/2 = 1.5$. Avremo allora

$$\begin{aligned} P(Y \geq 2) &= 1 - P(Y < 2) = 1 - [P(Y = 0) + P(Y = 1)] \\ &= 1 - \sum_{k=0}^{1} \frac{(1.5)^k \, e^{-1.5}}{k!} = 0.44 \end{aligned}$$

(ii) Dobbiamo considerare il numero di pazienti, W, in 45 minuti, che segue la distribuzione di Poisson con $\lambda_Z = 3 \times 3/4 = 2.25$. La probabilità richiesta è una probabilità condizionata, data da

$$\begin{aligned} P(W \geq 4 | W \geq 2) &= \frac{P(\{W \geq 4\} \cap \{W \geq 2\})}{P(W \geq 2)} = \frac{P(\{W \geq 4\})}{P(W \geq 2)} \\ &= \frac{1 - \sum_{k=0}^{3} P(W = k)}{1 - \sum_{k=0}^{1} P(W = k)} = \frac{1 - \sum_{k=0}^{3} (2.25)^k e^{-2.25}/k!}{1 - \sum_{k=0}^{1} (2.25)^k e^{-2.25}/k!} = 0.29 \end{aligned}$$

3. In un gruppo di 1000 individui, nell'80 % dei quali è presente una data malattia infettiva, un test ha dato 700 positivi tra le persone realmente malate e 20 tra quelle non malate.

   (i) Calcolare sensibilità e specificità del test;

   (ii) se il test viene effettuato su una popolazione in cui si sa che la diffusione della malattia è del 15 %, qualè la probabilità che un individuo positivo al test sia realmente malato?

Soluzione.

(i) Si chiede $P(T + |M)$ e $P(T - |S)$, dove $T+$ e $T-$ sono gli eventi "positivo al test e "negativo al test, M ="persona malata, S ="persona sana. Abbiamo: $P(M) = 0.8$; $P(S) = 1 - P(M) = 0.2$; $P(T + \cap M) = 0.7$; $P(T + \cap S) = 0.02$. Inoltre ci servirà $P(T - \cap S) = 180/1000 = 0.18$. Allora:

$$\begin{aligned} \text{sens.} = P(T + |M) &= \frac{P(T + \cap M)}{P(M)} = \frac{0.7}{0.8} = 0.875 \\ \text{spec.} = P(T - |S) &= \frac{P(T - \cap S)}{P(S)} = \frac{0.18}{0.2} = 0.9 \end{aligned}$$

(ii) Si chiede $P(M|T+)$ nel caso in cui la sensibilità e la specificità del test sono date dai valori riportati sopra e $P(M) = 0.15$. Ci servirà $P(T+)$ che è dato dalla formula delle probabilità totali:

$$P(T+) = P(T + |M) \, P(M) + P(T + |S) \, P(S) = 0.875 \times 0.15 + (1 - 0.9) \times 0.85 = 0.22$$

Abbiamo pertanto:

$$P(M|T+) = \frac{P(T + |M) \, P(M)}{P(T+)} = \frac{sens. \, 0.15}{0.22} = \frac{0.875 \times 0.15}{0.22} = 0.6.$$

4. Nella popolazione umana, la distribuzione dei globuli bianchi ha una media $\mu_0 = 7250$ per mm^3, con varianza non nota. Nei soggetti che hanno contratto una data infezione, si ritiene che il numero di globuli bianchi sia in media minore. Si misura la quantità di globuli bianchi ad un campione casuale di 20 soggetti sospetti di aver contratto l'infezione, ottenendo una media campionaria di 6000 per mm^3 ed una deviazione standard campionaria di 1500 per mm^3. Testare al 5% l'ipotesi di una minor concentrazione di globuli bianchi nei soggetti che si suppongono infetti.

Soluzione. L'ipotesi nulla H_0 è che la popolazione sia distribuita secondo una legge normale di media $\mu_0 = 7250$ per mm^3 con σ incognita, quindi dovremo usare la distribuzione di Student. L'ipotesi alternativa H_1 è che $\mu < \mu_0$; la regione critica per il rigetto dell'ipotesi nulla è pertanto $T < -t_{0.05}(19) = -1.729$, dove 19 è il numero di gradi di libertà (il campione è di rango $n = 20$). Non ci resta che calcolare T:

$$T = \frac{\overline{X}_n - \mu_0}{S} \sqrt{n} = \frac{6000 - 7250}{1500} \sqrt{20} = -3.73$$

che cade nella regione di rigetto. I dati quindi non confermano l'ipotesi nulla.