ANALISI 2 | Lu 25/5/20

Esercizi vari

1. Sono stati scelti 100 numeri in [0,1]. Si utilizzi il test del chi-quadro al livello 0.05 per stabilire se tale scelta sia basata sulla ripartizione uniforme, supponendo che vi siano

7 numeri fra 0 e 0.1
13 numeri fra 0.1 e 0.2
6 numeri fra 0.2 e 0.3
16 numeri fra 0.3 e 0.4
13 numeri fra 0.4 e 0.5
9 numeri fra 0.5 e 0.6
17 numeri fra 0.6 e 0.7
4 numeri fra 0.7 e 0.8
8 numeri fra 0.8 e 0.9
7 numeri fra 0.9 e 1.

Risposta: si ha $N=100$; dividiamo (ovviamente, visti i dati) [0,1] in 10 parti, quindi $m=10$; dunque, nel caso che i dati seguano la ripartizione uniforme, deve essere $p_j = \frac{1}{10}$ e quindi $Np_j = 10$. Perciò il test del chi-quadro darà risultati attendibili. La tabella è

k	O_k	E_k	$O_k - E_k$	$\frac{(O_k-E_k)^2}{E_k}$
1	7	10	−3	9/10
2	13	10	3	9/10
3	6	10	−4	16/10
4	16	10	6	36/10
5	13	10	3	9/10
6	9	10	−1	1/10
7	17	10	7	49/10
8	4	10	−6	36/10
9	8	10	−2	4/10
10	7	10	−3	9/10

Dunque $T = \sum_{k=1}^{10} \frac{(O_k-E_k)^2}{E_k} = 17.8$, mentre $\chi^2_{0.95}(9) = 16.92$.

Pertanto, al livello 0.05 dobbiamo respingere l'ipotesi che la scelta dei dati sia stata fatta a caso (secondo la ripartizione uniforme).

Se avessimo scelto il livello 0.01, avremmo ottenuto $\chi_{0.99}(9) = 21.66$ e quindi non avremmo potuto respingere l'ipotesi.

2. Un'urna contiene un numero sconosciuto N di monete. Per stabilirne il numero si effettua il bizzarro esperimento aleatorio che consiste nel lanciare ciascuna delle monete e contare il numero di teste uscite durante i lanci. Ripetendo 10 volte questo esperimento, si ottengono i seguenti risultati:

692, 695, 665, 674, 719, 680, 686, 658, 691, 645.

Stimare il numero N di monete presenti nell'urna.

<u>Risposta</u> Se le monete sono equilibrate, i nostri dati costituiscono un campione statistico $(X_1 \ldots X_{10})$ di taglia 10, con le X_i v.a. di legge $B(N, \frac{1}{2})$. Poiché quindi

$$\mathbb{E}[X_i] = \frac{N}{2}, \quad i = 1 \ldots 10,$$

si ha $N = 2\mathbb{E}[X_i] = 2m_1$ (m_1 = momento del 1° ordine). Possiamo stimare m_1 con la media empirica $\bar{X}$, che nel nostro caso vale 680.5, e dunque $N \simeq 1361$.

3. Si supponga che i tempi di attesa alla fermata di un autobus, in minuti, seguano una legge esponenziale. Nel corso di una giornata si raccolgono 20 misure del tempo di attesa, ottenendo un tempo medio $\bar{x} = 13$ min. Trovare un intervallo di fiducia per il parametro λ della legge esponenziale a un livello α del 95%.

<u>Risposta</u> Lo stimatore del parametro λ è $T = \frac{1}{\bar{X}}$. L'intervallo di fiducia è della forma $\left\{\left|\frac{1}{\bar{X}} - \lambda\right| \leq \delta\right\}$, e si deve scegliere δ in modo che

$$P^{\lambda}\left(\left|\frac{1}{\bar{X}} - \lambda\right| \leq \delta\right) = 0.05.$$

Posto $\bar{Y} = \lambda \bar{X}$, se $\bar{X}$ aveva legge $\mathcal{E}(\lambda)$ allora $\bar{Y}$ ha legge $\mathcal{E}(1)$. Perciò deve essere

$$P^{1}\left(|1-\bar{Y}| \leq \delta \bar{X}\right) = P^{1}\left(1-\delta\bar{X} \leq \bar{Y} \leq 1+\delta\bar{X}\right).$$

Sostituendo i dati, si ottiene

$$P^{1}\left(|1-\bar{Y}| \leq 13\,\delta\right) = \int_{1-13\delta}^{1+13\delta} e^{-t}\,dt = e^{-1+13\delta} - e^{-1-13\delta} = 0.05$$

se e solo se $\delta \cong 0.005223$. Quindi i tempi di attesa alla fermata seguono approssimativamente una legge $\mathcal{E}(\lambda)$ con

$$\frac{1}{13} - 0.005223 \leq \lambda \leq \frac{1}{13} + 0.005223,$$

almeno a livello $\alpha = 0.95$.

4. Il vettore aleatorio (X,Y) ha densità congiunta

$$f_{XY}(x,y) = \begin{cases} \frac{1}{8}(6-x-y) & \text{se } x \in]0,2[,\ y \in]2,4[, \\ 0 & \text{altrimenti.} \end{cases}$$

(i) Determinare le densità marginali f_X, f_Y.

(ii) Stabilire se le v.a. X, Y sono indipendenti.

Risposta:

(i) $$f_X(x) = \int_2^4 f_{XY}(x,y)\,dy = \int_2^4 \frac{1}{8}(6-x-y)\,dy =$$

$$= \frac{12}{8} - \frac{x}{4} - \left[\frac{y^2}{16}\right]_2^4 = \frac{3-x}{4} \quad \text{se } x \in]0,2[$$

mentre $f_X = 0$ per $x \notin]0,2[$.

$$f_Y(y) = \int_0^2 f_{XY}(x,y)\,dy = \int_0^2 \frac{1}{8}(6-x-y)\,dx =$$

$$= \frac{12}{8} - \left[\frac{x^2}{16}\right]_0^2 - \frac{y}{4} = \frac{5-y}{4} \quad \text{se } y \in]2,4[,$$

mentre $f_Y = 0$ per $y \notin]2,4[$.

(ii) Se $]a,b[\subset]0,2[$ e $]c,d[\subset]2,4[$ si ha

$$P(X \in]a,b[) = \int_a^b \frac{3-x}{4}\,dx = \frac{3}{4}(b-a) - \frac{1}{8}(b^2-a^2),$$

$$P(Y \in]c,d[) = \int_c^d \frac{5-y}{4}\,dy = \frac{5}{4}(d-c) - \frac{1}{8}(d^2-c^2),$$

mentre

$$P((X,Y) \in]a,b[\times]c,d[) = \int_a^b \int_c^d \frac{1}{8}(6-x-y)\,dy\,dx =$$

$$= \frac{3}{4}(b-a)(d-c) - \frac{d-c}{16}(b^2-a^2) - \frac{b-a}{16}(c^2-d^2),$$

e scegliendo ad esempio $b=1, a=0$, $d=3$, $c=2$, si ha

$$P(X\in]0,1[) = \frac{5}{8}$$

$$P(Y\in]2,3[) = \frac{5}{8}$$

$$P((X,Y)\in]0,1[\times]2,3[) = \frac{3}{8} \neq \frac{5}{8}\cdot\frac{5}{8} = P(X\in]0,1[)\, P(Y\in]2,3[).$$

5. Indagini medico-statistiche mostrano che lo 0.001% degli italiani è affetto da AIDS, mentre lo 0.01% degli italiani appartiene a una delle cosiddette "categorie a rischio". Si sa inoltre che l'80% dei malati di AIDS appartiene ad una categoria a rischio. Qual è la probabilità che un italiano appartenente ad una categoria a rischio sia malato di AIDS?

Risposta

Siano M, R gli eventi "essere malati di AIDS" e "essere in una categoria a rischio". Si sa che

$$P(M) = \frac{1}{100'000}, \quad P(R) = \frac{1}{10'000}, \quad P(R|M) = \frac{4}{5}.$$

Dobbiamo calcolare $P(M|R)$. Dalla formula di Bayes

$$P(M|R) = \frac{P(R|M)\,P(M)}{P(R)} = \frac{\frac{4}{5}\cdot\frac{1}{100.000}}{\frac{1}{10'000}} = \frac{4}{50} = \frac{2}{25} = 8\%.$$