ESERCIZI DI STATISTICA

Soluzioni degli esercizi sulla verifica di ipotesi.

A cura di Nazareno Maroni

Soluzione dell'esercizio 1.

a) Si ha che

$$f_Y(y) = \frac{1}{\sqrt{10\pi}} e^{-\frac{(y-\mu)^2}{10}} = \frac{1}{\sqrt{10\pi}} e^{-\frac{\mu^2}{10}} e^{-\frac{y^2}{10}} e^{\frac{\mu y}{5}}$$

La distribuzione appartiene alla famiglia esponenziale, applichiamo il teorema 9.5 pag. 423 (M.G.B.): $c(\mu) = \frac{\mu}{5}$, $d(y) = y$; la statistica da usare è $T = \sum_{i=1}^{n} Y_i$ e poiché $c(\mu)$ è una funzione monotona crescente la regione critica sarà $C = \{(Y_1, \dots, Y_n) | T > k\}$ con k tale che l'errore di I specie sia pari ad α.

$$0,05 = \mathbb{P}\Big(\sum_{i=1}^{n} Y_i > k \Big| \mu_0\Big) = \mathbb{P}\Big(\sum_{i=1}^{n} Y_i - n\mu_0 > k - n\mu_0 \Big| \mu_0\Big) =$$
$$= \mathbb{P}\Big(\frac{\sum_i Y_i - n\mu_0}{\sqrt{5n}} > \frac{k - n\mu_0}{\sqrt{5n}} \Big| \mu_0\Big) = \mathbb{P}\Big(N(0,1) > \frac{k - n\mu_0}{\sqrt{5n}}\Big),$$

quindi $\dfrac{k - n\mu_0}{\sqrt{5n}} = z_{1-\alpha} = 1,65 \quad \Rightarrow \quad k = n\mu_0 + \sqrt{5n}1,65 = 156,5$. Quindi il test uniformemente più potente di ampiezza $\alpha = 0,05$ è dato dalla regione critica $C = \{(Y_1, \dots, Y_n) | \overline{Y} > 7,825\}$.

b)

$$\pi(\mu) = \mathbb{P}(\text{Rifiutare } \mathbb{H}_0 | \mu) = \mathbb{P}(\overline{Y} > 7,825 | \mu) = \mathbb{P}\Bigg(\frac{\overline{Y} - \mu}{\frac{\sqrt{5}}{\sqrt{n}}} > \frac{7,825 - \mu}{\frac{\sqrt{5}}{\sqrt{n}}}\Bigg) =$$
$$= \mathbb{P}\Bigg(N(0,1) > \frac{7,825 - \mu}{\frac{\sqrt{5}}{\sqrt{n}}}\Bigg)$$

Abbiamo che:

$$\begin{aligned} \pi(7,5) &= \mathbb{P}(N(0,1) > 0,65) = 0,2578, \\ \pi(8) &= \mathbb{P}(N(0,1) > -0,35) = 0,6368, \\ \pi(8,5) &= \mathbb{P}(N(0,1) > -1,35) = 0,9115 \end{aligned}$$

e che

$$\pi(9) = \mathbb{P}(N(0,1) > -2,35) = 0,9906.$$

Soluzione dell'esercizio 2. Il test più potente è quello di Neymann-Pearson. Calcoliamo il rapporto tra le verosimiglianze:

$$
\begin{aligned}
\frac{L(\mu_0)}{L(\mu_1)} \leqslant k \quad &\Leftrightarrow \quad \frac{e^{-\frac{1}{50}\sum_i (x_i-\mu_0)^2}}{e^{-\frac{1}{50}\sum_i (x_i-\mu_1)^2}} \leqslant k \quad \Leftrightarrow \quad e^{-\frac{1}{50}(\sum_i (x_i-\mu_0)^2 - \sum_i (x_i-\mu_1)^2)} \leqslant k \\
&\Leftrightarrow \quad \sum_{i=1}^n (x_i-\mu_0)^2 - \sum_{i=1}^n (x_i-\mu_1)^2) \geqslant k_1 \\
&\Leftrightarrow \quad n\mu_o^2 - 2\mu_0 \sum_{i=1}^n x_i - n\mu_1^2 + 2\mu_1 \sum_{i=1}^n x_i \geqslant k_2 \\
&\Leftrightarrow \quad \sum_{i=1}^n x_i(2\mu_1 - 2\mu_0) \geqslant k_3 \\
&\Leftrightarrow \quad \sum_{i=1}^n x_i \leqslant k_4 \quad \Leftrightarrow \quad \frac{1}{n}\sum_{i=1}^n x_i \leqslant k^*
\end{aligned}
$$

$$
\begin{aligned}
0,025 &= \mathbb{P}(\overline{X} \leqslant k^* | \mu_0) = \mathbb{P}\left(\frac{\overline{X}-\mu_0}{\frac{5}{\sqrt{n}}} \leqslant \frac{k^*-\mu_0}{\frac{5}{\sqrt{n}}} \Big| \mu_0 \right) = \\
&= \mathbb{P}\left(N(0,1) \leqslant \frac{k^*-\mu_0}{\frac{5}{\sqrt{n}}} \right) \quad \Rightarrow \quad \frac{k^*-\mu_0}{\frac{5}{\sqrt{n}}} = z_{0,025} = -1,96 \\
&\Rightarrow \quad k^* = 10 - \frac{9,8}{\sqrt{n}}.
\end{aligned}
$$

$$
\begin{aligned}
0,025 &= \mathbb{P}(\text{Accettare } \mathbb{H}_0 | \mu_1) = \mathbb{P}(\overline{X} > k^* | \mu_1) = \mathbb{P}\left(\frac{\overline{X}-\mu_1}{\frac{5}{\sqrt{n}}} > \frac{k^*-\mu_1}{\frac{5}{\sqrt{n}}} \Big| \mu_1 \right) = \\
&= \mathbb{P}\left(N(0,1) > \frac{k^*-\mu_1}{\frac{5}{\sqrt{n}}} \right) \quad \Rightarrow \quad \frac{k^*-\mu_1}{\frac{5}{\sqrt{n}}} = z_{0,975} = 1,96 \\
&\Rightarrow \quad k^* = 5 + \frac{9,8}{\sqrt{n}} \\
&\Rightarrow \quad 10 - \frac{9,8}{\sqrt{n}} = 5 + \frac{9,8}{\sqrt{n}} \quad \Rightarrow \quad \sqrt{n} = 3,92 \quad \Rightarrow \quad n = 16.
\end{aligned}
$$

Il test più potente di ampiezza $\alpha = 0,025$ è quello con regione critica $C = \{(X_1, \ldots, X_n) | \overline{X} \leqslant 7,5\}$.

Soluzione dell'esercizio 3. Le ipotesi sono semplici e quindi il test più potente è quello di Neymann-Pearson. Calcoliamo il rapporto di verosimiglianza.

$$
\begin{aligned}
\frac{L(\sigma_0^2)}{L(\sigma_1^2)} &= \frac{\sigma_1 e^{-\frac{\sum (x_i-\mu)^2}{2\sigma_0^2}}}{\sigma_0 e^{-\frac{\sum (x_i-\mu)^2}{2\sigma_1^2}}} = \frac{\sigma_1}{\sigma_0} e^{-\frac{\sum (x_i-\mu)^2}{2\sigma_0^2} + \frac{\sum (x_i-\mu)^2}{2\sigma_1^2}} \leqslant k \\
&\Rightarrow \quad -\frac{1}{2\sigma_0^2}\sum_{i=1}^n (x_i-\mu)^2 + \frac{1}{2\sigma_1^2}\sum_{i=1}^n (x_i-\mu)^2 \leqslant k_1 \\
&\Rightarrow \quad \sum_{i=1}^n (x_i-\mu)^2 \left(\frac{1}{2\sigma_1^2} - \frac{1}{2\sigma_0^2} \right) \leqslant k_1 \\
&\Rightarrow \quad \sum_{i=1}^n (x_i-\mu)^2 \left(\frac{\sigma_0^2 - \sigma_1^2}{2\sigma_0^2\sigma_1^2} \right) \leqslant k_1 \quad \Rightarrow \quad \sum_{i=1}^n (x_i-\mu)^2 \geqslant k^*
\end{aligned}
$$

$$0,05 = \mathbb{P}\Big(\sum_{i=1}^{n}(x_i-\mu)^2 \geqslant k^* \Big| \sigma_0^2\Big) = \mathbb{P}\Big(\sum_{i=1}^{n}\Big(\frac{x_i-\mu}{\sigma_0}\Big)^2 \geqslant \frac{k^*}{\sigma_0^2}\Big|\sigma_0^2\Big) =$$
$$= \mathbb{P}\Big(\chi_n^2 \geqslant \frac{k^*}{\sigma_0^2}\Big) \quad \Rightarrow \quad \frac{k^*}{\sigma_0^2} = \chi^2_{n,0,95}$$

Notate che sotto l'ipotesi $\mathbb{H}_0$ le X_i sono $N(\mu,\sigma_0^2)$ e quindi $\dfrac{X_i-\mu}{\sigma_0} \sim N(0,1)$ e $\displaystyle\sum_{i=1}^{n}\Big(\frac{X_i-\mu}{\sigma_0}\Big)^2 \sim \chi_n^2$. Quindi il test più potente di livello $\alpha = 0,05$ è dato dal test con regione critica $C = \Big\{(X_1,\dots,X_n)\Big| \displaystyle\sum_{i=1}^{n}(X_i-\mu)^2 \geqslant \sigma_0^2\chi^2_{n,0,95}\Big\}$.

Soluzione dell'esercizio 4. La funzione di verosimiglianza è

$$L(\lambda) = \prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}e^{-\lambda} = \frac{\lambda^{\sum x_i}}{\prod x_i!}e^{-n\lambda} = \frac{e^{-n\lambda}}{\prod x_i!}e^{\sum x_i \ln\lambda}$$

La funzione di verosimiglianza appartiene alla famiglia esponenziale e la funzione $\ln\lambda$ è monotona crescente, applichiamo il teorema 9.5 pag. 423 (M.G.B.). La statistica da usare è $T = \displaystyle\sum_{i=1}^{n} X_i$.

$$0,05 = \mathbb{P}\Big(\sum_{i=1}^{n} X_i > k \Big| \lambda_0\Big) = \mathbb{P}(Po(n\lambda_0) > k) = \sum_{j=k+1}^{+\infty}\frac{(n\lambda_0)^j}{j!}e^{-n\lambda_0} =$$
$$= 1 - \sum_{j=0}^{k}\frac{(n\lambda_0)^j}{j!}e^{-n\lambda_0} \quad \Rightarrow \quad k = q_{0,95}^{Po(n\lambda_0)}$$

Quindi il test uniformemente più potente di livello $\alpha = 0,05$ è dato dalla regione critica $C = \Big\{(X_1,\dots,X_n)\Big| \displaystyle\sum_{i=1}^{n} X_i > k\Big\}$, dove k è il quantile di livello $0,95$ di una Poisson di parametro $n\lambda_0$.

Soluzione dell'esercizio 5. Riscriviamo la funzione di densità:

$$f(x,\theta) = (\theta+2)x^{\theta+1} = (\theta+2)e^{(\theta+1)\ln x},$$

quindi appartiene alla famiglia esponenziale e possiamo applicare il teorema: $c(\theta) = \theta+1$ è una funzione monotona crescente e la statistica da usare è $T = \displaystyle\sum_{i=1}^{n}\ln X_i$. Calcoliamo la distribuzione di $Y = -\ln X$.

$$f_Y(y) = (\theta+2)e^{-(\theta+1)y}e^{-y} = (\theta+2)e^{-(\theta+2)y} \sim Esp(\theta+2)$$

Quindi $\displaystyle\sum_{i=1}^{n} Y_i = -\sum_{i=1}^{n}\ln X_i \sim Gamma(n,\theta+2)$.

$$\alpha = \mathbb{P}\Big(\sum_{i=1}^{n}\ln X_i > k \Big|\theta_0\Big) = \mathbb{P}\Big(-\sum_{i=1}^{n}\ln X_i < k_1\Big|\theta_0\Big) =$$
$$= \mathbb{P}(Gamma(n,\theta_0+2) < k_1) = \int_0^{k_1}\frac{(\theta_0+2)^n}{\Gamma(n)}x^{n-1}e^{-(\theta_0+2)x}\,\mathrm{d}x$$

Il test uniformemente più potente di ampiezza α è dato dalla regione critica
$C = \left\{(X_1, \ldots, X_n) \Big| \sum_{i=1}^{n} \ln X_i > k\right\}$, dove $k_1 = -k$ è il quantile di livello α di una $Gamma(n, \theta_0 + 2)$.