Università degli studi Roma Tre - Corso di Laurea in Matematica
Tutorato di ST1 - A.A. 2006/2007
Docente: Prof.ssa E. Scoppola - Tutore: Dott. Nazareno Maroni

Soluzioni del tutorato n.4 del 27/3/2007

Esercizio 1. Sappiamo che, per una popolazione normale di ampiezza n, di media μ e varianza σ^2, lo stimatore di massima verosimiglianza di $\theta = (\mu, \sigma^2)$ è

$$\hat{\Theta} = \left(\overline{X}, \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X})^2 \right)$$

Indicando con μ_i, $i = 1, 2, 3, 4$ le medie delle quattro popolazioni e con $\sigma_i^2 = \sigma^2$, $i = 1, 2, 3, 4$ le rispettive varianze, si ha che:

$$\hat{\mu}_1 = \overline{X}_1 \quad \hat{\mu}_2 = \overline{X}_2 \quad \hat{\mu}_3 = \overline{X}_3 \quad \hat{\mu}_4 = \overline{X}_4$$

Per l'invarianza degli stimatori di massima verosimiglianza, possiamo ricavare gli stimatori di a, b e c dal seguente sistema:

$$\begin{cases} \hat{a} + \hat{b} + \hat{c} = \hat{\mu}_1 = \overline{X}_1 \\ \hat{a} + \hat{b} - \hat{c} = \hat{\mu}_2 = \overline{X}_2 \\ \hat{a} - \hat{b} + \hat{c} = \hat{\mu}_3 = \overline{X}_3 \end{cases} \Rightarrow \begin{cases} \hat{a} = \frac{\overline{X}_2 + \overline{X}_3}{2} \\ \hat{b} = \frac{\overline{X}_1 - \overline{X}_3}{2} \\ \hat{c} = \frac{\overline{X}_1 - \overline{X}_2}{2} \end{cases}$$

Ovviamente altri stimatori sono possibili se nel sistema si usa la relazione che lega i parametri alla media della quarta popolazione. Come stimatore di massima verosimiglianza della varianza possiamo prendere

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \overline{X}_4)^2$$

non avendo usato l'informazione derivante dalla quarta popolazione. Un altro stimatore può essere ottenuto utilizzando l'invarianza:

$$\sigma^2 = \frac{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \sigma_4^2}{4} \quad \Rightarrow \quad \hat{\sigma}^2 = \frac{1}{4n} \sum_{i=1}^{4} \sum_{j=1}^{n} (X_{ij} - \overline{X}_i)^2$$

Potrebbe essere interessante confrontare gli errori quadratici medi dei due stimatori e valutare se uno dei due è uniformemente minore.

Esercizio 2.

(a) $\mathbb{E}(t_1(X)) = \mathbb{E}(X) = \theta$, quindi è corretto (non distorto). $\mathbb{E}(t_2(X)) = \mathbb{E}\left(\frac{1}{2}\right) = \frac{1}{2}$, quindi è distorto.

(b) $MSE(t_1(X)) = \mathbb{E}((X-\theta)^2) = Var(X) = \theta(1-\theta)$. $MSE(t_2(X)) = \mathbb{E}\left(\left(\frac{1}{2}-\theta\right)^2\right) = \left(\frac{1}{2} - \theta\right)^2$. Non c'è uno tra i due stimatori che abbia l'errore quadratico medio uniformemente più piccolo.

(c)

$$
\begin{aligned}
L(\theta) &= \theta^{\sum_{i=1}^n x_i}(1-\theta)^{n-\sum_{i=1}^n x_i}\prod_{i=1}^n \mathbb{1}_{\{0,1\}}(x_i)\\
\ln L(\theta) &= \sum_{i=1}^n x_i \ln\theta + \left(n-\sum_{i=1}^n x_i\right)\ln(1-\theta)\\
\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta) &= \frac{1}{\theta}\sum_{i=1}^n x_i - \frac{1}{1-\theta}\left(n-\sum_{i=1}^n x_i\right) \quad \Rightarrow \quad \hat{\Theta}=\overline{X}\\
\mathbb{E}(\hat{\Theta}) &= \theta \quad \text{quindi lo stimatore è corretto,} \quad MSE(\hat{\Theta}) = Var(\overline{X}) = \frac{\theta-\theta^2}{n}
\end{aligned}
$$

Esercizio 3.

(a) $\mathbb{E}(X)=\theta$ e quindi lo stimatore ottenuto col metodo dei momenti è $\hat{\Theta}=\overline{X}$.

(b)

$$
\begin{aligned}
L(\theta) &= \frac{1}{2^n}\exp\left\{-\sum_{i=1}^n |x_i-\theta|\right\}\\
\ln L(\theta) &= -n\ln 2 - \sum_{i=1}^n |x_i-\theta| = -n\ln 2 - \sum_{i=1}^n \sqrt{(x_i-\theta)^2}\\
\frac{\mathrm{d}}{\mathrm{d}\theta}\ln L(\theta) &= \sum_{i=1}^n \frac{(x_i-\theta)}{|x_i-\theta|} = \sum_{i=1}^n sgn(x_i-\theta)
\end{aligned}
$$

Dobbiamo trovare un valore di θ che annulla la derivata, ovvero, essendo la derivata una somma di segni (± 1, con $sgn(0)=0$), si devono avere tanti 1 quanti -1. Quindi se $n=2k+1$ deve essere (indicando con $X_{(i)}$, $i=1,\ldots,n$ la statistica d'ordine) $\hat{\Theta}=X_{(k+1)}$, infatti così si hanno k -1 e k $+1$ e $sgn(X_{(k+1)}-\theta)=0$; se $n=2k$, qualunque valore di θ tale che $x_{(k)}<\theta<x_{(k+1)}$ è tale che ci sono k -1 e k $+1$. Riassumendo:

$$
\begin{cases}
\hat{\Theta} = X_{(k+1)} & \text{se } n=2k+1\\
\hat{\Theta} = \frac{1}{2}(X_{(k+1)}+X_{(k)}) & \text{se } n=2k
\end{cases}
$$

Esercizio 4.

(a) $\mathbb{E}(X)=\dfrac{\theta}{\theta+1}$, quindi lo stimatore trovato col metodo dei momenti è $\hat{\Theta}=\dfrac{\overline{X}}{1-\overline{X}}$.

(b)

$$L(\theta) = \theta^n \left(\prod_{i=1}^{n} x_i \right)^{\theta - 1}$$

$$\ln L(\theta) = n \ln \theta + (\theta - 1) \sum_{i=1}^{n} \ln x_i$$

$$\frac{\mathrm{d}}{\mathrm{d}\theta} \ln L(\theta) = \frac{n}{\theta} + \sum_{i=1}^{n} \ln x_i = 0 \quad \Leftrightarrow \quad \hat{\Theta} = -\frac{n}{\sum_{i=1}^{n} \ln X_i}$$

Per calcolare il suo valore atteso, troviamo prima la distribuzione di $Y = -\ln X$:

$$f_Y(y) = f_X(e^{-y})e^{-y} = \theta e^{-\theta y} \quad \Rightarrow \quad Y \sim Esp(\theta)$$

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

$$\begin{aligned}
\mathbb{E}(\hat{\Theta}) &= \mathbb{E}\left(-\frac{n}{\sum_{i=1}^{n} \ln X_i} \right) = \mathbb{E}\left(\frac{n}{Z} \right) = \int_0^{+\infty} \frac{n}{z} \frac{\theta^n}{\Gamma(n)} z^{n-1} e^{-\theta z} \, \mathrm{d}z = \\
&= \frac{n}{\Gamma(n)} \theta \int_0^{+\infty} \theta^{n-1} z^{(n-1)-1} e^{-\theta z} \, \mathrm{d}z = n\theta \frac{\Gamma(n-1)}{\Gamma(n)} = \frac{n}{n-1}\theta \\
\mathbb{E}(\hat{\Theta}^2) &= \int_0^{+\infty} \frac{n^2}{z^2} \frac{\theta^n}{\Gamma(n)} z^{n-1} e^{-\theta z} \, \mathrm{d}z = \frac{n^2 \theta^2}{(n-1)(n-2)} \\
Var(\hat{\Theta}) &= \frac{n^2 \theta^2}{(n-1)(n-2)}
\end{aligned}$$

Lo stimatore non è corretto, quindi possiamo correggerlo prendendo

$$\hat{\Theta}^* = \frac{n-1}{n} \hat{\Theta}$$

Infatti è

$$\mathbb{E}(\hat{\Theta}^*) = \frac{n-1}{n} \mathbb{E}(\hat{\Theta}) = \frac{n-1}{n} \frac{n}{n-1} \theta = \theta.$$