Politecnico di Milano – Facoltà di Ingegneria Informatica
Anno Accademico 2008/2009
Corso di Statistica (2L) per INF e TEL
Docente: Antonio Pievatolo; esercitazioni: Raffaele Argiento

Esercitazione del 27/03/09

Esercizio 1

Sia $X_1, \ldots, X_n$ un campione da una popolazione $\mathrm{N}(\mu, \sigma)$, con $\mu \in \mathbb{R}$ e $\sigma > 0$ entrambi incogniti. Si può mostrare che

$$S^2 := \frac{1}{n-1} \sum_{i=1}^{n} X_i$$

è lo stimatore UMVUE per la varianza di popolazione. S^2 è efficiente?

SOLUZIONE
Per verificare se S^2 è efficiente bisogna confrontare la sua varianza con il limite inferiore individuato dal teorema di Fréchet-Cramér-Rao. Innanzitutto bisogna mostrare che lo stimatore S^2 soddisfa le condizioni di regolarità del teorema per il modello Gaussiano:

i. σ^2 appartiene ad un intervallo aperto di $\mathbb{R}$.

ii. $\mathcal{S} = \{x : f_X(x; \mu, \sigma^2)\}$ non dipende da σ^2.

iii. la funzione $\sigma^2 \to f_X(x; \mu, \sigma^2)$ è derivabile rispetto a σ^2 per ogni x e μ fissati.

iv. $\mathbb{E}\left(\frac{\partial}{\partial \sigma^2} \log f_X(X; \mu, \sigma^2)\right) = 0$.

v. $\mathbb{E}\left(\left[\frac{\partial}{\partial \sigma^2} \log f_X(X; \mu, \sigma^2)\right]^2\right) \in (0, +\infty)$.

vi. $\mathbb{E}\left(S^2 \cdot L_{\sigma^2}(X_1, \ldots, X_n)\right) = 1$.

VERIFICA

i. Banale.

ii. Banale.

iii. Banale.

iv. Ricordiamo che $f_X(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$ e poniamo

$$Y = \frac{\partial}{\partial \sigma^2} \log f_X(X; \mu, \sigma^2) = \frac{\partial}{\partial \sigma^2}\left\{-\frac{1}{2}(\log(2\pi) + \log \sigma^2) - \frac{(X-\mu)^2}{2\sigma^2}\right\} = -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(X-\mu)^2.$$

Quindi

$$\mathbb{E}(Y) = E\left(-\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(X-\mu)^2\right) = -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^2} \cdot \mathbb{E}\left(\left[\frac{X-\mu}{\sigma}\right]^2\right) = -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^2} \cdot \mathbb{E}\left(Z^2\right),$$

dove $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$. Di conseguenza $Z^2 \sim \chi^2(1)$ e $\mathbb{E}(Z^2) = 1$. Si conclude che $\mathbb{E}(Y) = 0$.

v. Si osservi che $\mathbb{E}\left(\left[\frac{\partial}{\partial\sigma^2}\log f_X(X;\mu,\sigma^2)\right]^2\right) = \mathbb{E}(Y^2)$. Inoltre, dato che $\mathbb{E}(Y) = 0$, allora $\mathbb{E}(Y^2) = \mathrm{Var}(Y)$.

$$\mathrm{Var}(Y) = \mathrm{Var}\left(-\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(X-\mu)^2\right) = \mathrm{Var}\left(\frac{1}{2\sigma^2}(Z^2-1)\right) = \frac{1}{4\sigma^4}\mathrm{Var}(Z^2) = \frac{1}{2\sigma^4}.$$

vi. La funzione di verosimiglianza è

$$L_{\sigma^2}(X_1,\dots,X_n) = \prod_{i=1}^n f_X(X_i;\mu,\sigma^2) \Rightarrow \log L_{\sigma^2}(X_1,\dots,X_n) = \sum_{i=1}^n \log f_X(X_i;\mu,\sigma^2).$$

Quindi

$$\begin{aligned}
\frac{\partial}{\partial\sigma^2}\log L_{\sigma^2}(X_1,\dots,X_n) &= \sum_{i=1}^n \frac{\partial}{\partial\sigma^2}\log f_X(X_i;\mu,\sigma^2) \\
&= \sum_{i=1}^n \left\{-\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}(X_i-\mu)^2\right\} \\
&= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^n (X_i-\mu)^2 \\
&= -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\left((n-1)S^2 + n(\bar{X}-\mu)^2\right).
\end{aligned}$$

Nell'ultima uguaglianza, abbiamo usato la scomposizione

$$\sum_{i=1}^n (X_i-\mu)^2 = \sum_{i=1}^n (X_i-\bar{X}) + n(\bar{X}-\mu)^2.$$

Si ha dunque che:

$$\begin{aligned}
&\mathbb{E}\left(S^2 \cdot L_{\sigma^2}(X_1,\dots,X_n)\right) = \mathbb{E}\left(S^2\cdot\left(-\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\left((n-1)S^2 + n(\bar{X}-\mu)^2\right)\right)\right) = \\
&= \mathbb{E}\left(-\frac{n}{2(n-1)}\left((n-1)\frac{S^2}{\sigma^2}\right) + \frac{1}{2(n-1)}\left[\left((n-1)\frac{S^2}{\sigma^2}\right)^2 + (n-1)\frac{S^2}{\sigma^2}\cdot\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)^2\right]\right) = \\
&= -\frac{n}{2(n-1)}\mathbb{E}(Q) + \frac{1}{2(n-1)}\left[\mathbb{E}(Q^2) + \mathbb{E}(Q\cdot Z^2)\right],
\end{aligned}$$

dove abbiamo posto $Q = (n-1)\frac{S^2}{\sigma^2} \sim \chi^2(n-1)$ e $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$, e vale $Q \perp Z$.

Si valuta facilmente l'ultima espressione, ottenendo:

$$\mathbb{E}\left(S^2 \cdot L_{\sigma^2}(X_1,\dots,X_n)\right) = -\frac{n}{2(n-1)}(n-1) + \frac{1}{2(n-1)}\left[2(n-1) + (n-1)^2 + n - 1 +\right] = 1.$$

Tutte le condizioni (i)-(vi) sono verificate, possiamo dunque calcolare l'informazione di Fisher per confrontarla con la varianza di S^2: per l'informazione di Fisher si ottiene

$$I_n(\sigma^2) = nI(\sigma^2) = n\mathbb{E}\left(\left[\frac{\partial}{\partial\sigma^2}\log f_X(X;\mu,\sigma^2)\right]^2\right) = n\mathbb{E}(Y^2) = \frac{n}{2\sigma^4};$$

per la varianza

$$\mathrm{Var}(S^2) = \frac{\sigma^4}{(n-1)^2}\mathrm{Var}\left((n-1)\frac{S^2}{\sigma^2}\right) = \frac{\sigma^4}{(n-1)^2}\mathrm{Var}(Q) = \frac{2\sigma^4}{(n-1)}.$$

Dove, $Q = (n-1)\frac{S^2}{\sigma^2} \sim \chi^2(n-1)$. Si ha dunque che

$$\mathrm{Var}(S^2) > \frac{(\sigma^2)'}{nI(\sigma^2)} = \frac{1}{nI(\sigma^2)}$$

di conseguenza S^2 non è lo stimatore efficiente!

∎

Esercizio 2

Vedi Esempio 6.9 delle dispense della Proff.sa Epifani

Esercizio 3

Vedi Esempio 6.7 delle dispense della Proff.sa Epifani