Efficienza degli stimatori di ML

Oss

Per parlare di efficienza degli stimatori ML, osserviamo che questa sia strettamente legata alla matrice d'informazione

Def Stimatori B.U.E. $\rightsquigarrow$ best unbiased estimator

Uno stimatore T_n si dice B.U.E. se vale:

a) $E[T_n] = \theta_0$ $\rightsquigarrow$ Non distorto

b) $\forall T_n' : E[T_n'] = \theta_0$

$$Var(T_n) \leq Var(T_n')$$ $\rightarrow$ e' il piu' efficiente

Lemma

Se $\exists T_n : T_n$ e' B.U.E. Allora

T_n e' unico

Dim

Assumendo esistano oltre a T_n altri due stimatori B.U.E. T_n' e T_n''

Riscriviamo T_n'' come

$$T_n'' = \alpha T_n + (1-\alpha) T_n' \qquad \text{se } \alpha \in [0,1] \text{ e' non distorto}$$

Calcoliamo ora la varianza T_n''

$$Var(T_n'') = \alpha^2 Var(T_n) + (1-\alpha)^2 Var(T_n') + 2\alpha(1-\alpha) Cov(T_n, T_n')$$

Ponendo $Var(T_n) = Var(T_n') = \sigma^2$ abbiamo che

$$Var(T_n'') = \sigma^2\alpha^2 + (1-\alpha)^2\sigma^2 + 2\alpha(1-\alpha) Cov(T_n, T_n') =$$

$$= \sigma^2 \left(\alpha^2 + (1-\alpha)^2 + 2\alpha(1-\alpha) \frac{Cov(T_n, T_n')}{\sqrt{Var(T_n) Var(T_n')}} \right)$$

$\sigma^2 = \sqrt{\sigma^4} = \sqrt{\sigma^2 \cdot \sigma^2}$

$(\alpha + (1-\alpha))^2 = 1^2$

Per Cauchy-Swartz noi abbiamo che

$$-1 \leq \frac{Cov(T_n, T_n')}{\sqrt{Var(T_n) Var(T_n')}} \leq 1 \rightarrow ☆$$

$\left[Var(\text{entro})\right] \geq \frac{1}{I_n(\theta_0)} \rightarrow \frac{1}{n I(\theta_0)}$

Nuovo

Quindi se $☆ \leq 1$ abbiamo che $Var(T_n'') \leq Var(T_n')$ → Assurdo

↳ Per Hp T_n e' un B.U.E.

Sotto le condizioni di regolarita' emerge che la varianza minima degli stimatori non distorti coincida con la varianza minima degli stimatori di massima verosimiglianza

Oss

La questione ha senso solo per gli stimatori non distorti, altrimenti qualsiasi stimatore con valore identicamente costante non potrebbe essere migliorato

Teo Limite inferiore di Cramer-Rao

Sotto le condizioni di regolarita'

Prendendo $T_n : E[T_n] = \theta_0$

Allora

1) Stimatore non distorto

$$Var(T_n) \geq \frac{1}{I_n(\theta_0)} \quad \text{oppure} \quad Var(T_n) \leq \frac{1}{n I_1(\theta_0)}$$

2) Stimatore distorto

$$Var(T_n) \leq \frac{\left(\frac{\partial}{\partial \theta_0} E[T_n]\Big|_{\theta=\theta_0}\right)^2}{I_n(\theta_0)}$$

Dim

Sappiamo che

dis Cauchy-Schwarz

$$Cov(X,Y)^2 \leq Var(X)Var(Y)$$

Allora possiamo scrivere

$$Var(Y) \geq \frac{Cov(X,Y)^2}{Var(X)}$$

Ponendo

$$Y = T_n \quad e \quad X = S_n = \frac{\partial}{\partial\theta} \log L(\theta; X_1, \dots, X_n)$$

e ricordando che

$$E[S_n]\Big|_{\theta=\theta_0} = 0$$

$$E[S_n^2]\Big|_{\theta=\theta_0} = I_n(\theta_0) = -E\left[\frac{\partial}{\partial\theta} S_n\right]\Big|_{\theta=\theta_0}$$

Otteniamo ora che

$$Var(T_n) \geq \frac{Cov(S_n, T_n)^2}{\underbrace{I_n(\theta_0)}_{Var(S_n)}}$$

Dimostriamo ora che $Cov(S_n, T_n) = \frac{\partial}{\partial\theta} E[T_n]$

Osserviamo ora che

Si porta dentro la derivata come ammesso dalle condizioni di regolarità

$$\frac{\partial}{\partial\theta_0} E[T_n] = \frac{\partial}{\partial\theta_0} \int T_n(x_1, x_2, \dots, x_n) f(x_1, \dots, x_n; \theta)\, dx_1, \dots, dx_n = \int T_n(x_1, \dots, x_n) \frac{\partial}{\partial\theta_0} f(x_1, \dots, x_n; \theta)\, dx_1, \dots, dx_n =$$

Moltiplico e divido per $f(x_1, \dots, x_n; \theta)$

$$= \int T_n(x_1, \dots, x_n) \underbrace{\frac{\frac{\partial}{\partial\theta_0} f(x_1, \dots, x_n; \theta)}{f(x_1, \dots, x_n; \theta)}}_{\frac{\partial}{\partial\theta} \log L} f(x_1, \dots, x_n; \theta)\, dx_1, \dots, dx_n = E[T_n S_n]$$

Sapendo che

$$Cov(X, Y) = E[XY] - E[X]E[Y]$$

Allora

$$Cov(S_n, T_n) = E[S_n, T_n] - E[S_n]E[T_n]$$

Ma $E[T_n]\Big|_{\theta=\theta_0} = 0$ Quindi

$$Cov(S_n, T_n) = E[S_n, T_n] = \frac{\partial}{\partial\theta} E[T_n]$$

Quindi possiamo concludere che

$$Var(T_n) \geq \frac{\left(\frac{\partial}{\partial\theta} E[T_n]\right)^2}{I_n(\theta_0)}$$

Oss

Sotto le condizioni di regolarità, preso uno stimatore non distorto, questo raggiungerà il limite inferiore di Cramer-Rao se

Stiamo dicendo che lo stimatore e' efficiente se l'errore della stima e' linearmente proporzionale allo score

$$\exists \alpha(\theta) : \alpha(\theta)(T_n - \theta) = S_n(\theta)$$

Oss

Per il caso p>1 bisogna precisare che

T_n e' un vettore di dimensione $p \times 1$

$E[T_n] = \Psi(\theta)$ di dimensione $p \times 1$

Avremo quindi:

Matrice di varianza e covarianza

$$Var(T_n) = E\left[(T_n - \Psi(\theta))(T_n - \Psi(\theta))^T\right] \geq \tilde{J}\Psi(\theta_0)\, \tilde{I}_n^{-1}(\theta_0)\, \tilde{J}\Psi(\theta_0)^T$$