Kleine toevoeging aan Statistical Inference

Andreas

26 januari 2013

1 The Cramér-Rao Inequality

Stelling 1 (Cramér-Rao Inequality)**.** *Zij $\boldsymbol{g} : \mathbb{R}^d \to \mathbb{R}^m$ voldoende braaf in $\boldsymbol{\theta}$ en zij $\boldsymbol{T}$ een schatter voor $\boldsymbol{g}(\boldsymbol{\theta})$. Noteer met Φ de totale afgeleide van $E\boldsymbol{T}$ naar $\boldsymbol{\theta}$ in $\boldsymbol{\theta}$:*

$$\Phi_{ij} = \frac{\partial ET_i}{\partial \theta_j}(\boldsymbol{\theta}). \tag{1}$$

Zij $I_n(\boldsymbol{\theta})$ positief-definiet. Dan is

$$\mathcal{V}(\boldsymbol{T}) - \Phi I_n^{-1} \Phi^T \tag{2}$$

positief-definiet.

Bewijs. We tonen aan dat $\Phi = \mathrm{E}(\boldsymbol{T}\boldsymbol{S}_n^T)$:

$$\begin{aligned}
\Phi_{ij} &= \frac{\partial}{\partial \theta_j} \mathrm{E}T_i = \frac{\partial}{\partial \theta_j} \int_{\mathbb{R}} T_i(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{\theta}) \mathrm{d}\boldsymbol{x} \\
&= \int_{\mathbb{R}} T_i(\boldsymbol{x}) \frac{\partial}{\partial \theta_j} f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{\theta}) \mathrm{d}\boldsymbol{x} = \int_{\mathbb{R}} T_i(\boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{\theta}) \frac{\partial}{\partial \theta_j} \ln\left(f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{\theta})\right) \mathrm{d}\boldsymbol{x} \\
&= \mathrm{E}\left(T_i(\boldsymbol{x}) \frac{\partial}{\partial \theta_j} \ln\left(f_{\boldsymbol{X}}(\boldsymbol{x}, \boldsymbol{\theta})\right) \right) = \mathrm{E}\left(T_i(\boldsymbol{x}) S_{nj}(\boldsymbol{x}, \boldsymbol{\theta})\right) = \mathrm{E}\left(\boldsymbol{T}\boldsymbol{S}_n^T\right)_{ij}.
\end{aligned} \tag{3}$$

Definieer $\boldsymbol{Y} := \boldsymbol{T} - \Phi I_n^{-1} \boldsymbol{S}_n \in \mathbb{R}^m$. Merk op dat $\mathrm{E}\boldsymbol{S}_n = \boldsymbol{0}$, zodat $\mathrm{E}\boldsymbol{Y} = \mathrm{E}\boldsymbol{T}$. We berekenen de variantiematrix $\mathcal{V}(\boldsymbol{Y})$ van $\boldsymbol{Y}$:

$$\begin{aligned}
\mathcal{V}(\boldsymbol{Y}) &= \mathrm{E}\left[(\boldsymbol{Y} - \mathrm{E}\boldsymbol{Y})(\boldsymbol{Y} - \mathrm{E}\boldsymbol{Y})^T\right] = \mathrm{E}\left[(\boldsymbol{Y} - \mathrm{E}\boldsymbol{T})(\boldsymbol{Y} - \mathrm{E}\boldsymbol{T})^T\right] \\
&= \mathrm{E}\left[\left((\boldsymbol{T} - \mathrm{E}\boldsymbol{T}) - \Phi I_n^{-1} \boldsymbol{S}_n\right) \left((\boldsymbol{T} - \mathrm{E}\boldsymbol{T}) - \Phi I_n^{-1} \boldsymbol{S}_n\right)^T\right] \\
&= \mathcal{V}(\boldsymbol{T}) + \Phi I_n^{-1} \mathrm{E}(\boldsymbol{S}_n \boldsymbol{S}_n^T) I_n^{-1} \Phi^T - M - M^T \\
&= \mathcal{V}(\boldsymbol{T}) + \Phi I_n^{-1} \Phi^T - M - M^T,
\end{aligned} \tag{4}$$

waarbij we een vervelende term M hebben genoemd. We berekenen nu M, door eerst gebruik te maken van het feit dat $\mathrm{E}\boldsymbol{S}_n = \boldsymbol{0}$ en daarna van (3):

$$M = \Phi I_n^{-1} \mathrm{E}\left(\boldsymbol{S}_n(\boldsymbol{T} - \mathrm{E}\boldsymbol{T})^T\right) = \Phi I_n^{-1} \mathrm{E}(\boldsymbol{S}_n \boldsymbol{T}^T) = \Phi I_n^{-1} \Phi^T. \tag{5}$$

We vinden dus dat

$$\mathcal{V}(\boldsymbol{Y}) = \mathcal{V}(\boldsymbol{T}) - \Phi I_n^{-1} \Phi^T, \tag{6}$$

en omdat deze matrix een variantiematrix is, is hij positief semi-definiet. □

Gevolg 2 (Stelling zoals in de cursus)**.** *Als $m = d$, $\boldsymbol{g} = Id$ en $\boldsymbol{T}$ een onvertekende schatter is voor $\boldsymbol{\theta}$, dan is*

$$\mathcal{V}(\boldsymbol{T}) - I_n^{-1} \tag{7}$$

positief-definiet.

Bewijs. In dit geval is

$$\Phi_{ij} = \frac{\partial \mathrm{E} T_i}{\partial \theta_j} = \frac{\partial \theta_i}{\partial \theta_j} = \delta_{ij}, \tag{8}$$

zodat $\Phi = I$ (de eenheidsmatrix). $\square$

Gevolg 3 (Eén van de opmerkingen)**.** *Als $m = d = 1$ en $g = Id$, dan is*

$$Var_\theta T \geq \frac{(1 + b'(\theta))^2}{I_n(\theta)}. \tag{9}$$

Bewijs. In dit geval is

$$\Phi = \frac{\partial \mathrm{E} T}{\partial \theta} = \frac{\partial}{\partial \theta}(\theta + b(\theta)) = 1 + b'(\theta). \tag{10}$$

$\square$

Gevolg 4 (Een andere opmerking)**.** *Noteer met Δ de totale afgeleide van $\boldsymbol{g}$. Als $\boldsymbol{T}$ een onvertekende schatter is voor $\boldsymbol{g}(\boldsymbol{\theta})$, dan is*

$$\mathcal{V}(\boldsymbol{T}) - \Delta I_n^{-1} \Delta^T \tag{11}$$

positief-definiet.

Bewijs. $\Phi = \Delta$. $\square$

2 Theorem of Rao-Blackwell

I don't know how to translate "sufficient" so the rest is in English:

Stelling 5.

- *Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a sample of X with a distribution $f_{\boldsymbol{X}}(\boldsymbol{x}, \theta)$.*
- *Let $T(\boldsymbol{X})$ be a sufficient statistic for θ.*
- *Let $U(\boldsymbol{X})$ be an unbiased estimator for θ that does not depend on $\boldsymbol{X}$ solely through $T(\boldsymbol{X})$, i.e. U doesn't factorize as $U = \hat{U} \circ T$.*

Define $\varphi_\theta : \mathbb{R} \to \mathbb{R} : t \mapsto E_\theta(U|T = t)$. Then it holds that:

1. *The function φ_θ does not depend on θ, so we can just write $\varphi_\theta = \varphi$.*
2. *$\varphi(T)$ is an unbiased estimator for θ.*[1]
3. *$\forall \theta \in \Theta : Var_\theta U < \infty \Rightarrow Var_\theta \varphi(T) < Var_\theta U$.*

Bewijs.

1. This follows immediately from the definition of a sufficient statistic.
2. $\mathrm{E}_\theta \varphi(T) = \mathrm{E}_\theta(\mathrm{E}_\theta(U|T)) = \mathrm{E}_\theta U = \theta$.
3. We will prove that $\mathrm{Var}_\theta U - \mathrm{Var}_\theta \varphi(T) > 0$.

$$\begin{aligned}
\mathrm{Var}_\theta U - \mathrm{Var}_\theta \varphi(T) &= \mathrm{E}_\theta(U^2) - \theta^2 - \mathrm{E}_\theta(\varphi(T)^2) + \theta^2 \qquad \text{because } \mathrm{E}_\theta U = \mathrm{E}_\theta \varphi(T) = \theta \\
&= \mathrm{E}_\theta \left(U^2 - \varphi(T)^2\right) = \mathrm{E}_\theta \left(U^2 - \mathrm{E}_\theta(U|T)^2\right) \\
&= \mathrm{E}_\theta \left(\mathrm{E}_\theta \left[U^2 - \mathrm{E}_\theta(U|T)^2 \middle| T\right]\right) \qquad \text{because } \mathrm{E}A = \mathrm{E}(\mathrm{E}(A|B)) \\
&= \mathrm{E}_\theta \left(\mathrm{E}_\theta(U^2|T) - \mathrm{E}_\theta(U|T)^2\right) \\
&\qquad \text{because } \mathrm{E}_\theta(U|T) \text{ is constant if } T \text{ is known} \\
&= \mathrm{E}_\theta \left(\mathrm{Var}_\theta(U|T)\right) \geq 0.
\end{aligned}$$

Assume that $\mathrm{E}_\theta \left(\mathrm{Var}_\theta(U|T)\right) = 0$. Then $\mathrm{Var}_\theta(U|T) = 0$ almost surely, and hence $U = \mathrm{E}_\theta(U|T)$ almost surely if T is given. But then U only depends on $\boldsymbol{X}$ through T and this is in contradiction with the assumptions. Hence, $\mathrm{Var}_\theta \varphi(T) < \mathrm{Var}_\theta U$.

$\square$

[1] Begin NOOIT een zin met een symbool :)