Repetition week 43

Score statistic

$$S(\boldsymbol{X}|\theta) = \frac{\partial}{\partial\theta}\log f(\boldsymbol{X}|\theta)$$

$$E\left[S(\boldsymbol{X}|\theta)\right] = 0$$

$$Var\left[S(\boldsymbol{X}|\theta)\right] = I_{\boldsymbol{X}}(\theta) = -E\left[\frac{\partial}{\partial\theta}S(\boldsymbol{X}|\theta)\right] = -E\left[\frac{\partial^2}{\partial\theta^2}\log f(\boldsymbol{X}|\theta)\right]$$

Let $\tau(\theta) = E\left[W(\boldsymbol{X})\right]$

Cramer-Rao

$$Var\left[W(\boldsymbol{X})\right] \geq \frac{\left(\frac{\partial}{\partial\theta}\tau(\theta)\right)^2}{I_{\boldsymbol{X}}(\theta)}$$

Cramer-Rao iid

$$Var\left[W(\boldsymbol{X})\right] \geq \frac{\left(\frac{\partial}{\partial\theta}\tau(\theta)\right)^2}{nI_X(\theta)}$$

Equality

If and only if $S(\boldsymbol{X}|\theta) = a(\theta)\left[W(\boldsymbol{X}) - \tau(\theta)\right]$

Cramer-Rao in the multiparameter case

$\boldsymbol{\theta} = (\theta_1, \ldots \theta_k)^t$

Define the Score function $\boldsymbol{S}(\boldsymbol{X}|\boldsymbol{\theta}) = \begin{bmatrix} \frac{\partial}{\partial \theta_1} \log f(\boldsymbol{X}|\boldsymbol{\theta}) \\ \vdots \\ \frac{\partial}{\partial \theta_k} \log f(\boldsymbol{X}|\boldsymbol{\theta}) \end{bmatrix} = \nabla \log f(\boldsymbol{X}|\boldsymbol{\theta})$

Define the Fisher information $I(\boldsymbol{\theta}) = Cov\left[\boldsymbol{S}(\boldsymbol{X}|\boldsymbol{\theta})\right]$

We have as in the univariate case that $E\left[\boldsymbol{S}(\boldsymbol{X}|\boldsymbol{\theta})\right] = \mathbf{0}$ and

$I(\boldsymbol{\theta}) = E\left[\boldsymbol{S}(\boldsymbol{X}|\boldsymbol{\theta})\boldsymbol{S}(\boldsymbol{X}|\boldsymbol{\theta})^T\right] = -E\left[H(\boldsymbol{X}|\boldsymbol{\theta})\right]$ where

$h_{ij} = \frac{\partial}{\partial \theta_i}\frac{\partial}{\partial \theta_j} \log f(\boldsymbol{X}|\boldsymbol{\theta})$.

If $\boldsymbol{W}(\boldsymbol{X})$ is an unbiased estimator for $\boldsymbol{\theta}$. Then $\boldsymbol{I}(\boldsymbol{\theta})^{-1}$ is taken as an approximation to $Cov\left[\boldsymbol{W}(\boldsymbol{X})\right]$

Complete statistic

Definition 6.2.21

Let $f(t|\theta)$ be a family of pdfs/pmfs for a statistic $T(X)$. The family is complete if

$E_\theta\left[g(T)\right] = 0 \Rightarrow P_\theta\left(g(T) = 0\right) = 1$, for all θ.

Completeness and the exponential class

Let $X_1,\ldots,X_n$ be iid. from an exponential family i.e.

$$f(x|\boldsymbol{\theta}) = h(x)c(\boldsymbol{\theta})e^{\sum_{i=1}^{k} w(\theta_i)t_i(x)}$$

Then $T(\boldsymbol{X}) = \left(\sum_{i=1}^{n} t_1(X_i), \sum_{i=1}^{n} t_2(X_i), \ldots, \sum_{i=1}^{n} t_k(X_i),\right)$ is complete as long as the parameter space contains an open set in R^n.

Minimal sufficient if $w_i(\theta), i = 1,2,\ldots n$ are not linearly dependent

Complete if no functional relationship exists between $w_i(\theta), i = 1,2,\ldots n$

Then also the distribution of

$T(\boldsymbol{X}) = \left(\sum_{i=1}^{n} t_1(X_i), \sum_{i=1}^{n} t_2(X_i), \ldots, \sum_{i=1}^{n} t_k(X_i),\right)$ is within the exponential family.

Sufficiency and Unbiasedness

W unbiased estimator of $\tau(\theta)$.

T a sufficient statistic $E[W|T] = \tau(\theta)$ and $Var[W|T] \leq Var[W], \ \forall \theta$

T complete $\Rightarrow E[W|T]$ is the unique best unbiased estimator for