4. Cremer, Rao, and Fisher

GULP Winter 2015

February 4, 2015

1 Lower bound for the variance of unbiased estimators

The discussion is under the assumption of squared loss. Therefore we have the bias-variance decomposition discussed last time. If we have an unbiased estimator, then the loss is equivalent to the variance of the estimator. Here we propose a lower bound of the variance of the unbiased estimators, which is independent of the form of the estimators.

Suppose we have a random sample X and some unbiased estimator $\delta(X)$ for $g(\theta)$. Define a **score variable** $V = \frac{\partial}{\partial\theta} \ln f(X;\theta)$. We can derive several properties for this variable. Under some regular conditions (which enables us to switch the integration and the differential operator),

$$\mathbb{E}(V) = \int [\frac{\partial}{\partial\theta} \ln f(x;\theta)] f(x;\theta) dx = \int \frac{1}{f(x;\theta)} [\frac{\partial}{\partial\theta} f(x;\theta)] f(x;\theta) dx = \int \frac{\partial}{\partial\theta} f(x;\theta) dx = 0$$

$$\begin{aligned} \mathrm{Cov}\,(V,\delta) &= \mathbb{E}(V\delta(X)) = \int \delta(x) \frac{1}{f(x;\theta)} [\frac{\partial}{\partial\theta} f(x;\theta)] f(x;\theta) dx \\ &= \int \frac{\partial}{\partial\theta} \delta(x) f(x;\theta) dx = \frac{\partial}{\partial\theta} \int \delta(x) f(x;\theta) dx = g'(\theta) \end{aligned}$$

Using Cauchy-Schwartz inequality we have

$$\sqrt{\mathrm{Var}\,(V)\,\mathrm{Var}\,(\delta)} \geq \mathrm{Cov}\,(V,\delta)\,,$$

which is just

$$\mathrm{Var}\,(\delta) = \frac{\mathrm{Cov}\,(V,\delta)^2}{\mathrm{Var}\,(V)} = \frac{(g'(\theta))^2}{\mathbb{E}(V^2)}$$

We define the right hand side as the **Cremer-Rao Bound**, and $\mathbb{E}(V^2)$ as the **Fisher information**, denoted as $I(\theta)$. By definition,

$$I(\theta) = \mathbb{E}((\frac{\partial}{\partial\theta} \ln f(X,\theta))^2)$$

2 Some properties of the Fisher information

If you have n iid samples, your Fisher information will be n times of the information carried by a single observation. It is necessary to have iid setting here.

Another important equation which simplifies a lot of calculations is that

$$I(\theta) = -\mathbb{E}((\frac{\partial^2}{\partial\theta^2}\ln f(X,\theta)))$$

A sketch of proof is as follows. From the original definition we have

$$\begin{aligned} I(\theta) &= \mathbb{E}((\frac{\partial}{\partial\theta}\ln f(X,\theta))^2) \\ &= \int (\frac{1}{f(x;\theta)})^2[\frac{\partial}{\partial\theta}f(x;\theta)]^2 f(x;\theta)dx, \end{aligned}$$

while

$$\begin{aligned} -\mathbb{E}((\frac{\partial^2}{\partial\theta^2}\ln f(X,\theta))) &= -\int (\frac{\partial^2}{\partial\theta^2}\ln f(x,\theta))f(x,\theta)dx \\ &= -\int \frac{\partial}{\partial\theta}(\frac{1}{f(x;\theta)}\cdot\frac{\partial}{\partial\theta}f(x;\theta))f(x;\theta)dx \\ &= -\int \{(\frac{1}{f(x;\theta)})^2[\frac{\partial}{\partial\theta}f(x;\theta)]^2 + \frac{1}{f(x;\theta)}\frac{\partial^2}{\partial\theta^2}f(x;\theta)\}f(x;\theta)dx \\ &= \int (\frac{1}{f(x;\theta)})^2[\frac{\partial}{\partial\theta}f(x;\theta)]^2 f(x;\theta)dx. \end{aligned}$$

For the latter part, just move the second order differentiate operator out of the integration sign and you will get zero.