

Recap:

① Lehman-Scheffe theorem.

② Rao-Blackwell theorem.

There is a second way to find the UMVUE.

This comes from the famous Cramer-Rao inequality.

Let $\lambda(x) = \log f(x|\theta)$

Now, we call $u_\theta(x) = \frac{d}{d\theta} \lambda(x) = \frac{d}{d\theta} \log f(x|\theta)$ $\left[\frac{f(x|\theta)}{f_\theta(x)} \right]$

is called the score function.

$E_\theta[u_\theta(x)] = 0$. This can be seen from the fact that

$$\int f_\theta(x) dx = 1 \Rightarrow \frac{d}{d\theta} \int f_\theta(x) dx = 0$$

$$\Rightarrow \int \frac{d}{d\theta} f_\theta(x) dx = 0 \Rightarrow \int \frac{1}{f_\theta(x)} \left[\frac{d}{d\theta} f_\theta(x) \right] f_\theta(x) dx = 0$$

$$\Rightarrow \int \left[\frac{d}{d\theta} \log f_\theta(x) \right] f_\theta(x) dx = 0 \Rightarrow E_\theta[u_\theta(x)] = 0$$

We also define the Fisher information

$$I(\theta) = E[u_\theta(x)^2] = \text{Var}(u_\theta(x)). \text{ Note that}$$

it can be shown just by taking another derivative

$$\text{w.r.t. } \theta \text{ that } E_\theta[u_\theta(x)^2] = -E[u_\theta'(x)]$$

$$= -E \left[\frac{d}{d\theta} u_\theta(x) \right]$$

Information for location family

If $X \sim f(x-\theta)$, $f(x) > 0$ for all x , then

$$I(\theta) = \int_{-\infty}^{\infty} \frac{[f'(x)]^2}{f(x)} dx$$

Pf: Note that $u_{\theta}(x) = \frac{d}{d\theta} \log f(x-\theta) = -f'(x-\theta) \frac{1}{f(x-\theta)}$

$$\text{Thus } I(\theta) = \int_{-\infty}^{\infty} u_{\theta}(x)^2 f(x-\theta) dx = \int_{-\infty}^{\infty} \left[\frac{f'(x-\theta)}{f(x-\theta)} \right]^2 f(x-\theta) dx$$

$$z = x - \theta \text{ then } = \int_{-\infty}^{\infty} \left[\frac{f'(z)}{f(z)} \right]^2 f(z) dz$$

Information for the scale family

When $X \sim \frac{1}{\theta} f\left(\frac{x}{\theta}\right)$, then

$$I(\theta) = \frac{1}{\theta^2} \int \left[\frac{y f'(y)}{f(y)} + 1 \right]^2 f(y) dy$$

$$\begin{aligned} \text{Pf: } u_{\theta}(x) &= \frac{d}{d\theta} \log \left[\frac{1}{\theta} f\left(\frac{x}{\theta}\right) \right] \\ &= \left[-\frac{1}{\theta^2} f\left(\frac{x}{\theta}\right) + \frac{1}{\theta} \left(-\frac{x}{\theta^2}\right) f'\left(\frac{x}{\theta}\right) \right] \frac{1}{\theta} f\left(\frac{x}{\theta}\right) \\ &= - \left[\frac{1}{\theta^2} f\left(\frac{x}{\theta}\right) + \frac{x}{\theta^3} f'\left(\frac{x}{\theta}\right) \right] \frac{1}{\theta} f\left(\frac{x}{\theta}\right) \end{aligned}$$

$$\text{then } I(\theta) = \int_{-\infty}^{\infty} u_{\theta}(x)^2 \frac{1}{\theta} f\left(\frac{x}{\theta}\right) dx$$

$$= \int_{-\infty}^{\infty} \frac{\left[\frac{1}{\theta^{\nu}} f\left(\frac{x}{\theta}\right) + \frac{x}{\theta^{\nu+1}} f'\left(\frac{x}{\theta}\right) \right]^2}{\left[\frac{1}{\theta} f\left(\frac{x}{\theta}\right) \right]^{\nu}} \frac{1}{\theta} f\left(\frac{x}{\theta}\right) dx$$

$$y = \frac{x}{\theta} \Rightarrow dy = \frac{dx}{\theta}$$

$$= \int_{-\infty}^{\infty} \frac{\left[\frac{1}{\theta^{\nu}} f(y) + \frac{1}{\theta^{\nu}} y f'(y) \right]^2}{\left[\frac{1}{\theta} f(y) \right]^{\nu}} f(y) dy$$

$$= \frac{1}{\theta^{\nu}} \int_{-\infty}^{\infty} \frac{\left[f(y) + y f'(y) \right]^2}{\left[f(y) \right]^{\nu}} f(y) dy$$

$$= \frac{1}{\theta^{\nu}} \int_{-\infty}^{\infty} \left[1 + y \frac{f'(y)}{f(y)} \right]^2 f(y) dy$$

Information Inequality

Suppose $X \sim f_{\theta}(x)$ and $I(\theta) > 0$. Let $S(x)$ be any function of x with $E_{\theta}[S(x)^{\nu}] < \infty$, for which the derivative w.r.t. θ of $E_{\theta}[S(x)]$ exists and can be differentiated under the integral sign, i.e.

$$\frac{d}{d\theta} E_{\theta}[S(x)] = \int S(x) \frac{d}{d\theta} f_{\theta}(x) dx = \int S(x) u_{\theta}(x) f_{\theta}(x) dx.$$

Then

$$\text{Var}_{\theta}(S(x)) \geq \frac{\left[\frac{d}{d\theta} E_{\theta}[S(x)] \right]^{\nu}}{I(\theta)}$$

(3)

$$\text{Prf: } \text{cov}(S(x), u_\theta(x))^2 \leq \text{var}_\theta(S(x)) \text{var}_\theta(u_\theta(x)) \quad \dots (*)$$

$$\begin{aligned} \text{cov}(S(x), u_\theta(x)) &= \int \cancel{s(x) u_\theta(x) f_\theta(x) dx} \\ &= E[S(x) u_\theta(x)] - E[S(x)] E[u_\theta(x)] \\ &= E[S(x) u_\theta(x)] \quad \left[\text{as } E[u_\theta(x)] = 0 \right] \\ &= \int s(x) u_\theta(x) f_\theta(x) dx = \frac{d}{d\theta} E_\theta[S(x)] \end{aligned}$$

$$\text{var}(u_\theta(x)) = I(\theta)$$

$$\text{by } (*) \quad \left\{ \frac{d}{d\theta} E_\theta[S(x)] \right\}^2 \leq \text{var}_\theta(S(x)) I(\theta)$$

$$\Rightarrow \text{var}_\theta(S(x)) \geq \frac{\left\{ \frac{d}{d\theta} E_\theta[S(x)] \right\}^2}{I(\theta)}$$

Recap:

Information inequality

Suppose $X \sim f_\theta(x)$ and $I(\theta) > 0$. Let $s(x)$ be any function of X with $E_\theta[s(x)^2] < \infty$, for which the derivative w.r.t. θ of $E_\theta[s(x)]$ exists and can be differentiated under the integral sign i.e.

$$\frac{d}{d\theta} E_\theta[s(x)] = \int s(x) u_\theta(x) f_\theta(x) dx \quad \text{then}$$

$$\text{Var}_\theta(s(x)) \geq \frac{\left[\frac{d}{d\theta} E_\theta[s(x)] \right]^2}{I(\theta)}$$

Suppose $X_1, \dots, X_n \stackrel{iid}{\sim} f_\theta(x)$. The score function of a random sample $u_\theta(\underline{x}) = \frac{d}{d\theta} \log \left[\prod_{i=1}^n f_\theta(x_i) \right]$

$$= \sum_{i=1}^n \frac{d}{d\theta} \log f_\theta(x_i) = \sum_{i=1}^n u_\theta(x_i)$$

Fisher information contained in X_1, \dots, X_n , denoted by $I_n(\theta) = \text{Var}(u_\theta(\underline{x})) = \text{Var}\left(\sum_{i=1}^n u_\theta(x_i)\right)$

$$= \sum_{i=1}^n \text{Var}(u_\theta(x_i)) = n I(\theta)$$

Cramer-Rao Inequality:

Let X_1, \dots, X_n be iid from a distribution with pdf or pmf $f(x|\theta)$. Let $T(\underline{x})$ be any unbiased estimator of $m(\theta)$ i.e. $E[T(\underline{x})] = m(\theta)$. Assume that all regularity conditions hold then,

$\text{Var}(T(\underline{x})) \geq \frac{[m'(\theta)]^2}{n I(\theta)}$. When equality holds,

$T(\underline{x})$ must be of the form $T(\underline{x}) = \frac{m'(\theta)}{n I(\theta)} \sum_{i=1}^n u_\theta(x_i) + m(\theta)$.

Algo. Finding UMVUE:

- 1) a) Start with any unbiased estimator of $m(\theta)$.
Let that be $H(\underline{x})$.
- b) Find a complete sufficient statistic $T(\underline{x})$ for the family of distributions.
- c) Compute $E_0[H(\underline{x}) | T(\underline{x})]$

2) a) Given any problem compute of estimating $m(\theta)$

$$\frac{m'(\theta)}{n I(\theta)} \sum_{i=1}^n u_{\theta}(x_i) + m(\theta)$$

- b) check if it is free of the parameter θ .
- c) In that case, it is the UMVUE by Cramer-Rao Inequality.

Cramer-Rao lower bound is not useful when the range of the density depends on the parameter.

Let $x_1, \dots, x_n \stackrel{iid}{\sim} U(0, \theta)$.

$$\text{Then } f_{\theta}(x) = \frac{1}{\theta} \Rightarrow \frac{d}{d\theta} \log f_{\theta}(x) = -\frac{1}{\theta} = u_{\theta}(x)$$

$$\text{and } I(\theta) = \text{var}(u_{\theta}(x)) = E_0[u_{\theta}(x)^2] = \frac{1}{\theta^2}$$

Cramer-Rao lower bound for any unbiased estimator of θ is

$$m(\theta) = \theta, \quad \text{var}(T(\underline{x})) \geq \frac{[m'(\theta)]^2}{n I(\theta)} = \frac{1}{\frac{n}{\theta^2}} = \frac{\theta^2}{n}$$

$$\text{Let } T(\underline{x}) = \left(\frac{n+1}{n}\right) X_{(n)}$$

$$f_{X_{(n)}}(x) = \frac{n x^{n-1}}{\theta^n}, \quad 0 < x < \theta$$

$$E\left[\left(\frac{n+1}{n}\right) X_{(n)}\right] = \int_0^\theta \left(\frac{n+1}{n}\right) \frac{n x^n}{\theta^n} dx = \theta$$

$$\text{Var}\left(\left(\frac{n+1}{n}\right) X_{(n)}\right) = \frac{(n+1)^2}{n^2} \left[E[X_{(n)}^2] - E[X_{(n)}]^2 \right]$$

$$= \left(\frac{n+1}{n}\right)^2 \left[\frac{\theta^2}{(n+2)} - \left(\frac{\theta}{n+1}\right)^2 \right] = \frac{\theta^2}{n(n+2)}$$

Thus $\left(\frac{n+1}{n}\right) X_{(n)}$ is an unbiased estimator of θ which has variance less than the Cramer-Rao lower bound.

Example: X_1, \dots, X_n iid Poiss(λ). UMVUE of λ .

$$m(\lambda) = \lambda, \quad u_\lambda(x) = \frac{d}{d\lambda} \log f_\lambda(x) = \frac{d}{d\lambda} \log \left[\frac{e^{-\lambda} \lambda^x}{x!} \right]$$

$$= \frac{d}{d\lambda} \left[-\lambda + x \log \lambda - \log x! \right] = -1 + \frac{x}{\lambda}$$

$$I(\lambda) = E[u_\lambda(x)^2] = E\left[\left(\frac{x}{\lambda} - 1\right)^2\right] = \frac{1}{\lambda^2} E[(x - \lambda)^2] = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

$$T(\underline{x}) = \frac{m'(\lambda)}{n I(\lambda)} \sum_{i=1}^n u_\lambda(x_i) + u(\lambda) = \frac{1}{\lambda} \sum_{i=1}^n \left(\frac{x_i}{\lambda} - 1\right) + \lambda$$

$$= \frac{1}{n} \sum_{i=1}^n x_i - \frac{\lambda}{n} \cdot n + \lambda = \bar{x}$$

proof of the Cramer-Rao Inequality

Use Cauchy Schwartz inequality to obtain

$$\text{Cov}(T(\underline{x}), u_{\theta}(\underline{x}))^2 \leq \text{Var}(T(\underline{x})) \text{Var}(u_{\theta}(\underline{x}))$$

$\text{Cov}(T(\underline{x}), u_{\theta}(\underline{x})) = m'(\theta)$ [by ~~info~~ the proof of the information inequality]

$$\text{Var}(u_{\theta}(\underline{x})) = n I(\theta)$$

$$\Rightarrow m'(\theta)^2 \leq \text{Var}(T(\underline{x})) n I(\theta)$$

$$\Rightarrow \text{Var}(T(\underline{x})) \geq \frac{[m'(\theta)]^2}{n I(\theta)}$$

When equality holds in the Cauchy-Schwarz inequality, then $T(\underline{x})$ must be of the form

$$T(\underline{x}) = a(\theta) u_{\theta}(\underline{x}) + b(\theta) \quad [\text{for some } a(\theta) \text{ and } b(\theta)]$$

$$E[T(\underline{x})] = a(\theta) E[u_{\theta}(\underline{x})] + b(\theta) = b(\theta)$$

We started with the assumption that $E[T(\underline{x})] = m(\theta)$

$$\Rightarrow m(\theta) = b(\theta)$$

Now, $\text{Cov}(T(\underline{x}), u_{\theta}(\underline{x})) = m'(\theta)$ [we know already] (*)

$$\text{Again } T(\underline{x}) = a(\theta) u_{\theta}(\underline{x}) + m(\theta)$$

$$\text{thus } \text{Cov}(T(\underline{x}), u_{\theta}(\underline{x})) = \text{Cov}(a(\theta) u_{\theta}(\underline{x}) + m(\theta), u_{\theta}(\underline{x}))$$

$$= a(\theta) \text{Var}(u_{\theta}(\underline{x})) = a(\theta) n I(\theta) \quad \text{--- (**)}$$

$$\Rightarrow m'(\theta) = a(\theta) n I(\theta) \Rightarrow a(\theta) = \frac{m'(\theta)}{n I(\theta)}$$

Multiparameter Cramer-Rao lower bound

When $X \sim f_{\theta}(x)$ where $\underline{\theta} = (\theta_1, \dots, \theta_k)'$. Define a score vector instead of a scalar score. The score vector is defined as

$$\underline{u}_{\theta}(x) = \left(\frac{\partial}{\partial \theta_1} \log f_{\theta}(x), \frac{\partial}{\partial \theta_2} \log f_{\theta}(x), \dots, \frac{\partial}{\partial \theta_k} \log f_{\theta}(x) \right)'$$

Define the Fisher information matrix

$$I(\underline{\theta}) = \left((I_{ij}(\underline{\theta})) \right)_{i,j=1}^k \quad (\text{a } k \times k \text{ matrix})$$

$$\text{where } I_{ij}(\underline{\theta}) = E \left[\left\{ \frac{\partial}{\partial \theta_i} \log f_{\theta}(x) \right\} \left\{ \frac{\partial}{\partial \theta_j} \log f_{\theta}(x) \right\} \right]$$

Multiparameter information Inequality

Suppose that $I(\underline{\theta})$ is positive definite and

$\alpha_i = \frac{\partial}{\partial \theta_i} E_{\underline{\theta}}[S(x)]$ exists and differentiation w.r.t.

θ_i can be done under integration w.r.t. x .

Then $\text{Var}_{\underline{\theta}}(S(x)) \geq \underline{\alpha}' I^{-1}(\underline{\theta}) \underline{\alpha}$ where

$$\underline{\alpha} = (\alpha_1, \dots, \alpha_k)'$$

~~Example 2~~

Example: $X \sim N(\mu, \sigma^2)$, μ, σ^2 both unknown

$X \sim \text{Gamma}(\alpha, \beta)$ where α, β both unknown.