Lehmann-Scheffe

Lehmann-Scheffé theorem

Recall: Let T be complete, and suppose $\varphi(T)$ and $\psi(T)$ are two unbiased estimators of θ. Then
$\mathbb{E}_\theta[\varphi(T)] = \mathbb{E}_\theta[\psi(T)] = \theta \quad \forall\theta$, so
$\mathbb{E}_\theta[\varphi(T) - \psi(T)] = 0 \quad \forall\theta$. But T is complete, so
$\varphi(T) = \psi(T)$, and so $\varphi(T)$ is unique (there is only one unbiased estimator of θ which is a function of T).

Theorem: (Lehmann-Scheffé) Suppose $T \equiv T(X_1, \dots, X_n)$ is a complete, sufficient statistic, and $\varphi(T)$ is an unbiased estimator of θ. Then $\varphi(T)$ is a best unbiased estimator of θ.

Pf: Let W be some other unbiased estimator of θ
WTS $Var(\varphi(T)) \leq Var(W)$
Let $\psi(T) = \mathbb{E}[W \mid T]$
T is sufficient, so by Rao-Blackwell

$$Var(\psi(T)) \leq Var(W)$$

And $\psi(T)$ is also an unbiased estimator.
But T is complete, $\psi(T) = \varphi(T)$
$\Rightarrow Var(\varphi(T)) \leq Var(W)$. So $\varphi(T)$ is the best unbiased estimator. //

Example

Suppose $X_1, \ldots, X_n \overset{iid}{\sim} Uniform[0, \theta]$.

Recall: $X_{(n)}$ is a biased estimator of θ $\left(\mathbb{E}[X_{(n)}] = \frac{n\theta}{n+1} \right)$

$\left(\frac{n+1}{n}\right) X_{(n)}$ is an unbiased estimator

Is $\left(\frac{n+1}{n}\right) X_{(n)}$ the best unbiased estimator of θ?

- Can't use CRLB (regularity conditions don't hold)
- Lehmann-Scheffé: $\left(\frac{n+1}{n}\right) X_{(n)}$ is the best unbiased estimator if $X_{(n)}$ is sufficient & complete

Factorization $\Rightarrow$ $X_{(n)}$ is sufficient

complete: WTS if $\mathbb{E}_\theta[g(X_{(n)})] = 0 \quad \forall \theta$, then

$P_\theta(g(X_{(n)}) = 0) = 1 \quad \forall \theta$

Suppose $\mathbb{E}_\theta[g(X_{(n)})] = 0$ $\qquad f_{X_{(n)}}(t|\theta) = \dfrac{nt^{n-1}}{\theta^n}$ $t \in (0,\theta)$

$$\mathbb{E}_\theta[g(X_{(n)})] = \int_0^\theta g(t)\frac{nt^{n-1}}{\theta^n}\,dt = \theta^{-n}\int_0^\theta g(t)\,nt^{n-1}\,dt$$

if $\mathbb{E}_\theta[g(X_{(n)})] = 0$ then $\dfrac{\partial}{\partial\theta}\mathbb{E}_\theta[g(X_{(n)})] = 0$

$$\frac{\partial}{\partial\theta}\mathbb{E}_\theta[g(X_{(n)})] = \frac{\partial}{\partial\theta}\left[\theta^{-n}\int_0^\theta g(t)\,nt^{n-1}\,dt\right]$$

$$= \theta^{-n}\left[\frac{\partial}{\partial\theta}\int_0^\theta g(t)\,nt^{n-1}\,dt\right] + \left[\frac{\partial}{\partial\theta}\theta^{-n}\right]\underbrace{\int_0^\theta g(t)\,nt^{n-1}\,dt}_{0 \text{ if } \mathbb{E}_\theta[g(X_{(n)})]=0}$$

$$= \theta^{-n}\left[\frac{\partial}{\partial\theta}\int_0^\theta g(t)\,nt^{n-1}\,dt\right]$$

$$= \theta^{-n}\left[n\,g(\theta)\,\theta^{n-1}\right] \qquad \text{(FTC)}$$

$$= \theta^{-1}\,n\,g(\theta) = 0$$ only when $g(\theta) = 0$

we assumed that $\mathbb{E}_\theta[g(X_{(n)})] = 0 \;\forall\theta$

$\Rightarrow g(\theta) = 0 \;\forall\theta$

$\Rightarrow g(X_{(n)}) = 0$

$\Rightarrow X_{(n)}$ is complete!

Examples: normal, gamma, beta, binomial, Poisson, negative binomial

Exponential families

A distribution with probability function $f(x|\theta)$ is an <u>exponential family</u> distribution if we can write

$$f(x|\theta) = h(x)\, c(\theta) \exp\left\{ \sum_{i=1}^{k} w_i(\theta)\, t_i(x) \right\}$$

where $h(x) \geq 0$, $c(\theta) \geq 0$, and $w_i(\theta)$, $t_i(x) \in \mathbb{R}$

$h(x)$, $t_i(x)$ only depend on x (no θ)

$c(\theta)$, $w_i(\theta)$ only depend on θ (no x)

<u>Ex:</u> Poisson (λ) $\quad f(x|\lambda) = \dfrac{e^{-\lambda} \lambda^{x}}{x!}$

$$= \frac{1}{x!} e^{-\lambda} \exp\{ x \log(\lambda) \}$$

$$= h(x)\, c(\lambda) \exp\{ w_1(\lambda)\, t_1(\lambda) \}$$

$h(x) = \frac{1}{x!}$, $c(\lambda) = e^{-\lambda}$, $w_1(\lambda) = \log(\lambda)$, $t_1(x) = x$

Exponential families and completeness

Suppose $X_1, \ldots, X_n$ are iid from an exponential family with

$$f(x|\theta) = h(x) \underbrace{c(\theta) \exp\left\{ \sum_{i=1}^{k} w_i(\theta) t_i(x) \right\}}_{g(x|\theta)}$$

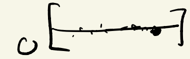
Sufficiency: By factorization theorem, $T(X_1, \ldots, X_n) = \left(\sum_{j=1}^{n} t_1(X_j), \ldots, \sum_{j=1}^{n} t_k(X_j) \right)$ is sufficient for θ

Completeness: If, in addition, $\{ (w_1(\theta), \ldots, w_k(\theta)) \}$ contains an open set in $\mathbb{R}^k$, then

$$T(X_1, \ldots, X_n) = \left(\sum_{j=1}^{n} t_1(X_j), \ldots, \sum_{j=1}^{n} t_k(X_j) \right)$$

is complete

Ex: $X_1, \dots, X_n \overset{iid}{\sim} \text{Poisson}(\lambda)$



$\frac{1}{n} \to 0$

$(0,1)$

$\uparrow$ exponential family with $w_1(\lambda) = \log(\lambda)$

$t_1(x) = x$

$$T(X_1, \dots, X_n) = \sum_{i=1}^{n} t_1(X_i) = \sum_{i=1}^{n} X_i$$

① $T = \sum_i X_i$ is sufficient ✓

② $\{\log(\lambda) : \lambda > 0\}$ contains an open set in $\mathbb{R}$

so $T = \sum_i X_i$ is complete ✓

By Lehmann-Scheffé, $T = \sum_i X_i$ is complete & sufficient,

and $\hat{\lambda} = \frac{1}{n} \sum_i X_i$ is unbiased for λ.

Thus $\hat{\lambda}$ is the best unbiased estimator of λ

Example

Suppose that $X_1, \ldots, X_n \overset{iid}{\sim} Bernoulli(p)$.

- Show the the Bernoulli is an exponential family distribution
- Find a complete, sufficient statistic
- Use the complete, sufficient statistic to find a best unbiased estimator of p