MATH 503: Mathematical Statistics

Lecture 5: More on Point Estimation

Reading: C&B Sec. 6.2, and HMC Sec. 7.7

Kimberly F. Sellers

Department of Mathematics & Statistics

Today's Topics

- Final comments connecting Rao-Blackwell and Lehmann-Scheffé
- Joint sufficiency
- Minimal sufficiency
- Ancillary statistics
- Sufficiency, Completeness & Independence

Rao-Blackwell Theorem

Let $X_1, \dots, X_n$, n a fixed positive integer, denote a random sample from a distribution that has pdf/pmf $f(x;\theta)$, $\theta \in \Omega$. Let $Y_1 = u_1(X_1, \dots, X_n)$ be a sufficient statistic for θ, and let $Y_2 = u_2(X_1, \dots, X_n)$, not a function of Y_1 alone, be an unbiased estimator of θ. Then $E(Y_2|y_1) = \varphi(y_1)$ defines a statistic $\varphi(Y_1)$. This statistic $\varphi(Y_1)$ is a function of the sufficient statistic for θ; it is an unbiased estimator of θ; and its variance is less than that of Y_2.

Lehmann-Scheffé Theorem

Let $X_1, \dots, X_n$, n a fixed positive integer, denote a random sample from a distribution that has pdf/pmf $f(x;\theta), \theta \in \Omega$, let $Y_1 = u_1(X_1, \dots, X_n)$ be a sufficient statistic for θ, and let the family $\{f_{Y_1}(y_1;\theta): \theta \in \Omega\}$ be complete. If there is a function of Y_1 that is an unbiased estimator of θ, then this function of Y_1 is the unique UMVUE of θ.

Theorem

Let $f(x;\theta), \gamma < \theta < \delta$, be a pdf/pmf of a rv X whose distribution is a regular case of the exponential class. Then if $X_1, X_2, \dots, X_n$ (where n is a fixed positive integer) is a random sample from the distribution of X, the statistic $Y_1 = \sum_{i=1}^n K(X_i)$ is a sufficient statistic for θ and the family $\{f_{Y_1}(y_1;\theta): \gamma < \theta < \delta\}$ of pdfs of Y_1 is complete. That is, Y_1 is a complete sufficient statistic for θ.

Implication: After determining the sufficient statistic, $Y_1 = \sum_{i=1}^n K(X_i)$, we form a function, $\varphi(Y_1)$, so that $\mathrm{E}(\varphi(Y_1)) = \theta \quad \Rightarrow \varphi(Y_1)$ is unique MVUE of θ.

Example

Let $X_1, \dots, X_n \sim$ Bernoulli(θ) iid, $0 < \theta < 1$. Find the UMVUE of θ.

$$f(x) = \theta^x (1-\theta)^{1-x} = \left(\frac{\theta}{1-\theta}\right)^x (1-\theta)$$

$$= \exp\left[x \ln\left(\frac{\theta}{1-\theta}\right) + \ln(1-\theta)\right]$$

$$= \exp\Big[\underbrace{x}_{K(x)}\underbrace{(\ln\theta - \ln(1-\theta))}_{p(\theta)} + \underbrace{\ln(1-\theta)}_{q(\theta)} + \underbrace{0}_{S(x)}\Big]$$

$\therefore Y = \sum_{i=1}^n K(x_i) = \sum x_i$ is complete sufficient for θ since pdf has exponential family form.

$\mathbb{E}(Y) = \sum_{i=1}^n \mathbb{E}(X_i) = \sum_{i=1}^n \theta = n\theta \Rightarrow \mathbb{E}(Y/n) = \mathbb{E}\left(\frac{\sum X_i}{n}\right) = \frac{n\theta}{n} = \theta$

$\therefore \frac{Y}{n} = \overline{X}$ is unique UMVUE of θ by Lehmann-Scheffe Thm.

Example

Let a random sample of size n, i.e. $X_1, \dots, X_n$, be taken from a distribution that has the pdf $f(x;\theta) = \frac{1}{\theta}\exp\left(-\frac{x}{\theta}\right) I_{(0,\infty)}(x)$. Find the MLE and the UMVUE of $P(X_1 \leq 2)$.

To find MLE of $\mathbb{P}(X \leq 2) = \int_0^2 \frac{1}{\theta} e^{-x/\theta} dx = \left(-e^{-x/\theta}\right)\Big|_0^2 = 1 - e^{-2/\theta}$, find MLE of θ and apply invariance property:

$\mathcal{L}(\theta; \underline{x}) = \prod_{i=1}^n f(x_i;\theta) = \frac{1}{\theta^n} \exp\left[-\frac{1}{\theta} \sum_1^n x_i\right]$

$\ln \mathcal{L}(\theta; \underline{x}) = -n \ln\theta - \frac{1}{\theta} \sum_1^n x_i$

$\frac{\partial \ln \mathcal{L}(\theta; \underline{x})}{\partial\theta} = \frac{-n}{\theta} + \frac{\sum x_i}{\theta^2} = 0 \Rightarrow \frac{\sum x_i}{\theta^2} = \frac{n}{\theta} \Rightarrow \hat{\theta} = \frac{\sum x_i}{n} = \bar{x}$

$\therefore Y_1 = 1 - e^{-2/\bar{x}}$ is MLE of $\mathbb{P}(X_1 \leq 2) = 1 - e^{-2/\theta}$ by invariance property.

SEE ATTACHED for UMVUE solution.

Joint Sufficiency

Let $X_1, \dots, X_n$ denote a random sample from a distribution that has pdf/pmf $f(x;\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Omega \subset R^p$. Let S denote the support of X. Let $\boldsymbol{Y}$ be an m-dimensional random vector of statistics, $\boldsymbol{Y} = (Y_1, \dots, Y_m)'$, where $Y_i = u_i(X_1, \dots, X_n)$, for $i = 1, \dots, m$. Denote the pdf/pmf of $\boldsymbol{Y}$ by $f_Y(\boldsymbol{y};\boldsymbol{\theta})$ for $\boldsymbol{y} \in R^m$. The random vector of statistics $\boldsymbol{Y}$ is <u>jointly sufficient</u> for $\boldsymbol{\theta}$ iff.

$$\frac{\prod_{i=1}^n f(x_i;\boldsymbol{\theta})}{f_Y(\boldsymbol{y};\boldsymbol{\theta})} = H(x_1, \dots, x_n) \ \forall x_i \in S$$

where $H(x_1, \dots, x_n)$ does not depend on $\boldsymbol{\theta}$.

To find the UMVUE,

$$f(x) = \frac{1}{\theta} e^{-x/\theta} = \exp\Big[\underbrace{-\ln\theta}_{q(\theta)} + \underbrace{\left(\tfrac{-1}{\theta}\right)}_{p(\theta)} \underbrace{x}_{K(x)} + \underbrace{0}_{S(x)}\Big]$$ is an exponential family

$\Rightarrow Y = \sum_{i=1}^{n} K(x_i) = \sum_{i=1}^{n} X_i$ is complete sufficient for θ.

Let $Z = \begin{cases} 1 & X_1 \le 2 \\ 0 & \text{ow} \end{cases}$. $E(Z) = \sum_{z=0}^{1} z\,p(z) = 1\cdot P(X_1 \le 2) + \cancel{0\cdot P(X_1 > 2)} = P(X_1 \le 2)$

$\therefore Z$ is unbiased for $P(X_1 \le 2)$

By Rao-Blackwellization, then $E(Z|Y)$ is unique UMVUE of $P(X_1 \le 2)$ (because Y complete sufficient and applying Lehmann-Scheffe Thm.), where

$$E(Z|Y=y) = \sum_{z=0}^{1} z\cdot p(z|y) = 1\cdot P(X_1 \le 2|Y) + \cancel{0\cdot P(X_1 > 2|Y)} = P(X_1 \le 2|Y=y)$$

where $P(X_1 \le 2 | Y=y) = \int_0^2 f(x_1|y)\,dx$ where

$$f(x_1|y) = \frac{f_{X_1,Y}(x_1,y)}{f_Y(y)} = \frac{f_{X_1}(x_1)\cdot f_{\sum_2^n X_i}(y-x_1)}{f_{\sum_1^n X_i}(y)}$$

where $X_1 \sim Exp(\theta)$

$\sum_1^n X_i \sim Gamma(n,\theta)$

$\sum_2^n X_i \sim Gamma(n-1,\theta)$

$$= \frac{\left(\frac{1}{\cancel{\theta}}\cancel{e^{-x/\theta}}\right)\left(\frac{1}{\Gamma(n-1)\theta^{n-1}}(y-x_1)^{(n-1)-1}\cancel{e^{-\frac{(y-x_1)}{\theta}}}\right)}{\frac{1}{\Gamma(n)\cancel{\theta^n}} y^{n-1}\cancel{e^{-y/\theta}}} = \frac{\Gamma(n)}{\Gamma(n-1)}\frac{(y-x_1)^{n-2}}{y^{n-1}}$$

$$= \frac{(n-1)\cancel{\Gamma(n-1)}}{\cancel{\Gamma(n-1)}}\frac{(y-x_1)^{n-2}}{y^{n-1}} = \frac{(n-1)(y-x_1)^{n-2}}{y^{n-1}}, \quad 0 < x_1 < y$$

$$\therefore P(X_1 \le 2|Y) = \int_0^2 \frac{(n-1)(y-x_1)^{n-2}}{y^{n-1}}\,dx_1 = \frac{-1}{y^{n-1}}(y-x_1)^{n-1}\Big|_0^2 = \frac{-1}{y^{n-1}}\left[(y-2)^{n-1} - y^{n-1}\right]$$

$= 1 - \left(\frac{y-2}{y}\right)^{n-1}$ where $y = \sum_1^n x_i$ is unique UMVUE of $P(X_1 \le 2)$ by Lehmann-Scheffe

The (Generalized) Factorization Thm

The vector of statistics $\boldsymbol{Y}$ is jointly sufficient for the parameter $\boldsymbol{\theta} \in \Omega$ iff we can find two nonnegative functions k_1 and k_2 s.t.

$$\prod_{i=1}^{n} f(x_i; \boldsymbol{\theta}) = k_1(\boldsymbol{y}; \boldsymbol{\theta}) k_2(x_1, \dots, x_n), \text{ for all } x_i \in S$$

where the function $k_2(x_1, \dots, x_n)$ does not depend on $\boldsymbol{\theta}$.

Example

Let $X_1, \dots, X_n$ be a random sample from a distribution having pdf

$$f(x; \theta_1, \theta_2) = \begin{cases} \frac{1}{2\theta_2} & \theta_1 - \theta_2 < x < \theta_1 + \theta_2 \\ 0 & \text{elsewhere,} \end{cases}$$

Find the joint sufficient statistics for θ_1 and θ_2.

$$\prod_{i=1}^{n} f(x_i; \theta_1, \theta_2) = \underbrace{\left(\frac{1}{2\theta_2}\right)^n I_{(\theta_1 - \theta_2, \theta_1 + \theta_2)}(X_{(1)}, X_{(n)})}_{k_1(X_{(1)}, X_{(n)}; \underset{\sim}{\theta})} \cdot \underbrace{1}_{k_2(\underset{\sim}{x})}$$

$\therefore$ by generalized NFFT, $(X_{(1)}, X_{(n)})$ are the joint sufficient statistics for (θ_1, θ_2).

(Extension of) Exponential Families

Let X be a rv with pdf/pmf $f(x; \boldsymbol{\theta})$ where the vector of parameters $\boldsymbol{\theta} \in \Omega \subset R^m$. Let S denote the support of X. If X is continuous assume that $S = (a, b)$, where a or b may be $-\infty$ or ∞, respectively. If X is discrete assume that $S = \{a_1, a_2, \ldots\}$. Suppose $f(x; \boldsymbol{\theta})$ is of the form

$$f(x;\boldsymbol{\theta}) = \begin{cases} \exp\left(\sum_{j=1}^{m} p_j(\boldsymbol{\theta})K_j(x) + S(x) + q(\theta_1, \theta_2, \ldots, \theta_m)\right) & \text{for all } x \in S \\ 0 & \text{elsewhere} \end{cases}$$

Then we say this pdf/pmf is a <u>member of the exponential class</u>.

(Ext. of) Exponential Families (cont.)

It is a <u>regular case of the exponential family</u> if, addition,

1) The support does not depend on the vector of parameters $\boldsymbol{\theta}$
2) The space Ω contains a nonempty, m-dimensional open rectangle,
3) The $p_j(\boldsymbol{\theta})$, $j = 1, \ldots, m$, are nontrivial, functionally independent, continuous functions of $\boldsymbol{\theta}$,
4) and
   (a) If X is a continuous r.v., then the m derivatives $K_j'(x)$, for $j = 1, \ldots, m$, are continuous for $a < x < b$ and no one is a linear homogeneous function of the others and $S(x)$ is a continuous function of x, $a < x < b$.
   (b) If X is discrete, the $K_j(x)$, $j = 1, \ldots, m$ are nontrivial functions of x on the support S and no one is a linear homogeneous function of the others.

Further Extensions

- Rao-Blackwell
- Lehmann-Scheffe
- Joint complete sufficient statistics for $\boldsymbol{\theta}$

Minimal Sufficiency

- **Goal**: reduce data contained in entire sample as much as possible without losing relevant information about important characteristics of underlying distribution

- **Definition**: a sufficient statistic, $T(\boldsymbol{X}) = T(X_1, \ldots, X_n)$, is called a minimal sufficient statistic if, for any other sufficient statistic $T'(\boldsymbol{X})$, $T(\boldsymbol{x})$ is a function of $T'(\boldsymbol{x})$ [i.e. if $T'(\boldsymbol{x}) = T'(\boldsymbol{y})$, then $T(\boldsymbol{x}) = T(\boldsymbol{y})$].

Theorem

Let $f(\boldsymbol{x}|\theta)$ be the pmf/pdf of a sample $X_1, \ldots, X_n$. Suppose there exists a function $T(\boldsymbol{x})$ s.t., for two sample points $\boldsymbol{x}$ and $\boldsymbol{y}$, the ratio

$$\frac{f(\boldsymbol{x}|\theta)}{f(\boldsymbol{y}|\theta)}$$

is constant as a function of θ iff $T(\boldsymbol{x})=T(\boldsymbol{y})$. Then $T(\boldsymbol{X})$ is a minimal sufficient statistic for θ.

Example

Let $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$ iid, both μ and σ^2 unknown. Let x and y denote two sample points and let $(\bar{x}, s_x^2)$ and $(\bar{y}, s_y^2)$ be the sample means and variances corresponding to the x and y samples, respectively. Then, the ratio of densities is

$$\frac{f(\boldsymbol{x} \mid \mu, \sigma^2)}{f(\boldsymbol{y} \mid \mu, \sigma^2)} = \frac{(2\pi\sigma^2)^{-n/2}\exp(-[n(\bar{x}-\mu)^2 + (n-1)s_x^2]/(2\sigma^2))}{(2\pi\sigma^2)^{-n/2}\exp(-[n(\bar{y}-\mu)^2 + (n-1)s_y^2]/(2\sigma^2))}$$

$$= \exp([-n(\bar{x}^2 - \bar{y}^2) + 2n\mu(\bar{x}-\bar{y}) - (n-1)(s_x^2 - s_y^2)]/(2\sigma^2))$$

This ratio will be constant as a function of μ and σ^2 iff $\bar{x} = \bar{y}$ and $s_x^2 = s_y^2$. Thus, $(\bar{X}, S^2)$ is a minimal sufficient statistic for (μ, σ^2).

SEE ATTACHED for details

$X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ where both μ, σ^2 unknown

$$f(x_i \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-1}{2\sigma^2}(x_i-\mu)^2} = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[\tfrac{-1}{2\sigma^2}(x_i-\mu)^2\right]$$

$$\therefore f(\underset{\sim}{x} \mid \mu, \sigma^2) = \prod_{i=1}^{n} f(x_i \mid \mu, \sigma^2)$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left[\tfrac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right]$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left[\tfrac{-1}{2\sigma^2}\sum_{i=1}^{n}\{(x_i-\bar{x})+(\bar{x}-\mu)\}^2\right]$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left[\tfrac{-1}{2\sigma^2}\Big\{\underbrace{\sum_{i=1}^{n}(x_i-\bar{x})^2}_{(n-1)S_x^2} + n(\bar{x}-\mu)^2 + 2(\bar{x}-\mu)\overset{0}{\cancel{\sum_{i=1}^{n}(x_i-\bar{x})}}\Big\}\right]$$

$$= (2\pi\sigma^2)^{-\frac{n}{2}} \exp\left[\tfrac{-1}{2\sigma^2}\{(n-1)S_x^2 + n(\bar{x}-\mu)^2\}\right]$$

$$\Rightarrow \frac{\prod_{i=1}^{n} f(x_i \mid \mu, \sigma^2)}{\prod_{i=1}^{n} f(y_i \mid \mu, \sigma^2)} = \frac{\cancel{(2\pi\sigma^2)^{-\frac{n}{2}}} \exp\left[\tfrac{-1}{2\sigma^2}\{(n-1)S_x^2 + n(\bar{x}-\mu)^2\}\right]}{\cancel{(2\pi\sigma^2)^{-\frac{n}{2}}} \exp\left[\tfrac{-1}{2\sigma^2}\{(n-1)S_y^2 + n(\bar{y}-\mu)^2\}\right]}$$

$$= \exp\left[\tfrac{-1}{2\sigma^2}\{(n-1)S_x^2 + n(\bar{x}-\mu)^2 - (n-1)S_y^2 - n(\bar{y}-\mu)^2\}\right]$$

$$= \exp\left[\tfrac{-1}{2\sigma^2}\{(n-1)(S_x^2 - S_y^2) + n(\bar{x}^2 - 2\bar{x}\mu + \cancel{\mu^2}) - n(\bar{y}^2 - 2\bar{y}\mu + \cancel{\mu^2})\}\right]$$

$$= \exp\left[\tfrac{-1}{2\sigma^2}\{(n-1)(S_x^2 - S_y^2) + n(\bar{x}^2 - \bar{y}^2) + 2n\mu(\bar{y}-\bar{x})\}\right]$$

is constant for μ and σ^2 $\Leftrightarrow$ $S_x^2 = S_y^2$ and $\bar{X} = \bar{Y}$

$\Rightarrow (\bar{X}, S_x^2)$ is a minimal sufficient statistic of (μ, σ^2).

Example

Suppose $X_1, \ldots, X_n \sim \text{Bernoulli}(\theta)$ iid, $0 \leq \theta \leq 1$. Find the MLE of θ and show that it is a sufficient statistic for θ and hence a minimal sufficient statistic for θ.

$$f(x) = \theta^x (1-\theta)^{1-x} \Rightarrow \mathcal{L}(\theta; \underline{x}) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} = \prod_{i=1}^{n} f(x_i)$$

$$\ln \mathcal{L} = (\sum x_i)\ln\theta + (n - \sum x_i)\ln(1-\theta)$$

$$\frac{\partial \ln \mathcal{L}}{\partial \theta} = \frac{\sum x_i}{\theta} - \frac{n-\sum x_i}{1-\theta} = 0 \Rightarrow \frac{\sum x_i}{\theta} = \frac{n - \sum x_i}{1-\theta}$$

$$\therefore \sum x_i \cancel{- \theta\sum x_i} = n\theta \cancel{- \theta \sum x_i} \Rightarrow \hat{\theta} = \frac{\sum x_i}{n} = \bar{x}$$

$$\frac{\prod f(x_i)}{\prod f(y_i)} = \frac{\theta^{\sum x_i}(1-\theta)^{n-\sum x_i}}{\theta^{\sum y_i}(1-\theta)^{n-\sum y_i}} = \frac{\theta^{n\bar{x}}(1-\theta)^{n-n\bar{x}}}{\theta^{n\bar{y}}(1-\theta)^{n-n\bar{y}}} = \theta^{n(\bar{x}-\bar{y})}(1-\theta)^{n(\bar{y}-\bar{x})}$$

is constant wrt $\theta \Leftrightarrow \bar{X} = \bar{Y}$ $\therefore \bar{X}$ is minimal sufficient for θ.

Example

Suppose $X_1, \ldots, X_n \sim \text{Unif}(\theta, \theta+1)$ iid. Determine the minimal sufficient statistic for θ.

$$f(x_i) = \begin{cases} 1 & \theta < x_i < \theta+1 \\ 0 & \text{ow.} \end{cases} = I_{(\theta, \theta+1)}(x_i)$$

$$\prod_{i=1}^{n} f(x_i) = I_{(\theta,\theta+1)}(\underline{x}) = I_{(\theta,\theta+1)}(x_{(1)}, x_{(n)})$$

$$\frac{\prod_1^n f(x_i)}{\prod_1^n f(y_i)} = \frac{I_{(\theta,\theta+1)}(x_{(1)}, x_{(n)})}{I_{(\theta,\theta+1)}(y_{(1)}, y_{(n)})}$$ is constant wrt $\theta \Leftrightarrow X_{(1)} = Y_{(1)}$ and $X_{(n)} = Y_{(n)}$

$\therefore (X_{(1)}, X_{(n)})$ minimal sufficient statistics for θ

Note: the dimension of minimal sufficient statistic doesn't have to equal number of parameters.

Ancillary Statistics

- **Definition:** A statistic $S(X)$ whose distribution does not depend on the parameter θ is an ancillary statistic.
- Alone, contains no information about θ
- Observation on a r.v. whose distribution is fixed and known, unrelated to θ
- When used in conjunction with other statistics, sometimes contain valuable information for inferences about θ

Example

Let $X_1, X_2 \sim \text{Gamma}(\alpha, \theta)$ (indpt), α known. Show that $Z = X_1/(X_1 + X_2)$ is an ancillary statistic for θ.

Let $\begin{cases} Y = X_1 + X_2 \\ Z = \frac{X_1}{X_1+X_2} = \frac{X_1}{Y} \end{cases} \Rightarrow \begin{matrix} X_2 = Y - YZ \\ X_1 = YZ \end{matrix}$ $\quad J = \begin{vmatrix} 1-z & -y \\ z & y \end{vmatrix} = y(1-z) + yz = y \geq 0$

$g(y,z) = f(yz, y-yz)\cdot|J| = f_{X_1}(yz) f_{X_2}(y-yz)\cdot y$

$= \left(\frac{1}{\Gamma(\alpha)\theta^\alpha}(yz)^{\alpha-1}\cancel{e^{-yz/\theta}}\right)\left(\frac{1}{\Gamma(\alpha)\theta^\alpha}(y-yz)^{\alpha-1}e^{-\frac{(y-\cancel{yz})}{\theta}}\right)y$

$= \left(\frac{1}{\Gamma(\alpha)\theta^\alpha}\right)^2 y^{2\alpha-1} z^{\alpha-1}(1-z)^{\alpha-1} e^{-y/\theta}$; $0<z<1$, $y \geq 0$ where $Y \sim \text{Gamma}(2\alpha, \theta)$

$g(z) = \int_0^\infty g(y,z)\,dy = \left(\frac{1}{\Gamma(\alpha)\cancel{\theta^\alpha}}\right)^2 z^{\alpha-1}(1-z)^{\alpha-1}\cdot\Gamma(2\alpha)\cancel{\theta^{2\alpha}} \cancel{\int_0^\infty \frac{1}{\Gamma(2\alpha)\theta^{2\alpha}} y^{2\alpha-1} e^{-y/\theta}\,dy}^{\,1}$

$= \frac{\Gamma(2\alpha)}{\Gamma(\alpha)\Gamma(\alpha)} z^{\alpha-1}(1-z)^{\alpha-1}$ ie $Z \sim \text{Beta}(\alpha,\alpha)$ which does not depend on θ $\therefore$ Z is ancillary statistic for θ.

Example

Consider $X_1, \dots, X_n$ random sample having the model $X_i = \theta + W_i$, $i = 1, \dots, n$, where $-\infty < \theta < \infty$ and $W_1, \dots, W_n$ are iid r.v.'s whose pdf does not depend on θ.

Let $Z = u(X_1, \dots, X_n)$ be a statistic s.t.

$u(x_1 + d, \dots, x_n + d) = u(x_1, \dots, x_n)$, for all real d.

Hence, $Z = u(W_1 + \theta, \dots, W_n + \theta) = u(W_1, \dots, W_n)$ is a function of Ws alone (ie, no θ).

⇒ Z has a distribution that doesn't depend on θ, therefore Z is ancillary.

Z is called a location-invariant statistic.

distributions
W_i don't depend on θ

$X_i = \theta + W_i$

$W_i = X_i - \theta$ doesn't depend on θ

⇒ $f_W(w) = f_X(\theta + w_i)$ does not depend on θ

$u(X_1+d, \dots, X_n+d) = u(W_1+\theta+d, \dots, W_n+\theta+d) = u(W_1, \dots, W_n)$ by location invariance $\therefore f(z) = f(u(W_1, \dots, W_n))$ where W's pdfs don't depend on θ $\therefore$ $f(z)$ doesn't depend on θ ⇒ Z ancillary.

Example (cont.)

- Location-invariant statistics are ancillary.
- Examples of location-invariant statistics:
  - Sample variance, $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left[(X_i+d) - \frac{\sum_{j=1}^{n}(X_j+d)}{n}\right]^2$
  - Sample range, $R = X_{(n)} - X_{(1)}$
  - Mean deviation from the sample median,

$$\frac{1}{n}\sum_{i=1}^{n} |X_i - \mathrm{med}(X_i)|$$

- Analogous arguments for scale-invariant statistics

Example

Consider $X_1, \ldots, X_n$ random sample having the model $X_i = \theta W_i, i = 1, \ldots, n$, where $\theta > 0$ and $W_1, \ldots, W_n$ are iid r.v.'s whose pdf does not depend on θ.

Let $Z = u(X_1, \ldots, X_n)$ be a statistic s.t.

$$u(dx_1, \ldots, dx_n) = u(x_1, \ldots, x_n), \text{ for all real } d.$$

Hence, $Z = u(\theta W_1, \ldots, \theta W_n) = u(W_1, \ldots, W_n)$ is a function of Ws alone (ie, no θ).

⇒ Z has a distribution that doesn't depend on θ, therefore Z is ancillary.

Z is called a <u>scale-invariant statistic</u>.

W_i distribution doesn't depend on θ

$X_i = \theta W_i$

$\Rightarrow W_i = X_i/\theta$

$\therefore g(w_i) = f(\theta w_i)\cdot\theta$ doesn't depend on θ

$Z = u(\theta w_1, \ldots, \theta w_n)$
$= u(w_1, \ldots, w_n)$
by scale invariance, where w_i pdf doesn't depend on θ $\therefore Z$ has pdf that doesn't depend on $\theta \Rightarrow Z$ ancillary.

Theorem

Let $X_1, \ldots, X_n$ denote a random sample from a distribution having a pdf $f(x;\theta), \theta \in \Omega$, where Ω is an interval set. Suppose that the statistic Y_1 is a complete and sufficient statistic for θ. Let $Z = u(X_1, \ldots, X_n)$ be any other statistic (not a function of Y_1 alone). If the distribution of Z does not depend upon θ, then Z is independent of the sufficient statistic Y_1.

Basu's Theorem

If $T(X)$ is a complete and minimal sufficient statistic, then $T(X)$ is independent of every ancillary statistic.

The point: Complete sufficient statistic is indpt of ancillary statistic.

Example

Let $X_1, \ldots, X_n$ be a random sample from a distribution having pdf $f(x;\theta) = \exp[-(x-\theta)]$, $\theta < x < \infty$, $-\infty < \theta < \infty$. Show that $X_{(1)}$ is independent of location-invariant statistics.

SEE ATTACHED

Example

Let X_1, X_2 be a random sample from a distribution having pdf $f(x;\theta) = (1/\theta)\exp(-x/\theta)$, $0 < x < \infty$, $0 < \theta < \infty$. Show that $Y = X_1 + X_2$ is a complete sufficient statistic for θ. Hence, Y is independent of scale-invariant statistics.

SEE ATTACHED

$X_1, \ldots, X_n$ $f(x) = e^{-(x-\theta)}$, $\theta < x < \infty$. Show $X_{(1)}$ indpT of location-invariant statistics.

① Show $X_{(1)}$ complete and minimal sufficient.

$$\prod_{i=1}^{n} f(x_i) = e^{-\Sigma(x_i - \theta)} I_{(\theta,\infty)}(x_{(1)}) = \underbrace{e^{-\Sigma x_i}}_{k_2(\underline{x})} \underbrace{e^{n\theta} I_{(\theta,\infty)}(x_{(1)})}_{k_1(x_{(1)};\theta)} \quad \therefore \text{by NFFT, } Y = X_{(1)} \text{ sufficient}$$

$$f_{X_{(1)}}(y) = n\left[1 - (1 - e^{-(y-\theta)})\right]^{n-1} e^{-(y-\theta)} = ne^{-n(y-\theta)}, \; y > \theta$$

Setting $E(g(Y)) = E(g(X_{(n)})) = 0 \; \forall \theta$:

$$E(g(Y)) = \int_{\theta}^{\infty} g(y) \cdot ne^{-n(y-\theta)} dy = 0 \; \forall \theta. \text{ Letting } w = y - \theta, \; dw = dy$$

$$= \cancel{n} \int_0^{\infty} g(w+\theta) e^{-nw} dw = 0$$

$\int_0^{\infty} h(w) e^{-nw} dw = 0$ is a Laplace transform so its integral equals zero $\Leftrightarrow h(w) = g(w+\theta) = 0 \; \forall w$. Because $g(w+\theta) = 0 \; \forall w, \forall \theta$

$\Rightarrow Y = X_{(1)}$ complete sufficient for θ.

To show minimal sufficient, $\dfrac{\prod f(x_i)}{\prod f(y_i)} = \dfrac{e^{-\Sigma x_i} \cancel{e^{n\theta}} I_{(\theta,\infty)}(x_{(1)})}{e^{-\Sigma y_i} \cancel{e^{n\theta}} I_{(\theta,\infty)}(Y_{(1)})}$

is constant wrt $\theta \Leftrightarrow X_{(1)} = Y_{(1)} \; \therefore X_{(1)}$ minimal sufficient

② Show location-invariant statistics are ancillary.

For a particular statistic that is location-invariant, we can show ancillary because, letting $W_i = X_i - \theta$, $f(w) = e^{-w}$, $0 < w < \infty$ does not depend on θ. Depending on the structure of $Z = u(X_1 + d, \ldots, X_n + d) = u(X_1, \ldots, X_n)$, we can show that Z can be written as a function of Ws where the Ws have pdf that doesn't depend on $\theta \Rightarrow Z$ has pdf that doesn't depend on $\theta \; \therefore Z$ ancillary.

$\Rightarrow$ By Basu's Thm., $X_{(1)}$ indpt of location-invariant statistics.

X_1, X_2 iid $\quad f(x;\theta) = \frac{1}{\theta} e^{-x/\theta}, \quad 0 < x < \infty, \quad 0 < \theta < \infty$

$$= \exp\Big[\underbrace{-\ln\theta}_{q(\theta)} \underbrace{-\tfrac{1}{\theta}}_{p(\theta)} \underbrace{x}_{K(x)} + \underbrace{0}_{S(x)}\Big]$$ has the form of an exponential family

$\Rightarrow Y = \sum_{i=1}^{2} K(x_i) = X_1 + X_2$ complete sufficient.

$$\prod_{i=1}^{2} f(x_i) = \left(\frac{1}{\theta} e^{-x_1/\theta}\right)\left(\frac{1}{\theta} e^{-x_2/\theta}\right) = \frac{1}{\theta^2} e^{\frac{-(x_1+x_2)}{\theta}}$$

$$\therefore \frac{\prod_{i=1}^{2} f(x_i)}{\prod_{i=1}^{2} f(y_i)} = \frac{\cancel{\frac{1}{\theta^2}}\, e^{\frac{-(x_1+x_2)}{\theta}}}{\cancel{\frac{1}{\theta^2}}\, e^{\frac{-(y_1+y_2)}{\theta}}} = \exp\left[\frac{-1}{\theta}\left\{(x_1+x_2) - (y_1+y_2)\right\}\right]$$

is constant wrt $\theta \iff \sum_{i=1}^{2} X_i = \sum_{i=1}^{2} Y_i \quad \therefore X_1 + X_2$ minimal sufficient for θ.

Meanwhile, for a statistic that is scale-invariant, we can show it is ancillary because, letting $W_i = X_i/\theta$, $f(w_i) = e^{-w_i}$, $0 < w_i < \infty$ does not depend on θ. Thus, given ~~some~~ $Z = u(dx_1, \dots, dx_n) = u(x_1, \dots, x_n)\ \forall d$, we can show that $Z = u(X_1, \dots, X_n) = u(\theta W_1, \dots, \theta W_n) = u(W_1, \dots, W_n)$, ie a function of Ws where the Ws have pdf that doesn't depend on $\theta \Rightarrow Z$ has pdf that doesn't depend on $\theta \ \therefore Z$ ancillary

$\Rightarrow$ by Basu's Thm., $X_1 + X_2$ indpt of scale-invariant statistics.