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 $X_1, \ldots, X_n$, n a fixed positive integer, denote a random sample from a distribution that has pdf/pmf $f(x;\theta), \theta \in \Omega$, let $Y_1 = u_1(X_1, \ldots, X_n)$ 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 $X_1, X_2, \ldots, X_n$ (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 <u>complete sufficient statistic</u> 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 $\mathrm{E}\big(\varphi(Y_1)\big) = \theta \quad \Rightarrow \varphi(Y_1)$ is unique MVUE of θ.

Example

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

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 $P(X_1 \le 2)$.

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 jointly sufficient 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 $\boldsymbol{\theta}$.

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

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

Find the joint sufficient statistics for θ_1 and θ_2.

(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 regular case of the exponential family 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 <u>minimal sufficient statistic</u> 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 $\left(\bar{y}, s_y^2\right)$ be the sample means and variances corresponding to the x and y samples, respectively. Then, the ratio of densities is

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

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

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).

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.

Example

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

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

Ancillary Statistics

- **Definition:** A statistic $S(\boldsymbol{X})$ 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 $X_1, X_2 \sim \mathrm{Gamma}(\alpha, \theta)$ iid, α known. Show that $Z = X_1/(X_1 + X_2)$ is an ancillary statistic for θ.

Example

Consider $X_1, \ldots, X_n$ random sample having the model $X_i = \theta + W_i, \;\; i = 1, \ldots, n$, where $-\infty < \theta < \infty$ 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.

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

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

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

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

Example (cont.)

- Location-invariant statistics are ancillary.
- Examples of location-invariant statistics:
  - Sample variance, S^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 $X_1, \dots, X_n$ random sample having the model $X_i = \theta W_i, i = 1, \dots, n$, where $\theta > 0$ 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.

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

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

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

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

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 $T(\boldsymbol{X})$ is a complete and minimal sufficient statistic, then $T(\boldsymbol{X})$ is independent of every 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.

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.