MATH 503: Mathematical Statistics

Lecture 4: Properties of Point Estimators II

Reading: Sections 6.1-6.2, 7.3

Kimberly F. Sellers

Department of Mathematics & Statistics

Today's Topics

- Recap: Sufficient statistics
- Uniform minimum variance unbiased estimators (UMVUEs)
  - Rao-Blackwell Theorem
  - Completeness
  - Lehmann-Scheffé Theorem
  - Uniqueness
- Exponential families
- Comments connecting Rao-Blackwell and Lehmann-Scheffé

Sufficiency

Let $X_1, \ldots, X_n$ denote a random sample of size n 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 statistic whose pdf/pmf is $f_{Y_1}(y_1;\theta)$. Then Y_1 is a sufficient statistic for θ iff.

$$\frac{f(x_1;\theta)f(x_2;\theta)\cdots f(x_n;\theta)}{f_{Y_1}[u_1(x_1, \ldots, x_n);\theta]} = H(x_1, \ldots, x_n),$$

where $H(x_1, \ldots, x_n)$ does not depend on $\theta \in \Omega$.

Neyman-Fisher Factorization Thm

Let $X_1, \ldots, X_n$ denote a random sample from a distribution that has pdf/pmf $f(x;\theta), \theta \in \Omega$. The statistic $Y_1 = u_1(X_1, \ldots, X_n)$ is a sufficient statistic for θ iff. we can find two nonnegative functions, k_1 and k_2, such that

$$f(x_1;\theta)f(x_2;\theta)\cdots f(x_n;\theta) = k_1[u_1(x_1, \ldots, x_n);\theta] \cdot k_2(x_1, \ldots, x_n)$$

where $k_2(x_1, \ldots, x_n)$ does not depend on θ.

Uniform Minimum Variance Unbiased Estimators (UMVUEs)

- For a given positive integer n, $Y = u(X_1, \ldots, X_n)$ is a uniform minimum variance unbiased estimator (UMVUE) of the parameter θ
  - if Y is unbiased, and
  - if the variance of Y is less than or equal to the variance of every other unbiased estimator of θ.

Rao-Blackwell Theorem

(Hogg, McKean, & Craig)

C.R. Rao

David Blackwell

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 $Y_2 = u_2(X_1, \ldots, X_n)$, not a function of Y_1 alone, be an unbiased estimator of θ. Then $E(\,Y_2 \mid y_1\,) = \varphi(y_1)$ 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.

Rao-Blackwell Theorem

(Casella & Berger)

C.R. Rao

David Blackwell

Let W be any unbiased estimator of $\tau(\theta)$, and let T be a sufficient statistic of θ. Define $\phi(T) = E(W|T)$. Then $E_\theta \phi(T) = \tau(\theta)$ and $\mathrm{Var}_\theta \phi(T) \leq \mathrm{Var}_\theta W$ for all θ, that is, $\phi(T)$ is a uniformly better unbiased estimator of $\tau(\theta)$.

Notes re. Rao-Blackwell Thm.

- If we know a sufficient statistic for the parameter exists, the MVUE will be a function of the sufficient statistic.
- This does not mean that we first need to find an unbiased statistic!
- Focus on functions of sufficient statistics

Theorem

- Let $X_1, \ldots, X_n$ denote a random sample from a distribution that has pdf/pmf $f(x; \theta), \theta \in \Omega$. If a sufficient statistic $Y_1 = u_1(X_1, \ldots, X_n)$ for θ exists and if a MLE $\hat{\theta}$ of θ, also exists uniquely, then $\hat{\theta}$ is a function of $Y_1 = u_1(X_1, \ldots, X_n)$.

- **The point:** MLEs are functions of sufficient statistics.

Example

Let $X_1, \ldots, X_n$ denote a random sample from a distribution that has pdf $f(x;\theta) = \theta e^{-\theta x},\ 0 < x < \infty$.

1. Find a sufficient statistic for θ.
2. Find the MLE of θ.
3. Determine a MVUE of θ.

Completeness

Let the random variable Z have a pdf/pmf that is one member of the family $\{h(z;\theta): \theta \in \Omega\}$. If the condition $E[u(Z)] = 0$, for every $\theta \in \Omega$, requires that $u(z)$ be zero except on a set of points that has probability zero for each $h(z;\theta): \theta \in \Omega$, then the family $\{h(z;\theta): \theta \in \Omega\}$ is called a <u>complete family</u> of pdfs/pmfs.

Note: One-to-one functions of complete sufficient statistics are themselves complete sufficient.

Example 1

Let $X_1, \dots, X_n$ ~Poisson(θ) iid.

1. Determine a sufficient statistic for θ.
2. What is the pdf associated with this statistic?
3. Show that this statistic is complete.

Example 2

Let $X_1, \ldots, X_n \sim$Uniform$(0, \theta)$ iid, $\theta > 0$. Show $X_{(n)}$ is complete sufficient for θ.

Example 3

Let $T \sim$Binomial(n, p), $0 < p < 1$. Show T is complete.

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.

Uniqueness

- In most instances, if there is one function $\varphi(Y_1)$ that is unbiased, then it is the only unbiased estimator based on the sufficient statistic Y_1
- Lehmann-Scheffe $\Rightarrow$ unbiased estimators based on complete sufficient statistics are unique.

How to Determine UMVUEs?

- Expected value of complete sufficient statistic
- Conditional expectation of unbiased estimate given sufficient statistic

Example 4

Let a random sample of size n be taken from a distribution of the discrete type with pmf $f(x;\theta) = \frac{1}{\theta},\ x = 1,2,\dots,\theta$, where θ is an unknown positive integer.

1. Show that the largest observation, say $Y = X_{(n)}$, of the sample is a complete sufficient statistic for θ.
2. Prove that $[Y^{n+1} - (Y-1)^{n+1}]/[Y^n - (Y-1)^n]$ is the unique UMVUE of θ.

Exponential Family/Class

A pdf of the form

$$f(x;\theta) = \exp[p(\theta)K(x) + S(x) + q(\theta)], x \in S^*$$

is said to be a member of the regular exponential class of probability density or mass functions if

1. S^*, the support of X, does not depend on θ
2. $p(\theta)$ is a nontrivial continuous function of $\theta \in \Omega$
3. Finally,
   - If X is a continuous rv then each of $K'(x) \not\equiv 0$ and $S(x)$ is a continuous function of $x \in S^*$
   - If X is a discrete rv then $K(x)$ is a nontrivial function of $x \in S^*$

Example 5

Show that the Normal($0,\sigma^2 = \theta$) distribution is a member of the regular exponential class.

Example 6

Is the Uniform$(0, \theta)$ distribution a member of the regular exponential class?

What about for a random sample?

Result: $Y_1 = \sum_{i=1}^{n} K(x_i)$ is a sufficient statistic for θ.

Theorem

Let $X_1, \ldots, X_n$, denote a random sample from a distribution that represents a regular case of the exponential class, with pdf/pmf given by

$$f(x;\theta) = \exp[p(\theta)K(x) + S(x) + q(\theta)], x \in S^*$$

Consider the statistic $Y_1 = \sum_{i=1}^{n} K(x_i)$. Then,

1. The pdf/pmf of Y_1 has the form,

   $$f_{Y_1}(y_1;\theta) = R(y_1)\exp[p(\theta)y_1 + nq(\theta)]$$

   for $y_1 \in S_{Y_1}^*$ and some function $R(y_1)$. Neither $S_{Y_1}^*$ nor $R(y_1)$ depend on θ.
2. $E(Y_1) = -nq'(\theta)/p'(\theta)$
3. $\mathrm{Var}(Y_1) = n[1/p'(\theta)]^3\{p''(\theta)q'(\theta) - q''(\theta)p'(\theta)\}$

Example 7

1. Consider $X \sim \text{Poisson}(\theta)$. Show that it is a member of the regular exponential class.
2. For a random sample, $X_1, \ldots, X_n \sim \text{Poisson}(\theta)$, determine the sufficient statistic, Y_1.
3. Use the above theorem to verify $E(Y_1)$ and $V(Y_1)$.

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 <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 $\mathrm{E}\big(\varphi(Y_1)\big) = \theta$ implies $\varphi(Y_1)$ is unique and UMVUE of θ.

Example 8

Consider $X_1, \dots, X_n$ ~Normal(θ, σ^2) iid, σ known.
Show that $Y_1 = \sum_{i=1}^{n} X_i$ is complete sufficient.
Determine the unique UMVUE of θ.

Example 9

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

Example 10

Let a random sample of size n, i.e. $X_1, \ldots, X_n$, 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 $P(X_1 \leq 2)$.