Ch 7. Sufficiency

Ch 7.1 A sufficient statistic for a parameter, p433

- Example 7.1

  - Experiment A: Observe IID sample X_1 and X_2 from $b(1, \theta)$, $0 < \theta < 1$

  - Experiment B: Observe X_1 and $X_1 + X_2$ where $X_1, X_2 \overset{iid}{\sim} b(1, \theta)$

  - Experiment C

    1st step: Observe $Y = X_1 + X_2$

    2nd step: Observe X_1 given $Y = y$

- Probability Structure of Experiment A

(x_1, x_2)	$(0,0)$	$(0,1)$	$(1,0)$	$(1,1)$
$P(x_1, x_2)$	$(1-\theta)^2$	$(1-\theta)\theta$	$\theta(1-\theta)$	θ^2

- Probability Structure of Experiment B

(x_1, x_2)	$(0,0)$	$(0,1)$	$(1,1)$	$(1,2)$
$P(x_1, x_2)$	$(1-\theta)^2$	$(1-\theta)\theta$	$\theta(1-\theta)$	θ^2

- Experiment C

  1st step: Observe $Y = X_1 + X_2$

  2nd step: Observe X_1 given $Y = y$

- Probability Structure of Experiment C
  1. $P((X_1, X_2) = (0,0)|Y = 0)) = 1$
  2. $P((X_1, X_2) = (0,1)|Y = 1)) = P((X_1, X_2) = (1,0)|Y = 1)) = 0.5$
  3. $P((X_1, X_2) = (1,1)|Y = 2)) = 1$

- From this probability structure, we can see that the conditional probability of (X_1, X_2) given Y does not depend on θ. This implies that if Y is given, knowing (X_1, X_2) does not give any useful information for the inference of θ.

Definition

Let $X_1, \ldots, X_n$ be a random sample from pdf $f(x;\theta)$, $\theta \in \Omega$. $T(\underline{X}) = T(X_1, \ldots, X_n)$ is called <u>sufficient statistic</u> for $\theta \in \Omega$ if

(i) Y is a statistic

(ii) $P_\theta\left[(X_1, \ldots, X_n) \in A | T(\underline{X}) = y\right]$ does not depend on θ for each y and all A.

- Example 7.2 (p.434)

For $X_1, \ldots, X_n \overset{iid}{\sim} b(1, \theta)$, $Y = X_1 + \ldots + X_n$ is a sufficient statistic for θ.

Theorem (Factorization theorem, p436)

Let $X_1, \ldots, X_n$ be a random sample from pdf $f(x;\theta)$, $\theta \in \Omega$. $Y = T(\underline{X})$ is a sufficient statistic if and only if we can find two nonnegative function k_1 and k_2 such that

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = k_1(y,\theta)k_2(\underline{x})$$

- Example 7.2 (revisited): $X_1, \ldots, X_n \overset{iid}{\sim} b(1, \theta)$

- Example 7.3: $X_1, \ldots, X_n \overset{iid}{\sim} N(\mu, 1)$

- Example 7.4 (p.438): $X_1, \ldots, X_n \overset{iid}{\sim} Beta(\theta, 1)$

- Example 7.5: $X_1, \ldots, X_n \overset{iid}{\sim} U[0, \theta]$

- Fact: one-to-one function of a sufficient statistic for θ is also a sufficient statistic for θ.

Theorem (Rao-Blackwell, p.441)

Let $X_1, \ldots, X_n$ *be a random sample from pdf* $f(x;\theta)$, $\theta \in \Omega$. *If* $Y = T(\underline{X})$ *is a sufficient statistic for* θ, *and* $\hat{\theta}_1$ *is an unbiased estimator of* θ, *then* $\hat{\theta}_2 = E(\hat{\theta}_1 | T(\underline{X})) = E(\hat{\theta}_1) = \theta$ *with* $V(\hat{\theta}_2) \leq V(\hat{\theta}_1)$.

Remarks

- For any unbiased estimator $\hat{\theta}_1$ for θ, we can always find a better unbiased estimator $\hat{\theta}_2 = E(\hat{\theta}_1 | u_1(\underline{X}))$ whe a sufficient statistic $u_1(\underline{X})$ is given.

- Although SS is not unique, we can know that a desirable estimator should be at least a function of a sufficient statistic.

Theorem

If a unique MLE of $\hat{\theta}$ exists, $\hat{\theta}$ should be a function of SS.

- Fact: MVUE with finite variance is unique when it exists.

Proof.

Let $\hat{\theta}_1$ and $\hat{\theta}_2$ be MVUE's of θ. We will show that $\hat{\theta}_1 = \hat{\theta}_2$ or $P(\hat{\theta}_1 = \hat{\theta}_2) = 1$. For this, define $\hat{\theta}_3 = (\hat{\theta}_1 + \hat{\theta}_2)/2$. Then $\hat{\theta}_3$ is unbiased and this means $V(\hat{\theta}_3) \geq V(\hat{\theta}_1) = V(\hat{\theta}_2)$. Hence, $Cov(\hat{\theta}_1, \hat{\theta}_2) \geq V(\hat{\theta}_1)$ and this implies $V(\hat{\theta}_1 - \hat{\theta}_2) = 0$. □

Definition

$X_1, \ldots, X_n$: random sample from pdf $f(x;\theta)$, $\theta \in \Omega$.

(a) $Y = u(\underline{\mathsf{X}})$ is a complete statistic for θ if and only if

$$E_\theta(g(Y)) = 0 \ \text{ for all } \ \theta \in \Omega \text{ implies } \ P_\theta(g(Y) = 0) = 1 \ \text{ for all } \ \theta \in \Omega$$

(b) $Y = u(\underline{\mathsf{X}})$ is a complete sufficient statistic (CSS) for θ if Y is complete and sufficient for θ.

Theorem (Lehmann & Scheffe, p.446)

If T is CSS for θ and $E(\phi(T)) = \theta$, then $\phi(T)$ is the MVUE of θ.

- Why this theorem works?

  - From Rao-Blackwell, we know that a good estimator should be a function of SS.
  - If a given sufficient statistic T is also complete, we can say that $h(T) = 0$ whenever $E_\theta(h(T)) = 0$ for all $\theta \in \Omega$.
  - When T is CSS for θ, suppose that there are two unbiased estimators for θ: $\phi(T)$ and $\psi(T)$. Then the completeness of T implies that $\phi(T) = \psi(T)$.
  - This means that there is only one unbiased estimator which is a function of CSS.
  - How to show the completeness? No general way exists..

- Example 7.5 (revisited): $X_1, \ldots, X_n \overset{iid}{\sim} U[0, \theta]$, $\theta > 0$

$\rightarrow$ From Example 7.5, we know that $Y_n = \max(X_1, \ldots, X_n)$ is a sufficient statistic for θ.

1. Y_n is also a complete statistic?
   (i) Assume $E_\theta(g(Y_n)) = 0$ for all $\theta > 0$
   (ii) Show that $g(y) = 0$

- Example 7.5 (revisited): $X_1, \ldots, X_n \overset{iid}{\sim} U[0, \theta]$, $\theta > 0$

$\rightarrow$ From Example 7.5, we know that $Y_n = \max(X_1, \ldots, X_n)$ is a sufficient statistic for θ.

1. Y_n is also a complete statistic? Yes!
2. What is the MVUE of θ?
3. Is there any easy way to show completeness? For some family of distributions, we can show the completeness.

- Example 7.6: $X_1, \ldots, X_n \overset{iid}{\sim} Gamma(1, \theta)$, $\theta > 0$

Definition (Laplace transform)

The Laplace transform of $f(x)$ is

$$\mathcal{L}f(x) = \int_0^\infty f(t)e^{-tx}dt$$

provided with $\int_0^\infty f(t)e^{-tx}dt < \infty$

Theorem (Uniqueness of Laplace transform)

$\mathcal{L}f(x) = \mathcal{L}g(x)$ *if and only if* $f(x) = g(x)$

- Example 7.6: $X_1, \ldots, X_n \overset{iid}{\sim} Gamma(1, \theta)$, $\theta > 0$
$\rightarrow \sum_{i=1}^{n} X_i$ is a complete and sufficient statistic

Definition (Exponential family of distributions, p. 449)

$f(x;\theta)$ is said to be a member of the regular exponential family if

(i) Support of $f(x;\theta)$ does not depend on θ.

(ii) $f(x;\theta) = \exp(p(\theta)K(x) + H(x) + q(\theta))$, where $p(\theta)$ is a nontrivial continuous function of θ.

(iii) $p(\theta)$ is a nontrivial continuous function of θ, and $K(x)$ is a nontrivial function of x

- Examples
  - Binomial distributions
  - Gamma distributions
  - Normal distributions
  - Beta distributions
  - Uniform distributions $\rightarrow$ NOT exponential family!

Theorem (p. 450)

If $f(x;\theta)$ belongs to the exponential family (i.e. $f(x;\theta) = \exp(p(\theta)K(x) + H(x) + q(\theta))$, then

(i) $q(\theta) = -\log\left(\int \exp(p(\theta)K(x) + H(x))dx\right)$

(ii) $E(K(X)) = -q'(\theta)/p'(\theta)$

(iii)

$$V(K(X)) = -\frac{p''(\theta)E(K(X)) + q''(\theta)}{(p'(\theta))^2} = \frac{p''(\theta)q'(\theta) - q''(\theta)p'(\theta)}{(p'(\theta))^3}$$

If $p(\theta) = \theta$, then $E(K(X)) = -q'(\theta)$ and $V(K(X)) = -q''(\theta)$

Proof.

Note that

$$\frac{d}{d\theta}f(x;\theta) = (p'(\theta)K(x) + q'(\theta))f(x;\theta)$$
$$\frac{d^2}{d\theta^2}f(x;\theta) = (p''(\theta)K(x) + q''(\theta))f(x;\theta) + (p'(\theta)K(x) + q'(\theta))^2 f(x;\theta)$$

- Example 7.7

(i) $X \sim b(n, \theta)$

- Example 7.7

(ii) X follows exponential distribution having mean λ.

Theorem (p.451)

For $X_1, \ldots, X_n \overset{iid}{\sim} f(x;\theta)$, if $f(x;\theta)$ is a member of exponential family, then $\sum K(X_i)$ is CSS

- Example 7.7 (revisited)

(i) If $X_1, \dots, X_n \overset{iid}{\sim} b(1, \theta)$, find the MVUE of θ.

- Example 7.7(revisited)

(ii) If $X_1, \ldots, X_n$ is a random sample from the exponential distribution having mean λ, find the MVUE of λ.

- Example 6.9 (revisited)

If $X_1, \ldots, X_n \overset{iid}{\sim} Beta(\theta, 1)$, find the MVUE of θ.
In Example 6.12, the MLE of θ is $\hat{\theta}^{MLE} = -\frac{n}{\sum \log X_i}$ and $E(\hat{\theta}^{MLE}) = \frac{n\theta}{n+1}$. What is the MVUE of θ?

- Rao-Blackwell+Lehmann-Scheffe
  If $X_1, \ldots, X_n \stackrel{iid}{\sim} f(x;\theta)$, $\hat{\theta}_1$ is unbiased, and Y is CSS, then $\hat{\theta}_2 = E(\hat{\theta}_1|Y)$ is the MVUE of θ.
- Remark
  1. If your estimator is a function of CSS and unbiased, then your estimator is the unique MVUE.
  2. If you have an unbiased estimator $\hat{\theta}$ which is not a function of CSS, $\hat{\theta}^* = E(\hat{\theta}|Y)$ is the MVUE.

- Example 7.8

If $X_1, \ldots, X_n \overset{iid}{\sim} b(1, p)$, find the MVUE of p.

- Example 7.9

If $X_1, \ldots, X_n$ is a random sample from the exponential distribution having mean θ, find the MVUE of $\eta = e^{-1/\theta} = P(X_1 > 1)$.