Inference 2, 2023, lecture 8

Rolf Larsson

November 23, 2023

Today

Chap. 4. Estimation (continued):

- The multidimensional case
- The univariate MLE revisited
- Rao-Blackwell
- Lehmann-Scheffé

The multidimensional case

- Suppose that the parameter space is k dimensional: $\theta \in \Theta \subseteq \mathcal{R}^k$.
- The parameter of interest, $\gamma = g(\theta) = \{g_1(\theta), ..., g_m(\theta)\}'$ is m dimensional: $g : \Theta \to \Gamma \subseteq \mathcal{R}^m$.
- An estimator $T(\mathbf{X}) = \{T_1(\mathbf{X}), ..., T_m(\mathbf{X})\}'$ is called unbiased for γ if

$$\mathrm{E}_\theta\{T_j(\mathbf{X})\} = g_j(\theta)$$

for all $\theta \in \Theta$ and all $j = 1, ..., m$.

The multidimensional case

Definition (4.10)

Let T and T^* be two unbiased estimators for γ. We say that T^* has smaller covariance matrix than T at $\theta \in \Theta$, if the covariance matrices of T and T^* satisfy

$$u^{\mathrm{T}}\{\mathrm{Cov}_\theta(T^*) - \mathrm{Cov}_\theta(T)\}u \leq 0$$

for all $u \in \mathcal{R}^m$. We write $\mathrm{Cov}_\theta(T^*) \preceq \mathrm{Cov}_\theta(T)$.

- We may also write $\mathrm{Cov}_\theta(T) \succeq \mathrm{Cov}_\theta(T^*)$.
- $\mathrm{Cov}_\theta(T^*) \preceq \mathrm{Cov}_\theta(T)$ if.f. $\mathrm{Var}(u^{\mathrm{T}} T^*) \leq \mathrm{Var}(u^{\mathrm{T}} T)$ for all u, because $\mathrm{Var}(u^{\mathrm{T}} T) = u^{\mathrm{T}}\mathrm{Cov}(T)u$.
- If $\mathrm{Cov}_\theta(T^*) \preceq \mathrm{Cov}_\theta(T)$ and $\mathrm{Cov}_\theta(T^*)$ and $\mathrm{Cov}_\theta(T)$ are diagonal, then all elements of $\mathrm{Cov}_\theta(T^*)$ are less than or equal to the corresponding elements of $\mathrm{Cov}_\theta(T)$.

The multidimensional case

- Let $(D_\theta g)(\theta)$ be the $m \times k$ matrix of partial derivatives

$$\frac{\partial g_j(\theta)}{\partial \theta_l},\ j = 1, ..., m,\ \ l = 1, ..., k.$$

- Let $V(\theta; \mathbf{X})$ be the k-dimensional vector of partial derivatives

$$\frac{\partial}{\partial \theta_l} \log\{L(\theta; \mathbf{X})\},\ \ l = 1, ..., k.$$

- Let the Fisher information matrix be

$$I_{\mathbf{X}}(\theta) = \mathrm{Cov}_\theta\{V(\theta; \mathbf{X})\}.$$

- Assume that $I_{\mathbf{X}}(\theta)$ is non singular. Under regularity conditions 1, 2' and 4', the Cramér-Rao inequality generalizes to

$$\mathrm{Cov}_\theta(T) \succeq (D_\theta g)(\theta) I_{\mathbf{X}}(\theta)^{-1} (D_\theta g)(\theta)^{\mathrm{T}} \quad \text{for all} \quad \theta \in \Theta.$$

The univariate MLE revisited

Theorem (4.4)

Under regularity conditions 1 and 2'-4', a maximum likelihood estimator (MLE) has the following properties:

1. *The MLE depends on the data only via the sufficient statistic.*
2. *If there exists an efficient unbiased estimator* $\tilde{\theta}$*, then* $\tilde{\theta} = \hat{\theta}_{\mathrm{MLE}}$ *with probability one.*

Rao-Blackwell

How can we improve on a given unbiased estimator?

Theorem (4.5)

The Rao-Blackwell theorem.
Let T be a sufficient statistic for the statistical model $\mathcal{P}$, and let $\tilde{\gamma}$ be an unbiased estimator for the parameter $\gamma = g(\theta) \in \mathcal{R}^k$.
Define $\hat{\gamma}(T) = \mathrm{E}_\theta(\tilde{\gamma}|T)$.

1. *The conditional expectation $\hat{\gamma}$ is independent on θ.*
2. *Furthermore, for all $\theta \in \Theta$,*
   $\mathrm{E}_\theta(\hat{\gamma}) = \gamma, \quad \mathrm{Cov}_\theta(\hat{\gamma}) \preceq \mathrm{Cov}_\theta(\tilde{\gamma})$
3. *If $\mathrm{trace}\{\mathrm{Cov}_\theta(\tilde{\gamma})\} < \infty$, then $\mathrm{Cov}_\theta(\hat{\gamma}) = \mathrm{Cov}_\theta(\tilde{\gamma})$ if.f.*
   $\mathrm{P}_\theta(\hat{\gamma} = \tilde{\gamma}) = 1$.

Consequence: To find an optimal statistic, only search among those that are functions of a sufficient statistic.

Rao-Blackwell

Example 1:

Suppose that $\mathbf{X} = (X_1, ..., X_n)$ where the X_i are independent $\mathrm{Po}(\lambda)$.

1. Show that an unbiased estimator of λ is given by X_1.
2. Show that a sufficient statistic is given by $T(\mathbf{X}) = \sum_{i=1}^n X_i$.
3. Use the Rao-Blackwell theorem to find an unbiased estimator for λ with smaller variance than X_1.

Lehmann-Sheffé

Is there any criterion under which the Rao-Blackwell theorem results in the *minimum* variance statistic (BUE)?

Definition (4.11)

A statistical model $\{\mathrm{P}_\theta : \theta \in \Theta\}$ is called **complete** if for any function $h : \mathcal{X} \to \mathcal{R}$,

$$\mathrm{E}_\theta\{h(\mathbf{X})\} = 0 \quad \text{for all } \theta \in \Theta$$

implies

$$\mathrm{P}_\theta\{h(\mathbf{X}) = 0\} = 1 \quad \text{for all } \theta \in \Theta.$$

A statistic $T \sim \mathrm{P}_\theta^T$ is called complete if the statistical model $\{\mathrm{P}_\theta^T : \theta \in \Theta\}$ is complete.

Example 2:
Show that the Binomial model with parameters $n = 2$ and p is complete.

Lehmann-Sheffé

Corollary (4.1)

Assume that P_θ belongs to a strictly k-parameter exponential family. Then the statistic $T(\mathbf{X})$ is sufficient and complete.

Theorem (4.7)

(Lehmann-Scheffé)
Let T be a sufficient and complete statistic for the statistical model $\mathcal{P}$, and let $\tilde{\gamma}_1$ be an unbiased estimator for the parameter $\gamma = g(\theta) \in \mathcal{R}^k$. Then the estimator

$$\hat{\gamma}(T) = \mathrm{E}_\theta(\tilde{\gamma}_1 | T)$$

has the smallest covariance matrix among all unbiased estimators for γ. That is, for all estimators $\tilde{\gamma}$ with $\mathrm{E}_\theta(\tilde{\gamma}) = \gamma$ we have

$$\mathrm{Cov}_\theta(\hat{\gamma}) \preceq \mathrm{Cov}_\theta(\tilde{\gamma}), \quad \textit{for all } \theta \in \Theta.$$

Lehmann-Sheffé

Example 3:

- Suppose that we observe $\mathbf{X} = (X_1, ..., X_n)$ where the X_i are independent with the same distribution as X.
- Show that the following statistics are BUE.

1. $\bar{X}$ where X is Bernoulli with parameter p.
2. $(\bar{X}, S^2)$ where $X \sim N(\mu, \sigma^2)$ (and note that S^2 does not attain the Cramér-Rao lower bound!).

News of today

- The multidimensional case
  - Unbiased estimator
  - Smaller covariance matrix
  - Score vector
  - Fisher information matrix
  - Cramér-Rao
- If there is an efficient unbiased estimator, it is the MLE.
- Rao-Blackwell: conditioning on a sufficient statistic improves an unbiased estimator.
- Completeness (always holds for members of the exponential family)
- Lehmann-Sheffé: under completeness, Rao-Blackwell gives the BUE.