# STAT 3100 Lecture Note

Week Five (Oct 4 & 6, 2022)

Zhiyang Zhou (zhiyang.zhou@umanitoba.ca, zhiyanggeezhou.github.io)

2022/Oct/06 14:13:00

---

## Evaluating estimators

### Mean squared error (MSE)

- Univariate: $\mathrm{E}(\hat{\theta}-\theta)^2 = \{\mathrm{E}(\hat{\theta})-\theta\}^2 + \mathrm{var}(\hat{\theta})$
- Multivariate: $\mathrm{E}\{(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})^\top(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\} = \{\mathrm{E}(\hat{\boldsymbol{\theta}})-\boldsymbol{\theta}\}^\top\{\mathrm{E}(\hat{\boldsymbol{\theta}})-\boldsymbol{\theta}\} + \mathrm{cov}(\hat{\boldsymbol{\theta}})$
- Best unbiased estimator (i.e., (uniform) minimun variance unbiased estimator, abbr. UMVUE/MVUE): if $\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}}(\mathbf{X})$ satisfies that
  - $\hat{\boldsymbol{\theta}}$ is unbiased for $\boldsymbol{\theta}$, i.e., $\mathrm{E}(\hat{\boldsymbol{\theta}}) = \boldsymbol{\theta}$
  - $\mathrm{var}(\hat{\boldsymbol{\theta}}) - \mathrm{var}(\hat{\boldsymbol{\theta}}^*) \le 0$ for all unbiased $\hat{\boldsymbol{\theta}}^*$
- UMVUE is unique (CB Thm 7.3.19)

### Cramer-Rao lower bound (CB Thm 7.3.9 & Lemma 7.3.11)

- Only consider the univariate case, i.e., one-dimensional unknown parameter $\theta$
  - Score function $S(\theta;\mathbf{X})$ and Hessian $H(\theta;\mathbf{X})$ both scalar
- Cramer-Rao lower bound: $\mathrm{var}(\hat{\theta}) \ge \{(\mathrm{d}/\mathrm{d}\theta)\mathrm{E}(\hat{\theta})\}^2/I(\theta)$ for $\hat{\boldsymbol{\theta}}$ satisfying regularity conditions
  - Fisher information: $I(\theta) = \mathrm{var}(S(\theta;\mathbf{X})) = \mathrm{E}[\{S(\theta;\mathbf{X})\}^2] = -\mathrm{E}\{H(\theta;\mathbf{X})\}$
  - Proof: Applying the Cauchy-Schwarz inequality (CB Thm 4.7.3)
- (CB Coro 7.3.15) $\hat{\theta}$ attains the lower bound $\Leftrightarrow$ there is $a(\theta)$ such that $S(\theta;\mathbf{X}) = a(\theta)\{\hat{\theta} - \mathrm{E}(\hat{\theta})\}$
- The unbiased $\hat{\theta}$ attaining the lower bound is UMVUE

### Example Lec8.1

- Find the lower bound for unbiased estimators for $\sigma^2$ in the following cases.
  a. iid $X_1,\ldots,X_n \sim \mathcal{N}(\mu_0,\sigma^2)$ with known $\mu_0$ and unknown $\sigma^2$.

### Sufficiency (CB Sec 6.2.1)

- A statistic $\mathbf{T} = \mathbf{T}(\mathbf{X})$ is sufficient for $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_p)$ $\Leftrightarrow$ the distribution of $\mathbf{X}$ conditioning on $\mathbf{T}$ and $\boldsymbol{\theta}$, say $f_{\mathbf{X}\mid\mathbf{T},\boldsymbol{\theta}}(\boldsymbol{x}\mid\boldsymbol{t},\boldsymbol{\theta})$, is free of $\boldsymbol{\theta}$.

- Fisher-Neyman factorization theorem (CB Thm 6.2.6; HMC Thm 7.2.1): $\mathbf{T}$ is sufficient for $\boldsymbol{\theta} \Leftrightarrow$ the likelihood function can be factored into two parts, one of them not depending on $\boldsymbol{\theta}$, i.e.,

$$L(\boldsymbol{\theta}; \boldsymbol{x}) = f_{\mathbf{X}}(\boldsymbol{x} \mid \boldsymbol{\theta}) = h(\boldsymbol{x}) g(\mathbf{T}(\boldsymbol{x}), \boldsymbol{\theta}), \text{for all } \boldsymbol{x} \text{ and } \boldsymbol{\theta}$$

- (HMC Thm 7.3.2) If $\mathbf{T}$ is sufficient for $\theta$ and $\hat{\theta}$ is the unique MLE of $\theta$, then $\hat{\theta}$ must be a function of $\mathbf{T}$.

---

- Nonuniqueness of sufficient statistics
  - Trivial examples
    - $\mathbf{X}$ is always sufficient.
    - $(X_{(1)}, \ldots, X_{(n)})$ is always sufficient if $X_i$'s are iid, with $X_{(1)} \leq \cdots \leq X_{(n)}$.
  - $\mathbf{T}$ is sufficient and $g(\cdot)$ is a one-to-one mapping $\Rightarrow g(\mathbf{T})$ is also sufficient.
- Minimal sufficiency: a sufficient statistc that is a function of all the other sufficient statistics.
  - (CB Thm 6.2.13) How to find a minimal sufficient sufficient statistic:
    1. Find the sufficient and necessary condition for $L(\boldsymbol{\theta}; \boldsymbol{x})/L(\boldsymbol{\theta}; \boldsymbol{y})$ to be free of $\boldsymbol{\theta}$;
    2. If the above condition is of the form $\mathbf{T}(\boldsymbol{x}) = \mathbf{T}(\boldsymbol{y})$, then $\mathbf{T}(\mathbf{X})$ is a minimal sufficient statistic for $\boldsymbol{\theta}$.

## Example Lec8.2

- Find the minimal sufficient statistics in the following scenarios.
  a. iid $X_1, \ldots, X_n \sim \text{Unif}\{1, \ldots, \theta\}$ with unknown positive integer $\theta$.
  b. iid $X_1, \ldots, X_n \sim \mathcal{N}(\mu, \sigma^2)$ with unknown $\mu$ and $\sigma^2$.

## Rao-Blackwellization (CB Thm 7.3.17)

- Only consider one-dimensional cases

- Improve the variance of statistic $W$, an estimator of $\theta$: take use of $\mathrm{E}(W \mid T)$ (a function of $T$ only) instead with sufficient $T$

- $T$ sufficient for $\theta \Rightarrow \mathrm{E}(W \mid T, \theta) = \mathrm{E}(W \mid T) \Rightarrow \mathrm{var}\{\mathrm{E}(W \mid T, \theta) \mid \theta\} = \mathrm{var}\{\mathrm{E}(W \mid T) \mid \theta\} \leq \mathrm{var}(W \mid \theta)$ for all $\theta \in \Theta$

  - No impact on the bias
  - Not working if $W$ is already a function of $T$

## Example Lec9.1

- Improve statistic $W$ in terms of variance.
  a. $W = X_1$, where iid $X_1, X_2 \sim \mathcal{N}(\theta, 1)$ with unknown $\theta$.
  b. $W = 2X_1 - X_2$, where iid $X_1, X_2 \sim f(x \mid \theta) = \theta^{-1} \exp(-x\theta^{-1}) \mathbf{1}_{\mathbb{R}^+ \times \mathbb{R}^+}(x, \theta)$ with unknown $\theta$.