Statistical Data Analysis

Dr. Jana de Wiljes

22. November 2021

Universität Potsdam

Unbiased estimators

Estimator

Def:

- An estimator is an arbitrary (Borel-measurable) function

$$\widehat{\theta} : \mathcal{X} \to \Theta, \quad x \mapsto \widehat{\theta}(x) \tag{1}$$

- An estimator $\widehat{\theta}$ is called unbiased, if

$$\mathbb{E}_\theta[\widehat{\theta}(X)] = \theta \tag{2}$$

  for all $\theta \in \Theta$.
- The bias of an estimator $\widehat{\theta}$ is

$$\mathrm{Bias}_\theta(\widehat{\theta}) = \mathbb{E}_\theta[\widehat{\theta}(X)] - \theta \tag{3}$$

Note: $\mathrm{Bias}_\theta(\widehat{\theta})$ is a function in $\widehat{\theta}$

Example

Example

Example

Mean square error

Def: Let $\Theta = (a, b) \subset \mathbb{R}$ be an interval. The mean square error (MSE) of an estimator $\widehat{\theta} : \mathcal{X} \to \Theta$

$$\mathsf{MSE}_\theta(\widehat{\theta}) = \mathbb{E}_\theta[(\widehat{\theta}(X) - \theta)^2] \tag{4}$$

Mean square error

Lemma: The relationship between the mean square error (MSE) of an estimator $\widehat{\theta} : \mathcal{X} \to \Theta$ and the BIAS is given by

$$\mathsf{MSE}_\theta(\widehat{\theta}) = \mathsf{Var}_\theta \widehat{\theta} + (\mathsf{Bias}_\theta(\widehat{\theta}))^2 \tag{5}$$

Proof

Proof

Consistently better

Def: Let $\widehat{\Theta}_1$ and $\widehat{\Theta}_2$ be two estimators. The estimator θ_1 is called consistently better than θ_2 if,

$$MSE_\theta(\widehat{\theta}_1) \leq MSE_\theta(\widehat{\theta}_2) \quad \forall\theta \in \Theta \tag{6}$$

Minimum-variance unbiased estimator

Def: An unbiased estimator $\widehat{\theta}$ is called minimum-variance unbiased estimator if all unbiased estimators $\tilde{\theta}$ the following inequality holds

$$\mathsf{Var}_\theta \widehat{\theta} \leq \mathsf{Var}_\theta \tilde{\theta} \tag{7}$$

for all $\theta \in \Theta$.

Minimum-variance unbiased estimator

Lemma: Let $\widehat{\theta}_1, \widehat{\theta}_2 : \mathcal{X} \to \Theta$ are two minimum-variance unbiased estimator the

$$\widehat{\theta}_1 = \widehat{\theta}_2 \quad \text{almost surely under } \mathbb{P} \text{ for all } \theta \in \Theta \tag{8}$$

for all $\theta \in \Theta$.

Bernoulli Experiment MVUE

Lemma: The estimator $\widehat{\theta}(x_1, \ldots, x_n) = \bar{x}_n$ is the minimum-variance unbiased estimator of θ in n Bernoulli experiments.

Proof

Proof

Proof

Sufficient statistic

Def: A function $T : \mathcal{X} \to \mathbb{R}^r$ is called a sufficient statistic if the function

$$\theta \mapsto \mathbb{P}_\theta[X = x | T(X) = t] \tag{9}$$

is constant for all $x \in \mathcal{X}$ and for all $t \in \mathbb{R}^r$,i.e.,

$$\mathbb{P}_{\theta_1}[X = x | T(X) = t] = \mathbb{P}_{\theta_2}[X = x | T(X) = t] \tag{10}$$

for all $t \in \mathbb{R}^r$ and all $\theta_1, \theta_2 \in \Theta$ with $\mathbb{P}_{\theta_1}[T(X) = t] \neq 0$ and $\mathsf{P}_{\theta_2}[T(X) = t] \neq 0$

Sufficient statistic

Lemma: Let $T_{\mathcal{X}} \to \mathbb{R}^r$ be a sufficient statistic and let $g : Im(T) \to \mathbb{R}^k$ an injective function. Then the concatenation

$$g \circ T : \mathcal{X} \to \mathbb{R}^k, \quad x \mapsto g(T(x)) \tag{11}$$

a sufficient statistic.

Example

Sufficient statistic

Lemma: Let $T_{\mathcal{X}} \to \mathbb{R}^r$ be a sufficient statistic and let $g : Im(T) \to \mathbb{R}^k$ an injective function. Then the concatenation

$$g \circ T : \mathcal{X} \to \mathbb{R}^k, \quad x \mapsto g(T(x)) \tag{12}$$

a sufficient statistic.

Rao-Blackwell

Proposition: Let $(\mathbb{P}_\theta)_{\theta \in \Theta}$ a family of probability measures on the sample space $(\mathcal{X}, \mathcal{A})$, where $\Theta \subset \mathbb{R}$ is an interval. Furthermore let

- $T : \mathcal{X} \to \mathbb{R}^m$ a sufficient statistic and
- $\hat{\theta} : \mathcal{X} \to \mathcal{R}$ an unbiased estimator of θ with $\mathbb{E}_\theta[\theta^2] \leq \infty$ for all $\theta \in \Theta$.

Define $\tilde{\theta} := \mathbb{E}_\theta[\hat{\theta}|T]$. Then $\tilde{\theta}$ is an unbiased estimator of θ and the following holds

$$Var_\theta \tilde{\theta} \leq Var_\theta \hat{\theta} \tag{13}$$

for all $\theta \in \Theta$.