AUA CS 108, Statistics, Fall 2019

Lecture 23

Michael Poghosyan
YSU, AUA
michael@ysu.am, mpoghosyan@aua.am

16 Oct 2019

Contents

Last Lecture ReCap

- State the Problem of the Point Estimation.

Last Lecture ReCap

- State the Problem of the Point Estimation.
- Define the Bias of an Estimator.

Last Lecture ReCap

- State the Problem of the Point Estimation.
- Define the Bias of an Estimator.
- What is the definition of the Unbiased/Biased Estimator?

Last Lecture ReCap

- State the Problem of the Point Estimation.
- Define the Bias of an Estimator.
- What is the definition of the Unbiased/Biased Estimator?
- What is the definition of the Asymptotically Unbiased Estimator?

Last Lecture ReCap

- State the Problem of the Point Estimation.
- Define the Bias of an Estimator.
- What is the definition of the Unbiased/Biased Estimator?
- What is the definition of the Asymptotically Unbiased Estimator?
- Give the Bias-Variance Decomposition

Example

Example: Assume we have a Random Sample for a some Distribution with the Mean μ and Variance σ^2:

$$X_1, X_2, ..., X_n \sim \mathcal{F}_{\mu,\sigma^2},$$

and we want to estimate the Parameters μ and σ^2.

Example

Example: Assume we have a Random Sample for a some Distribution with the Mean μ and Variance σ^2:

$$X_1, X_2, ..., X_n \sim \mathcal{F}_{\mu,\sigma^2},$$

and we want to estimate the Parameters μ and σ^2.

We consider the following Estimators:

$$\hat{\mu} = \overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n}$$

and

$$\widehat{\sigma^2} = \frac{\sum_{k=1}^{n}(X_k - \overline{X}_n)^2}{n} \qquad \text{and} \qquad \widehat{\sigma^2} = S^2 = \frac{\sum_{k=1}^{n}(X_k - \overline{X}_n)^2}{n-1}$$

Example

Example: Assume we have a Random Sample for a some Distribution with the Mean μ and Variance σ^2:

$$X_1, X_2, ..., X_n \sim \mathcal{F}_{\mu,\sigma^2},$$

and we want to estimate the Parameters μ and σ^2.

We consider the following Estimators:

$$\hat{\mu} = \overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n}$$

and

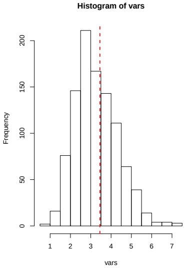
$$\widehat{\sigma^2} = \frac{\sum\limits_{k=1}^{n}(X_k - \overline{X}_n)^2}{n} \qquad \text{and} \qquad \widehat{\sigma^2} = S^2 = \frac{\sum\limits_{k=1}^{n}(X_k - \overline{X}_n)^2}{n-1}$$

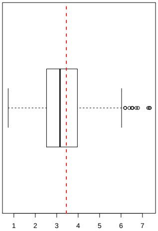
Let us see (OTB) which ones are Biased and which ones are not.

Example

- Biased Case, with *n* in the Denominator:

```
v <- 3.45; n <- 20; it <- 1000
x <- rnorm(n*it, mean = 2, sd = sqrt(v))
x <- as.data.frame(matrix(x, ncol = it))
my.var <- function(x){return((length(x)-1)*var(x)/length(x))}
vars <- sapply(x, my.var)
par(mfrow = c(1,2))
hist(vars, breaks = 15)
abline(v = v, col = "red", lty = 2, lwd = 2)
boxplot(vars, horizontal = T)
abline(v = v, col = "red", lty = 2, lwd = 2)
```





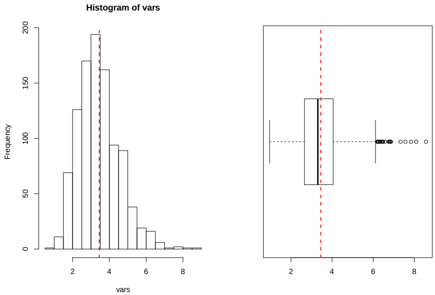

```
par(mfrow = c(1,1))
mean(vars) - v
```

```
## [1] -0.1696762
```

Example

- UnBiased Case, with $n - 1$ in the Denominator:

```
v <- 3.45; n <- 20; it <- 1000
x <- rnorm(n*it, mean = 2, sd = sqrt(v))
x <- as.data.frame(matrix(x, ncol = it))
vars <- sapply(x,var)
par(mfrow = c(1,2))
hist(vars, breaks = 15)
abline(v = v, col = "red", lty = 2, lwd = 2)
boxplot(vars, horizontal = T)
abline(v = v, col = "red", lty = 2, lwd = 2)
```




```
par(mfrow = c(1,1))
mean(vars) - v
```

```
## [1] -0.01123591
```

Bias-Variance Decomposition

Again, let's recall our BVD:

$$MSE(\hat{\theta}, \theta) = \left(Bias(\hat{\theta}, \theta)\right)^2 + Var_\theta(\hat{\theta}).$$

Bias-Variance Decomposition

Again, let's recall our BVD:

$$MSE(\hat{\theta}, \theta) = \left(Bias(\hat{\theta}, \theta)\right)^2 + Var_\theta(\hat{\theta}).$$

We interpret the RHS:

- Bias is the **Accuracy** of our Estimator $\hat{\theta}$

Bias-Variance Decomposition

Again, let's recall our BVD:

$$MSE(\hat{\theta}, \theta) = \Big(Bias(\hat{\theta}, \theta)\Big)^2 + Var_\theta(\hat{\theta}).$$

We interpret the RHS:

- Bias is the **Accuracy** of our Estimator $\hat{\theta}$
- Variance is the **Precision** of our Estimator $\hat{\theta}$

Bias-Variance Decomposition

Again, let's recall our BVD:

$$MSE(\hat{\theta}, \theta) = \Big(Bias(\hat{\theta}, \theta)\Big)^2 + Var_\theta(\hat{\theta}).$$

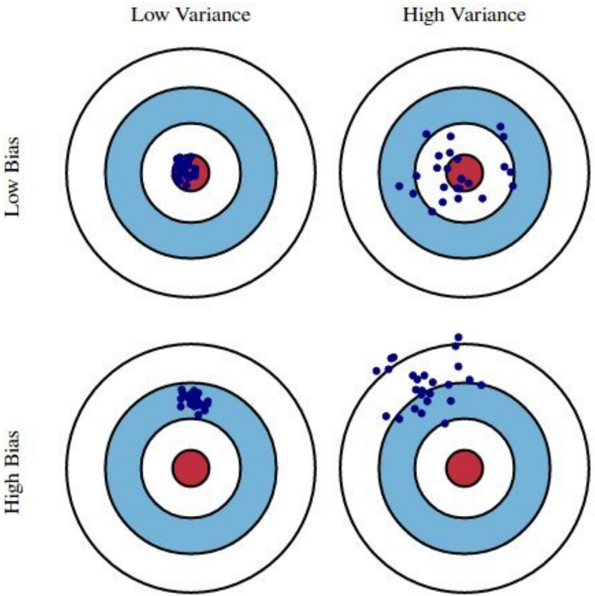
We interpret the RHS:

- Bias is the **Accuracy** of our Estimator $\hat{\theta}$
- Variance is the **Precision** of our Estimator $\hat{\theta}$

Nice Graphical Interpretation: Link, see also the next slide.

Bias-Variance Decomposition/Tradeoff

Bias-Variance Decomposition/Tradeoff



Standard Error and Estimated Standard Error

Definition: The Standard Deviation of the Estimator is called the **Standard Error** of the Estimator $\hat{\theta}$ and is denoted by

$$SE(\hat{\theta}) = SD(\hat{\theta}) = \sqrt{Var_{\theta}(\hat{\theta})}.$$

Standard Error and Estimated Standard Error

Definition: The Standard Deviation of the Estimator is called the **Standard Error** of the Estimator $\hat{\theta}$ and is denoted by

$$SE(\hat{\theta}) = SD(\hat{\theta}) = \sqrt{Var_\theta(\hat{\theta})}.$$

Usually, the Standard Error will depend on the unknown value of the Parameter θ. If we use the Estimator $\hat{\theta}$, then the **Estimated Standard Error** of $\hat{\theta}$, $\widehat{SE}(\hat{\theta})$ is the Standard Error, where after calculation we plug $\hat{\theta}$ instead of θ.

Standard Error and Estimated Standard Error

Definition: The Standard Deviation of the Estimator is called the **Standard Error** of the Estimator $\hat{\theta}$ and is denoted by

$$SE(\hat{\theta}) = SD(\hat{\theta}) = \sqrt{Var_{\theta}(\hat{\theta})}.$$

Usually, the Standard Error will depend on the unknown value of the Parameter θ. If we use the Estimator $\hat{\theta}$, then the **Estimated Standard Error** of $\hat{\theta}$, $\widehat{SE}(\hat{\theta})$ is the Standard Error, where after calculation we plug $\hat{\theta}$ instead of θ.

And statisticians, when reporting the Estimate, usually report also the Estimated Standard Error, as a measure how precise is the result. If the Standard Error is small (and we are using a nice Estimator, say, it is Unbiased), then this is a sign that the result is close to real/actual one.

Example

Example: Assume we are facing an election with Parties A and B, and we want to estimate the percentage of voters for A in advance. So we do a poll, asking 10 persons to give their preferences. Let the result be:

$$A, B, B, B, A, B, B, A, B, B.$$

Problem: Estimate the percentage of voters for the Party A, and give the Estimated Standard Error.

Solution: OTB.