YSU ASDS, Statistics, Fall 2019

Lecture 16

Michael Poghosyan

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 MSE.

Last Lecture ReCap

- State the Problem of the Point Estimation.
- Define the MSE.
- How we can define that $\hat{\theta}$ is close to θ?

Last Lecture ReCap

- State the Problem of the Point Estimation.
- Define the MSE.
- How we can define that $\hat{\theta}$ is close to θ?
- Define the Bias of an Estimator.

Last Lecture ReCap

- State the Problem of the Point Estimation.
- Define the MSE.
- How we can define that $\hat{\theta}$ is close to θ?
- Define the Bias of an Estimator.
- What is the definition of the Unbiased/Biased Estimator?

Example

Example: Now, let's do an experiment with Biased and Unbiased Estimators in **R**.

Example

Example: Now, let's do an experiment with Biased and Unbiased Estimators in **R**.

UnBiased Estimator Case

We consider the Poisson Model:

$$X_1, X_2, ..., X_{10} \sim Pois(\lambda)$$

and we want to estimate λ.

Example

Example: Now, let's do an experiment with Biased and Unbiased Estimators in **R**.

UnBiased Estimator Case

We consider the Poisson Model:

$$X_1, X_2, ..., X_{10} \sim Pois(\lambda)$$

and we want to estimate λ. We consider the following Estimator:

$$\hat{\lambda} = \frac{X_1 + X_2 + ... + X_{10}}{10}.$$

Example

Example: Now, let's do an experiment with Biased and Unbiased Estimators in **R**.

UnBiased Estimator Case

We consider the Poisson Model:

$$X_1, X_2, ..., X_{10} \sim Pois(\lambda)$$

and we want to estimate λ. We consider the following Estimator:

$$\hat{\lambda} = \frac{X_1 + X_2 + ... + X_{10}}{10}.$$

Easy to se that $\hat{\lambda}$ is an Unbiased Estimator for λ (OTB!).

Example, cont'd

Now, the code

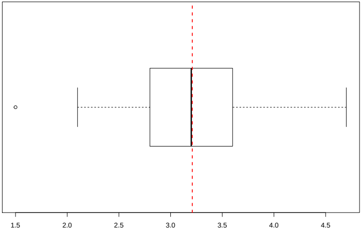
- observing once: ganarating a Sample just once and calculating one Estimate:

```r
lambda <- 3.21
x <- rpois(10, lambda = lambda)
lambda.hat <- mean(x)
lambda.hat
```

```
## [1] 2.4
```

- observing many times: ganarating Samples many times, calculating Estimates, and then averaging:

```
lambda <- 3.21; n <- 10; m <- 200
x <- rpois(n*m, lambda = lambda)
x <- as.data.frame(matrix(x, ncol = m))
lambda.hats <- sapply(x, mean)
boxplot(lambda.hats, horizontal = T);
abline(v = lambda, col="red", lwd = 2, lty = 2)
```

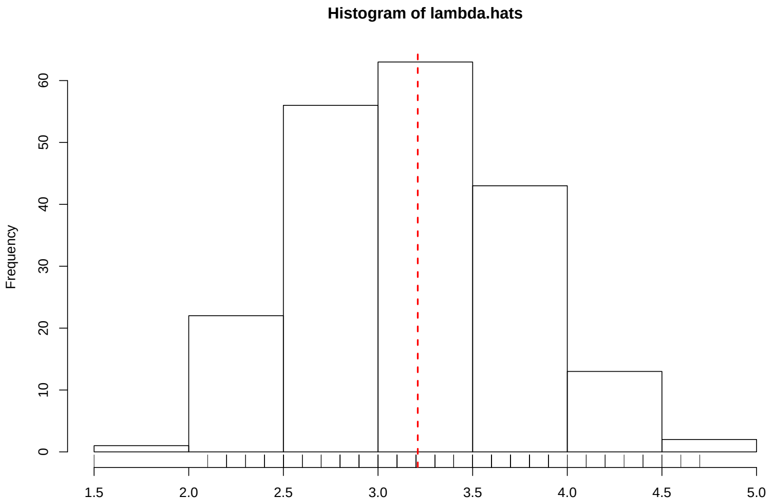



```
mean(lambda.hats)
```

```
## [1] 3.227
```

With a Histogram:

```r
hist(lambda.hats)
rug(lambda.hats)
abline(v = lambda, col="red", lwd = 2, lty = 2)
```



Example

Biased Estimator Case

Say, let us consider the Exponential Model:

$$X_1, X_2, ..., X_{10} \sim Exp(\lambda)$$

and we want to estimate λ.

Example

Biased Estimator Case

Say, let us consider the Exponential Model:

$$X_1, X_2, ..., X_{10} \sim Exp(\lambda)$$

and we want to estimate λ. We consider the following Estimator:

$$\hat{\lambda} = \frac{X_1 + X_2 + ... + X_{10}}{10}.$$

Example

Biased Estimator Case

Say, let us consider the Exponential Model:

$$X_1, X_2, ..., X_{10} \sim Exp(\lambda)$$

and we want to estimate λ. We consider the following Estimator:

$$\hat{\lambda} = \frac{X_1 + X_2 + ... + X_{10}}{10}.$$

Easy to se that $\hat{\lambda}$ is an Biased Estimator for λ (OTB!).

Example, cont'd

Now, the code:

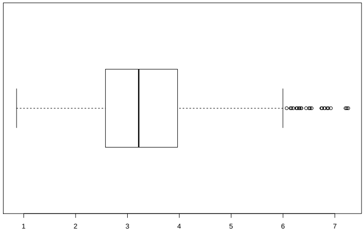
- observing once: ganarating a Sample just once and calculating one Estimate:

```
lambda <- 0.3
x <- rexp(10, rate = lambda)
lambda.hat <- mean(x)
lambda.hat
```

```
## [1] 3.059252
```

- observing many times: ganarating Samples many times, calculating Estimates, and then averaging:

```
lambda <- 0.3; n <- 10; m <- 2000
x <- rexp(n*m, rate = lambda)
x <- as.data.frame(matrix(x, ncol = m))
lambda.hats <- sapply(x, mean)
boxplot(lambda.hats, horizontal = T);
abline(v = lambda, col="red", lwd = 2, lty = 2)
```

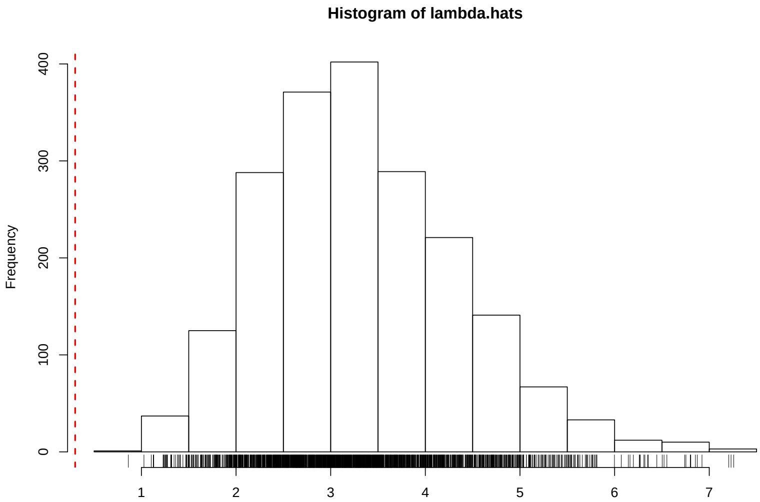



```
mean(lambda.hats)
```

```
## [1] 3.319164
```

With a Histogram:

```
hist(lambda.hats)
rug(lambda.hats)
abline(v = lambda, col="red", lwd = 2, lty = 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.

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

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

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

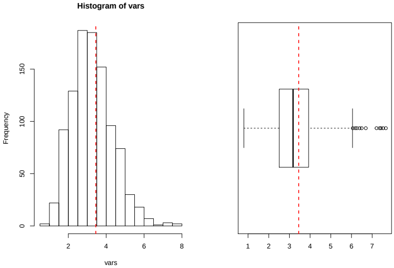
$$\widehat{\sigma^2} = \frac{\sum_{k=1}^{n}(X_k - \overline{X}_n)^2}{n} \qquad \text{and} \qquad \widehat{\sigma^2} = S^2 = \frac{\sum_{k=1}^{n}(X_k - \overline{X}_n)^2}{n-1}$$

Let us see (OTB) which ones are Biased and which ones are not, and calculate the Biases.

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)
```



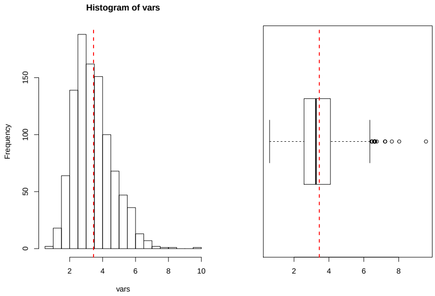

```
par(mfrow = c(1,1))
mean(vars) - v
```

```
## [1] -0.1747914
```

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] -0.02689461
```

Asymptotic Unbiasedness

Some years ago Unbiasedness was a very important, desirable property from any good Estimator.

Asymptotic Unbiasedness

Some years ago Unbiasedness was a very important, desirable property from any good Estimator. Nowadays, it is enough to have *Asymptotic Unbiasedness*:

Asymptotic Unbiasedness

Some years ago Unbiasedness was a very important, desirable property from any good Estimator. Nowadays, it is enough to have *Asymptotic Unbiasedness*:

Definition: Estimator $\hat{\theta}_n$ is called **Asymptotically Unbiased** for θ, if

$$Bias(\hat{\theta}_n, \theta) \to 0, \qquad \text{for any } \theta \in \Theta.$$

Asymptotic Unbiasedness

Some years ago Unbiasedness was a very important, desirable property from any good Estimator. Nowadays, it is enough to have *Asymptotic Unbiasedness*:

Definition: Estimator $\hat{\theta}_n$ is called **Asymptotically Unbiased** for θ, if

$$Bias(\hat{\theta}_n, \theta) \to 0, \qquad \text{for any } \theta \in \Theta.$$

Idea: If the Sample size is very large, then the behaviour of our Asymptotic Unbiased Estimator is close to an Unbiased one, $Bias(\hat{\theta}_n, \theta) \approx 0$

Asymptotic Unbiasedness

Some years ago Unbiasedness was a very important, desirable property from any good Estimator. Nowadays, it is enough to have *Asymptotic Unbiasedness*:

Definition: Estimator $\hat{\theta}_n$ is called **Asymptotically Unbiased** for θ, if

$$Bias(\hat{\theta}_n, \theta) \to 0, \qquad \text{for any } \theta \in \Theta.$$

Idea: If the Sample size is very large, then the behaviour of our Asymptotic Unbiased Estimator is close to an Unbiased one, $Bias(\hat{\theta}_n, \theta) \approx 0$

Example: Say, for the Mean μ of the Population,

$$\hat{\mu} = \frac{X_1 + X_2 + ... + X_n}{n+1}$$

is a Biased, but Asymptotically Unbiased Estimator. OTB, please!

Example

Example: Another one.

Example

Example: Another one. As we have calculated above, the following estimate for σ^2,

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

is Biased.

Example

Example: Another one. As we have calculated above, the following estimate for σ^2,

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

is Biased. But the Bias is

$$Bias(\widehat{\sigma^2}, \sigma^2) = -\frac{\sigma^2}{n} \to 0, \qquad n \to \infty,$$

or, equivalently,

$$\mathbb{E}(\widehat{\sigma^2}) \to \sigma^2,$$

so $\widehat{\sigma^2}$ is an Asymptotically Unbiased Estimator for σ^2.

Bias-Variance Decomposition

This is an important result:

Theorem(Bias-Variance Decomposition of the MSE): If $\hat{\theta}$ is an Estimator for θ, then

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

Bias-Variance Decomposition

This is an important result:

Theorem(Bias-Variance Decomposition of the MSE): If $\hat{\theta}$ is an Estimator for θ, then

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

Proof: OTB

Bias-Variance Decomposition

Again, let's recall our BVD:

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

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:

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

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

Bias-Variance Decomposition/Tradeoff

Bias-Variance Decomposition/Tradeoff

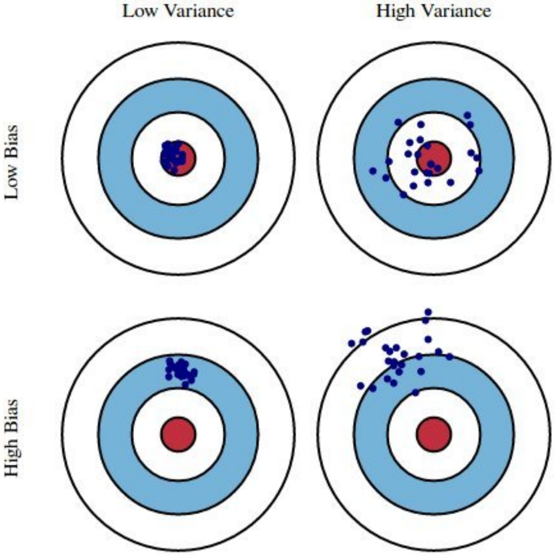


Figure 1: Credits to: http://scott.fortmann-roe.com

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.

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.

Bias-Variance Decomposition, Corollary

Recall from the last lecture the BVD:

Theorem(Bias-Variance Decomposition of the MSE): If $\hat{\theta}$ is an Estimator for θ, then

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

Bias-Variance Decomposition, Corollary

Recall from the last lecture the BVD:

Theorem(Bias-Variance Decomposition of the MSE): If $\hat{\theta}$ is an Estimator for θ, then

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

Corollary: If $\hat{\theta}_1$ and $\hat{\theta}_2$ are both *Unbiased* Estimators for the unknown Parameter θ, then $\hat{\theta}_1$ is preferable to $\hat{\theta}_2$ if and only if

$$Var(\hat{\theta}_1) \leq Var(\hat{\theta}_2), \qquad \text{for all } \theta,$$

and a strict inequality holds for at least one θ.

Bias-Variance Decomposition, Corollary

Recall from the last lecture the BVD:

Theorem(Bias-Variance Decomposition of the MSE): If $\hat{\theta}$ is an Estimator for θ, then

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

Corollary: If $\hat{\theta}_1$ and $\hat{\theta}_2$ are both *Unbiased* Estimators for the unknown Parameter θ, then $\hat{\theta}_1$ is preferable to $\hat{\theta}_2$ if and only if

$$Var(\hat{\theta}_1) \leq Var(\hat{\theta}_2), \qquad \text{for all } \theta,$$

and a strict inequality holds for at least one θ.

Idea: If we consider Unbiased Estimators only, then we prefer that one, which has the least variability.

Bias-Variance Decomposition, Corollary

Recall from the last lecture the BVD:

Theorem(Bias-Variance Decomposition of the MSE): If $\hat{\theta}$ is an Estimator for θ, then

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

Corollary: If $\hat{\theta}_1$ and $\hat{\theta}_2$ are both *Unbiased* Estimators for the unknown Parameter θ, then $\hat{\theta}_1$ is preferable to $\hat{\theta}_2$ if and only if

$$\mathit{Var}(\hat{\theta}_1) \leq \mathit{Var}(\hat{\theta}_2), \qquad \text{for all } \theta,$$

and a strict inequality holds for at least one θ.

Idea: If we consider Unbiased Estimators only, then we prefer that one, which has the least variability. Unbiased means that the values of our Estimators are centered around the true value of the Parameter.

Bias-Variance Decomposition, Corollary

Recall from the last lecture the BVD:

Theorem(Bias-Variance Decomposition of the MSE): If $\hat{\theta}$ is an Estimator for θ, then

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

Corollary: If $\hat{\theta}_1$ and $\hat{\theta}_2$ are both *Unbiased* Estimators for the unknown Parameter θ, then $\hat{\theta}_1$ is preferable to $\hat{\theta}_2$ if and only if

$$\mathit{Var}(\hat{\theta}_1) \leq \mathit{Var}(\hat{\theta}_2), \qquad \text{for all } \theta,$$

and a strict inequality holds for at least one θ.

Idea: If we consider Unbiased Estimators only, then we prefer that one, which has the least variability. Unbiased means that the values of our Estimators are centered around the true value of the Parameter. Variance of the Estimator is measuring how concentrated are the values of our Estimator around the true value of the Parameter.

Bias-Variance Decomposition, Corollary

Recall from the last lecture the BVD:

Theorem(Bias-Variance Decomposition of the MSE): If $\hat{\theta}$ is an Estimator for θ, then

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

Corollary: If $\hat{\theta}_1$ and $\hat{\theta}_2$ are both *Unbiased* Estimators for the unknown Parameter θ, then $\hat{\theta}_1$ is preferable to $\hat{\theta}_2$ if and only if

$$\mathit{Var}(\hat{\theta}_1) \leq \mathit{Var}(\hat{\theta}_2), \qquad \text{for all } \theta,$$

and a strict inequality holds for at least one θ.

Idea: If we consider Unbiased Estimators only, then we prefer that one, which has the least variability. Unbiased means that the values of our Estimators are centered around the true value of the Parameter. Variance of the Estimator is measuring how concentrated are the values of our Estimator around the true value of the Parameter. If variability, Variance is small, it will give better results.

B-V Decomposition, Again

Recall again the B-V D:

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

B-V Decomposition, Again

Recall again the B-V D:

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

We have talked before that it would be nice to be able to minimize $MSE(\hat{\theta},\theta)$, but, unfortunately, this is impossible, in general.

B-V Decomposition, Again

Recall again the B-V D:

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

We have talked before that it would be nice to be able to minimize $MSE(\hat{\theta},\theta)$, but, unfortunately, this is impossible, in general. So Statisticians consider the following restricted Problem:

Find an Unbiased Estimator with the Minimal Variance.

B-V Decomposition, Again

Recall again the B-V D:

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

We have talked before that it would be nice to be able to minimize $MSE(\hat{\theta}, \theta)$, but, unfortunately, this is impossible, in general. So Statisticians consider the following restricted Problem:

Find an Unbiased Estimator with the Minimal Variance.

Well, in general, there will be a lot of Unbiased Estimators for the same Parameter.

B-V Decomposition, Again

Recall again the B-V D:

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

We have talked before that it would be nice to be able to minimize $MSE(\hat{\theta}, \theta)$, but, unfortunately, this is impossible, in general. So Statisticians consider the following restricted Problem:

Find an Unbiased Estimator with the Minimal Variance.

Well, in general, there will be a lot of Unbiased Estimators for the same Parameter.Say, if $\hat{\theta}_0$ and $\hat{\theta}_1$ are Unbiased Estimators of θ, then for any $\alpha \in [0, 1]$, the Estimator

$$\hat{\theta}_\alpha = \alpha \cdot \hat{\theta}_1 + (1 - \alpha) \cdot \hat{\theta}_0$$

will be an Unbiased Estimator too.

MVUE

So the idea is to restrict our attention to only Unbiased Estimators. In that case, since $Bias(\hat{\theta}, \theta) = 0$,

$$MSE(\hat{\theta}, \theta) = Var_{\theta}(\hat{\theta}).$$

MVUE

So the idea is to restrict our attention to only Unbiased Estimators. In that case, since $Bias(\hat{\theta}, \theta) = 0$,

$$MSE(\hat{\theta}, \theta) = Var_{\theta}(\hat{\theta}).$$

Then we need to find the Estimator with the smallest Variance, i.e., with the highest precision.

MVUE

So the idea is to restrict our attention to only Unbiased Estimators. In that case, since $Bias(\hat{\theta}, \theta) = 0$,

$$MSE(\hat{\theta}, \theta) = Var_{\theta}(\hat{\theta}).$$

Then we need to find the Estimator with the smallest Variance, i.e., with the highest precision.

And we give the following

Definition: Estimator $\hat{\theta}$ is called the **MVUE (Minimum Variance Unbiased Estimator)** for θ, if

MVUE

So the idea is to restrict our attention to only Unbiased Estimators. In that case, since $Bias(\hat{\theta}, \theta) = 0$,

$$MSE(\hat{\theta}, \theta) = Var_{\theta}(\hat{\theta}).$$

Then we need to find the Estimator with the smallest Variance, i.e., with the highest precision.

And we give the following

Definition: Estimator $\hat{\theta}$ is called the **MVUE (Minimum Variance Unbiased Estimator)** for θ, if

- $\hat{\theta}$ is Unbiased Estimator for θ;

MVUE

So the idea is to restrict our attention to only Unbiased Estimators. In that case, since $Bias(\hat{\theta}, \theta) = 0$,

$$MSE(\hat{\theta}, \theta) = Var_\theta(\hat{\theta}).$$

Then we need to find the Estimator with the smallest Variance, i.e., with the highest precision.

And we give the following

Definition: Estimator $\hat{\theta}$ is called the **MVUE (Minimum Variance Unbiased Estimator)** for θ, if

- $\hat{\theta}$ is Unbiased Estimator for θ;
- $\hat{\theta}$ has the smallest Variance among all *Unbiased* Estimators of θ, i.e., for any Unbiased Estimator $\tilde{\theta}$,

$$Var_\theta(\hat{\theta}) \leq Var_\theta(\tilde{\theta}), \qquad \forall \theta \in \Theta.$$

MVUE

So the idea is to restrict our attention to only Unbiased Estimators. In that case, since $Bias(\hat{\theta}, \theta) = 0$,

$$MSE(\hat{\theta}, \theta) = Var_{\theta}(\hat{\theta}).$$

Then we need to find the Estimator with the smallest Variance, i.e., with the highest precision.

And we give the following

Definition: Estimator $\hat{\theta}$ is called the **MVUE (Minimum Variance Unbiased Estimator)** for θ, if

- $\hat{\theta}$ is Unbiased Estimator for θ;
- $\hat{\theta}$ has the smallest Variance among all *Unbiased* Estimators of θ, i.e., for any Unbiased Estimator $\tilde{\theta}$,

$$Var_{\theta}(\hat{\theta}) \leq Var_{\theta}(\tilde{\theta}), \qquad \forall \theta \in \Theta.$$

Later we will talk about how to find MVUE for a parameter for some cases.