AUA CS108, Statistics, Fall 2020

Lecture 28

Michael Poghosyan

30 Oct 2020

Contents

Risk, Mean Squared Error of the Estimator

Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta,$$

and we use the Estimator $\hat{\theta}$ to estimate θ.

Risk, Mean Squared Error of the Estimator

Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta,$$

and we use the Estimator $\hat{\theta}$ to estimate θ.

Definition: The **Mean Squared Error** or the **Quadratic Risk** of the estimator $\hat{\theta}_n$ of θ is

$$MSE(\hat{\theta}_n, \theta) = Risk(\hat{\theta}_n, \theta) = \mathbb{E}_\theta[(\hat{\theta}_n - \theta)^2].$$

Risk, Mean Squared Error of the Estimator

Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta,$$

and we use the Estimator $\hat{\theta}$ to estimate θ.

Definition: The **Mean Squared Error** or the **Quadratic Risk** of the estimator $\hat{\theta}_n$ of θ is

$$MSE(\hat{\theta}_n, \theta) = Risk(\hat{\theta}_n, \theta) = \mathbb{E}_\theta[(\hat{\theta}_n - \theta)^2].$$

Note: *MSE* calculates how close are, in the Quadratic Mean sense, possible values of the Estimator $\hat{\theta}$ to the actual (unknown) value of θ.

Risk, Mean Squared Error of the Estimator

Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta,$$

and we use the Estimator $\hat{\theta}$ to estimate θ.

Definition: The **Mean Squared Error** or the **Quadratic Risk** of the estimator $\hat{\theta}_n$ of θ is

$$MSE(\hat{\theta}_n, \theta) = Risk(\hat{\theta}_n, \theta) = \mathbb{E}_\theta[(\hat{\theta}_n - \theta)^2].$$

Note: *MSE* calculates how close are, in the Quadratic Mean sense, possible values of the Estimator $\hat{\theta}$ to the actual (unknown) value of θ. The smaller the value of *MSE*, the better, of course.

Risk, Mean Squared Error of the Estimator

Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim \mathcal{F}_\theta, \qquad \theta \in \Theta,$$

and we use the Estimator $\hat{\theta}$ to estimate θ.

Definition: The **Mean Squared Error** or the **Quadratic Risk** of the estimator $\hat{\theta}_n$ of θ is

$$MSE(\hat{\theta}_n, \theta) = Risk(\hat{\theta}_n, \theta) = \mathbb{E}_\theta[(\hat{\theta}_n - \theta)^2].$$

Note: *MSE* calculates how close are, in the Quadratic Mean sense, possible values of the Estimator $\hat{\theta}$ to the actual (unknown) value of θ. The smaller the value of *MSE*, the better, of course.

Definition: We say that the estimator $\hat{\theta}_n^1$ of θ is **preferable** to $\hat{\theta}_n^2$, another estimator of θ, if

$$MSE(\hat{\theta}_n^1, \theta) \leq MSE(\hat{\theta}_n^2, \theta), \qquad \forall \theta \in \Theta,$$

and there exists a θ s.t. $MSE(\hat{\theta}_n^1, \theta) < MSE(\hat{\theta}_n^2, \theta)$.

Useful Identity

A very useful identity for calculating $\mathbb{E}(X^2)$ for a r.v. X is the following:

Useful Identity

A very useful identity for calculating $\mathbb{E}(X^2)$ for a r.v. X is the following: because

$$Var(X) = \mathbb{E}(X^2) - \Big[\mathbb{E}(X)\Big]^2,$$

then

Useful Identity

A very useful identity for calculating $\mathbb{E}(X^2)$ for a r.v. X is the following: because

$$Var(X) = \mathbb{E}(X^2) - \left[\mathbb{E}(X)\right]^2,$$

then

$$\mathbb{E}(X^2) = Var(X) + \left[\mathbb{E}(X)\right]^2.$$

Example

Example: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim Pois(\lambda),$$

and we want to Estimate the unknown λ.

Example

Example: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim Pois(\lambda),$$

and we want to Estimate the unknown λ. We want to use one of the following Estimators:

$$\hat{\lambda}_1 = X_1 \qquad or \qquad \hat{\lambda}_2 = \frac{X_1 + X_n}{2}.$$

Example

Example: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim Pois(\lambda),$$

and we want to Estimate the unknown λ. We want to use one of the following Estimators:

$$\hat{\lambda}_1 = X_1 \qquad or \qquad \hat{\lambda}_2 = \frac{X_1 + X_n}{2}.$$

Question: Calculate the Risks of $\hat{\lambda}_1$ and $\hat{\lambda}_2$. Compare these two Estimators: which of $\hat{\lambda}_1$ and $\hat{\lambda}_2$ is preferable?

Example

Example: Assume we have a Random Sample

$$X_1, X_2, ..., X_n \sim Pois(\lambda),$$

and we want to Estimate the unknown λ. We want to use one of the following Estimators:

$$\hat{\lambda}_1 = X_1 \qquad or \qquad \hat{\lambda}_2 = \frac{X_1 + X_n}{2}.$$

Question: Calculate the Risks of $\hat{\lambda}_1$ and $\hat{\lambda}_2$. Compare these two Estimators: which of $\hat{\lambda}_1$ and $\hat{\lambda}_2$ is preferable?

Solution: OTB: Use the fact that for $X \sim Pois(\lambda)$,

$$\mathbb{E}(X) = \lambda = Var(X).$$

Best MSE Estimators

Now, having the idea of the Error of estimation for an Estimator, we can try to find the best Estimator in the sense of MSE, i.e., we can try to solve

Find $\hat{\theta}$ among all Estimators that minimizes $MSE(\hat{\theta}, \theta)$.

Best MSE Estimators

Now, having the idea of the Error of estimation for an Estimator, we can try to find the best Estimator in the sense of MSE, i.e., we can try to solve

Find $\hat{\theta}$ among all Estimators that minimizes $MSE(\hat{\theta}, \theta)$.

Unfortunately, this Minimization problem has no solution in general.

Best MSE Estimators

Now, having the idea of the Error of estimation for an Estimator, we can try to find the best Estimator in the sense of MSE, i.e., we can try to solve

Find $\hat{\theta}$ among all Estimators that minimizes $MSE(\hat{\theta}, \theta)$.

Unfortunately, this Minimization problem has no solution in general.

So we move forward considering other Properties of Estimators making them useful in the estimation process.

Bias, Biased and Unbiased Estimators

Definition: The **Bias** of Estimator $\hat{\theta}$ of θ is

$$\mathit{Bias}(\hat{\theta}, \theta) = \mathbb{E}_\theta(\hat{\theta} - \theta) = \mathbb{E}_\theta(\hat{\theta}) - \theta.$$

Bias, Biased and Unbiased Estimators

Definition: The **Bias** of Estimator $\hat{\theta}$ of θ is

$$Bias(\hat{\theta}, \theta) = \mathbb{E}_\theta(\hat{\theta} - \theta) = \mathbb{E}_\theta(\hat{\theta}) - \theta.$$

Definition: We say that our Estimator $\hat{\theta}$ (of the parameter θ) is **Unbiased**, if

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

Bias, Biased and Unbiased Estimators

Definition: The **Bias** of Estimator $\hat{\theta}$ of θ is

$$Bias(\hat{\theta}, \theta) = \mathbb{E}_\theta(\hat{\theta} - \theta) = \mathbb{E}_\theta(\hat{\theta}) - \theta.$$

Definition: We say that our Estimator $\hat{\theta}$ (of the parameter θ) is **Unbiased**, if

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

If $\hat{\theta}$ is not Unbiased, then we say that it is **Biased**.

Bias, Biased and Unbiased Estimators

Definition: The **Bias** of Estimator $\hat{\theta}$ of θ is

$$Bias(\hat{\theta}, \theta) = \mathbb{E}_\theta(\hat{\theta} - \theta) = \mathbb{E}_\theta(\hat{\theta}) - \theta.$$

Definition: We say that our Estimator $\hat{\theta}$ (of the parameter θ) is **Unbiased**, if

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

If $\hat{\theta}$ is not Unbiased, then we say that it is **Biased**.

In other words, the Estimator $\hat{\theta}$ is **Unbiased**, if

$$\mathbb{E}_\theta\left(\hat{\theta}\right) = \theta \qquad \text{for any} \qquad \theta \in \Theta.$$

Bias, Biased and Unbiased Estimators

Definition: The **Bias** of Estimator $\hat{\theta}$ of θ is

$$Bias(\hat{\theta}, \theta) = \mathbb{E}_\theta(\hat{\theta} - \theta) = \mathbb{E}_\theta(\hat{\theta}) - \theta.$$

Definition: We say that our Estimator $\hat{\theta}$ (of the parameter θ) is **Unbiased**, if

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

If $\hat{\theta}$ is not Unbiased, then we say that it is **Biased**.

In other words, the Estimator $\hat{\theta}$ is **Unbiased**, if

$$\mathbb{E}_\theta\left(\hat{\theta}\right) = \theta \qquad \text{for any} \qquad \theta \in \Theta.$$

Note: We require the above hold **for any parameter value** since, if, say, it is correct for *some* values of θ, then it can happen that the true value of our unknown θ is exactly that value, for whihc we do not have the equality.

Bias and Unbiasedness

- The idea of Unbiased Estimator is the following: if we will calculate Estimates many-many times using our Unbiased Estimator, and then average the obtained Estimates, we will obtain the (almost exact, exact when many-many$\rightarrow \infty$) value of our Parameter.

Bias and Unbiasedness

- The idea of Unbiased Estimator is the following: if we will calculate Estimates many-many times using our Unbiased Estimator, and then average the obtained Estimates, we will obtain the (almost exact, exact when many-many$\to \infty$) value of our Parameter.
- So we can say about an Unbiased Estimator as: **In average, it is Exact**

Bias and Unbiasedness

- The idea of Unbiased Estimator is the following: if we will calculate Estimates many-many times using our Unbiased Estimator, and then average the obtained Estimates, we will obtain the (almost exact, exact when many-many$\to \infty$) value of our Parameter.
- So we can say about an Unbiased Estimator as: **In average, it is Exact**
- Bias can be interpreted, in some sense, as the *accuracy* of the Estimator.

Example

Example: Assume we have a Sample of size n from some Population, and we want to estimate the Mean of that Population.

Example

Example: Assume we have a Sample of size *n* from some Population, and we want to estimate the Mean of that Population. Let $\mathcal{F}_{\mu}$ be the Distribution of our Population, and μ be the Mean (Expectation) of that Distribution.

Example

Example: Assume we have a Sample of size *n* from some Population, and we want to estimate the Mean of that Population. Let $\mathcal{F}_\mu$ be the Distribution of our Population, and μ be the Mean (Expectation) of that Distribution. So our unknown Parameter here is μ.

Example

Example: Assume we have a Sample of size n from some Population, and we want to estimate the Mean of that Population. Let $\mathcal{F}_\mu$ be the Distribution of our Population, and μ be the Mean (Expectation) of that Distribution. So our unknown Parameter here is μ.

We use size n Random Sample to Model the situation:

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

Example

Example: Assume we have a Sample of size n from some Population, and we want to estimate the Mean of that Population. Let $\mathcal{F}_\mu$ be the Distribution of our Population, and μ be the Mean (Expectation) of that Distribution. So our unknown Parameter here is μ.

We use size n Random Sample to Model the situation:

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

Now, we choose several Estimators to estimate μ:

$$\hat{\mu}_1 = X_1, \qquad \hat{\mu}_2 = \frac{X_1 + X_3}{2}, \qquad \hat{\mu}_3 = \frac{X_1 + X_4}{10},$$

$$\hat{\mu}_4 = \overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n}.$$

Example

Example: Assume we have a Sample of size n from some Population, and we want to estimate the Mean of that Population. Let $\mathcal{F}_\mu$ be the Distribution of our Population, and μ be the Mean (Expectation) of that Distribution. So our unknown Parameter here is μ.

We use size n Random Sample to Model the situation:

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

Now, we choose several Estimators to estimate μ:

$$\hat{\mu}_1 = X_1, \qquad \hat{\mu}_2 = \frac{X_1 + X_3}{2}, \qquad \hat{\mu}_3 = \frac{X_1 + X_4}{10},$$

$$\hat{\mu}_4 = \overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n}.$$

Problem: Check if each Estimator is Biased or Unbiased.

Solution: OTB

Example

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

Example

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

UnBiased Estimator Case

We consider the Poisson Model:

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

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

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

Example

Example: Let's do an experiment with Biased and Unbiased Estimators, with **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

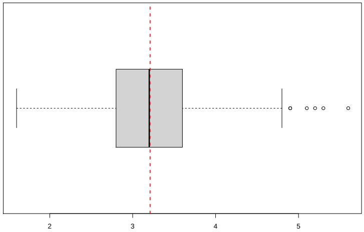
- observing once: ganarating a Sample just once and calculating one Estimate:

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

```
## [1] 3
```

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

```
lambda <- 3.21; n <- 10; m <- 2000
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] 3.22645
```

With a Histogram:

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