

# W1211 Introduction to Statistics

## Lecture 19

Wei Wang

Nov 12, 2012

# Point Estimation

- ▶ A point estimate of a parameter  $\theta$  is a single number that can be regarded as a sensible value for  $\theta$ . A **point estimate** is obtained by selecting a suitable statistic and computing its value from the given sample data. The selected statistic is called **the point estimator** of  $\theta$ , denoted by  $\hat{\theta}$ .

# Point Estimation

- ▶ A point estimate of a parameter  $\theta$  is a single number that can be regarded as a sensible value for  $\theta$ . A **point estimate** is obtained by selecting a suitable statistic and computing its value from the given sample data. The selected statistic is called **the point estimator** of  $\theta$ , denoted by  $\hat{\theta}$ .
- ▶ For example, if we flip a biased coin  $n$  times and get  $X$  heads, and we want to estimate the probability of getting heads  $p$ , it is intuitive to use sample proportion(mean)  $X/n$  as the estimator. If we observe  $n = 100$  and  $X = 73$ , then our estimate of  $p$  is 0.73.

# Measure of Good Estimators

- ▶ First of all estimator is a random variable. In the biased coin example, we use the sample proportion  $X/n$  as our estimator of  $p$ . Can you say something about the distribution of  $\hat{p}$ ?

# Measure of Good Estimators

- ▶ First of all estimator is a random variable. In the biased coin example, we use the sample proportion  $X/n$  as our estimator of  $p$ . Can you say something about the distribution of  $\hat{p}$ ?
- ▶ From the knowledge of sampling distribution, we know  $E(\hat{p}) = p$  and  $Var(\hat{p}) = p(1 - p)/n$ . And if  $n$  is large,  $\hat{p}$  approximately follows a normal distribution.

# Measure of Good Estimators

- ▶ First of all estimator is a random variable. In the biased coin example, we use the sample proportion  $X/n$  as our estimator of  $p$ . Can you say something about the distribution of  $\hat{p}$ ?
- ▶ From the knowledge of sampling distribution, we know  $E(\hat{p}) = p$  and  $Var(\hat{p}) = p(1 - p)/n$ . And if  $n$  is large,  $\hat{p}$  approximately follows a normal distribution.
- ▶ In principle, there are many estimators for a given parameters. Two properties of estimators are desired.
  - ▶ **Unbiasedness** This is about how faithful the estimator is. A point estimator  $\hat{\theta}$  is an unbiased estimator of  $\theta$  if  $E(\hat{\theta}) = \theta$  for every possible value of  $\theta$ . If  $\hat{\theta}$  is not unbiased, then  $E(\hat{\theta}) - \theta$  is called the bias of  $\hat{\theta}$ .

# Measure of Good Estimators

- ▶ First of all estimator is a random variable. In the biased coin example, we use the sample proportion  $X/n$  as our estimator of  $p$ . Can you say something about the distribution of  $\hat{p}$ ?
- ▶ From the knowledge of sampling distribution, we know  $E(\hat{p}) = p$  and  $Var(\hat{p}) = p(1 - p)/n$ . And if  $n$  is large,  $\hat{p}$  approximately follows a normal distribution.
- ▶ In principle, there are many estimators for a given parameters. Two properties of estimators are desired.
  - ▶ **Unbiasedness** This is about how faithful the estimator is. A point estimator  $\hat{\theta}$  is an unbiased estimator of  $\theta$  if  $E(\hat{\theta}) = \theta$  for every possible value of  $\theta$ . If  $\hat{\theta}$  is not unbiased, then  $E(\hat{\theta}) - \theta$  is called the bias of  $\hat{\theta}$ .
  - ▶ **Small Variance** This is about how stable the estimator is. Unbiased estimators are faithful in the long run, but might have large fluctuation when sample size is small.

# Principles of Selecting Estimators

- ▶ First, choose the estimators that are unbiased.
- ▶ Then, among the unbiased estimators, choose the one with the smallest variance.



# Two General Results about Unbiasedness

- ▶ Let  $X_1, X_2, X_3, \dots, X_n$  be a random sample from a distribution with mean  $\mu$  and variance  $\sigma^2$ , and if we use sample mean  $\bar{X}$  as the estimator of population mean  $\mu$ , then  $\hat{\mu} = \bar{X}$  is an unbiased estimator of  $\mu$

# Two General Results about Unbiasedness

- ▶ Let  $X_1, X_2, X_3, \dots, X_n$  be a random sample from a distribution with mean  $\mu$  and variance  $\sigma^2$ , and if we use sample mean  $\bar{X}$  as the estimator of population mean  $\mu$ , then  $\hat{\mu} = \bar{X}$  is an unbiased estimator of  $\mu$
- ▶ Let  $X_1, X_2, X_3, \dots, X_n$  be a random sample from a distribution with mean  $\mu$  and variance  $\sigma^2$ , and if we use sample variance  $S^2$  as the estimator of population variance  $\sigma^2$ , then  $\hat{\sigma}^2$  is an unbiased estimator of  $\sigma^2$ , i.e.,

$$E\left(\frac{\sum (X_i - \bar{X})^2}{n - 1}\right) = \sigma^2$$

# MVUE

- For unbiased estimators, what are their MSE's?

$$E(\hat{\theta} - \theta)^2 = E(\hat{\theta} - E(\hat{\theta}))^2 = \text{Var}(\hat{\theta})$$

- Among all estimators of  $\theta$  that are unbiased, choose the one that has minimum variance. The resulting  $\hat{\theta}$  is called the **minimum variance unbiased estimator (MVUE)** of  $\theta$ .
- One needs more knowledge to actually identify if some estimator is really MVUE. But in a special case, we have the following theorem.

Let  $X_1, X_2, \dots, X_n$  be an i.i.d. sequence of random samples from a **normal distribution** with mean  $\mu$  and  $\sigma$ . Then the estimator  $\hat{\mu} = \bar{X}$  is the **MVUE** for  $\mu$ .

# The Standard Error

- When reporting a point estimator, one also reports the **standard error** associated with it.
- The **standard error** of an estimator  $\hat{\theta}$  is its standard deviation  $\sigma_{\hat{\theta}} = \sqrt{\text{Var}(\hat{\theta})}$ . If the standard error itself involves unknown parameters whose values can be estimated, substitution of these estimates into  $\sigma_{\hat{\theta}}$  yields the **estimated standard error** of the estimator, which we denote as  $\hat{\sigma}_{\hat{\theta}}$ .
- The associated standard error gives us an idea of how good/accurate the estimators are.

# Estimator, Its Standard Error and Estimated Standard Error

- ▶ Now we are trying to estimate the probability of getting heads of a biased coin, so each flip  $X_i$  is a Bernoulli RV with parameter  $p$ , the estimator of parameter  $p$  is the sample mean/proportion

$$\hat{p} = \frac{\sum_{i=1}^n X_i}{n}$$

# Estimator, Its Standard Error and Estimated Standard Error

- ▶ Now we are trying to estimate the probability of getting heads of a biased coin, so each flip  $X_i$  is a Bernoulli RV with parameter  $p$ , the estimator of parameter  $p$  is the sample mean/proportion

$$\hat{p} = \frac{\sum_{i=1}^n X_i}{n}$$

- ▶ If we flipped 100 times and observed 75 heads, then our estimate of  $p$  is

$$\hat{p} = \frac{75}{100} = 0.75$$

# Estimator, Its Standard Error and Estimated Standard Error

- ▶ Now we are trying to estimate the probability of getting heads of a biased coin, so each flip  $X_i$  is a Bernoulli RV with parameter  $p$ , the estimator of parameter  $p$  is the sample mean/proportion

$$\hat{p} = \frac{\sum_{i=1}^n X_i}{n}$$

- ▶ If we flipped 100 times and observed 75 heads, then our estimate of  $p$  is

$$\hat{p} = \frac{75}{100} = 0.75$$

- ▶ Also, we need to report how good our estimator is through its Standard Error. This is also related to the Interval Estimation.

# Estimator, Its Standard Error and Estimated Standard Error

- ▶ The standard error is  $Var(\hat{p}) = \frac{p(1-p)}{n}$ , but we cannot report it since we don't know what  $p$  is.
- ▶ So we can only report the estimated standard error of the estimator  $\hat{p}$

$$\widehat{Var(\hat{p})} = \frac{\hat{p}(1 - \hat{p})}{n}$$



# Another Example

- ▶ Now we have  $X_1, X_2, \dots, X_N$  IID with mean  $\mu$  and variance  $\sigma^2$ , what's the estimator of  $\mu$ ?

# Another Example

- ▶ Now we have  $X_1, X_2, \dots, X_N$  IID with mean  $\mu$  and variance  $\sigma^2$ , what's the estimator of  $\mu$ ?

▶

$$\hat{\mu} = \bar{X}$$

# Another Example

- ▶ Now we have  $X_1, X_2, \dots, X_N$  IID with mean  $\mu$  and variance  $\sigma^2$ , what's the estimator of  $\mu$ ?

▶

$$\hat{\mu} = \bar{X}$$

- ▶ The standard error of  $\hat{\mu}$  is  $\sqrt{\text{Var}(\hat{\mu})} = \frac{\sigma}{\sqrt{n}}$ . Can we report  $\frac{\sigma}{\sqrt{n}}$ ?

# Another Example

- ▶ Now we have  $X_1, X_2, \dots, X_N$  IID with mean  $\mu$  and variance  $\sigma^2$ , what's the estimator of  $\mu$ ?

▶

$$\hat{\mu} = \bar{X}$$

- ▶ The standard error of  $\hat{\mu}$  is  $\sqrt{\text{Var}(\hat{\mu})} = \frac{\sigma}{\sqrt{n}}$ . Can we report  $\frac{\sigma}{\sqrt{n}}$ ?
- ▶ It really depends on whether or not we know  $\sigma$ . If we know it, then we can report  $\frac{\sigma}{\sqrt{n}}$ ; otherwise, we can only report  $\frac{\hat{\sigma}}{\sqrt{n}}$ .

# Examples

Ex. What is the standard error of the sample proportion?

Ex. What is the standard error of the sample mean? Given that we know the variance.

Ex. What is the standard error of the sample mean? Given that we don't know the variance.