BU551V - Applied Econometrics
Lecture Notes

Rodrigo Miguel

February 1, 2023

Contents

Chapter 1

Lecture 1

1.1 What is statistics?

Statistics is the science of uncertainty.

- What *could* be;
- What *might* be;
- What *probably* is.

Statistics can be divided into two:

- **Descriptive statistics:** Used to summarize information (how many), describe in a concise and easy to understand way.
- **Inferential statistics:** Asks about what we can learn from the population in the sample.

1.1.1 Example

If an airline is selling tickets to a flight with only 100 seats. A safe solution is to accept only 100 reservations. However, due to the human nature, cancellations and no-show-ups might happen. What would be the optimal number of reservations the airline should make?

1.2 Statistical Inference

1.2.1 Statistical Inference

Is used when we want to learn about a population given a sample/subset.

1.2.2 Steps to Make a Statistical Inference

- Identify the population of interest;
- Specify a model for the population relationship of interest that contains unknown parameters.
- Obtain a **random** sample from the population.
- Estimate the parameters using the sample.

1.2.3 Random Sample

A **Random Sample** is a set of random variables (RVs) that follow a probability density function (pdf) $f(y, \theta)$.
When the RVs are from a random sample with $f(y, \theta)$ density, they are said to be **independent and identically distributed**.
Random samples can be assumed to be drawn from a normal distribution. It's population can be characterized by two parameters:

- Mean μ;
- Variance σ^2.

Usually we are only interested in μ, however, in some cases, we also need to know σ^2.

1.2.4 Estimators

An estimator is a rule that assigns a value of θ to the sample.
For an actual realized sample, the estimate is just the average in the sample:

$$\overline{y} = \frac{(y_1 + y_2 + ... + y_n)}{n}$$

1.2.5 General Expression for Estimators

An estimator W of a parameter θ can be expressed as an abstract formula:

$$W = h(Y_1, Y_2, ..., Y_n)$$

, where h is an unknown function.
When we plug a set of actual observations $y_1, y_2, ...y_n$ into the function h, we obtain an estimate of θ.

1.3 Choosing an estimator

To choose an estimator, criteria needs to be followed so that an estimator with "desirable" properties is correctly chosen. Here's the several properties that are used as choosing criteria:

- Unbiasedness;
- Efficiency;
- Consistency.

1.4 Unbiasedness

An estimator of θ is **unbiased** if:

$$E(W) = \theta$$

This does, however, **not** mean that an estimate in a particular sample is equal to θ.
It means that:

1. We draw random samples from the population many times;
2. Compute an estimate each time;
3. Average these estimates over all random samples;
4. Then we obtain the true parameter θ.

1.4.1 Bias of an estimator

An estimator of θ, W, is a biased estimator if:

$$E(W) \neq \theta$$

Its bias is defined as:

$$Bias(W) \equiv E(W) - \theta$$

- An estimator being unbiased does not necessarily mean it's a good estimator.

1.5 Efficiency

1.5.1 Sampling Variance of Estimators

- Unbiasedness ensures that the distribution of an estimator has a mean value which is equal to the true parameter θ.
- It is a good property but not enough.
- We want an estimator that can show us a mean value equal to θ and centred around θ as tightly as possible.

To measure how spread out a distribution is, we use the variance, Var(W). To calculate it, we use the sample average as an estimator $w = \overline{Y}$.

Summary If $Y_1, ..., Y_n$ is a random sample from a population with mean μ and variance σ^2:

- $\overline{Y}$ has a mean μ and variance $\frac{\sigma^2}{n}$

1.5.2 Which estimator is better?

Among unbiased estimators, we prefer the estimator with the smallest variance.

1.6 Consistency

Consistency, is a property that captures how estimators improve as the sample size, n, increases.

1.6.1 Example

- An estimator $\overline{Y}$ has $Var(\overline{Y}) = \frac{\sigma^2}{n}$, meaning that as n gets larger, the estimator improves.
- An estimator Y_1 has $Var(Y_1) = \sigma^2$, meaning no improvement even if n gets larger.

1.6.2 Importance of Consistency

Consistency is a minimum requirement of an estimator used in econometrics.

1.6.3 Law of Large Number