Review - Mathematical Statistics

Caio Vigo

The University of Kansas
Department of Economics

Spring 2020

These slides were based on *Introductory Econometrics* by Jeffrey M. Wooldridge (2015)

Topics

Population, Parameters, and Random Sampling

- Statistical inference involves **learning** (or inferring) some thing about a population given the availability of a sample from that population.
- Inferring mainly comprises two tasks:
  1. estimation,
     - point estimate
     - interval estimate
  2. hypothesis testing

Population, Parameters, and Random Sampling

Population

Any well defined group of subjects, which would be individuals, firms, cities, or many other possibilities.

- **Examples:**
  - blood / blood test sample
  - preparing a pot of soup / a spoon of soup to try it
  - all working adults in US / a sample from it (it's impractical to collect data from the entire population)

- Let Y be a r.v. representing a population with p.d.f. $f(y;\theta)$
- The p.d.f. of Y is assumed to be known, except for the value of θ

Random Sample

If $Y_1, Y_2, \ldots, Y_n$ are independent r.v. with a common probability density function $f(y;\theta)$, then $\{Y_1, Y_2, \ldots, Y_n\}$ is said to be a **random sample** from $f(y;\theta)$ [or a random sample from the population represented by $f(y;\theta)$]

Sampling

- When $\{Y_1, Y_2, \ldots, Y_n\}$ is a random sample from the density $f(y;\theta)$, we also say that the Y_i are *independent, identically distributed* (or i.i.d.) r.v. from $f(y;\theta)$

- Whether or not it is appropriate to assume the sample came from a random sampling scheme requires knowledge about the actual sampling process.

Estimators and Estimates

- **Estimator** = Rule

Estimator

Given a population,

in which this population distribution depends of a parameter θ

you draw a random sample $\{Y_1, Y_2, \ldots, Y_n\}$.

Then an **estimator** of θ, say W, is a rule that assigns each outcome of the sample a value of θ.

- **Example** (on board) **sample average** and **sample variance**.

- **Attention!**

$$\text{Parameter} \neq \text{Estimator} \neq \text{estimate}$$

Estimator

Thus, an estimator is

$$W = h(Y_1, Y_2, \ldots, Y_n)$$

Unbiasedness

Unbiased Estimator

An estimator W of θ, is an **unbiased estimator** if

$$E(W) = \theta$$

- Unbiasedness does not mean that the **estimate** we get with any particular sample is equal to θ (or even close to θ).

Unbiasedness

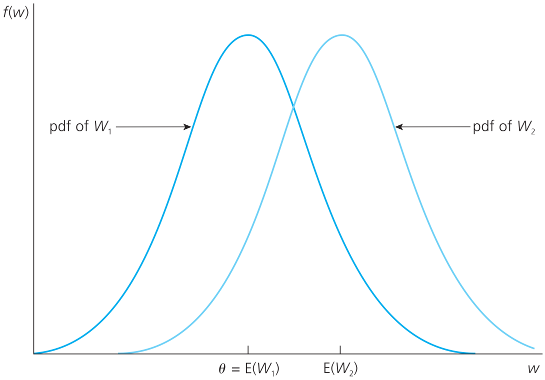
Bias

If W is **biased estimator** of θ, its bias is defined

$$\text{Bias}(W) = E(W) - \theta$$

- Some estimators can be shown to be unbiased quite generally.
- **Example** (on white board): sample average $(\bar{Y})$.

Figure: An unbiased estimator, W_1, and an estimator with positive bias, W_2



Source: Wooldridge, Jeffrey M. (2015). Introductory Econometrics: A Modern Approach.

Unbiasedness

- Even though being an unbiased estimator is a good quality for an estimator, we should not try to reach it at any cost. There are good estimators that are biased, and there are bad estimators that are unbiased (example: $W \equiv Y_1$)

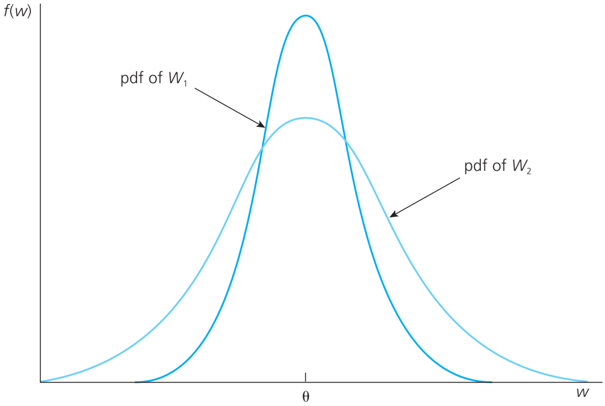
The Sampling Variance of Estimators

- Another criteria to evaluate estimators.
- We also would like to know how spread an estimator might be.

Sampling Variance: the variance of an estimator

The Sampling Variance of Estimators

Figure: The sampling distributions of two unbiased estimators of θ



Source: Wooldridge, Jeffrey M. (2015). Introductory Econometrics: A Modern Approach.

Efficiency

Efficiency (Relative)

If W_1 and W_2 are two unbiased estimators of θ, W_1 is efficient relative to W_2 when

$$\text{Var}(W_1) \leq \text{Var}(W_2)$$

for all θ, with strict inequality for at least one value of θ.

Efficiency

- One way to compare estimators that are not necessarily unbiased is to compute the **mean squared error (MSE)** of the estimators.

Mean Squared Error (MSE)

$$\begin{aligned} \text{MSE}(W) &= E\left[(W-\theta)^2\right] \\ &= Var(W) + [Bias(W)]^2 \end{aligned}$$

Consistency

- We can rule out certain silly/bad estimators by studying the *asymptotic* or *large sample* properties of estimators.

- It is related to the behavior of the sampling distribution when the sample size n gets large.

- If an estimator is not consistent (i.e., **inconsistent**), then it does not help us to learn about θ, even with with an unlimited amount of data.

- **Consistency:** minimal requirement of an estimator.

- Unbiased estimators are not necessarily consistent.

Consistency

Consistency

An estimator W of θ, is a **consistent** if

$$W_n \xrightarrow{\text{p}} \theta$$

Consistency

Let W_n be an estimator of θ based on a sample. Then, W_n is a **consistent estimator** of θ if for every $\epsilon > 0$,

$$\mathbb{P}(|W_n - \theta| > \epsilon) \to 0, \text{ as } n \to \infty$$

Law of Large Numbers (LLN)

- Under general conditions, $\bar{Y}$ will be near μ with very high probability when n is large.

Law of Large Numbers (LLN)

Let $Y_1, Y_2, \ldots, Y_n$ be i.i.d. random variables with mean μ. Then,

$$\bar{Y}_n \xrightarrow{\text{p}} \mu$$

Law of Large Numbers (LLN)

- The **LLN** does NOT say that the estimator $\bar{Y}$ will converge to any type of distribution. *(Don't confuse with the Central Limit Theorem).*

- The **LLN** just says that the estimator will converge to the true parameter, i.e, the sample average $\bar{Y}$ will get closer and closer to the true parameter μ as you increase the sample size.

Central Limit Theorem (CLT)

Central Limit Theorem (CLT)

Let $Y_1, Y_2, \ldots, Y_n$ be i.i.d. with mean μ and variance σ^2. Let,

$$Z_n = \frac{\bar{Y}_n - \mu}{\sigma/\sqrt{n}}$$

Then, Z_n will converge to a Normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$, i.e., to a $N(0, 1)$ as $n \to \infty$

Website Suggestions to Explore

- **Seeing Theory:** `https://seeing-theory.brown.edu/`
- **Statistics Web Apps:** `http://www.artofstat.com/webapps.html`