

Lecture 3

Chapters R.5-R.8

R.5 Sampling

Population – a group of items or events we would like to know about.

Parameter – a value that describes that population. The parameter of interest is the parameter that the researcher seeks to learn about.

Sample – a survey of a subset of the population.

Exercise 1: Come up with an example of a population, a parameter from that population, and a sample of that population.

Population –

Parameter –

Sample –

Exercise 2: Suppose we want to learn about the athleticism of college students. If we only survey football wide receivers, would we have a sample that represents the average college student? Why or why not?

Exercise 3: Suppose instead we decide to survey every single college student in the United States. Does this sound reasonable? Why or why not?

Exercise 4: How could you get a sample that represents the population of college students without surveying every college student?

Simple Random Sampling – each observation is selected at random from the entire set of possible observations.

Estimator – a rule (or formula) for estimating an unknown population parameter given a sample of data.

Note: Each observation of the sample is a random variable. The estimator is a combination of random variables, so it is also a random variable.

Estimate – the specific number obtained from the estimator.

Sample mean – the estimator for population mean.

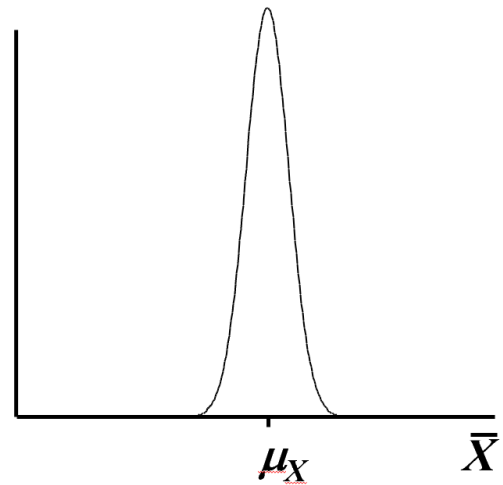
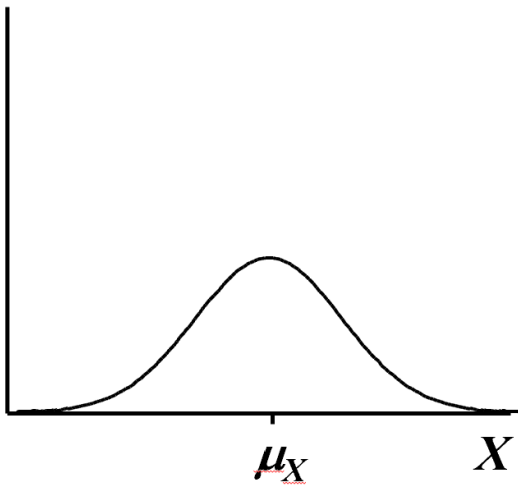
Formula – Sample mean

Exercise 5: Suppose we take a survey of individuals to find out what the average number of pets owned in Eugene is. Given the sample data below, what is the estimate?

Individual	Number of Pets
1	2
2	1
3	0
4	5
5	2

Exercise 6: suppose X has a distribution with variance σ_X^2 . What is the variance of the sample mean?

The probability density of an estimate of a parameter of X always has lower variance than distribution of X .

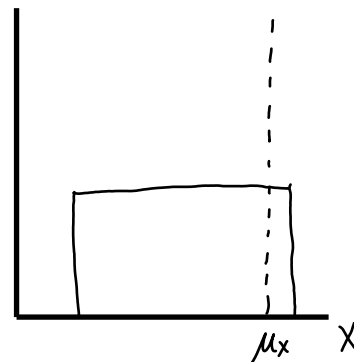
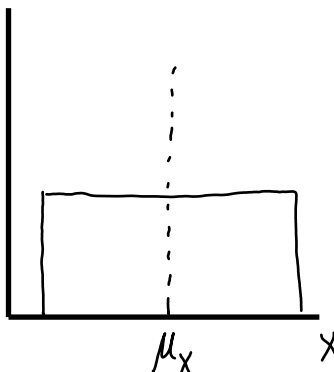
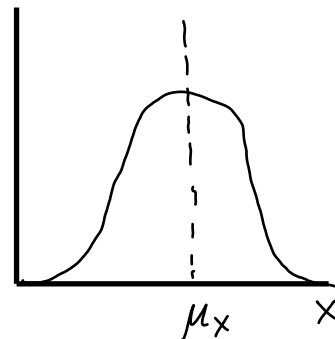
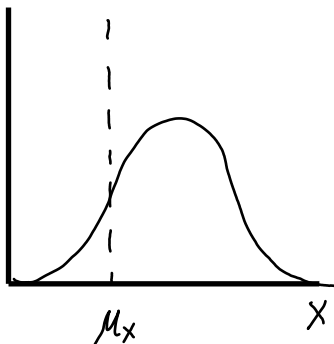


R.6 Unbiasedness and Efficiency

We need two things for a reliable estimator.

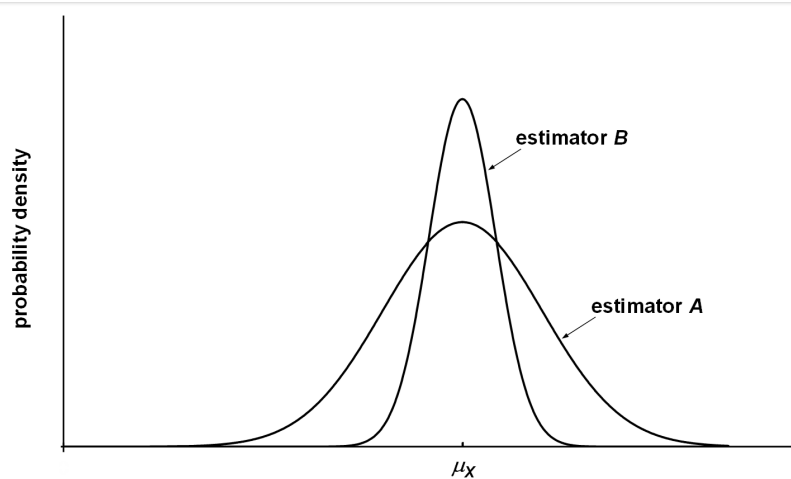
1. Unbiasedness – the expected value of the estimator of a parameter equals the true value of the parameter.
2. Efficiency – an unbiased estimator that has low variance.

Exercise 7: Label the following distributions as biased or unbiased.

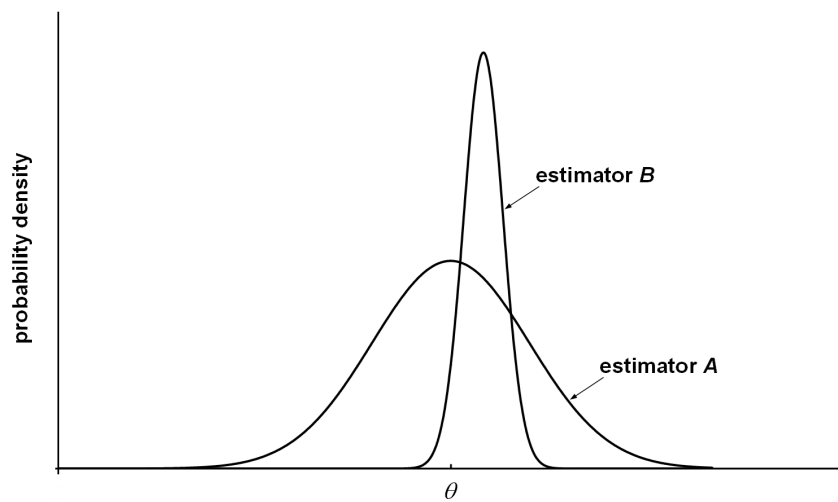


Exercise 8: Show that sample mean is an unbiased estimator of population mean.

Exercise 9: Which estimator would you rather use A or B? Why?



Exercise 10: Which estimator would you rather use, A or B? Why?



R.7 Estimators of variance, covariance, and correlation

Formula – sample variance

Formula – sample covariance

Formula – sample correlation

Exercise 11: Find the sample means, variances, covariance, and correlation of the following sample data.

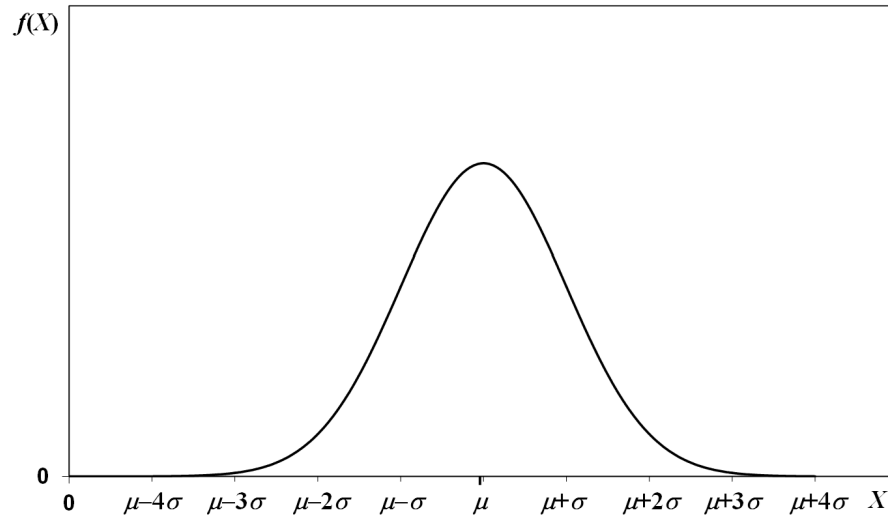
X	Y
1	4
2	5
3	6

R.8 The normal distribution

There are 4 continuous distributions that we will use in econometrics.

1. Normal distribution
2. t distribution
3. F distribution
4. χ^2 distribution (chi-squared)

The normal distribution – a distribution that is “bell shaped” around the population mean. Most of it’s probability lies close to the middle, with relatively little far away.



Standard normal distribution – a normal distribution with mean = 0 and standard deviation = 1.

To find probabilities for a standard normal distribution, we use a z-table.

Exercise 12: What is the probability that $X < 1$ for a standard normal distribution?

Exercise 13: What is the probability that $X > 2$ for a standard normal distribution?

Exercise 14: What is the probability that $-1 < X < 0.5$ for a standard normal distribution?

Exercise 15: What is Z if $\Pr(X < Z) = 0.5120$?

Exercise 16: What is Z if $\Pr(X < Z) = 0.877$?