GOV 1000/2000/2000e/E-2000 10/03/2013

Section 5: Interval Estimation and Hypothesis Testing

Solé Prillaman

This Week

Problem Set 4- Due by 1pm on Tuesday to course dropbox.
Problem Set 3 Corrections- Due by 1pm on Tuesday to course dropbox.
No Reading Quiz
Practice Midterms- Not due, but recommended as midterm preparation.

Next Week

Midterm- Posted October 8, due by 11:59pm on Sunday, October 13 with 5 hours to complete.
Office Hours- Cancelled Friday October 6. Rescheduled for Tuesday October 15 from 9:30am to 12pm.
Problem Set 5- Due by 1pm on Tuesday, October 15 to course dropbox.
Problem Set 4 Corrections- Due by 1pm on Tuesday, October 15 to course dropbox.
Reading Quiz (ALZ ch. 11 & 12)- Due by 9pm on Sunday, October 13 on course website.
Mid-semester Evaluation- Due by 9pm on Monday, October 14 on course website.

Sampling

Parameter- characteristic of the population distribution (the distribution of X_i)
ex. μ or σ^2
Estimand- parameter being estimated
ex. μ
Statistic- function of the sample used to estimate a parameter
ex. $\bar{X}_n$ or $\hat{\mu}$ or S^2
Estimator- a statistic (random variable) that describes the estimation procedure
ex. $\bar{X}_n$
Estimate- realized values of an estimator
ex. $\bar{x}_n$

Notation	Description	Formula[1]/Distribution	Random Variable	Observed?	Estimand/te/tor
n	Sample size		No	Yes	
N	Population size		No	?	
s	Number of samples		No	Yes	
Y_i	Random variable	$\sim ?(\mu, \sigma^2)$	Yes	No	
y_i	Realization of Y_i		No	Yes	
$\mu, E[Y_i], \bar{y}_N$	Population mean	$E[Y_i] = \frac{1}{N}\sum_{i=1}^{N} y_i$	No	No	Estimand
$\hat{\mu}$	Estimator for μ		Yes	No	Estimator
	Estimate of $\hat{\mu}$ (realization)		No	Yes	Estimate
$\bar{Y}_n$	Sample mean	$\sim N(\mu, \frac{\sigma^2}{n})$ for n large	Yes	No	Estimator
$\bar{y}_n$	Realization of $\bar{Y}_n$	$\frac{1}{n}\sum_{i=1}^{n} y_i$	No	Yes	Estimate
σ^2	Population variance	$V[Y_i] = \frac{1}{N}\sum_{i=1}^{N}(y_i - \mu)^2$	No	No	Estimand
σ	Population standard deviation	$SD[Y_i] = \sqrt{\sigma^2}$	No	No	Estimand
S_n^2	Sample variance	$\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$	Yes	No	Estimator
S_n	Sample standard deviation	$\sqrt{S_n^2}$	Yes	No	Estimator
s_n^2	Realization of S_n^2	$\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2$	No	Yes	Estimate
s_n	Realization of S_n	$\sqrt{s_n^2}$	No	Yes	Estimate
$SE[\bar{Y}_n]$; $SE[\hat{\mu}]$	Standard deviation of $\bar{Y}_n$	$\frac{\sigma}{\sqrt{n}}$	No	No	Estimand
$\hat{SE}[\bar{Y}_n]$; $\hat{SE}[\hat{\mu}]$	Estimator of standard error	$\frac{S_n}{\sqrt{n}}$	Yes	No	Estimator
	Estimated standard error	$\frac{s_n}{\sqrt{n}}$	No	Yes	Estimate

[1] The formulas in this table assume that Y_i is a discrete random variable and each outcome has equal probability. These can be adapted to more general formulas found in Lecture 3.

Practice Conceptual Questions

1. What is the difference between and estimand, estimator, and estimate? How do these relate to parameters and statistics?

2. What is the difference between a statistic and a test statistic?

3. Given that you have only one sample from the population, what are some point estimates you might calculate and how would you calculate them?

4. What do we do when we don't know σ? What are the implications of this?

5. What is bootstrapping and why do we do it?

6. Why can't we use the t-distribution with small samples unless the underlying population distribution is normal?

7. Why do we get our critical values (eg. 1.96) from the standard normal distribution?

8. Can we interpret our confidence intervals the same when we don't know σ as when we do? Do these confidence intervals truly cover the population mean in 95% of samples? Why is this not a problem in most instances?

9. Why and how is hypothesis testing related to interval estimation?

10. How are p-values related to hypothesis testing?