

636, Fall 2009

Homework 3,

Due September 28, before class begins.

1. Suppose there was a census taken in Armenia and the population mean income for people aged from 20-30 was \$2409. Some scientist not aware of a census or of the mean calculation samples randomly 20 people in the 20-30 age bracket asking them their income and converts it to dollars. The data is given below:

3000	1000	5500	6400	3000
6000	800	3300	5500	500
8000	16000	9000	7100	5200
300	5000	100	2500	2900

- a. Using Stata, calculate the mean and sample variance for this problem.
b. What is the critical “t-value” for determining a 95% confidence interval of the mean test statistic?
c. Determine a 95% confidence interval for the sample. Was the population mean within the confidence range?
2. Let X_1, X_2, \dots, X_n denote a random sample from a normal distribution with mean 0 and variance σ^2 . Prove that $\sum_{i=1}^n X_i^2 / n$ is an unbiased estimator of σ^2 .

To start a proof, start with something you know and progress through it stating the logic for each step:

Thus, for this problem:

Step 1: A statistic $\hat{\theta}$ is said to be an unbiased estimator of the parameter θ if $\mu_{\hat{\theta}} = E(\hat{\theta}) = \theta$.

Step 2. Let $\hat{\theta} = \sum_{i=1}^n X_i^2 / n$ be a candidate estimator for σ^2

Step 3. $E(\hat{\theta}) = \dots$ [state logic]

3. Suppose the random variables: X_1, X_2, \dots, X_n are drawn from a distribution having the following p.d.f.

$$f(x) = \theta x^{\theta-1}, \quad 0 < x < 1 \text{ and where } 0 < \theta < \infty$$

$$= 0 \text{ otherwise}$$

Find the maximum likelihood estimator of θ

4. Suppose the random variables: X_1, X_2, \dots, X_n are drawn from a distribution having the following p.d.f.

$$f(x) = \theta e^{-\theta x}, \quad 0 < x \text{ and where } 0 < \theta < \infty$$

$$= 0 \text{ otherwise}$$

Find the maximum likelihood estimator of θ

5. You are given the following results from a regression:

. regress meanwage meduc

Source	SS	df	MS	Number of obs =	15
Model	2.2875e+10	1	2.2875e+10	F(1, 13) =	_____
Residual	1.8566e+09	13	142812410	Prob > F =	0.0000
Total	2.4732e+10	14	1.7666e+09	R-squared =	_____
				Adj R-squared =	0.9192
				Root MSE =	11950

meanwage	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
meduc	1142.342	_____	12.66	0.000	_____ 1337.338
_cons	_____	17460.37	-5.88	0.000	_____ -64914.43

- a) Fill in the missing values (may need to peek ahead a little bit)
 b) How would you interpret the coefficient 1142.342
 c) At the mean (see summary command below) of the data, calculate the elasticity of education on wage.

. sum meanwage meduc

Variable	Obs	Mean	Std. Dev.	Min	Max
meanwage	15	114866.7	42030.37	34000	162000
meduc	15	190.4	35.38523	132	240

6. A manufacturer produces two types of cheap tires, one type with a life expectancy of 25,000 miles and one with a life expectancy of 30,000 miles. The known variation in durability around the expected lifetime is the same for both types of tires: the standard deviation being 3000 miles. The distribution can be assumed to be normal. Normally the tires have markings that identify the tire type. However, a case of 100 tires is found without markings and therefore appear identical to both types of tires. The examining engineer thinks that the tires are of the less durable kind but recommends a test. A test is to be performed on a sample of 4 tires. The scientist sets up a null hypothesis that the life expectancy is 25,000 and an alternative that it is equal to 30,000.

- a) Produce a diagram that shows the PDF around of both means in question. Make sure the distributions overlap. We will only be concerned with the intersecting tails
- b) Mark and label a 5% critical value reasonably well.
- c) Shade in and label what will amount to the Type I and Type II errors.
- d) Now, figure out the mileage for the Type I error for the mark you made in part b). What is the probability mass of the Type I error?
- e) Calculate the probability mass for the Type II error.
- f) Repeat d) and e) for 1% critical value.
- g) Discuss how a manager under pressure to maximize short-term revenue could seriously hurt the firm. Be specific about the relationship between 'harm' and the types of errors that are possible.

7. not graded, but I will read and give some comments:

A: (optional), I would like you to state what you plan to do with your project. Lay out your 8-step approach and indicate what you are thinking about for each step. Describe the model and data you are thinking about.

B: (do this one if you skipped A): List three topics from which you will likely pick one for your project. Briefly explain why each is important and why you are motivated to look at these for your research. By doing B, you may discover that one of the three is clearly the best. If so, then try to explain part of 8-step approach for the likely winning project, explain your data sources and begin describe a model.