Exam #1

Economics 435: Quantitative Methods

Spring 2005

Question 1 is worth 24 points, question 2 is worth 26, question 3 is worth 20, and question 4 is worth 30.

1 Regression with standardized variables

Suppose that we have a random sample of size n on the random variables y and x such that:

$$E(y|x) = \beta_0 + \beta_1 x$$

Now suppose that we standardize both the independent and dependent variable. Let $(\tilde{y}_i, \tilde{x}_i)$ be defined by:

$$\begin{aligned} \tilde{y}_i &= \frac{y_i - \bar{y}}{\sqrt{\hat{var}(y)}} \\ \tilde{x}_i &= \frac{x_i - \bar{x}}{\sqrt{\hat{var}(x)}} \end{aligned}$$

Let $\hat{\beta}_1$ be the coefficient on $\tilde{x}_i$ from an OLS regression of $\tilde{y}$ on $\tilde{x}$.

a) Find $\hat{\beta}_0$ as a function of $\hat{\beta}_1$.

b) Find the R^2 from the regression as a function of $\hat{\beta}_1$. In case you didn't write it down,

$$R^2 = \frac{\text{explained sum of squares}}{\text{total sum of squares}}$$

c) Find the sample correlation $\hat{corr}(\tilde{y}, \tilde{x})$ as a function of $\hat{\beta}_1$.

d) Can $\hat{\beta}_1$ possibly be less than -1 or greater than $+1$? Prove it.

2 Heteroskedasticity: A simple example

We are going to consider a very simple case of heteroskedasticity[1] in a linear regression setting. Let (x, y, u) be a set of random variables such that x is binary (either zero or one), $E(u|x) = 0$, and:

$$y = \beta_0 + \beta_1 x + u$$

Furthermore, we will assume that u is heteroskedastic:

$$\begin{aligned} E(u^2|x = 0) &= \sigma_0^2 \\ E(u^2|x = 1) &= \sigma_1^2 \end{aligned}$$

[1] Don't worry, this problem is *not* based on any material in the chapter on heteroskedasticity (chapter 8); it is just an application of tools discussed in lecture.

As usual we have a random sample $\{(y_i, x_i)\}_{i=1}^n$ with at least some variation in x_i.

Let $\bar{x}$, $\bar{y}_0$, and $\bar{y}_1$ be defined as follows:

$$\begin{aligned} \bar{x} &= \frac{1}{n}\sum_{i=1}^n x_i \\ \bar{y}_0 &= \frac{\sum_{i=1}^n y_i(1-x_i)}{n(1-\bar{x})} \\ \bar{y}_1 &= \frac{\sum_{i=1}^n y_i x_i}{n\bar{x}} \end{aligned}$$

In other words, $\bar{y}_0$ is the average value of y_i for all observations in which $x_i = 0$, and $\bar{y}_1$ is the average value of y_i for all observations in which $x_i = 1$.

Let $\hat{\beta}_0$ and $\hat{\beta}_1$ be the standard OLS estimators of β_0 and β_1:

$$\begin{aligned} \hat{\beta}_1 &= \frac{c\hat{o}v(y,x)}{v\hat{a}r(x)} \\ \hat{\beta}_0 &= \bar{y} - \hat{\beta}_1\bar{x} \end{aligned}$$

a) Show that for this case:

$$\begin{aligned} \hat{\beta}_1 &= \bar{y}_1 - \bar{y}_0 \\ \hat{\beta}_0 &= \bar{y}_0 \end{aligned}$$

b) Are the OLS estimators unbiased in this setting? Are they consistent?

c) Using this result, find $var(\hat{\beta}_0)$, $var(\hat{\beta}_1)$, and $cov(\hat{\beta}_0, \hat{\beta}_1)$ in terms of n, $\bar{x}$, σ_0 and σ_1. Note that $\bar{y}_0$ and $\bar{y}_1$ are independent conditional on the x's.

3 Hypothesis testing

Let x be a random variable with an unknown probability distribution, and suppose that we have a random sample of size n on that random variable: $\{x_i\}_{i=1}^n$. We are interested in testing several hypotheses about the probability distribution of x.

a) First, suppose we are testing the null that the median of x is zero against the alternative that it is not.

$$\begin{aligned} H_0: &\quad \Pr(x \leq 0) = 0.5 \\ H_1: &\quad \Pr(x \leq 0) \neq 0.5 \end{aligned}$$

To test this hypothesis, let:

$$y = I(x \leq 0)$$

where $I(.)$ is the indicator function, let:

$$\bar{y}_n = \frac{1}{n}\sum_{i=1}^n y_i$$

and let:

$$t = \sqrt{n}\frac{\bar{y}_n - 0.5}{0.5}$$

be our test statistic. Under the null H_0, the test statistic t has an asymptotic $N(0,1)$ distribution. We perform an asymptotic test of H_0 at the 5% significance level by rejecting H_0 if and only if $|t| > 1.96$.

Is it possible to reject the null based on a sample of size 2? What is the smallest sample size needed to have a chance at rejecting the null?

b) Now suppose we are interested in testing the null that x cannot possibly be larger than ten:

$$\begin{aligned} H_0 : &\quad \Pr(x \leq 10) = 1 \\ H_1 : &\quad \Pr(x \leq 10) < 1 \end{aligned}$$

Devise a testing procedure for this null hypothesis. In particular, define a test statistic t, and define the set of values for t under which you would reject the null at the 5% significance level.

4 Craps

If you have any questions about the rules below, please ask me - this is not a test on craps!

We are going to analyze the casino game of craps. Here are the basic rules. Craps is played with a pair of 6-sided dice, rolled by one of the gamblers (the "shooter"). There are many bets in craps, we will consider only one, called the "pass line" bet. Custom obligates the shooter to place at least some bet on the pass line.

The first roll of the game is called the "come out" roll. If the come out roll is a 7 or 11, the pass line wins and the game is over. If the come out roll is a 2, 3, or 12 the pass line loses, and the game is over. If the come out roll is a 4, 5, 6, 8, 9, or 10, then that number is called the "point" (for example, if a 4 is rolled, then the point is 4). The shooter then continues rolling the dice until either the point is rolled again (in which case the pass line wins) or a 7 is rolled (in which case the pass line loses.

a) Let the discrete random variable x be the total rolled on a pair of 6-sided dice. Describe the probability distribution of x by reporting its PDF. You need to get this right to do the rest, so be careful!

b) Calculate the probability of each of the following events:

1. Pass line wins on come out roll.
2. Pass line loses on come out roll.
3. Point of 4 established on come out roll.
4. Point of 5 established on come out roll.
5. Point of 6 established on come out roll.
6. Point of 8 established on come out roll.
7. Point of 9 established on come out roll.
8. Point of 10 established on come out roll.

c) Suppose a point has been established on the come out roll. Calculate the probability that the pass line will win conditional on each of the following events:

1. Point of 4 established on come out roll.
2. Point of 5 established on come out roll.
3. Point of 6 established on come out roll.
4. Point of 8 established on come out roll.

5. Point of 9 established on come out roll.

6. Point of 10 established on come out roll.

d) Calculate the (overall) probability that the pass line will win.

e) The pass line typically pays out at 1:1 odds. In other words, you win \$1 for every \$1 bet. Based on our calculations above, what is the house advantage (expected casino profit per \$1 bet)?