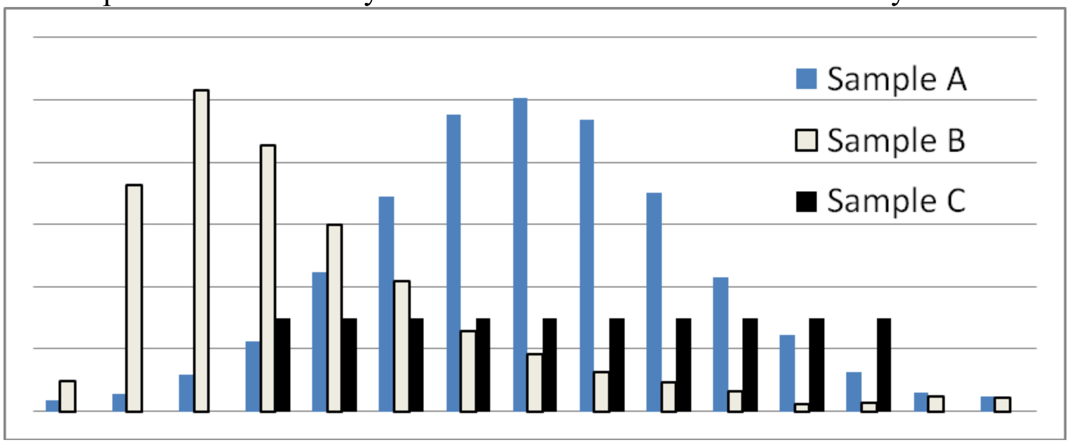
Econometrics Problem Sets
William M. Boal

Problem Set 10
"Two-Variable Regression: Normal Error Terms"

(10.1) [Normality assumption] Consider the histograms below from three samples. Which sample seems most likely drawn from a normal distribution? Why?



(10.2) [Confidence intervals, t-tests] Suppose the relationship between residential electricity consumption per capita and the price of electricity is estimated using a random sample of 20 cities with the following results. (The numbers on top are the least-squares estimates of the intercept and slope, and the numbers at the bottom in parentheses are standard errors.)

Electricity consumption	=	5273.1 (329.2)	-	450.5 (265.0)	Price of electricity

Assume that the error terms are normally-distributed.

a. Suppose electricity were free. According to these results, how much electricity would be consumed per capita?
b. Suppose the price of electricity were increased from \$0.05 to \$0.10. By how much would electricity consumption decrease?
c. For purposes of computing confidence intervals and tests, what are the degrees of freedom for this problem? (Give a number.)
d. Compute a 95 percent confidence interval for the intercept.
e. Test the null hypothesis that the price of electricity has no effect on consumption (that is, H_0: $\beta_2=0$) against the one-sided alternative hypothesis that it has a negative effect on consumption (that is, H_1: $\beta_2<0$) at 5% significance. Give
   - the *value* of the test statistic
   - the *critical point* from the appropriate table at the back of your textbook (or compute the *p-value* using a spreadsheet program)
   - your conclusion: whether you reject the null hypothesis at 5% significance.

(10.3) [Hypothesis test] Suppose the relationship between the price of a television and the size of its screen is estimated by least squares using a small random sample of 12 models, with the following results. (The numbers on top are the least-squares estimates of the intercept and slope, and the numbers at the bottom in parentheses are standard errors.) Assume that the error term is normally-distributed.

Price of television	=	8.61 (2.40)	+	6.09 (3.48)	Size of screen in inches

a. According to these results, what is the best unbiased prediction of the price of a television (not in the sample) that has a 19-inch screen?
b. According to these results, if the size of a television screen increases by 4 inches, by how much does the price increase?
c. For purposes of computing hypothesis tests, what are the degrees of freedom for this problem? (Give a number.)
d. Test the null hypothesis that the size of the screen has no effect on price (that is, H_0: $\beta_2=0$) against the two-sided alternative hypothesis that it does have an effect on price (that is, H_1: $\beta_2 \neq 0$) at 5% significance. Give
   - the *value* of the test statistic
   - the *critical points* from the appropriate table at the back of your textbook (or compute the *p-value* using a spreadsheet program)
   - your conclusion: whether you reject the null hypothesis at 5% significance.

(10.4) [Prediction] Suppose we wish to estimate the relationship between the size of a house and its price using a random sample of 22 houses. Let $size_i$ denote the size of each house (in square feet) and let $price_i$ denote the price of the same house (in thousands of dollars). Our purpose is to predict the price of a house with x=1500 square feet, so to simplify calculations, we should transform the data before estimation.

a. Which variable ($price_i$ or $size_i$) should be transformed? How?

Suppose the following equation has been estimated on the *transformed data*. (The numbers on top are the least-squares estimates of the intercept and slope, and the numbers at the bottom in parentheses are standard errors.)

price	=	147.2 (2.7)	+	0.053 (0.011)	size	$\hat{\sigma}^2 = 12.96$

Assume the error term is normally-distributed.

b. Predict the price of a house, not in our sample, with size=1500 square feet.
c. Compute the standard error of prediction error.
d. For the purpose of computing prediction intervals, what are the degrees of freedom for this problem? (Give a number.)
e. Compute a 95 percent prediction interval for the price of a house with size=1500 square feet.
f. Compute a 90 percent prediction interval for the price of a house with size=1500 square feet.

(10.5) [Prediction] The effect of the price of water on daily water consumption per capita is measured using a sample of n=18 cities. For each city i, let y_i denote its water consumption per capita and let x_i denote its price of water. We wish to predict water consumption per capita (y_{n+1}) when the price (x_{n+1}) is 0.04. So we first transform the data.

a. Which variable (x_i or y_i) should be transformed? How?

Suppose the following equation has been estimated on the *transformed data* with the following results. (Numbers on top are the least-squares intercept and slope. Numbers at the bottom in parentheses are standard errors.)

y_i	=	105.0	–	135	x_i
		(2.6)		(42.3)	

The estimated variance of the error term is $\hat{\sigma}^2 = 2.24$. Assume the error term is normally distributed.

b. Predict water consumption per capita (y_{n+1}) when the price (x_{n+1}) is 0.04.
c. For purpose of computing prediction intervals, what are the degrees of freedom for this problem. (Give a number.)
d. Compute the standard error of prediction error.
e. Compute a 95% prediction interval for water consumption when the price (x_{n+1}) is 0.04.

[end of problem set]

