Wilf Family Department of Politics
Quantitative Research in Political Science I, Professor Patrick Egan
Diagnostic Exam

Instructions. This exam should take you up to three hours to complete. You are welcome to use notes and books from previous classes as you need them. Show your work for all problems on separate sheets of paper.

1. The random variable Y has a density function

$$f(y) = \begin{cases} cy, & 0 \leq y \leq 4 \\ 0, & \text{elsewhere} \end{cases}$$

   (a) What value of c would make $f(y)$ a probability density function?

   (b) What is $F(y)$?

   (c) What is the probability that Y takes on a value in the interval $[1, 3]$?

2. 100 residents polled in Country A have a mean income of \$42,000 with a standard deviation of \$1,000. 100 residents polled in Country B have a mean income of \$44,000 with a standard deviation of \$2,400.

   (a) Would you accept the claim that Country B is more affluent than Country A at the 95% confidence level? Be sure to explain fully your choice of a test statistic and rejection region, and provide a diagram showing the rejection region, critical value, and test statistic.

   (b) Provide a **90%** confidence interval for the difference in mean income between the two countries.

3. The cdf of a random variable X is given by

$$F(X) = \frac{e^x}{1 + e^x}.$$

   (a) Is it possible to compute the probability density function of X? If so, what is it?

4. You believe that increases in education are associated with increases in GDP. You also believe that this association is stronger for democracies than non-democracies. In addition to a measure of GDP, you have continuous measures of education (*education*) and democracy (*democracy*), where *democracy* is measured on a scale of 1 to 10 (10 is highest). Assume the relationship between education and GDP is linear.

   (a) Specify an appropriate model.

   (b) Provide an equation for the for the marginal impact of education on GDP.

   (c) You estimate the model with OLS. Provide an expression for your estimate of the marginal impact of education on GDP.

   (d) Provide an expression for the standard error for your answer in (c).
   Note: Your answers to (e) and (f) should be in terms of quantities from your OLS estimates.

(e) For a country with a democracy level of 6, by how much would you expect GDP to increase if education rose from 12 to 16?

(f) Provide a 95% confidence interval for your answer in (e).

(g) You think it likely that there are diminishing returns to education in this relationship. How might you transform your data to account for this?

(h) Does this analysis demonstrate that education leads to increases in GDP? Why or why not?

5. You believe that the true model determining the dropout rate in school districts is given by

$$dropout_i = \beta_0 + \beta_1 perpupil_i + \beta_2 povertyrate_i + \beta_3 classsize_i + u_i,$$

where

- $dropout_i$ is the dropout rate in the ith district as a percentage of students enrolled;
- $perpupil_i$ is the per-pupil expenditure in the ith district (in thousands of dollars);
- $povertyrate_i$ is the percentage of students in the ith district with family incomes below the poverty line; and
- $classsize_i$ is the average class size (in number of students) in the ith district.

You run OLS and find that

$$\widehat{\beta}_0 = 5.78 \quad \sigma_{\widehat{\beta}_0} = 0.96$$
$$\widehat{\beta}_1 = -3.00 \quad \sigma_{\widehat{\beta}_1} = 0.48$$
$$\widehat{\beta}_2 = 1.00 \quad \sigma_{\widehat{\beta}_2} = 0.70$$
$$\widehat{\beta}_3 = 0.40 \quad \sigma_{\widehat{\beta}_3} = 0.84$$

$$R^2 = 0.35$$
$$N = 1000$$

(a) What is a theoretically justified hypothesis about the relationship between the percentage of a district's students below the poverty line and the dropout rate? Use this hypothesis in a test determining whether you are 95% certain that increases in the poverty rate lead to increases in the attrition rate. Include equations for H_0 and H_1.

(b) You find that there is a high correlation between a district's poverty rate and average class size. How might this affect your estimates of β_2 and β_3?

(c) You decide that you would be content to show that the effect of either poverty or class size is significant at the 95% level. Is either of these two variables significant at the 95% level (note this is a question about joint significance)? Write down H_0 and H_1, as well as equations for any test statistic you use. Big hint: You obtain an R^2 of .27 from the bivariate regression
$$dropout_i = \delta_0 + \delta_1 perpupil_i + v_i.$$

(d) Suppose per-pupil expenditure were measured as hundred dollars spent per student (rather than thousand dollars spent). Can you say what $\widehat{\beta}_1$ would be? Can you say what the t-statistic associated with $\widehat{\beta}_1$ would be?

(e) Assuming that the multivariate model specified above is the true model, is β_2 the total effect of poverty on the dropout rate? Why or why not?