Quantitative Methods for Lawyers
Problem Set #2

Professor Daniel Martin Katz
Michigan State University College of Law

March 14, 2012

Instructions: You may work in groups but each student must do their own work and submit his/her own problem set. If you work with others, please denote this fact in the upper right corner of your problem set (including the names of everyone in your group). Please show all of your work (**including Stata output where applicable)** in order to receive credit. **Due at Noon - March 27, 2012** (Note the Time Extension). Please submit a **hard copy** either to me or to Sue Nelson in Room 465. **Retain a photocopy or scanned copy for your records.**

Problem 1 *Worth 20pts*

Imagine you are interested in the pleading strategies of lawyers in a selected form of civil dispute. You are new to the field but have been told by an expert the following information: "let me tell you how things work in this particular area of law – there are some rookie lawyers (lawyers with less than 5 years of experience) that really do not know what they are doing when it comes to this type of a case. They think that they should try to hold onto every single theory for relief. Therefore, they choose at the pleading stage to pursue a large number of theories for recovery (i.e. engage in so called shotgun pleading). At the same time, the older/experienced lawyers choose to focus their attention on a smaller number of legal theories (i.e. engage in so called surgical pleading).

As a rookie lawyer who just been assigned such a case you are tempted to believe the claim of this expert. However, as a matter of due diligence, before pursuing a surgical pleading approach you decide to evaluate this claim by collecting a small sample of filings in the jurisdiction. Imagine you develop an appropriate sampling frame and collect 17 random pleadings. Next, you decide to quickly code those dispute as follows:

Is the Lawyer Older/Experienced (1 if True, 0 if False)?
How Many Theories are Contained in the Pleading?

Using the Dataset Below and the appropriate approach please test the claim offered by the expert. In addition, provide a qualitative interpretation of the results of your analysis.

Case Number	Experience	Number of Theories
Case 1	1	7
Case 2	1	3
Case 3	1	4
Case 4	1	2
Case 5	1	2
Case 6	1	9
Case 7	1	3
Case 8	1	4
Case 9	0	4
Case 10	0	9
Case 11	0	11
Case 12	0	2
Case 13	0	5
Case 14	0	7
Case 15	0	3
Case 16	0	8
Case 17	0	6

Table 1: Experience and Pleading Approach for Sample of 17 Cases

.....

Problem 2 *Worth 10pts*

Please Complete Question #4 on Page 286-287 of Lawless, Robbennolt & Ulen. Please show all of your work.

.....

Problem 3 *Worth 10pts*

Please Complete Question #4 on Page 332 of Lawless, Robbennolt & Ulen. Please show all of your work.

.....

Problem 4 *Worth 10pts*

Please Complete Question #5 on Page 332-333 of Lawless, Robbennolt & Ulen. Please show all of your work.

.....

Problem 5 *Worth 35pts*

This is an opportunity to become more familiar with Stata and extend some of the examples that I have demonstrated in class. Open a Log File and Do File so that you can use it to submit clean output when you turn in your assignment. Next, please access the version of the States.dta file that is on the course webpage. Using that dataset please complete the following tasks:

(a) Use the Describe Command and Review the Variable Labels.
Please Identify the variable label for "Waste"?

(b) Use Stata to Calculate the Mean, Standard Deviation and Kurtosis for "Area"?

(c) Generate a Scatter Plot of "Pop" on "Csat" and plot a regression line on that Scatter Plot. Ensure that the regression line starts at the intercept and carries through full range of X values. Please describe the nature of this relationship in words (which specific discussion of the correlation).

(d) Assume that your are interested in Student Scores on the Verbal Portion of the SAT. Specifically, you are interested in the Score on the **VERBAL** Portion as a Dependent Variable and Expense, Percent, Income, High, College and Region as Independent Variables. Please Run this Regression with Robust Standard Errors. Next, please provide an interpretation of the statistical significance of each of the independent variables.

(e) Write out the full regression equation with the proper Beta Coefficients and then use that Equation to Predict the Expected Mean Verbal SAT in a hypothetical state with the following attributes: Per Pupil Expenditures Primary & Secondary is $8000, Percent of HS of graduates taking SAT is 28%, Median Household Income is $38,000, Percent of Adults with HS Diploma is 80%, the Percentage of Adults with College Degree is 33% and the Region is the Northeast.

(f) You wonder whether the Population of a State might in some way contribute to average performance on the verbal portion of the SAT. Please generate a scatter plot for these two variables and interpret the resulting relationship.

(g) Please include this variable in your overall regression model. Please interpret the resulting model and identify whether population size has a statistically significant relationship on composite sat score (i.e the csat variable). In light of your findings should Pop be included in future runs of the SAT prediction model?

.....

Problem 6 *Worth 15pts*

Now we will work with a new dataset. This is a dataset of housing prices from 1981. We are going to use this to develop a rough prediction model using this data. Open a Log File and Do File so that you can use it to submit clean output when you turn in your assignment. Next, please access version of the hprice1981.dta file that is on the course webpage. Using that dataset please complete the following tasks:

(a) Use the Describe Command and Review the Variable Labels.
Please Identify the variable label for "Area"?

(b) Generate a Scatter Plot of "Area" on "Price" and plot a regression line on that Scatter Plot. Please describe the nature of this relationship in words (which specific discussion of the correlation).

(c) Assume that your are interested in how various factors predict the ultimate selling price of a house. Specifically, you are interested in the Price as a Dependent Variable and Number of Bedrooms, Number of Bathrooms, the Age of the House, the Number of Square Feet of the Home and the Size of the Lot as Independent Variables. Please Run this Regression with Robust Standard Errors. Please describe in words the sign (+ or -) on each of these independent variables and provide a qualitative account for why the sign might be in that particular direction. Finally, please provide an interpretation of the statistical significance of each of the independent variables.

(d) Please Write out the full regression equation with the proper Beta Coefficients and then use that Equation to Predict the Expected House Price for a hypothetical home with the following attributes: 4 Bedrooms, 2 Bathrooms, 35 Years Old, 1700 Square Feet and a Lot Size of 18,000 Square Feet.

.....