Problem Set 4 — Political Science 599

Due Thursday, 27 October 2005

Instructions: Please type this problem set single spaced. When presenting tables, do not cut and paste from statistical programs directly into your word processor. Make real tables with sensible significant digits. Embed any figures or tables into the text. Make equations, when necessary, that look nice, and that are well explained in terms of the symbols and variables used. Use numbers for equations, figures, and tables. If you have questions about how this should work, look at a recent copy of the APSR. In responding to the questions use your own words rather than quoting from books. If you quote from books, explain the quotes in your own words. Include an appendix at the end of the problem set containing the code you used to read, manipulate, and analyze the datasets. The appendix should have separate sections for each question. Someone should be able to "run" a given section of your appendix to produce the results for a given question. You may hand-write (neatly) the solutions to the math problems or use an equation editor or LaTeX. And of course, show your work.

1. For each of the following two articles, precisely what inferences does the author draw from precisely what numbers? Evaluate these inferences. Include measurement in your evaluation. Give precise reasons for your evaluation. Do all of this work using prose that others in the profession can easily understand.

   (a) Samuel J. Eldersveld. 1956. "Experimental Propaganda Techniques and Voting Behavior," *APSR* 50:1, p 154-165.

   (b) Robert E. Lane. 1955. "Political Personality and Electoral Choice," *APSR* 59:1, p 173-190

2. (a) What is maximum-likelihood estimation? (b) What is the relationship between probability and likelihood? (c) How exactly do you find a maximum-likelihood estimate? (Explain each step so that readers who are completely unfamiliar with the concept will understand). (d) Why would someone prefer a maximum-likelihood estimator to other kinds of estimators? Under what circumstances might you prefer to avoid a maximum-likelihood estimator?

3. The following table shows data on the length of time (in weeks) that Italian cabinets have lasted from about 1946 to about 1987.[1]

Weeks	2	8	12	16	20	24	28	32	40	44	48	52	56	60	64	72	76	88	92	108	180
Number of Cabinets	4	1	4	1	4	2	6	2	1	1	2	2	2	1	1	2	1	1	1	1	1

   Cioffi-Revilla (1984) uses these data for the 1946-1980 period to show that Italian governments end at a constant rate of .021 per week (i.e the mean length of an Italian cabinet from 1946 to 1980 was $1/.021 \approx 48$ weeks). Your job is to use maximum likelihood estimation to see if his estimated rate of government failure is still reasonable given the addition of 7 more years to the dataset. So, you'll need to first roughly calculate the maximum of the likelihood function using numerical approximations (you may find it useful to plot your approximations to eyeball the maximum), second using calculus, and third using a computer program that you write. Plot the exponential model you've just fitted against the observed data. Just from looking at the graph of observed versus predicted values, do you think the exponential is a reasonable model for government termination in Italy? What does your estimate mean when it comes to Italian government survival? Does your assessment change depending on the way you set up the bins in your histogram of the observed data? What is the confidence interval for your estimate? Given your confidence interval, do you think that Cioffi-Revilla's estimate of .021 per week is still reasonable to use for this expanded dataset? [2]

4. What is a confidence interval? Give an example and talk the reader through it.

5. Work 3 of the following 6 problems from Wonnacott and Wonnacott. Show all of your work. Give reasons for your conclusions in such a way that mere mortals who've not taken this class will understand. 18-12,8-19,8-20,8-32,9-6,9-8.

---

[1]So, 4 cabinets only lasted 2 weeks, 1 cabinet lasted 8 weeks, 4 cabinets lasted 12 weeks, and so on. These data were provided for replication of King, Alt, Burns, and Laver (1990).

[2]Hint: Use the results from the numerical approximation, calculus, and Cioffi-Revilla to guide your choice of starting values. It is standard practice to use a few different starting values, just to make sure that the computer is finding a global rather than a local maximum/minimum. This exponential with one parameter is globally concave & reasonably smooth, though, so it matters here only for how long it might take to find the maximum.

6. Let $\mathbf{X}$ be an $n \times k$ matrix such that $\mathbf{X}^T\mathbf{X} \neq \mathbf{0}$. Let $\mathbf{Q} = \mathbf{X}^T\mathbf{X}$, $\mathbf{A} = \mathbf{Q}^{-1}\mathbf{X}^T$, $\mathbf{N} = \mathbf{XA} = \mathbf{XQ}^{-1}\mathbf{X}^T$, and $\mathbf{M} = \mathbf{I} - \mathbf{N}$, $\mathbf{M} = \mathbf{M}^T = \mathbf{MM}$, $\mathbf{N} = \mathbf{N}^T = \mathbf{NN}$, and $\mathbf{AX} = \mathbf{I}$. Also let $\mathbf{b} = \mathbf{Ay}$, $\hat{\mathbf{y}} = \mathbf{Ny}$, and $\mathbf{e} = \mathbf{My}$ be vectors (symbolized by the lower case letters).[3]

   Prove any 7 of the 14 following, as concisely as you can.

   (a.) $\mathbf{AN} = \mathbf{A}$
   (b.) $\mathbf{AM} = \mathbf{0}$
   (c.) $\mathbf{MN} = \mathbf{0}$
   (d.) $\mathbf{A}\hat{\mathbf{y}} = \mathbf{b}$
   (e.) $\mathbf{Ae} = \mathbf{0}$
   (f.) $\mathbf{X}^T\hat{\mathbf{y}} = \mathbf{X}^T\mathbf{y}$
   (g.) $\mathbf{N}\hat{\mathbf{y}} = \hat{\mathbf{y}}$
   (h.) $\mathbf{M}\hat{\mathbf{y}} = \mathbf{0}$
   (i.) $\mathbf{X}^T\mathbf{e} = \mathbf{0}$
   (j.) $\mathbf{Ne} = \mathbf{0}$
   (k.) $\mathbf{Me} = \mathbf{e}$
   (l.) $\hat{\mathbf{y}}^T\hat{\mathbf{y}} = \mathbf{y}^T\mathbf{Ny} = \mathbf{y}^T\mathbf{Xb} = \mathbf{b}^T\mathbf{X}^T\mathbf{y} = \mathbf{b}^T\mathbf{Qb}$
   (m.) $\mathbf{e}^T\mathbf{e} = \mathbf{y}^T\mathbf{My}$
   (n.) $\hat{\mathbf{y}}^T\hat{\mathbf{y}} = \mathbf{y}^T\mathbf{y} - \mathbf{e}^T\mathbf{e}$

---

[3]These vectors are actually vectors of coefficients, fitted values, and residuals resulting from the regression of $\mathbf{y}_{n \times 1}$ on the matrix $\mathbf{X}_{n \times k}$, though you don't need to know that to do the problem.