EXAMINATION:
QUANTITATIVE EMPIRICAL METHODS

Yale University

Department of Political Science

January 2014

You have seven hours (and fifteen minutes) to complete the exam. You can use the points assigned to each question as a (rough) guide to allocation of your time. The exam is 100 points long, which is approximately 15 points per hour. This exam consists of two parts.

Part One (9:00am - 2:15pm; 70 points)
Part One will be distributed at 9am. Your answers for Part One will be collected at 2:15pm. The only aids permitted for Part One are (1) one page of double-sided notes, (2) a calculator, and (3) a word processor on one of the Statlab computers to write up your answers (you may also write up answers using pencil/pen and paper). At 2:15 all your answers for Part One must be submitted to the proctor. Hold on to the instructions for Part One of the exam.

Advice: Use algebra to back up your assertions wherever appropriate, and remember to show your work. Do not leave sub-parts of questions unanswered. Your answers should be succinct and to the point.

Part Two (2:30pm - 4:30pm; 30 points)
Part Two of the exam will be distributed at 2:30pm. Your answers will be collected at 4:30pm. Part Two may involve using statistical software. A complete answer to Part Two will include the code and output, as well as your written answers. For Part Two you are permitted (1) unrestricted use of your own computer with access to the internet or (2) use of a Statlab computer with access to the internet. The only restriction for Part Two is that you may not interact with anyone, online or otherwise. In addition, you must credit in your answer any sources (code or references) that help you.

Advice: Explain what you are trying to do in comments. Even if you are not able to execute your program correctly, you can receive partial grades for explaining clearly what you wanted to do and why. We recommend you use a program such as *knitr* (see here) for R or *log* for Stata to easily record the input and output from your analysis.

Part I

9:00-2:15. (70 points)

Rules: No references or aids permitted except one page of notes and a calculator. Answers may be written up by hand or word processor.

1 30 minutes, 7 points

Suppose some treatment (W) is thought to have an effect on some outcome (Y). The treatment is randomly assigned with $p = 0.5$. Observations are independent. Consider three possible ways of testing the null hypothesis of no treatment effect for every unit.
(1) A (one-sample) Student's t-test.
(2) A Wilcoxon signed rank test.

1.1

State the conditions under which you would recommend (1) over (2). Explain precisely why this is the case. You may want to give an example. Similarly, state the conditions under which you would recommend (2) over (1). Explain precisely why this is the case. You may want to give an example.

1.2

Now consider (3) a permutation test using a (one-sample) Student's t-test as the test statistic.
State the conditions under which you would recommend (3) over (1).
State the conditions under which you would recommend (3) over (2).

2 30 minutes, 8 points

A scholar is interested in the determinants of registration for elections. Specifically, the scholar is interested in the effects of an educational intervention to encourage registration. The scholar randomly selected $N = 10000$ citizens, and then randomly assigned this intervention to $m < N$ of them. Denote as $W_i \in \{0, 1\}$ the intervention for citizen i, $Y_i^t \in \{0, 1\}$ whether the citizen registered within t years since the intervention took place. The scholar measures the outcome at t =1 and $t = 5$ years after the intervention. Because of the nature of election registration, $Y_i^1 = 1 \implies Y_i^5 = 1$. There is no missing data or measurement error. Denote the set of individuals who receive treatment as $\mathbf{T}$ and control as $\mathbf{C}$. The scholar estimates the difference in means of the outcome one year after the intervention:

$$\hat{\delta_1} = \frac{\sum_{i \in \mathbf{T}} Y_i^1}{\sum_{i \in \mathbf{T}} \mathbf{1}} - \frac{\sum_{j \in \mathbf{C}} Y_j^1}{\sum_{j \in \mathbf{C}} \mathbf{1}} = 0.1$$

(The denominator is just a sum of 1s for each element of the set.)

The scholar then estimates the difference in means of the outcome five years after the intervention for those citizens who didn't register by $t = 1$ ($Y_i^1 = 0$). Denote the set of individuals who received treatment and didn't register in $t = 1$ as $\mathbf{T}'$ and those who received control and didn't register in $t = 1$ as $\mathbf{C}'$.

$$\hat{\delta_5} = \frac{\sum_{i \in \mathbf{T}'} Y_i^5}{\sum_{i \in \mathbf{T}'} \mathbf{1}} - \frac{\sum_{j \in \mathbf{C}'} Y_j^5}{\sum_{j \in \mathbf{C}'} \mathbf{1}} = -0.03$$

Both of these estimates are highly significant from a permutation test.

From this analysis, the scholar concludes:

> $\hat{\delta_1}$ provides an unbiased and consistent estimate of the effect of the intervention on registration within one year of the intervention. Since it is not possible to register again, or to deregister, we know that the treatment effect δ_5 is 0 for all individuals who had registered by $t = 1$; therefore in estimating the long term effect I exclude individuals who had registered by $t = 1$. $\hat{\delta_5}$ therefore provides an unbiased and consistent estimate of the effect of the intervention on registration between one to five years after the intervention. From this, it is apparent that this intervention has a significant positive short-term effect on registration, as well as a significant negative long term effect; the long term effect is weaker, however, than the short term effect. This is consistent with my theory that while educational interventions remind individuals to register for elections, in the long run it makes them more cynical about the electoral process and discourages them from registering. Since the estimate of the short-term effect is about three times larger than the long-term effect, the net effect of this intervention is probably positive.

Evaluate this scholar's conclusion. As usual, be careful and precise. If there is a flaw in the scholar's reasoning, explain the nature of this flaw formally and thoroughly.

3 60 minutes, 14 points

Suppose three regions are observed over the course of one year. One anti-government protest occurs in one of the three regions. No anti-government protests occur in the remaining regions. Thus, the data look like this: (0,0,1).

A. Suppose we model the probability of a region experiencing an anti-government protest as a binomial process. Call the probability that at least one event occurs p. State the assumptions about the DGP that justify modeling this process as binomial and whether these assumptions might be plausible. Write down the likelihood function. Guess the MLE, and then prove that this is correct. Graph the likelihood function for different values of p (five values is sufficient).

B. Now suppose we model the number of anti-government protests experienced by each region as being generated by a Poisson process. Call the 'intensity' parameter λ. State the assumptions about the DGP that justify modeling this process as Poisson and whether these assumptions might be plausible. Write down the likelihood function. Guess the MLE, and then prove that this is correct. Graph the likelihood function.

C. Now suppose we model the number of anti-government protests recorded in each region as following a normal process with mean of μ and variance of σ^2. State the assumptions about the DGP that justify modeling this process as normal and whether these assumptions might be plausible. Write down the likelihood function; as a reminder, the pdf of the Normal distribution is:

$$f(x|\mu,\sigma^2) = \frac{e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$$

Guess the MLE, and then prove that this is correct. Explain and/or sketch roughly what the likelihood function looks like.

D. Which model is theoretically most appealing for this application: binomial, Poisson, or Normal?

4 30 minutes, 8 points

How does random treatment assignment help scientists to draw reliable causal inference?

There are a number of non-redundant answers to this question. We are looking for all of them. In addition, explain how random treatment assignment assists a practitioner of Bayesian inference and Bayesian statistics to draw reliable causal inference.

5 30 minutes, 8 points

Suppose that a researcher runs an experiment in which X and Z are independently randomly assigned, Y is an outcome variable, X, Z, and Y are all continuous variables, and ϵ is all unobserved determinants of Y.

5.1

Suppose that a researcher then estimates the parameters in the following linear model by OLS:

$$Y_i = \alpha + \beta_1 X_i + \beta_2 Z_i + \beta_3 X_i Z_i + \epsilon_i$$

Explain what the four parameters of this model mean in the context of this regression. How would you evaluate if an increase in Z increases the causal effect of X? Be careful and precise in your answer, including stating any assumptions you make.

5.2

Now suppose Y is measured as a dichotomous variable. The researcher then estimates the parameters in the following model using MLE:

$$P(Y_i = 1|M) = \frac{e^M}{1+e^M}$$

$$M = \alpha + \beta_1 X_i + \beta_2 Z_i + \beta_3 X_i Z_i$$

What is this kind of model called? Explain what the four parameters of this model mean. How would you evaluate if an increase in Z increases the effect of X? Be careful and precise in your answer, including stating any assumptions you make.

6 2 hours, 25 points

Skim and evaluate the following article as if you were a (rushed) reviewer who has been asked by the editor to evaluate its methodological merit. Your review should focus on methodological issues. (The article should be attached to the back of the exam.)

A good answer will be an insightful, critical, and constructive review that offers a global evaluation of the work. Highlight the methodological strengths of the paper, for example by pointing out what problems they are able to overcome through their design or methods. Then identify potential (or actual flaws), discuss them in order of importance, evaluate the seriousness of each of them, and suggest *practical* ways of overcoming these potential flaws. A good answer will concisely summarize the main methodological *strengths and weaknesses* of the paper, will connect the work to other relevant literature, and will offer practical and detailed recommendations. Make your advice as precise and formal as possible; for example, if you propose a new design or model you should clearly state the required assumptions.

"The Violence We Do Not See: Reporting Bias in Conflict Event Data".
(Included on this exam with permission from the author.)
https://www.dropbox.com/s/vj9pvf6d16lbfgw/article1.pdf
Abstract: Scholars of civil war rarely collect data on violence themselves, and instead rely on other sources of information. One frequently-used source is media reports, which serve as the basis for many ongoing data projects in the discipline. However, media reports may be selective in the sense that they cover some events, but not others. For example, media accounts could systematically miss smaller events, since they are of little interest to a global audience. By linking a media-based dataset and one based on military records, this paper presents a systematic assessment of selective reporting of violent incidents. The paper starts by analyzing the effects of selective reporting using artificial data. This exercise shows that under certain conditions, selection can attenuate, amplify or even reverse the true effects when estimated on a subset of media- reported events. These effects are then examined on a real case. Using event data from Afghanistan on the real events on the ground, the analysis shows that selection can strongly bias the effects we estimate on a media-reported subsample, primarily when reporting is influenced by one of the independent variables that also affect violence. The results also show, however, that selective reporting is too weak to inflate or revert the effects we estimate on media-based violence data, and the real problem that remains is attenuation.

Part II

2:30-4:30. (30 points)

Randomization/permutation inference is typically used for testing the sharp null hypothesis of zero treatment effect for all units. However, some scholars are not interested in evaluating such a null hypothesis because they find it implausible. In social systems, almost everything has an effect on everything else (that is later in time), all be it often very small (Andy Gelman: "the hypothesis of zero effect is almost never true!").

This question will ask you to evaluate whether, and when, tests of the sharp null serve as a conservative test of a set of null hypotheses involving heterogeneous effects.

Denote the potential outcome for unit i under treatment as $Y_i(1)$ and under control as $Y_i(0)$. Denote $\delta_i = Y_i(1) - Y_i(0)$. The set of all units is $\mathbf{N}$. The sharp null (denoted $s0$) of a constant treatment effect is then

$$H_{s0} : \delta_i = k \quad \forall i \in \mathbf{N}$$

with the alternative hypothesis being that the effect is weakly greater than k for every unit and greater than k for at least one unit

$$H_a : \delta_i \geq k \quad \forall i \in \mathbf{N}; \quad \exists j \in \mathbf{N} \text{ such that } \delta_j > k$$

The general null hypothesis (denoted $g0$) is that the treatment effect is weakly less than k for each unit

$$H_{g0} : \delta_i \leq k \quad \forall i \in \mathbf{N}$$

A hypothesis test is *conservative* if, for any nominal significance level α, the true probability of incorrectly rejecting the null hypothesis is less than or equal to α.

Conjecture 1. A permutation test of H_{s0}, with H_a as the alternative, is a conservative test of H_{g0}.

(1) **Evaluate Conjecture 1.** You may do so using whatever method you find most appropriate, though we recommend you use simulations as part of your answer. If the conjecture is correct, you should try to demonstrate this (ideally prove it). If the conjecture is incorrect, you should explain why, ideally with one or more counter-examples; if possible, you should also refine the claim so that it is correct and demonstrate why your refined claim is correct.

(2) Explain how this conjecture matters for the use of permutation inference.

Advice: If you are having trouble getting started, you might want to begin by considering specific examples of these hypotheses, writing simulations to produce data consistent with them, and then evaluating them using permutation inference. Doing so will allow you to evaluate the conjecture for specific null hypotheses, perhaps leading you to an intuition to explain why the conjecture is correct, or a counter-example to demonstrate why it is not.