Applications and Midterm Exam Review

November 9, 2022

Today

- Brief review of statistical inference and hypothesis testing
- Preparation for second midterm
- Discussion of an application article (a quasi-experimental study using difference-of-means tests)

Statistical Inference

- Interested in population, but have data on a sample
- But, if a random sample, we know what a distribution of samples would look like
- Can determine spread of sampling distribution based on sample
  - Standard error—affected by variance of sample and square root of sample size
- Two approaches:
  - Assume test statistic is center of distribution
    - Then can examine confidence interval around test statistic
  - Assume null hypothesis is true
    - Can examine confidence interval around null value & identify the p-value

Hypothesis Testing

- Review of approach:
  - Our hypothesis (H_a)—generally a prediction about relationship between variables in population
  - Null hypothesis (H_0)—generally a prediction of "no effect," i.e., that IV does not have an effect on DV
  - Assume null hypothesis is true
  - Examine sampling distribution around test statistic
  - How likely would we be to get test statistic if null hypothesis were true?
  - Allows us to say how confidently we can reject null hypothesis, and thus the level of statistical significance we find

Means and Proportions

- One sample t-test
  - Test specific prediction about value of mean
- Ex.—Hypothesis: The average Norwegian adult male is taller than 1.7 meters
  - What's the null hypothesis?
  - If null is correct, what's the center of the sampling distribution?
  - I take a SRS of 225 Norwegian adult males and find that the average height in the sample is 1.73 meters with a standard deviation of 20 centimeters. Can I reject the null hypothesis with 95% confidence?

Means and Proportions

- Two-sample t-test
  - Test whether there is a difference between two means
- Ex—Hypothesis: Norwegian men have a higher average height than Swedish men
  - What's the null hypothesis?
  - Random sample of 100 Norwegians and 100 Swedes
  - $\bar{x}_{Norwegians}$: 1.81 meters, $s_{Norwegians}$: 0.1 meters
  - $\bar{x}_{Swedes}$: 1.79 meters, $s_{Norwegians}$: 0.08 meters
  - Can I reject the null hypothesis with 95% confidence?

Means and Proportions

- One-sample proportion test
  - Test specific prediction about proportion in population
- Ex—Hypothesis: proportion of Maryland drivers who are bad drivers is greater than 0.3
  - What's the null hypothesis?
  - I take a sample of 100 Maryland drivers, find that 0.38 of them are bad drivers.
    - What's the p-value?

Means and Proportions

- Two-sample proportion test
  - Test whether there is a difference in proportions
- Ex.—Hypothesis: Proportion of Maryland drivers who are bad drivers is lower than the proportion of Massachusetts drivers who are bad drivers
  - What's the null hypothesis?
  - I take a random sample of 100 Maryland drivers and find that 0.35 of them are bad drivers and take a random sample of 100 Massachusetts drivers and find that 0.4 of them are bad drivers
  - What is the p-value?

Cross-Tab and Chi-Square

- Ha: In comparing individuals, republicans are more likely to be morning people than independents, who are more likely than democrats
- IV: Party-id (self-identified)
- DV: Morning people, afternoon people, evening people
- In a SRS, of morning people, 78 are Republicans, 53 Independent, and 95 Democrats
- Of afternoon people: 27 are Republicans, 25 Independent, and 34 Democrats
- Of night people: 55 Republican, 62 Independent, 111 Democrats
- Make a cross-tab, find chi-square, and give me a range of the p-value

What We Have Not Discussed

- What to do with a very small sample
  - T-statistic is not robust to (very) small sample sizes
  - If your sample size is very small, use binomial (not normal) distribution to make inferences about p
  - If expected cell counts are very small for any cells, chi-square test not appropriate
- Punchline—if you have a very small sample (say, less than 30-40), either
  - Go get more data
  - Learn how to analyze it (you won't learn it in this class)

Control Variables

- We have examined three principal approaches: difference of means, difference of proportions, & chi-square
- Each examines the effect of some IV on a DV
- But, remember the danger of confounding variables
- Can incorporate control variables into these analyses, but is often unwieldy
  - Essentially split the sample

Control Variables

- In experimental design, with random assignment to control/treatment groups, these methods can be very powerful
  - Example—proportion of cancer patients surviving 5 years receiving treatment compared to placebo
- If observational research, can tell us something about differences, but want to think about control variables for hypothesis testing
- Regression allows for easy incorporation of control variables
  - Will discuss multiple linear regression after second exam

Exam 2

- Same format as last exam:
  - Two sections—conceptual & applied questions
- Same process:
  - Exam becomes available after lab session on Friday
  - Due by Monday at 11:59 p.m.
  - Open note, but you may not consult anyone else

What you should know

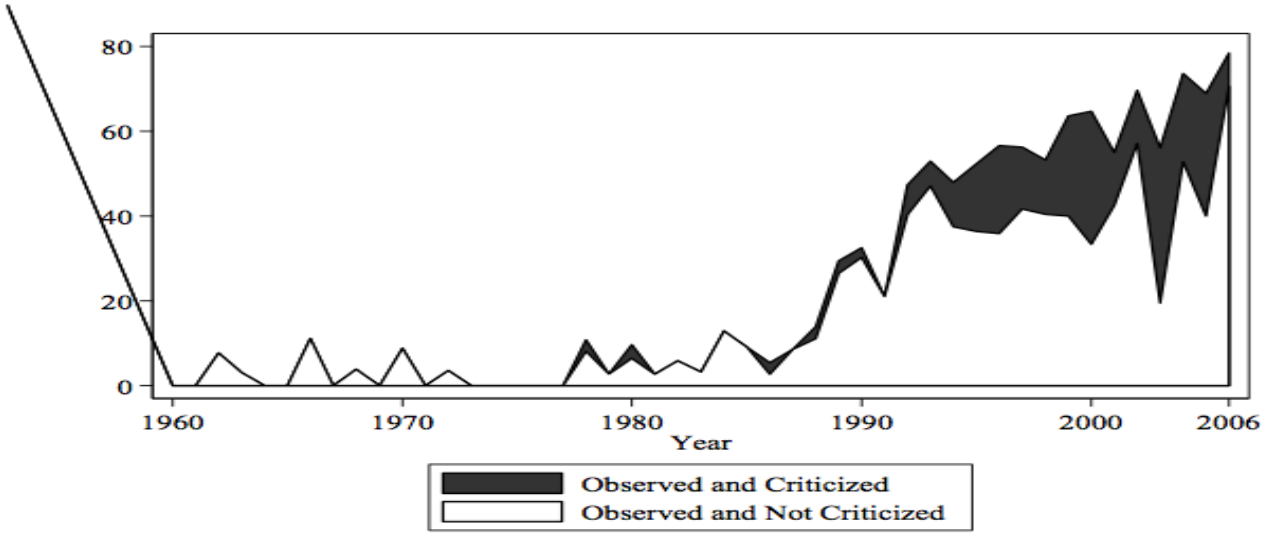
- Basic logic of probability theory and sampling distributions
  - E.g., Expect probability questions/problems to solve
- How to compute a standard error for a mean and a proportion
- How to find confidence intervals, p-values, t-statistics, and z-scores
- The logic of hypothesis testing and statistical significance
- How to interpret a difference of means and difference of proportions test
- Cross-tabs & chi-square statistics

Application—Election Monitoring and Fraud

- Election monitoring has become a "norm"
  - But, does election monitoring matter?
- Is there less fraud in elections that are monitored than those that are not?
- Study argues yes—election monitoring can deter fraud at observed polling places
  - Hyde, Susan D. (2007), "The Observer Effect in International Politics: Evidence from a Natural Experiment." *World Politics* 60(1): 37-63.
- Hypothesis: In comparing polling stations, those that are monitored by international observers will have less fraud than those that are not monitored by international observers

International Election Monitoring

Figure 1: Rate of Internationally Observed Elections, 1960-2006



Note: Includes 1754 election events in 157 independent states, excluding those with population < 250,000
Source: Author

Alternative Explanations

- One difficulty in assessing effect of election monitoring is controlling for alternative explanations
  - Election monitors may go to easier to reach polling stations
  - Fraud may be committed at harder to reach stations
- May mean we observe no relationship between monitoring and fraud
- Hyde (2007)'s approach:
  - Quasi-experimental design
  - Natural Experiment—Election monitors assigned essentially randomly in Armenian election of 2003
    - No advance warning given
  - Since assignment essentially random, no worry of alternative explanations

Research Design

- Unit of Analysis:
  - Polling Station
- Variables:
  - I.V.—Presence (or absence) of international election monitors
  - D.V.—Level of electoral fraud
- Predicted effect—lower levels of electoral fraud in polling stations with presence of election monitors
- How to measure fraud?
  - Hyde's measure—percentage of the vote won by the incumbent
  - Logic—incumbents more able to steal election, if fraud is lessened, incumbent should get less of the vote share

The 2003 Elections

- Two rounds of elections:
  - First round, several candidates for President on February 19
  - Second round, top two candidates on March 5
- Descriptive statistics:
  - 755 polling stations never monitored
  - 385 polling stations only monitored in 1st round
  - 260 polling stations only monitored in 2nd round
  - 363 polling stations monitored in both rounds
- Two rounds allows for examining if there was a lingering effect of monitoring
- Compare mean vote share for incumbent in polling stations based on presence of monitors

Results

TABLE 1
DIFFERENCE OF MEANS TESTS COMPARING "TREATMENT" AND "CONTROL" GROUPS*

Average Incumbent Vote Share among Polling Stations That Were. . .	*vs.*	*Average Incumbent Vote Share among Polling Stations That Were. . .*	*Difference*
1. Not observed in R1 54.2% (R1 vote share)	vs.	Observed in R1 48.3% (R1 vote share)	5.9% t(1762)=5.92 P>|t|=0.00
2. Not observed in R2 69.3% (R2 vote share)	vs.	Observed in R2 67.3% (R2 vote share)	2.0% t(1761)=2.47 P>|t|=0.014
3. Never observed 70.7% (R2 vote share)	vs.	Observed in both R1 and R2 66.2% (R2 vote share)	4.5% t(1116)=4.48 P>|t|=0.00
4. Never observed 62.8% (Average of R1 and R2 vote share)	vs.	Observed in both R1 and R2 57% (Average of R1 and R2 vote share)	5.8% t(1116)=5.36 P>|t|=0.00
5. Never observed 62.7% (Average of R1 and R2 vote share)	vs.	Observed in one or both rounds 58.1% (Average of R1 and R2 vote share)	4.6% t(1761)=5.65 P>|t|=0.00
6. Never observed 70.7% (R2 vote share)	vs.	Observed only in R1 66.3% (R2 vote share)	4.4% t(1138)=4.40 P>|t|=0.00
7. Never observed 70.7% (R2 vote share)	vs.	Observed only in R2 68.7% (R2 vote share)	2.0% t(1013)=1.73 P>|t|=0.084
8. Observed only in R2 68.7% (R2 vote share)	vs.	Observed in both R1 and R2 66.2% (R2 vote share)	2.5% t(621)=1.93 P>|t|=0.054
9. Observed in both R1 and R2 66.3% (R2 vote share)	vs.	Observed only in R1 66.2% (R2 vote share)	.11% t(746)=0.094 P>|t|=0.93
10. Observed only in R1 68.7% (R2 vote share)	vs.	Observed only in R2 66.3% (R2 vote share)	2.4% t(643)=1.83 P>|t|=0.067

* Reported results reflect two-sample t-tests with equal variances.

Rejecting the Null Hypothesis

- Ha: In comparing polling stations, those with international election monitors will have higher levels of fraud (measured as a higher percentage of votes for the incumbent) than those without monitors
- H0: In comparing polling stations, those with international election monitors will have the same level of fraud (measured as a higher percentage of votes for the incumbent) than those without monitors
- We have a sample that shows a difference between the vote share of incumbent between polling stations:
  - Not Observed (Round 1): 54.2%
  - Observed (Round 1): 48.3%
- Could this difference be due to random sampling error?

Differences between Polling Stations

- Not Observed (Round 1): 54.2%
- Observed (Round 1): 48.3%
- Difference: 5.9%
- Standard Error of the Difference: 1
- This means that the Difference is almost 6 times the S.E. of the Difference
- If there were no actual difference between monitored and unmonitored polling stations, it is extremely unlikely (less than 0.001% chance) we would get a difference like we find in our sample
- So, can reject the null hypothesis with more than 95% confidence

Results

TABLE 1
DIFFERENCE OF MEANS TESTS COMPARING "TREATMENT" AND "CONTROL" GROUPS*

Average Incumbent Vote Share among Polling Stations That Were. . .	*vs.*	*Average Incumbent Vote Share among Polling Stations That Were. . .*	*Difference*
1. Not observed in R1 54.2% (R1 vote share)	vs.	Observed in R1 48.3% (R1 vote share)	5.9% t(1762)=5.92 P>|t|=0.00
2. Not observed in R2 69.3% (R2 vote share)	vs.	Observed in R2 67.3% (R2 vote share)	2.0% t(1761)=2.47 P>|t|=0.014
3. Never observed 70.7% (R2 vote share)	vs.	Observed in both R1 and R2 66.2% (R2 vote share)	4.5% t(1116)=4.48 P>|t|=0.00
4. Never observed 62.8% (Average of R1 and R2 vote share)	vs.	Observed in both R1 and R2 57% (Average of R1 and R2 vote share)	5.8% t(1116)=5.36 P>|t|=0.00
5. Never observed 62.7% (Average of R1 and R2 vote share)	vs.	Observed in one or both rounds 58.1% (Average of R1 and R2 vote share)	4.6% t(1761)=5.65 P>|t|=0.00
6. Never observed 70.7% (R2 vote share)	vs.	Observed only in R1 66.3% (R2 vote share)	4.4% t(1138)=4.40 P>|t|=0.00
7. Never observed 70.7% (R2 vote share)	vs.	Observed only in R2 68.7% (R2 vote share)	2.0% t(1013)=1.73 P>|t|=0.084
8. Observed only in R2 68.7% (R2 vote share)	vs.	Observed in both R1 and R2 66.2% (R2 vote share)	2.5% t(621)=1.93 P>|t|=0.054
9. Observed in both R1 and R2 66.3% (R2 vote share)	vs.	Observed only in R1 66.2% (R2 vote share)	.11% t(746)=0.094 P>|t|=0.93
10. Observed only in R1 68.7% (R2 vote share)	vs.	Observed only in R2 66.3% (R2 vote share)	2.4% t(643)=1.83 P>|t|=0.067

* Reported results reflect two-sample t-tests with equal variances.

From Measure to Concept

- Hyde demonstrates that the incumbent president received a lower vote share in polling stations that were monitored than in those not monitored
  - And finds a "lingering effect" on polling stations in R2 that were monitored in R1
- Is this evidence that election monitors deter fraud?
- Concept—fraud
- Measure—vote share of the incumbent

Control Variables

- This difference of means test looks at difference between monitored and unmonitored polling stations
  - No control variables
  - Hyde does do a regression controlling for difficulty of access to polling places, rural/urban distinction, and whether the polling station is in Nagorno-Karabakh
- Random assignment of polling stations to monitoring means there should be no problem of lurking variables
- So, any difference between polling stations should only be due to random sampling error and effect of election monitoring