

-- No calculators, phones, or other devices permitted. --

First Name and Last Name \_\_\_\_\_

**This practice exam focuses on material for Exam 4, but material eligible for Exam 4 includes material from prior exams and from online notes for Chapters 12 and earlier.**

**Note: Exam 4 will not have as many items as this practice exam has.**

*[Directions for the actual exam are below, not this practice exam]*

1. Write your first and last name on this page.
2. Write your responses under the indicated item on the following pages. See the instructor if you need additional paper.
3. No need for your responses to be in complete sentences: feel free to write enough to signal to me that you understand the correct response.
4. Responses that the grader cannot read might not be marked as correct.
5. Each item will be worth up to 4 points.

## **POL 138-003 Quantitative Reasoning in Political Science**

### **KEY Practice Exam 4 · Fall 2025**

1. Write the specific definition of a p-value in terms of the probability that is defined by the p-value. Include the three major elements of the definition.

**The p-value indicates the probability that an outcome at least as extreme as the observed outcome would occur if the null hypothesis were true.**

-- No calculators, phones, or other devices permitted. --

2. Explain what information a p-value provides.

**A p-value is a measure of the strength of evidence that an analysis has provided against the null hypothesis. You can also get credit for a more technical definition, such as "the probability of obtaining the observed results or more extreme results if the null hypothesis were true". Note, though, that the p-value is not the probability that the null hypothesis is true.**

-- No calculators, phones, or other devices permitted. --

3. Suppose that an analysis indicated that, in elections that have two candidates, the candidate who raises more money wins a higher percentage of the elections. Suppose also that the p-value is  $p < 0.05$  for a test of the null hypothesis that the candidate who raised more money was equally likely to win an election as the candidate who raised less money was to win the election. Explain whether this would be sufficient evidence at the conventional level of political science to conclude that raising more money than the other candidate causes a candidate to be more likely to win an election.

**Not sufficient evidence, because alternate explanations should be addressed. For example, maybe the candidate who raises more money is more likely to be an incumbent, and this incumbency causes the candidate to be more likely to win an election.**

-- No calculators, phones, or other devices permitted. --

4. Suppose that researchers are interested in the extent to which the number of representatives in a legislature influences the number of laws that the legislature passes. Suppose that, in Freedonia at the start of the year 2021, the size of the legislature was reduced from 500 representatives to 300 representatives. In 2020, these 500 representatives had enacted 651 laws, and, in 2021, the 300 representatives enacted only 413 laws. Instead of merely reporting this decrease of 238 laws passed, the researchers also reported the change from 2020 to 2021 the number of laws enacted in the neighboring country of Oceania, which did not change the size of its legislature. Explain the benefit of this comparison to Oceania.

**The benefit is that we have a comparison group to better assess the extent to which the change in the size of the legislature mattered. If the number of laws enacted in Oceania also dropped between 2020 and 2021, then we would be less certain that the 238-law drop in Freedonia was due to the change in the size of the Freedonia legislature. But if the number of laws enacted in Oceania did not drop between 2020 and 2021, then we would be more certain that the 238-law drop in Freedonia was due to the change in the size of the Freedonia legislature.**

-- No calculators, phones, or other devices permitted. --

5. [A] For a randomized experiment involving human participants, indicate what must be randomized for the randomized experiment to identify whether the difference in treatment caused a difference in outcomes. [B] Explain how that randomization helps identify whether the difference in treatment caused a difference in outcomes.

**[A] For a randomized experiment involving human participants to identify whether the difference in treatment caused a difference in outcomes, the treatment must be randomly assigned to the participants. [B] This randomization helps identify whether the difference in treatment caused a difference in outcomes, because this randomization helps ensure, as much as possible, that the set of people who got the treatment are the same as the set of people who did not get the treatment. Therefore, the only other plausible reason for a difference in outcome between the two groups would be random assignment error.**

-- No calculators, phones, or other devices permitted. --

6. In the Friday night television show *Shark Tank*, people who have a business idea explain this idea to wealthy investors, and these investors decide whether to invest in the idea.

Data analysis from the American National Election Studies 2024 Time Series Study indicated that divorce rate is lower among U.S. adults who regularly watch *Shark Tank* than among U.S. adults who do not regularly watch *Shark Tank* ( $p < 0.05$ ).

Suppose that we wanted to test the theory that, at least among U.S. adults, regularly watching *Shark Tank* causes people to be less likely to get divorced. We therefore run a regression using a "regularly watch *Shark Tank*" variable to predict whether a person has been divorced, and this regression controls for the person's income. Explain how statistical control for income can help us better test the theory that regularly watching *Shark Tank* causes people to be less likely to get divorced.

**The reason why U.S. adults who regularly watch *Shark Tank* are less likely to be divorced might only be because higher income causes both regularly watching *Shark Tank* and a lower likelihood of divorce. For instance, maybe higher income causes people to think that *Shark Tank* is more interesting because higher income people have more money to invest, and maybe higher income reduces the risk of divorce because higher income people have fewer fights about money. If the regression controls for a person's income, then that the regression was address the alternate explanation that income causes the association between regularly watching *Shark Tank* and a lower likelihood of divorce. (The mechanism is that statistical control will compare people at the same level of income and thus eliminate income as a source of difference between people.)**

-- No calculators, phones, or other devices permitted. --

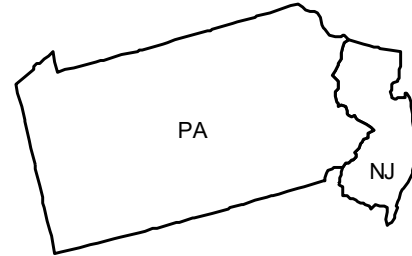
7. Suppose that we conduct a randomized experiment test whether a treatment affects an outcome. Our analysis does not provide sufficient evidence at the conventional level in political science that the treatment affects the outcome, but, in truth, the treatment does affect the outcome. [A] Identify one thing that could have caused our null result. [B] Then explain how that thing could have caused the null result.

**One potential response that would receive full credit: The null result could have been caused by a too-small sample size. Small sample sizes don't provide much information about the effect that we are testing for, so it's possible that a study with a small sample size merely did not detect an effect that exists.**

-- No calculators, phones, or other devices permitted. --

8. The passage below is from a famous study, "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania" by David Card and Alan B. Krueger, published in the American Economic Review:

On April 1, 1992, New Jersey's minimum wage rose from \$4.25 to \$5.05 per hour. To evaluate the impact of the law we surveyed 410 fast-food restaurants in New Jersey and eastern Pennsylvania before and after the rise. Comparisons of employment growth at stores in New Jersey and [eastern] Pennsylvania (where the minimum wage was constant) provide simple estimates of the effect of the higher minimum wage...We find no indication that the rise in the minimum wage reduced employment.



Card and Krueger could have compared stores in New Jersey to stores across the entire state of Pennsylvania. Explain the advantage of comparing stores in New Jersey to stores only in the eastern part of Pennsylvania.

**New Jersey shares a border with the eastern part of Pennsylvania, so -- on most characteristics -- New Jersey is plausibly more similar to the eastern part of Pennsylvania than to the entire state of Pennsylvania. Comparing regions that are similar to each other, except for the characteristic of interest (a change in the minimum wage, in this case), helps address the alternate explanation that any difference in outcome is due to a characteristic other than the characteristic of interest.**

-- No calculators, phones, or other devices permitted. --

9. Suppose that we want to estimate how much political ads influence vote choice in U.S. presidential elections. We could conduct a meta-analysis in which we combine the results from all studies that have ever been conducted on this research question, to produce an estimate of the average effect of how much political ads influence vote choice in U.S. presidential elections. Explain the benefit of the average effect size estimate in our meta-analysis being weighted by the sample size of the studies, instead of giving each study equal weight.

**We want to give more evidence to the larger sample studies, because, compared to smaller sample studies, larger sample studies provide more evidence about an effect (larger sample studies have more participants, more countries, etc).**

-- No calculators, phones, or other devices permitted. --

10. Suppose that a political science professor is interested in estimating the percentage of Illinois State University undergraduate students who will vote in the November 2028 election. In October 2028, the professor visits each 300-level political science course and gives each student a survey to complete, in which one of the items asks students whether they will vote in the November 2028 election. The professor uses the percentage of students who report on this survey that they will vote in the November 2028 election as the estimate of the percentage of Illinois State University undergraduate students who will vote in the November 2028 election. [A] Indicate whether you expect that – compared to the true percentage of ISU students who will vote in the November 2028 election – the professor's estimate of percentage of ISU students who will vote in the November 2028 election will be biased lower, higher, or not at all. [B] Explain the reason for your response.

**One potential response that would receive full credit: It's plausible that political science students are more interested in politics than the typical ISU student is interested in politics, so the professor's estimate is plausibly too high. It's also plausible that students (and people in general) will be more likely to falsely report that they will vote (compared to falsely reporting that they won't vote), because many people think that voting is expected of a good citizen.**

-- No calculators, phones, or other devices permitted. --

11. In the 2012 report "Who Attends Charter Schools and How Are Those Students Doing?", Naomi Chudowsky and Alan Ginsburg wrote:

When we examined four urban areas specifically (DC, Atlanta, Chicago, Milwaukee), students in charter schools significantly outperformed their peers in regular public schools in many of the subjects/grades analyzed.

It is possible that students in charter schools outperform students in public schools because charter schools teach students better than public schools teach students. But provide a different plausible explanation why students in charter schools outperform students in public schools.

**There are many acceptable responses for this item. For example, many charter schools might have the ability to select which students attend their school, and these charter schools might select only the best performing students, and that selection might be why students who attend charter schools outperform students in public schools: these students were better students on average even before the students got to the charter school.**

-- No calculators, phones, or other devices permitted. --

12. Suppose that a study indicated that U.S. residents who read the newspaper are more likely to vote, compared to U.S. residents who do not read the newspaper ( $p < 0.01$ ). Explain whether this is sufficient evidence to conclude at the conventional level in political science that, at least on average, reading the newspaper causes a U.S. resident to be more likely to vote.

**Not sufficient evidence, because the research design has not eliminated all plausible explanations. For example, having a high interest in politics might cause both reading the newspaper and voting, so that the association between reading the newspaper and voting is not necessarily caused by reading the newspaper.**

-- No calculators, phones, or other devices permitted. --

13. This past semester, Amy taught a POL 138 course in person, and Bob taught a POL 138 course online. Amy and Bob had a competition to see which was the better teacher, so they gave their students a 10-point test at the start of the semester and gave their students the same test at the end of the semester. Amy's students increased their average score by 7 points, from a 2 to a 9, but Bob's students only increased their score by 3 points, from a 6 to a 9. The p-value is  $p < 0.05$  for a test of the null hypothesis that the mean increase in scores among Amy's students was higher than the mean increase in scores among Bob's students. Explain whether, at least for this past semester, we can conclude at the conventional level in political science Amy was a better teacher that semester than Bob had been.

**Not sufficient evidence, because the research design has not eliminated all plausible explanations. For example, this could be due to a ceiling effect in which Amy and Bob did equally well, but -- because Bob's students started off higher in their scores (and mean of 6 compared to a mean of 2) -- Bob had less room to prove that he was a better teacher than Amy was.**

-- No calculators, phones, or other devices permitted. --

14. Suppose that researchers are testing for gender bias in student evaluations of teaching. Participants in an experiment were randomly drawn from students at a Midwestern university and were randomly assigned to rate the final exam for a course. Participants in Group 1 were given a final exam for a sociology course, in which the instructor's name was listed on the final exam as "Jennifer Daley". Participants in Group 2 were given a final exam for a physics course, in which the instructor's name was listed on the final exam as "Jack Daley". Participants in Group 2 rated Jack's final exam more positively than how participants in Group 1 rated Jennifer's final exam ( $p < 0.05$ ). Explain whether this means that we can conclude, at the conventional level in political science, that participants were biased against female instructors in their ratings of the final exam.

**Not sufficient evidence, because the research design has not eliminated all plausible explanations. For example, the female instructor's final exam was for sociology, and the male instructor's exam was for physics, so maybe participants were biased against sociology relative to physics.**

-- No calculators, phones, or other devices permitted. --

15. Suppose that a large random sample of college students indicated that college students who were randomly assigned to watch a video about Congress scored higher on a test about Congress than college students who were randomly assigned to not watch any video. The p-value for this comparison was  $p < 0.001$ . Is this sufficient evidence to conclude, at the conventional level in political science, that, at least on average, watching the video caused higher scores on the test about Congress?

**Yes, the random assignment for the experiment means that random assignment error is the only alternate explanation for the difference in test scores, and the p-value under  $p = 0.05$  is sufficient to eliminate the alternate explanation of random assignment error.**

-- No calculators, phones, or other devices permitted. --

16. A candidate for governor of Virginia stated during a debate before the 2021 election for governor of Virginia that "I don't think parents should be telling schools what they should teach". Suppose that a researcher was interested in whether hearing the candidate make that statement reduced support for the candidate. The day after the 2021 election for governor of Virginia, the researcher surveyed 1,000 randomly selected Virginians who had voted for governor of Virginia in the 2021 election. Suppose that the data indicated that persons in the sample who reported that they had heard the candidate make that statement were less likely to have voted for the candidate, compared to persons in the sample who reported that they had not heard the candidate make that statement ( $p < 0.001$ ). Explain whether that is sufficient evidence at the conventional level in political science to conclude that hearing the candidate make that statement reduced support for the candidate among Virginians who had voted for governor of Virginia in the 2021 election.

**Not sufficient evidence, because the research design has not eliminated all plausible explanations. For example, suppose that the candidate who made the statement was a Democrat, and conservative media such as Fox News was more likely than other media to repeat the statement. In this case, voting against the candidate would associate with hearing the statement, even if hearing the statement had no effect on vote choice.**

-- No calculators, phones, or other devices permitted. --

17. Suppose that a large random sample of college seniors indicated that political knowledge was higher among political science majors than among art majors. The p-value for this comparison was  $p < 0.001$ . Is this sufficient evidence at the conventional level in political science to conclude that being a political science major causes higher levels of political knowledge than being an art major does, at least on average?

**Not sufficient evidence, because the research design has not eliminated all plausible explanations. For example, people who choose political science as a major might have more political knowledge than people who choose art as a major, even before these people get to college and start their majors.**

-- No calculators, phones, or other devices permitted. --

18. Researchers A and B are interested in estimating the extent to which, compared to attending a public school, attending a charter school affects student scores on standardized tests.

Researcher A collected scores on state standardized tests taken at the end of the 2024 school year, from a random selection of 5,000 students in charter schools and from a random selection of 5,000 students in public schools. Researcher A tested the null hypothesis that the mean score among charter school students differed from the mean score among public school students.

Researcher B instead used a lottery that one charter school conducted to determine which applicants to admit for the 2023-24 school year. This charter school received applications from 800 potential students, but had space to admit only 200 students, so the charter school conducted a lottery to determine which applicants to admit. Researcher B collected scores for all 800 student applicants on state standardized tests taken at the end of the 2024 school year. Researcher B tested the null hypothesis that the mean score among the 200 students who randomly won the lottery and thus attended the charter school differed from the mean score among the 600 students who did not randomly win the lottery and thus attended a public school.

Indicate which researcher you think had the better research design for estimating the extent to which, compared to attending a public school, attending a charter school affects student scores on standardized tests. Then explain why.

**One acceptable response: Researcher B has a better research design, because the randomized experiment is better for causal inference: the lottery is random, so, at the time that the students start school, the 200 students who attended the charter school should be very similar to the 600 students who did not attend the charter school, so that the mean scores at the end of the year can be interpreted as largely due to the difference in school type.**

-- No calculators, phones, or other devices permitted. --

19. Barro et al. 2024 "Looking Backward: Long-term Religious Service Attendance in 66 Countries" reported results from an analysis of whether the reforms to the Catholic Church undertaken due to the Vatican II conference from 1962 to 1965 (such as not using Latin in the mass and eliminating the "fish on Fridays" requirement) affected religious service attendance rates among Catholics. Barro et al. 2024 could have merely reported that religious service attendance decreased in Catholic countries after 1965, but Barro et al. 2024 noted that religious service attendance after 1965 decreased more in Catholic countries than in Protestant or Orthodox countries that did not adopt the Vatican II reforms. Explain the benefit of this comparison to Protestant or Orthodox countries.

**The benefit is that we have a comparison group to better assess the extent to which the Vatican II reforms mattered. For all we know, religious service attendance could have decreased among all religions, so that – if we merely observe religious service attendance decrease among Catholics after Vatican II – we might incorrectly conclude that Vatican II caused the decrease in religious service attendance. But the comparison religions that did not have Vatican II reforms help us assess whether any change in religious service attendance in Catholic countries was due to Vatican II.**

-- No calculators, phones, or other devices permitted. --

20. In 2006, after Hurricane Katrina, the Time-sharing Experiments for the Social Sciences program ran a survey experiment proposed by van Boven et al. During the experiment, participants were randomly shown either the photo on the left of a Black man or the photo on the right of a White man. All participants were asked: "The person in the photograph is holding several packaged items. To what extent do you agree or disagree that this person is looting?", and all participants were given a scale that ranged from "strongly agree" to "strongly disagree".



On a scale in which 0 is "strongly disagree" and 1 is "strongly agree", the mean response was 0.38 among participants shown the photo of the Black man and was 0.48 among participants shown the photo of the White man. The p-value was  $p=0.0012$  for a test of the null hypothesis that these means equaled each other. Explain whether this would be sufficient evidence to conclude, at the conventional level in political science, that racial bias against Whites relative to Blacks was the only reason for this difference in means.

**In a randomized experiment, the only reasons for difference between groups is [1] the treatment difference or [2] random assignment error. The p-value less than  $p=0.05$  is sufficient evidence in political science to exclude random assignment error as a plausible explanation. So we can conclude that the treatment difference caused the difference in means. However, the men in the photos differ in an important way other than race, because the White man is struggling to carry the groceries and the Black man isn't; moreover, the Black man has his shirt tucked, and the White man doesn't. Therefore, it's plausible that the participants were more likely to report that the White man was looting than to report that the Black man was looting because of the non-racial differences.**