POL 138 Quantitative Reasoning in Political Science
KEY Practice Exam 1

Completing this practice exam is not worth points toward your POL 138 grade, but it might be useful to take the practice exam and check your responses against the key. Content on Exam 1 should closely match content on Practice Exam 1, but with potentially different ways of measuring understanding of concepts, such as if short answer items are converted to multiple-choice items or vice versa, or if understanding of a concept is measured a different way.

1 Basic tools of quantitative reasoning

1.1 Quantitative reasoning

1. Research focusing on numbers is ___.
   A. qualitative research
   B. quantitative research

2. Which of these is closest to what an inference is?
   A. a conclusion
   B. a prediction
   C. a reason for a prediction
   D. a flawed idea

1.2 Measures of central tendency

3. What is the mean of the set of numbers {0, 0, 1, 2, 7}?
   A. 0
   B. 1
   C. 2
   D. 10
   E. None of the above

   To calculate the mean of a set of numbers, add the numbers together and then divide that sum by the total number of numbers: 0+0+1+2+7=10, and then 10/5 = 2

4. What is the median of the set of numbers {0, 0, 1, 2, 7}?
   A. 0
   B. 1
   C. 2
   D. 10
   E. None of the above

   The median is a number that divides in half a set of numbers that is ordered from low to high. If the ordered set of numbers has an odd number of numbers, the

median is the middle number; if the ordered set has an even number of numbers, the median is the mean of the middle two numbers.

5. What is the median of the set of numbers {0, 1, 3, 8}?
   A. 0
   B. 1
   C. 3
   D. 12
   E. None of the above

   The median is a number that divides in half a set of numbers that is ordered from low to high. If the ordered set of numbers has an odd number of numbers, the median is the middle number; if the ordered set has an even number of numbers, the median is the mean of the middle two numbers.

1.3 Outliers

6. Which number or numbers is or are an outlier in the set {-1, 0, 0, 1, 1001}?
   A. -1 only
   B. 0 only
   C. -1 and 1001 only
   D. 1001 only
   E. There is no outlier

   An outlier is a number in a set that is very far away from most other numbers in the set.

7. Adding an outlier to a set of data would be expected to have more influence on the ___ of the data.
   A. mean
   B. median

8. Removing an outlier from a set of data would be expected to have more influence on the ___ of the data.
   A. mean
   B. median

1.4 Standard deviation

9. Standard deviation is a measure of ___.
   A. central tendency
   B. correctness
   C. reliability
   D. validity
   E. variation

10. Bob receives a 6 on each of his 4 exams. What is the standard deviation of Bob's exam scores?
A. 0
B. 4
C. 6
D. 24
E. None of the above

Standard deviation measures variation, and there is no variation in the set of numbers.

11. Bob recorded the temperatures in Celsius for the past few days, which were -2, -4, -5, and -1. What is known about the standard deviation of these temperatures?
A. The standard deviation of these temperatures is less than zero.
B. The standard deviation of these temperatures is zero.
C. The standard deviation of these temperatures is greater than zero.

Standard deviation measures variation, and there is some variation in the set of numbers.

12. Suppose that all students in a class take a 100-item test. The mean number of correct items on the test is 60, and the standard deviation of the number of correct items is 10. If the teacher counts each item correct as half of a point, the standard deviation of the test scores would be ___.
A. lower than 10
B. higher than 10
C. 10

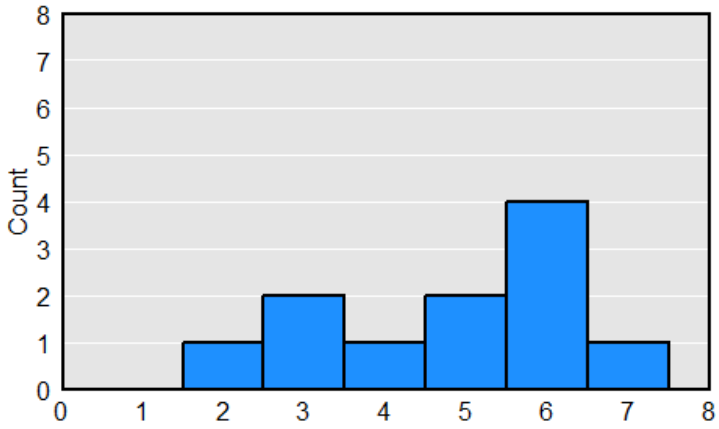
Multiplying numbers by 0.5 points (half of a point) reduces the variation in the set of numbers. Consider the numbers correct of 10 and 12, which are 2 units apart. Counting those as 5 points and 6 points makes those points only 1 unit apart, and thus the variation in the set of the numbers has decreased.

13. Suppose that students in a class have weights of 80 kilograms, 90 kilograms, 100 kilograms, and 110 kilograms. The teacher multiplies each weight in kilograms by 2.2 to estimate the students' weights in pounds. Compared to the students' weights in kilograms, the students' weights in pounds will have ___ number for the standard deviation.
A. the same
B. a lower
C. a higher

Multiplying numbers by 2.2 increases the variation in the set of numbers. Consider the weights of 100 kilograms and 110 kilograms, which are 10 units apart. Converting those numbers into pounds gives us 220 pounds and 242 pounds, which are 22 units apart, and thus the variation in the set of the numbers has increased.

1.5 Histograms

14. In the histogram below, which is true?



A. There is 1 observation of 2.
B. There are 5 observations of 2.

1.6 Proportions, percentages, and percentage points

15. Suppose that a sample has 10 Democrats, 5 Independents, and 7 Republicans. What proportion of the sample is Independent, to two decimal places?
A. 0.05
B. 0.23
C. 0.37
D. 0.50
E. 0.95

A proportion is a comparison of a part to a whole. In this case, the 5 Independents are compared to the full sample of 22 persons (10+5+7), so that the proportion is 5/22, which is 0.23 to two decimal places.

16. Suppose that, in 2023, 60% of students at a college are women, but that, in 2024, only 54% of students at the college are women. That change can be correctly expressed as a decrease of ___.
A. 6 percent
B. 6 percentage points

A percentage point change is calculated by subtracting the old percentage from the new percentage. In this case, the percentage point change is 54% minus 60%, to get a 6 percentage point decrease. A percent change is the percentage point change divided by the original percentage, which in this case is 6 percentage points divided by 60 percent, to get a 10% decrease.

17. Suppose that, in 2023, 60% of students at a college are women, but that, in 2024, only 54% of students at the college are women. That change can be correctly expressed as a decrease of ___.
A. 10 percent
B. 10 percentage points

A percentage point change is calculated by subtracting the old percentage from the new percentage. In this case, the percentage point change is 54% minus 60%, to get a 6 percentage point decrease. A percent change is the percentage point change divided by the original percentage, which in this case is 6 percentage points divided by 60 percent, to get a 10% decrease.

18. Suppose that, this year, 20% of students at a college are Republican, but that, next year, 40% of students at the college are Republican. That change can be correctly expressed as an increase of ___.
A. 20 percent
B. 20 percentage points

A percentage point change is calculated by subtracting the old percentage from the new percentage. In this case, the percentage point change is 40% minus 20%, to get a 20 percentage point increase. A percent change is the percentage point change divided by the original percentage, which in this case is 20 percentage points divided by 20 percent, to get a 100% increase.

19. Suppose that, this year, 20% of students at a college are Republican, but that, next year, 40% of students at the college are Republican. That change can be correctly expressed as an increase of ___.
A. 100 percent
B. 100 percentage points

A percentage point change is calculated by subtracting the old percentage from the new percentage. In this case, the percentage point change is 40% minus 20%, to get a 20 percentage point increase. A percent change is the percentage point change divided by the original percentage, which in this case is 20 percentage points divided by 20 percent, to get a 100% increase.

1.7 Percentiles

20. Suppose that a score of 70 is at the 80th percentile for scores on a test. What does this mean?
A. 80 percent of scores are above 70.
B. 80 percent of scores are below 70.
C. 80 percent of scores were 70.
D. A test with a score of 70 was a test with 80 percent of items correct.
E. None of the above

This POL 38 course will define percentile so that scoring at the Nth percentile indicates scoring above N% of scores

21. Suppose that the eight scores on a test are: 18, 34, 65, 75, 78, 81, 89, and 91. What percentile would the score of 89 be at, to the nearest whole percentile?
A. 11th percentile
B. 25th percentile
C. 37th percentile
D. 50th percentile
E. 75th percentile
F. 89th percentile

The score of 89 is above 6 scores, and there are 8 scores, so the score of 89 is above 6/8 or 75% of scores. So 89 is at the 75th percentile.

22. Which score below indicates a higher degree of political knowledge for a political knowledge test?
A. scoring at the 1st percentile on the test
B. scoring at the 99th percentile on the test

This POL 38 course will define percentile so that scoring at the Nth percentile indicates scoring above N% of scores, so that scoring above 99 percent of scores on a political knowledge test indicates a higher degree of political knowledge, compared to scoring above 1 percent of scores.

23. NBA basketball players tend to be taller than the average U.S. resident. Suppose that Bob is at the 80th percentile of height among U.S. residents. Bob's percentile height among NBA basketball players is likely ___.
A. less than the 80th percentile
B. at the 80th percentile
C. greater than the 80th percentile

Bob is taller than 80 percent of U.S. residents, but, if we move Bob to a taller-on-average group, Bob will not be as relatively tall.

24. NBA basketball players tend to be taller than the average U.S. resident. Suppose that Bob is at the 80th percentile of height among NBA basketball players. Bob's percentile height among U.S. residents is likely ___.
A. less than the 80th percentile
B. at the 80th percentile
C. greater than the 80th percentile

Bob is taller than 80 percent of NBA players, but, if we move Bob to a shorter-on-average group, Bob will be relatively taller.

1.8 Weighted means

25. Suppose that a course has three exams: Exam 1 is worth 10% of the overall grade for the course, Exam 2 is worth 30% of the overall grade for the course, and the Final Exam is worth 60% of the overall grade for the course. If a student scored 80% on Exam 1, 70% on Exam 2, and 90% on the Final Exam, what would be that student's overall percentage for the course?

We can calculate the student's final grade accounting for the fact that the final exam gets more weight than the midterm, by calculating the areas for each part of the final grade and them adding the areas together, as follows:

(0.10*80) + (0.30*70) + (0.60*90) = 83

2 Sampling

2.1 Sampling error

[Items 26 and 27] Suppose that a researcher is interested in estimating the percentage of students at U.S. universities who voted in the 2022 midterm elections, so the researcher samples 400 students at Illinois State University.

26. The researcher's population is best described as ___.
A. students at U.S. universities
B. the 400 students at Illinois State University

A population is the set of things of interest for a study. A sample is the set of things that were studied for the study.

27. The researcher's sample is best described as ___.
A. students at U.S. universities
B. the 400 students at Illinois State University

A population is the set of things of interest for a study. A sample is the set of things that were studied for the study.

2.2 Law of Large Numbers

28. Amy and Bob are estimating the mean political ideology of students at Illinois State University. Amy emails all 3,800 business majors at ISU, asks them to rate their political ideology on a scale from 0 for very liberal to 100 for very conservative, and then calculates the mean political ideology among these students; each of these 3,800 business majors responded to Amy's email. Bob randomly emails 600 ISU students, asks them to rate their political ideology on a scale from 0 for very liberal to 100 for very conservative, and then calculates the mean political ideology among these students; each

of these 600 students responded to Bob's email. Who has a more credible estimate of the mean political ideology of ISU students?

A. Amy, because her sample size is larger
B. Amy, because her sample included all business majors
C. Bob, because his sample size is smaller
D. Bob, because he used a random sample of the population

In a random sample of a population, each member of the population has an equal chance of being sampled. The benefit of this random sampling is that it tends to produce samples that are representative of the population, especially if the sample is large.

29. Amy flips a fair coin 20 times. Bob flips a fair coin 300 times. Compared to the percentage heads among Amy's coin flips, the percentage heads among Bob's coin flips is expected to be ___.

A. higher
B. lower
C. closer to 50%
D. farther from 50%

The Law of Large Numbers is that, as the number of randomly selected observations in a sample increases, the characteristics of the sample will tend to approach the characteristics of the population. So, compared to Amy's 20 flips, Bob's 300 flips are expected to produce a percentage heads closer to the true fair coin percentage of 50 percent.

30. Suppose that a population is 40% male. Which sample below is more likely to be closer to 40% male?

A. a random sample of 40 members of the population
B. a random sample of 120 members of the population

The Law of Large Numbers is that, as the number of randomly selected observations in a sample increases, the characteristics of the sample will tend to approach the characteristics of the population. So, compared to smaller samples, larger samples are expected to have characteristics that are closer to the characteristics of the population.

[Items 31 and 32] Suppose that a school randomly assigns students to a small class of 10 students or to a large class of 50 students, so that 20 classes in the school are small classes and 20 classes in the school are large classes. The school calculates the mean weight of students for each of these 40 classes, so that each class has one mean weight and there are a total of 40 mean class weights.

31. Across these 40 mean class weights, which would be most likely about the class that has the *lowest* mean class weight?

A. It is likely to be a small class.

B. It is likely to be a large class.
C. It is just as likely to be a small class as a large class.

Because of the Law of Large Numbers, the mean of a smaller random sample is more likely to be extreme, compared to the mean of a larger random sample.

32. Across these 40 mean class weights, which would be most likely about the class that has the *highest* mean class weight?
A. It is likely to be a small class.
B. It is likely to be a large class.
C. It is just as likely to be a small class as a large class.

Because of the Law of Large Numbers, the mean of a smaller random sample is more likely to be extreme, compared to the mean of a larger random sample.

---

33. Amy randomly samples 80 ISU students and asks them to rate the U.S. president on a scale from 0 for very cold to 100 for very warm. Bob randomly samples 250 ISU students and asks them to rate the U.S. president on a scale from 0 for very cold to 100 for very warm. Which of the following, if any, should be expected due to this difference in sample size?
A. The mean support for the president is likely to be lower in Amy's sample than in Bob's sample.
B. The mean support for the president is likely to be higher in Amy's sample than in Bob's sample.
C. Neither of the above

All else equal, larger random samples are better than smaller random samples, because larger samples provide more precision for our estimates. Compared to the mean of a larger sample, the mean a smaller sample is expected to be farther from the true population mean, but the mean of a smaller sample is not expected to biased higher or lower than the true population mean.

2.3 Imbalanced sample sizes

34. Suppose that we want to test the null hypothesis that the percentage of Illinois Democrats who support lowering the voting age to 12 equals the percentage of Illinois Republicans who support lowering the voting age to 12. Which set of samples below would be better for testing this null hypothesis?
A. a random sample of 300 Illinois Democrats and a random sample of 300 Illinois Republicans
B. a random sample of 400 Illinois Democrats and a random sample of 600 Illinois Republicans

The difference in sample size won't be expected to bias estimates of the gap between Illinois Democrats and Illinois Republicans.

35. Amy randomly samples 30 ISU students and asks them to rate the president on a scale from 0 for very cold to 100 for very warm. Bob randomly samples 200 ISU students and asks them to rate the president on a scale from 0 for very cold to 100 for very warm. Which of the following, if any, should be expected due to this difference in sample size?
    A. The mean support for the president is likely to be lower in Amy's sample than in Bob's sample.
    B. The mean support for the president is likely to be higher in Amy's sample than in Bob's sample.
    C. Neither of the above

Smaller random sample sizes are not expected to bias the mean of the sample to be necessarily higher or necessarily lower than the mean of the population. But, compared to the mean of a larger random sample, the mean of a smaller random sample is expected to be farther from the mean of the population.

2.4 Relatively small samples can be useful

[No exam items]

2.5 Sampling weights

36. Political scientists weight survey data for which of the following reasons?
    A. Because the sample is too small
    B. Because the population is much larger than the sample
    C. Because the sample characteristics do not match the population characteristics

37. Suppose that men are 70% of a population and women are 30% of the population. Men are 50% of our sample of the population, and women are 50% of our sample of the population. Men in our sample had a mean height of 180 cm, and women in our sample had a mean height of 170 cm. What would be the best estimate for the mean height of the population, assuming that the samples are representative?
    A. 172 cm
    B. 173 cm
    C. 175 cm
    D. 177 cm
    E. 178 cm

We are estimating the population mean height, so we can ignore the sample percentages. The problem then becomes a weighted mean problem, so:

0.70*180cm + 0.30*170cm = 177cm

38. If Asians are 10 percent of a sample and 6 percent of a population, what weight should be applied to each Asian person in the sample, if weighting on only race?
A. 6 + 10
B. 6 / 10
C. 10 + 6
D. 10 / 6
E. None of the above

The survey weight formula in general is population divided by sample, so, for this item, the survey weight for Asians is 6/10.

39. If the mean survey weight for a group is 0.4, then that means that the group was ___.
A. undersampled, relative to the group's percentage of the population
B. oversampled, relative to the group's percentage of the population
C. neither undersampled nor oversampled, relative to the group's percentage of the population

Generally speaking, if a person is underrepresented in the sample, then the sampling weight will be greater than 1, because multiplying by a number greater than 1 will increase the emphasis on that observation. And if a person is overrepresented in the sample, then the sampling weight will be less than 1, because multiplying by a number less than 1 will increase the emphasis on that observation. And if a person is correctly represented in the sample, then the sampling weight will be 1, because multiplying by 1 will not change the emphasis on that observation.

40. If the mean survey weight for a group is 1, then that means that the group was ___.
A. undersampled, relative to the group's percentage of the population
B. oversampled, relative to the group's percentage of the population
C. neither undersampled nor oversampled, relative to the group's percentage of the population

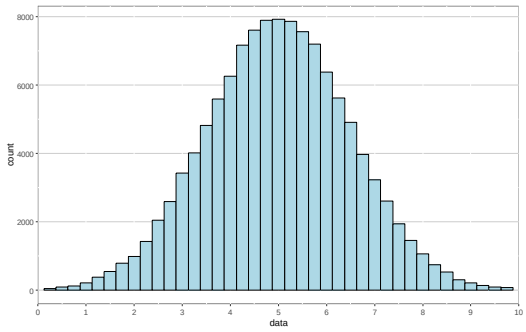
Generally speaking, if a person is underrepresented in the sample, then the sampling weight will be greater than 1, because multiplying by a number greater than 1 will increase the emphasis on that observation. And if a person is overrepresented in the sample, then the sampling weight will be less than 1, because multiplying by a number less than 1 will increase the emphasis on that observation. And if a person is correctly represented in the sample, then the sampling weight will be 1, because multiplying by 1 will not change the emphasis on that observation.

41. If the mean survey weight for a group is 1.3, then that means that the group was ___.
A. undersampled, relative to the group's percentage of the population
B. oversampled, relative to the group's percentage of the population
C. neither undersampled nor oversampled, relative to the group's percentage of the population

Generally speaking, if a person is underrepresented in the sample, then the sampling weight will be greater than 1, because multiplying by a number greater than 1 will increase the emphasis on that observation. And if a person is overrepresented in the sample, then the sampling weight will be less than 1, because multiplying by a number less than 1 will increase the emphasis on that observation. And if a person is correctly represented in the sample, then the sampling weight will be 1, because multiplying by 1 will not change the emphasis on that observation.

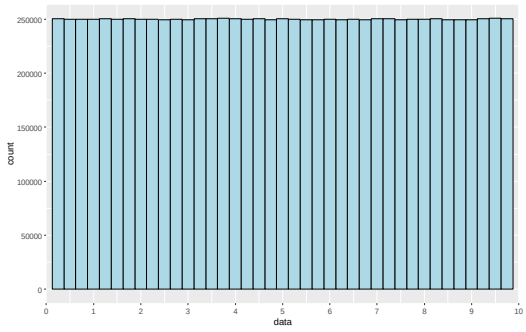
2.6 The normal distribution

42. The image below is an example of a ___.



A. uniform distribution
B. normal distribution

43. The image below is an example of a ___.



A. uniform distribution
B. normal distribution

44. Suppose that, in a set of 101 adults, each whole number weight from 150 lbs to 250 lbs appears exactly once (e.g., 150, 151, 152, …, 249, 250). We randomly select ten of these adults and plot the mean weight of this sample of ten adults. We randomly select another ten of these adults (which might include some adults already selected) and plot the mean weight of this second sample of ten adults. We continue until we plot 2,000 means. A histogram of the means is expected to be ___.
A. a uniform distribution
B. a non-uniform distribution

If one person is randomly selected, then each weight has the same chance of being selected. But, because of the Law of Large Numbers, if more than one person is randomly selected and the weights are averaged, then the sample average weight is expected to be closer to the population average, compared to other weights in the sample.

45. Suppose that we have a set of the ten whole numbers from 0 to 9, of {0,1,2,3,4,5,6,7,8,9}. We randomly select a number from this set and then add that number to a brand new empty database. We randomly select another number from this set (which might or might not have already been selected) and then add that number to the same database, so that the database now has two numbers. We do this over and over again until the database has 40,000 numbers. A histogram of the numbers in the database is expected to be ___.
A. a uniform distribution
B. a non-uniform distribution

If one number is randomly selected, then each number has the same chance of being selected.

46. Suppose that a test has a mean of 100 and a standard deviation of 10. Scores on the test follow a normal distribution. About 95% of scores should fall within which two scores?
A. 10 and 100
B. 60 and 120
C. 70 and 130
D. 80 and 120
E. 90 and 110

The normal distribution has the useful property that about 95% of points fall within 2 standard deviations of the mean.

47. The 600 scores in Group A follow a normal distribution and have a mean of 100 and a standard deviation of 5. The 600 scores in Group B follow a normal distribution and have a mean of 100 and a standard deviation of 20. Based on these statements, which one of the following statements is true?
A. It is more likely that Group A has the highest score, and not Group B.
B. It is more likely that Group B has the highest score, and not Group A.
C. The probability that Group A has the highest score is the same as the probability that Group B has the highest score.

Scores for Group A have the same mean as scores for Group B have, and both sets of scores are for a normal distribution, so the shapes of the distribution are both bell-shaped curves. But scores for Group B have a higher standard deviation than scores for Group A have, so scores for Group B are farther from the mean on average than scores from Group A. Thus, the highest score is more likely to be from Group B.

48. The 600 scores in Group A follow a normal distribution and have a mean of 100 and a standard deviation of 5. The 600 scores in Group B follow a normal distribution and have a mean of 100 and a standard deviation of 20. Based on these statements, which one of the following statements is true?

A. It is more likely that Group A has the lowest score, and not Group B.
B. It is more likely that Group B has the lowest score, and not Group A.
C. The probability that Group A has the lowest score is the same as the probability that Group B has the lowest score.

Scores for Group A have the same mean as scores for Group B have, and both sets of scores are for a normal distribution, so the shapes of the distribution are both bell-shaped curves. But scores for Group B have a higher standard deviation than scores for Group A have, so scores for Group B are farther from the mean on average than scores from Group A. Thus, the highest score is more likely to be from Group B.

49. Suppose that scores on a national test follow a normal distribution and have a mean of 100 and a standard deviation of 10. If Student A raises her score from 90 to 100, and Student B raises her score from 120 to 130, which of the following statements is true?

A. Student A had the higher percentile increase on the test.
B. Student B had the higher percentile increase on the test.
C. Student A had the same percentile increase on the test as Student B had.

In a normal distribution, there are more cases near the middle of the distribution. In this case, Student A's 10-point increase is closer to the middle of the distribution that Student B's 10-point increase, so that Student A will jump over more scores than Student B does and thus Student A will have a higher percentile increase than Student B does.

50. Suppose that scores on a national test follow a normal distribution and have a mean of 100 and a standard deviation of 10. If Student A raises her score from 90 to 100, and Student B raises her score from 100 to 110, which of the following statements is true?

A. Student A had the higher percentile increase on the test.
B. Student B had the higher percentile increase on the test.
C. Student A had the same percentile increase on the test as Student B had.

In a normal distribution, there are more cases near the middle of the distribution. In this case, Student A's 10-point increase is just as close to the middle of the distribution as Student B's 10-point increase is, so that Student A will jump over as many scores as Student B does and thus Student A will have the same percentile increase that Student B does.

2.7 Confidence intervals

51. Which of the following is expected to be wider?

A. the 95% confidence interval for the mean weight of a random sample of 10 U.S. residents
B. the 95% confidence interval for the mean weight of a random sample of 200 U.S. residents

All else equal, larger samples produce thinner confidence intervals, because a larger amount of data from random sampling better helps us "close in" on the true characteristic of the population.

52. For a given estimate, all else equal, which of the following would be the wider?
A. 90% confidence interval
B. 99% confidence interval

The higher the percentage for a confidence interval, the wider the confidence interval must be to contain the true population mean. For example, a 99% confidence interval must contain the true mean 99 percent of the time, so a 99% confidence interval must be wider than a 95% confidence interval.

53. Amy randomly samples 30 ISU students and asks them to rate the president on a scale from 0 for very cold to 100 for very warm. Bob randomly samples 200 ISU students and asks them to rate the president on a scale from 0 for very cold to 100 for very warm. Which of the following, if any, should be expected due to this difference in sample size?
A. The 95% confidence interval for mean support for the president is thinner in Amy's sample than in Bob's sample.
B. The 95% confidence interval for mean support for the president is wider in Amy's sample than in Bob's sample.
C. Neither of the above

All else equal, larger samples produce thinner confidence intervals, because a larger amount of data from random sampling better helps us "close in" on the true characteristic of the population.

2.8 Margin of error

54. Suppose that, in a random sample of 1600 Illinois residents, the mean rating about the governor on a scale from 0 to 100 was 59, and the standard deviation of the ratings was 18. Use the margin of error formula below to calculate to two decimal places the margin of error for the mean rating of the governor, in which *s* is the sample standard deviation and *n* is the sample size. Show all of your work.

$$MOE = 1.960 \times \frac{s}{\sqrt{n}}$$

Plug in and calculate: $MOE = 1.960 \times \frac{18}{\sqrt{1600}} = 0.88$

55. Use the margin of error formula below to calculate to two decimal places the margin of error for an estimated percentage of persons who approve of the governor in a population, based on a random sample of the population in which 576 of 720 residents approved of the governor, in which p is the sample percentage and n is the sample size. Show all of your work.

$$MOE = 1.960 \times \sqrt{\frac{(p)(100-p)}{n}}$$

Plug in and calculate: $\boldsymbol{MOE = 1.960 \times \sqrt{\frac{(80)(100-20)}{720}} = 2.92}$

3 p-values

3.1 The null hypothesis

56. Which best indicates what the null hypothesis is?
A. The hypothesis being tested
B. The hypothesis that is true
C. The hypothesis that the effect is not zero
D. The hypothesis that is most supported by the evidence

57. Suppose that the null hypothesis is that a treatment will have a negative effect. Which of the following would be the alternate hypothesis?
A. The treatment will have no effect.
B. The treatment will have a positive effect.
C. The treatment will not have a negative effect.

Everything that is not included in the null hypothesis is included in the alternative hypothesis.

3.2 p-values

58. Of the following, which best describes what a p-value measures?
A. the precision of an estimate
B. the strength of evidence against the null hypothesis
C. the size of an association controlling for other model factors

59. Of the p-values below, which p-value is the strongest evidence that an observed difference between the percentage of heads and the percentage of tails from a set of coin flips would have been unlikely to have occurred due to random chance, if the coin is fair?
A. 0.01
B. 0.05
C. 0.99
D. 1.00

A p-value is a measure of the strength of the evidence that an analysis has provided against the null hypothesis. If an analysis has provided no evidence against the null hypothesis, the p-value is 1. The lower the p-value, the more evidence the analysis has provided against the null hypothesis. A p-value of zero would indicate that the analysis has provided infinitely strong evidence against the null hypothesis.

60. If we flipped a coin 12 times and got 6 heads and 6 tails, what would be the p-value for a statistical test of the null hypothesis that the coin is fair?
A. 0
B. 1
C. something between 0 and 1

A p-value is a measure of the strength of the evidence that an analysis has provided against the null hypothesis. If an analysis has provided no evidence against the null hypothesis, the p-value is 1.

In this case, there is no evidence against the null hypothesis that the coin is fair.

61. If we flipped a coin 12 times and got 2 heads and 10 tails, what would be the p-value for a statistical test of the null hypothesis that the coin is fair?
A. 0
B. 1
C. something between 0 and 1

A p-value is a measure of the strength of the evidence that an analysis has provided against the null hypothesis. If an analysis has provided no evidence against the null hypothesis, the p-value is 1. The lower the p-value, the more evidence the analysis has provided against the null hypothesis. A p-value of zero would indicate that the analysis has provided infinitely strong evidence against the null hypothesis.

In this case, there is some evidence against the null hypothesis that the coin is fair.

62. If we flipped a coin 12 times and got 0 heads and 12 tails, what would be the p-value for a statistical test of the null hypothesis that the coin is fair?
A. 0
B. 1
C. something between 0 and 1

A p-value is a measure of the strength of the evidence that an analysis has provided against the null hypothesis. If an analysis has provided no evidence against the null hypothesis, the p-value is 1. The lower the p-value, the more evidence the analysis has provided against the null hypothesis. A p-value of zero would indicate that the analysis has provided infinitely strong evidence against the null hypothesis.

In this case, there is some evidence against the null hypothesis that the coin is fair.

63. Suppose that, in an experiment, the mean for the control group was 2, the standard deviation for the control group was 2, the mean for the treatment group was 2, and the standard deviation for the treatment group was 3. What would be the p-value for a test of the null hypothesis that the control group mean equals the treatment group mean?
A. 0
B. 1
C. something between 0 and 1

A p-value is a measure of the strength of the evidence that an analysis has provided against the null hypothesis. If an analysis has provided no evidence against the null hypothesis, the p-value is 1.

In this case, there is no evidence against the null hypothesis that the control group mean equals the treatment group mean.

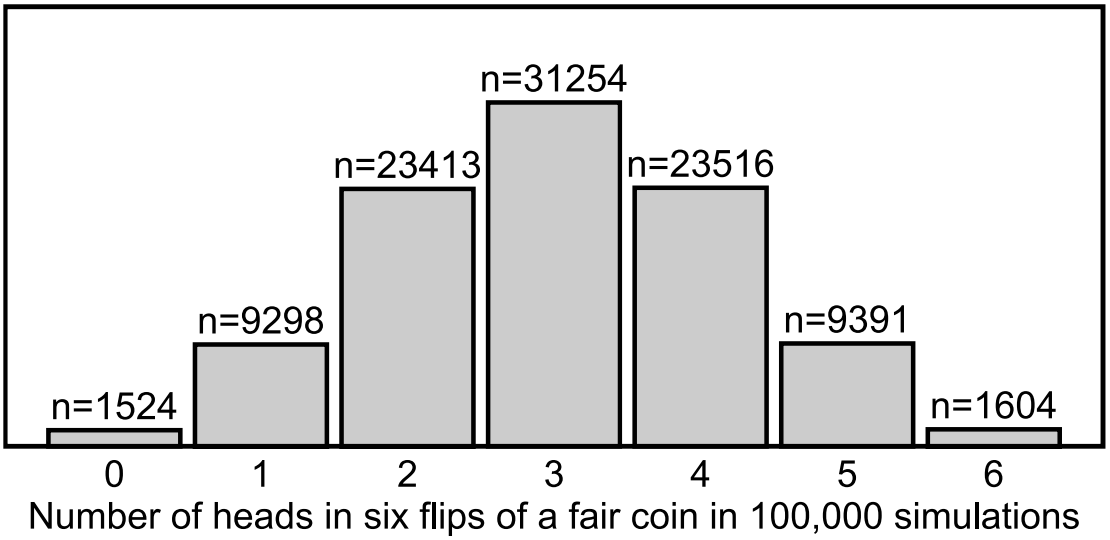
64. Suppose that, in an experiment, the mean for the control group was 4, the standard deviation for the control group was 3, the mean for the treatment group was 2, and the standard deviation for the treatment group was 1. What would be the p-value for a test of the null hypothesis that the control group mean equals the treatment group mean?
A. 0
B. 1
C. something between 0 and 1

A p-value is a measure of the strength of the evidence that an analysis has provided against the null hypothesis. If an analysis has provided no evidence against the null hypothesis, the p-value is 1. The lower the p-value, the more evidence the analysis has provided against the null hypothesis. A p-value of zero would indicate that the analysis has provided infinitely strong evidence against the null hypothesis.

In this case, there is some evidence against the null hypothesis that the control group mean equals the treatment group mean.

3.3 Estimating p-values

65. The histogram below is from a simulation that had 100,000 trials. For each trial, a fair coin was flipped six times. The horizontal x-axis indicates each number of times that the fair coin could land on heads in the six flips, and the height of the columns indicates the number of times the fair coin landed on that number of heads in the six flips. For example, the n=31254 for the x-axis value of 3 indicates that the fair coin landed on heads exactly 3 times in the 6 flips in 31,254 of the 100,000 trials.



Based on the above simulation data, which of the following calculations indicates the p-value that would occur for a test of the null hypothesis that a coin is fair, if the coin landed on 5 heads in 6 flips?

A. 9391 ÷ 100,000
B. (9391 + 1604) ÷ 100,000
C. (9391 + 1604 + 9298 + 1524) ÷ 100,000
D. (23413 + 31254 + 23516) ÷ 100,000

The p-value indicates the amount of evidence that an analysis has provided against the null hypothesis, so, to estimate a p-value, we first simulate – over and over again – what would happen if the null hypothesis were true. Then, we calculate the percentage of simulated outcomes that are at least as extreme as the observed outcome. That percentage is our p-value, which indicates the amount of evidence that our analysis has provided against the null hypothesis.

3.4 p-values if the null hypothesis is true

66. Suppose that we conduct 900 well-designed independent tests of a null hypothesis. In reality, the null hypothesis is true. What is the expected percentage of these tests that are expected to have a p-value of $p<0.05$?

A. 0%
B. 5%
C. 50%
D. 95%
E. 100%
F. Cannot be determined without more information

The distribution of p-values when the null hypothesis is true is expected to be a uniform distribution, so that, for example, 5% of p-values are equal to or lower than p=0.05.

3.5 p-values if the null hypothesis is not true

[Items 67 through 69] In the original setup for an experiment, the outcome variable is coded from 0 through 20, the mean of the outcome variable is 8 for Group A and is 11 for Group B, the standard deviation of the outcome variable is 9 for Group A and is 9 for Group B, and the sample size is 100 for Group A and is 100 for Group B. The p-value is p=0.02 for a test of the null hypothesis that the mean of the outcome variable for Group A equals the mean of the outcome variable for Group B.

67. Suppose that everything else were the same as in the original setup, but the sample sizes were 200 for Group A and 200 for Group B. Which of the following, if any, would we know about the p-value for a test of the null hypothesis that the mean of the outcome variable for Group A equals the mean of that outcome variable for Group B?
A. The p-value would be p=0.02
B. The p-value would be less than p=0.02
C. The p-value would be greater than p=0.02

If the null hypothesis is not true, then the size of a p-value can be affected by sample size: all else equal, larger samples provide more evidence, so – if the null hypothesis is not true – larger samples are expected to associate with more evidence against the null hypothesis and thus to associate with smaller p-values, all else equal.

68. Suppose that everything else were the same as in the original setup, but the standard deviation of responses were 6 for Group A and 6 for Group B. Which of the following, if any, would we know about the p-value for a test of the null hypothesis that the mean of the outcome variable for Group A equals the mean of that outcome variable for Group B?
A. The p-value would be p=0.02
B. The p-value would be less than p=0.02
C. The p-value would be greater than p=0.02

If the null hypothesis is not true, then the size of a p-value can be affected the standard deviation of measurements: the less variation in the measurements, the more certainty we have about the center of the measurements, so – if the null hypothesis is not true – smaller standard deviations are expected to associate with smaller p-values, all else equal.

69. Suppose that everything else were the same as in the original setup, but the mean response was 8 for Group A and 14 for Group B. Which of the following, if any, would we know about the p-value for a test of the null hypothesis that the mean of the outcome variable for Group A equals the mean of that outcome variable for Group B?
A. The p-value would be p=0.02
B. The p-value would be less than p=0.02

C. The p-value would be greater than p=0.02

If the null hypothesis is not true, then the size of a p-value can be affected by the size of the association: all else equal, larger associations are easier to detect, so – if the null hypothesis is not true – larger associations are expected to associate with smaller p-values, all else equal.

3.6 Hypothesis testing

70. What is the conventional p-value threshold in political science?
A. 0
B. 0.01
C. 0.05
D. 0.50
E. 0.95
F. 0.99
G. 1

71. If the p-value for a test of a null hypothesis is p=0.02, then we should do which of the following?
A. accept the null hypothesis and accept the alternative hypothesis
B. reject the null hypothesis and reject the alternative hypothesis
C. accept the null hypothesis and reject the alternative hypothesis
D. reject the null hypothesis and accept the alternative hypothesis
E. none of the above

If the p-value is less than the p-value threshold (we will use p=0.05 for the conventional p-value threshold in political science), then there is sufficient evidence to reject the null hypothesis; and, because the null hypothesis is the complement of the alternative hypothesis, if the null hypothesis is rejected, then the alternative hypothesis can be accepted.

72. A researcher tested the null hypothesis that an association is zero. The p-value for this test p=0.91. Based on this p-value, which of the following should the researcher do, using the conventional level in political science?
A. conclude that the association is zero
B. conclude that the association is not zero
C. neither of the above

If the p-value is NOT less than the p-value threshold (we will use p=0.05 for the conventional p-value threshold in political science), then there is NOT sufficient evidence to reject the null hypothesis; however, we don't accept the null hypothesis, because the p-value doesn't tell us anything directly about the plausibility of the null hypothesis.

73. A researcher tested the null hypothesis that an association is zero. The p-value for this test p=0.01. Based on this p-value, which of the following should the researcher do, using the conventional level in political science?
A. conclude that the association is zero
B. conclude that the association is not zero
C. neither of the above

If the p-value is less than the p-value threshold (we will use p=0.05 for the conventional p-value threshold in political science), then there is sufficient evidence to reject the null hypothesis; and, because the null hypothesis is the complement of the alternative hypothesis, if the null hypothesis is rejected, then the alternative hypothesis can be accepted.

74. A researcher tested the null hypothesis that an association is zero. The p-value for this test p=0.30. Based on this p-value, which of the following should the researcher do, using the conventional level in political science?
A. conclude that the association is zero
B. conclude that the association is not zero
C. neither of the above

If the p-value is NOT less than the p-value threshold (we will use p=0.05 for the conventional p-value threshold in political science), then there is NOT sufficient evidence to reject the null hypothesis; however, we don't accept the null hypothesis, because the p-value doesn't tell us anything directly about the plausibility of the null hypothesis.

3.7 Selecting a p-value threshold

75. Suppose that we are testing patient blood samples for evidence of the presence of a new virus. If there is sufficient evidence in the blood sample for the presence of the new virus, we will prescribe the patient a pill that has no negative effects and that can help combat the new virus. Our null hypothesis is that the blood sample does not have the new virus. Which p-value threshold below would be more appropriate, if we prefer to avoid not prescribing the pill to patients whose blood contains the new virus?
A. p=0.01
B. p=0.10

If we want to better avoid a false negative in which we incorrectly do not reject a false null hypothesis, then we can raise the p-value to something such as p=0.10, so that we require less evidence to reject the null hypothesis.

76. Suppose that we are testing for gender bias among ISU students in student evaluations of teachers. Our null hypothesis is that there is no gender bias among ISU students in student evaluations of teachers. Which p-value threshold below would be more appropriate, if we prefer to avoid falsely concluding that ISU students have a gender bias in student evaluations of teachers?

A. p=0.01
B. p=0.10

If we want to better avoid a false positive in which we incorrectly reject a true null hypothesis, then we can lower the p-value to something such as p=0.01 or p=0.001, so that we require more evidence to reject the null hypothesis

77. Suppose that we are testing for gender bias among ISU students in student evaluations of teachers. Our null hypothesis is that there is no gender bias among ISU students in student evaluations of teachers. Which p-value threshold below would be more appropriate, if we prefer to avoid falsely concluding that ISU students have a gender bias in student evaluations of teachers?
A. p=0.01
B. p=0.10

If we want to better avoid a false positive in which we incorrectly reject a true null hypothesis, then we can lower the p-value to something such as p=0.01 or p=0.001, so that we require more evidence to reject the null hypothesis

3.8 Statistical and substantive significance

78. For a test of the null hypothesis that there is no association, "statistically significant evidence" for the association refers to ___.
A. sufficient evidence that a particular association exists
B. sufficient evidence that a particular association is large

79. If the p-value is p=0.03 for a single statistical test of a null hypothesis that there is no association, do we have enough evidence to claim that there is statistically significant evidence for the detected association?
A. Yes
B. No

80. If the p-value is p=0.00001 for a single statistical test of a null hypothesis that there is no association, do we have enough evidence to claim that there is statistically significant evidence for the detected association?
A. Yes
B. No

81. If the p-value is p=0.00001 for a single statistical test of a null hypothesis that there is no association, do we have enough evidence to claim that there is substantively significant evidence for the detected association?
A. Yes
B. No

Even if the effect is substantively small, it's possible to get a p-value indicating a lot of evidence against the null hypothesis. For instance, if a coin flipped 1 million times

landed on heads 50.2% of time, the coin isn't very unfair, but the associated p-value would be very small, at p=0.000064

3.9 Hypothesis tests involving random sampling

82. Suppose that we ask each resident in a random sample of 1,000 Illinois residents whether the resident approves of the Illinois governor. Results indicate that 535 sampled Illinois residents reported approving of the Illinois governor and 465 sampled Illinois residents reported not approving of the Illinois governor. The p-value is p=0.40 for a test of the null hypothesis that this 53.5% approval equals 50% approval. Is this sufficient evidence at the conventional level in political science to reject the null hypothesis that 50% of the population of Illinois residents approve of the Illinois governor?
A. Yes
B. No

The p-value of p=0.40 is NOT sufficient evidence to reject the null hypothesis that 50% of the population of Illinois residents approve of the Illinois governor.

83. Suppose that we ask each resident in a random sample of 1,000 Illinois residents whether the resident approves of the Illinois governor. Results indicate that 535 sampled Illinois residents reported approving of the Illinois governor and 465 sampled Illinois residents reported not approving of the Illinois governor. The p-value is p=0.04 for a test of the null hypothesis that this 53.5% approval equals 50% approval. Is this sufficient evidence at the conventional level in political science to reject the null hypothesis that 50% of population of Illinois residents approve of the Illinois governor?
A. Yes
B. No

The p-value of p<0.05 is sufficient evidence to reject the null hypothesis that 50% of the population of Illinois residents approve of the Illinois governor.

3.10 Caution about p-values for causal inference

84. Suppose that researchers in Freedonia propose a theory that getting married will cause men to work more hours and thus increase their income. Researchers collect data from a representative sample of Freedonia men, and the data indicate that income is 11% higher for married Freedonia men than for never married Freedonia men. The p-value is p<0.05 for a test of the null hypothesis that these means equal each other. Does this analysis contain sufficient evidence to conclude, at the conventional level in political science, that, at least among men in this analysis and at least on average, getting married causes men to have a higher salary?
A. Yes
B. No

p-values are useful for making descriptive inferences. But to make a causal inference, we need to consider other potential explanations. In this case, married

men might have a higher salary than unmarried men, for reasons other than marriage. Maybe married men are older or maybe – even before getting married – married men are smarter or harder working.

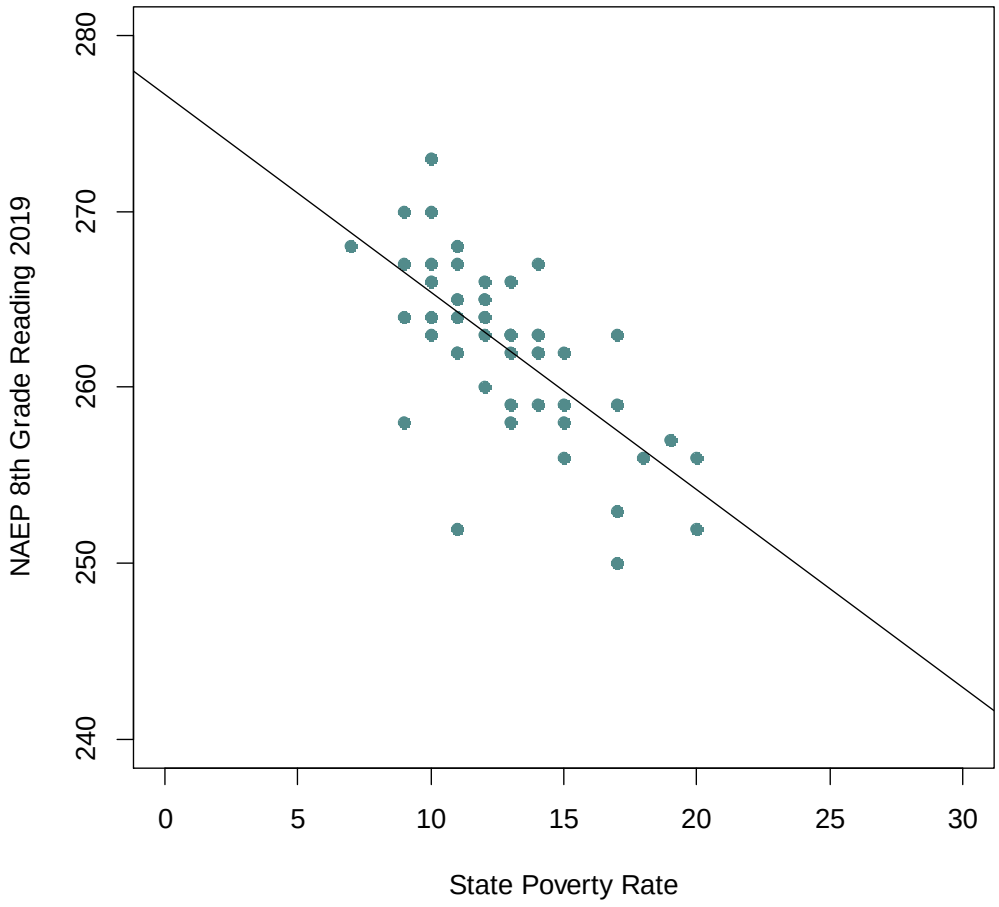
4 Linear regression

4.1 Linear regression line of best fit using OLS

[No exam items]

4.2 Simple linear regression

[Items 85 through 89] Use the image and statistical output below to answer items 30 to 34. The image and output are from a linear regression that used the poverty rate in a state (X) to predict the average eighth grade reading score in that state (Y). The Poverty Rate predictor is in whole number percentages, running from about 6 percent to about 20 percent.




```
Coefficients:
```

	Estimate	p-value
(Intercept)	276.6	<0.0001
Poverty Rate	-1.1	<0.0001

85. What does the 276.6 intercept coefficient indicate?
A. The predicted eighth grade reading score in a state with no poverty is 276.6.
B. The average eighth grade reading score in a state is 276.6.
C. For each one-unit increase in the poverty rate, a state's eighth grade reading score is predicted to increase by 276.6 units.
D. The highest observed eighth grade reading score in any state is 276.6.
E. For each one-unit increase in eighth grade reading score, a state's poverty rate is predicted to increase by 276.6 units.

The intercept for a linear regression is the predicted outcome when all predictors are set to zero.

86. What does the -1.1 coefficient for poverty rate indicate?
A. The predicted eighth grade reading score in a state with no poverty is -1.1.
B. The average eighth grade reading score in a state is -1.1.
C. For each one-unit increase in the poverty rate, a state's eighth grade reading score is predicted to decrease by 1.1 units.
D. The highest observed eighth grade reading score in any state is 1.1.
E. For each one-unit increase in eighth grade reading score, a state's poverty rate is predicted to decrease by 1.1 units.

For a linear regression, the estimate for a predictor can be thought of as a slope: for a one-unit increase in the predictor, the predicted outcome changes by the coefficient for the predictor.

87. Which of the following is a correct linear regression equation for the output in the image, using X and Y?
A. $Y = (276.6-1.1)X$
B. $Y = -1.1(276.6X)$
C. $Y = 276.6X - 1.1$
D. $Y = -1.1X + 276.6$

For the line of best fit, the intercept does not get multiplied by anything but gets added to a predictor times the coefficient for the predictor.

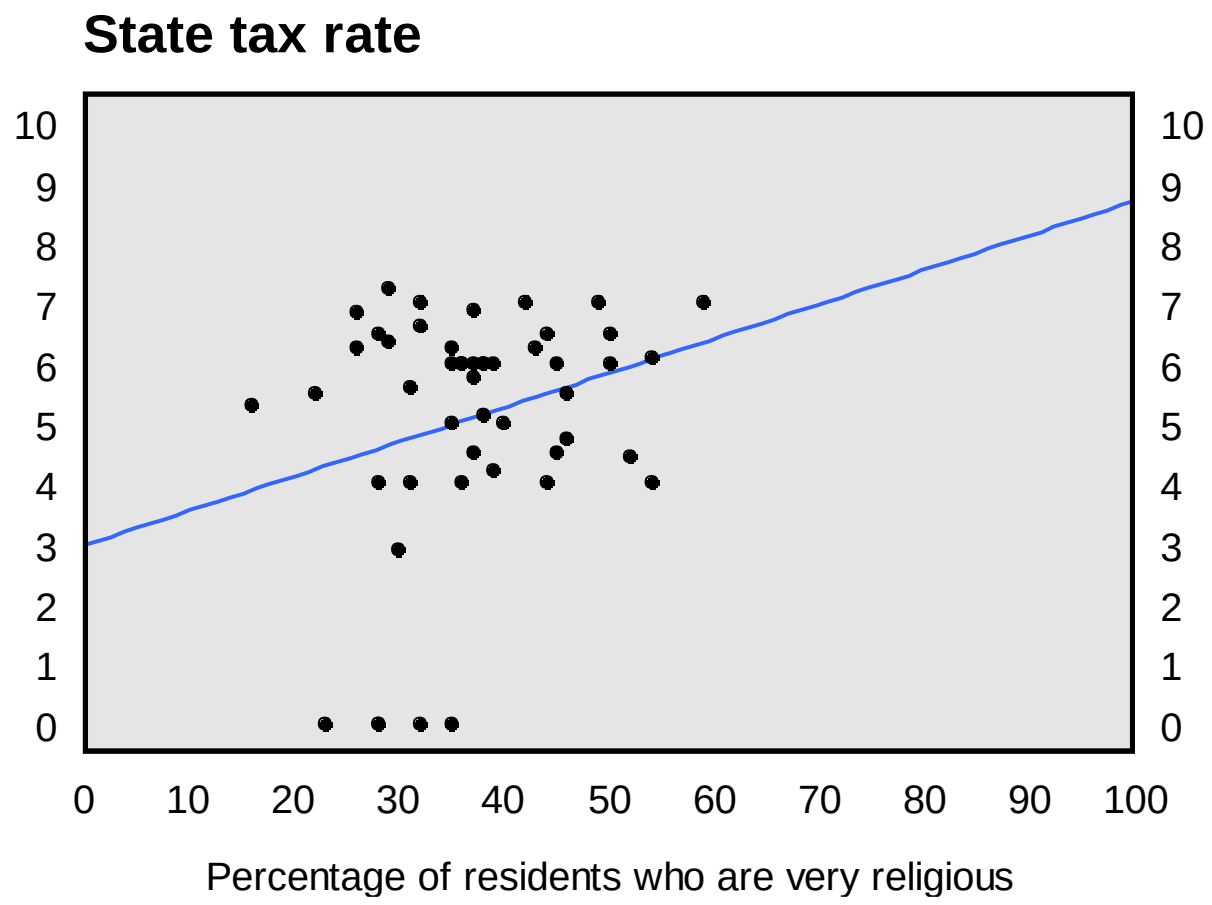
88. Which of the following would be closest to the predicted eighth grade reading score for a state that had a 16% poverty rate?
A. 245
B. 250
C. 255
D. 259
E. 262

Take the formula for the line of best fit (Y = -1.1X + 276.6) and plug in 16 for X, to get: Y = -1.1*16 + 276.6, which equals 259.

89. Is there is enough evidence in the data and output to conclude that a higher poverty rate in a state caused a lower eighth grade reading score in the state, at least on average?
A. Yes
B. No

p-values are useful for making descriptive inferences. But to make a causal inference, we need to consider other potential explanations. In this case, students and others in the state having a relatively low reading ability might cause the higher levels of poverty (instead of the effect being the other way around).

[For items 90 through 94] In the plot below, each point is a state in the United States. The horizontal x-axis is the percentage of state residents who are very religious (variable name `PctReligVery`, and the vertical y-axis is the tax rate in the state (variable name `StateTax`).



The statistical output for the analysis in the plot is:

```
                  Estimate   p-value
(Intercept)         2.96      0.014
PctReligVery        0.06      0.067
```

The statistical output for the analysis in the plot on the prior page is:

```
                  Estimate   p-value
(Intercept)         2.96      0.014
PctReligVery        0.06      0.067
```

90. What does the 2.96 intercept coefficient indicate?
    A. The percentage of very religious residents in a state with an average state tax rate is predicted to be 2.96.
    B. The state tax rate in a state with an average percentage of very religious residents is predicted to be 2.96.
    C. The percentage very religious in a state with no state taxes is predicted to be 2.96.
    D. The state tax rate in a state with no very religious residents is predicted to be 2.96.
    E. For each one-unit increase in a state's percentage of very religious residents, the predicted state tax rate is 2.96 units higher.
    F. For each one-unit increase in the state tax rate, the predicted percentage of very religious residents in the state is 2.96 units higher.

The intercept for a linear regression is the predicted outcome when all predictors are set to zero.

91. What does the 0.06 coefficient for PctReligVery indicate?
    A. The percentage of very religious residents in a state with an average state tax rate is predicted to be 0.06.
    B. The state tax rate in a state with an average percentage of very religious residents is predicted to be 0.06.
    C. The percentage very religious in a state with no state taxes is predicted to be 0.06.
    D. The state tax rate in a state with no very religious residents is predicted to be 0.06.
    E. For each one-unit increase in a state's percentage of very religious residents, the predicted state tax rate is 0.06 units higher.
    F. For each one-unit increase in the state tax rate, the predicted percentage of very religious residents in the state is 0.06 units higher

For a linear regression, the estimate for a predictor can be thought of as a slope: for a one-unit increase in the predictor, the predicted outcome changes by the coefficient for the predictor.

92. Which of the following is a correct linear regression equation for the output in the image, using X and Y?
    A. Y = 2.96 + 0.06 + X
    B. Y = 2.96*X + 0.06
    C. Y = 2.96 + 0.06*X

D. Y = (2.96/0.06)*X

For the line of best fit, the intercept does not get multiplied by anything but gets added to a predictor times the coefficient for the predictor.

93. Which of the following would be closest to the predicted state tax rate for a state that had 50% of its residents who are very religious?
A. -0.04
B. 0.06
C. 3.00
D. 5.96

Take the formula for the line of best fit (Y = 2.96 + 0.06*X) and plug in 50 for X, to get: Y = 2.96 + 0.06*50, which equals 5.96.

94. Is there is enough evidence in the data and output to conclude at the conventional level in political science that having a higher percentage of state residents who are very religious causes the tax rate in the state to be higher, at least on average?
A. Yes
B. No

p-values are useful for making descriptive inferences. But to make a causal inference, we need to consider other potential explanations. In this case, religiousness in a state might be correlated with political views or racial views or other views, and these other views (and not religion per se) might affect state tax rates.

4.3 Drawing the line of best fit

[Items 95 through 99] Below are coefficient estimates from a linear regression of data from residents in a hypothetical country. The linear regression used the education level of a resident (X, coded from 1 for less than a high school education to 6 for a post-graduate degree) to predict the resident's support for the country's president (Y).

```
Coefficients:
                Estimate
(Intercept)      40.00
Education         6.00
```

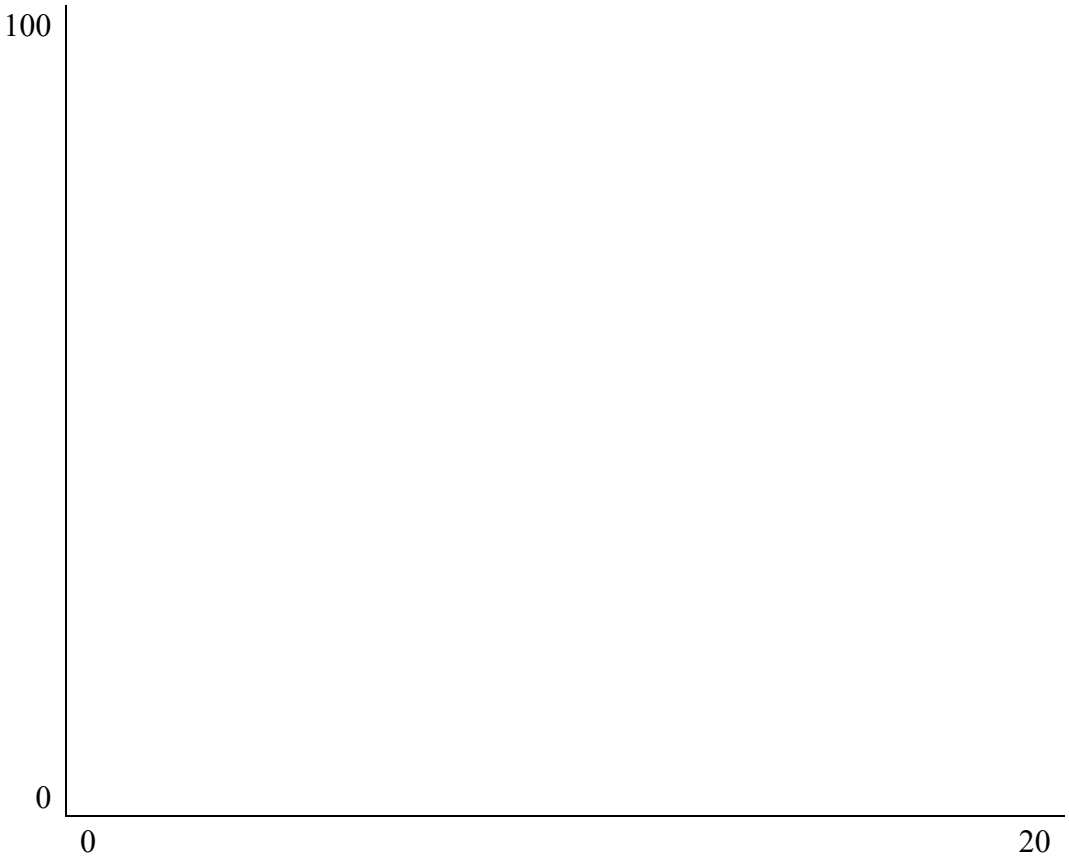
95. Below, write the equation to predict Y using X. **Y = 40 + 6X**
96. Label the Y-axis and the X-axis on the graph. **Y = Support, X = Education**
97. Draw and label a point at the value of Y for which the X variable is 1 (the lowest observed level of education). **Y = 40 + 6X = 40 + (6*1) = 46. So plot a point at X=1, Y=46**

98. Draw and label a point at the value of Y for which the X variable is 6 (the highest observed level of education). **Y = 40 + 6X = 40 + (6*6) = 76. So plot a point at X=6, Y=76**
99. Draw a line between the above two points.

[Items 100 to 104] Below are coefficient estimates from a linear regression of data from residents in a hypothetical country. The linear regression used the years of education of a resident (X) to predict the resident's support for the country's president (Y).

```
Coefficients:
                Estimate
(Intercept)      80.00
Education        -2.00
```

100. Below, write the equation to predict Y using X. **Y = -2X + 80**
101. Label the Y-axis and the X-axis on the graph. **Y = Support, X = Education**
102. Draw and label a point at the value of Y for which the X variable is 10 (the lowest observed years of education). **Y = -2*10 + 80 = 60. So plot a point at X=10, Y=60**
103. Draw and label a point at the value of Y for which the X variable is 20 (the highest observed years of education). **Y = -2*20 + 80 = 40. So plot a point at X=20, Y=40**
104. Draw a line between the above two points.



4.4 Linear regression with categorical predictors

[Items 105 and 106] Below is output from a linear regression using data from the ANES 2020 Time Series Study, predicting respondent ratings about the #MeToo movement (FTMETOO), using a predictor for the marital status of the respondent, with categories of married, widowed, divorced, separated, and never married, with "married" as the omitted category.

FTMETOO	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
(intercept)	56	1	108.32	0.000	55	57
Widowed	2	2	1.12	0.264	-1	5
Divorced	5	1	4.24	0.000	3	7
Separated	5	3	1.78	0.075	-1	11
Never married	10	1	10.65	0.000	8	12

105. What does the 56 coefficient estimate for the intercept indicate?
   A. The mean rating about the #MeToo movement is predicted to be 56 among the average respondent.

B. The mean rating about the #MeToo movement is predicted to be 56 among married respondents.
C. The mean rating about the #MeToo movement is predicted to increase by 56 for a one-unit increase in participant marital status.
D. The mean rating about the #MeToo movement is predicted to be 56 units higher for married respondents than for nonmarried respondents.

The intercept for a linear regression is the predicted outcome when all predictors are set to zero. In this case, the 56 coefficient is for a respondent who is not widowed, is not divorced, is not separated, and is not never married. That respondent is in the omitted category of married.

106. What does the 10 coefficient estimate for the "Never married" category indicate?
A. The mean rating about the #MeToo movement is predicted to be 10 among never married respondents.
B. The mean rating about the #MeToo movement is predicted to be 10 higher among never married respondents than among all other respondents.
C. The mean rating about the #MeToo movement is predicted to be 10 higher among never married respondents than among married respondents.

For a categorical predictor, the coefficient always refers to a comparison with the omitted category.