

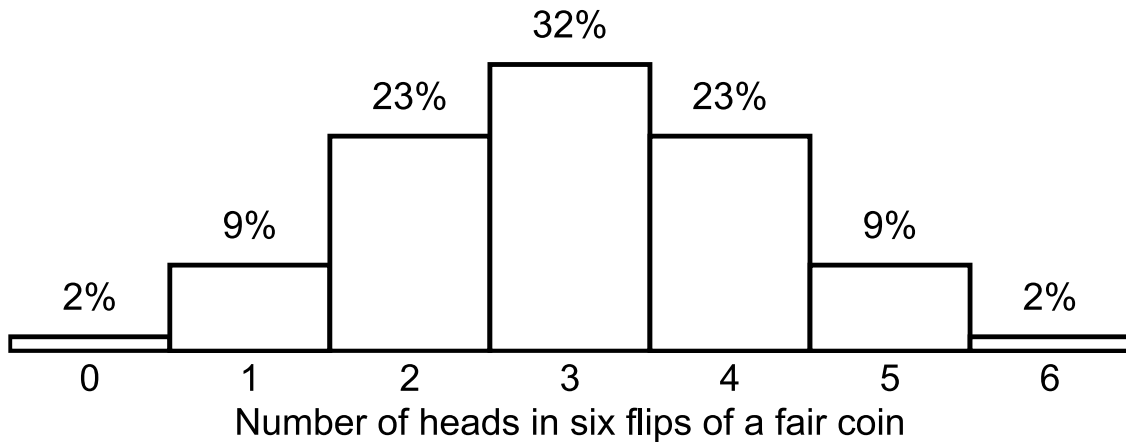
POL 138 Quantitative Reasoning in Political Science Practice · p-values

Selected p-values:

- 1 no evidence against the null hypothesis
- 0.99 a very small amount of evidence against the null hypothesis
- 0.98 a bit more evidence against the null hypothesis
- 0.50 even more evidence against the null hypothesis
- 0.05 enough evidence to reject the null hypothesis at the conventional level in political science
- 0 infinitely strong evidence against the null hypothesis

1. Suppose that we flip a coin and get 2 heads and 2 tails. What would be the p-value for a test of the null hypothesis that the coin is fair?
 0
 1
 something between 0 and 1
2. Suppose that we flip a coin and get 3 heads and 1 tail. What would be the p-value for a test of the null hypothesis that the coin is fair?
 0
 1
 something between 0 and 1
3. Suppose that we flip a coin and get 4 heads and 0 tails. What would be the p-value for a test of the null hypothesis that the coin is fair?
 0
 1
 something between 0 and 1
4. Suppose that we flip a coin and get 3 heads and 4 tails. What would be the p-value for a test of the null hypothesis that the coin is fair?
 0
 1
 something between 0 and 1
5. If the p-value is $p=0.03$ for a test of the null hypothesis, we should ____.
 accept the null hypothesis
 reject the null hypothesis
 not reject the null hypothesis
6. If the p-value is $p=0.99$ for a test of the null hypothesis, we should ____.
 accept the null hypothesis
 reject the null hypothesis
 not reject the null hypothesis
7. Suppose that the null hypothesis is that the mean height of Group A equals the mean height of Group B. If the mean height of a random sample of Group A is 180cm and the mean height of a random sample of Group B is 180cm, then the p-value will be ____.
 0
 1
 something between 0 and 1
8. Suppose that the null hypothesis is that the mean score of Group A equals the mean score of Group B. If the mean score of a random sample of Group A is 0 and the mean score of a random sample of Group B is 1, then the p-value will be ____.
 0
 1
 something between 0 and 1

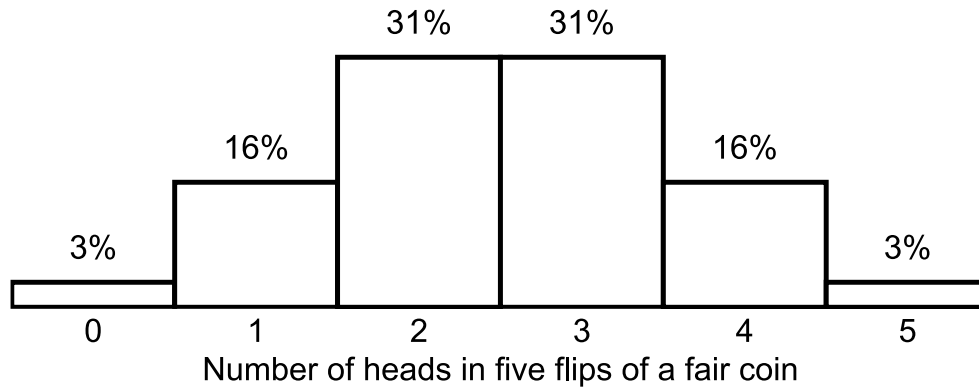
9. The histogram below indicates the percentage of times that a fair coin flipped six times is expected to land on the indicated number of heads. For example, the 23% for the column labeled 2 indicates that, if a fair coin is flipped six times over and over again, the fair coin is expected to land on heads exactly two times in the six flips 23% of the time.



A p-value is a measure of the strength of evidence that an analysis has provided against the null hypothesis. Suppose that we have a coin and that we flip the coin six times to test the null hypothesis that our coin is fair. In that case, a p-value can be calculated as the percentage of the time that a fair coin flipped six times is expected to produce an outcome that is at least as far from fairness as the outcome that we observed in the six flips of our coin. **The logic of the p-value is that, if the outcome that we observe for our coin – or a more unfair outcome – is unlikely to occur if the coin is fair, then it is unlikely that our coin is fair.** Based on this, let's complete the table below:

# of heads	Corresponding p-value for a test of the null hypothesis that the coin is fair
0 heads in 6 flips 6 heads in 6 flips	$2\% + 2\% = 4\% \rightarrow p=0.04$
1 head in 6 flips 5 heads in 6 flips	$9\% + 9\% + 2\% + 2\% = 22\% \rightarrow p=0.22$
2 heads in 6 flips 4 heads in 6 flips	$23\% + 9\% + 2\% + 23\% + 9\% + 2\% \rightarrow p=0.68$
3 heads in 6 flips	$32\% + 23\% + 9\% + 2\% + 23\% + 9\% + 2\% \rightarrow p=1.00$

[Items 10 through 15] The histogram below indicates the percentage of times that a fair coin flipped five times is expected to land on the indicated number of heads. For example, the 31% for the column labeled 2 indicates that, if a fair coin is flipped five times over and over again, the fair coin is expected to land on heads exactly two times in the five flips 31% of the time.



10. What would be the p-value for a test of the null hypothesis that the coin is fair, for a coin that landed on heads 4 times in 5 flips?

$$16\% + 16\% + 3\% + 3\% \rightarrow p = 0.38$$

12. What would be the p-value for a test of the null hypothesis that the coin is fair, for a coin that landed on heads 0 times in 5 flips?

$$3\% + 3\% \rightarrow p = 0.06$$

11. What would be the p-value for a test of the null hypothesis that the coin is fair, for a coin that landed on heads 2 times in 5 flips?

$$31\% + 16\% + 3\% + 31\% + 16\% + 3\% \rightarrow p = 1.00$$

13. What would be the p-value for a test of the null hypothesis that the coin is fair, for a coin that landed on heads 4 times in 5 flips?

$$16\% + 16\% + 3\% + 3\% \rightarrow p = 0.38$$

[Note: No items 14 and 15]

[Items 16 through 21] Data to the right are from the U.S. Department of Education's National Assessment of Educational Progress (NAEP), for the math test for Grade 4 students in 2022, based on representative samples from each jurisdiction:

Jurisdiction	Average Grade 4 Math Score
Illinois	237
Indiana	239
Massachusetts	242*
Oregon	228*
Texas	239

For each jurisdiction other than Illinois, the average Grade 4 math score for the jurisdiction was compared to the average Grade 4 math score for Illinois, to test the null hypothesis that the average Grade 4 math score for that jurisdiction equaled the average Grade 4 math score for Illinois. The table includes an asterisk (*) if and only if the p-value was $p < 0.05$ for the test of the null hypothesis regarding the comparison of that jurisdiction to Illinois.

16. Does the table contain sufficient evidence to conclude at the conventional level in political science that, at least on average and at least for Grade 4 students, schools in Massachusetts teach math better than schools in Illinois teach math?

- Yes
 No

17. Does the table contain sufficient evidence to conclude at the conventional level in political science that, in 2022, the average Grade 4 math score among students in Massachusetts differed from the average Grade 4 math score among students in Illinois?

- Yes
 No

18. Does the table contain sufficient evidence to conclude at the conventional level in political science that, in 2022, the average Grade 4 math score among students in Indiana differed from the average Grade 4 math score among students in Illinois?

- Yes
 No

19. Does the table contain sufficient evidence to conclude at the conventional level in political science that, in 2022, the average Grade 4 math score among students in Indiana equaled the average Grade 4 math score among students in Illinois?

- Yes
 No

20. Does the table contain sufficient evidence to conclude at the conventional level in political science that, in 2022, the average Grade 4 math score among students in Massachusetts differed from the average Grade 4 math score among students in Indiana?

- Yes
 No

21. Does the table contain sufficient evidence to conclude at the conventional level in political science that, in 2022, the average Grade 4 math score among students in Massachusetts differed from the average Grade 4 math score among students in Oregon?

- Yes
 No

22. Of the p-values below, which p-value indicates the strongest evidence against the null hypothesis?
- 0.01
 - 0.05
 - 0.50
 - 0.99
 - 1.00
23. Amy and Bob each conduct the same randomized experiment using samples from the same population and testing the same hypothesis in the same way. Amy has a sample size of 200 and Bob has a sample size of 900. If their experiments are testing for an effect that truly exists, which of the following would most likely be true?
- The p-value from Amy's experiment is lower than the p-value from Bob's experiment.
 - The p-value from Bob's experiment is lower than the p-value from Amy's experiment.
 - There is no reason to expect the p-value from Amy's experiment to differ from the p-value from Bob's experiment.
24. Amy and Bob each conduct the same randomized experiment using samples from the same population and testing the same hypothesis in the same way. Amy has a sample size of 200 and Bob has a sample size of 900. If their experiments are testing for an effect that does not exist, which of the following would most likely be true?
- The p-value from Amy's experiment is lower than the p-value from Bob's experiment.
 - The p-value from Bob's experiment is lower than the p-value from Amy's experiment.
 - There is no reason to expect the p-value from Amy's experiment to differ from the p-value from Bob's experiment.
25. True or false? If an experiment has a p-value of 0.000001, this p-value means that the effect size from the experiment is large.
- True
 - False
26. Suppose that we test the null hypothesis that an athlete is not using steroids. Which p-value threshold below would be more appropriate if we wanted to better avoid falsely concluding that the athlete is using steroids?
- $p=0.01$
 - $p=0.10$
27. Data from the 2022 National Assessment of Educational Progress indicated that, for the United States as a whole, the Grade 4 average math score for charter schools was 232, which was lower than the Grade 4 average math score for other schools of 236. The p-value was $p<0.05$ for a test of the null hypothesis that these scores equal each other. Is this sufficient evidence at the conventional level in political science to conclude that, compared to other schools, charter schools do worse teaching students math, at least up to and including Grade 4?
- Yes
 - No
28. If the p-value is $p=0.02$ for a test of the null hypothesis, which, if any, of the following should we do?
- accept the null hypothesis
 - reject the null hypothesis
 - neither of the above
29. if the p-value is $p=0.84$ for a test of the null hypothesis, which, if any, of the following should we do?
- accept the null hypothesis
 - reject the null hypothesis
 - neither of the above

30. Suppose that we conduct a $N=500$ randomized experiment to assess whether a treatment has an effect. In reality, the treatment has no effect. The probability that our experiment produces a p-value less than $p=0.05$ ____.

- is 0%
- is 5%
- is 50%
- is 95%
- is 100%
- can't be estimated with any certainty without additional information

31. A citation is a reference that one publication makes to another publication. Suppose that researchers develop a theory that sharing a publication on Twitter will increase the number of citations that the publication receives. Suppose also that data from all publications over the past five years indicate that the mean number of citations for a publication that has been shared on Twitter is higher than the mean number of citations for a publication that has not been shared on Twitter, with a p-value of $p < 0.001$ for a test of the null hypothesis that these means equal each other.

Would this be this sufficient evidence at the conventional level in political science to conclude that sharing a publication on Twitter increased the number of citations that the publication received?

- Yes
- No

32. Suppose that we conduct an $N=500$ randomized experiment to assess whether a treatment has an effect. In reality, the treatment has an effect. The probability that our experiment produces a p-value less than $p=0.05$ ____.

- is 0%
- is 5%
- is 50%
- is 95%
- is 100%
- can't be estimated with any certainty without additional information

33. Suppose that researchers develop a theory that sharing a publication on Facebook will increase the number of citations that the publication receives. Bob lists all publications that have been published over the past five years that have not yet been shared on Facebook. Bob randomly selects half of these publications and shares these publications on his Facebook feed, and Bob does not share the other half of these publications on his Facebook feed. After six years, Bob discovers that the mean number of citations for the publications that he shared on Facebook was exactly the same as the mean number of citations for the publications that he did not share on Facebook, with a p-value of $p=1.00$ for a test of the null hypothesis that these means equal each other.

Would this be sufficient evidence at the conventional level in political science to conclude that sharing a publication on Facebook does not increase the number of citations that the publication receives?

- Yes
- No