

Comparing Cuddling Preferences



In the *Cuddling Preferences* activity, you computed interval estimates to infer the percentage of women who prefer cuddling with their dog rather than with their partner. The same poll of British pet owners also collected data about the extent to which male pet owners preferred cuddling with their pets more than their partners. In this activity, you will be exploring the following research question:

Is the percentage of women who prefer cuddling with their dog higher than the percentage of men who prefer cuddling with their dog?

Bootstrap Test

To answer this question, you will carry out a bootstrap test using the data from the files *men-cuddle-20.tp* and *women-cuddle-20.tp*.

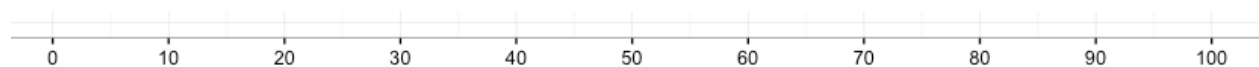
1. Compute the sample estimates for: (1) the percentage of men who prefer to cuddle with their dog rather than their partner, (2) the percentage of women who prefer to cuddle with their dog rather than their partner, and (3) the difference between the two percentages.
2. Carry out 500 trials of the bootstrap test assuming no differences between the percentage of men and women who prefer to cuddle with their dog rather than their partner. Sketch the distribution of bootstrapped differences.
3. Compute and report the p -value based on the observed result.

4. Based on the p-value you computed, how compatible is the observed difference in means with the results produced by the model specified in the null hypothesis? What does this suggest about the answer to the research question? Explain.

Compatibility Intervals

Compatibility intervals can also be used to evaluate whether there are **statistical differences** between two groups. In the *Cuddling Preferences* activity, you computed the interval estimate for the percentage of women who prefer cuddling with their dog rather than with their partner using the *women-cuddle-20.tp* data.

5. Draw the range of the compatibility interval for women (from the previous *Cuddling Preferences* activity) using the axis below.



6. Use the data from the file *men-cuddle-20.tp* to provide an answer to the research question: What percentage of men prefer cuddling with their dog rather than with their partner? To do this, estimate the percentage from the observed data, and then carry out 500 bootstrap trials to estimate the uncertainty in the estimate. Fill in the information from this analysis in the table.

Sample Size	Observed Percentage	Standard Error	Compatibility Interval

7. Draw the range of the interval estimate for men on the axis in Question #5.

Remember that the compatibility interval for women gives the percentage of women who prefer cuddling with their dogs after accounting for sampling uncertainty. Similarly the interval estimate for men gives the percentage of men who prefer cuddling with their dogs after accounting for sampling uncertainty. If both intervals include some of the same values (the intervals overlap), it provides evidence that the two groups could have the same level of preference (i.e., evidence of no difference)...or at least the uncertainty is too great for us to differentiate which group has the higher percentage.

8. Do the two intervals you drew in Question #5 overlap each other?
9. Explain using your drawing whether there is evidence that the percentage of women who prefer cuddling with their dog is higher than the percentage of men who prefer cuddling with their dog.

Effect of Sample Size: $n = 200$ (100 per Group)

How would our results change if we had used a different sample size?

Bootstrap Test

Carry out a bootstrap test using the data from the files *men-cuddle-100.tp* and *women-cuddle-100.tp*.

10. Compute the sample estimates for: (1) the percentage of men who prefer to cuddle with their dog rather than their partner, (2) the percentage of women who prefer to cuddle with their dog rather than their partner, and (3) the difference between the two percentages.

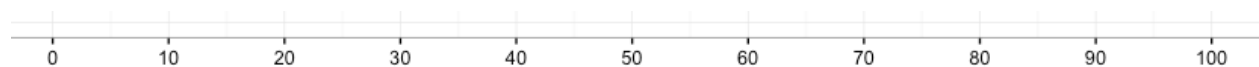
11. Carry out 500 trials of the bootstrap test assuming no differences between the percentage of men and women who prefer to cuddle with their dog rather than their partner. Sketch the distribution of bootstrapped differences.

12. Compute and report the p -value based on the observed result.

13. Based on the p -value you computed, how compatible is the observed difference in means with the results produced by the model specified in the null hypothesis? What does this suggest about the answer to the research question? Explain.
14. How does the p -value for the bootstrap test based on a sample size of 40 (20 in each group) compare to the p -value for the bootstrap test based on a sample size of 200 (100 in each group)?

Bootstrap Intervals

15. In the *Cuddling Preferences* activity, you computed the compatibility interval for the percentage of women who prefer cuddling with their dog rather than with their partner using the *women-cuddle-100.tp* data. Draw the range of this compatibility interval estimate using the axis below.



16. Use the data from the file *men-cuddle-100.tp* to provide an answer to the research question: What percentage of men prefer cuddling with their dog rather than with their partner? To do this, estimate the percentage from the observed data, and then carry out 500 bootstrap trials to estimate the uncertainty in the estimate. Fill in the information from this analysis in the table.

Sample Size	Observed Percentage	Standard Error	Compatibility Interval

17. Draw the range of the compatibility interval for men on the axis in Question #15.
18. Based on whether or not the two intervals you drew in Question #15 overlap each other, explain whether there is evidence that the percentage of women who prefer cuddling with their dog is higher than the percentage of men who prefer cuddling with their dog.

Effect of Sample Size: $n = 1000$ (500 per Group)

19. How would you expect the p -value for the bootstrap test to change if we used a larger sample size?
20. How would you expect the uncertainty in the compatibility intervals for men and women to change if we used a larger sample size?

Bootstrap Test

Carry out a bootstrap test using the data from the files *men-cuddle-500.tp* and *women-cuddle-500.tp*.

21. Compute the sample estimates for: (1) the percentage of men who prefer to cuddle with their dog rather than their partner, (2) the percentage of women who prefer to cuddle with their dog rather than their partner, and (3) the difference between the two percentages.

22. Carry out 500 trials of the bootstrap test assuming no differences between the percentage of men and women who prefer to cuddle with their dog rather than their partner. Sketch the distribution of bootstrapped differences.

23. Compute and report the p -value based on the observed result.

24. Based on the p -value you computed, how compatible is the observed difference in means with the results produced by the model specified in the null hypothesis? What does this suggest about the answer to the research question? Explain.
25. How does the p -value for the bootstrap test based on a sample size of 1000 (500 in each group) compare to the p -value for the bootstrap test based on a sample size of 200 (100 in each group)?

Compatibility Intervals

26. In the *Cuddling Preferences* activity, you computed the interval estimate for the percentage of women who prefer cuddling with their dog rather than with their partner using the *women-cuddle-500.tp* data. Draw the range of the compatibility interval for women using the axis below.



27. Use the data from the file *men-cuddle-500.tp* to provide an answer to the research question: What percentage of men prefer cuddling with their dog rather than with their partner? To do this, estimate the percentage from the observed data, and then carry out 500 bootstrap trials to estimate the uncertainty in the estimate. Fill in the information from this analysis in the table.

Sample Size	Observed Percentage	Standard Error	Compatibility Interval

28. Draw the range of the compatibility interval for men on the axis in Question #26.
29. Based on whether or not the two intervals you drew in Question #26 overlap each other, explain whether there is evidence that the percentage of women who prefer cuddling with their dog is higher than the percentage of men who prefer cuddling with their dog.
30. How is the result of the bootstrap test (i.e., the p -value) effected by sample size? Explain by comparing the p -values for the three bootstrap tests.

31. You learned how sample size effects statistical uncertainty (i.e., the range of the compatibility interval) in the *Cuddling Preferences* activity. Based on your answer to the previous question, how is statistical uncertainty related to p -value?