Feedback Activity

Permutation Question

Data + Context: These data used for this question were collected in late 2021. Students were to describe the simulation scheme for conducting a hypothesis test between US air quality and Canada air quality. Air quality was measured as a quantitative variable from 1 - 100. There were 842 observations in the US, and 155 observations in Canada.

Question: In as much detail as possible, DESCRIBE the simulation scheme for:

(a) conducting a hypothesis test for evaluating whether the average air pollution are different between the United States and Canada.

Solution: Take all air pollution values from US and Canada, and mix them all together because, under the assumption of the null hypothesis, we can assume that country is independent from air quality. Randomly shuffle observations into new US and Canada groups of the same sizes as the original US (842) and Canada group (155). Next, take the mean air pollution for each group, and subtract them in the same order as your sample statistic. Do this process a large amount of times to create a null distribution.

Student Answer: To determine whether the average air pollution levels are different between the US and Canada, we can conduct a hypothesis test. The first step in our process is to define our null and alternative hypothesis. We could then specifically use the simulation scheme of a bootstrap technique. This involves resampling the data and calculating the mean differences for both the US and Canada from the resampled data.

Feedback

– missing mixing data together or assume the null hypothesis to be true (-1)

– missing randomly shuffle or permute observations to two new groups (-1)

– missing size 842 and 155 (-1)

What questions may this student have after receiving this feedback?

Your Feedback

Notes