18.4: Randomization Association

skills to develop

- Compute a randomization test for Pearson's r

A significance test for Pearson's r is described in the section inferential statistics for b and r. The significance test described in that section assumes normality. This section describes a method for testing the significance of r that makes no distributional assumptions.

Table 18.4.1: Example data

X	1.0	2.4	3.8	4.0	11.0
Y	1.0	2.0	2.3	3.7	2.5

The approach is to consider the X variable fixed and compare the correlation obtained in the actual data to the correlations that could be obtained by rearranging the Y variable. For the data shown in Table $18.4.1$, the correlation between X and Y is 0.385. There is only one arrangement of Y that would produce a higher correlation. This arrangement is shown in Table $18.4.2$ and the r is 0.945. Therefore, there are two arrangements of Y that lead to correlations as high or higher than the actual data.

Table $18.4.1$: The example data arranged to give the highest r

X	Y
1.0	1.0
2.4	2.0
3.8	2.3
4.0	2.5
11.0	3.7

The next step is to calculate the number of possible arrangements of Y. The number is simply $N!$, where N is the number of pairs of scores. Here, the number of arrangements is $5! = 120$. Therefore, the probability value is $2/120 = 0.017$. Note that this is a one-tailed probability since it is the proportion of arrangements that give an r as large or larger. For the two-tailed probability, you would also count arrangements for which the value of r were less than or equal to -0.385. In randomization tests, the two-tailed probability is not necessarily double the one-tailed probability.