ST 516: Foundations of Data Analytics

Permutation Test

Permutation Test

- Example
- Details

Example

On January 29, 1986, the space shuttle Challenger exploded a little over a minute after it was launched.

- It was subsequently determined that the explosion was caused by the failure of an O-ring—a large, rubber seal that connected two pieces of one of the solid rocket boosters.
- The failure of the O-ring was directly linked to the cold temperature on the morning of the launch—the air temperature was between -2.2° and -1.7° C (28° to 29° F) that morning.
- Case Study 4.1.1 in *The Statistical Sleuth* contains data from shuttle launches prior to this one.

The Challenger Data

Launch Temperature	**Number of o-ring incidents**									
Below $18.3°$C	1	1	1	3						
	0	0	0	0	0	0	0	0	0	0
Above $18.3°$C	0	0	0	0	0	0	0	1	1	2

(the temperature cutoff is 65° F)

The Challenger Data

- The sample sizes are small ($n_1 = 4,\ n_2 = 20$).
- We can't reasonably say that these data are from Normal populations.
- A t-based procedure will probably not be valid for these data.
- There are lots of zeros, so a log-transformation is not recommended.
- Instead, use an alternative procedure called a **permutation test**.

Permutation Test: The Intuition

1. We assume the null hypothesis is true: The two population distributions are the same. Then, the grouping of the observations into two samples (or groups) is simply an arbitrary *labeling* of those observations into one group or the other, and every possible labeling is equally likely.

2. For a relevant statistic, T, the null distribution is the histogram of T values calculated under every permutation of the group labels.

3. We then compare the value of T calculated from the observed data (i.e., with the observed groupings)—call this T_{obs}—to this null distribution.

4. If T_{obs} is unusual relative to the null distribution, it provides evidence that the null hypothesis, upon which the null distribution is based, is suspect.

Permutation Test: Details

We have to consider how to actually perform the permutation test. We could:

(a) have a computer generate all of these permutations,

(b) have a computer randomly generate some representative proportion of these permutations,

(c) in some cases it's possible to simply list the needed permutations.

In all of these approaches, we will obtain a p-value by counting the number of permutation that result in a value of T_{obs} that is as unusual or more unusual than T_{obs} is relative to the null distribution.

Challenger Example

In the Challenger data, there are two groups, one of size $n_1 = 4$ and one of size $n_2 = 20$.

- A calculation from the area of mathematics known as combinatorics tells us that there are a total of

$$\binom{24}{4} = 10,626$$

  permutations.

- That's not actually very many—a computer can perform all of them fairly quickly.

Challenger Example

Recall that the Challenger data are:

Launch Temperature	**Number of o-ring incidents**									
Below 18.3°C	1	1	1	3						
	0	0	0	0	0	0	0	0	0	0
Above 18.3°C	0	0	0	0	0	0	0	1	1	2

If we use $\overline{X}_1 - \overline{X}_2$ as a test statistic, then

$$T_{obs} = \overline{X}_1 - \overline{X}_2 = 1.5 - 0.2 = 1.3.$$

Let's examine T under a few permutations.

The Challenger Data

Here's one permutation:

Launch Temperature	**Number of o-ring incidents**									
Below $18.3°$C	0	0	0	0						
	0	0	0	0	0	0	0	0	0	0
Above $18.3°$C	0	0	0	1	1	1	1	1	2	3

Here, $T = 0 - 0.5 = -0.5$. Notice that there are several permutations out of the possible 10,626 that result in the table above—since there are a total of 17 zeros, there are several ways that 4 of them can end up the in 'low' temperature group.

The Challenger Data

By using combinatorics, you can calculate that there are only 105 out of the total 10,626 possible permutation that result in values of T that are as large or larger than $T_{obs} = 1.3$.

- That's not very many!
- In fact, we can calculate a p-value by simply taking 105/10626 = 0.0099.
- This provides strong evidence against the null hypothesis that the two populations are the same.

Permutation Test in R

There are a number of packages in R that will perform two-sample permutation tests:

coin
perm
exactRankTests

You will investigate some of these packages in the computer lab activity for this Module.