Statistical Hypothesis Testing and Baseball

A *t-test* is a statistical test designed to accept or reject a null hypothesis. For example, in major-league baseball conventional wisdom holds that left-handed batters have an advantage over right-handed batters, but *can show that this is statistically true with a 95 percent confidence level?*

For a given data set, the t-test will generate a *test statistic*, which is a single number that will help you accept or reject the hypothesis. *Welch’s t-test* is designed specifically to compare sample sets from two populations.

For this question, our two data samples are: batting averages for right-handed and left-handed batters. We assume that the samples from each population follow a normal distribution. So, our null hypothesis is: the population mean from the left-handed set is equal to the population mean from the right-handed set. Or in plain English: *there is no statistically significant difference between batting averages of left-handed and right-handed players.*



According to Welch’s test, the t-statistic is calculated as follows:



We also need to calculate the degrees of freedom for the sample sets. Once we have these values, we can calculate the p-value. Conceptually, the p-value is the probability of obtaining a t-statistic *at least* as extreme as the one that was actually observed assuming the null-hypothesis is true. For the test, we will select a criticalvalue of p, such that if p < Pcritical, we *reject* the null hypothesis. If p > Pcritical, we *accept* the null hypothesis.

Now, you are probably saying, “is there a easy way to do this test in Python?” Indeed, the answer is *yes* we can use the pre-defined function, *ttest\_ind*, in SciPy.

After coding this up in Python and importing a baseball statistics data set consisting of 18,177 records, courtesy of Sean Lahman (<http://www.seanlahman.com>) we get the following result:

RH average = 0.163788652864

LH average = 0.185389054514

t-statistic = 9.9357022262420944, p-value = 3.8102742258887383e-23

result = False

As we can see from the result, we must *reject* our null hypothesis at a 95% confidence level. (In other words, in major-league baseball, *left-handed batters ARE statistically better than right-handed batters.*

Source code and data set can be found on github at:

https://github.com/bwblock/baseball