BEHIND STATISTICAL HYPOTHESIS TESTING

by Zeynep Kalaylioglu
Department of Statistics, METU

Following script assumes that following basic concepts of statistical hypothesis testing, namely test statistic, p-value, statistical distribution, Type I and II errors and significance level are known to the reader.

I will try to illustrate "behind the scene" of statistical hypothesis testing and my target audience is undergraduate students of any discipline. What lies behind the scene of statistical hypothesis testing is proof by contradiction, one of the finest methods of *logic (and of mathematics, needless to say)* to prove a proposition.

Proof by contradiction is a powerful method to prove a proposition. You start with assuming that the proposition is false. Then if it leads to a contradictory situation, this implies that the assumption with which you had started must be false implying that your proposition was true. Proof is complete.

The great Avicenna (Ibn Sina) used this method to prove important philosophical theorems, for instance he provided a proof of incorruptibility of intellect using proof by contradiction method.

What does this have to do with statistical hypothesis testing? Everything.

We have a hypothesis to test. That is, we have a proposition to prove (based on the observed data). Let's denote this proposition by H_1.

Now let us apply the method of proof by contradiction to test the truthness of H_1.

We start with assuming that our proposition is false. That is to say that we assume complement of H_1, i.e. H_1^C. Let's use a shorthand notation for that, e.g. H_0.

First of all, under H_0, i.e. in a world in which H_0 is the true situation, we have a test statistic with a particular distribution. We compute a p-value based on our observed data using this test statistic, assuming this distribution. This leads to a p-value which is either greater or smaller than α (nominal Type I error). Under the assumption of H_0, p-value being the probability of occurrence of observations supporting H_1 due to chance alone, is a large value, it can not be small. So, if starting assumption (H_0) leads to a large p-value, there is no contradiction. If, on the other hand, it leads to a small p-value, then this is a contradiction with the starting assumption. That means, if you find out that your p-value is small, your starting assumption must be false. That is to say you reject H_0 if your p-value is small (there you go: that's our usual statistical testing procedure). That proves that your proposition (i.e. H_1) is true. In the face of statistical hypothesis testing in which Type I and Type II errors are present, it is better to say that "that provides evidence for your proposition (i.e. H_1) being true".