

On the Language of Hypothesis Testing

Many statisticians will tell you that the first step in hypothesis testing is to rip out of the manager's head all prior knowledge of relevance, and all cost-related knowledge, and summarize all that in a single number called the "critical significance level," or "the significance level of the test."

Then, they'll tell you to **accept** the null hypothesis if the actual significance level of the data (the p-value) is greater than the critical significance level, and **reject** the null hypothesis otherwise.

This means, for example, that if the actual significance level of the data with respect to some null hypothesis was 2.35%, the statistician might summarize his/her work by merely telling you that "the data was significant at the 5% level," or "we rejected the null hypothesis at the 5% level."

As I noted in class, no one knows how to pull all the knowledge of relevance out of a manager's head, so it's hard to carry out this program in practice.

One of the reasons this language was developed was to avoid needing to print hundreds of t-tables, one for every different number of degrees of freedom. You'd never be told the precise significance level (or p-value) of the data. Instead, tables would just say how large the standardized observation ((estimated mean – hypothesized mean) / standard error of the mean) had to be in order to cross the 5% threshold. This reason is completely obviated by today's computer technology.

The old language ends up equating significance levels of, for example, 4.7%, 2.34%, and 1.175%, where in actuality they are as different as are observations of 4, 5, or 6 consecutive Tails in our coin-tossing story.

On your exam for my section of this course, any hypothesis testing questions will focus only on determining the significance level (p-value) of the data. Several of the problems in the Session-4 "practice problems" are stated in "significance level of the test" format. Look at them before you start DECS-431: Translating to a focus on the data (instead of the "test") becomes easy.