Lessons 10-3 and 10-4 Using Significance Tests and Inference as a Decision

# Type I and Type II Errors

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | *Truth about the population* | |
|  |  | Ho True | Ho False |
| *Decision based on our sample* | Reject Ho | Type I Error | Correct Decision |
| “Accept” Ho | Correct Decision | Type II Error |

# Error probabilities

The significance level α of any fixed level test is the **probability of a Type I** error. That is, α is the probability that the test will reject the null hypothesis Ho when, in fact, Ho is true.

In many textbooks, the **probability of a Type II error** is denoted by β. The **power** of a test against any alternative is 1 minus the probability of a Type II error or 1 - β. In the significance test setting, it is more common to report the probability that a test *does* reject *H*0 when an alternative is true. This probability is called the ***power*** of the test against that specific alternative. The higher this probability is, the more sensitive the test is. Power calculations are important in planning studies. Using a significance test with low power makes it unlikely that you will find a significant effect even if the truth is far from the null hypothesis. A null hypothesis that is in fact false can become widely believed if repeated attempts to find evidence against it fail because of low power. The best advice for maximizing the power of a test is to choose as high an α level (Type I error probability) as you are willing to risk *and* as large a sample size as you can afford.

# Making decisions

P-values are more informative than the reject-or-not result of a fixed alpha level. Beware of placing too much emphasis on the traditional values of alpha such as » = .05. Plot the data to display the effect you are seeking and use confidence intervals to estimate the actual values of the parameters.

# What can go wrong?

Plot your data and examine them carefully. Are there outliers or other deviations from a consistent pattern? A few outlying observations can produce highly significant results if you blindly apply common significance tests. Outliers can also destroy the significance of otherwise-convincing data. **The foolish user of statistics who feeds the data to a calculator or computer without exploratory analysis will often be embarrassed.**

*P-*values are more informative than the reject-or-not result of a fixed level α test. Beware of placing too much weight on traditional values of α, such as α = 0.05. On the other hand, lack of significance does not imply that *H*0 is true, especially when the test is based on just a few observations.