# Week 7

Refer to the course map

This week we refine our study of statistical tests by introducing the notion of *power*. As we will see, this can be of great help, especially when analyzing a non significant test.

## Objectives for this week

* Understand the notion of power, and its use
* Compute the power function for simple tests, and discuss the implications.
* **Optional, "advanced" work**: Compute the power function for non Gaussian tests

## Reading Material

We discuss the notion of "power" as it illuminates what the result of a test (especially, when the result is "not significant") means. In particular, we will discuss the notion, we will informally call "resolution power" of a test. (here is the PDF version)

Our companion course is less expansive on this than it is on other subjects, but we can look up a couple of examples , to help us understand this notion better.

To gain a better understanding of this subject, which is not as well represented as others in most textbooks, please, check this additional file (PDF format). You can also look at a possible setup for a one-tailed test (including power calculations), both using a normal distribution and a Student distribution in this spreadsheet: the yellow-marked cells are for data entry, while the others are calculated cells. The point is not in the numbers, but in the formulas. There are version in Gnumeric, Open Document Format (OpenOffice/LibreOffice/Kspread), and Excel format.

#### **Please Take Note Of A Problem With The ODS and XLS Versions**

The way the spreadsheet function tdist (the "survival function" of the Student distribution) is implemented in OpenOffice (and, as far as I can tell, in Excel), produces an error when the input "*x*" is negative. This used to be the case in the previous versions of Gnumeric too, but it seems to have been fixed there. If you open these files, the power column, for the Student case has many entries showing an error. To me, this is a bug, even if it is easily managed, by a small change in the formulas, relying on the symmetry of the t-distribution (thus tdist(-x,...)=1-tdist(x,...)), but it's there, so be aware of it.

## Assignments

* Apply the theory to the simulated data you downloaded in week 1: compute the power function of your tests, as developed in week6. In particular, concentrate on the non significant ones, and discuss the "resolution" of your tests. Repeat using only part of the full samples (e.g., the first half), and compare the power functions you obtain.
* Look up this PDF file containing a data set, and questions for this and the last three weeks. Answer the questions under "Week 7". **Do not type the data into a spreadsheet: that is both tedious, and terribly prone to errors**. Rather, use the summary data in the file to do the computations you need (you may want to use a calculator, but the cheapest one will do)

**Optional, "advanced" work**: test the "advanced" data sets, comparing how a "naive" application of normality-based tests performs in non normal cases. If possible, repeat for the data you simulated in week 4. In particular, you could try to see how the power function improves, as you work with larger samples.

**Please, turn in your work by Monday, Week 8**

### Quizzes

Unfortunately, there were no quizzes available on this week's material on WAMAP.