# Week 5

Refer to the course map

This week we take on our first inferential problem: estimating the mean and variance of a distribution, from a sample. As usual, we first try to focus on the core ideas. Next, we can browse over a number of situations. You may notice that practically all cases are normally distributed, or assumed to be close to normal. This is not the only case you may encounter, but it is the most common (and, just to reiterate, it is sometimes assumed, more out of habit than out of solid grounds)

## Objectives for this week

* Understand the notion of "interval estimate", and "confidence level": apply to simple estimation cases
* Calculate interval estimation at various confidence levels, and calculate *p* - values for expected values of normal variables, with known and unknown variance
* **Optional, "advanced" work**: Verify how tools developed with normality assumptions work when that assumption is unwarranted. Apply the general method to rigorously estimate the mean of exponential variables. Explore the effect of using methods for exponential variables to Weibull variables. If you can, apply this also to the data simulated in week 4.

## Reading Material

We start with a very short general introduction on inferential statistics

We then move to discuss how to use our knowledge of probability tools to provide *quantitative* estimates of interesting features of the random variables we are using to model our observations (here is the PDF version)

And now, we can go into more detailed "how-to" for a number of situations.

## Assignments

* Refer to the simulated data you downloaded back in weeks 1 or 2. Use the techniques discussed in this chapter to find interval estimates for the expectations of the underlying distributions. Try different confidence levels, and compare with the "theoretical" means (that is, the ones fed to the simulating program)
* Look up this PDF file containing a data set, and questions for this and the following two weeks. Answer the questions under "Week 5". **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)

**Please, turn in your work by Monday, Week 6**

**Optional, "advanced" work:** refer to the "advanced" simulated data we downloaded in week 1, estimating the mean of the data (note that the Cauchy data does not have a theoretical mean at all). Also, if possible, compare the results form a "normal-distribution" based estimate, to the "rigorous" "chi-square" based method when working with exponential data (please, contact your instructor for details).

### Quizzes

You can get credit for participation by taking the *quizzes 36 to 50* on the [WAMAP](http://www.wamap.org/) site: this group is about interval estimation for normal (or approximated by normal) variables.