# Week 4

Refer to the course map

This week we conclude our brief exploration of *Probability Theory*, the main mathematical tool used in statistical modelling. We look at useful ways to work with several random variables at once.

## Objectives for this week

* Understand the concept of Independent vs. Dependent Random Variables: recognize simple examples and distinguish between apparently similar, but actually very different (almost opposite) concepts, such as *independent events* vs. *disjoint events*.
* Formulate clearly the basic limit theorems (in their simplest form), and recognize when they apply: the *Weak Law of Large Numbers*, and the *Central Limit Theorem* for independent, identically distributed random variables. Apply them to simple examples, as in stating the limit distribution for a sequence of Binomial experiments.
* **Optiona, "advanced" work**: recognize how the two main limit theorems work when applied to skewed or "fat tailed" distributions)

## Reading Material

In addition to last week's items, we look at the file on multiple random variables HTML/MathML (PDF format)

The ideas introduced there can now be applied to the study of *sampling distributions* (what you have to work with when you take a sample for your statistical work)

## Assignments

Since this is a theoretical chapter, we keep working on a few "pencil and paper" problems

**Please, turn in your work by Monday, Week 5**

**Optional, "advanced" work**: Use your spreadsheet to simulate samples from different distributions, and build histograms *for the means* you obtain. **Get in touch with your instructor for more detailed instructions.** A quick introduction to simulation is in this file (PDF). What do you observe when simulating

* A big (say, 50, or 100) sample of Bernoulli (yes/no) variables with *p* close to 0.5?
* What if the sample is now from variables with *p* = 0.01?
* What if the sample is form an exponential distribution?
* What if the sample is from a Cauchy Distribution? You may wan to check the comments on "fat-tailed" distributions (XHTML), (PDF version).

The goal is to explore, "experimentally", how the Law of Large Numbers and the Central Limit Theorem work (or do not work) in different circumstances. We will keep this data, and apply the statistical tools coming up in the next few weeks, to "experiment" how these behave on different types of data.

### Quizzes

The quizzes were already introduced last week. You can get credit for participation by taking the

1. *quizzes from 16 to 25* from the quizzes on the [WAMAP](http://www.wamap.org) site: this group is about finite probabilities
2. *quizzes from 26 to 35* of quizzes on the [WAMAP](http://www.wamap.org) site: this group is about the normal and, occasionally. continuous distributions. Note that many problems use the normal distribution, motivated by the Central Limit Theorem.