1. Suppose you have $n$ $t$-stats $t_1, \ldots, t_n$. These might be, for example, the values of $n$ different autocorrelation coefficients, or the outputs of $n$ different models applied to the same set of data. We learned in class that a $t$-stat whose absolute value is larger than about 2 indicates statistical significance.

Suppose that all of these $n$ models are insignificant. What is the probability that at least one of the $t$-statistics will indicate significance?

2. I will put three data series on the course Web page. These will all be samples generated from a process of the form

$$x_j = x_{j-1} + \sigma_j \epsilon_j.$$ 

But I will use three different rules for generating the data sequences:

(a) $\sigma_j \equiv \sigma$ is constant, and the $\epsilon_j$ are i.i.d. standard normal.

(b) The $\epsilon_j$ are i.i.d. standard normal, but $\sigma_j$ depends on $j$ in a “smooth” fashion: it has some consistent trend through the data sample but it does not vary randomly.

(c) $\sigma_j \equiv \sigma$ is constant, and the $\epsilon_j$ are i.i.d. but not normal.

Which sample is which, and how can you tell?

3. We have a course account at WRDS:


The login and password will be given in class.

From the CRSP database, download at least 5 years of IBM daily price and return ("holding period return"). Find out how CRSP defines return; why is it not the same as the change in price, or log price? Plot the price, and plot the returns. Test the returns for normality and independence as discussed in class.