EXERCISE ON BREAKS, PROBABILITY AND CONFIDENCE INTERVALS

Due Thursday, 9/22

On the course web site (sims.princeton.edu/yftp/Times16) there are two files named BreakX.txt and BreakX.RData. Each contains 100 values of a time series we'll call $x_t, t = 1, \dots, 100$, which was drawn from a $N(\mu_1, 1)$ distribution for $t \leq S$ and from $N(\mu_2, 1)$ for $t > S$. Conditional on μ_1, μ_2 and S, the draws are independent across t.

(1) Write down the likelihood as a function of μ_1, μ_2 and S.

(2) Taking the prior distribution as uniform across 1 to 99 for S and uniform over $\mathbb{R}^2$ for μ_1, μ_2, calculate the marginal density for S and plot the result. This can be done analytically. It involves completing the square for the exponent in the likelihood, so that it is recognizable as shaped like a normal pdf for μ_1 and μ_2.

    The uniform prior on the two μ's is of course not a proper prior, since it doesn't integrate to one, but results with it will be almost the same as if we had just used a normal prior on them with very large variance.

(3) Calculate the HPD (highest posterior density) 90% posterior probability interval for S. This interval is the shortest interval that has posterior probability .9.

(4) Describe how you could construct a 90% confidence set for μ_1, μ_2, and S jointly. This probably requires discretizing μ_1, μ_2 space and doing numerical integration or making simulated draws from the data distribution. It will imply a test for the μ_1, μ_2 S null hypothesis at each such point in the parameter space, and for the interval to be useful the test must be constructed well. For extra credit, you can actually do the calculations.

(5) Discuss whether what you do, or propose, in 4 can be used to provide a confidence interval for S alone, and how such a confidence interval might best be calculated or approximated.

Date: September 18, 2016.