EXERCISE ON BAYES BASICS

Here's an i.i.d. sample of ten observations from a Gamma(n, α) distribution, where neither n nor α is known.

2.1195487	0.6376608	2.1838401	0.7427399	1.3834716
2.7273786	3.3355102	4.2340466	1.6721151	0.6676172

Using a flat prior:

(1) Write out the posterior density (i.e., since the prior is flat, the likelihood function).

(2) Plot its level curves. Treat both the degrees of freedom parameter n and the rate parameter α as continuously distributed on $(0, \infty)$. It is probably best to first calculate the log likelihood on a grid of n and α values, then exponentiate it to get the level curves. In R you would want to use the **`lgamma`** and **`contour`** functions as you do this.

(3) Find 60% and 90% joint posterior probability regions for n and α and plot them. (One approach: Sort the points in your grid by the height of the posterior density (in such a way that you can undo the sort — e.g. **`order`** in R); find the point in the cumulative sum (**`cumsum`** in R), starting at the high-posterior-density end, where you hit the desired 90% or 60% level, then plot with big dots all the points in the original grid that correspond to the high-posterior density values.)

(4) Find 60% and 90% marginal posterior probability intervals for n and α separately and plot the two parameters' marginal posterior densities.

(5) Think about whether there is any way to get frequentist confidence regions or intervals for this problem. The best I could come up with is the asymptotic approximations that one gets by using the second-order Taylor expansion of the log likelihood around its peak to arrive at a Gaussian density for the MLE. If you can't think of anything better, calculate the confidence region and intervals at the 60% and 90% levels that come from that approximation. Note that with a sample of size 10, relying on the asymptotic approximation is probably not a good idea. Another approach would be to do a bootstrap — generate many artificial samples of 10 by sampling with replacement from the observed sample, calculating MLE's for these artificial samples, and using the results to form confidence regions and intervals. The bootstrap also has at best asymptotic justification, however, and could be quite unreliable here.

Date: September 24, 2015.