Chapter 2 Portfolio Contribution

Required for undergraduate and graduate students:

Assuming a non-informative prior, please obtain posterior distributions for the probability of successful free throw shooting based on a sample of size 5 (with 1 successful free throw and 4 misses) and on a sample of size 20 (with 4 successful free throws and 16 misses). Create a plot which displays the non-informative prior and posterior distributions. Write a paragraph to describe your plot, in your own words. Comment on what you perceive to be the consequences of a non-informative prior regarding posterior modes and posterior uncertainties, versus the informative prior from which I created {Chapter2bNotesFig1.pdf}. Please include an appendix with your R code.

To accomplish the above, you can proceed most easily by referring to the file {R code for Chapter2bNotes.txt} and adapting the R code leading to {Chapter2bNotesFig1.pdf}. Changing 2 and 5 in that R code, when dbeta is called, to 1 and 1 will make the prior non-informative. If proceeding in this manner, please (as a matter of academic propriety) acknowledge such adaptation of the existing R code in the appendix.

Required for graduate students only:

Use R to obtain a 95% highest posterior density credible interval for p, corresponding to the non-informative prior and the larger data set. In this problem, the posterior probability that p falls between a and b is given by pbeta( b, 1+4,1+16) - pbeta( a, 1+4,1+16). So, you want to find a and b such that

pbeta( b, 5, 17) - pbeta( a, 5, 17) = 0.95 [Condition 1]

and

dbeta( b, 5, 17) = dbeta( a, 5, 17). [Condition 2]

Condition 1 guarantees that you have a 95% interval, and condition 2 guarantees that the interval is highest posterior density. A way to enforce Condition 1 is to set a = qbeta(c, 5, 17) and b = qbeta(c + 0.95, 5, 17), where c is any number between 0 and 0.05. Then you can satisfy Condition 2 by finding c such that b - a is minimal.\*

\*If you have studied calculus, I can explain why minimizing b - a satisfies Condition 2. If you have not studied calculus, you may simply take for granted that it is so. Here is the explanation: Condition 1 entails that the integral of the posterior density from a to b is constant (indeed, equal to 0.95), so that changing a requires b to change as well. In other words, b is a function of a. We can minimize b - a by taking its derivative with respect to a and setting that equal to 0. So, we obtain db/da - 1 = 0 or db/da = 1. On the other hand, the Leibniz integral rule tells us that the derivative of the aforementioned integral is the posterior density at b, multiplied by db/da, minus the posterior density at a. Because the aforementioned integral is constant, its derivative is 0. Combining this with db/da = 1 gives that the posterior density at b minus the posterior density at a equals 0, which is the same as Condition 2.