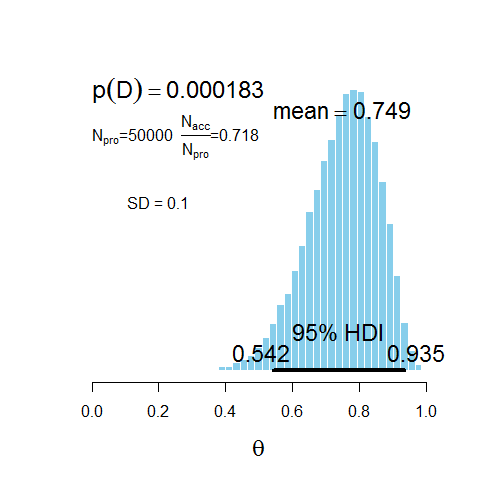
Jon Janelle

MAT 500

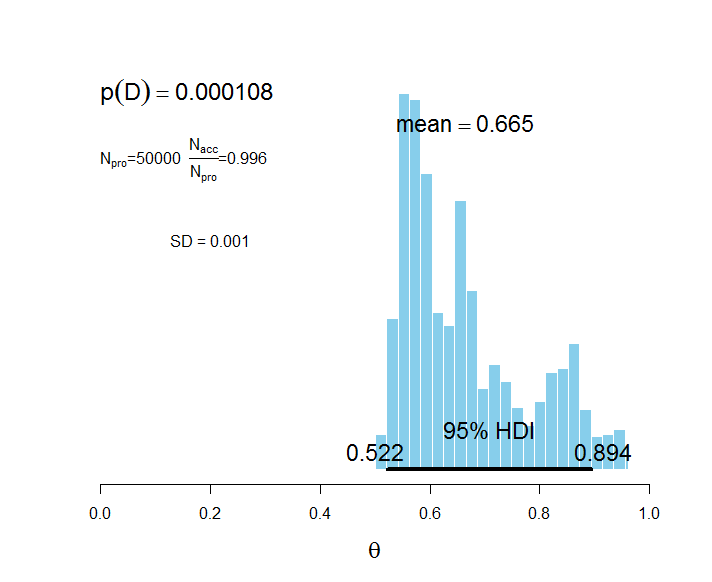
7.1, 7.3, 7.5

**Chapter 7 Homework**

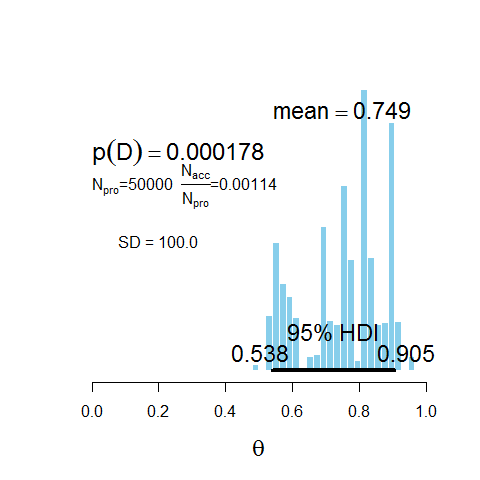
**(7.1A)**  The graphs shown in parts A, B, and C show the results of the metropolis algorithm when proposed jumps are generated with mean 0 and standard deviations of 0.1, 0.001, and 100, respectively. In each case the starting value is 0.5, 55556 jumps are attempted, and the first 10% of attempted jumps are discarded as a burn-in period.



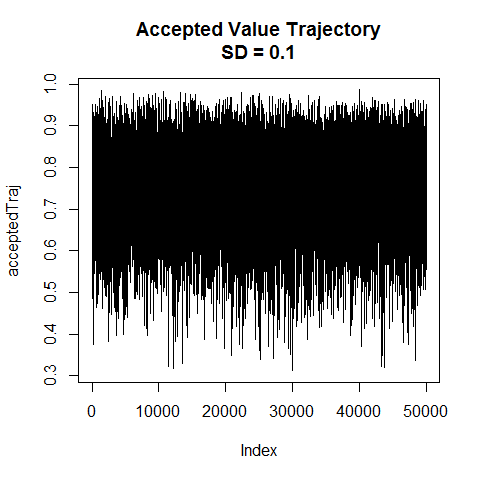
**B)**



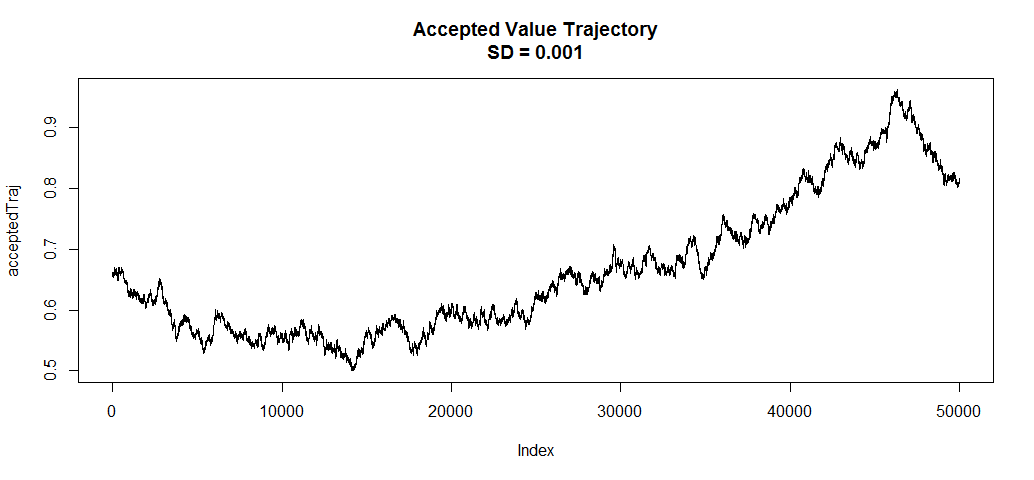
**(C)**



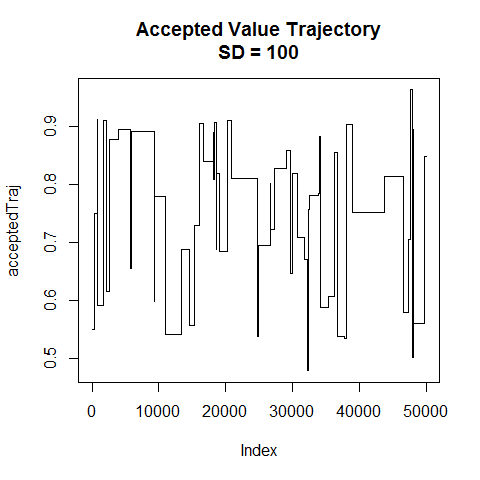
**(D)** The proposal distribution in part A, where SD = 0.1, is the most accurate representation of the posterior. As the plot below shows, the range of the posterior distribution is well covered by the accepted values, and there do not appear to be high levels of correlation between points.



The proposal distribution in part B had the fewest rejections with 99.6% of the proposed jumps being accepted. The small standard deviation caused all of the proposed values to be very close to their preceding values, which greatly reduces the probability that proposed jumps fall outside the acceptable range. Consequently, the majority of proposed values were accepted. This distribution does not represent the posterior well because the small jump sizes result in the random walk moving very little at a time, and thus the entire range of the distribution is not represented. The plot of accepted values below shows that each accepted value is highly correlated with the previous values and that the range of the distribution is not well covered by the accepted values.



The proposal distribution in part C had the greatest number of rejections with only 0.114% of proposed jumps accepted. This occurred because the large standard deviation caused many proposed values to fall outside of the acceptable range. The plot of accepted values shows that there are frequently intervals during which the accepted trajectory remains constant. This occurs when many consecutive proposed values are rejected, and the result is that the distribution is poorly represented and successive points are highly correlated.

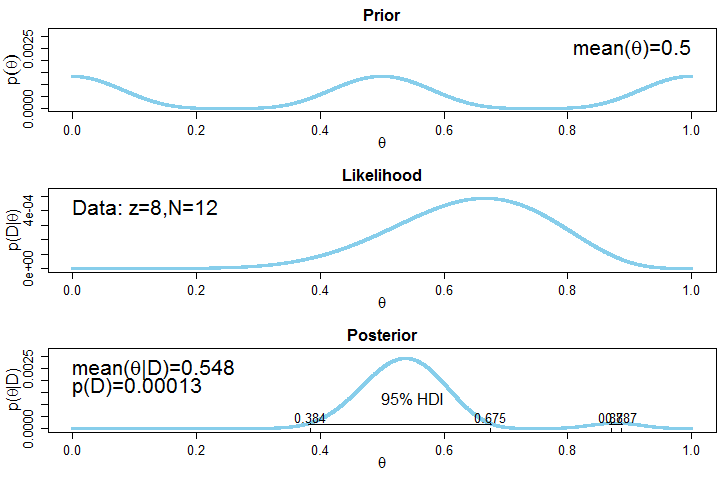


**(E)** To determine which proposal distribution generates the most accurate representation of the posterior, the correlation between accepted values in the trajectory can be examined. One way to do this is by creating line graphs, like those shown in part D, of the trajectory of accepted values. If the graphs show regular intervals of little movement, as is the case when SD = 0.001 and SD = 100, then this is indicative of a high levels of autocorrelation. The trajectory with the lowest level of autocorrelation, the SD = 0.1 case in this example, is most likely to be an accurate representation of the posterior.

**(7.3A)** We have that , where Z is the appropriate normalizing constant. Given that a sample of 8 heads and 4 tails was observed, the likelihood is . Putting these together yields:

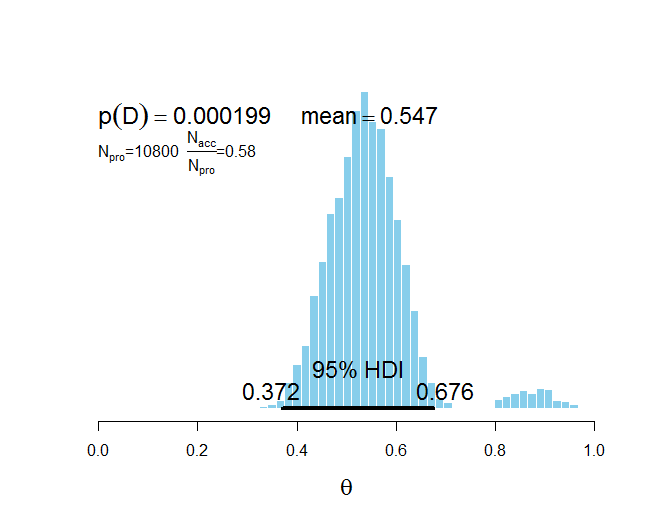
I do not believe that the prior and likelihood are conjugate because the posterior will not likely resemble the prior at all. Finding an analytic solution to verify this requires solving and, if this is possible, it will result is something far different from the prior.

**(B)** The bin width for the grid is , and the relative weighting of each grid point was calculated as . Forming the prior requires the appropriate normalizing constant, which is the sum of the relative weightings, so the prior was calculated as at each grid point. The plot shows the prior, the likelihood given a sample of N = 12 flips with Z = 8 heads, and the posterior distribution. The 95% HDI shows that it is credible to believe that the coin is fair or, to a lesser extent, weighted toward heads, but a strong tails bias is unlikely.



**(C)** Using the Metropolis algorithm with the prior , the posterior histogram below was generated. 12000 jumps were attempted, and the first 10% of these jumps were discarded as a burn-in period.

It is not necessary to first normalize the prior. To decide when to accept or reject a proposed value , the ratio is calculated, where K is a normalizing constant and the current value. This simplifies to , which demonstrates that the normalizing constant is not relevant.



The value of p(D) shown on the histogram, and the shape of the posterior distribution in general, is very similar to that obtained in part B using a grid approximation. Both methods generate correct values of p(D).

**(D)** Based on the JAGS 3.3 User Manual, there is not a distribution built into JAGS that allows the prior distribution used in this example to be specified.

**(7.5)** The plots below show that p(D|M1)=1.018e-06, p(D|M2)=9.049e-05, and p(D|M3)=0.00043. Model 3, the heads biased model, is preferred since the sample of 14 flips with 11 heads is much more probable within it than within the other two models.

