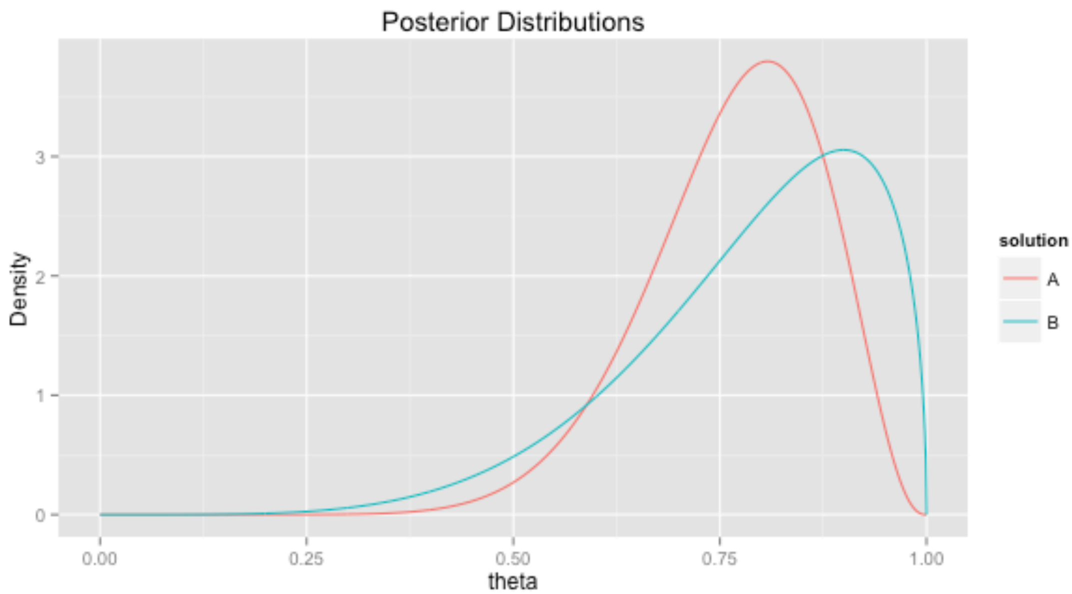
Kelvin Niu

Lab 3

Question 1

$\theta_A \mid data_A \sim \text{Beta}(0.5 + 11, 0.5 + 3) = \mathbf{Beta(11.5, 3.5)}$

$\theta_B \mid data_B \sim \text{Beta}(0.5 + 5, 0.5 + 1) = \mathbf{Beta(5.5, 1.5)}$



Question 2

$P(\theta_A > 0.8 \mid data_A) = 1 - \text{pbeta}(0.8, 11.5, 3.5) = \mathbf{0.4204}$

$P(\theta_B > 0.8 \mid data_B) = 1 - \text{pbeta}(0.8, 5.5, 1.5) = \mathbf{0.5355}$

Question 3

To estimate $P(\theta_B - \theta_A > 0 \mid data_A, data_B)$, I performed a Monte Carlo simulation. First, I drew 10000 samples from the posterior of A and 10000 samples from the posterior of B. Then, I took the difference between the posterior of B samples and the posterior of A samples and calculated the proportion for which the difference was positive. Thus,

$P(\theta_B - \theta_A > 0 \mid data_A, data_B) \approx \mathbf{0.5733}$

Question 4

The data suggests that solution B has a higher success rate than solution A. It also suggests that solution B's success rate is more likely to be greater than 80% compared to solution A's success rate. Thus, since we have no additional information, I would suggest that, in general, Jennifer choose solution B.

However, for lower values of theta, the densities for the posterior of B are greater than those for the posterior of A (looking at the plot in Question 1). Thus, because solution B's sample size is small (thus making its posterior prediction less stable), it may be a better choice for Jennifer to choose solution A if errors are penalized heavily.

The best thing to do is to conduct more trials of solution A and (especially) solution B, if possible. With more data and greater sample sizes, we can get a more accurate idea of both solutions' success rates and compare them.