STA 6360, Report 1.4

Carson Slater *Baylor University*

1

The MLE for θ, the true proportion of those who think device B is more irritable than device A, is $\theta \approx 0.8125$. We employ JAGS to find th posterior interval estimate.

Table 1: 95% credible set for Theta using a Beta(2, 2) prior.

	2.5%	97.5%
Theta	0.6321991	0.89646

Because the 95% credible set does not contain 0.60, but contains values that are all greater than 0.60, we can reasonable conclude that device B is 'substantially more irritating' than device A.

2

Table 2: 95% credible set for Theta using a uniform prior.

	2.5%	97.5%
Theta	0.6464031	0.9105104

If the cost of a type I error is $19\times$ more costly than the cost of making a type II error, we can use the following decision rule to reject M_1 in favor of M_2.

$$\Pr(\theta \in \Theta_0|\mathbf{x}) = \Pr(\theta \geq 0.6\,|\mathbf{x}) < \frac{L_{01}}{L_{01} + L_{10}} = 0.05.$$

Prob: 0.9937
Everytime I run this script, I get a different value for $\Pr(\theta \in \Theta_0|\mathbf{x})$, but it hovers around 0.99, which is greater than 0.05. Therefore, we would conclude that device B is 'substantially more irritating' than device A.

3

Table 3: 95% credible set for Theta using a Beta(2,3) prior.

	2.5%	97.5%
Theta	0.6068483	0.875278

Prob: 0.9795
Similarly to (2), get a different value for $\Pr(\theta \in \Theta_0|\mathbf{x})$, but it hovers around 0.98, which is greater than 0.05. Therefore, we would conclude that device B is 'substantially more irritating' than device A.

4

Table 4: 95% credible set for Theta using a Beta(3,2) prior.

	2.5%	97.5%
Theta	0.6412141	0.8959734

Prob: 0.9937

Similarly to (3), get a different value for $\Pr(\theta \in \Theta_0|\mathbf{x})$, but it hovers around 0.99, which is greater than 0.05. Therefore, we would conclude that device B is 'substantially more irritating' than device A.

5

Table 5: Bayes Factors for Problems 1-4.

	1	2	3	4
Bayes Factor	205.0065	236.5834	218.8753	142.8198
Evidence for M1	Very Strong	Very Strong	Very Strong	Strong

6

The MLE for θ, the true proportion of those who think device B is more irritable than device A, is still $\theta \approx 0.8125$. We first compute credible sets from JAGS in light of $y = 52$ and $n = 64$ for all four priors. The results are in the tables below.

[1] 1 "Prob: 0.9998"
[2] 1 "Prob: 1"
[3] 1 "Prob: 0.9999"
[4] 1 "Prob: 0.9993"

For all four posterior probabilities, all of them exceed 0.05, which would lend to the conclusion that given all four priors on θ, device B is 'substantially more irritating' than device A.

Table 6: Bayes Factors for Problem 6, but redone for new data.

	6.1	6.2	6.3	6.4
Bayes Factor	9202.872	Inf	45803.94	1292.622

All of these bayes factors for each of the four priors indicate very strong evidence in favor of M1 ($\theta > 0.6$). Ultimately, the change in sample size drastically overwhelmed the prior in each case, and highly inflated the Bayes factors.

7

Given a sample $Y_1, \ldots, Y_n \sim \text{Ber}(\theta)$, we conduct the appropriate frequentist hypothesis test for $H_0 : \theta \geq 0.6$ against $H_1 : \theta < 0.6$, with an $\alpha = 0.05$ significance level. We know that given H_0, we can compute the probability that $\Pr(Y < y|\, \theta = 0.6)$.

For $y = 26$, $n = 32$:

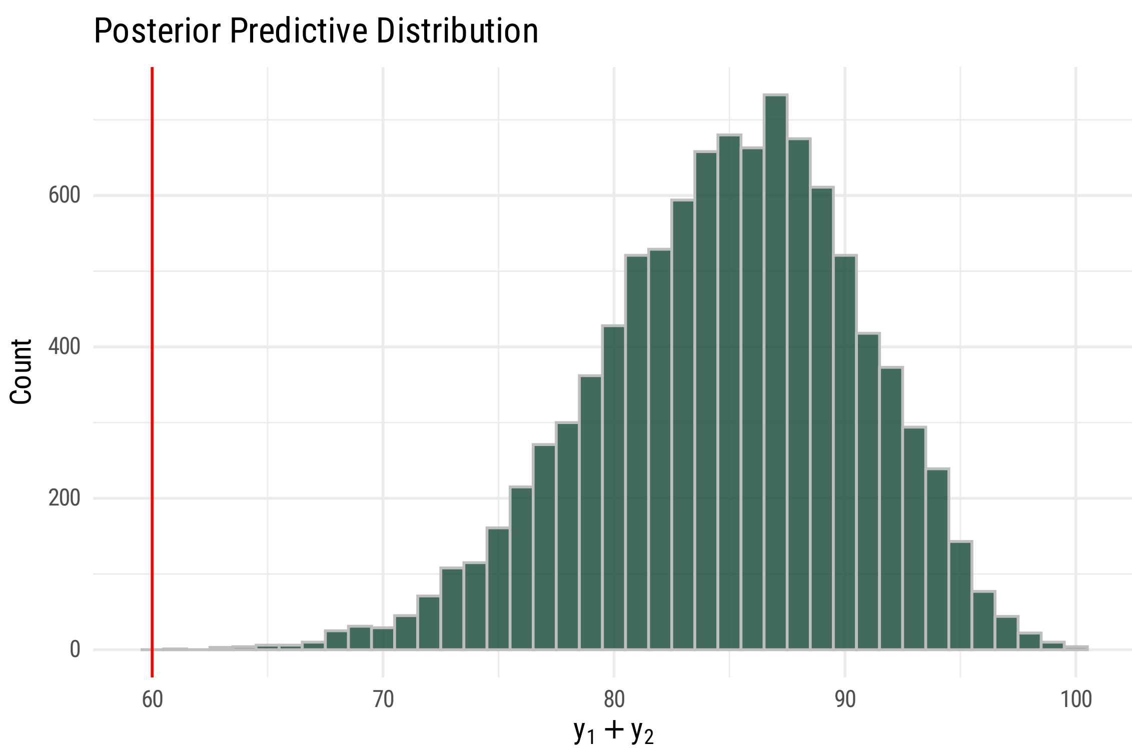
Then, we find a p-value by testing $P(Y < 26) = 0.9972$. We used `pbinom(26, size = 32, prob = 0.6)` to compute this. In this case, we would fail to reject H_0 in this case.

For $y = 52$, $n = 64$:

Additionally we find a p-value for this new case by testing $P(Y < 52) = 0.999$. We used `pbinom(52, size = 64, prob = 0.6)` to compute this. In this case, we would fail to reject H_0 in this case.

8

Now, we have that the 32 subjects described above are used for an interim analysis in a trial designed to enroll 100 patients. When the trial is complete, a frequentist analysis will be used to decide if Device A is sufficiently better (i.e., Device B is more irritating) to warrant further study. Using the frequentist method you employed in part 7, perform an interim analysis using the posterior predictive distribution. (Use a Beta(2, 2) prior.)



To test these hypotheses using the frequentist paradigm, we want to find a p-value using the posterior predictive distribution. We know that we have observed 26 successes, meaning for θ to be at least 0.6, we would need to observe 34 more observations out of 68. Therefore, we calculate the p-value as using the closed form for the posterior predictive distribution,

$$\Pr(s_n \geq 34 | s_0 = 26, f_0 = 6) = \sum_{s_n=34}^{68} \binom{68}{s_n} \frac{B(s_n + 26 + 2, f_n + 6 + 2)}{B(26 + 2, 6 + 2)}.$$

Which yields a p-value of 0.998. Hence, we again would fail to reject H_0 in favor of H_1. Ultimately, the p-value is so large that it would lend to the notion that the trial could be stopped and that there would be sufficient information to conclude that device B is far more irritating than device A, if each trial was really expensive.

```
## [1] 0.9983067
```