Homework 1

Casey Gibson

Problem 1

$$p(\theta) = \frac{1}{21} \text{ for } \theta = 0, 0.05, 0.1, ..., 1$$
$$Y|\theta \sim Binom(n, \theta)$$

$$P(Y = 5|\theta = .05) = \binom{n}{5}\theta^5(1-\theta)^{n-5} = \binom{100}{5}.05^5(1-.05)^{100-5} \approx .180$$

$$P(Y = 5|\theta = .5) = \binom{n}{5}\theta^5(1-\theta)^{n-5} = \binom{100}{5}.5^5(1-.5)^{100-5} \approx 5.94 \text{ x } 10^{-23}$$

$$P(\theta = .05|Y = 5) = \frac{P(Y = 5|\theta = .05)P(\theta = .05)}{P(Y = 5)}$$

If we denote $\Theta = \{0, 1, 2.., 20\}$ then $\theta \in \frac{\Theta}{20}$

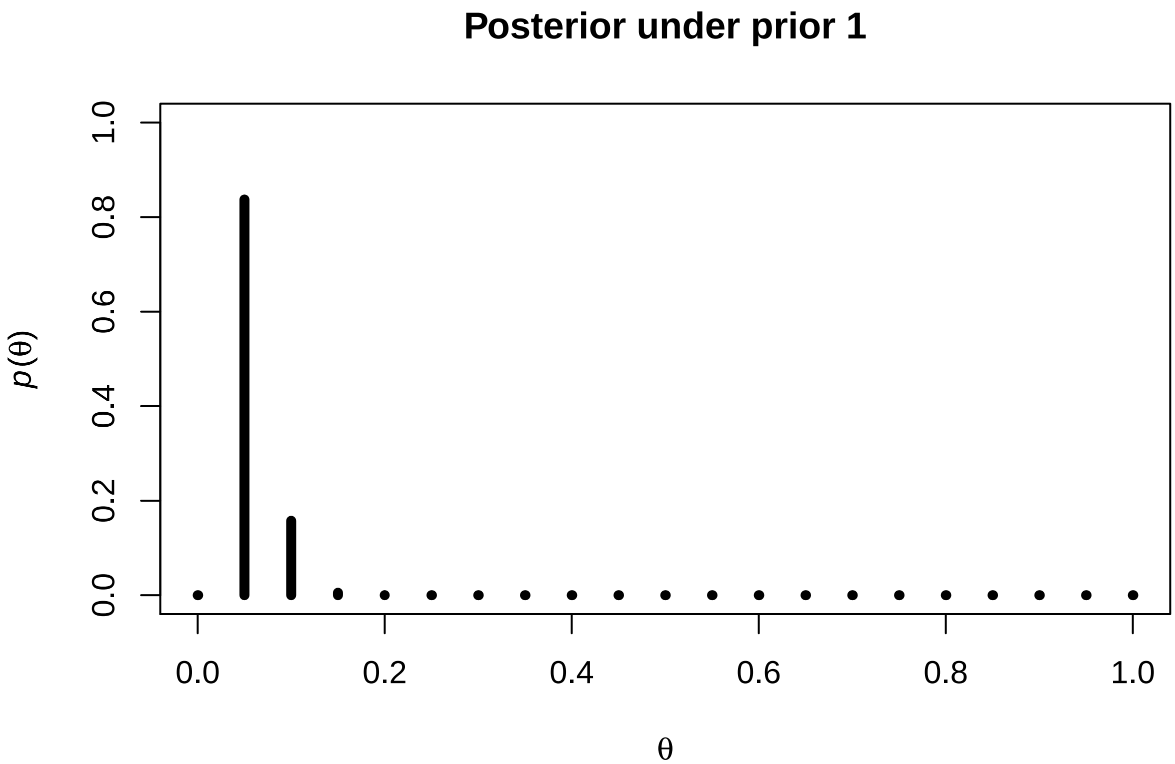
$$P(Y = 5) = \sum_{\theta' \in \Theta} P(Y, \theta') = \sum_{\theta' \in \Theta} P(Y|\theta')P(\theta') = \sum_{i=0}^{20} P(Y = 5|\theta' = \frac{i}{20})P(\theta' = \frac{i}{20})$$

$$= \sum_{i=0}^{20} \binom{100}{5}(\frac{i}{20})^5(1-(\frac{i}{20}))^{100-5}\frac{1}{21} \approx .0102$$

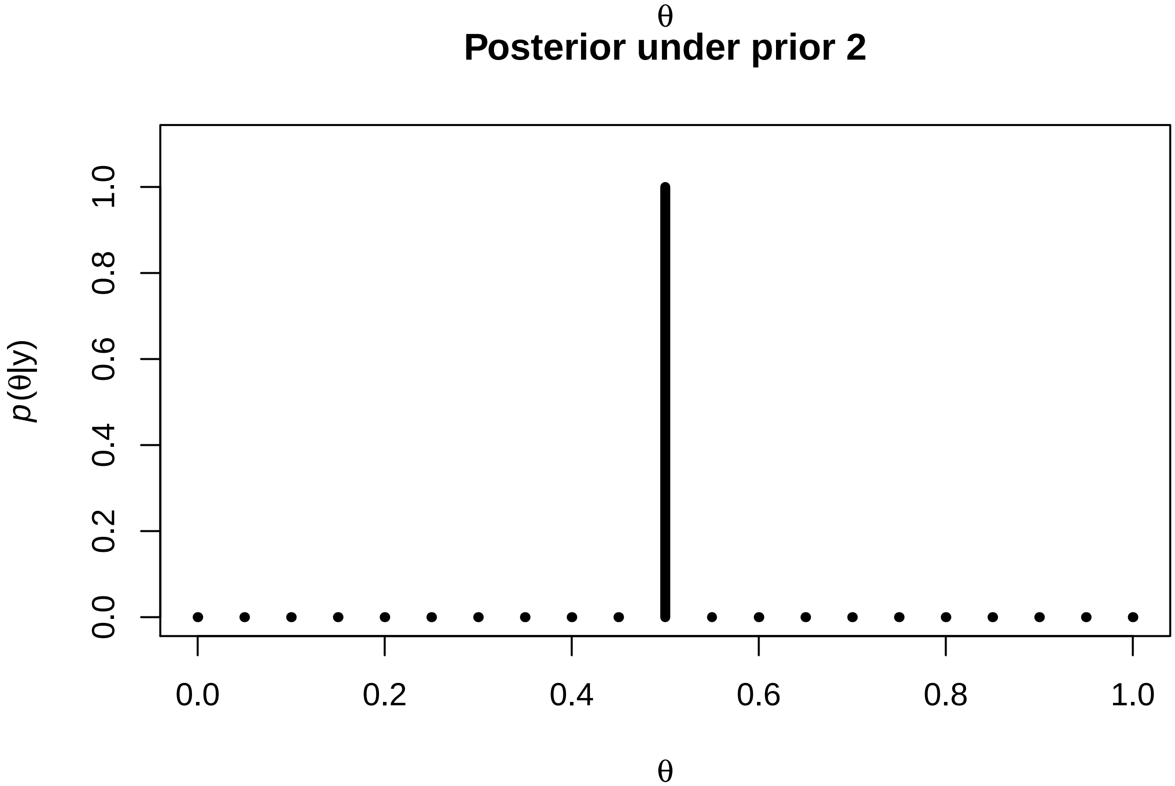
$$P(\theta = .05|Y = 5) = \frac{.1800178 * \frac{1}{21}}{.0102393} \approx .837$$

$$P(\theta = .5|Y = 5) = \frac{5.93914e^{-23}\frac{1}{21}}{.0102393} \approx 2.76 \ x \ 10^{-22}$$

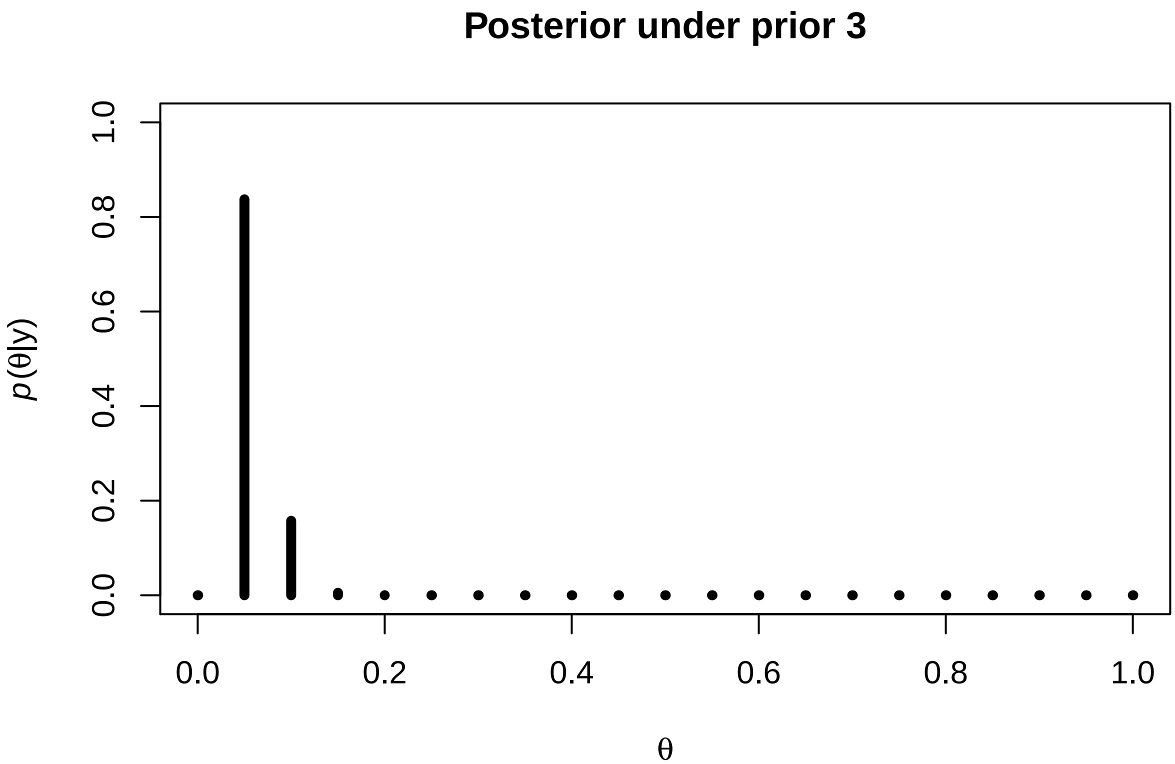
Problem 2

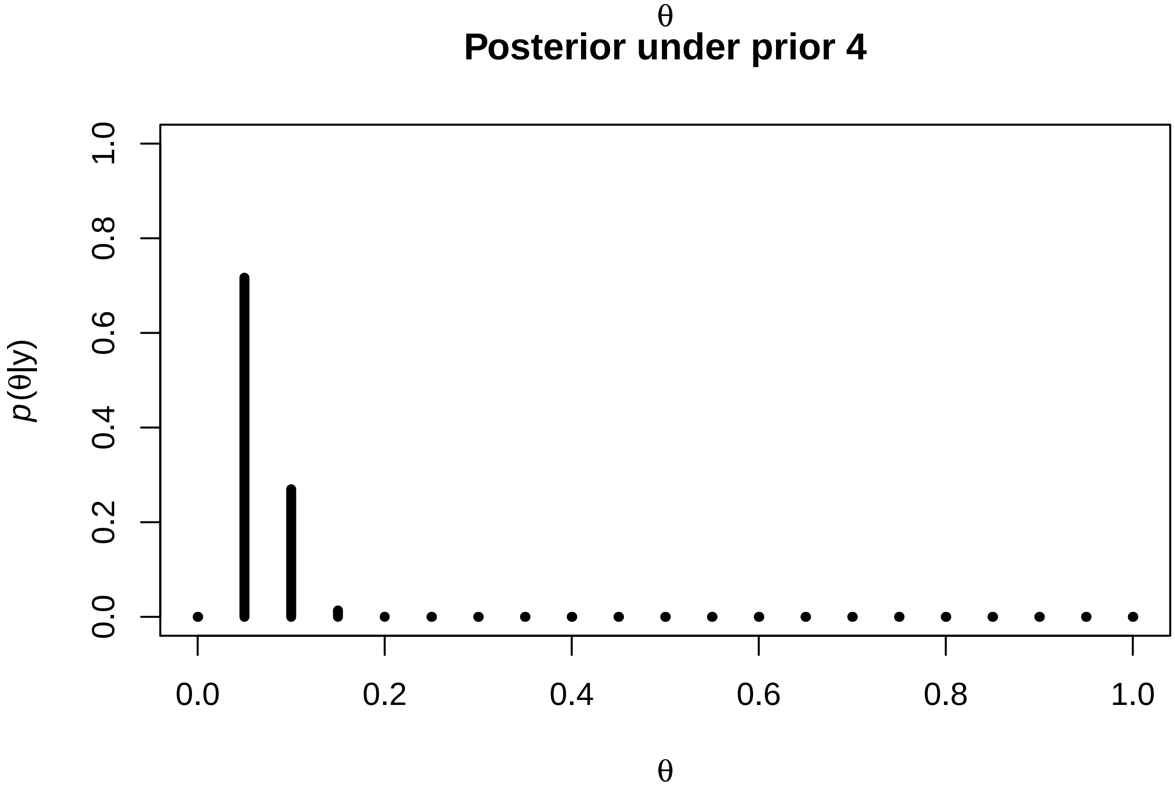


Posterior under prior 1
p(θ)
0.0
0.2
0.4
0.6
0.8
1.0
θ



Posterior under prior 2
p(θ|y)
0.0
0.2
0.4
0.6
0.8
1.0
θ



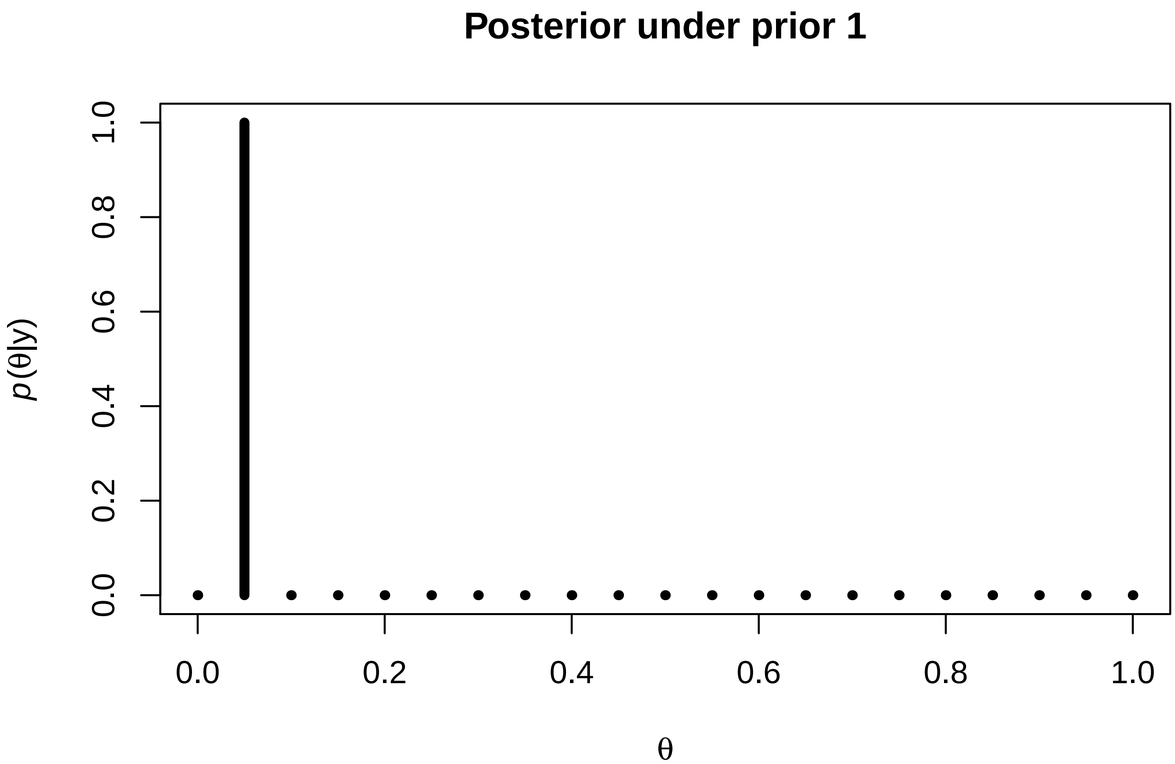


Expected Value Prior 1	Expected Value Prior 2	Expected Value Prior 3	Expected Value Prior 4
0.0584	0.5	0.0584	0.0649

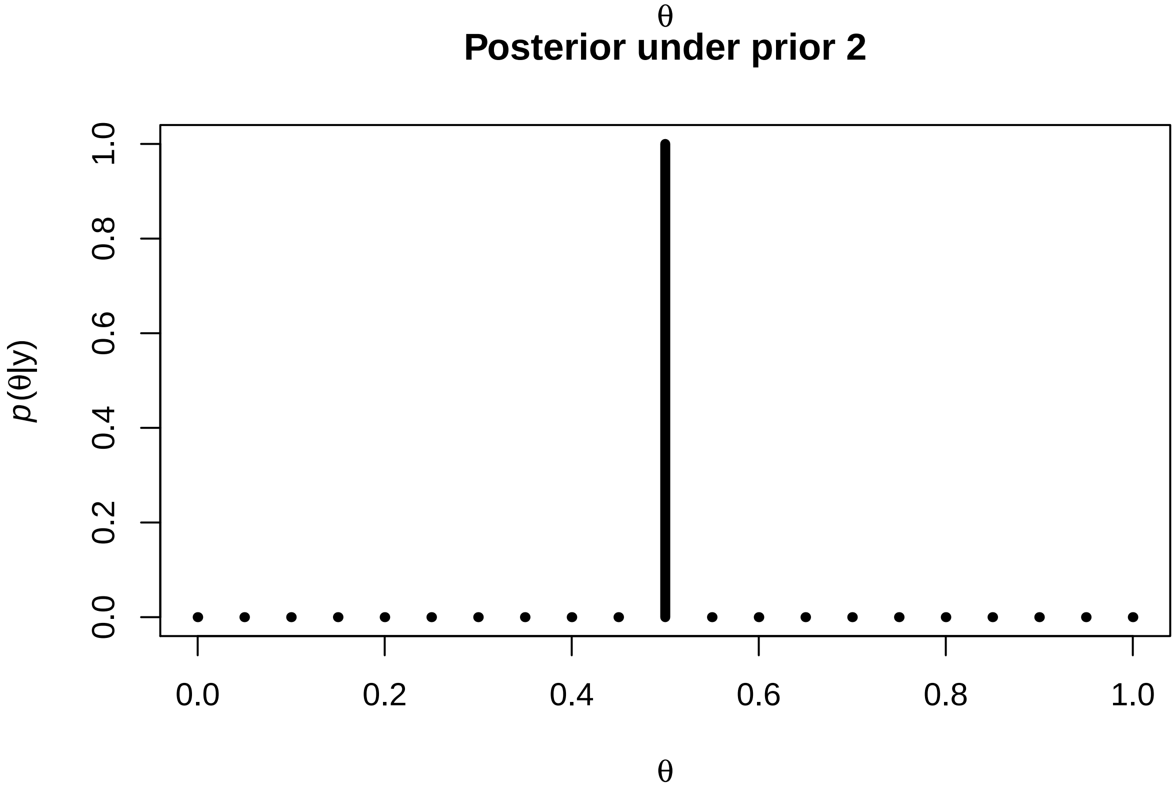
Results

Prior 2 clearly pulls the posterior towards the right because it forces all $\theta's$ less than .5 to have 0 probability. Prior 3 yeilds essentially the same results as prior 1, which makes sense because the likelihood is so skewed towards the range $0, .5$, so excluding $.5, 1$ does not change the posterior much. Prior 4 pushes the posterior slightly upwards, which again makes sense because the prior assigns higher probabilities to higher values of theta, pushing the posterior to the right.

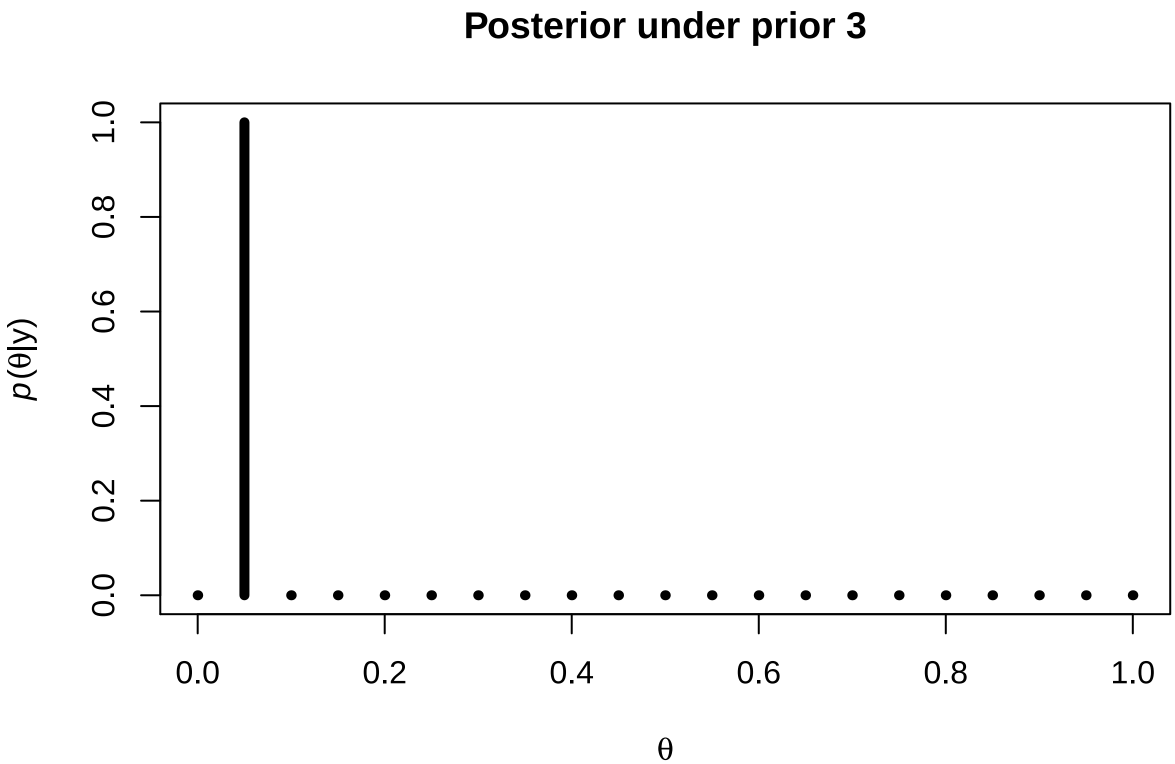
Problem 3

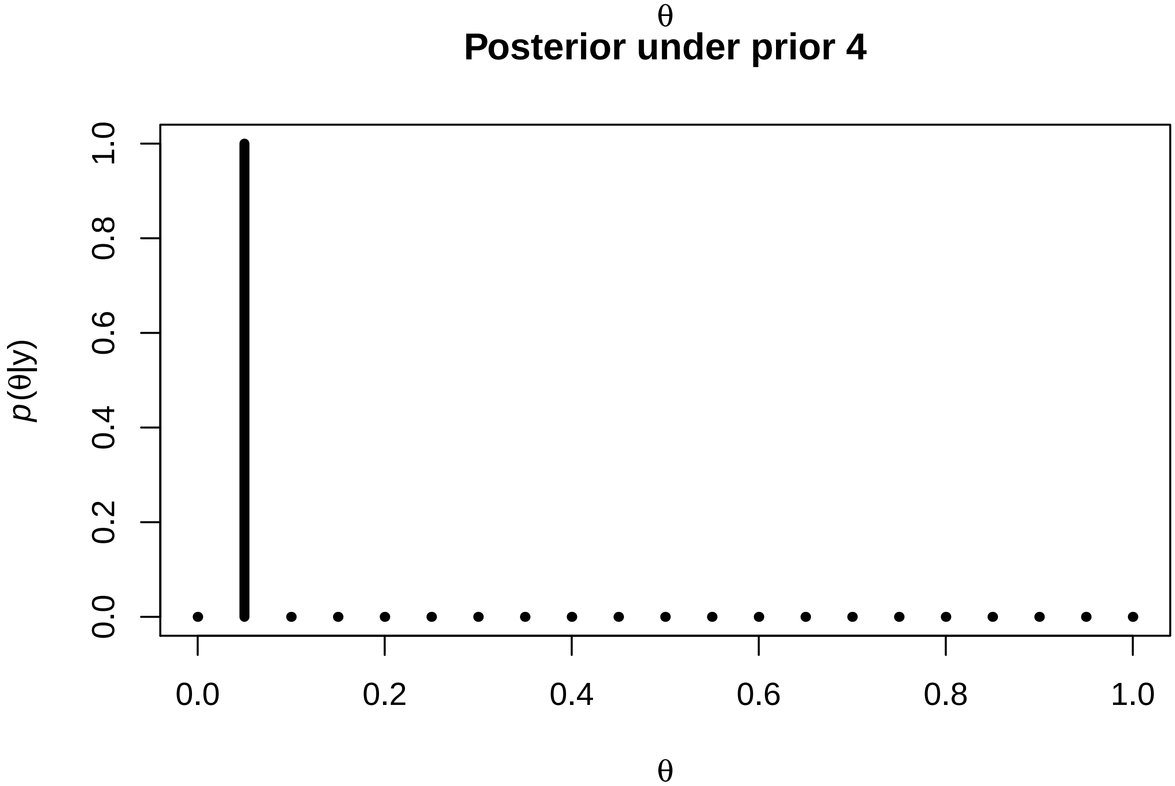


Posterior under prior 1
p(θ|y)
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
θ



Posterior under prior 2
p(θ|y)
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
θ





Expected Value Prior 1	Expected Value Prior 2	Expected Value Prior 3	Expected Value Prior 4
0.05	0.5	0.05	0.05

Results

All the priors become much more concentrated around the maximum likelihood estimate of θ.

As we get more data the effect of the prior diminishes (which is called “swamping” I think?). This makes more sense, to me at least, if we consider the log posterior.

$$p(\theta|y) \propto p(y|\theta)p(\theta)$$

taking logs we see

$$log\ p(\theta|y) \propto log\ p(y|\theta) + log\ p(\theta)$$

$$log/p(\theta|y) \propto log\ \prod_{i=1}^{n} L(\theta|y) + log\ p(\theta)$$

where $L(\theta|y)$ is the likelihood function.

This becomes

$$log\ p(\theta|y) \propto \sum_{i=1}^{n} log\ L(\theta|y) + log\ p(\theta)$$

Now we can clearly see, as n increases the effect of the $log\ p(\theta)$ dimishes since the sum term contributes more than the prior. It is harder for me to see this clearly in the non-log space because we are dealing with multiplication of small numbers.