# Problem Set 3.3

$\alpha\beta$

## Exercise 3.11

a) Specify and plot a Beta model that reflects the staff's prior ideas about π

The staff previous ideas are a mean of 1/4 and a mode of 5/22.

$$E(\pi) = \frac{\alpha}{\alpha + \beta}$$
$$Mode(\pi) = \frac{\alpha - 1}{\alpha + \beta - 2}$$

$$\begin{matrix} E(\pi) = \dfrac{\alpha}{\alpha + \beta} = \dfrac{1}{4} \\ Mode(\pi) = \dfrac{\alpha - 1}{\alpha + \beta - 2} = \dfrac{5}{22} \end{matrix} \Rightarrow \begin{matrix} \dfrac{\alpha}{\alpha + \beta} = \dfrac{1}{4} \\ \dfrac{\alpha - 1}{\alpha + \beta - 2} = \dfrac{5}{22} \end{matrix} \Rightarrow \begin{matrix} 3\alpha = \beta \\ 17\alpha = 5\beta + 12 \end{matrix}$$

$$\begin{matrix} 3\alpha = \beta \\ 17\alpha = 5\beta + 12 \end{matrix} \Rightarrow \begin{matrix} 3\alpha = \beta \\ 17\alpha = 5(3\alpha) + 12 \end{matrix} \Rightarrow \begin{matrix} 3\alpha = \beta \\ 2\alpha = 12 \end{matrix} \Rightarrow \begin{matrix} 3\alpha = \beta \\ \alpha = 6 \end{matrix} \Rightarrow \begin{matrix} 3 * 6 = \beta \\ \alpha = 6 \end{matrix} \Rightarrow$$

$$\begin{matrix} \beta = 18 \\ \alpha = 6 \end{matrix}$$

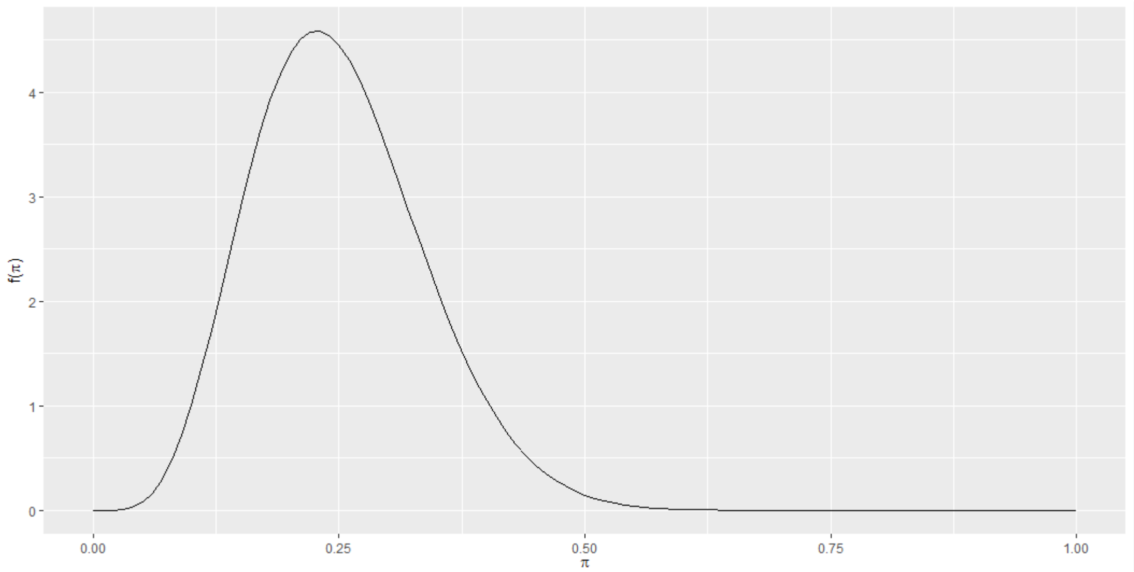
We now have the alpha and beta values for the prior model:

$$\pi \sim Beta(6, 18)$$

We can plot this distribution:



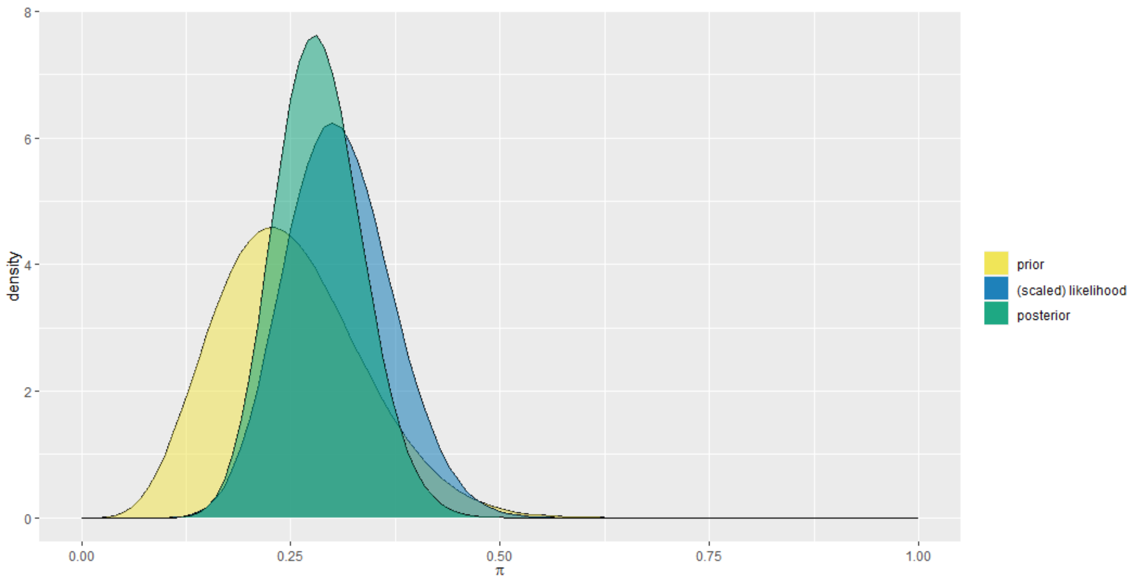
Alejandro C Parra Garcia




b) Among 50 surveyed students, 15 are regular bike riders. What is the posterior model for π?

n=50 | y=15

```
plot_beta_binomial(alpha = 6, beta = 18, y = 15, n = 50)
```




```
> summarize_beta_binomial(alpha = 6, beta = 18, y = 15, n = 50)
      model alpha beta      mean      mode         var         sd
1     prior     6   18 0.2500000 0.2272727 0.007500000 0.08660254
2 posterior    21   53 0.2837838 0.2777778 0.002710007 0.05205773
> 
```

The posterior model is:

$$\pi \sim Beta(21, 53)$$

c) What is the mean, mode, and standard deviation of the posterior model?

We can obtain this numbers form the function summarize_beta_binomial(), as we can see in the previous image. The mean is 0.2837, the mode is 0.2777, the standard deviation is 0.052

But we can calculate these values:

$$E(\pi) = \frac{\alpha}{\alpha + \beta} = \frac{21}{21 + 53} = 0.2837$$

$$Mode(\pi) = \frac{\alpha - 1}{\alpha + \beta - 2} = \frac{21 - 1}{21 + 53 - 2} = 0.2777$$

$$Var(\pi) = \frac{\alpha * \beta}{(\alpha + \beta)^2 * (\alpha + \beta + 1)} = \frac{21 * 53}{(21 + 53)^2 * (21 + 53 + 1)} = 0.00271$$

$$Std(\pi) = \sqrt{Var(\pi)} = \sqrt{0.00271} = 0.052$$

d) Does the posterior model more closely reflect the prior information or the data? Explain your reasoning.

In the prior model the expected value of people using bikes was 1/4=0.25, and from the data capture 15/50 people used the bike, meaning 0.3, this makes the posterior model change the expected value from the original 0.25 to 0.2837, or from 25% to 28.37%

$$E(\pi) = 0.25 \; vs \; E(\pi|y = 15) = 0.2837$$

In the case of the standard deviation, we observed a similar result, the variability in the results gets narrower, from 0.086 in the prior to 0.052 in the posterior.

$$Std(\pi) = 0.0866 \; vs \; Std(\pi|y = 15) = 0.052$$

The new posterior model represents a middle ground between the original prior model and the data capture during the survey.

Since the values of alpha and beta for the prior model are smaller than the ones added by the survey data, that means that the posterior model is going to be more influence by the data than from the prior model, but still the prior is relevant in the posterior.