Lab 3

Roxanne Li

2024-01-27

```
library(ggplot2)
```

Q3

The new prior Beta(10, 10) will be more peaked around the center of the interval between 0 and 1, compared to the Beta (1, 1) prior which is basically an uniform distribution. This means that with the Beta(10, 10) prior we believe θ is very likely to be around 0.5.

We are assuming more information with the Beta(10, 10) prior because having a peak in the distribution is more informative than a flat prior like Beta(1, 1). A flat prior means it is equally likely for the true θ to lie anywhere between 0 and 1, whereas a peaked prior suggests that we have more confidence about where the true value of θ lies.

Q4

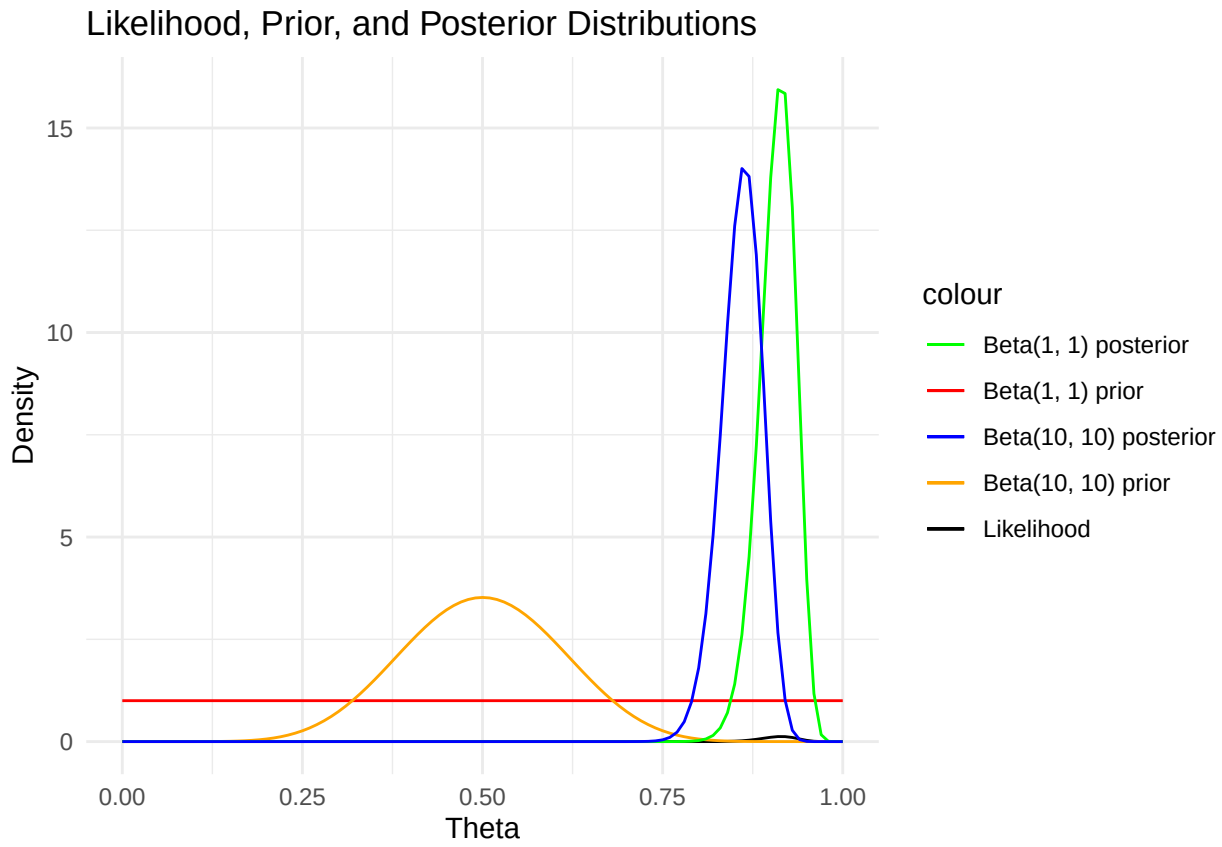
Setting the parameters:

```
n <- 129
y <- 118
theta <- seq(0, 1, by = 0.01)

beta11_prior_a <- 1
beta11_prior_b <- 1
beta11_posterior_a <- y + beta11_prior_a
beta11_posterior_b <- n - y + beta11_prior_b

beta1010_prior_a <- 10
beta1010_prior_b <- 10
beta1010_posterior_a <- y + beta1010_prior_a
beta1010_posterior_b <- n - y + beta1010_prior_b
```

Plotting the likelihood, prior, and posterior distributions:

```
likelihood <- dbinom(y, size = n, prob = theta)
df <- data.frame(theta = theta, likelihood = likelihood)
#prior <- dbeta(theta, shape1 = beta11_prior_a, shape2 = beta11_prior_b)
#posterior <- dbeta(theta, shape1 = beta11_posterior_a, shape2 = beta11_posterior_a)

ggplot(df, aes(x = theta)) +
  geom_line(aes(y = likelihood, color = "Likelihood")) +
```

```r
  stat_function(fun = dbeta, args = list(shape1 = beta11_prior_a, shape2 = beta11_prior_b),
                aes(color = "Beta(1, 1) prior")) +
  stat_function(fun = dbeta, args = list(shape1 = beta1010_prior_a, shape2 = beta1010_prior_b),
                aes(color = "Beta(10, 10) prior")) +
  stat_function(fun = dbeta, args = list(shape1 = beta11_posterior_a, shape2 = beta11_posterior_b),
                aes(color = "Beta(1, 1) posterior")) +
  stat_function(fun = dbeta, args = list(shape1 = beta1010_posterior_a, shape2 = beta1010_posterior_b),
                aes(color = "Beta(10, 10) posterior")) +
  #geom_line(aes(y = prior, color = "Prior")) +
  #geom_line(aes(y = posterior, color = "Posterior")) +
  labs(title = "Likelihood, Prior, and Posterior Distributions",
       x = "Theta", y = "Density") +
  scale_color_manual(values = c("Likelihood" = "black",
                                "Beta(1, 1) prior" = "red", "Beta(10, 10) prior" = "orange",
                                "Beta(1, 1) posterior" = "green", "Beta(10, 10) posterior" = "blue")) +
  theme_minimal()
```



The Beta(1, 1) prior is flat whereas the Beta(10, 10) prior has a peak around 0.5. The peak of the posterior distribution resulted from the Beta(10, 10) prior is closer to the peak of its prior. Beta(1, 1) prior is flat and doesn't provide much information regarding its belief of where the true θ lies, so the resulting posterior has a peak around the observed $\theta = \frac{y}{n} = \frac{118}{129} \approx 0.91$.

Q6

A noninformative prior does not provide any belief of the average improvement in success probability before seeing the final measurements.

A subjective/informative prior provides numeric information or beliefs about the average improvement in success probability before seeing the final measurements.