

# Advanced Statistical Modelling - Exercise 1.4

Filip Chmielowski

2024-10-02

EuroVegas (Binomial)

Two students are modeling their prior belief about the proportion of residents of Baix Llobregat that support building a macro casino in Baix Llobregat. Anna lives in Gavà, and her prior mean is 0.2 and her prior standard deviation is 0.08. Sam is an Erasmus student from Scotland, so he is not aware of the local feeling for or against the proposed casino. He decides to use a uniform prior. The two students take a random sample of  $n = 100$  Baix Llobregat residents and ask for their views of the casino. Out of the random sample  $y = 26$  said they support building a casino in Baix Llobregat. a) Draw these two prior distributions in the same graph. b) Draw the two posterior distributions in the same graph. c) Repeat a) and b) assuming that the data were:  $n = 1000$  and  $y = 260$ , and think about the difference in using different sample sizes.

a) Draw these two prior distributions in the same graph.

```
# mean = alpha / (alpha + beta)
# variance = (alpha * beta) / [(alpha + beta)^2 * (alpha + beta + 1)]
# alpha = ((1 - mean) / var - 1 / mean) * mean^2
# beta = alpha * (1 / mean - 1)

# Anna: prior is a Beta distribution with a mean of 0.2 and standard deviation of 0.08
anna_mean = 0.2
anna_sd = 0.08
anna_var = anna_sd^2
anna_alpha = ((1 - anna_mean) / anna_var - 1 / anna_mean) * anna_mean^2
anna_beta = anna_alpha * (1 / anna_mean - 1)

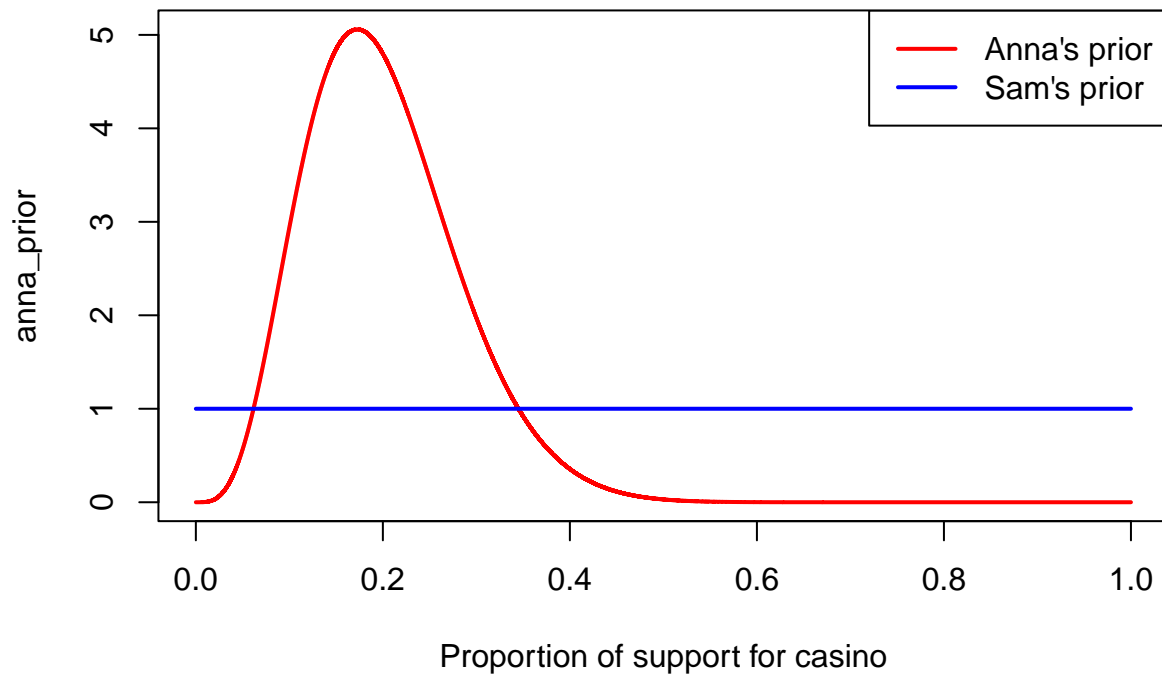
# Sam: prior is uniform, so that is Beta(1, 1)
sam_alpha = 1
sam_beta = 1

# Generate regular sequence of data
x = seq(0, 1, length.out = 100000)

# Distributions calculation
anna_prior = dbeta(x, anna_alpha, anna_beta)
sam_prior = dbeta(x, sam_alpha, sam_beta)

# Draw the prior distributions' plots
plot(x, anna_prior, type = "l", lwd = 2, col = "red",
     main = "Prior distributions for Anna and Sam", xlab = "Proportion of support for casino")
lines(x, sam_prior, lwd = 2, col = "blue")
legend("topright", legend = c("Anna's prior", "Sam's prior"), lwd = 2, col = c("red", "blue"))
```

## Prior distributions for Anna and Sam



b) Draw the two posterior distributions in the same graph.

```
# Observations
n = 100
y = 26

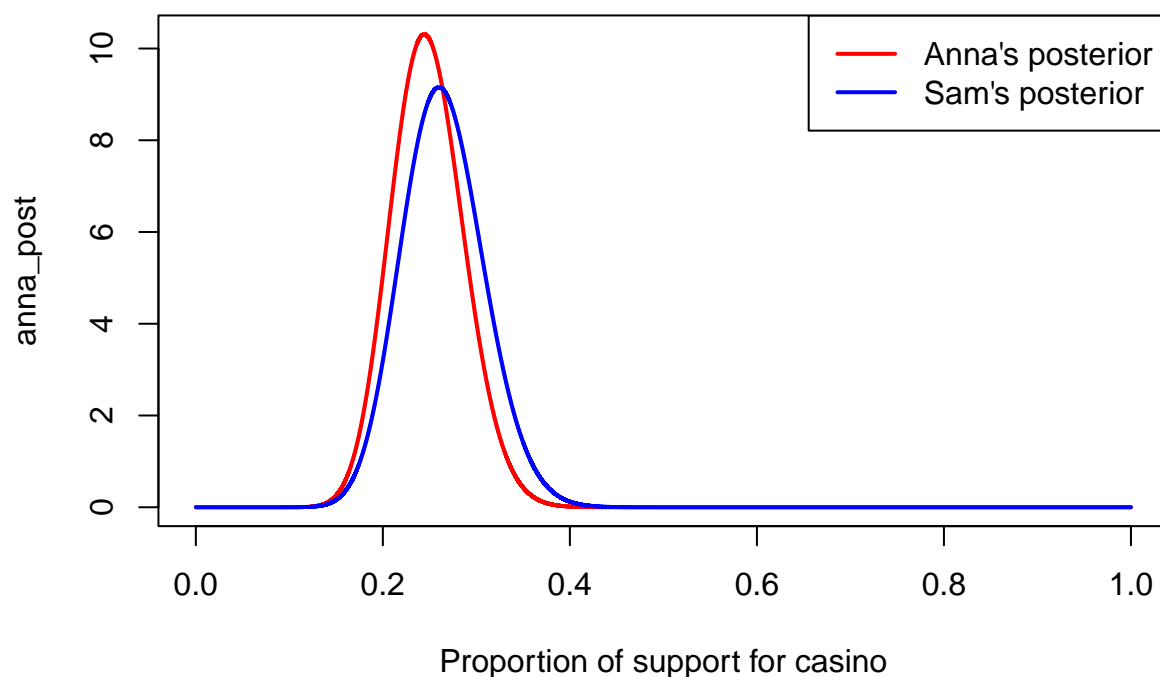
# Anna's prior update
anna_alpha_post = anna_alpha + y
anna_beta_post = anna_beta + n - y

# Sam's prior update
sam_alpha_post = sam_alpha + y
sam_beta_post = sam_beta + n - y

# Distributions' calculation
anna_post = dbeta(x, anna_alpha_post, anna_beta_post)
sam_post = dbeta(x, sam_alpha_post, sam_beta_post)

# Draw the posterior distributions' plots
plot(x, anna_post, type = "l", lwd = 2, col = "red",
     main = "Posterior distributions for Anna and Sam (n = 100, y = 26)", xlab = "Proportion of support",
     lines(x, sam_post, lwd = 2, col = "blue")
legend("topright", legend = c("Anna's posterior", "Sam's posterior"), lwd = 2, col = c("red", "blue"))
```

### Posterior distributions for Anna and Sam ( $n = 100$ , $y = 26$ )



- c) Repeat a) and b) assuming that the data were:  $n = 1000$  and  $y = 260$ , and think about the difference in using different sample sizes.

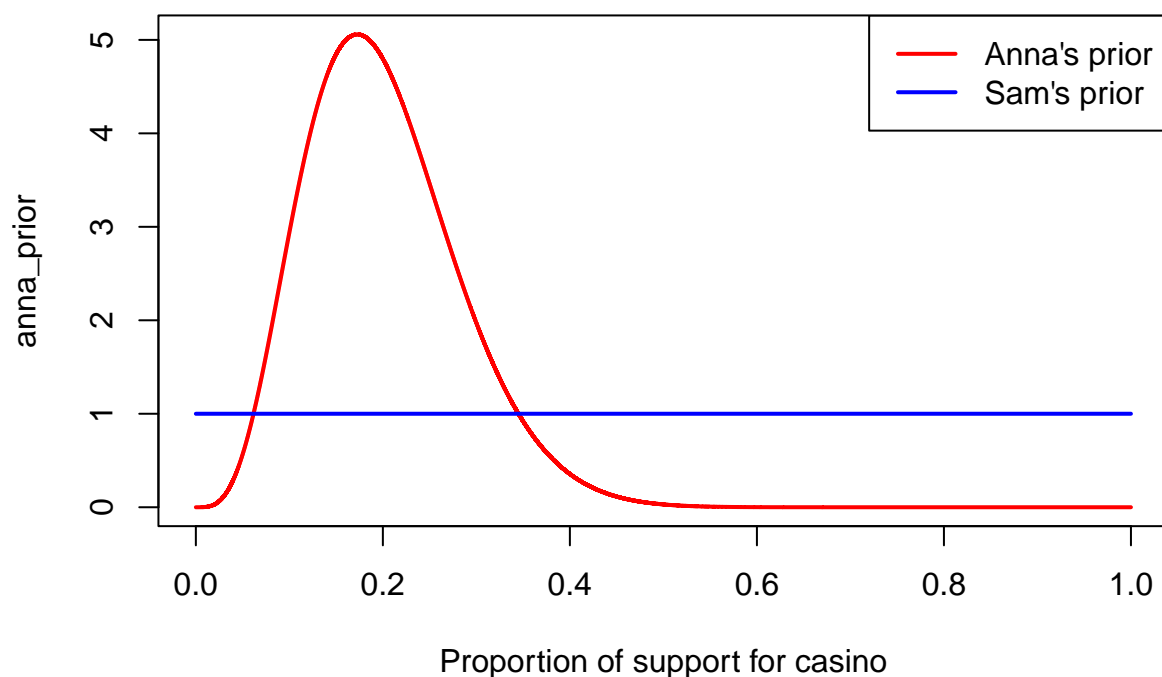
```
# New observations
n_new = 1000
y_new = 260

# Anna's prior update with the new data
anna_alpha_post_new = anna_alpha + y_new
anna_beta_post_new = anna_beta + n_new - y_new

# Sam's prior update with the new data
sam_alpha_post_new = sam_alpha + y_new
sam_beta_post_new = sam_beta + n_new - y_new

# Draw the new prior distributions' plots
plot(x, anna_prior, type = "l", lwd = 2, col = "red",
     main = "Prior distributions for Anna and Sam (n = 1000, y = 260)", xlab = "Proportion of support for casino",
     lines(x, sam_prior, lwd = 2, col = "blue")
legend("topright", legend = c("Anna's prior", "Sam's prior"), lwd = 2, col = c("red", "blue"))
```

### Prior distributions for Anna and Sam ( $n = 1000$ , $y = 260$ )



```
# Distributions' calculation
anna_post_new = dbeta(x, anna_alpha_post_new, anna_beta_post_new)
sam_post_new = dbeta(x, sam_alpha_post_new, sam_beta_post_new)

# Draw the new posterior distributions' plots
plot(x, anna_post_new, type = "l", lwd = 2, col = "red",
      main = "Posterior distributions for Anna and Sam (n = 1000, y = 260)", xlab = "Proportion of support for casino",
      lines(x, sam_post_new, lwd = 2, col = "blue")
legend("topright", legend = c("Anna's posterior", "Sam's posterior"), lwd = 2, col = c("red", "blue"))
```

**Posterior distributions for Anna and Sam ( $n = 1000$ ,  $y = 260$ )**

