Homework 2

Due on October 26, 2025 at 11:59 pm

Your name here

Submit your Rmarkdown (.qmd) and the compiled pdf on Gauchospace.

1. Trend in Same-sex Marriage

A 2017 Pew Research survey found that 10.2% of LGBT adults in the U.S. were married to a same-sex spouse. Now it's the 2020s, and Bayard guesses that π, the percent of LGBT adults in the U.S. who are married to a same-sex spouse, has most likely increased to about 15% but could reasonably range from 10% to 25%.

1a. Identify a Beta model that reflects Bayard's prior ideas about π by specifying the parameters of the Beta, α and β.

```r
alpha <- NULL # YOUR CODE HERE
beta <- NULL # YOUR CODE HERE
```

```r
. = ottr::check("tests/q1a.R")
```

1b. Bayard wants to update his prior, so he randomly selects 90 US LGBT adults and 30 of them are married to a same-sex partner. What is the posterior model for π?

```r
posterior_alpha <- NULL # YOUR CODE HERE
posterior_beta <- NULL # YOUR CODE HERE
```

1c. Use R to compute the posterior mean and standard deviation of π.

```r
posterior_mean <- NULL # YOUR CODE HERE
posterior_sd <- NULL # YOUR CODE HERE

print(sprintf("The posterior mean is %f", posterior_mean))
print(sprintf("The posterior sd is %f", posterior_sd))
```

1d. Does the posterior model more closely reflect the prior information or the data? Explain your reasoning. Hint: in the recorded lecture we showed a special way in which we can write the posterior mean in a Beta-Binomial model. How can this help? Check the lectures notes.

```
# YOUR CODE HERE
```

Type your answer here, replacing this text.

2. Cancer Research in Laboratory Mice

A laboratory is estimating the rate of tumorigenesis (the formation of tumors) in two strains of mice, A and B. They have tumor count data for 10 mice in strain A and 13 mice in strain B. Type A mice have been well studied, and information from other laboratories suggests that type A mice have tumor counts that are approximately Poisson-distributed. Tumor count rates for type B mice are unknown, but type B mice are related to type A mice. Assuming a Poisson sampling distribution for each group with rates θ_A and θ_B. Based on previous research you settle on the following prior distribution:

$$\theta_A \sim \text{gamma}(120, 10),\ \ \theta_B \sim \text{gamma}(12, 1)$$

2a. Before seeing any data, which group do you expect to have a higher average incidence of cancer? Which group are you more certain about a priori? You answers should be based on the priors specified above.

Type your answer here, replacing this text.

2b. After you the complete of the experiment, you observe the following tumor counts for the two populations:

$$y_A = (12, 9, 12, 14, 13, 13, 15, 8, 15, 6)$$
$$y_B = (11, 11, 10, 9, 9, 8, 7, 10, 6, 8, 8, 9, 7)$$

Compute the posterior parameters, posterior means, posterior variances and 95% quantile-based credible intervals for θ_A and θ_B. Same them in the appropriate variables in the code cell below. You do not need to show your work, but you cannot get partial credit unless you do show work.

```
. = ottr::check("tests/q2b.R")
```

2c. Compute and plot the posterior expectation of θ_B given y_B under the prior distribution gamma$(12 \times n_0, n_0)$ for each value of $n_0 \in \{1, 2, ..., 50\}$. As a reminder, n_0 can be thought of as the number of prior observations (or pseudo-counts).