ICMU Bayes Intro, Homework 1
Due on September 26th, 2025

1. Exit poll results from a local election, where 2 candidates were competing, found that 56% of the respondents voted in favor of Candidate A. They also estimated that of those who did vote in favor for Candidate A, 36% had a college degree, while 48% of those who voted for Candidate B had a college degree. Suppose we randomly sampled a person who participated in the exit poll and found that they had a college degree. What is the probability that they voted in favor of Candidate A? Show all your work.

2. Let's revisit our example of estimating the proportion of people who had COVID-19 in Los Angeles from a random sample. As in class, we use Beta(2,20) prior for the θ — unknown proportion and have $y = 3$ individuals who tested positive for SARS-CoV-2 antibodies out of $n = 100$ recruited study participants. We are interested in predicting how many positive individuals we will have if we conduct the study one more time: y^{new}. Using the posterior distribution for θ, $p(\theta \mid y)$,

   (a) Show that the posterior predictive distribution $p(y^{\text{new}} \mid y)$ is a probability mass function of a Beta-Binomial distribution (`https://en.wikipedia.org/wiki/Beta-binomial_distribution`). Compute parameters of this distribution.
   Hint: When working with $p(y^{\text{new}} \mid y)$, work out the joint density $p(y^{\text{new}}, \theta \mid y)$ first. You can use the fact that $p(y^{\text{new}} \mid \theta, y) = p(y^{\text{new}} \mid \theta)$ — this says that if I know my model parameters, I don't need to know observed data to make predictions, I can just use my model (binomial in this case).

   (b) Plot the probability mass function of the Beta-Binomial distribution above in R. Install `extraDistr` package in R and use the `dbbinom()` function from this package when making this plot.

   (c) Computing posterior predictive mean $\text{E}(y^{\text{new}} \mid y)$ and $\text{Var}(y^{\text{new}} \mid y)$.

3. (adapted from chapter 5 material of the "Bayes Rules!" book) If a likelihood-prior pair results in the posterior distribution that belongs to the same family of distributions as the prior, this pair of distributions is called **conjugate**. We have seen that Beta prior and Binomial likelihood form one such conjugate pair. Another such pair is Gamma prior and Poisson likelihood.

   Let $\lambda > 0$ be an unknown rate parameter of a Poisson distribution and $Y_1, \ldots, Y_n$ be iid samples from this Poisson distribution. We assume the following prior distribution for λ:

   $$\lambda \sim \text{Gamma}(s, r); \text{ in this parameterization } \text{E}(\lambda) = \frac{s}{r}.$$

   (a) Prove that $p(\lambda \mid y_1, \ldots, y_n)$ is density of a Gamma distribution and find parameters of this Gamma distribution.

   (b) Let λ be the rate of text/sms messages people receive per hour. Suppose our prior belief is that people receive around 5 messages per hour, with standard deviation of 0.25 messages. Find a Gamma prior that approximately encodes this prior belief and plot its density.

   (c) You collect data from 10 friends and find that during the most recent hour they received 0, 1, 1, 1, 3, 3, 2, 6, 5, and 2 messages respectively. Find the posterior distribution of λ conditional on this dataset and plot the posterior density of λ together with its prior in the same plot.