Individual HW05

Your Name Here

```r
library(bayesrules) # R package for our textbook
library(tidyverse) # Collection of packages for tidying and plotting data
library(janitor) # Helper functions like tidy and tabyl
library(rstan) # for MCMC
library(bayesplot) # for plotting
```

1 BR Exercise 8.7

For each situation, find the appropriate credible interval using the "middle" approach

(a) A 99% credible interval of λ where $\lambda|y \sim$ Gamma(1,5)
(b) A 95% credible interval of μ where $\mu|y \sim N(10, 2^2)$
(c) An 80% credible interval of μ where $\mu|y \sim N(-3, 1^2)$

2 BR Exercise 8.8

There's more than one approach to constructing a 95% credible interval. The "middle 95%" approach reports the range of the middle 95% of the posterior density, from the 2.5th to the 97.5th percentile. The "highest posterior density" approach reports the 95% of posterior values with the highest posterior densities.

(a) Let $\lambda|y \sim$Gamma(1,5). Construct the 95% highest posterior density credible interval for λ. Represent this interval on a sketch of the posterior pdf. *Hint:* The sketch itself will help you identify the appropriate CI. Do not try to find a solution that will generalize to any posterior density, just focus on this specific density.
(b) Repeat part a using the middle 95% approach.
(c) Compare the two intervals from parts a and b. Are they the same? If not, how do they differ and which is more appropriate here?
(d) Let $\mu|y \sim N(-13, 2^2)$. Construct the 95% highest posterior density credible interval for μ.
(e) Repeat part d using the middle 95% approach.

(f) Compare the two intervals from parts d and e. Are they the same? If not, why not?

3 BR Exercise 8.21

The loon is a species of bird common to the Ontario region of Canada. Let λ denote the typical number of loons observed by a birdwatcher across a 100-hour observation period. To learn about λ, we'll utilize bird counts $(Y_1, ..., Y_n)$ collected in n different outings.

(a) Explain which Bayesian model is appropriate for this analysis: Beta-Binomial, Gamma-Poisson, or Normal-Normal.
(b) Your prior understanding is that the typical rate of loon sightings is 2 per 100 hours with a standard deviation of 1 per 100-hours. Specify an appropriate prior model for λ and explain your reasoning.
(c) The `loons` data in the {bayesrules} package contains loon counts in different 100-hour observation periods. How many data points do we have and what's the average loon count per 100 hours?
(d) In light of your prior and data, calculate and interpret a (middle) 95% posterior credible interval for λ. NOTE: You'll first need to specify your posterior model of λ

4 BR Exercise 8.22

Let's continue our analysis of λ, the typical rate of loon sightings in a 100-hour observation period. You hypothesize that birdwatchers should anticipate a rate of less than 1 loon per observation period.

(a) State this as a formal hypothesis test (using H_0, H_a, and λ notation)

(b) What decision might you make about these hypotheses utilizing the credible interval from the previous exercise?
(c) Calculate and interpret the posterior probability that your hypothesis is true.
(d) Putting this together, explain your conclusion about λ

5 8.24: Posterior predictive distribution of Loon sightings (modified slightly)

(a) Simulate the posterior model of λ, the typical rate of loon sightings per observation period, with {rstan} using 4 chains and 10000 iterations per chain. Perform some MCMC diagnostics to confirm that your simulation has stabilized.

(b) Use your MCMC simulation to approximate the posterior predictive model of Y', the number of loons that a birdwatcher will spy in their *next* observation period. Construct a visualization of this model.

(c) Summarize your observations/conclusions of the posterior predictive model of Y'

(d) Approximate the probability that the birdwatcher observes 0 loons in their next observation period.

6 8.13: Posterior predictive Normal-Normal with calculus

Let $Y = y$ be an observed data point from a $N(\mu, \sigma^2)$ model. Further, suppose σ is known and that μ is unknown with a $N(\theta, \tau^2)$ prior.

(a) Identify the posterior pdf of μ given the observed data, $f(\mu|y)$

(b) Let $Y' = y'$ be the value of a *new* data point. Identify the conditional PDF of Y' given μ, $f(y'|\mu)$

(c) Identify the posterior predictive pdf of Y', $f(y'|y)$

(d) Suppose $y = -10$, $\sigma = 3$, $\theta = 0$ and $\tau = 1$. Specify and sketch the posterior predictive pdf of Y'