STAT 7630: Homework 6
(Due: Tuesday, 11/05/2024)

Note: Show all your work for the necessary steps to receive full credit.

Please turn in the HW on paper, hand-written and/or typed. For computational problems, return only the relevant parts of the output with comments/annotations. Questions taken from the textbook are marked with BR for "Bayes Rules!". From the code you are using to answer the problems, turn in the relevant output, and the figures (if requested), preferably printed from the output. No need to turn in your code or long lists of generated samples.

Please disclose any us of AI in your solutions. Regardless of you use it or not, make sure you submit your own work, not copy from other source(s). Any suspicion of AI use will result in automatic 0 or substantial point loss in any question.

BR Chapter 8:
Do Exercises 8.8 and 8.9.

Hint: In 8.8, you do not have to provide the sketches of the intervals on the posterior pdf. And you can use the `hpd` function from the `TeachingDemos` package to get the HPD intervals.

Additional Questions (AQs):
AQ1. In class (for the Prussian cavalry example) we derived the posterior predictive distribution in the case where $Y_1, \ldots, Y_n \sim$ i.i.d. Poisson(λ), with the Gamma prior $\lambda \sim$ Gamma$(2, 4)$. Instead of deriving the posterior predictive distribution analytically, we could have sampled from it using Monte Carlo methods.

(a) Randomly sample $\lambda^{[1]}, \ldots, \lambda^{[J]}$ from a Gamma distribution with shape parameter $\sum y_i + 2$ and rate parameter $n + 4$. Using these samples, generate $Y^{[1]}, \ldots, Y^{[J]}$ from Poisson distributions with mean $\lambda^{[j]}$, for $j = 1, \ldots, J$.

(b) Plot the approximate posterior predictive distribution, similar to our in-class example. How does it appear to compare to the observed Prussian army data distribution?

AQ2. In a NASA experiment, 14 male rats were sent into space. Upon their return, the red blood cell mass (in ml) of each rat was measured. Additionally, 14 other male rats were kept on earth for the same period, and their red blood cell mass was also measured. Assume that the red blood cell masses for both groups can be modeled by a normal distribution with equal variances across the two groups. The data are as follows:

Space rats: 8.59, 8.64, 7.43, 7.21, 6.87, 7.89, 9.79, 6.85, 7.00, 8.80, 9.30, 8.03, 6.39, 7.24.

Earth rats: 8.65, 6.99, 8.40, 9.66, 7.62, 7.44, 8.55, 8.70, 7.33, 8.58, 9.88, 9.94, 7.14, 9.04.

(a) Suppose the research question of interest is to test whether the mean red blood cell mass differs between the two groups. Conduct a Bayesian hypothesis test to address this question. Clearly specify your prior settings, using $\mu_\Delta = 0$ and $\sigma^2_\Delta = \frac{1}{5}$. Calculate and provide the posterior probability for each hypothesis being true.

(b) Now, suppose the research question of interest is to test whether the mean red blood cell mass for the space group is lower than that of the control group. Conduct a Bayesian hypothesis test to address this question. Clearly specify your prior settings. (The researcher believed *a priori* that the rat population's average red blood cell mass might be around 7 ml, but was uncertain about the effect of space travel.) Calculate and provide the posterior probability for each hypothesis being true.

AQ3. A physician is interested in determining whether the mean systolic blood pressure of a certain group of patients is less than 130. She takes a random sample of 17 patients and measures their systolic blood pressure. Assume the measurements follow a $N(\mu, \sigma^2)$ distribution, with μ unknown and $\sigma^2 = 225$ known. (You can use the normal-normal results from Chapter 5 to obtain the posterior distribution for μ.) The physician states *a priori* that she is 95% certain that the true mean systolic blood pressure is between 120 and 140. The data are:

$$118, 140, 90, 150, 128, 112, 134, 140, 112, 126, 112, 148, 124, 130, 142, 105, 125.$$

(a) Conduct a Bayesian hypothesis test of $H_0 : \mu \geq 130$ vs. $H_a : \mu < 130$, basing your conclusions on the posterior distribution for μ. Clearly state your prior specifications.

(b) Conduct a classical t-test using $\alpha = 0.05$. Are the substantive conclusions any different from those in part (a)?