Required Problems

1. Suppose you are consulting for a manufacturer of tea bags. Each bag is supposed to contain 5.5 grams of tea, and the filling process is known to have a standard deviation of 0.106 grams per bag (assume a normal distribution). Of interest is whether there is a drift (change) in the mean amount of tea per bag.

(a) Determine a prior distribution that has a mean matching the process specification and contains the information of about one hundred observations.

(b) A sample of 40 bags finds an average of 5.494 grams per bag.
   i. What is the posterior probability that the process is within 0.001 grams of its specification? (Be sure to keep enough decimal places in your computations.)
   ii. What is the posterior predictive probability that a new tea bag will contain at least the labeled amount of 5.5 grams of tea?

2. A sample of 29 cereal boxes is randomly selected and the amount of potassium (in mg per serving) is recorded, and can be summarized with the statistics $\bar{x} = 92.1$ and $s = 52.8$. Suppose you model the potassium content of each cereal as iid normal with unknown mean $\mu$ and unknown variance $\sigma^2$, and suppose you use the standard noninformative prior, $f(\mu, \sigma^2) \propto \frac{1}{\sigma^2}$ (this is like choosing $\alpha = \beta = 0$ in the parameterization from class).

(a) Compute the frequentist 95% confidence interval for $\mu$: $\bar{y} \pm q(.975, n - 1) \frac{s}{\sqrt{n}}$.

(b) Find the joint posterior distribution of $\mu$ and $\sigma^2$ by finding the posterior marginal for $\sigma^2|y$ and the posterior conditional distribution for $\mu|\sigma^2, y$. You can use the results from class, rather than re-deriving any of the equations.

(c) Find a 95% credible interval for the mean $\mu$, and compare your result to that of part (a). See Section 3.8 for the derivation of the $t$ distribution as the marginal posterior distribution for $\mu|y$ (and just use the result in the book or from class, don’t re-derive it). You can get the numeric results from R by transforming between the location-scale $t$ and the standard $t$. (Hint: doing the comparing part early on may save you some numerical effort.)

3. Elicit your personal prior on the probability that Schwarzenegger is elected to the U.S. senate in 2010. Use the Beta family to find an appropriate distribution to approximate your beliefs. [You may want to do this by trial and error by making plots in R.] Note that he may not decide to run for the senate, and even if he does, he is likely to face strong Democratic opposition, so you will need to have a definite opinion to have a point estimate that is 0.5 or higher (i.e., if you put equal probability on whether he runs, and equal probability on him winning, then your point estimate is 0.25; you can easily justify a point estimate of 0.5, but please do not justify it as “ignorance”), and your full distribution is unlikely to be uniform. You will have to decide exactly what it is. For this question, state the specific Beta distribution you choose, and plot its density with R. [Only include the plot in your homework, not any of the other R commands for this problem.]
4. Suppose we are going to observe data that are from a uniform distribution with unknown upper endpoint, i.e.,

$$Y_i \overset{iid}{\sim} \text{Unif}[0, \theta], \quad i \in \{1, \ldots, n\}$$

The conjugate prior for this uniform distribution is the Pareto distribution:

$$\theta \sim \text{Pareto}(\alpha, \beta) \Rightarrow f(\theta|\alpha, \beta) = \alpha \beta^\alpha \theta^{-(\alpha+1)} I(\theta \geq \beta)$$

(i.e., the density if zero if \( \theta \) is smaller than \( \beta \)). Show that the Pareto is indeed conjugate for this uniform by finding the posterior distribution (and stating it as a Pareto, giving the updated parameters).

Optional Problem

5. Assuming \( \alpha > 2 \), the Pareto distribution has mean \( \frac{\alpha \beta}{\alpha - 1} \) and variance \( \frac{\alpha \beta^2}{(\alpha - 1)^2(\alpha - 2)} \). Continuing with the uniform/Pareto problem:

(a) Compare the posterior mean to the prior mean and the mean of the likelihood (a Pareto\((n - 1, m)\) distribution, where \( n \) is the sample size and \( m \) is the maximum of the observations). In particular:

i. Show that if \( \beta < m \) (which we would generally want to be true) and if \( \alpha > 2 \) (which rules out the weirder parts of the family), then the posterior mean is always less than the mean of the likelihood.

ii. Show by example that the posterior mean can be smaller than the prior mean (i.e., find reasonable values of \( \alpha, \beta, n, \) and \( m \) such that this is true).

iii. Is the posterior mean a weighted average of the prior mean and the likelihood in this case?

(b) Write out the posterior variance. When considered as a function of the sample size, \( n \), what is unusual about this expression compared to most standard cases?