STAT 7630: Homework 4
(Due: Thursday, 10/08/2024)

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.

Q1. Assume the data $Y_1, Y_2, \ldots, Y_n$ are independent and identically distributed (iid) $N(\mu, \sigma^2)$ with mean μ known and variance σ^2 unknown.

(a) Show that the gamma prior for $1/\sigma^2$ is a conjugate prior (use Gamma(α, β) distribution with shape and rate parameters).

(b) Derive the posterior distribution of σ^2 by transforming $1/\sigma^2$ to σ^2.

(c) Compare the posterior you obtain in part (b) with the posterior obtained (in lecture slides) for σ^2 when an Inverse Gamma(α, β) prior is used.

Q2.: Let θ represent the probability of success in a Bernoulli distribution, and let $\tau = \frac{\theta}{1-\theta}$ be the odds ratio.

(a) Find the Jeffreys prior for the log-odds ratio, $\log(\tau)$.

(b) Verify that this prior is invariant under reparameterization.

Q3. Suppose $X_1, X_2, \ldots, X_n$ are iid $N(\mu, \sigma^2)$ with both μ and σ^2 unknown. Simulate data using `rnorm` in `R` with $n = 50$, $\mu = 5$, and $\sigma = 2$ (use `setseed(123)` to have the same data for all). We will perform Bayesian analysis of this model using three different prior settings using this simulated data as your observed data set.

(a) **Normal-Inverse Gamma Priors**: Assume the conjugate priors where $\mu|\sigma \sim N(\delta, \sigma^2/s_0)$ and $\sigma^2 \sim$ Inverse-Gamma(α, β).

   (i) Calculate the posterior distribution of μ and σ^2 under the conjugate prior. Hint: Assume the prior hyperparameters are $\delta = 5$, $s_0 = 1$, $\alpha = 3$, $\beta = 2$.

   (ii) Report the posterior means and 95% credible intervals for both μ and σ^2.

(b) **Vague Proper Priors**: Assume an almost vague but proper prior where $\mu|\sigma \sim N(0, \sigma^2/.001)$ and $\sigma^2 \sim$ Inverse-Gamma$(0.001, 0.001)$.

(i) Calculate the posterior distribution of μ and σ^2 under the vague priors.

(ii) Report the posterior means and 95% credible intervals for both μ and σ^2.

(c) **Vague Priors**: Assume vague priors $p(\mu) \propto 1$ and $p(\sigma) \propto \frac{1}{\sigma}$.

(i) Calculate the posterior distribution of μ and σ^2 under the vague priors.

(ii) Report the posterior means and 95% credible intervals for both μ and σ^2.

(d) **Comparison**: Compare the results from (a), (b), and (c) in terms of the posterior means, variances, and credible intervals for μ and σ^2. Comment on how the choice of priors influences the posterior results. Visualize the posterior distributions of μ under all three priors using plots.

Q4. A researcher aims to estimate the mean number of accidents per month occurring within 100 feet of the Gervais Street/Assembly Street intersection in Columbia. The researcher models the number of accidents Y per month using a Poisson(λ) distribution, where the density function for Y given λ is

$$p(y|\lambda) = \frac{\lambda^y e^{-\lambda}}{y!}, \quad y = 0, 1, 2, \ldots, \quad \lambda \geq 0.$$

(a) The researcher adopts a gamma distribution with shape parameter α and rate parameter β). Derive the general form of the posterior distribution for λ given a random sample $y_1, \ldots, y_n$ from n months of data.

(b) The following accident counts were collected over 15 randomly selected months:

$$1, 0, 4, 1, 4, 2, 5, 3, 0, 3, 1, 2, 2, 4, 1$$

For each of

(i) $\alpha = 1$ and $\beta = .5$

(ii) $\alpha = 5$ and $\beta = .5$,

calculate the posterior mean and a 95% credible interval for λ. Compute both a quantile-based (i.e. equal-tail) interval and a highest posterior density (HPD) interval, using the standard exponential prior and the provided data.

Q5. In Q4, for $\alpha = 5$ plot the posterior mean for values of $\beta \in [1, 50]$ (use a grid of 500 beta values), and also add to the plot the end points of the 95% HPD credible intervals for λ.