STAT 4640/7640 **Homework 4**

Due: February 24, 2022

1. Problem 12 of Chapter 2: Say $Y|\mu \sim N(\mu, 1)$ and μ has an improper prior $\pi(\mu) = 1$ for all μ. Prove that the posterior distribution of μ is a proper probability distribution function.

2. Use the `cars` dataset that is built in to R for this problem. If you type "cars" into your R terminal, you will see the data for $n = 50$ observations. The first column is the speed (mph) at which the car is traveling and the second column is the distance (meters) it took the car to stop.

   (a) Plot speed versus distance. Comment on the relationship between the two variables.

   (b) Using the `lm` function, fit a linear regression model with distance as the response variable and speed as the explanatory variable. What is the fitted regression line? Interpret your results in context. As a refresher:

   ```
   fit=lm(dist~speed,data=cars)
   summary(fit)
   ```

   (c) Now, let's assume σ^2 is known by setting it equal to the estimate, $\hat{\sigma}^2$, fom part (b). Therefore, $\Sigma = \hat{\sigma}^2\mathbf{I}_n$. Using your class notes (and section 2.1.6), conduct Bayesian linear regression using the prior $\boldsymbol{\beta} \sim N(\mathbf{0}, 100^2\mathbf{I}_2)$ where $\boldsymbol{\beta} = (\beta_0, \beta_1)'$. Compute the posterior mean and variance of $\boldsymbol{\beta}$. How do these estimates compare to those from part (b)?

   (d) Use Monte Carlo sampling to test the hypothesis $\beta_1 < 3$. Justify you answers and interpret your results in context of the problem. Note: the R package `mvtnorm` has the function `rmvnorm` which allows you to sampling from a multivariate normal distribution.

   (e) Now, let's assume $\boldsymbol{\beta}$ is known by setting each coefficient equal to its estimate in part (b). Assume the conjugate prior $\sigma^2 \sim \mathit{Inverse\ Gamma}(2, 2)$ and obtain the posterior distribution of $\sigma^2|\mathbf{y}$. Using Monte Carlo sampling, obtain draws from the posterior distribution and compute the posterior mean, variance, and 95% credible interval. Note that the Inverse Gamma distribution is in the R package `MCMCpack` and the following code draws 100 independent samples from an $\mathit{Inverse\ Gamma}(2, 2)$.

   ```
   library(MCMCpack)
   rinvgamma(100,2,2)
   ```

   (f) Choose two additional Inverse Gamma priors (e.g., select different values for a and b). Using Monte Carlo sampling, obtain draws from the posterior for each prior and plot the densities together in the same figure along with the posterior density from the $\mathit{Inverse\ Gamma}(2, 2)$ prior. Compare the results and comment on the sensitivity of the prior specification. Make sure to report what priors you chose.