MSc in Biostatistics
Bayesian Inference Project

For the data set `RegressionData` consider the simple linear regression model

$$y_i = \alpha + \beta x_i + \epsilon_i, \quad \epsilon_i \sim \text{Normal}(0, \tau^{-1}), \quad i = 1, \ldots, 100.$$

1. Set $\alpha = 2$ and consider the hypothesis test

$$\begin{aligned} H_0 &: \beta = 0 \\ H_1 &\quad \beta \neq 0 \end{aligned}$$

Let $\pi_1, 1 - \pi_1$ be the prior probabilities of hypotheses H_0 and H_1, repectively, and assume the following prior distributions for the model parameters.

$$\begin{aligned} \beta|\tau &\sim \text{Normal}(\mu_2, (c_2\tau)^{-1}) \\ \tau &\sim \text{Gamma}(p, q). \end{aligned}$$

(i) Consider equal prior probabilities for the two hypotheses and set $\mu_2 = 0$, $c_2 = 0.5$ and $p = q = 0.01$. Compare the two hypotheses and comment on the results. Perform this analysis using R.
(ii) Now, consider $\pi_1 = 0.2$, $\mu_2 = 1$, $c_2 = 2$ and $p = q = 1$. Compare the two hypotheses and comment on the results. Perform a prior sensitivity analysis and discuss your findings.

2. Consider Bayesian Inference for the simple linear regression model with the following prior structure.

$$\begin{aligned} \alpha|\tau &\sim \text{Normal}(\mu_1, (c_1\tau)^{-1}) \\ \beta|\tau &\sim \text{Normal}(\mu_2, (c_2\tau)^{-1}) \\ \tau &\sim \text{Gamma}(p, q). \end{aligned}$$

Calculate the joint posterior distribution of the model parameters, and the full conditional posterior distributions of α, β and τ.
(i) Construct an MCMC algorithm for Bayesian inference based on the full conditional posterior distributions. Program your algorithm in R and use your code to analyze the provided data set. Summarize your posterior inferences and comment on the results.
(ii) Perform the analysis in (i) using STAN and compare your results with those obtained in (i).