1. Let \( y = (y'_1, y'_2, \ldots, y'_K)' \) be an \( N \times 1 \) vector of outcomes that has been partitioned into \( K \) blocks (or subvectors) such that \( y_k \) is \( n_k \times 1 \) and \( \sum_{k=1}^{K} n_k = N \). Suppose we assume that each block is independent and consider the following two models:

Model 1: \( IG(\sigma^2 \mid a, b) \times N(\beta_0 \mid \mu_0, \sigma^2 V_0) \times \prod_{k=1}^{K} N(\beta_k \mid \beta_0, \sigma^2 V_\beta) \times \prod_{k=1}^{K} N(y_k \mid X_k \beta_k, \sigma^2 V_k) \)

Model 2: \( IG(\sigma^2 \mid a, b) \times N(\beta_0 \mid \mu_0, \sigma^2 V_0) \times \prod_{k=1}^{K} N(y_k \mid X_k \beta_k, \sigma^2 V_k) \)

where \( a, b \) are constant (known) scalars, \( \mu_0 \) and \( V_0 \) are known and fixed, each \( X_k \) is the \( n_k \times p \) matrix of predictors associated with outcome \( y_k \), each \( \beta_k \) is \( p \times 1 \), \( V_\beta \) and each \( V_k \) are known and fixed. For each of Model 1 and Model 2, answer the following questions:

(a) Find the marginal posterior distribution \( p(\sigma^2 \mid y) \)

(b) Find the conditional posterior distribution \( p(\beta_0 \mid y, \sigma^2) \) in Model 1 and \( p(\beta \mid y, \sigma^2) \) in Model 2.

(c) Explain clearly how you can draw exact samples from the marginal posterior distributions of \( \sigma^2 \) and \( \beta \) using only normal and gamma random number generators.

(d) For Model 1, explain clearly how we can obtain exact samples from the marginal posterior distributions \( p(\beta_k \mid y) \) for \( k = 1, 2, \ldots, K \).

(e) Examine separately each of the following cases: (i) \( V_0^{-1} = I \) in both models; (ii) \( V_\beta^{-1} = I \) in Model 1, and (iii) the prior on \( \sigma^2 \) is the improper prior \( p(\sigma^2) \propto 1/\sigma^2 \). Report on whether you can compute the above marginal posteriors.

Consider the linear regression data set available from: [http://www.biostat.umn.edu/~ph7440/pub/h7440/LinearModelExample.txt](http://www.biostat.umn.edu/~ph7440/pub/h7440/LinearModelExample.txt). Ignore the variables “latitude” and “longitude”. The total number of observations on the outcome is \( N = 1369 \). Let \( \beta \)'s be known and fixed. For each of Model 1 and Model 2, answer the following questions:

Model 3: \( IG(\sigma^2 \mid a, b) \times N(\beta \mid \mu_0, \sigma^2 V_0) \times N(y \mid X \beta, \sigma^2 V_y) \)

where \( a, b, \mu_0 \) and \( V_\beta \) are constants as before, \( X \) is the \( N \times p \) matrix of predictors associated with \( y \), \( \beta \) is the \( p \times 1 \) vector of slopes and \( V_y \) is a known \( N \times N \) matrix. Find 95% credible intervals and the posterior median and mean for \( \sigma^2 \) and each component of \( \beta \) in Model 3.

Write an R program to subset the data into \( K = 9 \) subvectors with \( n_1 = n_2 = \cdots = n_8 = 150 \) and \( n_9 = 169 \). Subject to these sample sizes, you can choose these subvectors in any way you like. Assume that \( V_y = I_N \), the \( N \times N \) identity matrix, \( V_\beta = \alpha I_p \) and \( V_0 = \delta I_p \). Consider the values of \( \alpha \) and \( \delta \) in \{10, 1000, 100000\}. For each of the 9 combinations of \((\alpha, \delta)\) present the 95% credible intervals for every unknown parameter in Models 1 and 2. Draw side-by-side box-plots of the marginal posterior distribution of the \( \beta \)'s for \( k = 1, 2, \ldots, 9 \). Compare the estimates of \( \sigma^2, \beta \) and \( \beta_0 \) among Models 1–3.
2. **Prior elicitation on the winner of the 2012 Presidential Elections.** The US President is elected on the basis of electoral college votes that are assigned to each of the 50 states. The candidate who wins the most number of votes in a state carries the state and is allotted all the electoral college votes for that state (barring a few idiosyncrasies in Maine and Nevada, where there is proportional allocation, which we will ignore for this exercise). The sum total of electoral college votes is 538. The candidate who wins 270 or more electoral votes is declared the President. It is a tie if each candidate splits 269 votes each.

About 15 days before the Presidential elections, the state of the race was such that Obama would surely win 237 electoral college votes, while Romney was sure to win 206 electoral college votes. There were 8 battleground states that were clearly “toss-ups”. These 8 battleground states and their assigned electoral college votes were: NH = 4, IA = 6, NV = 6, CO = 9, WI = 10, VA = 13, OH = 18 and FL = 29.

Assuming you believe all of the above information (i.e. there is no doubt about the safe-Obama and safe-Romney states) and you have no other information (polls, experts, etc) on the toss-ups available to you, what would be your computed probability of (i) an Obama win, (ii) a Romney win, and (iii) a tie.

*Please submit your R or any other computer programs to the TA as a part of this exercise.*