The data for this assignment are in the R workspace Gambia.RData on the class webpage. These data are taken from the R package GeoR. For more information, you can visit http://artax.karlin.mff.cuni.cz/r-help/library/geoR/html/gambia.html

The data consist of 2,035 children from 65 villages from The Gambia. For child number $i$, \texttt{pos[i]} is the binary indicator that the child tested positive for malaria (coded so that a one means the child has the disease) and \texttt{village[i]} is the index of the child’s village. Averaging over villages, malaria prevalence is 36%.

Denote $n_j$ as the number of children from village $j$ in the dataset, $Y_j \in \{0, \ldots, n_j\}$ as the number of positive cases in village $j$, and $\theta_j \in [0,1]$ as the true proportion infected in village $j$. We will consider the model

$$Y_j|\theta_j \sim \text{Binomial}(n_j, \theta_j) \quad \theta_j \sim \text{Beta}(a, b).$$

Village are analyzed separately.

1. Find the values of $a$ and $b$ so that all values of $\theta_j \in (0,1)$ are equally likely \textit{a priori}.
2. Find the values of $a$ and $b$ so that the prior mean is 0.36 and the prior standard deviation is $\sqrt{0.36(1-0.36)/2,035}$. Explain why this is a reasonable prior.
3. Extract the data for the 33 children from village 1. Plot and summarize in a table the posterior distribution $\theta_1|Y_1$ for both priors given above. Are the results sensitive to the prior? Which posterior would you use and why?
4. Compute and plot the 90% posterior interval for $\theta_1, \ldots, \theta_{65}$ for both priors. From this plot, can we conclude that the malaria rate varies by village? Are the results sensitive to the prior?
5. For each village test the hypotheses $H_0 : \theta_j < 0.36$ versus $H_1 : \theta_j > 0.36$ for both priors. Can we conclude that the malaria rate varies by village? Are the results sensitive to the prior?

You should turn in your responses to these questions in 2-3 pages. You should also turn in a separate file with carefully commented code. Only output in the 2-3 page document will be graded. Please staple both documents together.