STAT 578 – Advanced Bayesian Modeling – Spring 2020

Assignment 5

File `usparkvisits.csv` contains annual numbers of recreational visits to 30 different parks managed by the US National Park Service for the years 2006 through 2018.[1] You will build and compare two different varying-coefficient hierarchical normal regression models for the *log*-scale numbers, using JAGS and `rjags`.

(a) (i) [2 pts] On the *same* set of axes, plot segmented lines, one for each park, representing the number of visits versus the year (2006 through 2018). Distinguish the lines for different parks by using different colors or line types. (You should label the axes, but no legend is needed.)

(ii) [2 pts] Repeat the previous part using the <u>*natural logarithm*</u> of the number of visits.

Let y_{ij} be the <u>*natural logarithm*</u> of the number of visits to park j in year i ($i = 1, \ldots, 13$, $j = 1, \ldots, 30$). For each park, model the log-number as a simple linear regression on the *centered* year number:

$$y_{ij} \mid \beta^{(j)}, \sigma_y^2, X \quad \sim \quad \text{indep. } \mathrm{N}\big(\beta_1^{(j)} + \beta_2^{(j)}(x_i - \bar{x}),\, \sigma_y^2\big)$$

where

$$\beta^{(j)} \;=\; \begin{pmatrix} \beta_1^{(j)} \\ \beta_2^{(j)} \end{pmatrix} \qquad j = 1, \ldots, 30 \qquad\qquad x_i \;=\; i \qquad i = 1, \ldots, 13$$

Note that the coefficients are allowed to depend on the park, but the variance is not.

(b) Let $\hat{\beta}_1^{(j)}$ and $\hat{\beta}_2^{(j)}$ be the *ordinary least squares* estimates of $\beta_1^{(j)}$ and $\beta_2^{(j)}$, estimated for park j. You may use the `lm` function in R to compute these estimates. (For this part, the coefficient pairs are estimated completely separately for each park.)

(i) [1 pt] Produce a scatterplot of the pairs $\big(\hat{\beta}_1^{(j)}, \hat{\beta}_2^{(j)}\big)$, $j = 1, \ldots, 30$.

(ii) [1 pt] Give the average (sample mean) of $\hat{\beta}_1^{(j)}$ and also of $\hat{\beta}_2^{(j)}$.

(iii) [1 pt] Give the sample variance of $\hat{\beta}_1^{(j)}$ and also of $\hat{\beta}_2^{(j)}$.

(iv) [1 pt] Give the sample correlation between $\hat{\beta}_1^{(j)}$ and $\hat{\beta}_2^{(j)}$.

(c) Consider the bivariate prior

$$\beta^{(j)} \mid \mu_\beta, \Sigma_\beta \quad \sim \quad \text{iid } \mathrm{N}(\mu_\beta,\, \Sigma_\beta)$$

$$\mu_\beta \;=\; \begin{pmatrix} \mu_{\beta_1} \\ \mu_{\beta_2} \end{pmatrix} \qquad\qquad \Sigma_\beta \;=\; \begin{pmatrix} \sigma_{\beta_1}^2 & \rho\,\sigma_{\beta_1}\sigma_{\beta_2} \\ \rho\,\sigma_{\beta_1}\sigma_{\beta_2} & \sigma_{\beta_2}^2 \end{pmatrix}$$

with hyperpriors

$$\mu_\beta \quad \sim \quad \mathrm{N}\big(0,\, 1000^2 I\big)$$

$$\Sigma_\beta^{-1} \quad \sim \quad \mathrm{Wishart}_2\big(\Sigma_0^{-1}/2\big)$$

[1] Data from `https://irma.nps.gov/STATS`

in the notation used in the lecture videos. For your analysis, use

$$\Sigma_0 \;=\; \begin{pmatrix} 10 & 0 \\ 0 & 0.01 \end{pmatrix}$$

based on preliminary analyses. Let the prior on σ_y^2 be

$$\sigma_y^2 \;\sim\; \text{Inv-gamma}(0.0001, 0.0001)$$

(i) [2 pts] List an appropriate JAGS model. Make sure to create nodes for Σ_β, ρ, and σ_y^2.

Remember that the numbers of visits are to be analyzed on the *log* scale.

Now run your model using `rjags`. Make sure to use multiple chains with overdispersed starting points, check convergence, and monitor μ_β, Σ_β, σ_y^2, and ρ (after convergence) long enough to obtain effective sample sizes of at least 4000 for each parameter.

(ii) [2 pts] Display the `coda` summary of the results for the monitored parameters.

(iii) [2 pts] Give an approximate 95% central posterior interval for the correlation parameter ρ, and also produce a graph of its (estimated) posterior density.

(iv) [2 pts] Approximate the posterior probability that $\rho < 0$. Also, compute the Bayes factor favoring $\rho < 0$ versus $\rho > 0$. (You may use the fact that $\rho < 0$ and $\rho > 0$ have equal prior probability.) Describe the level of data evidence for $\rho < 0$.

(v) [1 pt] Your model implies that, over the 12 years from 2006 to 2018, the (population) median number of visits should have changed by a factor of

$$e^{12\mu_{\beta_2}}$$

Form an approximate 95% central posterior interval for this quantity.

(vi) [2 pts] Use the `rjags` function `dic.samples` to compute the effective number of parameters ("penalty") and Plummer's DIC ("Penalized deviance"). Use at least 100,000 iterations.

(d) Now consider a different model with "univariate" hyperpriors for the model coefficients, which do not allow for a coefficient correlation parameter:

$$\begin{aligned} \beta_1^{(j)} \mid \mu_{\beta_1}, \sigma_{\beta_1} &\sim \text{iid N}\big(\mu_{\beta_1},\, \sigma_{\beta_1}^2\big) \\ \beta_2^{(j)} \mid \mu_{\beta_2}, \sigma_{\beta_2} &\sim \text{iid N}\big(\mu_{\beta_2},\, \sigma_{\beta_2}^2\big) \end{aligned}$$

with hyperpriors

$$\begin{aligned} \mu_{\beta_1}, \mu_{\beta_2} &\sim \text{iid N}\big(0,\, 1000^2\big) \\ \sigma_{\beta_1}, \sigma_{\beta_2} &\sim \text{iid U}(0, 1000) \end{aligned}$$

Let the prior on σ_y^2 be the same as in the previous model.

(i) [4 pts] Draw a complete DAG for this new model.

(ii) [2 pts] List an appropriate JAGS model. Make sure that there are nodes for $\sigma^2_{\beta_1}$, $\sigma^2_{\beta_2}$, and σ^2_y.

Remember that the numbers of visits are to be analyzed on the *log* scale.

Now run your model using `rjags`. Make sure to use multiple chains with overdispersed starting points, check convergence, and monitor μ_{β_1}, μ_{β_2}, $\sigma^2_{\beta_1}$, $\sigma^2_{\beta_2}$, σ^2_y (after convergence) long enough to obtain effective sample sizes of at least 4000 for each parameter.

(iii) [2 pts] Display the `coda` summary of the results for the monitored parameters.

(iv) [2 pts] Recall the (population) median change factor for the number of visits,

$$e^{12\mu_{\beta_2}}$$

as considered in the previous analysis. Form an approximate 95% central posterior interval for this quantity, and compare it with the previous results.

(v) [2 pts] Use the `rjags` function `dic.samples` to compute the effective number of parameters ("penalty") and Plummer's DIC ("Penalized deviance"). Use at least 100,000 iterations.

(vi) [1 pt] Compare the (Plummer's) DIC values for this model and the previous one. Which is preferred?

Total: 32 pts