

MI 602 A - Métodos Computacionais em Estatística
Primeiro semestre de 2012
Lista de Exercícios V
Data da entrega: 27/06/2012 (todas as questões)

OBS: TODAS AS QUESTÕES DEVERÃO SER IMPLEMENTADAS NO R2WINBUGS. TODOS OS CÓDIGOS DO R2WINBUGS DEVERÃO SER APRESENTADOS.

1. Suspeita-se que o conjunto de observações contidas no arquivo snack.xls pode ser modelado apropriadamente por uma distribuição $Weibull(\alpha, \beta)$ com valores apropriados para (α, β) , relativos à força de cisalhamento de snacks (veja o livro do Prof. Gilberto Paula, Modelo de Regressão e aplicações, para mais detalhes). O interesse é estimar, via Inferência Bayesiana usando algoritmos MCMC, os parâmetros (α, β) . Assuma, à priori, que $p(\alpha, \beta) = p(\alpha)p(\beta)$, e $\alpha \sim \text{gama}(r_1, \lambda_1), \beta \sim \text{gama}(r_2, \lambda_2)$. À escolha dos hiperparâmetros fica à seu critério (não se esqueça de justificar à escolha). Sugestão: use prioris vagas. Comente sobre a convergência do algoritmo e sobre o comportamento das distribuições à posteriori. Não se esqueça de fornecer estimativas pontuais, os desvios-padrão à posteriori e intervalos de credibilidade. O modelo exponencial seria adequado para modelar os dados? Justifique, adequadamente, sua conclusão.
2. Os dados do arquivo pesca.dat correspondem a dados sobre técnicas de pesca nos municípios de Santos e Ubatuba. Maiores detalhes podem ser encontrados no livro “Modelos de regressão com apoio computacional”, do Prof. Gilberto Paula. Considere, além da variável reposta (CPUE), as covariáveis frota e ano (notem que elas são categorizadas). Proponha um modelo (que considere interação entre os fatores) para explicar a CPUE em função dessas duas variáveis categorizadas, com distribuição **gama** e função de ligação log. Assuma prioris vagas para os parâmetros, respeitando o espaço paramétrico.
 - a) Implemente o algoritmo MCMC no winbugs para o ajuste do modelo em questão usando prioris apropriadas.
 - b) Comente sobre a convergência do algoritmo e sobre o comportamento das distribuições à posteriori. Não se esqueça de fornecer estimativas pontuais, os desvios-padrão à posteriori e intervalos de credibilidade.
 - c) De acordo com os resultados do item b), o que você pode dizer sobre a interação e os efeitos principais?
3. Considere o modelo de regressão probito usual e os dados sobre a idade de ocorrência de menarca de garotas de Varsóvia (arquivo meninas.dat). Assumindo prioris do tipo $\beta_i \sim N(0, 10^6)$, independentes, implemente no WinBUGS o ajuste Bayesiano do modelo de regressão logístico com distribuição binomial. Ou seja $Y_i | p_i \sim \text{Binomial}(m_i, p_i)$, m_i : é o número de garotas entrevistadas com idade i , Y_i : é o número de meninas que já apresentaram menarca com idade i e $\text{logito}(p_i) = \beta_0 + \beta_1(x_i - \bar{x})$, $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$, $n = 25$ e x_i : é a idade média do grupo i . Comente sobre a convergência do algoritmo, o comportamento das distribuições à posteriori, e sobre a significância dos parâmetros do modelo. Não se esqueça de fornecer estimativas pontuais, os desvios-padrão à posteriori e intervalos de credibilidade. O aumento na probabilidade de menarca aumenta, substancialmente, com o aumento em uma ano na idade média? Para mais detalhes veja as páginas 235 e 236 do livro do Prof. Gilberto Paul, Modelos de Regressão com apoio computacional.

4. Considere o modelo de regressão de Poisson usual (com função de ligação log) e os dados de número de clientes de uma determinada loja (variável resposta) e a distância até a loja mais próxima em milhas (variável explicativa) de 110 áreas de uma determinada cidade. Utilize a variável explicativa centrada na média amostral dela. Considere a primeira e a última colunas do arquivo store.dat. Assumindo prioris do tipo $\beta_i \sim N(0, 10^6)$, independentes, implemente no WinBUGS o ajuste Bayesiano do modelo. Comente sobre a convergência do algoritmo, o comportamento das distribuições à posteriori, e sobre a significância dos parâmetros do modelo. Não se esqueça de fornecer estimativas pontuais, os desvios-padrão à posteriori e intervalos de credibilidade. O aumento no número de clientes aumenta, substancialmente, com o aumento em uma milha na distância até a loja mais próxima?