Final Homework

Due: Noon Wednesday, December 18, 2019

1. Modify the ESS function to also estimate a 95% HPD interval. Include your function in the printed version of the homework.

2. Consider the following model:

$$\begin{aligned} y_i|\kappa_i &\overset{\text{ind}}{\sim} \text{exponential}(\kappa_i) \\ \kappa_i &= \prod_j \theta_j^{x_{ij}} \\ \theta_j &\overset{\text{ind}}{\sim} \text{gamma}(\alpha_j, \lambda_j) \end{aligned}$$

   where x_{ij} are known covariates, and α_j and λ_j are known hyperparameters. In some cases the y_i are censored at time t_i so the data are the pairs (t_i, w_i) where $w_i = 1$ if t_i is an uncensored time and $w_i = 0$ if t_i is a censored time yielding

$$f_{\mathbf{t},\mathbf{w}|\boldsymbol{\kappa}}(\mathbf{t}, \mathbf{w}|\boldsymbol{\kappa}) = \prod_i \kappa_i^{w_i} e^{-t_i \kappa_i}.$$

   Note: Consistent with the book we are using the parameterization where κ_i and λ_j are rate parameters as opposed to scale parameters.

   Note: For this problem you may assume that $x_{ij} \in \{0, 1\}$. Which implies that

$$\kappa_i = \prod_{J_i} \theta_j$$

   where $J_i = \{j : x_{ij} = 1\}$. Other useful sets are $I_j = \{i : x_{ij} = 1\}$ and $K_{ij} = \{k : x_{ik} = 1 \cap k \neq j\}$.

   (a) Derive the score function and Hessian matrix necessary to compute the MLE estimates of $\boldsymbol{\theta}$ using the Newton-Raphson algorithm. Note: This will not involve the prior distributions, gamma(α_j, λ_j).

   (b) Write a function to compute MLE estimates of $\boldsymbol{\theta}$ along with their approximate standard errors given $\mathbf{t}$, $\mathbf{w}$, and $\mathbf{X}$. Include your function in the printed version of the homework.

   (c) Derive the conditional distributions necessary to implement the Gibbs sampler for $\boldsymbol{\theta}$.

   (d) Write a function that to implement your Gibbs sampler given $\mathbf{t}$, $\mathbf{w}$, $\mathbf{X}$, $\boldsymbol{\alpha}$, and $\boldsymbol{\lambda}$. Include your function in the printed version of the homework.

   (e) Using the data and model described in problem 7.5.

      i. Run your MLE function to obtain maximum likelihood estimates and approximate standard errors of $\boldsymbol{\theta}$.

      ii. Run and evaluate the performance of your Gibbs sampler using a single chain.

      iii. Using the `doParallel` and `foreach` packages run multiple chains of your Gibbs sampler and evaluate the performance of your Gibbs sampler.

      iv. Compute summary statistics of the estimated joint posterior distribution of the $\boldsymbol{\theta}$ along with the mean remission times for control and treated patients, including marginal means, standard deviations, and 95% probability intervals.

3. Using the data and change point model for problem 7.6. For the prior assume $\lambda_i \sim \text{Gamma}(\gamma_1, \alpha)$ for $i = 1, 2$ and $\alpha \sim \text{Gamma}(\gamma_2, \gamma_3)$ where γ_1, γ_2, and γ_3 are known hyperparameters.

   (a) Derive the conditional distributions necessary to implement a change point model Gibbs sampler.

   (b) Write a function that to implement a change point model Gibbs sampler given $\mathbf{X}$ and the three gamma distribution hyperparameters. Include your function in the printed version of the homework.

   (c) Run your Gibbs sample using the coal mine data and compute summary statistics of the estimated joint posterior distribution of the θ, λ_1, and λ_2, including marginal means, standard deviations, and 95% probability intervals.