Homework III (2022)

1. Consider the following hierarchical changepoint model for the number of occurrences Y_i of some event during time interval i:

$$Y_i \sim \begin{cases} Poisson(\theta), & i = 1, \ldots, k; \\ Poisson(\lambda), & i = k+1, \ldots, n, \end{cases}$$

   $\theta \sim Gamma(a_1, 1/b_1)$, $\lambda \sim Gamma(a_2, 1/b_2)$, θ, and λ independent; $b_1 \sim IGamma(c_1, 1/d_1)$, $b_2 \sim IGamma(c_2, 1/d_2)$, b_1 and b_2 independent, where *Gamma* denotes the gamma and *IGamma* the inverse gamma distributions.

   Apply this model to the data (coal_data.txt), which gives counts of coal mining disasters in Great Britain by year from 1851 to 1962. (Here, “disaster” is defined as an accident resulting in the deaths of 10 or more miners.) Set $a_1 = a_2 = .5$, $c_1 = c_2 = 1$, and $d_1 = d_2 = 1$ (a collection of “moderately informative” values). Also assume $k = 40$ (corresponding to the year 1890), and write an R program to obtain marginal posterior means and variances for θ, λ, and $R = \theta/\lambda$ with output from the Gibbs sampler.

2. The Gamma distribution has two parameters: the shape α, which does not have a conjugate form, and the rate β, which has a Gamma conjugate form. MCMC samplers for the Gamma must take this into account. Set the prior distributions of α and β to both be $Gamma(1, 0.01)$ (that is, $Exp(0.01)$, with mean 100.)

   (a) Generate 100 variates $(Y_1, \ldots, Y_{100})$ from a Gamma distribution with shape parameter 5 and rate parameter 5. Report the mean and variance of the sample and contrast these with the known properties of Gamma variates with these parameters. Save these values for use in the remaining problems.

   (b) Construct a grid approximation of the distribution $(\alpha, \beta|Y)$ given the data and the prior distribution. Generate 10,000 variates from this approximation. What are the posterior probabilities that the samples for each parameter are less than the true values used to generate the distribution - that is, $p(\alpha < 5|Y)$ and $p(\beta < 5|Y)$?

   (c) Use the grid approximation you have constructed in the previous part to estimate the normalizing constant for $p(\alpha, \beta|Y)$; that is, given the joint distribution $p(Y|\alpha, \beta)p(\alpha)p(\beta)$ and conditioning on the data you have generated, estimate $p(y) = \int\int p(Y|\alpha, \beta)p(\alpha)p(\beta)d\alpha d\beta$ by calculating the joint density at each grid point α, β and summing over the density at each point times the grid area $(\Delta\alpha\Delta\beta)$.

   (d) Keeping the value of α fixed at 5, code the conjugate sampler for the posterior distribution of the parameter $\beta|\alpha, Y$. Take 1000 direct draws and report the mean, variance and posterior probability $p(\beta < 5|Y, \alpha = 5)$.

(e) Keeping the value of β fixed at 5, code a Metropolis or Metropolis-Hastings sampler for the posterior distribution of the parameter $\alpha|\beta, Y$. Choose a proposal distribution for α that is similar to the target distribution; describe how you made your choice. After thinning the chain to remove any autocorrelation, and discarding a sufficient number of draws as burn-in, obtain 1000 draws and report the mean, variance and posterior probability $p(\alpha < 5|Y, \beta = 5)$.

(f) Given the pieces you have put together in the previous two parts, construct the Gibbs sampler for the joint distribution $(\alpha, \beta|Y)$. After thinning the chain to remove any autocorrelation, and discarding a sufficient number of draws as burn-in, obtain 1000 draws and report the mean, variance and posterior probability $p(\alpha < 5|Y)$. Compare your draws from the sampler to those from the grid approximation using a Q-Q plot for each parameter, and comment on the values of the correlations between the two parameters for each sampler.