**STAT853: Assignment 1**

## 1. Genetic Linkage Model

Assume that 197 animals are distributed into four categories as $Y = (y_1, y_2, y_3, y_4) = (125, 18, 20, 34)$ with cell probabilities $\left(\frac{1}{2} + \frac{\theta}{4}, \frac{1}{4}(1-\theta), \frac{1}{4}(1-\theta), \frac{\theta}{4}\right)$.
We assume that $\theta \sim \mathcal{U}[0,1]$, i.e. we set a uniform prior on $[0,1]$ for $\theta$.

- Design and implement an accept/reject procedure to sample from the posterior $p(\theta|\, y)$.
- Design and impement an importance sampling method to sample from the posterior $p(\theta|\, y)$.
- Compare the efficiency of both algorithms.

## 2. Mixture of Poisson Distributions

The data in Table 1 were extracted from The London Times during the years 1910-1912. The two columns labeled "Deaths $i$" refer to the number of deaths to women 80 years and older reported by day. The columns labeled "Frequency $n_i$" referer to the number of days with $i$ deaths.

| Deaths $i$ | Frequency $i$ | Deaths $i$ | Frequency $i$ |
|---|---|---|---|
| 0 | 162 | 5 | 61 |
| 1 | 267 | 6 | 27 |
| 2 | 271 | 7 | 8 |
| 3 | 185 | 8 | 3 |
| 4 | 111 | 9 | 1 |

A Poisson distribution gives a poor fit to these data, possibly because of different patterns of deaths in winter and summer. A mixture of two Poisson distributions provides a much better fit. Under such a model, the likelihood of the $N$ observations is given by

$$\begin{aligned} p(y_1, ..., y_N|\, \theta) &= \prod_{i=1}^{N} \left[\alpha \exp(-\lambda_1) \frac{\lambda_1^{y_i}}{y_i!} + (1-\alpha)\exp(-\lambda_2)\frac{\lambda_2^{y_i}}{y_i!}\right] \\ &= \prod_{i=0}^{\max_j y_j} \left[\alpha \exp(-\lambda_1) \frac{\lambda_1^{y_i}}{y_i!} + (1-\alpha)\exp(-\lambda_2)\frac{\lambda_2^{y_i}}{y_i!}\right]^{n_i} \end{aligned}$$

where $\theta = (\alpha, \lambda_1, \lambda_2)$. $\alpha \in [0,1]$ is the mixture parameter, $\lambda_1$ and $\lambda_2$ are the means of the two components, $N$ is the total numbers of days, $n_i = \sum_{j=1}^{N} 1(y_j = i)$ is the total number of days with $i$ deaths, and $\max_j y_j = 9$ for this example. We use the following prior $\alpha \sim \mathcal{U}[0,1]$ (uniform distribution on $[0,1]$) and $\lambda_i \sim \mathcal{G}(0.1, 0.1)$ (Gamma distribution with the shape parameter 0.1 and rate parameter 0.1) for $i = 1, 2$.

a) Propose and implement a Metropolis-Hastings algorithm to sample from $p\left(\theta\right| y_1,...,y_N) .$

b) Propose a Gibbs sampler to sample from $p\left(\theta\right| y_1,...,y_N)$ [Hint: It requires introducing some missing data].

c) Present meaningful point estimates of $\lambda_1$ and $\lambda_2$.