STK4021 - Applied Bayesian Analysis and Numerical Methods

Supplemental exercises for Autumn 2016 (by Ida Scheel)

Exercise 1

Assume that $\theta \sim N(10, 1)$. For different values of S ($S = 10, 100, 1000, 10000$), sample S draws from this distribution, so that you get S samples of θ: $\theta^1, \theta,^2, \ldots, \theta^S$. For each value of S:

- Make a histogram of $\theta^1, \theta,^2, \ldots, \theta^S$ and compare it to the true distribution of θ
- Calculate the average of $\theta^1, \theta,^2, \ldots, \theta^S$ and compare it to the true mean of θ
- Calculate the 2.5th and 97.5th percentiles of $\theta^1, \theta,^2, \ldots, \theta^S$, and compare with the 95% true probability interval of θ

Exercise 2

Consider $y = (y_1, y_2, \ldots, y_n)$ where the $y_i,\ i = 1, \ldots, n$ are assumed iid with $p(y_i) = N(y_i \mid \theta, \sigma)$, $p(\theta) = N(\theta \mid \mu_0, \tau_0^2)$ and σ, μ_0, τ_0 are fixed constants. Show that

$$p(\theta \mid y) = N(\theta \mid \mu_n, \tau_n)$$

where

$$\mu_n = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{n}{\sigma^2}\bar{y}}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}}$$
$$\tau_n^2 = \frac{1}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}}$$

This results from the fact that the prior density is of the same functional form as the likelihood, hence it is an example of a prior distribution being naturally conjugate to the sampling distribution.

Exercise 3: The multivariate normal distribution

The bivariate Normal distribution for two dependent variables should be known. If x_1 has mean μ_1, variance σ_1^2, and x_2 has mean μ_2, variance σ_2^2, and the correlation between x_1 and x_2 is ρ, then (x_1, x_2) has the bivariate normal distribution if the joint density of x_1 and x_2 is

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{ -\frac{\frac{(x_1-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}}{2(1-\rho^2)} \right\} \tag{1}$$

Now let $x = (x_1, x_2)^T$, $\mu = (\mu_1, \mu_2)^T$ and

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

Show that we can write

$$f(x) = f(x_1, x_1) = \frac{1}{\sqrt{(2\pi)^2|\Sigma|}} \exp\left\{ -\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) \right\}$$

This generalizes to k variables $x = (x_1, \ldots, x_k)^T$ with respective means given by the mean vector $\mu = (\mu_1, \ldots, \mu_k)^T$ and variances and covariances given by the $k \times k$ symmetric, positive definite covariance matrix Σ, where element (i, j) corresponds to $\text{Cov}(x_i, x_j)$ (and hence element (i, i) is the variance of x_i). Then x has the multivariate normal distribution if the joint density is

$$f(x) = \frac{1}{\sqrt{(2\pi)^{-k}|\Sigma|}} \exp\left\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}$$

More on properties of this distribution can be found in Appendix A in the textbook, page 580.

Exercise 4: The full Bayesian model for the meta-analysis example

The meta-analysis data in Table 5.4 the textbook [1] was analysed in Ch. 5.6 and Exercise 15. The data are from $J = 22$ different clinical trials, with trial j having n_j^C subjects in the control group and n_j^T subjects in the treatment group, resulting in y_j^C and y_j^T number of deaths in the control and treatment group, respectively. Remember from the lectures that we put up a more complete Bayesian model, which was

$$\begin{aligned}
&y_j^C \sim \text{Bin}(n_j^C, p_j^C),\ j = 1, \ldots, J \\
&y_j^T \sim \text{Bin}(n_j^T, p_j^T),\ j = 1, \ldots, J \\
&\text{logit}(p_j^C) = \beta_j,\ j = 1, \ldots, J \\
&\text{logit}(p_j^T) = \beta_j + \theta_j,\ j = 1, \ldots, J \\
&\theta_j \sim N(\mu, \tau^2),\ j = 1, \ldots, J \\
&\beta_j \sim N(0, 10^5),\ j = 1, \ldots, J \\
&\mu \sim N(0, 10^6) \\
&\tau^2 \sim \text{Inv-Gamma}(10^{-3}, 10^{-3})
\end{aligned}$$

a. Write the full conditional distributions (up to a constant of proportionality) for all the parameters
b. For which parameters is the full conditional distribution a known distribution (i.e. Gibbs sampling is straightforward), and for which parameters is it not (we need a more advanced method)?
c. Perform posterior analysis for this model using the data in Table 5.4 by using MCMC. Use overdispersed starting values and calculate $\hat{R}$ for the estimands of particular interest, which are mu, τ and the effect $\tilde{\theta}$ for a future study
d. Compare results with the corresponding from the simpler analysis in Chapter 5 in the textbook, Exercise 15.

References

[1] A. Gelman, J. B. Carlin, H. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian Data Analysis*. Chapman&Hall/CRC Texts in statistical science, third edition, 2014.