STAT40380/40390 - Bayesian Analysis

Tutorial Sheet 4

Question 1

A linear regression model is proposed as:

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\boldsymbol{y}$ is a vector of n response values, $\boldsymbol{X}$ a $n \times (p+1)$ design matrix of known values, $\boldsymbol{\beta}$ a vector of $p+1$ unknown regression parameters, and $\boldsymbol{\epsilon} \sim N(0, \phi \boldsymbol{I})$. A prior distribution $p(\boldsymbol{\beta}, \phi) \propto \frac{1}{\phi}$ is proposed. Show that the posterior distribution of $\boldsymbol{\beta}|\phi, \boldsymbol{y}$ is:

$$N\left(\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{y}, \phi\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\right).$$

Similarly, show that the posterior distribution of $\phi|\boldsymbol{\beta}, \boldsymbol{y}$ is $IG(n/2, S/2)$ where $S = (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})^T(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})$.
Solutions:

$$\begin{aligned}
p(\boldsymbol{\beta}, \phi|\boldsymbol{y}) &\propto \phi^{-1}\phi^{-n/2}e^{-\frac{1}{2\phi}(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})^T(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{\beta})} \\
&= \phi^{-n/2-1}e^{-\frac{1}{2\phi}\left(\boldsymbol{y}^T\boldsymbol{y}-2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{y}+\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta}\right)} \\
p(\boldsymbol{\beta}|\boldsymbol{y}, \phi) &\propto \exp\left[-\frac{1}{2\phi}\left(\boldsymbol{y}^T\boldsymbol{y} - 2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{y} + \boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta}\right)\right] \\
&\propto \exp\left[-\frac{1}{2\phi}\left(-2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{y} + \boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta}\right)\right]
\end{aligned}$$

Note that:

$$\begin{aligned}
\left[\boldsymbol{\beta} - \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{y}\right]^T\left(\boldsymbol{X}^T\boldsymbol{X}\right)\left[\boldsymbol{\beta} - \left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{y}\right] &= \boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{y}^T\boldsymbol{X}\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{y} - 2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{y} \\
&\propto -2\boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{y} + \boldsymbol{\beta}^T\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta}
\end{aligned}$$

so $\boldsymbol{\beta}|\phi, \boldsymbol{y} \sim N\left(\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\boldsymbol{X}^T\boldsymbol{y}, \phi\left(\boldsymbol{X}^T\boldsymbol{X}\right)^{-1}\right)$. Similarly:

$$p(\phi|\boldsymbol{\beta}, \boldsymbol{y}) \propto \phi^{-n/2-1}e^{-\frac{1}{2\phi}S}$$

so $\phi|\boldsymbol{\beta}, \boldsymbol{y} \sim IG(n/2, S/2)$.

Question 2

The energy of particles being emitted form a radioactive source are modelled using a $N(\mu, \rho\mu^2)$ distribution where ρ is set at 1. A nuclear physicist said that he believes that particles will be emitted with a mean energy of 80MeV but he's not certain about that and it could be anywhere between 50MeV and 110MeV.

(a) Specify a gamma prior that reflects the nuclear physicist's opinions on the average energy of the particles. Assume that the range specified is $\pm$ 3 standard deviations from the mean.

(b) The energy of eight particles was recorded as: 50, 60, 60, 80, 40, 40, 80 and 70MeV. Show that the likelihood is of the form:

$$p(\boldsymbol{x}|\mu) \propto \frac{1}{\mu^8} \exp\left[-\frac{15300}{\mu^2} + \frac{480}{\mu} - 4\right]$$

(c) Find the posterior density, up to a constant.

(d) Describe, in detail, a suitable procedure for sampling values from the posterior. Comment on the efficiency of the method you propose,

(e) Suppose that the physicist want to extend the model so that ρ is treated as unknown. Outline the extra steps that would need to be taken to complete a Bayesian inference for this extended model

Solutions:
(a) Should get $\alpha = 64,\ \beta = 0.8$ (or $1/0.8$ if using other version of gamma)
(b)

$$\begin{aligned}
p(\boldsymbol{x}|\mu) &\propto \mu^{-n} e^{-\frac{1}{2\mu^2}\sum(x_i-\mu)^2} \\
&= \mu^{-n} e^{-\frac{1}{2\mu^2}\left(\sum x_i^2 - 2\mu\sum x_i + n\mu^2\right)} \\
&= \mu^{-8} \exp\left[-\frac{1}{2\mu^2} 30600 - 2.480.\mu + n\mu^2\right] \\
&= \mu^{-8} \exp\left[-\frac{15300}{\mu^2} + \frac{480}{\mu} - 4\right]
\end{aligned}$$

(c) The prior distribution is $p(\mu) \propto \mu^{63} e^{-0.8\mu}$ so posterior is:

$$p(\mu|\boldsymbol{x}) \propto \mu^{55} \exp\left[-\frac{15300}{\mu^2} + \frac{480}{\mu} - 4 - 0.8\mu\right]$$

(d) Need to talk about Rejection sampling or grid-based methods
(e) Needs to specify a prior distribution and then talk about Gibbs or MCMC based methods