University College Dublin
An Coláiste Ollscoile, Baile Átha Cliath

SEMESTER 2 EXAMINATION 2013/2014

STAT 40380/40390

Bayesian Analysis

Dr Patrick Murphy

Professor Tom Fearn

Professor Nial Friel*

Time Allowed: 2 hours

Instructions for Candidates

Full marks for *four* complete questions.
Marks are indicated for each question.

Instructions for Invigilators

Non-programmable calculators are permitted.

1. Data x_i arise from a normal distribution, $N(\mu, \sigma^2)$, for $i = 1, \ldots, n$ where σ^2 is known.

(a) Write down the likelihood for the model, up to the constant of proportionality.
[3 marks]

(b) Define the Jeffreys prior distribution. Why would a Jeffreys prior distribution be chosen for a parameter?
[5 marks]

(c) Show that the Jeffreys prior for μ is flat.
[7 marks]

(d) Show that the posterior distribution for $\mu | x_1, \ldots, x_n$ is also Normally distributed when the Jeffreys prior is used.
[5 marks]

(e) An alternative model is suggested such that $x_i \sim N(\mu_i, \sigma^2)$, for $i = 1, \ldots, n$. What does this new model represent? Suggest a suitable prior distribution for μ_i.
[7 marks]

[Total 25 marks]

2. A hospital manager is concerned that the number of deaths in a particular ward is too high. Over the previous 10 quarters, the number of deaths x_i, $i = 1, \ldots, 10$ are given below:

$$11,\ 11,\ 10,\ 8,\ 15,\ 11,\ 3,\ 7,\ 6,\ 8.$$

The hospital manager believes a Poisson distribution is appropriate for the data. Of key interest is the rate parameter λ at which deaths occur in the ward in each quarter.

(a) Show that the likelihood for the above data can be written as:

$$p(\boldsymbol{x}|\lambda) \propto e^{-10\lambda}\lambda^{90}$$

[3 marks]

(b) The hospital manager believes a log-normal prior distribution is appropriate for the parameter λ. From looking at other wards, she believes that the parameters of the log-normal distribution should be $\mu = 1.8$ and $\sigma = 1$. Use this information to write down the posterior distribution for $\lambda|\boldsymbol{x}$ up to a suitable constant. (Hint: the log-normal pdf is given at the end of the paper.)

[6 marks]

(c) By looking at the prior mean and the sample mean of the data, is it conceivable that the ward is experiencing higher than normal death rates? How might such a hypothesis be checked in a Bayesian setting?

[7 marks]

(d) Suggest in detail an appropriate procedure for sampling values from the posterior distribution of $\lambda|\boldsymbol{x}$.

[10 marks]

[Total 26 marks]

3. A Bayesian linear regression model is proposed for data (y_i, x_i) for $i = 1, \ldots, n$ where:

$$y_i = \alpha + \beta(x_i - \bar{x}) + \epsilon_i,$$

where $\bar{x} = \frac{1}{n}\sum x_i$ and $\epsilon_i \sim N(0, \phi^{-1})$ with ϕ an unknown precision parameter. A reference prior distribution is proposed so that $p(\alpha, \beta, \phi) \propto \frac{1}{\phi}$.

(a) Show that the posterior distribution can be written as:

$$p(\alpha, \beta, \phi|\boldsymbol{y}, \boldsymbol{x}) \propto \phi^{n/2-1} e^{-\frac{\phi}{2}\left[S_{yy}+n(\bar{y}-\alpha)^2+\beta^2 S_{xx}-2\beta S_{xy}\right]}$$

where $S_{xy} = \sum(x_i - \bar{x})(y_i - \bar{y})$. **[7 marks]**

(b) What is a complete conditional distribution? Of what use are they in fitting a Bayesian model like the above?

[4 marks]

(c) Show that the complete conditional distribution of α is such that:

$$\alpha| \ldots \sim N\left(\bar{y}, (n\phi)^{-1}\right)$$

[4 marks]

(d) Show that the complete conditional distribution of ϕ is such that:

$$\phi| \ldots \sim Ga\left(\frac{n}{2}, \frac{S_{yy} + n(\bar{y} - \alpha)^2 + \beta^2 S_{xx} - 2\beta S_{xy}}{2}\right)$$

[4 marks]

(e) Show that the complete conditional distribution of β is such that:

$$\beta| \ldots \sim N\left(\frac{S_{xy}}{S_{xx}}, (\phi S_{xx})^{-1}\right)$$

[6 marks]

[Total 25 marks]

4. Write short notes on **three** of the following topics.

(a) Convergence and convergence diagnostics for MCMC.

[8 marks]

(b) Conjugacy in Bayesian models.

[8 marks]

(c) Summarising posterior distributions.

[8 marks]

(d) Methods for finding posterior modes.

[8 marks]

[Total 24 marks]

Probability distributions

Normal distribution

$$p(x|\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]$$

$$\mathbb{E}(x) = \mu,\ Var(x) = \sigma^2$$

Log-normal distribution

$$p(x|\mu,\sigma) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(\log(x)-\mu)^2\right]$$

$$\mathbb{E}(x) = e^{\mu+\sigma^2/2},\ Var(x) = (e^{\sigma^2}-1)e^{2\mu+\sigma^2}$$

Gamma distribution

$$p(x|\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1}\exp(-\beta x)$$

$$\text{where } \Gamma(u) = \int_0^\infty t^{u-1}\exp(-t)\ dt$$

$$\mathbb{E}(x) = \frac{\alpha}{\beta},\ Var(x) = \frac{\alpha}{\beta^2}$$