University College Dublin
An Coláiste Ollscoile, Baile Átha Cliath

SEMESTER 2 EXAMINATION 2010/2011

STAT 40390
Bayesian Analysis

Dr Michaél Ó Searcóid

Professor Roger Payne

Dr Andrew Parnell*

Time Allowed: 2 hours

Instructions for Candidates

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

Instructions for Invigilators

Calculators are permitted.

1. Some data x_i arise from a Binomial distribution $Bin(k, \theta)$ for $i = 1, \ldots, n$ with k fixed.

(a) Write down the likelihood for the model, up to the constant of proportionality. **[2 marks]**

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

[4 marks]

(c) Show that the Jeffreys prior for θ is proportional to a $Be(0.5, 0.5)$ distribution. **[6 marks]**

(d) Show that the posterior distribution for $\theta | x_1, \ldots, x_n$ is also Beta distributed when the Jeffreys prior is used.

[4 marks]

(e) Suppose that k is now also considered to be an unknown parameter. What kind of prior distributions would be appropriate to use to fit the model? How might you estimate this prior distribution?

[4 marks]

(f) Write sample WinBUGS code to fit such a model. How might such code be changed to compare different prior structures?

[5 marks]

[Total 25 marks]

2. The expenditure of $n = 10$ different departments at a major Dublin hospital is recorded as x_i (in millions of €) below:

$$2.77,\ 1.01,\ 2.00,\ 3.92,\ 2.76,\ 2.77,\ 1.53,\ 2.85,\ 2.41,\ 1.81,$$

The observations are believed to come from a normal distribution with unknown mean τ but fixed variance 0.5.

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

$$p(\boldsymbol{x}|\tau) \propto e^{-n(\bar{x}-\tau)^2}$$

[3 marks]

(b) The hospital manager believes that the mean expenditure should be centered around 2 with variance 1.5. Estimate the parameters of a gamma distribution which represent the manager's beliefs.

[3 marks]

(c) Write down the posterior distribution for the manager up to a suitable constant.

[2 marks]

(d) Suggest in detail an appropriate procedure for sampling values from the posterior.

[8 marks]

(e) Define the Deviance Information Criterion as:

$$DIC = 2\overline{D(\tau)} - D(\bar{\tau})$$

where D is the deviance and $\bar{\tau}$ is the posterior mean of τ. Why is the DIC used? If $\overline{D(\tau)} = 24.667$, $\bar{\tau} = 2.35$ and $\sum(x_i - \bar{\tau})^2 = 6.12$, calculate the DIC for the above model.

[5 marks]

[Total 21 marks]

3. A model is defined as being in the 2-parameter exponential family if the likelihood can be written as proportional to:

$$h(\theta, \phi)^n \exp\left[\sum t(x_i)\psi(\theta, \phi) + \sum u(x_i)\chi(\theta, \psi)\right]$$

for parameters (θ, ϕ), data $x_1, \ldots, x_n$ and functions h, t, u, ψ and χ.

(a) Show that the Gamma $Ga(\theta, \phi)$ distribution is a member of the 2-parameter exponential family with $\sum t(x_i) = \sum \log x_i$ and $\sum u(x_i) = \sum x_i$.

[4 marks]

(b) What role to t and u play in estimating the parameters of the distribution?

[3 marks]

(c) What is a conjugate prior distribution? Show that the prior distribution defined by:

$$p(\theta, \phi) \propto h(\theta, \phi)^\nu \exp\left[\tau\psi(\theta, \phi) + \omega\chi(\theta, \phi)\right]$$

with constants ν, τ and ω, is conjugate for the two-parameter exponential family.

[3 marks]

(d) Find the form of the conjugate prior for the Gamma distribution with parameters (θ, ϕ).

[2 marks]

(e) Some observations are made to give $n = 5$, $\sum x_i = 12$ and $\log(\prod x_i) = 3$. With $\nu = -4$, $\tau = -2$, and $\omega = -2$, write out the posterior distribution up to a constant.

[2 marks]

(f) Hence show that the marginal posterior distribution for $\theta|\boldsymbol{x}$ is:

$$p(\theta|\boldsymbol{x}) \propto \frac{\theta e^\theta}{10^{\theta+1}}$$

(Hint: remember $\Gamma(y + 1)/\Gamma(y) = y$ if $y > 0$)

[6 marks]

[Total 20 marks]

4. Write short notes (approximately 200 words) on **three** of the following topics.

(a) The likelihood, conditionality, and sufficiency principles.

[8 marks]

(b) The use of Bayes factors in comparing models.

[8 marks]

(c) The error of the transconditional.

[8 marks]

(d) Stopping rules.

[8 marks]

[Total 24 marks]

oOo

Probability distributions

Normal distribution

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

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

Binomial distribution

$$P(x|n, p) = \binom{n}{x} p^x (1-p)^{n-x}$$

$$\mathbb{E}(x) = np,\ Var(x) = np(1-p)$$

Beta distribution

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

$$\text{where } B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\, dt$$

$$\mathbb{E}(x) = \frac{\alpha}{\alpha+\beta},\ Var(x) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Gamma distribution

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

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

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