University College Dublin
An Coláiste Ollscoile, Baile Átha Cliath

SEMESTER 2 EXAMINATION 2014/2015

STAT 40380/40390

Bayesian Analysis

Professor Gary McGuire

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. Suppose that we take a Bayesian approach to make inference for a continuously varying unknown parameter θ in a statistical model, given a vector $y = (y_1, \ldots, y_n)$ of independent identically distributed observations.

(a) Discuss briefly the interpretation of the *prior density* $\pi(\theta)$ for θ. Indicate how it may be used to incorporate prior belief about about this parameter, and discuss the use of informative and noninformative prior distributions.

[6 marks]

(b) Suppose that each observation y_i has density function $f(y_i|\theta)$. What is the definition of the likelihood function $L(y|\theta)$ **[5 marks]**

(c) Describe carefully how the *posterior distribution* $\pi(\theta|y)$ is determined from the prior density and the likelihood function. **[6 marks]**

(d) Describe carefully how, given the posterior distribution of θ, you would determine an *equi-tailed* 95% *credible interval* for θ. Give the interpretation of this interval, and contrast it with that of a 95% confidence interval for θ obtained using classical statistical techniques. **[8 marks]**

[Total 25 marks]

2. An insurance company assumes that motor insurance claims arise as a Poisson process with rate λ per week. The company would like to carry out a Bayesian estimation of λ. It collects data over a 25 week period observing a total of 130 claims.

(a) Previous experience suggests the use of a gamma prior distribution, $Ga(20, 5)$, for λ. What is the posterior distribution for λ? **[8 marks]**

(b) Compare the mean and standard deviation of the prior to that of the posterior and comment briefly. **[4 marks]**

(c) Use the normal approximation to the gamma distribution to determine a 95% credible region for λ based on its posterior distribution. **[4 marks]**

(d) For a given λ, what is the probability (under the insurance company's assumption) that, in a given period of 1 week, there will be no claims? **[4 marks]**

(e) Write down an expression for the predictive probability, under the above posterior distribution, that, in a given period of 1 week, there will be no claims? (Note, you do not need to evaluate this probability). **[6 marks]**

[Total 26 marks]

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

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

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 $\sum u(x_i) = \sum x_i$.

[5 marks]

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

[4 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.

[4 marks]

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

[3 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.

[3 marks]

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

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

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

[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

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

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

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

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

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

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

Poisson distribution

$$p(x|\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}$$

$$\mathbb{E}(x) = \lambda,\ Var(x) = \lambda.$$

Gamma distribution

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

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

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