

MAS3902

NEWCASTLE UNIVERSITY

SCHOOL OF MATHEMATICS, STATISTICS & PHYSICS

Specimen Exam Paper 2

MAS3902

Bayesian Inference

Time allowed: 2 hours

Candidates should attempt all questions. Marks for each question are indicated. However you are advised that marks may be adjusted in accordance with the University's Moderation and Scaling Policy.

There are THREE questions in Section A and TWO questions in Section B.

Calculators may be used.

SECTION A

A1. A two-parameter model was deemed appropriate for a sample of lifetimes, in months, for patients receiving a new drug. In the absence of a conjugate prior specification, and without the availability of the full conditional distributions for each parameter, a Metropolis-Hastings Markov chain Monte Carlo (MCMC) scheme was devised to obtain a sample from the marginal posteriors for each parameter.

Figure 1 below shows output from a pilot run of the MCMC scheme. The parameters, μ and σ , represent parameters of location and scale respectively.

Look at the plots in Figure 1, and comment on anything interesting.

[8 marks]

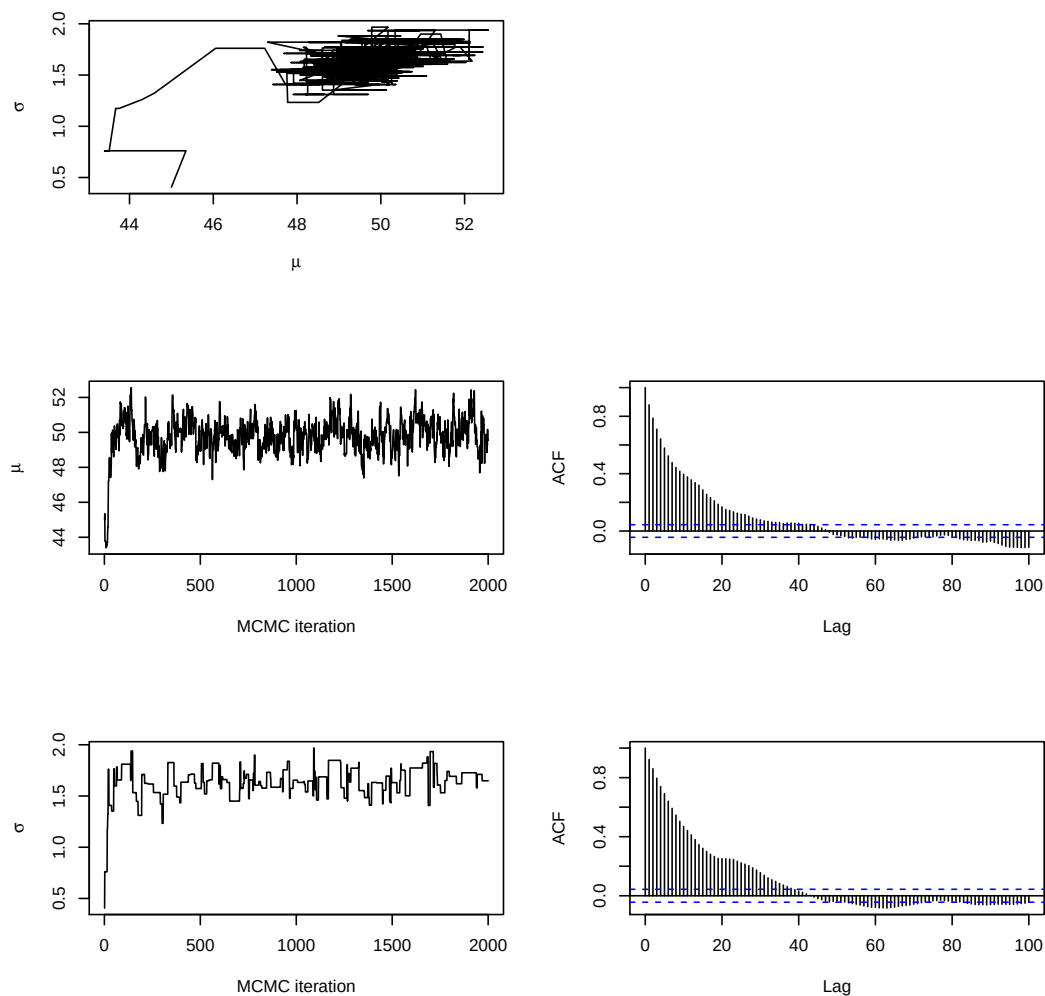


Figure 1: MCMC output for parameters in a model for lifetimes of patients.

A2. A book printer has a contract to deliver their product to a publisher. In order to assess the quality of the products they receive, the publisher examines a random sample of n books and finds the number of pages in each book with errors: $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$.

Suppose that, in each book, the number of pages with errors has a Poisson distribution with mean θ . Also, the publisher expresses his prior beliefs about θ using a gamma $Ga(g, h)$ distribution.

(a) Show that the likelihood function is

$$f(\mathbf{x}|\theta) \propto \theta^{n\bar{x}} e^{-n\theta},$$

where $\bar{x}$ is the mean number of errors per book in the sample.

- (b) Show that the posterior distribution for θ is a $Ga(G, H)$ distribution where $G = g + n\bar{x}$ and $H = h + n$.
- (c) The publisher is now interested in the number of pages Y with errors in the next book sent by the printer. Assuming that $Y|\theta \sim Po(\theta)$ and that G and H are integers, show that the predictive probability function determined by the publisher is

$$f(y|\mathbf{x}) = \binom{G+y-1}{G-1} \left(\frac{H}{H+1} \right)^G \left(\frac{1}{H+1} \right)^y \quad y = 0, 1, \dots$$

Hints: Candidate's formula is

$$f(y|\mathbf{x}) = \frac{f(y|\theta)\pi(\theta|\mathbf{x})}{\pi(\theta|\mathbf{x}, y)}$$

and $\Gamma(m) = (m-1)!$ if m is an integer.

[12 marks]

A3. A random sample $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ is taken from a log-normal $LN(\mu, \sigma^2)$ distribution.

(a) Show that the likelihood function is

$$f(\mathbf{x}|\mu, \sigma) \propto \sigma^{-n} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (\log x_i - \mu)^2 \right\}.$$

(b) Show that

$$\begin{aligned} \frac{\partial^2}{\partial \mu^2} \log f(\mathbf{x}|\mu, \sigma) &= -\frac{n}{\sigma^2} \\ \frac{\partial^2}{\partial \mu \partial \sigma} \log f(\mathbf{x}|\mu, \sigma) &= -\frac{2}{\sigma^3} \sum_{i=1}^n (\log x_i - \mu) \\ \frac{\partial^2}{\partial \sigma^2} \log f(\mathbf{x}|\mu, \sigma) &= \frac{n}{\sigma^2} - \frac{3}{\sigma^4} \sum_{i=1}^n (\log x_i - \mu)^2. \end{aligned}$$

(c) Hence show that the Jeffreys prior is

$$\pi(\mu, \sigma) \propto \frac{1}{\sigma^2}, \quad -\infty < \mu < \infty, \quad \sigma > 0.$$

Hint: For $i = 1, 2, \dots, n$, $\log X_i \sim N(\mu, \sigma^2)$ and so $E(\log X_i) = \mu$ and $Var(\log X_i) = \sigma^2$.

(d) Is this a proper prior distribution? Explain your answer.

(e) Are μ and σ independent in this prior distribution? Explain your answer.

[20 marks]

SECTION B

- B4.** A random sample $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ is taken from a normal $N(\mu, 1/\tau)$ distribution, with sample mean $\bar{x}$ and variance $s^2 = \sum_{i=1}^n (x_i - \bar{x})^2/n$. Suppose your prior beliefs about $(\mu, \tau)^T$ follow a normal-gamma $NGa(b, c, g, h)$ distribution.

(a) Show that the likelihood function is

$$f(\mathbf{x}|\mu, \tau) = \left(\frac{\tau}{2\pi}\right)^{n/2} \exp \left[-\frac{n\tau}{2} \{s^2 + (\bar{x} - \mu)^2\} \right].$$

[6 marks]

- (b) Show that your posterior distribution for $(\mu, \tau)^T$ is a $NGa(B, C, G, H)$ distribution where

$$B = \frac{bc + n\bar{x}}{c + n}, \quad C = c + n, \quad G = g + \frac{n}{2}, \quad H = h + \frac{cn(\bar{x} - b)^2}{2(c + n)} + \frac{ns^2}{2}.$$

[6 marks]

Hint:

$$c(\mu - b)^2 + n(\bar{x} - \mu)^2 + 2h + ns^2 = C(\mu - B)^2 + 2H.$$

- (c) Use this joint posterior distribution to determine your marginal posterior density for μ . Name the posterior distribution for μ .

[7 marks]

Hint:

$$\int_0^\infty \theta^{a-1} e^{-b\theta} d\theta = \frac{\Gamma(a)}{b^a}.$$

- (d) Hence, obtain a $100(1 - \alpha)\%$ posterior highest density interval for μ in terms of the upper percentage points $t_{\nu;p}$ of the standard t_ν distribution, where $\Pr(t_\nu > t_{\nu;p}) = p$.

[6 marks]

- (e) Briefly explain how you might attempt to construct a joint credible region for $\begin{pmatrix} \mu \\ \tau \end{pmatrix} \Big| \mathbf{x}$.

[5 marks]

[Total: 30 marks]

B5. The *Pareto* distribution is commonly used by geologists to describe the size of meteorites that land on earth. Suppose you have a random sample $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$ of meteorite lengths assumed from a Pareto $Pa(\theta, \alpha)$ distribution, where θ and α are unknown.

Discussions with a geologist lead to independent prior distributions such that $\alpha \sim Ga(a, b)$ and $\theta \sim Ga(c, d)$, with a, b, c and d known.

(a) Show that the likelihood function is

$$f(\mathbf{x}|\alpha, \theta) = \alpha^n \theta^{n\alpha} \bar{x}_g^{-n(\alpha+1)}, \quad 0 < \theta < x_{(1)},$$

where $\bar{x}_g = \sqrt[n]{\prod x_i}$ is the geometric mean and $x_{(1)}$ is the minimum of the data. **[3 marks]**

(b) Show that your posterior density is

$$\pi(\alpha, \theta|\mathbf{x}) \propto \alpha^{a+n-1} \theta^{c+n\alpha-1} \bar{x}_g^{-n(\alpha+1)} e^{-b\alpha-d\theta}, \quad \alpha > 0, \quad 0 < \theta < x_{(1)}.$$

[4 marks]

(c) Using the posterior from part (b),

(i) show that the posterior conditional distribution for $\alpha|\theta$ is a Gamma distribution. State the parameters of this Gamma distribution [Hint: $x^y = e^{y \log x}$]; **[5 marks]**

(ii) determine the posterior conditional density for $\theta|\alpha$ up to a multiplicative constant. **[3 marks]**

(d) You develop an MCMC algorithm with separate steps for each parameter to simulate realisations from the posterior distribution.

(i) Why might a Metropolis within Gibbs algorithm be appropriate for simulating realisations from the posterior? **[2 marks]**

(ii) You decide to use a normal $N(\theta, \Sigma_\theta)$ random walk proposal for θ , that is, make proposals using $\theta^*|\theta \sim N(\theta, \Sigma_\theta)$. What is the acceptance probability for the proposed value θ^* ? **[4 marks]**

(iii) Write down the steps in your MCMC algorithm for simulating realisations from the posterior. **[9 marks]**

[Total: 30 marks]

Distributions for data

Binomial distribution

If $X|\theta \sim \text{Bin}(k, \theta)$ then it has probability function

$$f(x|\theta) = \binom{k}{x} \theta^x (1 - \theta)^{k-x}, \quad x = 0, 1, \dots, k,$$

where k is a positive integer and $0 < \theta < 1$. Also, $E(X) = k\theta$ and $\text{Var}(X) = k\theta(1 - \theta)$.

Exponential distribution

If $X|\lambda \sim \text{Exp}(\lambda)$ then it has density

$$f(x|\lambda) = \lambda e^{-\lambda x}, \quad x > 0$$

where $\lambda > 0$. Also, $E(X) = 1/\lambda$ and $\text{Var}(X) = 1/\lambda^2$.

Gamma distribution

If $X|\alpha, \lambda \sim \text{Ga}(\alpha, \lambda)$ then it has density

$$f(x|\alpha, \lambda) = \frac{\lambda^\alpha x^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}, \quad x > 0,$$

where $\alpha > 0$ and $\lambda > 0$. Also, $E(X) = \alpha/\lambda$, $\text{Var}(X) = \alpha/\lambda^2$ and (for $\alpha \geq 1$) $\text{Mode}(X) = (\alpha - 1)/\lambda$; and $\Gamma(A) = (A - 1)!$ for integer A .

Inverse Gaussian distribution

If $X|\mu, \lambda \sim \text{IG}(\mu, \lambda)$ then it has density

$$f(x|\mu, \lambda) = \frac{\sqrt{\lambda}}{\sqrt{2\pi x^3}} \exp \left\{ -\frac{\lambda(x - \mu)^2}{2\mu^2 x} \right\}, \quad x > 0,$$

where $\mu > 0$ and $\lambda > 0$. Also, $E(X) = \mu$ and $\text{Var}(X) = \mu^3/\lambda$.

Laplace distribution

If $X|\mu, \sigma \sim \text{La}(\mu, \sigma^2)$ then it has density

$$f(x|\mu, \sigma) = \frac{1}{\sqrt{2} \sigma} \exp \left\{ -\frac{\sqrt{2}|x - \mu|}{\sigma} \right\}, \quad x > 0,$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$. Also, $E(X) = \mu$ and $\text{Var}(X) = \sigma^2$.

Log-normal distribution

If $X|\mu, \sigma \sim LN(\mu, \sigma^2)$ then it has density

$$f(x|\mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp \left\{ -\frac{1}{2\sigma^2}(\log x - \mu)^2 \right\}, \quad x > 0$$

where $\mu \in \mathbb{R}$ and $\sigma > 0$. Also, $E(X) = e^{\mu+\sigma^2/2}$, $Var(X) = (e^{\sigma^2} - 1)e^{2\mu+\sigma^2}$. Further $E(\log X) = \mu$ and $Var(\log X) = \sigma^2$.

Negative binomial distribution

If $X|\theta \sim NegBin(k, \theta)$ then it has probability function

$$f(x|\theta) = \binom{x-1}{k-1} \theta^k (1-\theta)^{x-k}, \quad x = k, k+1, \dots,$$

where k is a positive integer and $0 < \theta < 1$. Also $E(X) = k/\theta$ and $Var(X) = k(1-\theta)/\theta^2$.

Normal distribution

If $X|\mu, \tau \sim N(\mu, 1/\tau)$ then it has density

$$f(x|\mu, \tau) = \left(\frac{\tau}{2\pi}\right)^{1/2} \exp \left\{ -\frac{\tau}{2}(x - \mu)^2 \right\}, \quad x \in \mathbb{R}$$

where $\mu \in \mathbb{R}$ and $\tau > 0$. Also, $E(X) = \mu$ and $Var(X) = 1/\tau$. The distribution has the following quantiles

x	1.2816	1.6449	1.9600	2.3263	2.5758
$\Pr(X < x)$	0.9	0.95	0.975	0.99	0.995

Pareto distribution

If $X|\theta, \alpha \sim Pa(\theta, \alpha)$ then it has density

$$f(x|\alpha, \theta) = \frac{\alpha\theta^\alpha}{x^{\alpha+1}}, \quad x > \theta$$

where $\alpha > 0$ and $\theta > 0$. Also $E(X) = \alpha\theta/(\alpha - 1)$, $\alpha > 1$ and $Var(X) = \alpha\theta^2/\{(\alpha - 1)^2(\alpha - 2)\}$, $\alpha > 2$.

Poisson distribution

If $X|\theta \sim Po(\theta)$ then it has probability function

$$f(x|\theta) = \frac{\theta^x e^{-\theta}}{x!}, \quad x = 0, 1, \dots,$$

where $\theta > 0$. Also, $E(X) = \theta$ and $Var(X) = \theta$.

Rayleigh distribution

If $X|\theta \sim R(\theta)$ then it has density

$$f(x|\theta) = 2x\theta e^{-\theta x^2}, \quad x > 0,$$

where $\theta > 0$. Also $E(X) = \sqrt{\pi/(4\theta)}$ and $Var(X) = (4 - \pi)/(4\theta)$.

Uniform distribution

If $X|\phi, \theta \sim U(\phi, \theta)$ then it has density

$$f(x|\phi, \theta) = \frac{1}{\theta - \phi}, \quad \phi < x < \theta,$$

where $\phi < \theta$. Also, $E(X) = (\phi + \theta)/2$ and $Var(X) = (\theta - \phi)^2/12$.

von Mises distribution

If $X|\mu, \lambda \sim vM(\mu, \lambda)$ then it has density

$$f(x|\mu, \lambda) = \frac{1}{2\pi I_0(\lambda)} \exp\{\lambda \cos(x - \mu)\}, \quad 0 \leq x < 2\pi$$

where $0 \leq \mu < 2\pi$ and $\lambda > 0$. Also $I_0(\cdot)$ is the modified Bessel function of order zero.

Weibull distribution

If $X|\beta, \lambda \sim Wei(\beta, \lambda)$ then it has density

$$f(x|\beta, \lambda) = \beta \lambda^\beta x^{\beta-1} e^{-\lambda^\beta x^\beta}, \quad x > 0,$$

where $\beta > 0$ and $\lambda > 0$. Also, $E(X) = \Gamma(1 + 1/\beta)/\lambda$ and $Var(X) = \Gamma(1 + 2/\beta)/\lambda^2 - E(X)^2$.

Distributions for prior beliefs

Beta distribution

If $\theta \sim \text{Beta}(g, h)$ then it has density

$$\pi(\theta) = \frac{\theta^{g-1}(1-\theta)^{h-1}}{B(g, h)}, \quad 0 < \theta < 1,$$

where $g > 0$ and $h > 0$. Also, $E(\theta) = g/(g+h)$, $\text{Var}(\theta) = gh/\{(g+h)^2(g+h+1)\}$ and (when $g, h > 1$) $\text{Mode}(\theta) = (g-1)/(g+h-2)$; and $B(g, h) = \{\Gamma(g)\Gamma(h)\} / \Gamma(g+h)$, with $\Gamma(G) = (G-1)!$ for integer G .

Exponential distribution

If $\theta \sim \text{Exp}(h)$ then it has density

$$\pi(\theta) = he^{-h\theta}, \quad \theta > 0,$$

where $h > 0$. Also, $E(\theta) = 1/h$ and $\text{Var}(\theta) = 1/h^2$.

Gamma distribution

If $\theta \sim \text{Ga}(g, h)$ then it has density

$$\pi(\theta) = \frac{h^g \theta^{g-1} e^{-h\theta}}{\Gamma(g)}, \quad \theta > 0,$$

where $g > 0$ and $h > 0$. Also, $E(\theta) = g/h$, $\text{Var}(\theta) = g/h^2$ and (when $g \geq 1$) $\text{Mode}(\theta) = (g-1)/h$; and $\Gamma(G) = (G-1)!$ for integer G .

Generalised t distribution

If $\mu \sim t_a(b, c)$ then it has density

$$\pi(\mu) = \frac{\Gamma\left(\frac{a+1}{2}\right)}{\sqrt{ac\pi} \Gamma\left(\frac{a}{2}\right)} \left\{1 + \frac{(\mu - b)^2}{ac}\right\}^{-\frac{a+1}{2}}, \quad \mu \in \mathbb{R},$$

where $b \in \mathbb{R}$, $a > 0$ and $c > 0$. Also, $E(\mu) = b$ and $\text{Var}(\mu) = ac/(a-2)$ if $a \geq 2$.

Inverse Chi distribution

If $\sigma \sim \text{Inv-Chi}(a, b)$ then it has density

$$\pi(\sigma|a, b) = \frac{2b^a \sigma^{-2a-1} e^{-b/\sigma^2}}{\Gamma(a)}, \quad \sigma > 0,$$

where $a > 0$, $b > 0$ and $\Gamma(a)$ is the gamma function. Also $E(\sigma) = \sqrt{b} \Gamma(a - 1/2)/\Gamma(a)$ and $\text{Var}(\sigma) = b/(a - 1) - E(\sigma)^2$ if $a > 1$. The name of the distribution comes from the fact that $1/\sigma^2 \sim \text{Ga}(a, b) \equiv \chi_{2a}^2/(2b)$.

Log-normal distribution

If $\theta \sim \text{LN}(b, c^2)$ then it has density

$$\pi(\theta) = \frac{1}{\sqrt{2\pi} c \theta} \exp \left\{ -\frac{1}{2c^2} (\log \theta - b)^2 \right\}, \quad \theta > 0$$

where $b \in \mathbb{R}$ and $c > 0$. Also, $E(\theta) = e^{b+c^2/2}$, $\text{Var}(\theta) = (e^{c^2} - 1)e^{2b+c^2}$. Further $\log \theta \sim N(b, c^2)$ and so $E(\log \theta) = b$ and $\text{Var}(\log \theta) = c^2$.

Normal distribution

If $\mu \sim N(b, 1/d)$ then it has density

$$\pi(\mu) = \left(\frac{d}{2\pi} \right)^{1/2} \exp \left\{ -\frac{d}{2} (\mu - b)^2 \right\}, \quad \mu \in \mathbb{R},$$

where $b \in \mathbb{R}$ and $c > 0$. Also, $E(\mu) = b$ and $\text{Var}(\mu) = 1/d$.

Normal-gamma distribution

If $\begin{pmatrix} \mu \\ \tau \end{pmatrix} \sim \text{NGa}(b, c, g, h)$ then it has density

$$\pi(\mu, \tau) \propto \tau^{g-\frac{1}{2}} \exp \left\{ -\frac{\tau}{2} [c(\mu - b)^2 + 2h] \right\}, \quad \mu \in \mathbb{R}, \tau > 0$$

where $b \in \mathbb{R}$ and $c, g, h > 0$. Also, $\mu|\tau \sim N\left(b, \frac{1}{c\tau}\right)$, $\tau \sim \text{Ga}(g, h)$ and has marginal distribution $\mu \sim t_{2g}\left(b, \frac{h}{gc}\right)$.

Uniform distribution

If $\theta \sim U(a, b)$ then it has density

$$\pi(\theta) = \frac{1}{b-a}, \quad a < \theta < b,$$

where $a < b$. Also, $E(\theta) = (a+b)/2$ and $Var(\theta) = (b-a)^2/12$.

THE END