NEWCASTLE UNIVERSITY

SCHOOL OF MATHEMATICS, STATISTICS & PHYSICS

Specimen Exam Paper

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 TWO questions in Section A and TWO questions in Section B.

There are 100 marks on this exam, 40 marks for Section A and 60 marks for Section B.

Calculators may be used.

SECTION A

A1. On any day, nitrogen oxide (NO_X) emissions can reach dangerous levels in cities in China with probability θ. For ten randomly chosen days in 2018, the number of days x on which NO_X emissions reached dangerous levels is shown in the table below.

City	Beijing	Shanghai	Xi'an	Guilin	Chengdu	Guangzhou
Days x	7	6	5	1	6	3

(a) Show that the likelihood function is

$$f(\boldsymbol{x}|\theta) \propto \theta^{28}(1-\theta)^{32}.$$ [**4 marks**]

(b) (i) Show that, under the assumption of vague prior knowledge, $\theta \sim Beta(g, h)$ with $g, h \to 0$.

[**6 marks**]

(ii) Assuming the vague prior in part (i), and the likelihood from part (a), show that $\theta|\boldsymbol{x} \sim Beta(28, 32)$.

[**3 marks**]

(c) An air pollution monitoring team plan to observe NO_X emissions in Hong Kong on ten days in 2019. Show that the predictive probability function for Y, the number of days on which emissions reach dangerous levels, is

$$f_Y(y|\boldsymbol{x}) = \binom{10}{y} \frac{B(y+28, 42-y)}{B(28, 32)}, \quad y = 0, 1, \ldots, 10,$$

where $B(a, b) = \dfrac{(a-1)!(b-1)!}{(a+b-1)!}$.

[**5 marks**]

(d) Find the posterior predictive probability that NO_X emissions will reach dangerous levels in Hong Kong on exactly four of the ten monitoring days in 2019.

[**2 marks**]

[**Total: 20 marks**]

A2. The wages X of a random sample of n workers in a city follow a reparameterised log-normal $RLN(\mu, \alpha)$ distribution, where $(\mu, \alpha) \in \mathbb{R}^2$.

(a) Show that the likelihood function is

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

[**4 marks**]

(b) Show that

$$\frac{\partial^2}{\partial\mu^2}\log f(\boldsymbol{x}|\mu, \alpha) = -ne^\alpha$$
$$\frac{\partial^2}{\partial\mu\partial\alpha}\log f(\boldsymbol{x}|\mu, \alpha) = e^\alpha\sum_{i=1}^{n}(\log x_i - \mu)$$
$$\frac{\partial^2}{\partial\alpha^2}\log f(\boldsymbol{x}|\mu, \alpha) = -\frac{e^\alpha}{2}\sum_{i=1}^{n}(\log x_i - \mu)^2.$$

[**6 marks**]

(c) Hence show that the Jeffreys prior is, for $(\mu, \alpha) \in \mathbb{R}^2$

$$\pi(\mu, \alpha) \propto e^{\alpha/2}.$$

Hint: If $X \sim RLN(\mu, \alpha)$ then $E(\log X) = \mu$ and $Var(\log X) = e^{-\alpha}$.

[**8 marks**]

(d) Are μ and α independent in this prior distribution? Explain your answer.

[**2 marks**]

[**Total: 20 marks**]

SECTION B

B3. A random sample $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$ 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

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

[**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 $\left.\begin{pmatrix} \mu \\ \tau \end{pmatrix}\right| \boldsymbol{x}$.

[**5 marks**]

[**Total: 30 marks**]

B4. Wind directions $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$ at a location in the Peak District are assumed to follow a von Mises distribution $vM(\mu, \lambda)$, where μ is the mean wind direction ($0 \leq \mu < 2\pi$) and λ is known as the concentration parameter ($\lambda > 0$).

A meteorologist assumes independent priors for μ and λ:

$$\mu \sim U(0, 2\pi) \quad \text{and} \quad \lambda \sim Ga(g, h).$$

(a) Show that the likelihood function is

$$f(\boldsymbol{x}|\mu, \lambda) \propto I_0(\lambda)^{-n} \exp\left\{\lambda \sum_{i=1}^{n} \cos(x_i - \mu)\right\},$$

where $I_0(\cdot)$ is the modified Bessel function. [**3 marks**]

(b) Show that the meteorologist's posterior density is, for $0 \leq \mu < 2\pi$ and $\lambda > 0$, [**4 marks**]

$$\pi(\mu, \lambda|\boldsymbol{x}) \propto \lambda^{g-1} I_0(\lambda)^{-n} \exp\left\{\lambda \left[\sum_{i=1}^{n} \cos(x_i - \mu) - h\right]\right\}.$$

(c) Determine the posterior conditional densities for $\mu|\lambda$ and $\lambda|\mu$ up to a multiplicative constant. [**8 marks**]

You suggest the meteorologist develops a Markov chain Monte Carlo (MCMC) scheme with separate steps for each parameter to simulate realisations from the posterior distribution.

(d) A normal random walk proposal is used for μ, with variance Σ_μ; that is, $q(\mu^*|\mu) = N(\mu, \Sigma_\mu)$. The acceptance probability for the proposed value μ^* can be written as $\min(1, A_\mu)$. Determine A_μ, when $0 \leq \mu^* < 2\pi$. [**2 marks**]

(e) A normal random walk proposal is used for λ, with variance Σ_λ; that is, $q(\lambda^*|\lambda) = N(\lambda, \Sigma_\lambda)$. The acceptance probability for the proposed value λ^* can be written as $\min(1, A_\lambda)$. Determine A_λ, when $\lambda^* > 0$. [**2 marks**]

(f) Explain how you would initialise the MCMC scheme. [**2 marks**]

(g) Write down the steps in this MCMC scheme for simulating realisations from the posterior distribution. [**9 marks**]

[**30 marks**]

Distributions for data

Binomial distribution

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

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

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

Gamma distribution

If $X|\alpha, \lambda \sim Ga(\alpha, \lambda)$ 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, $E(X) = \alpha/\lambda$ and $Var(X) = \alpha/\lambda^2$.

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 $Var(\log X) = \sigma^2$.

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

Reparameterised log-normal distribution

If $X \sim RLN(\mu, \alpha)$ then it has density

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

where $-\infty < \mu < \infty$ and $-\infty < \alpha < \infty$. Also, $E(X) = \exp\{\mu + e^{-\alpha}/2\}$, $Var(X) = \{\exp(e^{-\alpha}) - 1\}\exp(2\mu + e^{-\alpha})$. Further $E(\log X) = \mu$ and $Var(\log X) = e^{-\alpha}$.

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.

Distributions for prior beliefs

Beta distribution

If $\theta \sim Beta(g, h)$ 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, $E(\theta) = g/(g+h)$ and $Var(\theta) = gh/\{(g+h)^2(g+h+1)\}$.

Gamma distribution

If $\theta \sim Ga(g, h)$ 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, $E(\theta) = g/h$ and $Var(\theta) = g/h^2$.

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 $Var(\mu) = ac/(a-2)$ if $a \geq 2$.

Inverse Chi distribution

If $\sigma \sim \textit{Inv-Chi}(a, b)$ 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 $Var(\sigma) = b/(a-1) - E(\sigma)^2$ if $a > 1$. The name of the distribution comes from the fact that $1/\sigma^2 \sim Ga(a, b) \equiv \chi^2_{2a}/(2b)$.

Log-normal distribution

If $\theta \sim LN(b, c^2)$ 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, $E(\theta) = e^{b+c^2/2}$, $Var(\theta) = (e^{c^2} - 1)e^{2b+c^2}$. Further $\log\theta \sim N(b, c^2)$ and so $E(\log\theta) = b$ and $Var(\log\theta) = c^2$.

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 $Var(\mu) = 1/d$.

Normal-gamma distribution

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

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

where $b \in \mathbb{R}$ and $c, g, h > 0$.

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