Queen Mary
University of London

Main Examination period 2017

MTH709U / MTHM042 / MTH776P: Bayesian Statistics

Duration: 3 hours

Apart from this page, you are not permitted to read the contents of this question paper until instructed to do so by an invigilator.

You should attempt ALL questions. Marks available are shown next to the questions.

Only non-programmable calculators that have been approved from the college list of non-programmable calculators are permitted in this examination. Please state on your answer book the name and type of machine used.
Statistical functions provided by the calculator may be used provided that you state clearly where you have used them.
The New Cambridge Statistical Tables 2nd Edition are provided.
A table of common distributions is provided as an appendix.

Complete all rough work in the answer book and cross through any work that is not to be assessed.

Possession of unauthorised material at any time when under examination conditions is an assessment offence and can lead to expulsion from QMUL. Check now to ensure you do not have any notes, mobile phones, smartwatches or unauthorised electronic devices on your person. If you do, raise your hand and give them to an invigilator immediately.

It is also an offence to have any writing of any kind on your person, including on your body. If you are found to have hidden unauthorised material elsewhere, including toilets and cloakrooms, it shall be treated as being found in your possession. Unauthorised material found on your mobile phone or other electronic device will be considered the same as being in possession of paper notes. A mobile phone that causes a disruption in the exam is also an assessment offence.

Exam papers must not be removed from the examination room.

Examiners: L I Pettit, H Maruri-Aguillar

Question 1. [5 marks] A random sample of n observations are available from a distribution with density $p(x|\theta)$. One model for the data is that $\theta = 3$. An alternative model is that θ has a prior $p(\theta)$.

(a) Define the **marginal likelihood** for each model and the **Bayes Factor**. [**3**]

(b) Explain how the Bayes Factor could be used to test which model is better supported by the data. [**2**]

Question 2. [20 marks] A radioactive device gives hourly counts which are assumed to have a Poisson distribution with mean λ. In fifteen hours the total count was 35.

(a) Scientist A has no experience with this device and so adopts a Jeffreys' prior for λ. Compute this prior and the resulting posterior. [**6**]

(b) Scientist B has more experience and adopts a Gamma prior with mean 2.0 and variance 0.4. Find this prior and show that the resulting posterior has a mean which is a weighted average of the prior mean and the data mean. [**6**]

(c) Scientist C adopts a mixture prior

$$p(\lambda) = 0.9\, Ga(20, 10) + 0.1\, Ga(2, 1).$$

Compute her posterior distribution.

You should leave the posterior weights as an expression that can be evaluated. [**6**]

(d) What is the advantage of using a mixture of conjugate priors over a single conjugate prior? [**2**]

Question 3. [17 marks]

(a) Show that a single Bernoulli trial x with probability of success θ can be written in exponential family form. [**3**]

(b) Identify the natural parameter. [**1**]

(c) If a random sample of n Bernoulli trials $x_1, x_2, \ldots, x_n$ is carried out, write down a minimal sufficient statistic for θ. [**1**]

(d) Show that a Beta distribution is a conjugate prior for θ by finding the density of the posterior distribution of θ. [**3**]

(e) Define a $100(1-\alpha)\%$ Highest Posterior Density (HPD) interval for θ and show that if the posterior is unimodal then it is shorter than any other $100(1-\alpha)\%$ credible interval. [**5**]

(f) If ten Bernoulli trials result in eight successes and the prior is $Be(5, 3)$, find the 95% HPD interval for θ using the appropriate table in the New Cambridge Statistical Tables. [**4**]

Question 4. [18 marks] A random variable X has a Galenshore(a, θ) distribution if its density is given by

$$p(x|a, \theta) = \frac{2}{\Gamma(a)} \theta^{2a} x^{2a-1} \exp(-\theta^2 x^2) \quad x > 0$$

where a and θ are positive parameters. For this distribution

$$\mathrm{E}[X] = \frac{\Gamma(a + 0.5)}{\theta \Gamma(a)}.$$

Assume that a is known and consider $X_1, \ldots, X_n$ independent observations from a Galenshore(a, θ).

(a) Suppose that a Galenshore(α, β) prior is assumed for the parameter θ. Show that the posterior distribution of θ given $X_1, \ldots, X_n$ is a Galenshore$(a^\star, \theta^\star)$ distribution where $a^\star = na + \alpha$ and $\theta^\star = \sqrt{\beta^2 + \sum X_i^2}$. **[6]**

(b) If $a = 1$, find the posterior mean of θ. **[2]**

(c) Suppose now that $a = 1$, $\alpha = 1$ and $\beta = 1$. Find the density of the predictive distribution of a new observation Y from a Galenshore$(1, \theta)$ distribution. **[6]**

(d) If $n = 2$ and $X_1 = 0.5$ and $X_2 = 1$, find the probability that the new observation Y satisfies $Y > 1$. **[4]**

Question 5. [22 marks]

(a) A two stage linear model is given by

$$\begin{aligned} \underline{y}|\underline{\theta}_1 &\sim N(A_1\underline{\theta}_1, C_1) \\ \underline{\theta}_1|\underline{\mu} &\sim N(\underline{\mu}, C_2) \end{aligned}$$

where $\underline{y}$ is a $n \times 1$ vector, $\underline{\theta}_1$ a $p \times 1$ vector and A_1, C_1, C_2, and $\underline{\mu}$ are assumed known.

(i) Find the marginal distribution of $\underline{y}$. **[4]**

(ii) Show that the posterior distribution of $\underline{\theta}_1$ can be written in the form $N(B\underline{b}, B)$ where

$$\begin{aligned} B^{-1} &= A_1^T C_1^{-1} A_1 + C_2^{-1} \\ \underline{b} &= A_1^T C_1^{-1} \underline{y} + C_2^{-1} \underline{\mu} \end{aligned}$$

[4]

(iii) Hence show that the posterior mean is a weighted average of the least squares estimate and the prior mean. Discuss the form of the weights. **[4]**

(b) In a clinical trial to test a new drug which is thought to reduce blood pressure five patients were given a placebo and five patients the new drug. The reduction in blood pressure six hours after giving the drug was recorded as follows

Placebo	3	-2	4	-1	2
Drug	11	7	12	10	14

The reduction in blood pressure may be modelled as $y \sim N(\gamma, 3)$ for patients receiving the placebo and $y \sim N(\gamma + \delta, 3)$ for patients receiving the new drug. Thus δ represents the effect of the drug over any placebo effect. Based on previous studies the prior for γ is $N(2, 1)$ and for δ is $N(6, 6)$ and they are independent.

(i) Show that this problem can be written as a two-stage linear model and find the posterior distributions of γ and δ. **[7]**

(ii) A physician says that he will only use the new drug if the posterior probability that the effect is more than 6 is 0.99 because of the possibility of side effects. Would this study convince the physician to use the drug? **[3]**

Question 6. [18 marks]

(a) It is desired to estimate the posterior means of three parameters, θ, δ and ϕ. The full conditional distributions,

$$p(\theta|\delta, \phi, \text{data}), \quad p(\delta|\phi, \theta, \text{data}), \quad p(\phi|\theta, \delta, \text{data})$$

are known. Explain how the posterior means may be estimated using the Gibbs sampling method. [**6**]

(b) Lifetimes of electric light bulbs in the ith production run are described by a distribution with density

$$p(x|\theta_i) = \theta_i \exp(-\theta_i x) \qquad x > 0, \theta_i > 0.$$

The parameter θ_i varies from run to run depending on the quality of tungsten used for the filaments, and this variability in θ_i is described by a Gamma distribution, $Ga(\alpha, \beta)$. The parameter α is known but β is unknown and is assumed to have a Gamma distribution, $Ga(\gamma, \delta)$. From each of p production runs, n bulbs are chosen at random and have lifetimes x_{ij} for $i = 1, \ldots, p$ and $j = 1, \ldots, n$.

(i) Compute the full conditional distributions necessary for Gibbs sampling and explain how this would proceed. [**8**]

(ii) Explain how you would amend the Gibbs sampling procedure if any bulb with a lifetime of d or more were censored at d. [**4**]

End of Paper—An appendix of 2 pages follows.

Bayesian Statistics – Common Distributions

Discrete Distributions

Distribution	Density	Range of Variates	Mean	Variance
Uniform	$\frac{1}{N}$	$N = 1, 2, \ldots$ $x = 1, 2, \ldots, N$	$\frac{N+1}{2}$	$\frac{N^2-1}{12}$
Bernoulli	$p^x(1-p)^{1-x}$	$0 \le p \le 1,\ x = 0, 1$	p	$p(1-p)$
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$	$0 \le p \le 1,\ n = 1, 2, \ldots$ $x = 0, 1, \ldots n$	np	$np(1-p)$
Poisson	$\frac{\exp(-\lambda)\lambda^x}{x!}$	$\lambda > 0,\ x = 0, 1, 2, \ldots$	λ	λ
Geometric	$p(1-p)^x$	$0 < p \le 1,\ x = 0, 1, 2, \ldots$	$\frac{(1-p)}{p}$	$\frac{(1-p)}{p^2}$
Negative Binomial	$\binom{r+x-1}{x} p^r (1-p)^x$	$0 < p \le 1,\ r > 0$ $x = 0, 1, 2, \ldots$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$

Continuous Distributions

Distribution	Density	Range of Variates	Mean	Variance
Uniform	$\frac{1}{b-a}$	$-\infty < a < b < \infty$ $a < x < b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal $N(\mu, \sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp[\frac{-(x-\mu)^2}{2\sigma^2}]$	$-\infty < \mu < \infty$ $\sigma > 0,\ -\infty < x < \infty$	μ	σ^2
Normal $No(\mu, h)$	$\frac{\sqrt{h}}{\sqrt{2\pi}} \exp[\frac{-h(x-\mu)^2}{2}]$	$-\infty < \mu < \infty$ $h > 0,\ -\infty < x < \infty$	μ	h^{-1}
Exponential	$\lambda \exp(-\lambda x)$	$\lambda > 0,\ x \ge 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gamma (α, β)	$\frac{\beta^\alpha x^{\alpha-1} \exp(-\beta x)}{\Gamma(\alpha)}$	$\beta > 0,\ \alpha > 0,\ x > 0$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$

Distribution	Density	Range of Variates	Mean	Variance
Beta (a, b)	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}$	$a > 0,\, b > 0,\, 0 < x < 1$	$\frac{a}{a+b}$	$\frac{ab}{(a+b+1)(a+b)^2}$
$t_\nu(m, g)$	$\frac{g^{1/2}\Gamma((\nu+1)/2)}{\sqrt{(\nu\pi)}\Gamma(\nu/2)}$	$-\infty < x < \infty$	location m	precision g,
	$\times \left[1 + \frac{g}{\nu}(x-m)^2\right]^{-(\nu+1)/2}$	dof ν		
F_n^m	$\frac{\Gamma[(m+n)/2]}{\Gamma(m/2)\Gamma(n/2)} \left(\frac{m}{n}\right)^{\frac{m}{2}}$	$m, n = 1, 2, \ldots$	$\frac{n}{n-2}$	$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$
	$\times \frac{x^{(m-2)/2}}{[1+(m/n)x]^{(m+n)/2}}$	$x \geq 0$	for $n > 2$	for $n > 4$
χ_k^2	$\frac{1}{\Gamma(k/2)2^{k/2}} x^{k/2-1} \exp(-\frac{x}{2})$	$k = 1, 2, \ldots,\, x > 0$	k	$2k$
Pareto	$\frac{\alpha\beta^\alpha}{x^{\alpha+1}}$	$\alpha > 0, \beta > 0, x > \beta$	$\frac{\beta\alpha}{(\alpha-1)}$	$\frac{\beta^2\alpha}{(\alpha-1)^2(\alpha-2)}$

End of Appendix.