Remote Assessment Rubric/Cover Sheet

University of Bath
Department of Mathematical Sciences
MA40189 Topics in Bayesian statistics

MA40189 Open Book Examination

Assessment available from: 10:30am, 4th June 2020
Latest submission time: 10:30am, 5th June 2020
All timings are given in British Summer Time (BST)

Between these times you must complete and submit your completed assessment. This assessment is designed to take approximately 3 hours to complete.

This is an open book examination. You may refer to your own course and revision notes and look up information in offline or online resources, for example textbooks or online journals. However, you may not communicate with any person or persons about this assessment before the submission deadline unless explicitly permitted to do so in the instructions below. When you submit your assignment, you will be asked to agree to an academic integrity declaration and confirm the work is your own. The use of the work of others, and your own past work, must be referenced appropriately. It is expected that you will have read and understood the Regulations for Students and your programme handbook, including the references to and penalties for unfair practices such as plagiarism, fabrication or falsification.

Which questions should be answered: Answer THREE of the first four questions, 1 - 4, AND question 5. If you submit solutions to all of questions 1 - 4 only the first three of these solutions in your submission will be marked.

Additional materials needed to complete the assessment: You may use the table of distributions given at https://people.bath.ac.uk/masss/ma40189/distributions.pdf. A calculator is not necessary for this assessment.

Any further instructions: In marking, emphasis will be placed upon clear mathematical argument and precision. Do NOT submit your rough work.

Submitting your assessment: When you have completed this assessment, you must submit your work in PDF format as a single file, uploaded to the Moodle Submission Point. You should follow the additional guidance on how to submit your assessment available at: https://teachinghub.bath.ac.uk/teaching-online-options-and-considerations/support-for-students-alternative-assessment/

--
Last update 220420

1. (a) State the definition of a conjugate family for a parameter θ with respect to a likelihood $f(x \mid \theta)$.

(b) Suppose that the likelihood $f(x \mid \theta)$ belongs to the k-parameter regular exponential family. Show that a conjugate prior can be found in the $(k+1)$-parameter regular exponential family.

Let $X_1, \ldots, X_n$ be exchangeable so that the X_i are conditionally independent given a parameter θ. Let $\theta = (\mu, \nu)$ and suppose that each $X_i \mid \theta \sim N(\mu, \nu)$.

(c) Show that

$$f(x \mid \theta) \;=\; (2\pi)^{-n/2}\nu^{-n/2} \exp\left\{-\frac{1}{2\nu}\left[(n-1)s^2 + n(\overline{x}-\mu)^2\right]\right\}$$

where $x = (x_1, \ldots, x_n)$ and $\overline{x}$, $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2$ are respectively the sample mean and variance. Hence, or otherwise, explain why $\overline{X}$ and S^2 are sufficient for $X = (X_1, \ldots, X_n)$ for learning about θ.

(d) It is judged that the improper joint prior distribution $f(\mu, \nu) \propto \nu^{-1}$ is appropriate.

(i) By first deriving the posterior distribution $\mu, \nu \mid x$ where $x = (x_1, \ldots, x_n)$, or otherwise, show that $\nu \mid x \sim \text{Inv-gamma}((n-1)/2, (n-1)s^2/2)$.

(ii) Let $Y = (n-1)s^2/\nu$. Using the result of part (d)(i), or otherwise, show that $Y \mid x \sim \chi^2_{n-1}$, the chi-squared distribution with $n-1$ degrees of freedom. Comment on why this result suggests that the prior $f(\mu, \nu) \propto \nu^{-1}$ may be viewed as noninformative.
(Hint: You may use, without proof, the property that if $A \sim \text{Gamma}(a, b)$ then $cA \sim \text{Gamma}(a, b/c)$ for any constant $c > 0$. You should clearly state any further properties of the Gamma distribution you use.)

(e) Suppose now that the prior for θ is given hierarchically and it is judged that $\mu \mid \nu \sim N(\mu_0, \nu/\lambda)$, where μ_0 and λ are known constants, and $\nu \sim \text{Inv-gamma}(\alpha, \beta)$ for known constants α and β. Show that, with respect to the normal likelihood given in part (c), the prior distribution for $\theta = (\mu, \nu)$ is a conjugate prior.
(Hint: You may use, without proof, the result that for all μ,

$$a(\mu-b)^2 + c(\mu-d)^2 \;=\; (a+c)\left(\mu - \frac{ab+cd}{a+c}\right)^2 + \left(\frac{ac}{a+c}\right)(b-d)^2$$

for any $a, b, c, d \in \mathbb{R}$ with $a \neq -c$.)

[20]

2. (a) Let $X_1, \ldots, X_n$ be exchangeable so that the X_i are conditionally independent given a parameter θ. Suppose that $X_i \,|\, \theta \sim \text{Exp}(\theta)$ so that $E(X_i \,|\, \theta) = 1/\theta$.

(i) Find the Jeffreys prior for θ and show that the posterior distribution for θ given $x = (x_1, \ldots, x_n)$ is Gamma(α_n, β_n), stating the values of α_n and β_n.

(ii) By considering the posterior mean, or otherwise, briefly explain how, in this case, the Jeffreys prior can be viewed as noninformative.

(iii) Consider a further observation Z which is exchangeable with $X = (X_1, \ldots, X_n)$ where $n > 2$. Without calculating the predictive distribution of Z given $X = x$, find $E(Z \,|\, X)$ and $Var(Z \,|\, X)$.

(iv) Suppose now that the prior for θ is instead given by the probability density function

$$f(\theta) \;=\; \frac{\beta^\alpha}{4\Gamma(\alpha)}\theta^{\alpha-1}\exp(-\beta\theta) + \frac{3\beta^{\alpha+1}}{4\Gamma(\alpha+1)}\theta^{\alpha}\exp(-\beta\theta).$$

for known constants $\alpha, \beta > 0$. Show that the posterior probability density function can be written as

$$f(\theta \,|\, x) \;=\; cf_1(\theta) + (1-c)f_2(\theta)$$

where

$$c \;=\; \frac{\alpha\left(\beta + \sum_{i=1}^n x_i\right)}{\alpha\left(\beta + \sum_{i=1}^n x_i\right) + 3\beta\left(\alpha + n\right)}$$

and $f_1(\theta)$ and $f_2(\theta)$ are probability density functions.

(b) (i) State the definition of finite exchangeability.

(ii) Let $X_1, \ldots, X_n$ be conditionally independent given a parameter $\theta = (\theta_1, \ldots, \theta_n)$ and suppose that each $X_i \,|\, \theta$ is from the same family of distributions with likelihood $f(x_i \,|\, \theta) = f(x_i \,|\, \theta_i)$. It is judged that $\theta_1, \ldots, \theta_n$ are independent and identically distributed with each θ_i having probability density function $\pi(\theta_i)$. Show that $X_1, \ldots, X_n$ are finitely exchangeable.

[20]

3. Let $X_1, \ldots, X_n$ be conditionally independent given a parameter $\theta = (\theta_1, \ldots, \theta_n)$ and suppose that $X_i \,|\, \theta \sim \text{Po}(s_i\theta_i)$ where each s_i is known. It is judged that $\theta_1, \ldots, \theta_n$ are conditionally independent given ϕ with $\theta_i \,|\, \phi \sim \text{Exp}(\phi)$, so that $E(\theta_i \,|\, \phi) = 1/\phi$, and $\phi \sim \text{Gamma}(\alpha, \beta)$, where α and β are known.

(a) Show that the posterior density $f(\theta_1, \ldots, \theta_n, \phi \,|\, x)$, where $x = (x_1, \ldots, x_n)$, can be expressed as

$$f(\theta_1, \ldots, \theta_n, \phi \,|\, x) \quad \propto \quad \left(\prod_{i=1}^{n} \theta_i^{x_i}\right) \phi^{\alpha+n-1} \exp\left\{-\phi\left(\beta + \sum_{i=1}^{n} \theta_i\right) - \sum_{i=1}^{n} s_i\theta_i\right\}.$$

(b) Describe how to use the Gibbs sampler to sample from the posterior distribution of $\theta_1, \ldots, \theta_n, \phi \,|\, x$, deriving any required conditional distributions.

Consider a general problem in which we will observe data X, where the distribution of X depends upon an unknown parameter θ, and we wish to make inference about θ.

(c) We wish to use the Metropolis-Hastings algorithm to sample from the posterior distribution of $\theta \,|\, x$ with density $\pi(\theta) = f(\theta \,|\, x)$. Letting $q(\phi \,|\, \theta)$ denote the proposal distribution when the current state is θ and $\alpha(\theta, \phi)$ the probability of accepting a move from θ to ϕ, describe how the algorithm works in this case.

(d) Suppose that you can sample from a distribution $q(\theta)$ which is an approximation to the posterior distribution $f(\theta \,|\, x)$. Explain how to use the method of importance sampling to estimate $E(g(\theta) \,|\, X)$ for some function $g(\theta)$.

(e) Let $\hat{I}$ denote the importance sampling estimator of $E(g(\theta) \,|\, X)$ constructed in part (d). By considering the variance of $\hat{I}$ with respect to the distribution $q(\theta)$, or otherwise, explain why a sensible choice of $q(\theta)$ is that which minimises $E\left(\frac{g^2(\theta) f(\theta \,|\, x)}{q(\theta)} \,\middle|\, X\right)$.

[20]

4. Consider a statistical decision problem $[\Theta, \mathcal{D}, \pi(\theta), L(\theta, d)]$ for a univariate parameter θ and loss function

$$L(\theta, d) \quad = \quad g(\theta)(\theta - d)^2$$

where $g(\theta) > 0$ and d is a point estimate of θ.

(a) Show that the Bayes decision rule against $\pi(\theta)$ is

$$d^* \quad = \quad \frac{E(g(\theta)\theta)}{E(g(\theta))}$$

with corresponding Bayes risk

$$\rho^*(\pi) \quad = \quad E(g(\theta)\theta^2) - \frac{\{E(g(\theta)\theta)\}^2}{E(g(\theta))},$$

interpreting the results for the case when $g(\theta) = 1$.

Let $X_1, \ldots, X_n$ be exchangeable so that the X_i are conditionally independent given a parameter θ. Suppose that $X_i \,|\, \theta$ is distributed as a Weibull distribution with known shape parameter $\lambda > 0$ and unknown scale parameter $\theta > 0$, denoted $\text{WE}(\lambda, \theta)$, with probability density function

$$f(x_i \,|\, \theta) \quad = \quad \begin{cases} \theta \lambda x_i^{\lambda - 1} \exp\left(-\theta x_i^{\lambda}\right) & 0 < x_i < \infty \\ 0 & \text{otherwise,} \end{cases}$$

and $\theta \sim \text{Gamma}(\alpha, \beta)$ where $\alpha > 3$ and β are known. We wish to produce a point estimate d for θ, with loss function

$$L(\theta, d) \quad = \quad \frac{(\theta - d)^2}{\theta^3}.$$

(b) Find the Bayes rule and Bayes risk of an immediate decision.

(c) Find the Bayes rule and Bayes risk after observing $x = (x_1, \ldots, x_n)$.

(d) Suppose that $\lambda = 1$ so that $X_i \,|\, \theta \sim \text{Exp}(\theta)$. If each observation costs a fixed amount c then R_n, the total risk of a sample of size n, is the sum of the sample cost and the Bayes risk of the sampling procedure. Find R_n and thus the optimal choice of n.

Let $\theta \in \mathbb{R}$ be a continuous univariate random variable with finite expectation. We wish to construct an interval estimate $d = (\theta_1, \theta_2)$, where $\theta_1 \leq \theta_2$, for θ. Let $\mathcal{D}_1$ denote the set of all such intervals d. Consider the loss function

$$L_1(\theta, d) \quad = \quad \theta_2 - \theta_1 + \begin{cases} \frac{2}{\alpha}(\theta_1 - \theta) & \text{if } \theta < \theta_1 \\ 0 & \text{if } \theta_1 \leq \theta < \theta_2 \\ \frac{2}{\alpha}(\theta - \theta_2) & \text{if } \theta_2 < \theta. \end{cases}$$

(e) By considering derivatives of the expected loss with respect to θ_1 and θ_2, or otherwise, show that, for the decision problem $[\Theta, \mathcal{D}_1, \pi(\theta), L_1(\theta, d)]$, the Bayes rule is $d^* = (\theta_1^*, \theta_2^*)$ where $P(\theta \leq \theta_1^*) = \alpha/2 = P(\theta \geq \theta_2^*)$.

[20]

5. Discuss, in approximately 500 words, the advantages and the disadvantages of the Bayesian approach to statistics. Your answer should include references to results and ideas explored in the course. [15]