Queen Mary University of London

Main Examination period 2020 – May/June – Semester B
Online Alternative Assessments

MTH709U / MTH776P: Bayesian Statistics

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

In completing this assessment, you may use books, notes, and the internet. You may use calculators and computers, but you should show your working for any calculations you do. You must not seek or obtain help from anyone else.

At the start of your work, please **copy out and sign** the following declaration:

> I declare that my submission is entirely my own, and I have not sought or obtained help from anyone else.

All work should be **handwritten**, and should **include your student number**.

You have **24 hours** in which to complete and submit this assessment. When you have finished your work:

- scan your work, convert it to a **single PDF file** and upload this using the upload tool on the QMplus page for the module;
- e-mail a copy to maths@qmul.ac.uk with your student number and the module code in the subject line;
- with your e-mail, include a photograph of the first page of your work together with either yourself or your student ID card.

You are not expected to spend a long time working on this assessment. We expect you to spend about **3 hours** to complete the assessment, plus the time taken to scan and upload your work. Please try to upload your work well before the end of the assessment period, in case you experience computer problems. **Only one attempt is allowed – once you have submitted your work, it is final**.

Examiners: L I Pettit, J Griffin

A table of common distributions is provided as an appendix.

Question 1 [14 marks].

(a) Explain what is meant by a **conjugate prior**. [**2**]

(b) Give one advantage and one disadvantage of using a conjugate prior. [**2**]

(c) State a conjugate prior distribution for each of the following data distributions.

(i) A normal distribution with known mean and unknown precision. [**1**]

(ii) A continuous uniform distribution on the interval $(0, \theta)$. [**1**]

(d) Show why ONE of your answers in (c) is correct. [**2**]

(e) A random variable X has a distribution denoted $H(a, \theta)$ 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. Assume that a is known and consider $X_1, \ldots, X_n$ independent observations from an $H(a, \theta)$ distribution.

Suppose that an $H(\alpha, \beta)$ prior is assumed for the parameter θ. Show that the posterior distribution of θ given $X_1, \ldots, X_n$ has an $H(\alpha^\star, \beta^\star)$ distribution and find expressions for $\alpha^\star$ and $\beta^\star$. [**6**]

Question 2 [17 marks].

Three scientists are interested in the decay rate of a rare isotope. They do a series of 20 independent experiments giving observations $X_1, X_2, \ldots, X_{20}$ with each X_i having a Poisson distribution with mean θ. They find $\sum x_i = 30$.

(a) The first scientist has little prior knowledge about θ and decides to use the Jeffreys prior. Derive the form of this prior and find the posterior distribution. [**5**]

(b) The second scientist has more experience and uses a prior which is Gamma$(4, 2)$. Find the posterior distribution for the second scientist. [**2**]

(c) One way to find an interval estimate would be to approximate the posterior distribution by a normal distribution with mean the posterior mode and variance the reciprocal of the observed information evaluated at the posterior mode. Find the parameters of this normal distribution and hence an approximate 95% highest posterior density interval for the second scientist. [**4**]

(d) The third scientist decided a mixture prior would be better and chooses

$$p(\theta) = 0.75Ga(4, 2) + 0.25Ga(1, 1).$$

Find the posterior for the third scientist. [**6**]

Question 3 [23 marks].

(a) Show that an exponential distribution with mean θ^{-1} is an exponential family. [**3**]

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

(c) If a random sample of n exponentially distributed observations is collected write down a minimal sufficient statistic for θ. [**1**]

(d) Six batteries are put on test for 200 hours. Four fail after 96, 130, 160 and 180 hours. The other two batteries are still working after 200 hours. Assuming the lifetime of these batteries has an exponential distribution with mean θ^{-1} find the likelihood function. [**3**]

(e) The prior distribution of θ is given by a gamma distribution with density

$$p(\theta) = \frac{\beta^\alpha \theta^{\alpha-1} \exp(-\beta\theta)}{\Gamma(\alpha)}, \quad \theta \geq 0.$$

It is thought the lifetime of the batteries has mean 180 and the probability it lies between 110 and 470 is approximately 0.95. Show by finding an approximate 95% probability interval for θ that a prior with $\alpha = 10$ and $\beta = 1800$ is appropriate. [**4**]

(f) Find the posterior distribution of θ using the prior from (e). [**2**]

(g) Find the Bayes estimate under squared error loss and the expected loss. [**2**]

(h) Find the density of the predictive distribution of another independent battery and find the probability it has a lifetime more than 200 hours. [**4**]

(i) Comment on the use of the exponential distribution as a lifetime distribution. [**3**]

Question 4 [17 marks].

(a) You want to investigate the proportion (θ) of defective items manufactured at a production line. An inspector is told to take a random sample of 30 items and finds 3 are defective. Assume a uniform prior for θ on the unit interval $(0, 1)$. Compute the posterior of θ. [**3**]

(b) The inspector now tells you that he did not decide on the sample size before the sampling was performed. His sampling plan was to keep on sampling items until he had found three defective ones. It happened that the 30th item was the third one to be defective. Redo the posterior calculation, this time under the new sampling scheme. Discuss the results. [**4**]

(c) Find the form of the Bayes factor for testing the hypothesis $H_0 : \theta = 0.05$ against $H_1 : \theta \sim Be(\alpha, \beta)$. [**4**]

(d) Evaluate the Bayes factor and discuss the conclusion when [**6**]

(i) $\alpha = \beta = 1$.

(ii) $\alpha = 1$ and $\beta = 2 + sn$ where sn is the sixth digit of your student number.

(iii) $\alpha = 10$ and $\beta = 100$.

Question 5 [15 marks].
The two-stage linear model is given by

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

where A_1, C_1, C_2 are known matrices of appropriate size and $\underline{\mu}$ is a known vector.
The posterior distribution of $\underline{\theta}_1$ is $N(B\underline{b}, B)$ where

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

The yields of chemical in two different processes were recorded at 5 different temperatures $40°, 50°, 60°, 70°, 80°$. It is expected that the effect of temperature will be the same in both processes so a parallel regressions model is adopted

$$y_{ij} = \alpha_i + \beta(x_j - \bar{x}) + \varepsilon_{ij} \qquad i = 1, 2 \; j = 1, \ldots, 5$$

where ε_{ij} are assumed to be independent and distributed as $N(0, 3)$. The priors for the unknown parameters are taken as $\alpha_1 \sim N(40, 1)$, $\alpha_2 \sim N(50, 1)$ and $\beta \sim N(0.5, 0.1)$ and all parameters are assumed to be independent. The data are given in the table below.

Temperature	Process 1 yield	Process 2 yield
40°	28.7	37.2
50°	38.2	44.6
60°	42.5	54.9
70°	43.4	60.5
80°	50.2	62.0

(a) Write this model as a two stage linear model. [**5**]

(b) Find the posterior distributions of α_1, α_2 and β. [**6**]

(c) Find the posterior distribution of $\alpha_2 - \alpha_1$. [**2**]

(d) Explain briefly how you would approach this problem if the variance of ε_{ij} was not assumed to be known. [**2**]

Question 6 [14 marks].
The inverted Pareto distribution has p.d.f.

$$p(x|\theta, \delta) = \theta\delta^{\theta}x^{\theta-1} \quad 0 < x < \delta^{-1}, \;\; \theta > 0, \delta > 0$$

and zero otherwise. We will denote this by $IP(\theta, \delta)$.
Suppose n observations are available from this distribution. The prior for θ is taken as $Ga(\alpha, \beta)$ and $p(\delta) \propto \delta^{-1}$ and they are assumed independent.

(a) Show that the conditional posterior distribution $p(\theta|\underline{x}, \delta)$ is Gamma and identify the parameters. [**5**]

(b) Show that the conditional posterior distribution $p(\delta|\underline{x}, \theta)$ is inverted Pareto and identify the parameters. [**5**]

(c) Assuming that any values can be simulated from required distributions, explain how Gibbs sampling would proceed. [**4**]

End of Paper – An appendix of 1 page follows.

Bayesian Statistics – Common Distributions

Discrete Distributions

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

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)	$\frac{\beta^\alpha x^{\alpha-1}\exp(-\beta x)}{\Gamma(\alpha)}$	$\beta > 0,\ \alpha > 0,\ x > 0$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$
Beta (a, b)	$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}x^{a-1}(1-x)^{b-1}$	$a > 0,\ b > 0,\ 0 < x < 1$	$\frac{a}{a+b}$	$\frac{ab}{(a+b+1)(a+b)^2}$
F_n^m	$\frac{\Gamma[(m+n)/2]}{\Gamma(m/2)\Gamma(n/2)}\left(\frac{m}{n}\right)^{\frac{m}{2}}$	$m, n = 1, 2, \dots$	$\frac{n}{n-2}$	$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$
	$\times\frac{x^{(m-2)/2}}{[1+(m/n)x]^{(m+n)/2}}$	$x \ge 0$	for $n > 2$	for $n > 4$
χ_k^2	$\frac{1}{\Gamma(k/2)2^{k/2}}x^{k/2-1}\exp(-\frac{x}{2})$	$k = 1, 2, \dots,\ x > 0$	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.