Queen Mary
University of London

Main Examination period 2018

MTH6909: Bayesian Statistics

Duration: 2 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, J Griffin

Question 1. [23 marks] A biased coin, which has probability θ of landing heads, is tossed repeatedly until the first head is seen. We call this one trial.

(a) Show that the total number of tails, x, in one trial has a geometric distribution. **[3]**

(b) If a Beta $Be(\alpha, \beta)$ distribution is chosen as the prior distribution find the posterior distribution if there are n trials. **[5]**

(c) If a $Be(3, 3)$ prior is chosen and there are a total of 20 tosses in 5 trials find the posterior mean of θ and a 95% highest posterior density interval. **[7]**

(d) Show that the Jeffreys' prior is proportional to $1/\theta\sqrt{(1-\theta)}$ and hence find the corresponding posterior distribution. **[8]**

Question 2. [20 marks]

(a) Show that if $t(\underline{x})$ is sufficient for the family $p(\underline{x}|\theta)$ then for any prior distribution the posterior distributions given $\underline{x}$ and $t(\underline{x})$ are the same. **[4]**

(b) The observations $x_1, x_2, \ldots, x_n$ are a random sample from a uniform distribution on the interval $(0, \theta)$. The prior distribution for θ is Pareto with density

$$p(\theta) = \frac{\alpha\theta_0^\alpha}{\theta^{\alpha+1}} \qquad \theta \geq \theta_0$$

and zero otherwise.

(i) Find the posterior distribution of θ. **[5]**

(ii) Deduce the sufficient statistic for θ. **[2]**

(iii) Find the Bayes estimate of θ under the loss function

$$L(t, \theta) = \frac{(t-\theta)^2}{t},$$

where t is any estimate of θ. **[9]**

Question 3. [22 marks]

(a) Show that the Poisson distribution is an exponential family, identifying all the necessary functions. [**5**]

(b) Show that a Gamma $Ga(\alpha, \beta)$ distribution gives a conjugate prior distribution. [**3**]

(c) A radioactive device gives hourly counts which are assumed to have a Poisson distribution with mean θ. The device is observed for three hours and emits x particles. Two models have been suggested. Firstly that the emissions will have a Poisson distribution with $\theta = 2$. Secondly that the emissions will have a Poisson distribution with θ unknown but given a $Ga(2, 1)$ prior.

(i) Show that the first model is preferred if

$$\frac{2^{3x+4}}{(x+1)!} > e^6.$$

[**9**]

(ii) Hence show that if $x = 10$ the first model is preferred but if $x \geq 11$ the second model is. [**5**]

Question 4. [8 marks] The Laplace approximation to the integral

$$\int \exp[-nh(\theta)]d\theta$$

over the range of θ, is given by

$$\exp[-nh(\hat{\theta})]\sqrt{2\pi}\hat{\sigma}/\sqrt{n}$$

where $\hat{\theta}$ maximises $-h(\theta)$ and $\hat{\sigma} = [h''(\theta)]^{-\frac{1}{2}} \mid_{\theta=\hat{\theta}}$.
Explain how to approximate the posterior mean and variance of a distribution using the Laplace approximation. Why are these approximations often very accurate? [**8**]

Question 5. [27 marks]
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.

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

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

[6]

(b) The yields of chemical in two different processes were recorded at 5 different temperatures $50°, 60°, 70°, 80°, 90°$. 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 $No(0, \xi)$, where the precision ξ is known.. The priors for the unknown parameters are taken as $\alpha_1 \sim No(24, \xi)$, $\alpha_2 \sim No(28, \xi)$ and $\beta \sim No(0.5, 5\xi)$ and all parameters are assumed to be independent. The data are given in the table below.

Temperature	Process 1 yield	Process 2 yield
50°	14.6	18.9
60°	18.8	23.5
70°	23.7	29.7
80°	27.9	31.5
90°	33.4	37.4

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

(ii) Find the posterior distributions of α_1, α_2 and β. **[7]**

(c) Suppose now that ξ is unknown and given a $Ga(2, 4)$ distribution.

(i) Find the conditional distribution of ξ given the other parameters and the data. **[4]**

(ii) Hence explain how Gibbs sampling could be carried out. **[5]**

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	$p^x(1-p)^{1-x}$	$0 \leq p \leq 1,\ x = 0, 1$	p	$p(1-p)$
Binomial	$\binom{n}{x} p^x (1-p)^{n-x}$	$0 \leq p \leq 1,\ n = 1, 2, \ldots$ $x = 0, 1, \ldots n$	np	$np(1-p)$
Poisson	$\frac{\exp(-\lambda)\lambda^x}{x!}$	$\lambda > 0,\ x = 0, 1, 2, \ldots$	λ	λ
Geometric	$p(1-p)^x$	$0 < p \leq 1,\ x = 0, 1, 2, \ldots$	$\frac{(1-p)}{p}$	$\frac{(1-p)}{p^2}$
Negative Binomial	$\binom{r+x-1}{x} p^r (1-p)^x$	$0 < p \leq 1,\ r > 0$ $x = 0, 1, 2, \ldots$	$\frac{r(1-p)}{p}$	$\frac{r(1-p)}{p^2}$

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	$\lambda \exp(-\lambda x)$	$\lambda > 0,\ x \geq 0$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
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^m_n	$\frac{\Gamma[(m+n)/2]}{\Gamma(m/2)\Gamma(n/2)}\left(\frac{m}{n}\right)^{\frac{m}{2}}$	$m,n=1,2,\ldots$	$\frac{n}{n-2}$	$\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$
	$\times\frac{x^{(m-2)/2}}{[1+(m/n)x]^{(m+n)/2}}$	$x\geq 0$	for $n>2$	for $n>4$
χ^2_k	$\frac{1}{\Gamma(k/2)2^{k/2}}x^{k/2-1}\exp(-\frac{x}{2})$	$k=1,2,\ldots,\ x>0$	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.