MATH 559: Bayesian Theory and Methods

Practice Final Exam
3 hours

First name (please write as legibly as possible within the boxes)

Last name

Student ID number

MATH 559: Specific Instructions.

Please write your answers on the exam in the boxes provided; only answers written inside the boxes will be marked. Please write clearly. You may request a regular blank answer booklet (not a Crowdmark booklet) for rough working, but answers written on that booklet will not be marked.

This paper contains five questions. Each question carries 20 marks. Credit will be given for all questions attempted. The total mark available is 100 but rescaling of the final mark may occur.

A formula sheet is provided. No crib sheet is allowed.

Only non-programmable calculators may be used. Dictionaries and Translation dictionaries are permitted.

The usual abbreviations are used: pmf indicates the probability mass function; pdf indicates the probability density function; cdf indicates the cumulative distribution function; i.i.d. indicates independent and identically distributed. The function $\mathbb{1}_A(x)$ is the indicator function for set A that takes the value 1 if $x \in A$ and zero otherwise.

This page intentionally left blank.

1. Random variables $Y_1, Y_2, \ldots$ are infinitely exchangeable.

    (a) Suppose for each i, Y_i is binary.

(i) Write down the de Finetti representation for the joint pmf for $Y_1, \ldots, Y_n$, $n \geq 1$ explaining the meaning of each of the terms, including the definition of the *parameter*, denoted θ, its *true value*, denoted θ_0, and the *prior distribution* π_0. 6 MARKS

Q1(a)(i):

(ii) Suppose that the prior distribution on θ is $Uniform(0, 1)$. Find the *prior predictive distribution* for Y_1.

3 MARKS

Q1(a)(ii):

(iii) Find the *posterior predictive distribution* for Y_2 given $Y_1 = y_1$ under the $Uniform(0, 1)$ prior.
3 MARKS

Q1(a)(iii):

(b) Suppose that data $y_1, \ldots, y_n$ are observed values of random variables that are assumed to be conditionally independent $Poisson(\theta)$ random variables. Show that the posterior distribution for θ depends only on a summary, s_n, of the observed values, and identify the limiting behaviour of this posterior as $n \longrightarrow \infty$ if the data generating model is $Poisson(\theta_0)$ with $\theta_0 = 1$. 8 MARKS

Q1(b):

2. Random variables $Y_1, \ldots, Y_n$ are presumed conditionally independent with pdf

$$f_Y(y;\theta) = \mathbb{1}_{(\theta,\infty)}(y)\lambda\exp\{-\lambda(y-\theta)\} \qquad y \in R$$

where $\theta \in \Theta \equiv (-1,1)$, and $\lambda > 0$.

(a) Find the form of the posterior distribution for θ if λ is a fixed constant, $\lambda = \lambda_0$, and the prior on θ is a Uniform distribution. 6 MARKS

Q2(a):

(b) Find the form of the *joint posterior* for (θ, λ) if θ and λ have independent prior distributions, if the prior on θ is Uniform and the prior on λ is $Gamma(2, 1)$. 6 MARKS

Q2(b):

(c) From the joint posterior in (b), find the marginal posteriors for θ and λ. 8 MARKS

Q2(c):

3. Suppose that $Y_1, \ldots, Y_n$ are presumed to be conditionally independent and distributed as $Normal(\mu, \sigma^2)$. A sample of data $y_1, \ldots, y_n$ are observed.

   (a) Derive the joint posterior for μ and σ^2 under a conjugate prior specification. 10 MARKS

Q3(a):

(b) For the posterior in (a), derive the Bayesian estimate of μ, $\widehat{\mu}_{nB}$ under *quadratic loss*, defined by

$$\widehat{\mu}_{nB} = \arg\min_{t} \int_{0}^{\infty} \int_{-\infty}^{\infty} L(t,\mu)\pi_n(\mu,\sigma^2)\, d\mu\, d\sigma^2$$

where $L(t,\mu) = (t-\mu)^2$. 5 MARKS

Q3(b):

(c) Suppose that the prior for σ^2 is *degenerate* at $\sigma^2 = 1$. Give an expression for a 95 % credible interval for μ in terms of the standard Normal cdf, $\Phi(\cdot)$ and its inverse. 5 MARKS

Q3(c):

4. Suppose that $Y_1, \ldots, Y_n$ are conditionally independent discrete random variables with conditional mass function

$$f_Y(y; \boldsymbol{\theta}) = \theta_y \qquad y = 1, 2, \ldots, N$$

for integer $N > 0$, and zero otherwise, where for each y, $0 \leq \theta_y \leq 1$, and

$$\sum_{y=1}^{N} \theta_y = 1.$$

A sample of data $y_1, \ldots, y_n$ is collected.

(a) Write down the likelihood for this model. 4 MARKS

Q4(a):

(b) Compute the joint posterior for $(\theta_1, \ldots, \theta_N)$ under a *conjugate prior*. 6 MARKS

Q4(b):

(c) In a reparameterization of the model, suppose that $\phi_1 = \theta_1$ and, for $k = 2, \ldots, N$,

$$\phi_k = \frac{\theta_k}{1 - \theta_1 - \cdots - \theta_{k-1}}$$

that is

$$\phi_2 = \frac{\theta_2}{1 - \theta_1} \qquad \phi_3 = \frac{\theta_3}{1 - \theta_1 - \theta_2}$$

etc. Find a conjugate prior for the new parameterization, and the posterior under this prior.
10 MARKS

Q4(c):

5. (a) Consider the pdf

$$f(x) = c\mathbb{1}_{(0,1)}(x)x(1-x)(1-\sqrt{x})^2 \quad x \in \mathbb{R}$$

for some constant $c > 0$. It is proposed to estimate the expectation $\mathbb{E}_f[X]$ using Monte Carlo.

(i) Give details of the Monte Carlo procedure that uses *importance sampling* based on the proposal pdf $f_0(x) \equiv Uniform(0,1)$. 3 MARKS

Q5(a)(i):

(ii) Give details of a Monte Carlo procedure that uses *importance sampling* based on a proposal $f_0(x)$ designed to produce an estimator whose variance is smaller than the variance of the Monte Carlo estimator from (a) (you do not have to verify that the variance is smaller). 4 MARKS

Q5(a)(ii):

(iii) Describe how to carry out *rejection sampling* from $f(x)$ based on proposals from a $Beta(2, 2)$ distribution. 3 MARKS

Q5(a)(iii):

(b) Consider a linear regression model for observations $y_1, \ldots, y_n$ that are considered conditionally independent with

$$Y_i | x_i; \beta_0, \beta_1, \sigma^2 \sim Normal(\beta_0 + \beta_1 x_i, \sigma^2)$$

for $i = 1, \ldots, n$. In a Bayesian analysis, an *independence* prior

$$\pi_0(\beta_0, \beta_1, \sigma^2) = \pi_0(\beta_0)\pi_0(\beta_1)\pi_0(\sigma^2)$$

is specified, where

$$\pi_0(\beta_0) \equiv \pi_0(\beta_1) \equiv Normal(0, M_0) \qquad \pi_0(\sigma^2) \equiv InvGamma(a_0/2, b_0/2)$$

for hyperparameters M_0, a_0, b_0.

(i) Describe how to implement a *Gibbs sampler* for this model, giving details of the full conditional posterior distributions. 6 MARKS

Q5(b)(i):

Q5(b)(i) continued:

(ii) Suppose for $k = 1, 2$, independent priors

$$\pi_0(\beta_k) = \frac{1}{2}\exp\{-|\beta_k|\} \quad \beta_k \in \mathbb{R}$$

are used. Describe how to implement a *Metropolis-within-Gibbs* MCMC algorithm for this model.

4 MARKS

Q5(b)(iii):

Additional space for answers or rough working

Additional space for answers or rough working

Additional space for answers or rough working

Additional space for answers or rough working