THE UNIVERSITY OF NEW SOUTH WALES

MATH3871/MATH5960
Assignment 1

Assignment 1

This assignment covers material in Lectures 1–3 and is worth 20% of the final course grade. **Please refer to the following instructions:**

- Assignment to be submitted via by 13 October 11:55PM AEDT
- Include in your assignment, any relevant R code, R output, and mathematical derivations. Embed the code and plots into your assignment (please don't attach R markdown or other R script files) separately.
- The total number of submitted pages should not exceed 6 A4 pages. Any pages submitted in excess of 6 pages will not be graded.
- Print, sign and attach this cover sheet with your assignment (not included in page count).
- Refer to course handout for grading of late submissions

Plagiarism Statement

I declare that this assessment item is my own work, except where acknowledged, and has not been submitted for academic credit elsewhere. I acknowledge that the assessor of this item may, for the purpose of assessing this item reproduce this assessment item and provide a copy to another member of UNSW; and/or communicate a copy of this assessment item to a plagiarism checking service (which may then retain a copy of the assessment item on its database for the purpose of future plagiarism checking).

I certify that I have read and understood UNSW Rules in respect of Student Academic Misconduct.

Name (print clearly):

Student Number:

Signature:

Date:

1. Researchers interested in understanding peoples' level of life satisfication conducted a simple survey. 1278 people were randomly selected from the general population in June 2019 (before COVID-19 lockdowns) and a *different* group of 1278 people were selected in June 2023 (after COVID-19 lockdowns). Both groups were asked to report their level of life satisfication using the categories "good", "bad" and "neutral". The resulting data is given in the table below.

Survey	good	bad	neutral	total
before COVID-19	588	614	76	1278
after COVID-19	576	664	38	1278

Assuming the two surveys are independent ranom samples, we can model the data with two different multinomial distributions with parameters $\{\theta_1^j, \theta_2^j, \theta_3^j\}$, where the superscript j denotes the jth survey, $j = 1, 2$. Let α^j denote the proportion of respondents who reported "good" out of those who had a strong opinion (i.e. "good" or "bad"). That is, $\alpha^j = \frac{\theta_1^j}{\theta_1^j + \theta_2^j}$. Let $\beta^j = \theta_1^j + \theta_2^j$ denote the proportion of people with a strong opinion.

   (a) **[5 marks]** Assume a Dirichlet prior on the multinomial parameters, i.e. $\pi(\theta_1^j, \theta_2^j, \theta_3^j) = \text{Dirichlet}(a_1, a_2, a_3)$ (note the same prior is used for both datasets). Analytically calculate the joint posterior $\pi(\alpha^j, \beta^j | y^j)$. Give your answer in abstract form (without substituting values from the table). *Hint: use the change of variables formula.*

   (b) **[2 marks]** Using your result in (a), give the marginal posterior $\pi(\alpha^j | y^j)$.

   (c) **[5 marks]** Select values of a_1, a_2 corresponding to an uninformative prior and then using the data in the table, generate samples from the posterior density for $\alpha^1 - \alpha^2$ in R and plot the histogram. Provide the code you have used to generate the histogram in your report.

   (d) **[2 marks]** Using the results from (c), estimate the posterior probability that there was an increase from before COVID to after COVID in the amount of people with a "good" attitude within the people with a strong opinion.

2. Suppose you are given a density $f(x)$, $x \in \mathbb{R}$ from which you need to generate samples. Let the unnormalised form of $f(x)$ be $\tilde{f}(x)$, i.e. $f(x) \propto \tilde{f}(x)$, and you know that

$$l(x) \leq \tilde{f}(x) \leq K\tilde{g}(x)$$

where $l(x)$ is a non-negative function, $K > 0$ and $g(x) = \frac{\tilde{g}(x)}{Z_g}$ is a density that we can easily sample from. A modified version of the standard rejection sampling method proceeds as follows:

   - Step 1. Generate independent random draw X from $g(x)$ and U from $U(0, 1)$.
   - Step 2. Accept X if $U \leq \frac{l(X)}{K\tilde{g}(X)}$
   - Step 3. If X was rejected, draw $V \sim U(0, 1)$ and accept X if

$$V \leq \frac{\tilde{f}(X) - l(X)}{K\tilde{g}(X) - l(X)}$$

(a) **[5 marks]** Show that the probability of accepting a proposed X in step 2 or 3 is

$$\frac{\tilde{f}(X)}{K\tilde{g}(X)}$$

In other words, calculate the probability that the proposed value is accepted in a single iteration of the algorithm.

(b) **[2 marks]** Show that the probability that step (3) must be carried out is

$$1 - \frac{\int l(x)dx}{KZ_g}$$

(c) **[7 marks]** Implement the above algorithm in R for $\tilde{f}(x) = \exp\left(-\frac{(x-4)^2}{2}\right)$ and $\tilde{g}(x) = \exp(-|x-4|)$ by choosing an appropriate $l(x)$ and K. Plot a histogram of your generated samples and overlay the target density f as a line. Provide the code you have used to generate the histogram in your report.

(d) **[2 marks]** Explain why this algorithm would be beneficial compared to a standard rejection sampling algorithm using the result in (b).

3. Suppose you have data x coming from a $N(\mu, \sigma^2)$ distribution where the variance σ^2 is known but you want to perform Bayesian inference on the mean μ. We are going to consider two different scenarios of prior knowledge:

- The first scenario involves using a normal prior i.e. $\pi(\mu) = N(\nu, w)$ and letting $w \to \infty$ to represent lack of knowledge.

- In the second scenario, you know that there is probability p that $\mu = 0$, but there is little prior knowledge about $\mu \neq 0$. This prior information will be represented by a mixture distribution, with a discrete probability at $\mu = 0$ (i.e. a probability atom) and $\pi(\mu) = N(0, w)$ for $\mu \neq 0$ and let $w \to \infty$ to represent our lack of knowledge for $\mu \neq 0$.

(a) **[6 marks]** Analytically calculate the posterior probability $\pi(\mu|x)$ for the first scenario. Comment on whether the prior and posterior are proper or improper.

(b) **[5 marks]** Analytically calculate the posterior probability $\pi(\mu = 0|x)$ for the second scenario. Give an explanation in words (1-2 sentences) of what this posterior indicates and whether this is sensible.

(c) **[3 marks]** Give an explanation as to what properties of the prior in (b) give rise to the behaviour of the posterior in (b), compared to the prior in (a).

(d) **[2 marks]** Suggest an alternative prior in the spirit of (b) and show that it gives a more sensible posterior $\pi(\mu = 0|x)$ than the one calculated in (b).