STAT 431 — Applied Bayesian Analysis — Fall 2022

Exam

October 13, 2022

Full Name: Key ID/Email: ________________

- This is a 75 minute exam. There are 6 problems for everyone, and one additional problem for the graduate section only.
- The exam is worth a total of 45 points for the undergraduate section and 55 points for the graduate section.
- You may use your own personal notes and a standard scientific calculator. (You may *not* share these items with anyone else.) No other aids or devices are permitted!
- *Write all answers in the spaces provided.* If you require more space to write your answer, you may use the back side of the page.
- You are not allowed to communicate with anyone except the instructor or proctors before you submit this exam.

PDF	=	probability density function
PMF	=	probability mass function
MAP	=	maximum a posteriori
HPD	=	highest posterior density
PPD	=	posterior predictive distribution

1. The residential populations of Champaign, Urbana, and Savoy are approximately 90,000, 38,000, and 9,000, respectively.

   (a) What is the probability that a randomly selected resident of Champaign-Urbana-Savoy lives in Champaign? In Urbana? In Savoy? [3 pts]

$$P(C) = \frac{90000}{90000+38000+9000} = \frac{90000}{137000} = \frac{90}{137} \approx 0.657$$

$$P(U) = \frac{38000}{137000} = \frac{38}{137} \approx 0.277$$

$$P(S) = \frac{9000}{137000} = \frac{9}{137} \approx 0.066$$

   (b) The home ownership rate (percentage of residents who live in their own home) is approximately 45% for Champaign, 35% for Urbana, and 55% for Savoy.

      (i) What is the probability that a randomly selected resident from Champaign-Urbana-Savoy (taken together) lives in their own home? [4 pts]

$$P(H) = P(H|C)\,P(C) + P(H|U)\,P(U) + P(H|S)\,P(S)$$

$$= 0.45 \cdot \frac{90}{137} + 0.35 \cdot \frac{38}{137} + 0.55 \cdot \frac{9}{137}$$

$$\approx 0.429$$

      (ii) Consider a randomly selected resident from Champaign-Urbana-Savoy (taken together) who lives in their own home. What is the probability that the resident lives in Savoy? [3 pts]

$$P(S|H) = \frac{P(H|S)\,P(S)}{P(H)}$$

$$\approx \frac{0.55 \cdot \frac{9}{137}}{0.429} \approx 0.084$$

2. In a Bayesian statistical model for data $Y = y$, continuous scalar parameter θ has a posterior distribution with PDF $p(\theta \mid y)$. From this posterior is drawn an independent simulation sample of size 10000

$$\theta^{(1)},\ \theta^{(2)},\ \ldots,\ \theta^{(10000)}$$

which is used to form an approximation

$$\frac{1}{10000}\sum_{s=1}^{10000} \theta^{(s)}$$

(a) Write out the **integral** (with respect to θ) that is being approximated. [3 pts]

$$\int_{-\infty}^{\infty} \theta \, p(\theta \mid y) \, d\theta$$

(b) If the **sample variance** of the simulation sample is 25, compute the **Monte Carlo error**. [3 pts]

$$\frac{\sqrt{25}}{\sqrt{10000}} = \frac{5}{100} = 0.05$$

(c) **Name** a different type of method that might produce an approximation **without** using any random simulation. [1 pt]

deterministic numerical integration

3. A continuous parameter θ, which can have any positive scalar value, has Jeffreys' prior density proportional to $1/\sqrt{\theta}$ (for $\theta > 0$).

(a) Write a specific function of θ that is proportional to the (expected) Fisher information $I(\theta)$. [2 pts]

$$1/\theta$$

(b) Is this Jeffreys' prior **proper**? Using calculus, **precisely** explain why it is or is not. [3 pts]

For $u > 0$,

$$\int_0^u 1/\sqrt{\theta}\, d\theta = \Big[2\sqrt{\theta}\Big]_0^u$$

$$= 2\sqrt{u} \longrightarrow \infty \quad \text{as} \quad u \to \infty$$

So it is <u>not</u> proper.

4. For each part below, CIRCLE the ONE BEST answer. [1 pt each]

(a) A model for (scalar) data Y has (scalar) parameter θ. Which of these represents a **prior mean**?

$\mathrm{E}(Y \mid \theta)$ $\mathrm{E}(Y)$ $\mathrm{E}(\theta)$ $\mathrm{E}(\theta \mid Y)$

(b) In the general expression $\mathrm{Prob}(M \mid D) \propto \mathrm{Prob}(D \mid M)\,\mathrm{Prob}(M)$, the proportionality constant may depend on

D M both neither

(c) In contrast to the Bayesian perspective, the usual **frequentist** perspective regards a parameter in a data model as

random and unobserved fixed and unknown fixed and known random and observed

(d) Which names a **conjugate** distribution family for the parameter θ in the data model $Y \mid \theta \sim \mathrm{Binomial}(10, \theta)$?

beta uniform both neither

(e) Let θ be a scalar parameter in a Bayesian model for data $\boldsymbol{Y} = \boldsymbol{y}$. If $I(\theta < 0.5)$ is the indicator function of the event that $\theta < 0.5$, which of the following equals $\mathrm{E}(I(\theta < 0.5) \mid \boldsymbol{y})$?

$\mathrm{E}(\theta \mid \boldsymbol{y}) - 0.5$ the posterior median of θ $\mathrm{Prob}(\theta < 0.5 \mid \boldsymbol{y})$ none of these

(f) A posterior predictive distribution **cannot** be

continuous discrete either continuous or discrete none of these is correct

(g) If (0.2, 0.7) is an exact 95% **HPD** posterior credible interval for a scalar parameter θ that has posterior PDF $p(\theta \mid \boldsymbol{y})$, which **must** be true?

$\int_{-\infty}^{0.2} p(\theta \mid \boldsymbol{y})\, d\theta = 0.025$ $\int_{0.2}^{0.7} p(\theta \mid \boldsymbol{y})\, d\theta = 0.95$ both of these neither of these

(h) Which is a **conjugate** prior family for a rate parameter in a Poisson rate model?

normal beta inverse gamma gamma

5. List one way in which **sensitivity analysis** and **posterior predictive checking** are **similar**, and also one way in which they **differ**. [2 pts]

Both are meant to assess suitability of a prior. (or several priors).

But sensitivity analysis assesses how inference varies as the prior changes, while posterior predictive checking assesses how well a (single) prior serves to predict the observed data.

6. The Gamma(α, β) distribution with shape $\alpha > 0$ and rate (reciprocal scale) $\beta > 0$ has **mean** α/β, **variance** α/β^2, and **PDF**

$$\frac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\beta x}, \qquad x > 0.$$

Consider a model for (observed) data $Y = y > 0$ that has one parameter $\theta > 0$ and a **likelihood** proportional to

$$\theta^5\, y^4\, e^{-\theta y}$$

(a) Fully demonstrate (using Bayes' rule) that Gamma(α, β) distributions form a **conjugate** prior family for θ, and express the posterior's shape and rate in terms of the prior's shape α_0 and rate β_0 (and anything else needed). [7 pts]

$$p(\theta|y) \propto \theta^5 e^{-\theta y} \cdot \underbrace{\theta^{\alpha_0 - 1} e^{-\beta_0 \theta}}_{\text{kernel of Gamma}(\alpha_0, \beta_0)} \qquad (\theta > 0)$$

$$\propto \underbrace{\theta^{\alpha_0 + 5 - 1} e^{-(\beta_0 + y)\theta}}_{\text{kernel of Gamma}(\alpha_0 + 5,\ \beta_0 + y)} \qquad (\theta > 0)$$

posterior shape

posterior rate

(b) Assuming a Gamma(α_0, β_0) prior, express the **posterior mean** of θ as a **weighted average** of its **prior mean** and a statistic $\hat{\theta}$ that depends only on y (and not on α_0 or β_0). (Remember, weights sum to 1.) Also, what is the expression for $\hat{\theta}$? [4 pts]

$$E(\theta|y) = \frac{\alpha_0 + 5}{\beta_0 + y} = \frac{\alpha_0}{\beta_0 + y} + \frac{5}{\beta_0 + y}$$

$$= \frac{\beta_0}{\beta_0 + y} \cdot \underbrace{\frac{\alpha_0}{\beta_0}}_{E(\theta)} + \frac{y}{\beta_0 + y} \cdot \underbrace{\frac{5}{y}}_{\hat{\theta}}$$

(c) What choices of α_0 and/or β_0 apparently make the prior **less informative**? Why? [2 pts]

Less informative for <u>smaller β_0</u> and correspondingly <u>smaller α_0</u>.

For example, in that case $E(\theta|y) \approx \hat{\theta}$ (less affected by prior mean)

GRADUATE SECTION ONLY

7. The exponential distribution with **mean** $\theta > 0$ has PDF

$$f(y \mid \theta) \;=\; \frac{1}{\theta}e^{-y/\theta}, \qquad y > 0$$

In a supply of a certain machine component, 80% of the components are genuine, and the remaining 20% are counterfeit. The lifetime of a genuine component is exponentially distributed with a mean of 4 years, while the lifetime of a counterfeit component is exponentially distributed with a mean of only 1 year.

(a) Let θ be 4 for a genuine component and 1 for a counterfeit component. Express the **PMF** of θ for a randomly chosen component. [2 pts]

$$\pi(\theta) = \begin{cases} 0.2, & \theta = 1 \\ 0.8, & \theta = 4 \\ 0, & \text{o.w.} \end{cases}$$

(b) Express the marginal PDF of the lifetime Y of a randomly chosen component as a **mixture density** $m(y)$. [Note: This mixture density is for Y, not θ.] [4 pts]

$$m(y) = \sum_{\theta} f(y \mid \theta)\, \pi(\theta)$$

$$= \pi(\theta=1)\, f(y \mid \theta=1) + \pi(\theta=4)\, f(y \mid \theta=4)$$

$$= 0.2 \cdot e^{-y} + 0.8 \cdot \frac{1}{4} e^{-y/4} \qquad (y>0)$$

(c) A randomly chosen component was observed to have a lifetime of only 0.5 years. Compute the (conditional) probability that it was genuine. [4 pts]

$$p(\theta=4 \mid y=0.5) = \frac{f(y=0.5 \mid \theta=4)\, \pi(\theta=4)}{m(y=0.5)}$$

$$= \frac{\frac{1}{4} e^{-0.5/4} \cdot 0.8}{0.2\, e^{-0.5} + 0.8 \cdot \frac{1}{4} e^{-0.5/4}} = \frac{1}{1 + e^{-0.375}}$$

$$\approx 0.593$$