Student Name: ______________________________

STAT 544 Mid-term Exam
Tuesday 5 March 3:40-4:55

Instructor: Jarad Niemi

2024-03-05

INSTRUCTIONS

Please check to make sure you have 3 pages with writing on the front and back.

On the following pages you will find short answer questions related to the topics we covered in class for a total of 100 points. Please read the directions carefully.

You are allowed to use any resource except real-time help from another individual which includes the use of any messaging platform as well as posting on any discussion board. Cheating will not be tolerated. Anyone caught cheating will receive an automatic F on the exam. In addition the incident will be reported, and dealt with according to University's Academic Dishonesty regulations. Please refrain from talking to your peers, exchanging papers, writing utensils or other objects, or walking around the room. All of these activities can be considered cheating. **If you have any questions, please raise your hand.**

You will be given only 1 hour and 15 minutes (the time allotted for the course); no extra time will be given.

Good Luck!

1. *Congenital amusia*, a musical disability typically referred to as *tone deafness*, affects 4% of the population. A researcher has developed a test to identify whether a subject is tone deaf. The test involves 5 questions. For a tone deaf individual, the probability of getting each question correct is 0.2 while for a non-tone deaf individual, the probability of getting each question correct is 0.8. The researcher is willing to assume the probability of obtaining a correct answer on one question is independent of getting the correct answer on any other question. For a subject that gets 1 question correct, what is the probability the subject is tone deaf? (20 pts)

Answer: Let D indicate the subject is tone deaf while D^C indicates the subject is not tone deaf. Since 4% of the population is tone deaf, we have $P(D) = 0.04 = 1 - P(D^C)$.

```
pD  <- 0.04
pDc <- 1 - pD
```

Let Y be the number of questions the individual answered correctly. Assume $Y|D \sim Bin(5, 0.2)$ and $Y|D^C \sim Bin(5, 0.8)$.

```
y   <- 1
n   <- 5

like_D  <- dbinom(y, n, 0.2)
like_Dc <- dbinom(y, n, 0.8)

post_D <- (1 + (like_Dc * pDc) / (like_D * pD) )^-1
```

Calculate

$$\begin{aligned} P(D|Y=1) &= \frac{P(Y=1|D)P(D)}{P(Y=1|D)P(D)+P(Y=1|D^C)P(D^C)} \\ &= \left[1 + \frac{P(Y=1|D^C)P(D^C)}{P(Y=1|D)P(D)}\right]^{-1} \\ &= \left[1 + \frac{0.0064 \cdot 0.96}{0.4096 \cdot 0.04}\right]^{-1} \\ &= 0.7272727 \end{aligned}$$

2. Let Y be a random variable with the following probability density function:

$$p(y|\lambda) = \frac{ky^{k-1}}{\lambda} \exp\left(-y^k/\lambda\right)$$

for $y > 0$ and unknown $\lambda > 0$.

(a) Derive Jeffreys' prior for λ. Note: $E[Y^k] = \lambda$. (20 points)

Answer: This pdf is an exponential family since

$$p(y|\lambda) = \exp\left(-y^k/\lambda - \log(\lambda) + \log(ky^{k-1})\right).$$

Thus we can find

$$\begin{aligned} I(\lambda) &= -E\left[\tfrac{d^2}{d\lambda^2} \log p(y|\lambda)\right] & &= -E\left[\tfrac{d^2}{d\lambda^2}\left(-y^k/\lambda - \log(\lambda)\right)\right] \\ &= -E\left[\tfrac{d}{d\lambda}\left(y^k/\lambda^2 - 1/\lambda\right)\right] & &= -E\left[-2y^k/\lambda^3 + 1/\lambda^2\right] \\ &= 2E[y^k]/\lambda^3 - 1/\lambda^2 & &= 2\lambda/\lambda^3 - 1/\lambda^2 \\ &= 1/\lambda^2 \end{aligned}$$

Thus Jeffreys prior is

$$p(\lambda) \propto \sqrt{|I(\lambda)|} = 1/\lambda$$

(b) Assume we have n independent observations, $Y_1, \ldots, Y_n$, from the pdf on the previous page. Derive the posterior for λ assuming $\lambda \sim IG(a, b)$, i.e.

$$p(\lambda) = \frac{b^a}{\Gamma(\alpha)} \lambda^{-a-1} \exp\left(-b/\lambda\right)$$

for $\lambda > 0$. (20 pts)

Answer: Derive

$$\begin{array}{ll} p(\lambda|y) & \propto p(\lambda) \prod_{i=1}^{n} p(y_i|\lambda) \\ & \propto \lambda^{-a-1} \exp\left(-b/\lambda\right) \lambda^{-n} \exp\left(-\frac{1}{\lambda} \sum_{i=1}^{n} y_i^k\right) \\ & = \lambda^{-(a+n)-1} \exp\left(-\frac{1}{\lambda}\left[b + \sum_{i=1}^{n} y_i^k\right]\right) \end{array}$$

This is the kernel of an inverse gamma and thus $\lambda|y \sim IG\left(a+n, b+\sum_{i=1}^{n} y_i^k\right)$.

3. Consider the following regression model

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \epsilon_i, \quad \epsilon_i \overset{ind}{\sim} N(0, \sigma^2).$$

(a) What value of X_i either maximizes or minimizes $E[Y_i]$? (8 pts) (Show your reasoning.)

Answer: We have $E[Y] = \beta_0 + \beta_1 X + \beta_2 X^2$ and

$$\frac{d}{dX} E[Y] = \beta_1 + 2\beta_2 X \overset{set}{=} 0 \implies X = -\frac{\beta_1}{2\beta_2}$$

is either a maximum or a minimum.

(b) Under what condition is the value above a maximum? (4 pts)

Answer:

$$\frac{d^2}{dX^2} E[Y] = 2\beta_2$$

and thus $-\frac{\beta_1}{2\beta_2}$ maximizes $E[Y]$ if β_2 is negative.

(c) Suppose you have Monte Carlo samples $\beta_0^{(m)}, \beta_1^{(m)}, \beta_2^{(m)}, \sigma^{2(m)}$ for $m = 1, \ldots, M$ from the joint posterior for these parameters. Explain how you would obtain an estimate of the posterior expectation for quantity in part a.? (4 pts)

Answer: Let $\chi = -\frac{\beta_1}{2\beta_2}$ and $\chi^{(m)} = -\frac{\beta_1^{(m)}}{2\beta_2^{(m)}}$.

$$E[\chi] \approx \frac{1}{M} \sum_{m=1}^{m} \chi^{(m)}.$$

(d) Explain how would you use these Monte Carlo samples to "test the hypothesis" that the value is a maximum rather than a minimum. (4 pts)

Answer: Calculate

$$P(\beta_2 < 0|y) \approx \frac{1}{M} \sum_{m=1}^{M} \mathrm{I}\left(\beta_2^{(m)} < 0\right).$$

This probability provides a measure of the value being a maximum.

4. Consider the following model for $g = 1, \ldots, G$ with independent Y_{gi}:

$$P(Y_{gi} = k) = 1/\theta_g \, \forall \, k = 1, 2, \ldots, \theta_g \in \mathbb{N}, i = 1, 2, \ldots, n_g \quad \theta_g \stackrel{ind}{\sim} Po(\lambda), \quad \lambda \sim Ga(a, b)$$

(a) Derive the conditional posterior $p(\theta_g|\lambda, y)$ where y is the set of all observations. Recall that conditional posteriors are proportional to the full posterior of all parameters. (10 pts)

Answer:

$$\begin{aligned} p(\theta_g|\lambda, y) &\propto p(\theta, \lambda|y) \\ &\propto p(y|\theta)p(\theta|\lambda)p(\lambda) \\ &= \theta_g^{-n_g} \mathrm{I}(\max_i y_{gi} \le \theta_g) \tfrac{\lambda^{\theta_g} e^{-\lambda}}{\theta_g!} \\ &\propto \tfrac{\lambda^{\theta_g}}{\theta_g^{n_g} \theta_g!} \mathrm{I}(\max_i y_{gi} \le \theta_g) \end{aligned}$$

(b) Derive the conditional posterior $p(\lambda|\theta, y)$ where $\theta = (\theta_1, \ldots, \theta_G)$. (10 pts)

Answer: Let $G\overline{\theta} = \sum_{g=1}^{G} \theta_g$.

$$\begin{aligned} p(\lambda|\theta, y) &\propto p(\theta, \lambda|y) \\ &\propto p(\theta|\lambda)p(\lambda) \\ &\propto \lambda^{G\overline{\theta}} e^{-G\lambda} \lambda^{a-1} e^{-b\lambda} \\ &= \lambda^{a+G\overline{\theta}-1} e^{-(b+G)\lambda} \end{aligned}$$

This is the kernel of a Gamma, so $\lambda|\theta, y \sim Ga(a + G\overline{\theta}, b + G)$ and is clearly independent of y.