ST559 Homework 2

Nick Sun

April 14, 2020

Abstract

Problems 2, 8, 10, 12, 18, 19 from Bayesian Data Analysis

1 Q2

Predictive distributions: consider two coins C_1 and C_2 with the following characteristics:

1. $P(\text{heads}|C_1) = .6$
2. $P(\text{heads}|C_2) = .4$

Choose one of the coints at random and imagine spinning it repeatedly. Given that the first two spins from the chosen coin are tails, what is the expectation of the numer of additional spins until a head shows up?

Answer We should answer this question by first getting the posterior probability that we flipped C_1 given two tail flips. This will also give us the posterior probability that we flipped C_2 since there are only two possible coins. Here we have to express this probability using Bayes' rule.

$$\begin{aligned} P(C_1|TT) &= \frac{P(TT|C_1)P(C_1)}{P(TT)} \\ &= \frac{.4^2\,(.5)}{.26} \\ &\approx .307 \end{aligned}$$

Note that we computed $P(TT) = .5\,(.4^2 + .6^2) = .26$.

Then the expected number of flips N until we get heads will just be a weighted sum of the geometric means of the two coins:

$$\begin{aligned} E[N|TT] &= \frac{1}{.6}\,(.307) + \frac{1}{.4}\,(1 - .307) \\ &= 2.2441 \end{aligned}$$

2 Q8

Normal distribution with unknown mean: a random sample of n students is drawn from a normal population and their weights are measured. The average weight of the n sampled students is $\bar{y} = 150$ pounds. Assume the weights in the population are normally distributed with unknown means θ and a known standard deviation of 20 pounds. Suppose your prior distribution for θ is normal with mean 180 and standard deviation of 40.

2.1 part a.

Give your posterior distribution for θ

Answer I will first derive a general formula where given $\bar{X}$ and a known variance for μ and Y_i, the posterior probabilty will be

$$\begin{aligned} f(\bar{X}|\mu)f(\mu) &= \frac{1}{\sqrt{2\pi\sigma_0^2}} exp\left(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}\right) \frac{1}{\sqrt{2\pi(\sigma^2/n)}} exp\left(-\frac{n(\bar{X}-\mu)^2}{2\sigma^2}\right) \\ &\propto exp\left(\frac{-\mu^2(\sigma^2+n\sigma_0^2)+2\mu(n\bar{X}\sigma_0^2+\mu_0\sigma^2+(\mu_0^2\sigma^2-n\bar{X}\sigma_0^2)}{2\sigma^2\sigma_0^2}\right) \\ &\propto exp\left(\frac{-(\mu-\frac{\mu_0\sigma^2+n\bar{X}\sigma_0^2}{\sigma^2+n\sigma_0^2})^2}{2\frac{\sigma^2\sigma_0^2}{\sigma^2+n\sigma_0^2}}\right) \end{aligned}$$

Note that we had to make use of completing the square. Inputting $\mu_0 = 180, \bar{X} = 150, \sigma^2 = 20^2, \sigma_0^2 = 40^2$ to the above formula, we get the final posterior probability of $N\left(\frac{180+600n}{4n+1}, \left(\frac{1}{40^2}+\frac{n}{20^2}\right)^{-1}\right)$.

2.2 part b.

A new student is sampled at random from the same population and has a weight of $\tilde{y}$ pounds. Give a posterior predictive distribution for $\tilde{y}$.

The posterior predictive distribution will be centered at the same value as in part (a), but the variance will be inflated by the variance of a single new observation 20^2.

Answer $N\left(\frac{180+600n}{4n+1}, \left(\frac{1}{40^2}+\frac{n}{20^2}\right)^{-1}+20^2\right)$

2.3 part c.

For $n = 10$ give a 95% posterior interval for θ and a 95% posterior predictive interval for $\tilde{y}$.

Answer Plugging in $n = 10$ and using $z^* = 1.96$ into our estimates for the posterior variance, we get $150.7 \pm 1.96 * (6.25)$ for θ and $150.7 \pm 1.96 * (20.95)$ the new observation.

2.4 part d.

Do the same for $n = 100$.

Answer Now plugging in $n = 100$, we get $150.7 \pm 1.96 * (2)$ for θ and $150.7 \pm 1.96 * (20.1)$ for the new observation.

3 Q10

Suppose there are N cable cars in San Francisco, numbered from $1, \ldots, N$. You see a cable car at random and it is numbered 203. You wish to estimate N.

3.1 part a.

Assume your prior distribution on N is geometric with mean 100. What is your posterior distribution for N?

Answer If we assume that every cable car of the N cars in San Francisco is equally likely to be observed, then we arrive at the following posterior:

$$p(N|\text{Observe } 203) = \begin{cases} \frac{1}{N}(.01)(.99)^{N-1} & \text{for } N \geq 203 \\ 0 & \text{o.w.} \end{cases}$$

3.2 part b.

What are the posterior mean and standard deviation of N? Sum the infinite series analytically or approximate them on the computer.

Answer Here is some `R` code showing the answer.

```
fn <- function(n) {
     return((1/n)*(.01)*(.99)^(n-1))
}

c <- 1/sum(sapply(seq(203, 5000, by = 1), fn))

fn2 <- function(n) {
     return((.01)*(.99)^(n-1))
}

c*sum(sapply(seq(203, 5000, by = 1), fn2))

fn3 <- function(n) {
  return(n*(.01)*(.99)^(n-1))
}

sqrt((c\*sum(sapply(seq(203, 5000, by = 1), fn3))) -
    (c\*(sum(sapply(seq(203, 5000, by = 1), fn2))))\^2)
```

The first function calculates a normalizing constant c and the second function calcuates the expected value of N: 279.0885. The last function calculates the second moment of N and uses this to calculate standard deviation using the formula $\sqrt{E[N^2] - E[N]^2}$. This function returns 79.964 as the estimated standard deviation.

3.3 part c.

Choose a reasonable "noninformative" prior distribution for N and give the resulting posterior distribution, mean, and standard deviation for N.

We can suppose a noninformative prior that N is equiprobable amongst all possible values, so $\pi(N) = \frac{1}{N}$. We can also consider that $f(y|N) = \frac{1}{N}$ as well, since we are equally likely to observe any of the N cable cars. This gives us the following:

$$\begin{aligned}\pi(N|y) &\propto f(y|N)\pi(N)\\ &\propto \frac{1}{N^2}\end{aligned}$$

The next step is normalizing this density $c = (\sum_{i=203}^{\infty} \frac{1}{N^2})^{-1}$. I will use `R` to calculate an approximate normalizing constant.

```
1/sum(sapply(seq(203, 5000, by =1), function(x) 1/(x^2)))
```

This code gives $c \approx 211.05$. Therefore, the posterior distribution with this noninformative prior is $\pi(N|y) = \frac{211.05}{N^2}$. However, the first moment of this distribution diverges to infinity and the second moment of this distribution does not exist.

4 Q12

Suppose $y|\theta \sim$ Poisson(θ). Find Jeffrey's prior density for θ and then find α and β for which the *Gamma*(α, β) density is a close match to Jeffrey's density.

Answer First we need to find the Fisher information for θ. This is easy enough to do:

$$-E\left(\frac{\partial^2}{\partial\theta} ln\left(p(y|\theta)\right)\right) = \frac{E[y]}{\theta^2} = \frac{1}{\theta}$$

$$\text{So Jeffrey's prior is } \frac{1}{\theta^{1/2}}$$

This is equivalent to a Gamma distribution of $p(\alpha = 1/2, \beta = 0)$ in the parameterization $p(\theta|\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}, \theta > 0$.

5 Q18

Derive the gamma posterior distribution for the Poisson model parameterized in term of rate and exposure with conjugate prior distribution.A

Answer Given that $y_i \sim$ Poisson($x_i\theta$) and that $\theta \sim$ Gamma(α, β), we first have to find $p(y|\theta)$. This is given on page 45 of the textbook. Here is the derivation:

$$
\begin{aligned}
p\left(y|\theta\right) &= \prod_{i=1}^{n} \frac{(x_i\theta)^{y_i} e^{x_i\theta}}{y_i!} \\
&= \frac{\prod x_i^{y_i} \prod \theta^{y_i} \prod e^{x_i\theta}}{\prod y_i!} \\
&\propto \theta^{\sum y_i} e^{-\theta \sum x_i}
\end{aligned}
$$

Now to calculate $p(\theta|y)$, we combine $p\left(\theta\right)$ and $p(y|\theta)$:

$$
\begin{aligned}
p(\theta|y) &\propto p(y|\theta)p(\theta) \\
&\propto \theta^{\sum y_i} e^{-\theta \sum x_i} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta} \\
&\propto \theta^{\alpha + \sum y_i - 1} e^{-\theta(\sum x_i + \beta)} \\
&\text{which is the Gamma distribution} \left(\alpha + \sum_{i=1}^{n} y_i, \beta + \sum_{i=1}^{n} x_i\right)
\end{aligned}
$$