

HW2

STAT 651

RJ Cass

2026-01-26

1: BDA3 2.10ab

$$\pi(N) = (.01)(.99)^{N-1} \quad (N > 1, 2, 3, \dots)$$

a)

We see 203. What is the posterior for actual number of roadcars? Thinking about what the distribution of $P(203|N)$, if there were 203 cars, $P(203|N) = 1/203$. If there were 204, it would be $1/204$, etc. For any N less than 203 we couldn't have seen a 203, so it's 0. Thus:

$$P(203|N) = \frac{1}{N} \text{ for } N \geq 203$$

$$P(N|203) \propto P(203|N)\pi(N) = \frac{1}{N}(.01)(.99)^{N-1} = \frac{1}{N}(.01)(.99)(.99)^N \propto \underline{\frac{1}{N}(.99)^N}$$

b)

What are mean and sd of N ?

To determine those, we must first find the normalizing constant. We know $\sum_N P(N|203) = \sum_N \frac{c}{N}(.99)^N = 1 \implies \frac{1}{c} = \sum_N \frac{1}{N}(.99)^N \implies c = \frac{1}{\sum_N \frac{1}{N}(.99)^N}$. This is solved via R below:

```
N_list <- c(203:1000)
c <- round(1/sum(sapply(N_list, \ (N) 1/N*.99^N)), 4)
```

Giving us $c \approx 21.4701$.

$$E(N|203) = \sum_{N=203}^{\infty} N * P(N|203) = \sum_{N=203}^{\infty} \frac{cN}{N}(.99)^N = c \sum_{N=203}^{\infty} (.99)^N$$

```
E <- round(c*sum(sapply(N_list, \ (N) .99^N)), 4)
```

$$E(N|203) \approx 279.0203$$

$$\sigma(N|203) = \sqrt{\sum_{N=203}^{\infty} (N - E(N|203))^2 P(N|203)} = \sqrt{\sum_{N=203}^{\infty} (N - 279.0203)^2 \frac{21.4701}{N} (.99)^N}$$

```
sd <- round(sqrt(sum(sapply(N_list, \(N) (N - E)^2 * c / N * .99^N))), 4)
```

$$\sigma(N|203) \approx 79.6154.$$

c)

Find the MLE and compare to the Bayes estimate.

Just thinking about this, as explained above, for $N = 203$, $P(203|N) = \frac{1}{203} = \frac{1}{N}$. As N increases, the probability only decreases, meaning the MLE for N is the smallest possible value, which in our case, is 203.

I think both methods have their benefits, but the Bayes method allows a bit more leeway for there to be more than the exact number we saw and is likely a better reflection of reality (as long as our prior is actually reasonable).

2

Assume $\theta \sim \text{Beta}(\alpha, \beta)$

a)

We observe $x_1|\theta \sim \text{Binom}(n_1, \theta)$. Find $\pi(\theta|x_1)$.

$$\begin{aligned}\pi(\theta|x_1) &= f(x_1|\theta)f(\theta) = \text{Binom}(n_1, \theta) * \text{Beta}(\alpha, \beta) = \binom{n}{x_1} \theta^{x_1} (1-\theta)^{n-x_1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} \\ &= \frac{\binom{n}{x_1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha+x_1-1} (1-\theta)^{\beta+n-x_1-1}}{\end{aligned}$$

b)

We observe $x_2|\theta \sim \text{Binom}(n_2, \theta)$. Use $\pi(\theta|x_1)$ as the prior for the second day. Find $\pi(\theta|x_1, x_2)$

$$\begin{aligned}\pi(\theta|x_1, x_2) &= f(x_2|\theta)f(\theta|x_1) = \text{Binom}(n_2, \theta) * \pi(\theta|x_1) = \binom{n}{x_2} \theta^{x_2} (1-\theta)^{n-x_2} \binom{n}{x_1} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha+x_1-1} (1-\theta)^{\beta+n-x_1-1} \\ &= \frac{\binom{n}{x_1+x_2} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha+x_1+x_2-1} (1-\theta)^{\beta+2n-x_1-x_2-1}}{\end{aligned}$$

c)

Let $y = x_1 + x_2$. What is $\pi(\theta|y)$. How does this compare to answer from (b)?

$$f(y|\theta) = \binom{n}{x_1+x_2} \theta^{x_1+x_2} (1-\theta)^{n-(x_1+x_2)}$$

Thus,

$$\begin{aligned} \pi(\theta|y) &= f(y|\theta) * f(\theta) = \binom{n}{x_1+x_2} \theta^{x_1+x_2} (1-\theta)^{n-(x_1+x_2)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} \\ &= \binom{n}{x_1+x_2} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha+x_1+x_2-1} (1-\theta)^{\beta+2n-x_1-x_2} \end{aligned}$$

This is the same answer which, as stated in the problem, indicates that using a hierarchical model gives the same answer as multiple step Bayesian if you use the posterior of step 1 as the prior for step 2.

3

$$f(y_i|\lambda) = \frac{\lambda^{y_i} e^{-\lambda}}{y_i!} \text{ for } y_i \in 0, 1, 2, \dots$$

a)

Show the conjugate prior for λ is the gamma distribution

$$\begin{aligned} L(y_i|\lambda) f(\lambda) &\propto \lambda^{\sum y_i} e^{-n\lambda} [\lambda^{\alpha-1} e^{-\lambda\beta}] \\ &= \lambda^{\sum y_i + \alpha - 1} e^{-\lambda(n+\beta)} \end{aligned}$$

This is another Gamma distribution with $\alpha = \sum y_i + \alpha, \beta = n + \beta$

b)

Identify the sufficient statistic for λ

The sufficient statistic for λ is just $\sum y_i$

c)

Derive the prior predictive distribution for one month's failure count y

$$\begin{aligned}\mathbb{P}(y_i) &= \int_{\Lambda} f(y_i|\lambda)\pi(\lambda) \\ &= \int_{\Lambda} \frac{\lambda^{y_i} e^{-\lambda}}{y_i!} \frac{\beta^{\alpha}}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} d\lambda \\ &= \frac{\beta^{\alpha}}{\Gamma(\alpha)y_i!} \int_{\Lambda} \lambda^{\alpha+y_i-1} e^{-\lambda(\beta+1)} d\lambda \quad \text{This is a Gamma Kernel} \\ &= \frac{\beta^{\alpha}}{\Gamma(\alpha)y_i!} \frac{\Gamma(\alpha+y_i)}{(\beta+1)^{\alpha+y_i}}\end{aligned}$$

In talking with classmates, they helped me see that this reduces down to a $Neg.Binom.(\alpha, \frac{\beta}{1+\beta})$

d)

Verify the prior predictive by simulating (λ, y) pairs

FINISH MEEEEEE

e)

Use my prior predictive to calibrate the against the expert's opinion

FINISH MEEEEEE

f)

Find posterior using: $y = (3, 4, 3, 1, 1, 4, 1, 4, 2, 1, 1, 3, 1, 1, 4)$

FINISH MEEEEEE

g)

Use posterior to predictive to find prob. that num. failures in any given month exceeds 3

FINISH MEEEEEE

4

a)

$$f(y_i|\theta) = \theta e^{-y_i\theta}, y_i > 0$$

i. Conjugate Prior

The conjugate prior for this is the Gamma distribution.

ii. Prove conjugate prior

$$\begin{aligned} p(y_i|\theta) * \Gamma(\alpha, \beta) &= \prod \theta e^{-\theta y_i} \Gamma(\alpha, \beta) \\ &\propto \theta^n e^{-\theta \sum y_i} \theta^{\alpha-1} e^{-\beta \theta} \\ &= \theta^{n+\alpha-1} e^{-\theta(\beta + \sum y_i)} \end{aligned}$$

This is a $\Gamma(n + \alpha, \beta \sum y_i)$

iii. Choose values for prior and justify

Looking online, Google says that many of the major rivers fall in the 600-1000 mile range, so I'll pick a mean of ~750, and a standard deviation of 350 (the max values are around 2500 miles). This gives what I believe to be a reasonable representation of the expected distribution.

This gives an $\alpha \approx 4.6, \beta \approx 163.3$

iv. Find MLE and Bayesian Posterior Mean

v. Plot prior and posterior densities with MLE and Bayesian mean included

vi. Find Posterior Mean, Median, Mode, Var, SD, and 95% CI for parameter

vii. Propose 2 other priors, plot all 3 priors on same graph, find the same summary stats, put all in table

b)

$$f(y_i|\mu, \tau) = \tau^{1/2} (2\pi)^{-1/2} \exp(-\frac{\tau}{2}(y_i - \mu)^2), y_i \in \mathbb{R}$$

μ unknown, $\tau = 1/81$

i. Conjugate Prior

The conjugate prior is the Normal

ii. Prove conjugate prior

$$\begin{aligned} p(y_i|\mu, \tau) * N(\mu, \sigma) &\propto \prod e^{-\frac{\tau}{2}(y_i - \mu)^2} * e^{-1/(2\sigma^2)(x - \mu)^2} \\ &= \prod e^{-\frac{\tau}{2}(y_i^2 - 2\mu y_i + \mu^2)} * e^{-1/(2\sigma^2)(x^2 - 2\mu x + \mu^2)} \\ &= e^{-\frac{n\tau}{2}(y_i^2 - 2\mu y_i + \mu^2)} * e^{-1/(2\sigma^2)(x^2 - 2\mu x + \mu^2)} \\ &= e^{-\mu^2(-\frac{n\tau}{2} + \frac{1}{2\sigma^2}) + 2\mu(n\tau + 1/\sigma^2)} \end{aligned}$$

iii. Choose values for prior and justify

iv. Find MLE and Bayesian Posterior Mean

v. Plot prior and posterior densities with MLE and Bayesian mean included

vi. Find Posterior Mean, Median, Mode, Var, SD, and 95% CI for parameter

vii. Propose 2 other priors, plot all 3 priors on same graph, find the same summary stats, put all in table

c)

$$f(y_i|\mu, \tau) = \tau^{1/2}(2\pi)^{-1/2} \exp(-\frac{\tau}{2}(y_i - \mu)^2), y_i \in \mathbb{R}$$

τ unknown, $\mu = 87$

i. Conjugate Prior

The conjugate prior is the Gamma.

ii. Prove conjugate prior

iii. Choose values for prior and justify

iv. Find MLE and Bayesian Posterior Mean

v. Plot prior and posterior densities with MLE and Bayesian mean included

vi. Find Posterior Mean, Median, Mode, Var, SD, and 95% CI for parameter

vii. Propose 2 other priors, plot all 3 priors on same graph, find the same summary stats, put all in table

5

$$Y|\theta, Z \sim \text{Binom}(n, \theta Z)$$

$$\pi(\theta) \sim \text{Unif}(0, 1)$$

$$\pi(Z) \sim \text{Bern}(.5)$$

When $Z = 0$, $Y = 0$. Otherwise, for $Z = 1$, $Y|\theta, Z \sim \text{Binom}(n, \theta)$

a)

What is posterior that the tree is in the forest ($\mathbb{P}(Z = 1|Y)$)

FINISH MEEEEEEEEEE

b)

What is the smallest sample size such that $\mathbb{P}(Z = 1|Y = 0) < .05$

FINISH MEEEEEEEEEE

Homework Statements

Estimate of Time Taken

I estimate this homework took about 10 hours in total.

Disclosure of Resources Used

I used some online helps (Google, StackOverflow) for some help with debugging. In particular, I used Google to identify the Conjugate distributions for (4). Doing Priors and everything is still really hard for me (this is my first exposure to them), so I also used them to help me understand how to setup the problems (ie. get the Likelihood, know what to multiply, etc.)