STAT 201B Homework 4

Hanze Yao
3033093286

1. (a) $f(\lambda | x^n) \propto f(x^n | \lambda) f(\lambda)$

$$= \left(\frac{1}{\lambda}\right)^n e^{-\frac{1}{\lambda}\sum_{i=1}^{n} X_i} \frac{b^a}{\Gamma(a)} \lambda^{-a-1} e^{-b/\lambda}$$

$$\propto \lambda^{-n} e^{-\frac{1}{\lambda}\sum_{i=1}^{n} X_i} \lambda^{-a-1} e^{-b/\lambda}$$

$$= \lambda^{-(n+a)-1} e^{-\frac{1}{\lambda}\left(\sum_{i=1}^{n} X_i + b\right)}$$

$\therefore$ $f(\lambda | x^n)$ also satisfies Inverse Gamma distribution (a', b') where $a' = n+a$ $\quad b' = \sum_{i=1}^{n} X_i + b$

(b) Now the prior mean is $\frac{b}{a-1}$

$$\frac{\partial}{\partial \lambda} \log\left(\lambda^{-n} e^{-\frac{\sum X_i}{\lambda}}\right) = \frac{\partial}{\partial \lambda}\left[(-n)\log\lambda - \frac{\sum_{i=1}^{n} X_i}{\lambda}\right] = -n \times \frac{1}{\lambda} + \frac{\sum_{i=1}^{n} X_i}{\lambda^2} = 0$$

$$\lambda = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}_n$$

Check the $\frac{\partial^2}{\partial \lambda^2} \log\left(\lambda^{-n} e^{-\frac{\sum X_i}{\lambda}}\right) = \frac{n}{\lambda^2} - \frac{2\sum X_i}{\lambda^3}$

When $\lambda = \bar{X}_n$ $\quad \frac{\partial^2}{\partial \lambda^2} \log\left(\lambda^{-n} e^{-\frac{\sum X_i}{\lambda}}\right) = \frac{n^3}{(\sum_{i=1}^{n} X_i)^2} - \frac{2n^3}{(\sum_{i=1}^{n} X_i)^2} = -\frac{n^3}{(\sum_{i=1}^{n} X_i)^2} < 0$

$\therefore$ $\hat{\lambda} = \bar{X}_n$ is the MLE of λ.

The posterior mean is $\frac{b'}{a'-1} = \frac{\sum_{i=1}^{n} X_i + b}{n+a-1} = \left(\frac{a-1}{n+a-1}\right)\frac{b}{a-1} + \left(\frac{n}{n+a-1}\right)\bar{X}_n$

$\because$ When $n \to \infty$, $\frac{a-1}{n+a-1} \to 0$ $\quad \frac{n}{n+a-1} \to 1$

$\therefore$ $\frac{b'}{a'+1} \longrightarrow \bar{X}_n$

$\therefore$ When $n \to \infty$ the posterior mean will be close to $\bar{X}_n$.

2. (a) $m = \frac{b}{a-1}$ $\quad V = \frac{b^2}{(a-1)^2(a-2)}$

$$(m(a-1))^2 = V(a-1)^2(a-2)$$

$$m^2(a-1)^2 = V(a-1)^2(a-2)$$

$\because$ $a > 2$ $\therefore$ $m^2 = V(a-2)$ $\quad a-2 = \frac{m^2}{V}$ $\quad a = \frac{m^2+2V}{V}$

$\therefore$ $b = (a-1)m = \left(\frac{m^2+V}{V}\right)m = \frac{m^3+mV}{V}$ $\quad \therefore \begin{cases} a = \frac{m^2+2V}{V} \\ b = \frac{m^3+mV}{V} \end{cases}$

(b) Based on my current knowledge, I have lived in Berkeley for 15 month now and I have felt two earthquakes. The time difference between these two is about 1 year. So I would guess m to be 365. But I completely have no knowledge about the prior variance v, So I guess this variance to be $v = 1000$

(c) The R code and the plot of the prior and posterior PDF are on the last page.

From the plot I can see that the mean λ value is a little smaller then my knowledge which is about 310 days. the variance is smaller than my guess too.

3. According to the definition of rejection sampling, the probability of the acceptance of θ^{cand} is:

$$P\left(u \le \frac{f(x^n \mid \theta^{cand})}{f(x^n \mid \hat{\theta}_n)}\right) = \frac{f(x^n \mid \theta^{cand})}{\underbrace{f(x^n \mid \hat{\theta}_n)}_{\hat{\theta}_n \to \text{MLE of } \theta} \Rightarrow M}$$ (Because $u \sim \text{Unif}(0,1)$)

$\because M = f(x^n \mid \hat{\theta}_n) = \sup_{\theta} f(x^n \mid \theta^{cand})$

$\therefore$ If $M < f(x^n \mid \hat{\theta}_n)$

then $\frac{f(x^n \mid \theta^{cand})}{M}$ may > 1

$\therefore$ the probability of the acceptance of θ^{cand} will be $\min\left\{1, \frac{f(x^n \mid \theta^{cand})}{M}\right\}$

Then when $M < f(x^n \mid \hat{\theta}_n)$

This probability may not be $\frac{f(x^n \mid \theta^{cand})}{M}$

$\because f(\theta^{accep}) \propto f(\theta^{cand})\, P(\text{acceptance of } \theta^{cand})$

$$= f(\theta^{cand}) \min\left\{1, \frac{f(x^n \mid \theta^{cand})}{M}\right\}$$

Then $f(\theta^{accep})$ may not $\propto f(\theta^{cand}) \times \frac{f(x^n \mid \theta^{cand})}{M}$

may not $\propto f(\theta^{cand})\, f(x^n \mid \theta^{cand})$

may not $\propto f(\theta^{cand} \mid x^n)$

$\therefore$ When $M < f(x^n \mid \theta^{cand})$, θ^{accep} may not satisfy the posterior distribution

4. (a) Obviously $f(x \mid \lambda)$ belongs to exponential family.

$$f(\lambda) \propto I(\lambda)^{\frac{1}{2}} = \left[\frac{I_n(\lambda)}{n}\right]^{\frac{1}{2}} = \left[-\frac{E_\lambda\left(\frac{\partial^2}{\partial \lambda^2} \log\left[\left(\frac{1}{\lambda} e^{-\frac{x_1}{\lambda}}\right)\left(\frac{1}{\lambda} e^{-\frac{x_2}{\lambda}}\right) \cdots \left(\frac{1}{\lambda} e^{-\frac{x_n}{\lambda}}\right)\right]\right.}{n}\right]^{\frac{1}{2}}$$

$$= \left[\frac{-E_\lambda \left(\frac{\partial^2}{\partial \lambda^2} \log(\lambda^{-n} e^{-\frac{\sum_{i=1}^{n} X_i}{\lambda}}) \right)}{n} \right]^{\frac{1}{2}} = \left[\frac{-E_\lambda \left(\frac{n}{\lambda^2} - \frac{2\sum_{i=1}^{n} X_i}{\lambda^3} \right)}{n} \right]^{\frac{1}{2}}$$

$$= \left[-E_\lambda \left(\frac{1}{\lambda^2} - \frac{2\bar{X}_n}{\lambda^3} \right) \right]^{\frac{1}{2}}$$

$\because X_1 \cdots X_n$ are i.i.d $\therefore$ $E_\lambda \bar{X}_n = E_\lambda X_1 = \int_0^{+\infty} \frac{1}{\lambda} e^{-\frac{X_1}{\lambda}} dx_1$

$$= -x_1 e^{-\frac{x_1}{\lambda}} \Big|_0^{+\infty} - \int_0^{+\infty} (-e^{-\frac{x_1}{\lambda}}) dx_1 = 0 - (-\lambda) = \lambda$$

$\therefore f(\lambda) \propto \left[-\frac{1}{\lambda^2} + \frac{2\lambda}{\lambda^3} \right]^{\frac{1}{2}} = \left[\frac{1}{\lambda^2} \right]^{\frac{1}{2}}$

$\because f(X|\lambda) \geq 0 \quad \therefore \lambda > 0$

$\therefore f(\lambda) \propto \frac{1}{\lambda}$

(b) This Jefferys Prior is not proper.

$\because \int_0^{+\infty} \frac{1}{\lambda} d\lambda = \log \lambda \Big|_0^{+\infty} = +\infty$

$\therefore$ This Jefferys Prior is not proper

5. (a) $f(x, y | H_1) = \int f(x, y, P_1, P_2 | H_1) \, dP_1 dP_2$

$$= \int f(x, y | P_1, P_2, H_1) f(P_1, P_2 | H_1) \, dP_1 P_2$$

$\because P_1$ and P_2 are independent, x and y are independent

$$\therefore = \int_0^1 f(x|P_1) f(y|P_2) \cdot 1 \cdot 1 \cdot dP_1 \, dP_2$$

$$= \int_0^1 \binom{n}{x} P_1^x (1-P_1)^{n-x} \binom{m}{y} P_2^y (1-P_2)^{m-y} \, dP_1 dP_2$$

$$= \binom{n}{x} \left(\int_0^1 P_1^x (1-P_1)^{n-x} dP_1 \right) \binom{m}{y} \left(\int_0^1 P_2^y (1-P_2)^{m-y} dP_2 \right)$$

$$= \binom{n}{x} B(x+1, n-x+1) \binom{m}{y} B(y+1, m-y+1)$$

$$= \frac{1}{(n+1)(m+1)}$$

(b) $f(x, y | H_0)$ $\quad \because$ Under H_0 $P_1 = P_2 \sim \text{Unif}(0, 1)$, X and Y are independent

$$\therefore = \int_0^1 \binom{n}{x} P_1^x (1-P_1)^{n-x} \binom{m}{y} P_1^y (1-P_1)^{m-y} \, dP_1$$

$$= \int_0^1 \binom{n}{x} \binom{m}{y} P_1^{x+y} (1-P_1)^{m+n-x-y} \, dP_1 = \binom{n}{x}\binom{m}{y} B(x+y+1, m+n-x-y+1)$$

$$= \binom{n}{x}\binom{m}{y}\frac{\Gamma(x+y+1)\,\Gamma(m+n-x-y+1)}{\Gamma(m+n+2)}$$

$$= \frac{n!}{x!\,(n-x)!}\;\frac{m!}{y!\,(m-y)!}\;\frac{(x+y)!\,(m+n-x-y)!}{(m+n+1)!}$$

$$= \frac{(x+y)!}{x!\;y!}\;\frac{(m+n-x-y)!}{(n-x)!\,(m-y)!}\;\frac{m!\,n!}{(m+n+1)!}$$

$$= \binom{x+y}{x}\binom{m+n-x-y}{n-x}\frac{1}{(m+n+1)\binom{m+n}{n}}$$

(c) $$BF_{10} = \frac{f(x,y\mid H_1)}{f(x,y\mid H_0)}$$

$$= \frac{\dfrac{1}{(m+1)(n+1)}}{\binom{x+y}{x}\binom{m+n-x-y}{n-x}\dfrac{1}{(m+n+1)\binom{m+n}{n}}}$$

$$= \frac{1}{(m+1)(n+1)} \times \frac{(m+n+1)\binom{m+n}{n}}{\binom{x+y}{x}\binom{m+n-x-y}{n-x}}$$

$$= \frac{1}{m+n+2}\;\frac{(m+n+2)(m+n+1)(m+n)!\,/\,(m+1)\,m!\,(n+1)\,n!}{\binom{x+y}{x}\binom{m+n-x-y}{n-x}}$$

$$= \frac{1}{m+n+2}\;\frac{\binom{m+n+2}{m+1}}{\binom{x+y}{x}\binom{m+n-x-y}{n-x}}$$

(d) Pujols : $\overset{n}{5146}$ at bats , $\overset{x}{1717}$ hits

Suzuki : $\overset{m}{6099}$ at bats , $\overset{y}{2030}$ hits

$$\therefore BF_{10} = \frac{1}{5146+6099+2}\;\frac{\binom{5146+6099+2}{6099+1}}{\binom{1717+2030}{1717}\binom{5146+6099-1717-2030}{5146-1717}} \approx 0{,}022 < 1$$

$\therefore$ The evidence against H_0 and favor H_1 is weak

So it favors H_0

Q2 R code & Plot

```
install.packages("MCMCpack")
library(MCMCpack)

## Specify the prior distribution

m <- 365    # Prior mean
v <- 1000 # Prior variance

a <- (m*m + 2*v)/v
b <- (m*m*m + m*v)/v

## Plot the prior density

x <- seq(200, 499, length = 300)
prior <- dinvgamma(x, a, b) # the prior pdf
plot(x, prior, type = "l", xlab = expression(lambda),
        ylab = expression(f(lambda)), main = "Prior density")

## The data

load("~/Documents/STAT_201_B/Homework/HW4/BerkeleyEarthquakes.RData") # Load the data
of EQs in Berkeley.
head(earthquakes)
y <- earthquakes$Lag[-1] # Extract the waiting time between each EQs.
n <- length(y)

## Update the parameters to get the posterior distribution

a.star <- n + a
b.star <- n*mean(y) + b
posterior <- dinvgamma(x, a.star, b.star)

## Plot the prior and posterior densities

plot(x, posterior, type = "l", xlab = expression(lambda), ylab = expression(f(lambda)))
lines(x, prior, col = 2)
legend("topleft", lty = rep(1, 2), col = 1:2,
           legend = c("Posterior", "Prior"), bty = "n")

## Exercise: calculate the equal tail and HPD credible intervals
```

(They will be close, since the posterior is nearly symmetric)