Assignment Report

Computational Methods for Bayesian Statistics

Trilokinath Modi

September 18, 2020

1 Question Set 1

1.1 Q1 - a

Function $C(\theta)$ can be found by

$$\int_0^\infty C(\theta)\, x^2\, \exp(-\theta\, x^3) = 1$$

Integrating by parts

$$\implies C(\theta) = 3\theta$$

1.2 Q1 - b

The likelihood function is

$$\begin{aligned} \pi(x \mid \theta) &= 3\, x^2\, \theta \exp(-\theta\, x^3) \\ &= h(x)\, g(\theta)\, \exp(\theta\, u(x)) \end{aligned}$$

Assuming a prior family to be of type

$$\pi(\theta \mid \alpha, \beta) \propto_\theta g(\theta)^\alpha\, \exp(\theta, \beta)$$

and one such distribution is gamma distribution. Hence, we can assume the prior to be gamma distributed.

$$\pi(\theta \mid \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1}\, e^{-\beta\theta}$$

The posterior density will take the form,

$$\begin{aligned} \pi(\theta \mid x) &\propto_\theta \pi(x \mid \theta)\, \pi(\theta \mid \alpha, \beta) \\ &\propto_\theta c(x) \theta^{\alpha+1-1} e^{-\theta(\beta + x^3)} \\ &= c(x)\ \mathrm{Gam}(\theta;\ \alpha + 1, \beta + x^3) \\ &= c(x)\ \mathrm{Gam}(\theta;\ \alpha^{'}, \beta^{'}) \end{aligned}$$

The variable involved in posterior is θ and hence c(x) can be considered as constant. We can see that the posterior density belongs to the conjugate family of prior density i.e. Gamma distribution.

1.3 Q1 - c

The prior predictive density $\pi(x)$ can be computed as

$$\begin{aligned}
\pi(x) &= \int_0^\infty \pi(x \mid \theta)\pi(\theta) d\theta \\
&= \int_0^\infty 3x^2\theta\, e^{-3} \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\theta\beta} d\theta \\
\text{Assume } Y &= X^3 \text{ as } x \in [0, \infty) \text{ then,} \\
\implies \pi(y^{1/3}) &= \frac{y^{2/3}\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty \theta^\alpha e^{-\theta(y+\beta)} d\theta \\
&= \frac{y^{2/3}\beta^\alpha}{\Gamma(\alpha)(y+\beta)^{(\alpha+1)}} \int_0^\infty z^{\alpha+1-1} e^{-z} dz \\
&= \frac{y^{2/3}\beta^\alpha}{\Gamma(\alpha)(y+\beta)^{(\alpha+1)}} \Gamma(\alpha+1) \\
\implies \pi(x) &= \frac{x^2\beta^\alpha}{\Gamma(\alpha)(x^3+\beta)^{(\alpha+1)}} \Gamma(\alpha+1)
\end{aligned}$$

We can use prior predictive to compute exact posterior density formula.

$$\begin{aligned}
\pi(\theta \mid x) &= \frac{\pi(x \mid \theta)\pi(\theta)}{\pi(x)} \\
&= \frac{(x^3+\beta)^{(\alpha+1)}}{\Gamma(\alpha+1)} \theta^\alpha e^{-\theta(\beta+x^3)} \\
&= \mathrm{Gam}(\theta\ ;\ \alpha+1\, ,\ \beta+x^3)
\end{aligned}$$

Now, if we assume $\alpha = 1$ and $\beta = 1$, then we get prior predictive density and posterior density respectively as

$$\begin{aligned}
\text{Prior predictive density, } \pi(x) &= \frac{x^2}{(x^3+1)^2} \\
\text{Posterior density, } \pi(\theta \mid x) &= \mathrm{Gam}(\theta\ ;\ 2\, ,\ 1+x^3) \\
&= (1+x^3)^2 \theta\, e^{-\theta(1+x^3)}
\end{aligned}$$

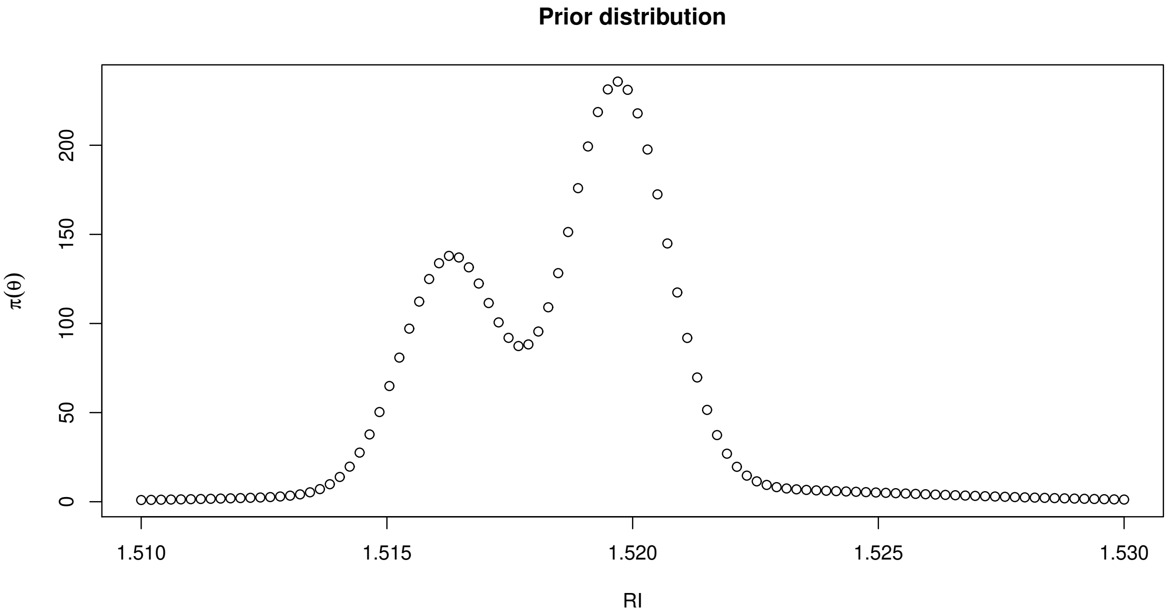

Figure 1: *Prior probability plot for RI*

2 Question Set 2

2.1 Q2 - a

The density plot looks like fig 1.

2.2 Q2 - b

Normal-Normal conjugacy along with the linear combination of prior can be used to compute the prior predictive distribution.

$$\pi(x) = \sum_{i=1}^{3} \gamma_i f_i(x)$$

Where,

$$f_i(x) = N(\mu_i \,, \sigma_i^2 + \sigma_L^2)$$

Hence,

$$\pi_x = 0.33N(1.5163 \,, 2 \times 10^-6) + 0.57N(1.5197 \,, 2 \times 10^-6) + 0.10N(1.5203 \,, 2.6 \times 10^-5)$$

The density plot for prior predictive looks like fig 2.

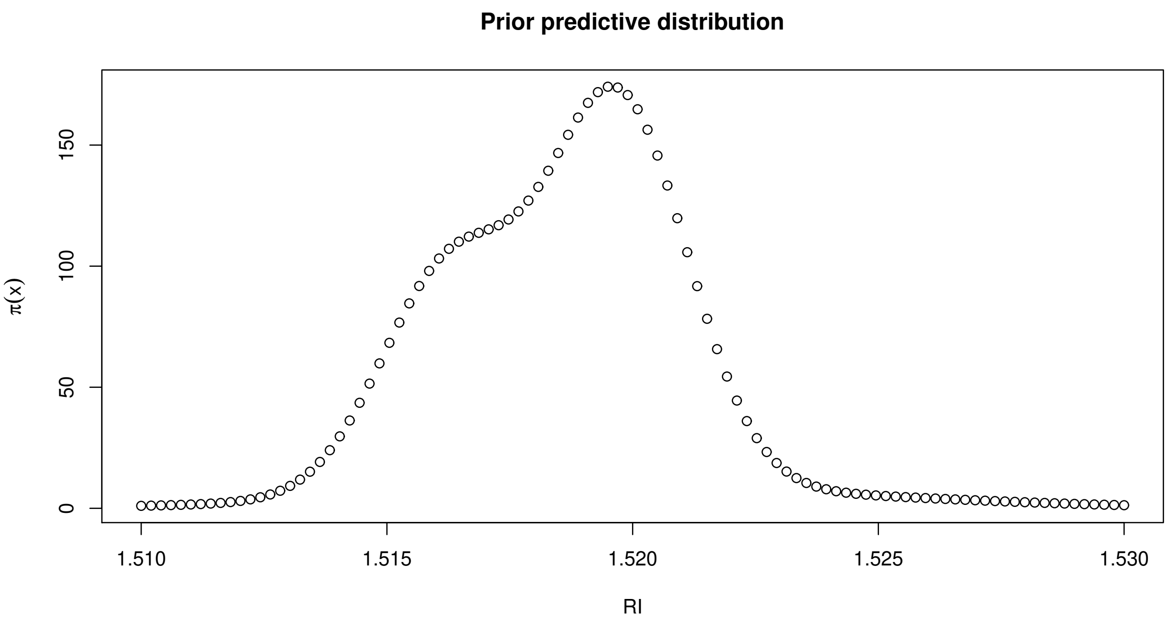

Figure 2: *Prior predictive distribution plot*

2.3 Q2 - c

The posterior density can be computed as

$$\pi(\theta \mid x) = \sum_{i=1}^{3} \gamma_i^{'} g_i(\theta \mid x)$$

where,

$$g_i(\theta \mid x) = N\left(\theta \,; \frac{x\sigma_i + \mu_i\sigma_L}{\sigma_i + \sigma_L}\,, \frac{\sigma_i\sigma_L}{\sigma_i + \sigma_L}\right)$$

$$\gamma_i^{'} = \frac{\gamma_j f_j(x)}{\sum_{i=1}^{3}\gamma_i f_i(x)}$$

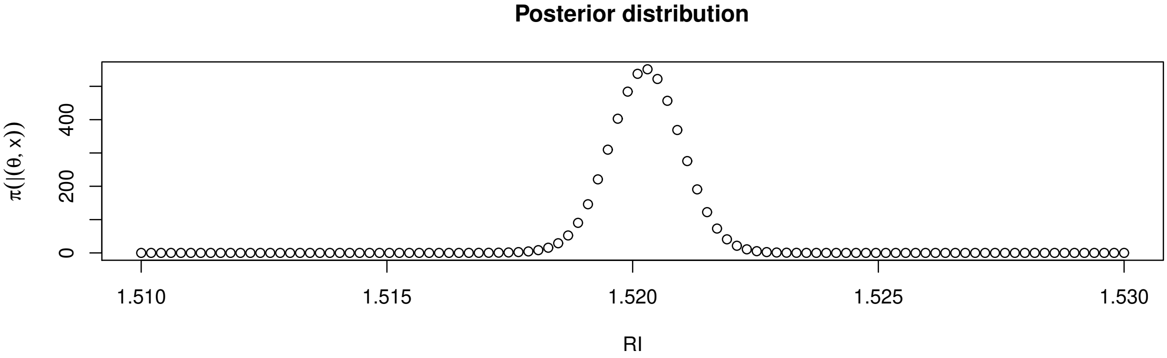

Figure 3: *Posterior density,* $\pi(\theta \mid x = 1.52083)$

Substituting the values we get,

$$\pi(\theta \mid x) = \frac{\sum_{i=1}^{3} \frac{\gamma_i}{2\pi\sigma_i\sigma_L} \exp\left[\frac{-1}{2}\left\{\frac{(x-\mu_i)^2}{\sigma_i^2+\sigma_L^2} + \frac{\left(\theta - \frac{\mu_i\sigma_i^2 + x\sigma_L^2}{\sigma_i^2+\sigma_L^2}\right)^2}{\frac{\sigma_i^2\sigma_L^2}{\sigma_i^2+\sigma_L^2}}\right\}\right]}{\sum_{i=1}^{3} \frac{\gamma_i}{\sqrt{2\pi}} \frac{1}{\sqrt{\sigma_i^2+\sigma_L^2}} \exp\left[\frac{-1}{2}\left\{\frac{(x-\mu_i)^2}{\sigma_i^2+\sigma_L^2}\right\}\right]}$$

$$= \frac{\frac{1}{\sqrt{2\pi}\sigma_L} \sum_{i=1}^{3} \frac{\gamma_i}{\sigma_i} \exp\left[\frac{-1}{2}\left\{\frac{(x-\mu_i)^2}{\sigma_i^2+\sigma_L^2}\right\}\right] \exp\left[\frac{-1}{2}\left\{\frac{\left(\theta - \frac{\mu_i\sigma_i^2 + x\sigma_L^2}{\sigma_i^2+\sigma_L^2}\right)^2}{\frac{\sigma_i^2\sigma_L^2}{\sigma_i^2+\sigma_L^2}}\right\}\right]}{\sum_{i=1}^{3} \frac{\gamma_i}{\sqrt{\sigma_i^2+\sigma_L^2}} \exp\left[\frac{-1}{2}\left\{\frac{(x-\mu_i)^2}{\sigma_i^2+\sigma_L^2}\right\}\right]}$$

$$\implies \pi(\theta \mid x) = \sum_{i=1}^{3} w_i N\left(\theta\,;\frac{\mu_i\sigma_i^2 + x\sigma_L^2}{\sigma_i^2+\sigma_L^2}\,,\,\frac{\sigma_i^2\sigma_L^2}{\sigma_i^2+\sigma_L^2}\right)$$

Where new weights,

$$w_i = \frac{\frac{\gamma_i}{\sqrt{\sigma_i^2+\sigma_L^2}} \exp\left[\frac{-1}{2}\frac{1}{\sigma_i^2\sigma_L^2}\left\{x^2\sigma_L^2 + \mu_i^2\sigma_i^2 - \left(\frac{(\mu_i\sigma_i^2 + x\sigma_L^2)^2}{\sigma_i^2+\sigma_L^2}\right)\right\}\right]}{\sum_{i=1}^{3} \frac{\gamma_i}{\sqrt{\sigma_i^2+\sigma_L^2}} \exp\left[\frac{-1}{2}\left\{\frac{(x-\mu_i)^2}{\sigma_i^2+\sigma_L^2}\right\}\right]}$$

If $x = 1.52083$ then, the posterior density will be, $\pi(\theta \mid x = 1.52083)$. The density plot is shown in fig 3. The posterior mean, standard deviation and weights are shown in table 1. 95% credibility interval can be computed as,

$$J_\theta = \sum_{i=1}^{3} \gamma_i^{(p)}(\mu_i^{(p)} \pm 1.96\sqrt{\sigma_i^{(p)}})$$
$$= [1.467566\ ,\ 1.572956]$$

Table 1: *Posterior density parameters*

i	$\gamma_i^{(p)}$	$\mu_i^{(p)}$	$\sigma_i^{(p)}$
1	4.39×10^-3	1.518565	7.071×10^-4
2	0.933	1.520265	7.071×10^-4
3	0.062	1.520320	9.805×10^-4

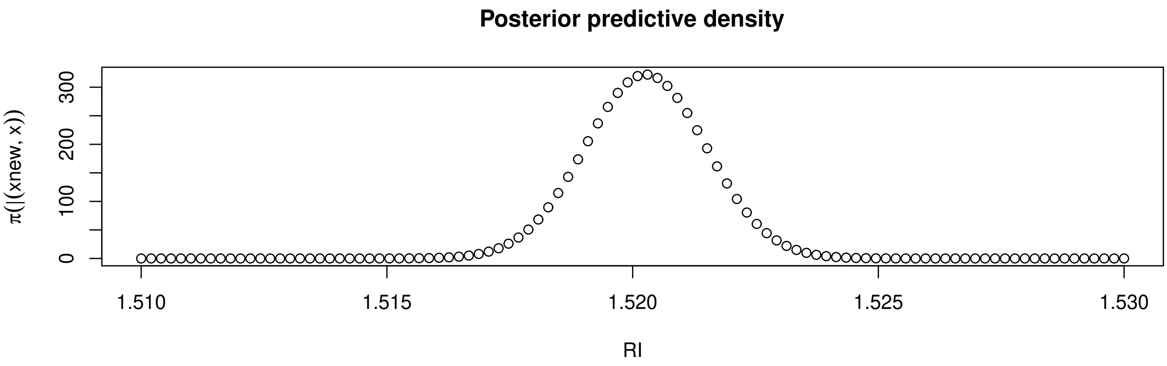

Figure 4: *Posterior predictive density,* $\pi(y \mid x = 1.52083)$

2.4 Q2 - d

The posterior predictive density can be computed as,

$$
\begin{aligned}
\pi(y \mid x) &= \int_\theta \pi(y \mid \theta)\pi(\theta \mid x)d\theta \\
&= \sum_{i=1}^{3} \gamma_i^{(p)} \int_\theta N(y\,;\theta\,,\sigma_L^2) N\left(\theta\,;\frac{\mu_i\sigma_i^2 + x\sigma_L^2}{\sigma_i^2+\sigma_L^2}\,,\,\frac{\sigma_i^2\sigma_L^2}{\sigma_i^2+\sigma_L^2}\right)d\theta \\
&= \sum_{i=0}^{3} \gamma_i^{(p)} N\left(y\,;\,\frac{\mu_i\sigma_i^2 + x\sigma_L^2}{\sigma_i^2+\sigma_L^2}\,,\,\frac{2\sigma_i^2\sigma_L^2+\sigma_L^4}{\sigma_i^2+\sigma_L^2}\right)
\end{aligned}
$$

The posterior density function is plotted in figure 4.

2.5 Q2 - e

The likelihood ratio is computed as

$$
LR = \frac{\text{Posterior odds}}{\text{prior odds}}
$$

$$
\begin{aligned}
\frac{\pi(Data \mid H_0)}{\pi(Data \mid H_1)} &= \frac{\pi(H_0 \mid Data)}{\pi(H_1 \mid Data)}\frac{\pi(H_1)}{\pi(H_0)} \\
&= \frac{\pi(\theta_0 \mid x_s)}{\pi(\theta_1 \mid x_c)}\frac{\pi(\theta_1)}{\pi(\theta_0)} \\
&= \frac{\pi(\theta_0 \mid x_s)}{\pi(\theta_1 \mid x_c)}
\end{aligned}
$$

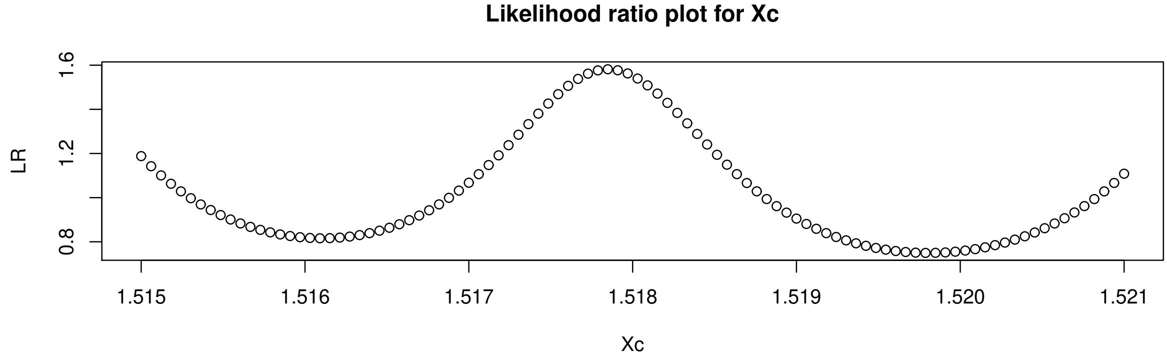

Figure 5: *Likelihood Ratio plot for X_c*

2.6 Q2 - f

For $x_s = 1.52083$, we have

$$LR = \frac{404.82}{\sum_{i=1}^{3} w_i N\left(\theta_1 \; ; \frac{\mu_i\sigma_i^2 + x_c\sigma_L^2}{\sigma_i^2+\sigma_L^2} \; , \; \frac{\sigma_i^2\sigma_L^2}{\sigma_i^2+\sigma_L^2}\right)}$$

Where weights,

$$w_i = \frac{\frac{\gamma_i}{\sqrt{\sigma_i^2+\sigma_L^2}} \exp\left[\frac{-1}{2}\frac{1}{\sigma_i^2\sigma_L^2}\left\{x_c^2\sigma_L^2 + \mu_i^2\sigma_i^2 - \left(\frac{(\mu_i\sigma_i^2+x\sigma_L^2)^2}{\sigma_i^2+\sigma_L^2}\right)\right\}\right]}{\sum_{i=1}^{3}\frac{\gamma_i}{\sqrt{\sigma_i^2+\sigma_L^2}} \exp\left[\frac{-1}{2}\left\{\frac{(x_c-\mu_i)^2}{\sigma_i^2+\sigma_L^2}\right\}\right]}$$

The range for x_c is chosen as $[1.515, 1.521]$ and the corresponding plot is displayed in figure 5. The plot can be interpreted in a way that if the LR is > 1 then the odds favoring null hypothesis is higher than odds favouring alternate hypothesis. For e.g. at $x_c = 1.518$, the $LR \approx 1.8$ implies if the RI measured at crime scene is 1.518 then the acceptance of null hypothesis is 1.8 times higher than alternate hypothesis provided that the RI measured on suspect jacket is 1.52083.

2.7 Q2 - g

With each successive measurement, the posterior density can be updated as it is dependent on likelihood and prior. With each additional measurement, the likelihood function which is a resultant of product of distribution of observations will be updated. Also, the prior can be updated with previous posterior. In this way, Multiple observations will lead to accuracy in likelihood ratio.