Homework 3

Yaqi Dai Mengjiao Wu

Question 1

$$
\begin{aligned}
ln^c(\mathbf{\Psi}) &= \sum_{i=1}^{n}\sum_{j=1}^{m} z_{ij}log\{\pi_j\phi(y_i-\mathbf{X_i^T}\beta^{\mathbf{j}};0,\sigma^2)\} \\
Q(\mathbf{\Psi}|\mathbf{\Psi^{(k+1)}}) &= E_z(ln^c(\mathbf{\Psi})) \\
&= \sum_{z} E[\boldsymbol{z}|\boldsymbol{y},\mathbf{\Psi}^{(k+1)}]lnE[\boldsymbol{y},\boldsymbol{z}|\mathbf{\Psi}] \\
&= \sum_{j=1}^{m}\sum_{i=1}^{n} E[z_{ij}|y_i,\mathbf{\Psi}^{(k+1)}]lnE[z_{ij},y_i|\mathbf{\Psi}]
\end{aligned}
$$

let

$$
\begin{aligned}
p_{ij}^{(k+1)} &= E[z_{ij}|y_i,\mathbf{\Psi}^{(k+1)}] \\
&= \frac{\pi_j^{(k+1)}\phi(y_i-\mathbf{X_i^T}\beta_{\mathbf{j}}^{\mathbf{(k+1)}};0,\sigma^{2(k+1)})}{\sum_{j=1}^{m}\pi_j^{(k+1)}\phi(y_i-\mathbf{X_i^T}\beta_{\mathbf{j}}^{\mathbf{(k+1)}};0,\sigma^{2(k+1)})}
\end{aligned}
$$

with

$$
\begin{aligned}
E[z_{ij},y_i|\mathbf{\Psi}^{(k+1)}] &= E[z_{ij}|\mathbf{\Psi}^{(k+1)}]E[y_i|z_{ij},\mathbf{\Psi}^{(k+1)}] \\
&= \pi_j^{(k+1)}\phi(y_i-\mathbf{X_i^T}\beta_{\mathbf{j}}^{\mathbf{(k+1)}};0,\sigma^{2(k+1)})
\end{aligned}
$$

$$
\begin{aligned}
Q(\mathbf{\Psi}|\mathbf{\Psi}^{(k+1)}) &= \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)}log(\pi_j^{(k+1)}\phi(y_i-\mathbf{X_i^T}\beta_{\mathbf{j}}^{\mathbf{(k+1)}};0,\sigma^{2(k+1)})) \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)}(log\pi_j^{(k+1)}+log(\phi(y_i-\mathbf{X_i^T}\beta_{\mathbf{j}}^{\mathbf{(k+1)}};0,\sigma^{2(k+1)})))
\end{aligned}
$$

$$
\begin{aligned}
Q(\mathbf{\Psi}|\mathbf{\Psi}^{(k+1)}) = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)}log((2\pi)^{\frac{d}{2}}\pi_j^{(k+1)}) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)}log|\sigma^{2(k+1)}| \\
-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m}\frac{p_{ij}^{(k+1)}}{\sigma^{2(k+1)}}(y_i-\boldsymbol{X}_i^T\boldsymbol{\beta}_j^{(k+1)})'(y_i-\boldsymbol{X}_i^T\boldsymbol{\beta}_j^{(k+1)}) \\
= I_1-\frac{1}{2}I_2-\frac{1}{2}I_3
\end{aligned}
$$

where

$$
\begin{aligned}
I_1 &= \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)}log((2\pi)^{-0.5d}\pi_j^{(k+1)}) \\
I_2 &= \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)}log|\sigma^{2(k+1)}| \\
I_3 &= \sum_{i=1}^{n}\sum_{j=1}^{m}\frac{p_{ij}^{(k+1)}}{\sigma^{2(k+1)}}(y_i-\boldsymbol{X}_i^T\boldsymbol{\beta}_j^{(k+1)})'(y_i-\boldsymbol{X}_i^T\boldsymbol{\beta}_j^{(k+1)})
\end{aligned}
$$

We can find that only I_3 contains $\beta_j^{(k+1)}$, which is a quardratic form.

$$I_{3j} = \sum_{i=1}^{n} \frac{p_{ij}^{(k+1)}}{\sigma^{2(k)}} (y_i - \boldsymbol{X}_i^T \boldsymbol{\beta}_j^{(k+1)})'(y_i - \boldsymbol{X}_i^T \boldsymbol{\beta}_j^{(k+1)})$$

We only need to find $\beta_j^{(k+1)}$ to minimize each I_{3j}

$$\frac{\partial I_{3j}}{\partial \beta_j^{(k+1)}} = \frac{-2}{\sigma^{2(k+1)}} \sum_{i=1}^{n} p_{ij}^{(k+1)} \boldsymbol{X}_i (y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})$$

Let it equals 0

$$\sum_{i=1}^{n} p_{ij}^{(k+1)} \boldsymbol{X}_i (y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)}) = 0$$

$$\sum_{i=1}^{n} p_{ij}^{(k+1)} \boldsymbol{X}_i y_i = \sum_{i=1}^{n} p_{ij}^{(k+1)} \boldsymbol{X}_i \boldsymbol{X}_i^T \beta_j^{(k+1)}$$

$$\beta_j^{(k+1)} = (\sum_{i=1}^{n} p_{ij}^{(k+1)} \boldsymbol{X}_i \boldsymbol{X}_i^T)^{-1} \sum_{i=1}^{n} p_{ij}^{(k+1)} \boldsymbol{X}_i y_i$$

Then we find only I_2, I_3 contains $\sigma^{2(k)}$.

$$I_2 + I_3 = I_2 = \sum_{i=1}^{n} \sum_{j=1}^{m} p_{ij}^{(k+1)} log|\sigma^{2(k+1)}| + \sum_{i=1}^{n} \sum_{j=1}^{m} \frac{p_{ij}^{(k+1)}}{\sigma^{2(k)}} (y_i - \boldsymbol{X}_i^T \boldsymbol{\beta}_j^{(k+1)})'(y_i - \boldsymbol{X}_i^T \boldsymbol{\beta}_j^{(k+1)})$$

$$\sigma_j^{2(k+1)} = \frac{\sum_{i=1}^{n} p_{ij}^{(k+1)} (y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})(y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})'}{\sum_{i=1}^{n} p_{ij}^{(k+1)}}$$

$$\sigma_j^{2(k+1)} \sum_{i=1}^{n} p_{ij}^{(k+1)} = \sum_{i=1}^{n} p_{ij}^{(k+1)} (y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})(y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})'$$

We know that $\sigma_j^{(2k)}$ are same for all j, so we can have:

$$\sum_{i=1}^{n} \sigma_j^{2(k+1)} p_{ij}^{(k+1)} = \sum_{i=1}^{n} p_{ij}^{(k+1)} (y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})(y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})'$$

$$\sum_{j=1}^{m} \sum_{i=1}^{n} \sigma_j^{2(k+1)} p_{ij}^{(k+1)} = \sum_{j=1}^{m} \sum_{i=1}^{n} p_{ij}^{(k+1)} (y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})(y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})'$$

For each i, we know that $\sum_{j=1}^{m} p_{ij}^{(k+1)} = 1$, and $\sigma_j^{2(k+1)} = \sigma^{2(k+1)}$, so :

$$\sigma^{2(k+1)} \sum_{i=1}^{n} 1 = \sum_{j=1}^{m} \sum_{i=1}^{n} p_{ij}^{(k+1)} (y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})(y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})'$$

$$n\sigma^{2(k+1)} = \sum_{j=1}^{m} \sum_{i=1}^{n} p_{ij}^{(k+1)} (y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})(y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})'$$

$$\sigma^{2(k+1)} = \frac{\sum_{j=1}^{m} \sum_{i=1}^{n} p_{ij}^{(k+1)} (y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})(y_i - \boldsymbol{X}_i^T \beta_j^{(k+1)})'}{n}$$

Finanlly, onlyI_1 contains $\pi_j^{(k+1)}$

$$I_1 = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} log((2\pi)^{-0.5d}\pi_j^{(k+1)})$$

$$I_1 = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)} log((2\pi)^{-0.5d} + \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)}\pi_j^{(k+1)}$$

It suffieces to minimize $\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(k+1)}\pi_j^{(k+1)}$. Let $P_j = \sum_{i=1}^{n} p_{ij}$,$L(\pi_1^{(k+1,)}, \pi_2^{(k+1,)}, \dots, \pi_m^{(k+1)}) = \sum_{j=1}^{m} P_j ln\pi_j^{(k+1)} - \lambda(\sum_{j=1}^{m}\pi_j^{(k+1)} - 1)$ with λ is a Lagrange multiplier. We only need to finde the solution for$L(\pi_1^{(k+1,)}, \pi_2^{(k+1,)}, \dots, \pi_m^{(k+1,)})' = 0$

$$\pi_j^{(k+1)} = \frac{P_j}{\sum_{j=1}^{m} P_j}$$

. For each i $\sum_{j=1}^{m} P_j = 1$, then

$$\pi_j^{(k+1)} = \frac{P_j}{n} = \frac{\sum_{i=1}^{n} p_{ij}}{n}$$

Question 2

Part (a)

As Gamma function is $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,\mathrm{d}x$, the given equation can be expressed as,

$$C\int_0^\infty (2x^{\theta-1} + x^{\theta-1/2})e^{-x}\,\mathrm{d}x = C(\int_0^\infty 2x^{\theta-1}e^{-x}\,\mathrm{d}x + \int_0^\infty x^{\theta-1/2}e^{-x}\,\mathrm{d}x) = C(2\Gamma(\theta) + \Gamma(\theta + \frac{1}{2})) = 1$$

Hence,

$$C = \frac{1}{2\Gamma(\theta) + \Gamma(\theta + \frac{1}{2})}$$

Therefore, g cam be expressed as a mixture of Gamma distributions which is,

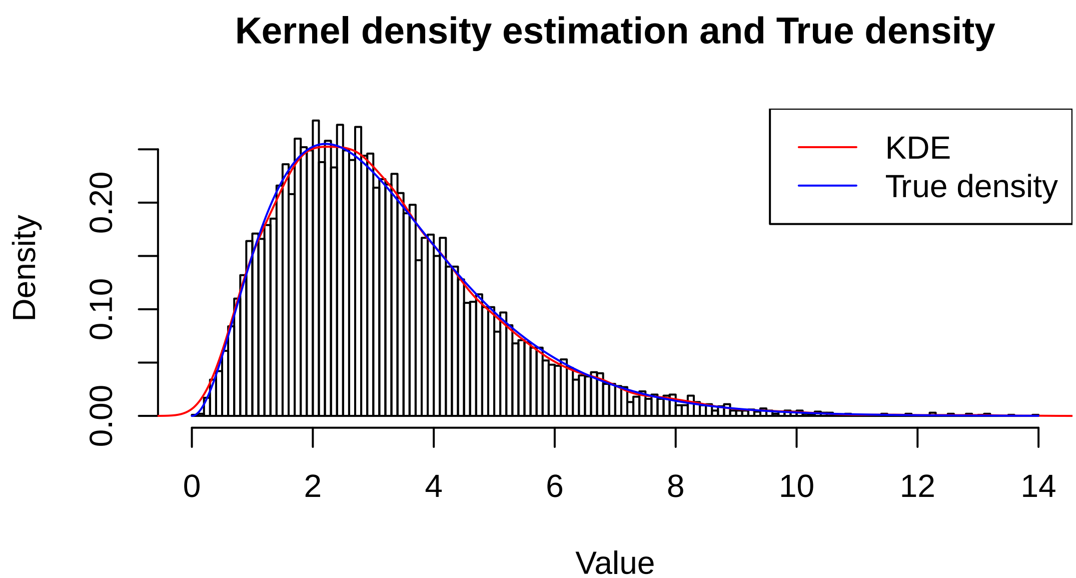
$$\begin{aligned} g(x) &= \frac{1}{2\Gamma(\theta) + \Gamma(\theta + \frac{1}{2})}(2x^{\theta-1}e^{-x} + x^{\theta-\frac{1}{2}}e^{-x}) \\ &= \frac{1}{2\Gamma(\theta) + \Gamma(\theta + \frac{1}{2})}(2\Gamma(\theta)\cdot d(x;\theta,1) + \Gamma(\theta + \frac{1}{2})\cdot d(x;\theta + \frac{1}{2}, 1)) \\ &= \frac{2\Gamma(\theta)}{2\Gamma(\theta) + \Gamma(\theta + \frac{1}{2})}\cdot d(x;\theta,1) + \frac{\Gamma(\theta + \frac{1}{2})}{2\Gamma(\theta) + \Gamma(\theta + \frac{1}{2})}\cdot d(x;\theta + \frac{1}{2}, 1) \end{aligned}$$

where $d(x;\theta,1)$ is density function of Gamma$(\theta,1)$ and $d(x;\theta+\frac{1}{2},1)$ is density function of Gamma$(\theta+\frac{1}{2},1)$. Consequently, $g(x)$ is a mixture of two Gamma distributions Gamma$(\theta,1)$ and Gamma$(\theta+\frac{1}{2},1)$, of which the weights are respectively,

$$p_1 = \frac{2\Gamma(\theta)}{2\Gamma(\theta) + \Gamma(\theta + \frac{1}{2})} \quad and \quad p_2 = \frac{\Gamma(\theta + \frac{1}{2})}{2\Gamma(\theta) + \Gamma(\theta + \frac{1}{2})}$$

Part (b)

By defining $\theta = 3$, a sample of size $n = 10000$ is generated. Kernel Density Estimation of g from the sample and the true density are shown as,

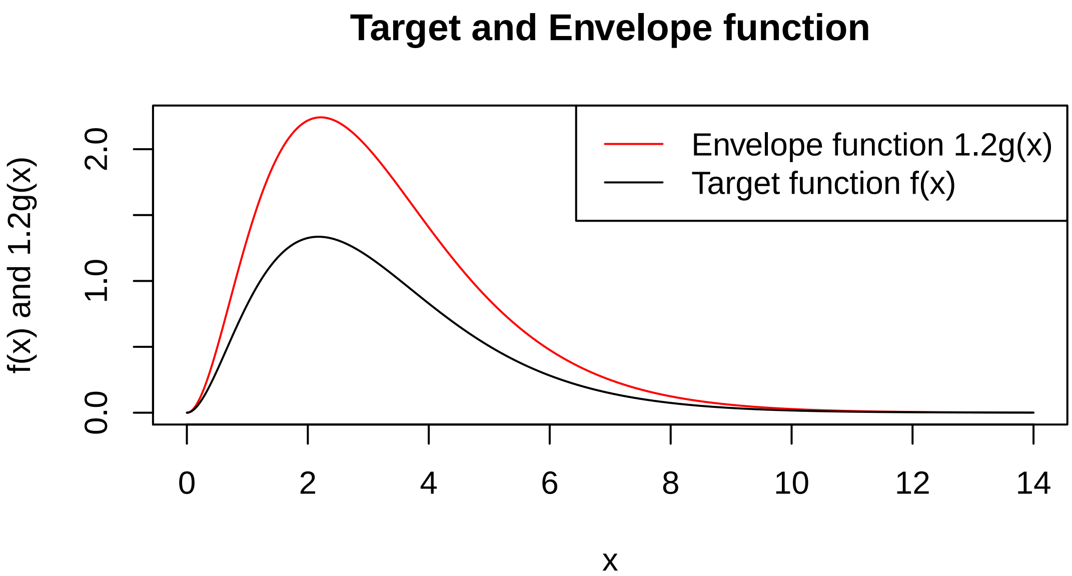


Part (c)

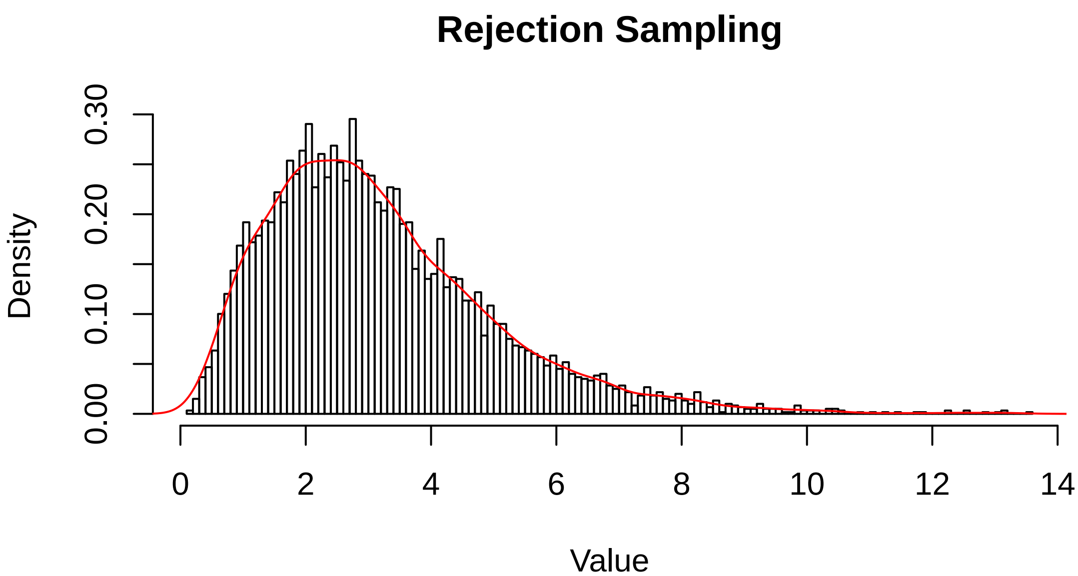
Acoording to the rejection sampling, envelope function $\alpha g(x)$ should be firstly derived and α is calculated as,

$$\alpha = \sup \frac{\mathrm{f(x)}}{\mathrm{g(x)}} = \sup \frac{\sqrt{4 + \mathrm{x}}}{2 + \sqrt{\mathrm{x}}}$$

For programming, the value of α is chosen to be 1.2. Graphs of envelpoe function and target function are presented as,



It can be found that $f(x)$ is absolutely located below the envelope function $1.2g(x)$, which is consistent with the requirement of rejection sampling.
The estimated density of a random sample generated by f is shown as,



Question 3

Part (a)

In this question, rejection sampling method is used so that envelope function should be initially found. According to the requirement, mixture of Beta distributions $g(x)$ is the instrumental function. So, $g(x)$ can be directly seen as the envelope function.
Based on $f(x)$, mixture of Beta distributions can be specified as: since,

$$f(x) \propto \frac{x^{\theta-1}}{1+x^2} + \sqrt{2+x^2}(1-x)^{\beta-1}$$

it can be obtained that when $0 < x < 1$,

$$\frac{x^{\theta-1}}{1+x^2} = \frac{1}{1+x^2} \cdot x^{\theta-1}(1-x)^{1-1} = \frac{B(\theta,1)}{1+x^2} \cdot d(x;\theta,1) < B(\theta,1) \cdot d(x;\theta,1)$$

$$\sqrt{2+x^2}(1-x)^{\beta-1} = \sqrt{2+x^2} \cdot x^{1-1}(1-x)^{\beta-1} = \sqrt{2+x^2}B(1,\beta) \cdot d(x;1,\beta) < 2B(1,\beta) \cdot d(x;1,\beta)$$

where $d(x;\theta,1)$ is density function of Beta$(\theta,1)$ and $d(x;1,\beta)$ is density function of Beta$(1,\beta)$.
Hence, $g(x)$ is expressed as,

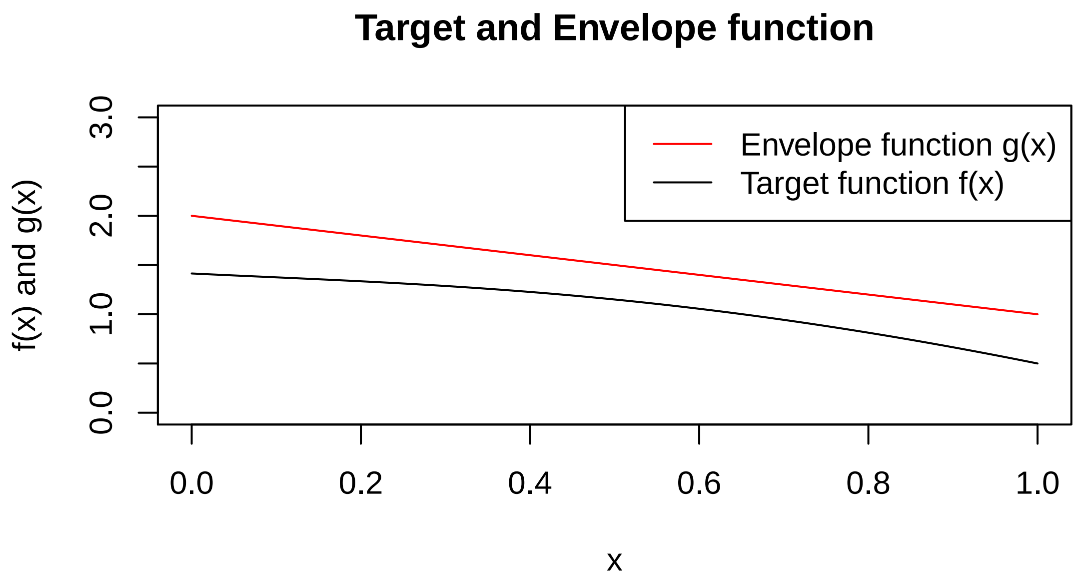
$$g(x) = B(\theta,1) \cdot d(x;\theta,1) + 2B(1,\beta) \cdot d(x;1,\beta) = x^{\theta-1} + 2(1-x)^{\beta-1}$$

and,

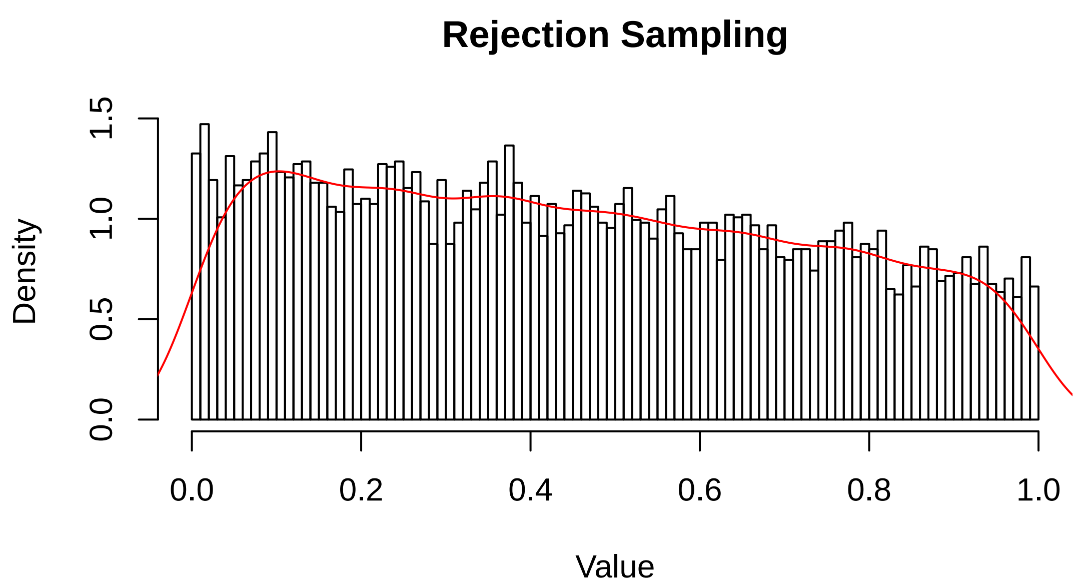
$$g_1(x) = d(x;\theta,1) = \frac{x^{\theta-1}}{B(\theta,1)} \quad and \quad g_2(x) = d(x;1,\beta) = \frac{(1-x)^{\beta-1}}{B(1,\beta)}$$

$$p_1 = B(\theta,1) \quad and \quad p_2 = 2B(1,\beta)$$

Graphs of envelpoe function and target function are presented as,



It can be found that $f(x)$ is absolutely located below the envelope function $g(x)$, which is consistent with the requirement of rejection sampling.
By defining $\theta = 2$ and $\beta = 2$, the estimated density of a random sample generated by f is shown as,

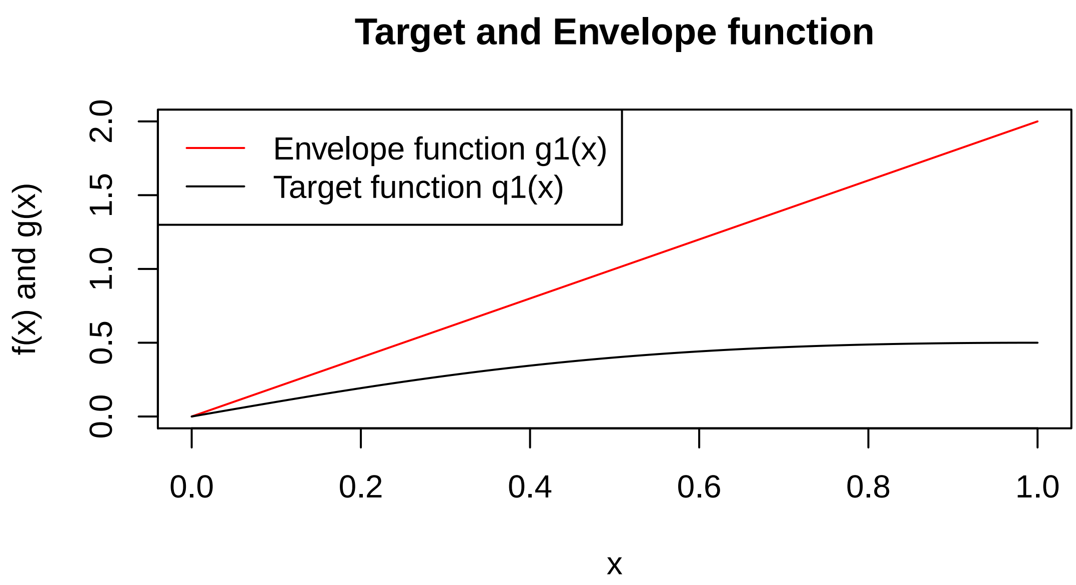


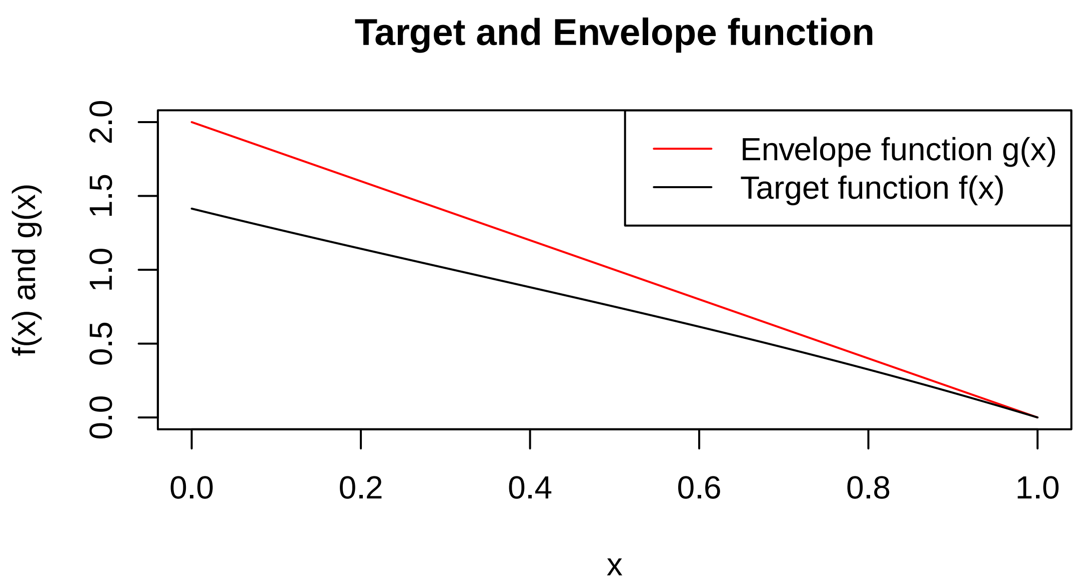
part(b)

When dealing with the two components of f(x), we first define $q_1(x) = \frac{x^{\theta-1}}{1+x^2}$,$q_2(x) = \sqrt{(2+x^2)}(1-x)^{\beta-1}$. So we can have $f(x) \propto q_1(x) + q_2(x)$.

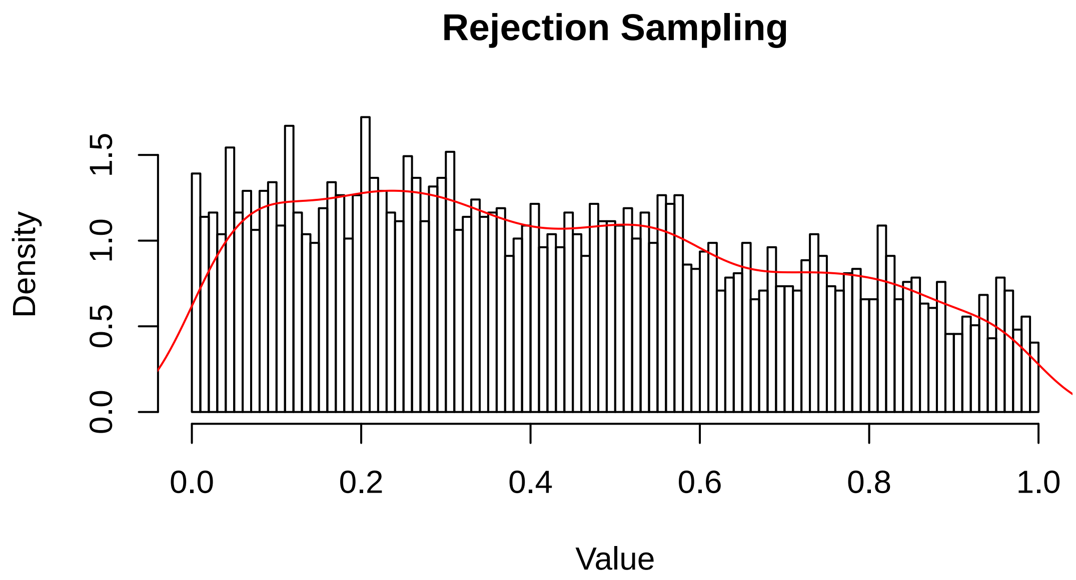
Then we define$g_1(x) = 2 * x^{\theta-1}$ and $g_2(x) = 2 * (1-x)^{\beta-1}$, we can know that $g_1(x), g_2(x)$ is the density function of $Beta(\theta, 1)$ and $Beta(1, \beta)$.

To find the envelope functions, we need to find α_1, α_2 that can make $q_1(x) \leq \alpha_1 g_1(x)$ and $q_2(x) \leq \alpha_2 g_2(x)$. $\alpha_1 \geq sup\frac{q_1(x)}{g_1(x)} = 0.5$ and $\alpha_2 \geq sup\frac{q_2(x)}{g_2(x)} = 0.5\sqrt{2}$. So we define $\alpha_1 = 1.2, \alpha_2 = 1.5$, and also define $\theta = 2$ and $\beta = 2$ as in part (a). Graphs of envelpoe function and target function are presented as,





The estimated density of a random sample generated by f is shown as


Rejection Sampling
Density
0.0
0.5
1.0
1.5
0.0
0.2
0.4
0.6
0.8
1.0
Value