Problem 1

(a) $l(\theta) = x_1 \log(2+\theta) + (x_2+x_3)\log(1-\theta) + x_4 \log\theta + c$

$$\frac{\partial l(\theta)}{\partial\theta} = \frac{x_1}{2+\theta} + \frac{x_2+x_3}{1-\theta}(-1) + \frac{x_4}{\theta}$$

$$= \frac{125}{2+\theta} + \frac{41}{\theta-1} + \frac{33}{\theta} = 0$$

$$\Rightarrow 125\theta(\theta-1) + 41\theta(2+\theta) + 33(\theta-1)(2+\theta) = 0$$

$$\Rightarrow 199\theta^2 - 10\theta - 66 = 0 \quad (\hat{\theta} > 0)$$

$$\hat{\theta} = \frac{10+\sqrt{100+4\times 66\times 199}}{2\times 199} \approx 0.60157$$

(b) E-step.

$$Q(\theta, \theta^{(t)}) = E[l_c(\theta) \mid x; \theta^{(t)}]$$

$$= E[(x_{12}+x_4)\log(\theta) + (x_2+x_3)\log(1-\theta) \mid x; \theta^{(t)}]$$

$$= E(x_{12} \mid x; \theta^{(t)})\log(\theta) + x_4\log(\theta) + (x_2+x_3)\log(1-\theta)$$

$$= x_1 \cdot \left(\frac{\frac{1}{4}\theta^{(t)}}{\frac{1}{2}+\frac{1}{4}\theta^{(t)}}\right)\log(\theta) + x_4\log(\theta) + (x_2+x_3)\log(1-\theta)$$

M-step:

$$\theta^{(t+1)} = \underset{\theta}{\arg\max}\, Q(\theta, \theta^{(t)})$$

$$\frac{\partial Q}{\partial\theta} = \frac{x_1\left(\frac{\frac{1}{4}\theta^{(t)}}{\frac{1}{2}+\frac{1}{4}\theta^{(t)}}\right)}{\theta} + \frac{x_4}{\theta} + \frac{x_2+x_3}{\theta-1} = 0.$$

$$\Rightarrow (\theta-1)\left(x_1\left(\frac{\frac{1}{4}\theta^{(t)}}{\frac{1}{2}+\frac{1}{4}\theta^{(t)}}\right) + x_4\right) + \theta(x_2+x_3) = 0.$$

$$\Rightarrow \theta^{(t+1)} = \frac{x_1\left(\frac{\frac{1}{4}\theta^{(t)}}{\frac{1}{2}+\frac{1}{4}\theta^{(t)}}\right) + x_4}{x_1\left(\frac{\frac{1}{4}\theta^{(t)}}{\frac{1}{2}+\frac{1}{4}\theta^{(t)}}\right) + x_2+x_3+x_4}$$

(c) in (a), $\hat{\theta} \approx 0.60157$

in (b), for $\hat{\theta}_0 = -50, -1, 0, 1, 100$,

all $\hat{\theta} \approx 0.60157$

so the answer in (a) and (b) are the same.

Problem 2:

$$\log L_c = -\log(|\Sigma|) - \Sigma E[(x_i-\mu)^T\Sigma^{-1}(x_i-\mu) \mid y_{obs}, \Sigma]$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \text{—}, x_{1,p+1}, \dots, x_{1,p+q}, x_{1,p+q+1}, \dots, x_{1,p+q+r} \\ x_{21}, \dots, x_{2p}, \text{—}, x_{2,p+q+1}, \dots, x_{2,p+q+r} \end{pmatrix}$$

Let $n = p+q+r$

$$E[x_{i1}|x_{i2}] = \mu_1^{(k)} + \sigma_{12}^{(k)}\sigma_2^{(k)-2}(x_{i2}-\mu_2^{(k)})$$

$$E[x_{i2}|x_{i1}] = \mu_2^{(k)} + \sigma_{12}^{(k)}\sigma_1^{(k)-2}(x_{i1}-\mu_1^{(k)})$$

$$E[x_{i1}^2|x_{i2}] = (\mu_1^{(k)} + \sigma_{12}^{(k)}\sigma_2^{(k)-2}(x_{i2}-\mu_2^{(k)}))^2 + \sigma_1^{(k)2} - \sigma_{12}^{(k)2}\sigma_2^{(k)-2}$$

$$E[x_{i2}^2|x_{i1}] = (\mu_2^{(k)} + \sigma_{12}^{(k)}\sigma_1^{(k)-2}(x_{i1}-\mu_1^{(k)}))^2 + \sigma_2^{(k)2} - \sigma_{12}^{(k)2}\sigma_1^{(k)-2}$$

for i from 1 to p, $E[(x_i-\mu)^T\Sigma^{-1}(x_i-\mu) \mid x_{obs}, \Sigma]$

$$= \underbrace{\left[\frac{\sigma_{12}^{(k)}(x_{i2}-\mu_2^{(k)})}{\sigma_2^{(k)2}} + \mu_1^{(k)} - \mu_1,\ x_{i2}-\mu_2\right]}_{A}\Sigma^{-1}A^T$$

for i from $p+1$ to $p+q$, $E[(x_i-\mu)^T\Sigma^{-1}(x_i-\mu) \mid x_{obs}, \Sigma]$

$$= \underbrace{\left[x_{i1}-\mu_1,\ \frac{\sigma_{12}^{(k)}(x_{i1}-\mu_1^{(k)})}{\sigma_1^{(k)2}} + \mu_2^{(k)} - \mu_2\right]}_{B}\Sigma^{-1}B^T$$

for i from $p+q+1$ to $p+q+r$, $E[(x_i-\mu)^T\Sigma^{-1}(x_i-\mu) \mid x_{obs}, \Sigma]$

$$= (x_i-\mu)^T\Sigma^{-1}(x_i-\mu)$$

$$\text{so } Q = -\frac{n}{2}\log(|\Sigma|) - \frac{1}{2}\sum_{i=1}^{p} A\Sigma^{-1}A^T - \frac{1}{2}\sum_{i=p+1}^{p+q} B\Sigma^{-1}B^T - \frac{1}{2}\sum_{i=p+q+1}^{p+q+r}(x_i-\mu)^T\Sigma^{-1}(x_i-\mu)$$

M-step:

$$\frac{\partial Q}{\partial \mu_1} = \left(\sum_{i=1}^{p}\left[\frac{\sigma_{12}^{(k)}(x_{i2}-\mu_2^{(k)})}{\sigma_2^{(k)2}} + \mu_1^{(k)}\right] + \sum_{i=p+1}^{p+q}[x_{i1}] + \sum_{i=p+q+1}^{p+q+r} x_{i1} - \sum_{i=1}^{p+q+r}\mu_1\right) = 0$$

$$\therefore \mu_1^{(k+1)} = \frac{1}{n}\left[\sum_{i=1}^{p}\left[\frac{\sigma_{12}^{(k)}(x_{i2}-\mu_2^{(k)})}{\sigma_2^{(k)2}} + \mu_1^{(k)}\right] + \sum_{i=p+1}^{p+q+r} x_{i1}\right]$$

$$\frac{\partial Q}{\partial \mu_2} = \left(\sum_{i=1}^{p} x_{i2} + \sum_{i=p+1}^{q+p}\left[\frac{\sigma_{12}^{(k)}(x_{i1}-\mu_1^{(k)})}{\sigma_1^{(k)2}} + \mu_2^{(k)}\right] + \sum_{i=p+q+1}^{p+q+r} x_{i2} - \sum_{i=1}^{p+q+r}\mu_2\right) = 0.$$

$$\therefore \mu_2^{(k+1)} = \frac{1}{n}\left[\sum_{i=1}^{p} x_{i2} + \sum_{i=p+1}^{q+p}\left[\frac{\sigma_{12}^{(k)}(x_{i1}-\mu_1^{(k)})}{\sigma_1^{(k)2}} + \mu_2^{(k)}\right] + \sum_{i=p+q+1}^{p+q+r} x_{i2}\right].$$

$$\frac{\partial Q}{\partial \Sigma^{-1}} = \frac{n}{2}\Sigma - \frac{1}{2}\left[\sum_{i=1}^{p} AA^T + \sum_{i=p+1}^{p+q} BB^T + \sum_{i=p+q+1}^{p+q+r}(x_i-\mu)(x_i-\mu)^T\right] = 0$$

$$\therefore \Sigma^{(k+1)} = \frac{1}{n}\Bigg\{\sum_{i=1}^{p}\left[\frac{\sigma_{12}^{(k)}(x_{i2}-\mu_2^{(k)})}{\sigma_2^{(k)2}} + \mu_1^{(k)} - \mu_1^{(k+1)},\ x_{i2}-\mu_2^{(k+1)}\right]^T\left[\frac{\sigma_{12}^{(k)}(x_{i2}-\mu_2^{(k)})}{\sigma_2^{(k)2}} + \mu_1^{(k)} - \mu_1^{(k+1)},\ x_{i2}-\mu_2^{(k+1)}\right]$$

$$+ \sum_{i=p+1}^{p+q}\left[x_{i1}-\mu_1^{(k+1)},\ \frac{\sigma_{12}^{(k)}(x_{i1}-\mu_1^{(k)})}{\sigma_1^{(k)2}} + \mu_2^{(k)} - \mu_2^{(k+1)}\right]^T\left[x_{i1}-\mu_1^{(k+1)},\ \frac{\sigma_{12}^{(k)}(x_{i1}-\mu_1^{(k)})}{\sigma_1^{(k)2}}\right]$$

$$+ \sum_{i=p+q+1}^{p+q+r}(x_i-\mu^{(k+1)})(x_i-\mu^{(k+1)})^T, \quad \mu^{(k+1)} = \begin{pmatrix} \mu_1^{(k+1)} \\ \mu_2^{(k+1)} \end{pmatrix}$$

Thus we can use this update formular to estimate μ and Σ.

Problem 3

Since $f(x) \propto e^{-x}$, $0<x<2$

Let $f(x) = ce^{-x}$

$$\int_R f(x)\,dx = 1 \Rightarrow \int_0^2 ce^{-x}\,dx = 1$$

$$\Rightarrow -ce^{-x}\Big|_0^2 = 1$$

$$-c(e^{-2}-1) = 1$$

$$c = \frac{1}{1-e^{-2}} = \frac{e^2}{e^2-1}$$

$$\therefore f(x) = \frac{e^{2-x}}{e^2-1}, \quad 0<x<2$$

$$F(x) = \int_0^x f(x)\,dx = \frac{1}{e^2-1}\int_0^x e^{2-x}\,dx = \frac{1}{e^2-1}(-1)\,e^{2-x}\Big|_0^x$$

$$= \frac{e^{2-x}-e^2}{1-e^2}$$

$$= \frac{1-e^{-x}}{1-e^{-2}} \qquad 0<x<2$$

So
$$F(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{1-e^{-x}}{1-e^{-2}} & 0<x<2 \\ 1 & x \geq 1. \end{cases}$$

Inverse transformation:
$$(1-e^{-2})F(x) = 1-e^{-x}$$

$$\Rightarrow 1-(1-e^{-2})F(x) = e^{-x}$$

$$\Rightarrow \log\{1-(1-e^{-2})F(x)\} = -x$$

$$\Rightarrow x = -\log\{1-(1-e^{-2})F(x)\}$$

$$\Leftrightarrow F^{-1}(x) = -\log\{1-(1-e^{-2})x\}$$

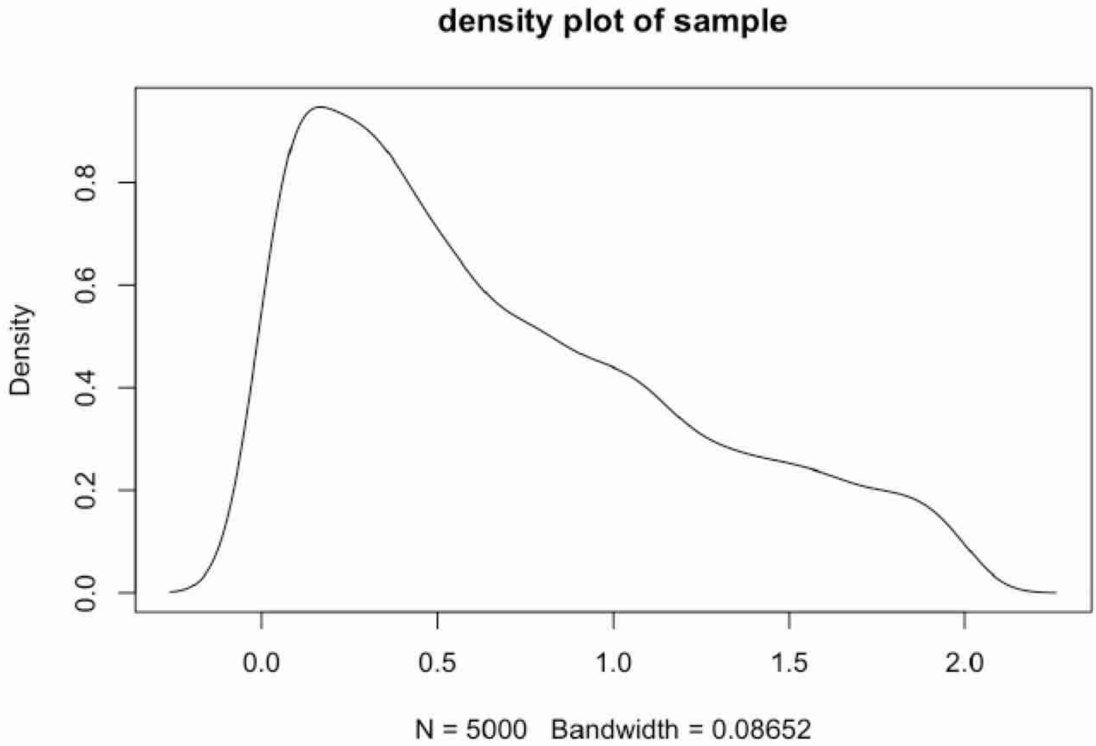
So sample $Y \sim Unif(0,1)$ with 5000 observations

$$\Rightarrow X = -\log\{1-(1-e^{-2})Y\}$$

\# the histogram of sample X and estimated density plot are in R code.

Problem 1: The θ is: 0.6015713

Problem 3:



Problem 4

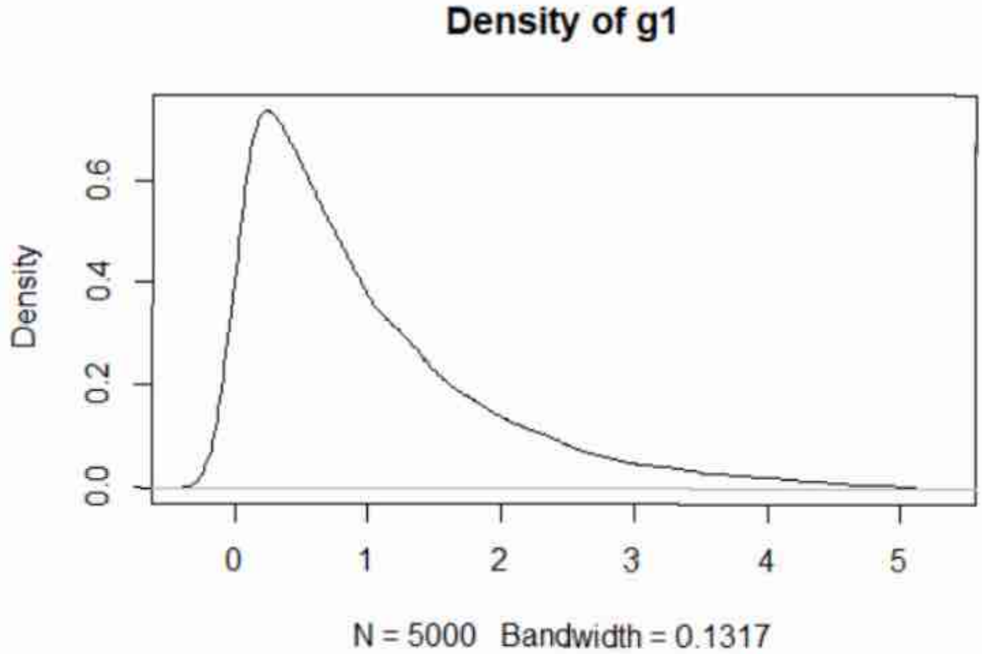
1.

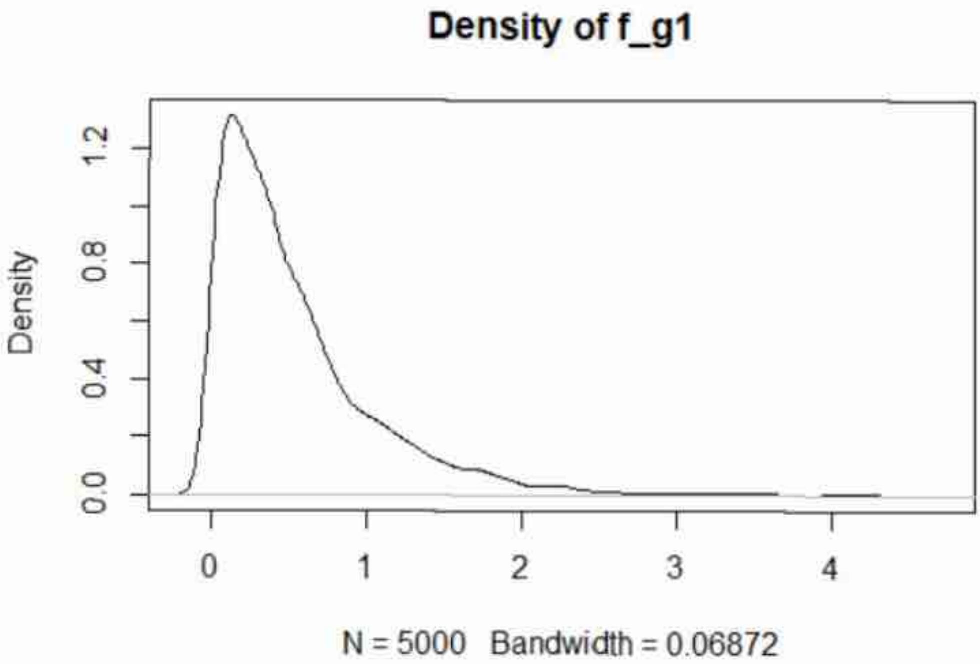
(1) $g_1(x) = e^{-x}$

We can get $X = -\log(1-U)$, $U \sim U(0,1)$;

Then, getting a sampling of 5000 random observations under the condition of range 0 to 5.

If $U \le \frac{q(x)}{\partial g_1(x)}$, $\partial = \sup \frac{q(x)}{g_1(x)} = 1$,we return X as the random variable from f(x).



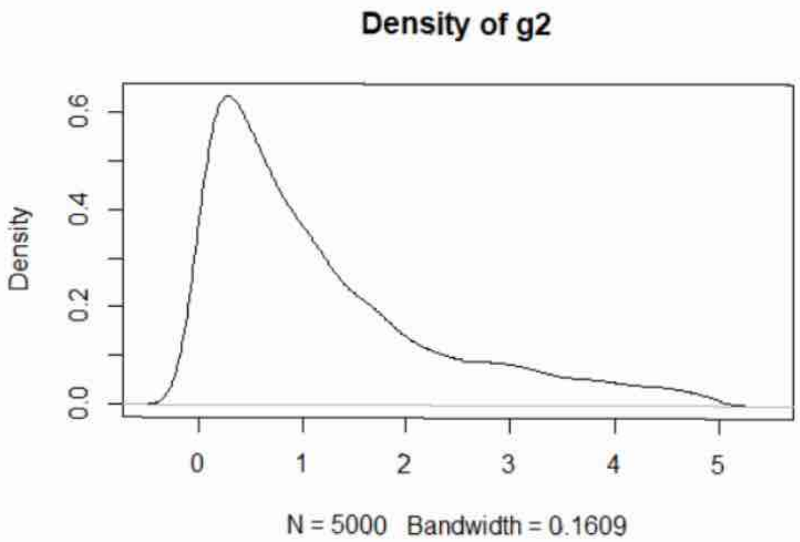


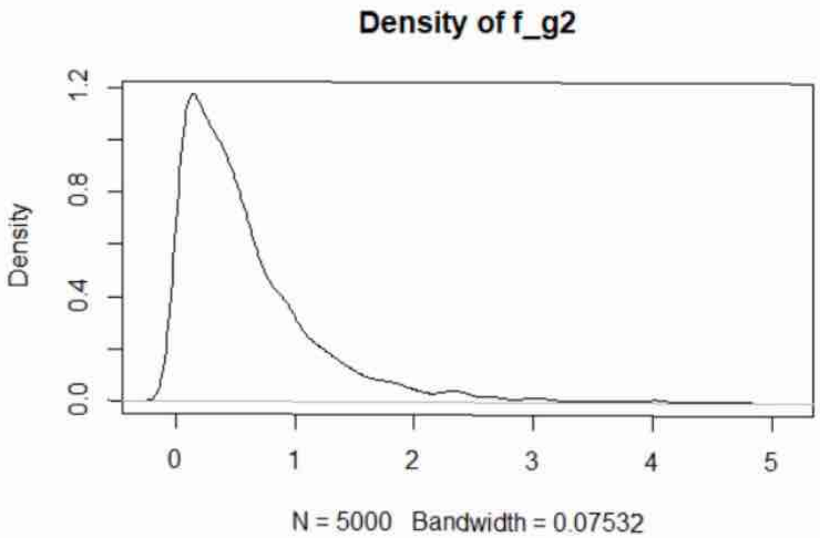
(2) $g_2(x) = \frac{2}{\pi(1+x^2)}$

We can get $X = \tan((\frac{\pi}{2}) * U)$, $U \sim U(0,1)$;

Then, getting a sampling of 5000 random observations under the condition of range 0 to 5.

If , $U \le \frac{q(x)}{\partial g_2(x)}$, $\partial = \sup \frac{q(x)}{g_2(x)}$ (we just choose $\partial = \frac{\pi}{2}$ to have p_a close to 1), we return X as the random variable from f(x).





2.

The speed of sampling: to sample 5000 observations, g1 is faster than g2. And g2 needs more iterations than g1.

The results: the ratio of accepted value of g1 is: 0.6214496

the ratio of accepted value of g2 is: 0.3956271

Problem 5

(a) $g(x) \propto (2x^{\theta-1} + x^{\theta-\frac{1}{2}})e^{-x} \quad x>0$

let $g(x) = c\,(2x^{\theta-1} + x^{\theta-\frac{1}{2}})\,e^{-x}$

$\int_{\mathbb{R}} g(x)\,dx = 1 \Rightarrow \int_0^{+\infty} c(2x^{\theta-1} + x^{\theta-\frac{1}{2}})e^{-x}\,dx = 1$

$\Rightarrow 2c\int_0^{+\infty} x^{\theta-1}e^{-x}\,dx + c\int_0^{+\infty} x^{\theta-\frac{1}{2}}e^{-x}\,dx = 1$

$\Rightarrow 2c\,\Gamma(\theta)\underbrace{\int_0^{+\infty}\frac{1}{\Gamma(\theta)\cdot 1^{\theta}}x^{\theta-1}e^{-\frac{x}{1}}dx}_{=1} + c\,\Gamma(\theta+\tfrac{1}{2})\underbrace{\int_0^{+\infty}\frac{1}{\Gamma(\theta+\frac{1}{2})\cdot 1^{\theta+\frac{1}{2}}}x^{\theta-\frac{1}{2}}e^{-\frac{x}{1}}dx}_{=1} = 1$

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

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

(b) $\frac{1}{\Gamma(\theta)\cdot 1^{\theta}}x^{\theta-1}e^{-\frac{x}{1}}$ is pdf of Gamma $(\theta,1)$ $\rightarrow A$

$\frac{1}{\Gamma(\theta+\frac{1}{2})\cdot 1^{\theta+\frac{1}{2}}}x^{\theta-\frac{1}{2}}e^{-\frac{x}{1}}$ is pdf of Gamma $(\theta+\frac{1}{2},1)$ $\rightarrow B$

$g(x) = \dfrac{1}{2\Gamma(\theta)+\Gamma(\theta+\frac{1}{2})}2x^{\theta-1}e^{-x} + \dfrac{1}{2\Gamma(\theta)+\Gamma(\theta+\frac{1}{2})}x^{\theta-\frac{1}{2}}e^{-x}$

$= mA+nB$ (m, n are constant).

$m = \dfrac{2\Gamma(\theta)}{2\Gamma(\theta)+\Gamma(\theta+\frac{1}{2})}, \quad n = \dfrac{\Gamma(\theta+\frac{1}{2})}{2\Gamma(\theta)+\Gamma(\theta+\frac{1}{2})}$

so $g(x)$ is a mixture of Gamma distributions

(c) 1. sample $U \sim$ uniform $(0,1)$

2. if $U \leq m$, sample $X \sim$ Gamma $(\theta,1)$, return X

otherwise, sample $X \sim$ Gamma $(\theta+\frac{1}{2},1)$.

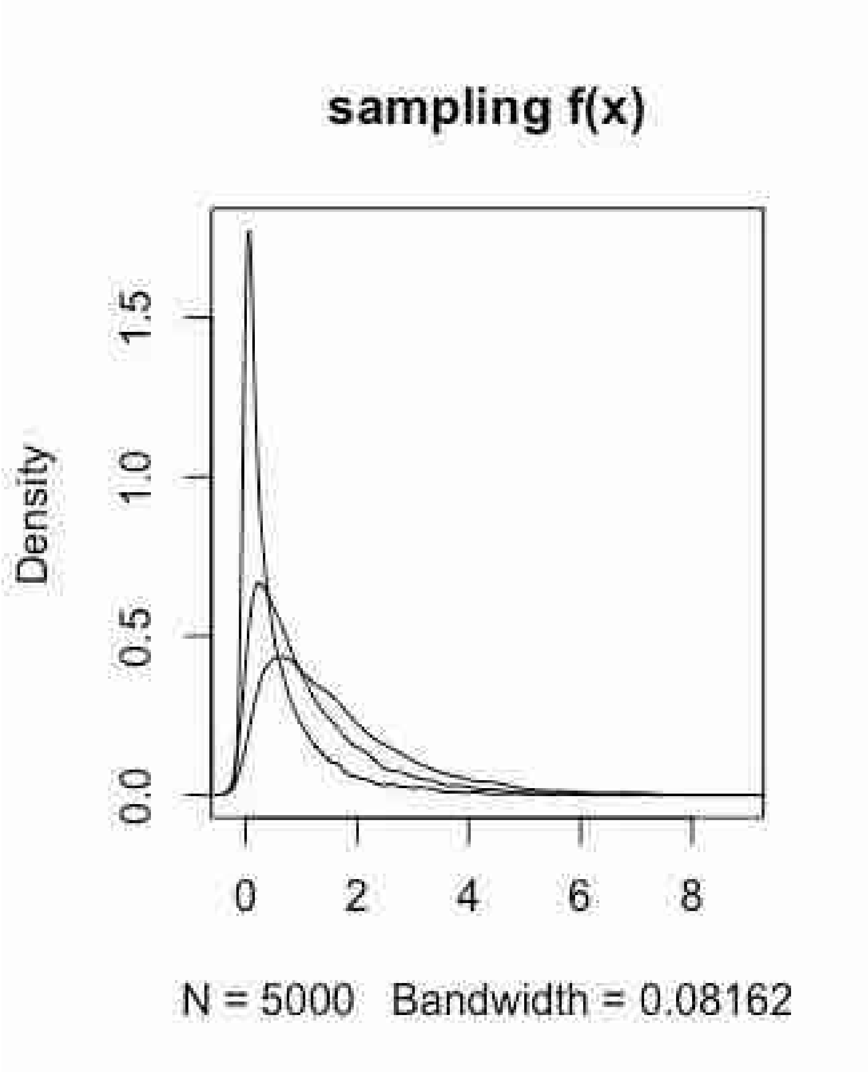
(d) $q(x) = \sqrt{4+x}\,x^{\theta-1}e^{-x}$

$\dfrac{q(x)}{g(x)} = \dfrac{\sqrt{4+x}}{\frac{2}{2\Gamma(\theta)+\Gamma(\theta+\frac{1}{2})} + \frac{1}{2\Gamma(\theta)+\Gamma(\theta+\frac{1}{2})}x^{\frac{1}{2}}} \leq \alpha.$

$\lim\limits_{\substack{x\to\infty \\ x>0}} \dfrac{q(x)}{g(x)} = \left(\dfrac{1}{2\Gamma(\theta)+\Gamma(\theta+\frac{1}{2})}\right)^{-1} \therefore \alpha = 2\Gamma(\theta)+\Gamma(\theta+\tfrac{1}{2}).$

1. sample $X \sim g(x)$, $U \sim$ uniform $(0,1)$

2. if $U > \dfrac{q(x)}{\alpha g(x)}$, then go to step 1, otherwise return X.


sampling f(x)
Density
0.0
0.5
1.0
1.5
0
2
4
6
8
N = 5000   Bandwidth = 0.08162

Problem 6

According to the density function, it will not be valid if a <= -1, because it is not meaningful for $\int_0^1 x^a dx$.

If a > -1, we have $\int_0^1 \int_0^{\sqrt{1-x^2}} x^a y dy dx = \frac{1}{(a+1)(a+3)}$, a valid density.

And also, according to calculation, we can find that when x is from $beta(\alpha+1,1)$ and y is from $beta(2,1)$, and also they are independent from each other, we can get the joint distribution $f(x,y) \propto x^\alpha y$. The calculation is as following:

$$\frac{\Gamma(\alpha+2)}{\Gamma(\alpha+1)\Gamma(1)} x^\alpha (1-x)^{1-1} * \frac{\Gamma(3)}{\Gamma(2)\Gamma(1)} y^{2-1}(1-y)^{1-1} = (\alpha+1)x^\alpha y \propto x^\alpha y$$

$$\text{Let } g(x,y) = \frac{\Gamma(\partial+2)}{\Gamma(\partial+1)\Gamma(1)} x^{\partial} (1-x)^0 \cdot \frac{\Gamma(3)}{\Gamma(2)\Gamma(1)} y^1 (1-y)^0$$

$$\partial = \sup \frac{q(x,y)}{g(x,y)} = \sup \frac{x^{\partial} y}{f(x,y)/k} = \sup \frac{\frac{f(x,y)}{(\partial+1)(\partial+3)}}{f(x,y)/k} = \frac{k}{(\partial+1)(\partial+3)}$$

$$\Rightarrow \frac{q(x,y)}{\partial \cdot g(x,y)} = \frac{x^{\partial} y}{\frac{k}{(\partial+1)(\partial+3)} \cdot \frac{f(x,y)}{k}} = 1 \qquad \therefore \forall u \le \frac{q(x,y)}{\partial \cdot g(x,y)}$$

Thus, we can get the rejection algorithm:

1、 sample X ~ $beta(\alpha+1,1)$, Y~ $beta(2,1)$, independently.

2、 if $X^2+Y^2>1$, then go to step 1, otherwise return (X,Y).