Homework3

Ruoyu_Jia & Xi_Chen

2/24/2018

1

$$\phi(y_i - x_i^\top \beta_j; 0, \delta^2) = \frac{1}{\sqrt{2\pi}\ \delta} e^{-\frac{(y_i - x_i^\top \beta_j)^2}{2\delta^2}}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{k+1}\{\log \pi_i + \log \phi(y_i - x_i^\top \beta_i; 0, \delta^2)\} = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{k+1}\{\log(\frac{1}{\sqrt{2\pi}\ \delta} e^{-\frac{(y_i - x_i^\top \beta_j)^2}{2\delta^2}})\}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{k+1} \log(\frac{1}{\sqrt{2\pi}}\pi_j) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{k+1}\log(\delta^2) - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{k+1}\frac{(y_i - x_i^\top \beta_j)^2}{\delta^2} = I_1 - \frac{1}{2}I_2 - \frac{1}{2}I_3$$

Firstly, only I_3 contains β_j , $j = 1, ...m.$~ We only need β_j to minimize each I_{ij}

$$I_{ij} = \sum_{i=1}^{n} p_{ij}^{k+1}\ \frac{(y_i - x_i^\top \beta_i)^2}{\delta^2}$$

$$\frac{\partial I_{ij}}{\partial \beta_j} = -\frac{2\sum_{i=1}^{n} p_{ij}^{k+1} x_i(y_i - x_i^\top \beta_j)}{\delta^2} = 0$$

$$\sum_{i=1}^{n} p_{ij}^{k+1} x_i x_i^\top \beta_j = \sum_{i=1}^{n} p_{ij}^{k+1} x_i yi$$

$$\beta_j = (\sum_{i=1}^{n} x_i x_i^\top p_{ij}^{k+1})^{-1}(\sum_{i=1}^{n} x_i p_{ij}^{k+1} y_i)$$

Next, only I_2 and I_3 contains δ^2, we need δ^2 to minimize $(I_2 + I_3)$

$$\frac{\partial(I_2 + I_3)}{\partial(\delta^2)} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{k+1}}{\delta^2} - \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}(y_i - x_i^\top \beta_j)^2}{\delta^4} = 0$$

$$\delta^2 = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}(y_i - x_i^\top \beta_j)^2}{\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{k+1}} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{m}(y_i - x_i^\top \beta_j)^2}{n}$$

Lastly, only I_1 contains π_j

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

So we need to minimize $\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{k+1}$. Note that $\pi_1 + \pi_2 + ... + \pi_m = 1$,

$$L = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{k+1} - \lambda(\sum_{j=1}^{m}\pi_j - 1)$$

λ is a Lagrange multipter,

$$\pi_j = \frac{\sum_{i=1}^{n} p_{ij}^{k+1}}{\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{k+1}} = \frac{\sum_{i=1}^{n} p_{ij}^{k+1}}{n}$$

2(a)

$$\int_0^\infty 2X^{\theta-1}e^{-x}\ dx = 2\Gamma(\theta)\int_0^\infty \frac{X^{\theta-1}e^{-x}}{\Gamma(\theta)}\ dx = 2\Gamma(\theta)Gamma(\theta,1)$$

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

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

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

$$\begin{aligned} g(x) &= \frac{(2X^{\theta-1}\ X^{\theta-1/2})e^{-x}}{2\Gamma(\theta)Gamma(\theta,1) + \Gamma(\theta+1/2)Gamma(\theta+1/2,1)} \\ &= 2C\Gamma(\theta)\frac{X^{\theta-1}e^{-x}}{\Gamma(\theta)} + C\Gamma(\theta+1/2)\frac{X^{\theta-1/2}e^{-x}}{\Gamma(\theta+1/2)} \\ &= 2C\Gamma(\theta)g_1(x) + C\Gamma(\theta+1/2)g_2(x) \end{aligned}$$

Therefore, g is a mixture of Gamma distributions, which are $Gamma(\theta,1)$ and $Gamma(\theta+1/2,1)$. the weights of them are $\frac{2\Gamma(\theta)}{2\Gamma(\theta)+\Gamma(\theta+1/2)}$ and $\frac{2\Gamma(\theta+1/2)}{2\Gamma(\theta)+\Gamma(\theta+1/2)}$

2(b)

step 1

sample k from {1,2}

the prob of k=1 is $2 * \Gamma(\theta) * C$

the prob of k=2 is $\Gamma(\theta+1/2) * C$

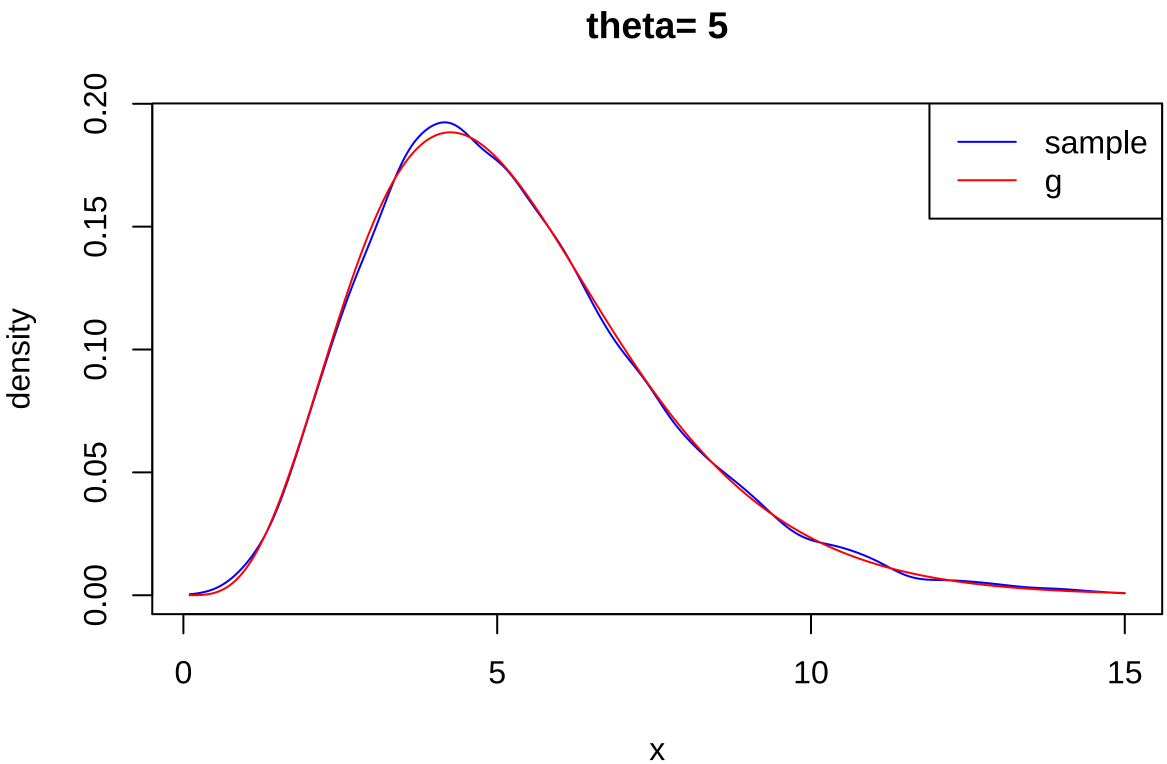
step 2

if k=1, get 1 sample from $Gamma(1,\theta)$

if k=2, get 1 sample from $Gamma(1,\theta+1/2)$

step 3

repeat step1 and step2 10000 times



2(c)

$f(x) = \alpha g(x)$, $f(x)/g(x) = \alpha = \frac{\sqrt{4+x}\ x^{\theta-1}e^{-x}}{(2x^{\theta-1}+x^{\theta-1/2})e^{-x}} = \frac{\sqrt{4+x}}{2+\sqrt{x}} = F(x)$

$$Let \quad F'(x) = 0$$

Then we have $x = 4$. Therefore, $\alpha = \sqrt{2}/2$ and accepting probability is $\frac{\sqrt{4+x}}{\sqrt{2}/2(2+\sqrt{x})}$

step 1

get 1 sample from function g

step 2

u ~ unif(0,1)

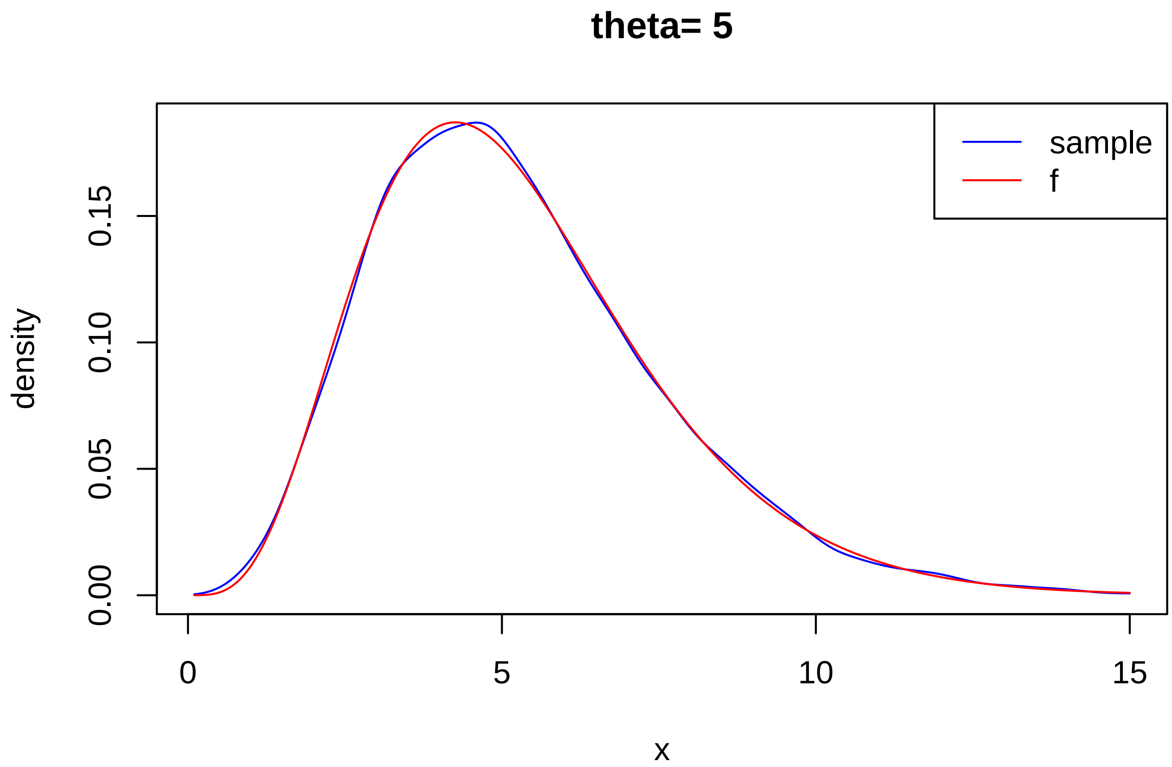
if u < $f(x)/\alpha g(x)$

accept it, go to step3

otherwise, reject it, return step1

step 3

repeat step1 and step2 10000 times



3(a)

Let $g(x) = (1-\lambda)\theta x^{\theta-1} + \lambda\beta(1-x)^{\beta-1} = (1-\lambda)g_\theta(x) + \lambda g_\beta(x),\ 0 < \lambda < 1$. Here $(1-\lambda)$ *and* λ can be considered as the weight of g_θ *and* g_β.

$$\alpha = \sup\left(\frac{\frac{x^{\theta-1}}{1+x^2} + \sqrt{2+x^2}(1-x)^{\beta-1}}{(1-\lambda)\theta x^{\theta-1} + \lambda\beta(1-x)^{\beta-1}}\right)$$

As the expression of α is complicated, we can use scaling method to estimate the value of α, which is

$$\alpha \le \sup\left(\frac{\frac{1}{1+x^2} + \sqrt{2+x^2}(1-x)^{\beta-1}x^{1-\theta}}{\lambda\beta(1-x)^{\beta-1}x^{1-\theta}}\right)$$

$$= \sup\left(\frac{\frac{1}{1+x^2}}{\lambda\beta(1-x)^{\beta-1}x^{1-\theta}} + \frac{\sqrt{2+x^2}}{\lambda\beta}\right)$$

$$= \frac{1}{\lambda\beta}\sup\left(\frac{(1-x)^{1-\beta}x^{\theta-1}}{1+x^2} + \sqrt{2+x^2}\right)$$

$$\le \frac{1}{\lambda\beta}\sup\left((1-x)^{1-\beta}x^{\theta-1} + \sqrt{2+x^2}\right)$$

$$= \frac{\sqrt{3}}{\lambda\beta}$$

Therefore, the accepting probability is $\frac{\frac{x^{\theta-1}}{1+x^2} + \sqrt{2+x^2}(1-x)^{\beta-1}}{\frac{\sqrt{3}}{\lambda\beta}((1-\lambda)\theta x^{\theta-1} + \lambda\beta(1-x)^{\beta-1})}$

step 1

get 1 sample from function g

step 2

u ~ unif(0,1)

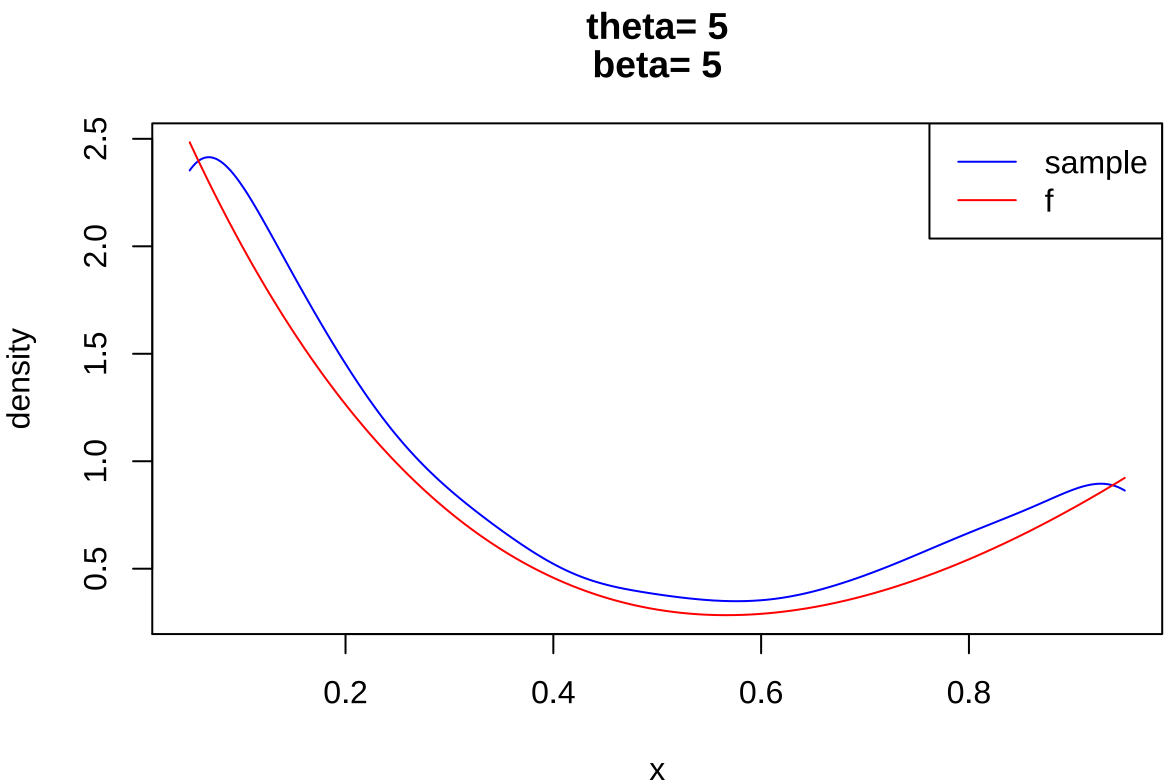
if u < $f(x)/\alpha g(x)$

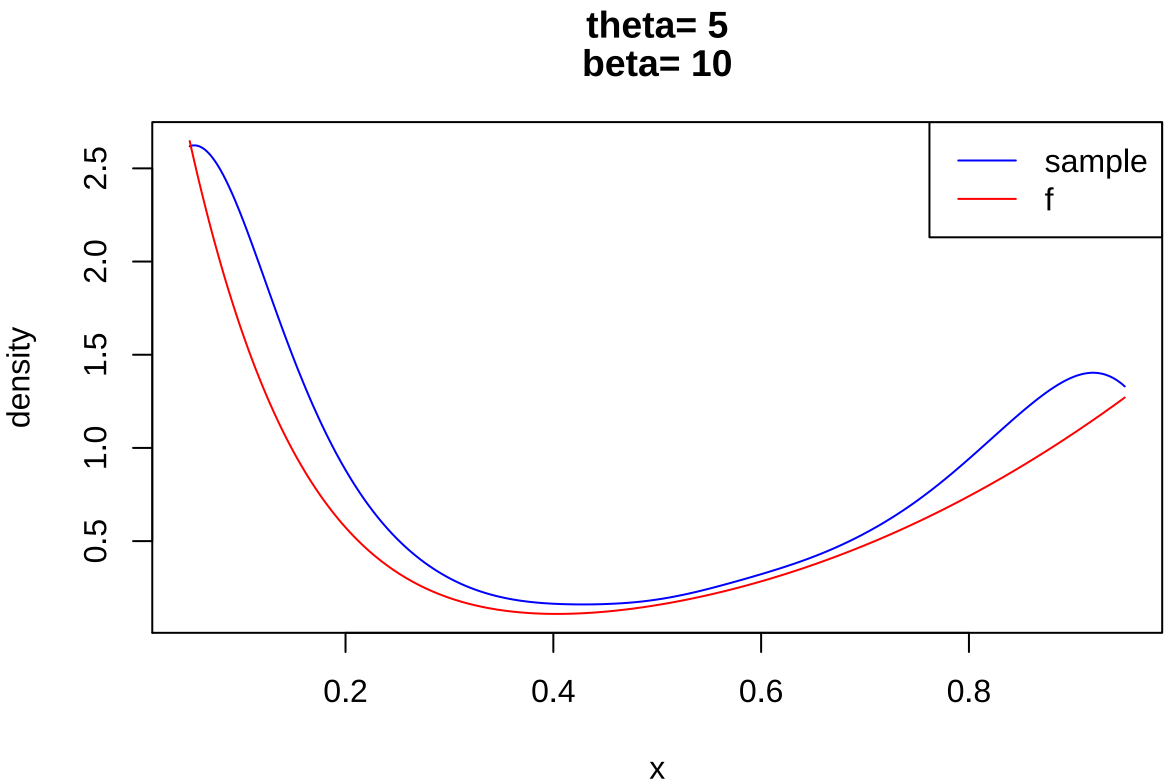
accept it, go to step3

otherwise, reject it, return step1

step 3

repeat step1 and step2 10000 times





3(b)

$$f(x) \propto \frac{x^{\theta-1}}{1+x^2} + \sqrt{2+x^2}(1-x)^{\beta-1} = q_1(x) + q_2(x), 0 < x < 1$$

$$q_1(x) = \frac{x^{\theta-1}}{1+x^2} \leq \alpha_1 g_1(x) = \alpha \frac{x^{\theta-1}}{B(\theta,1)}$$

$$\alpha = \sup(\frac{B(\theta,1)}{1+x^2}) = B(\theta,1)$$

So the accepting probability is $\frac{1}{1+x^2}$

$$q_2(x) = \sqrt{2+x^2}\ (1-x)^{\beta-1} \leq \alpha_2 g_2(x) = \alpha_2 \frac{(1-x)^{\beta-1}}{B(1,\beta)}$$

$$\alpha_2 = \sup(B(1,\beta)\sqrt{2+x^2}) = \sqrt{3}B(1,\beta)$$

So the accepting probability is $\frac{\sqrt{2+x^2}}{\sqrt{3}}$

step 1

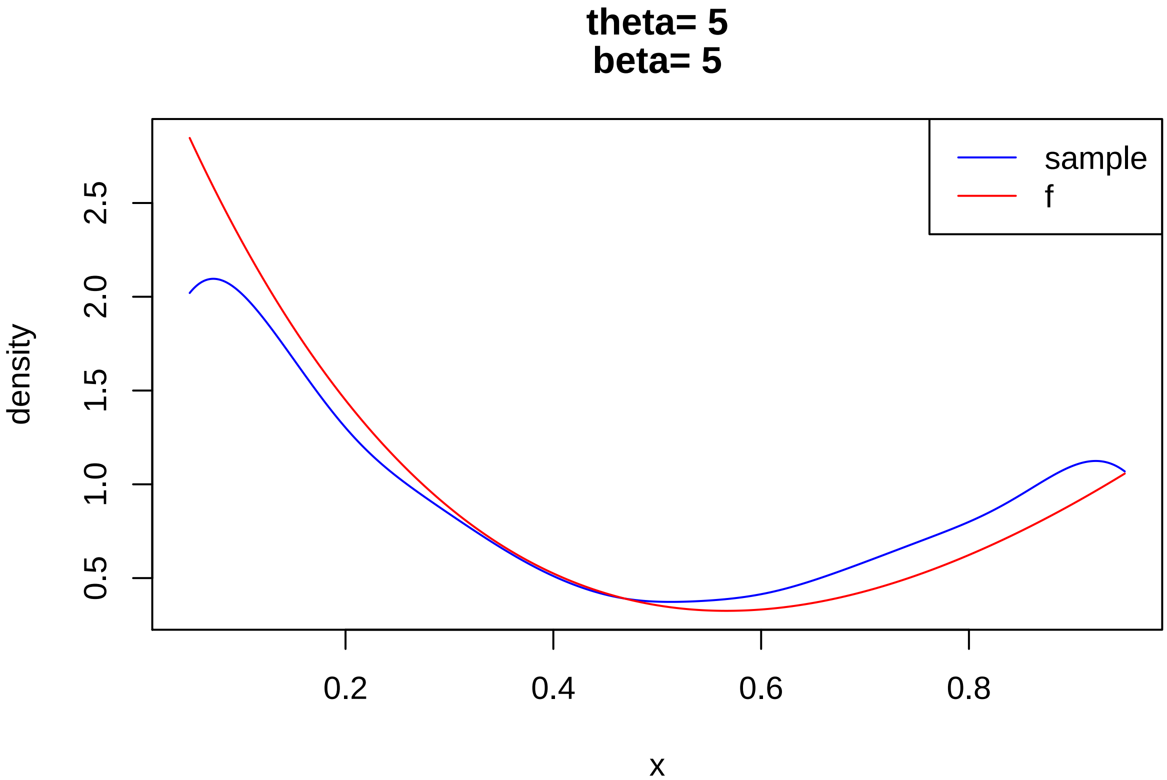
sample k from {1,2} with probabilities $p_k \propto \alpha_k$

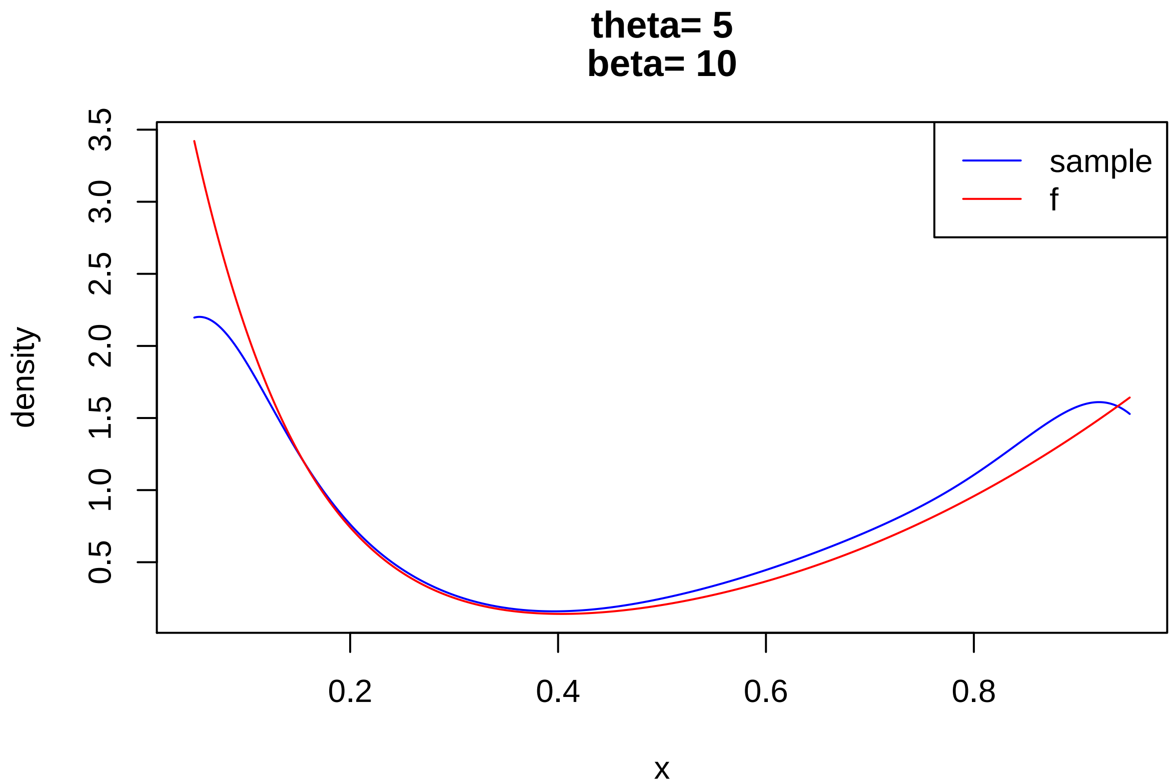
step 2

get 1 sample from $g_k(x)$
u ~ unif(0,1) if u < $f_k(x)/\alpha g_k(x)$
accept it
otherwise, reject it, repeat step2 until 1 sample is accepted

step 3

repeat step1 and step2 10000 times




theta= 5
beta= 10
sample
f
density
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.2
0.4
0.6
0.8
x