

STAT 5443-Computational Statistics-Midterm I

Md Shahbaz Alam; ID: 011015340

2023-09-28

Solution: Q1

Given that,

$x_1, x_2, \dots, x_n \stackrel{\text{iid}}{\sim} \text{Bernoulli}(\theta)$. So, the PDF of x is,

$$f(x) = \theta^x (1-\theta)^{1-x} \quad \text{where } x \in \{0, 1\} \text{ and } 0 < \theta < 1$$

We know the mean of Bernoulli random variable x is,

$$E[x] = \theta \quad \text{and} \quad \text{var}(\theta) = \theta(1-\theta)$$

a) The likelihood function can be written as,

$$\begin{aligned} L(\theta | x_1, \dots, x_n) &= \prod_{i=1}^n f(x_i | \theta) \\ &= \prod_{i=1}^n \theta^{x_i} (1-\theta)^{1-x_i} \\ &= \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i} \end{aligned}$$

Now, by taking the natural log on both side, we will get the loglikelihood,

$$\begin{aligned} \ell(\theta) &= \ln \left[\theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i} \right] \\ &= \sum_{i=1}^n x_i \ln(\theta) + (n - \sum_{i=1}^n x_i) \ln(1-\theta) \end{aligned}$$

Now, the first derivative of $\ell(\theta)$ is,

$$\ell'(\theta) = \frac{\partial}{\partial \theta} \ell(\theta) = \frac{\sum_{i=1}^n x_i}{\theta} - \frac{n - \sum_{i=1}^n x_i}{1-\theta}$$

$$\ell'(\theta) = \frac{\sum_{i=1}^n x_i}{\theta} - \frac{n - \sum_{i=1}^n x_i}{1-\theta}$$

then, the 2nd derivative of $\ell(\theta)$ will be,

$$\begin{aligned}\ell''(\theta) &= \frac{\partial}{\partial \theta} [\ell'(\theta)] = \frac{\partial}{\partial \theta} \left[\frac{\sum_{i=1}^n x_i}{\theta} - \frac{n - \sum_{i=1}^n x_i}{1-\theta} \right] \\ &= - \frac{\sum_{i=1}^n x_i}{\theta^2} - (-1) \frac{n - \sum_{i=1}^n x_i}{(1-\theta)^2} \quad (-1) \quad [\text{applying chain rule}] \\ &= - \frac{\sum_{i=1}^n x_i}{\theta^2} - \frac{n - \sum_{i=1}^n x_i}{(1-\theta)^2}\end{aligned}$$

Now,

Fisher information $I(\theta) = E[-\ell''(\theta)]$

$$\therefore I(\theta) = E \left[\frac{\sum_{i=1}^n x_i}{\theta^2} + \frac{n - \sum_{i=1}^n x_i}{(1-\theta)^2} \right]$$

$$= \frac{E[\sum_{i=1}^n x_i]}{\theta^2} + \frac{n - E[\sum_{i=1}^n x_i]}{(1-\theta)^2}$$

$$= \frac{n\theta}{\theta^2} + \frac{n - n\theta}{(1-\theta)^2}$$

[since $E[\sum x_i] = \sum E[x_i]$
and x_i 's are iid
so, $\sum E[x_i] = n\theta$]

$$= \frac{n}{\theta} + \frac{n}{1-\theta}$$

Thus, $I(\theta) = \frac{n - n\theta + n\theta}{\theta(1-\theta)} = \frac{n}{\theta(1-\theta)}$

b) Given that,

Prior on θ is Beta (α, β)

$$\therefore f(\theta) = \frac{1}{\text{Beta}(\alpha, \beta)} \cdot \theta^{\alpha-1} (1-\theta)^{\beta-1} ; 0 < \theta < 1 \text{ and } \alpha, \beta > 0$$

Since, the normalizing part $\frac{1}{\text{Beta}(\alpha, \beta)}$ is parameter (θ) free,

we can write $f(\theta)$ as,

$$f(\theta) \propto \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

The likelihood is we already get from part (a) which is,

$$L(\theta | x_1, \dots, x_n) = \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i}$$

we know that,

posterior \propto likelihood \times prior

$$\therefore f(\theta | \text{data}) \propto L(\theta | \text{Data}) \cdot f(\theta)$$

$$\propto \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i} \cdot \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

$$\propto \theta^{(\alpha + \sum_{i=1}^n x_i) - 1} (1-\theta)^{(\beta + n - \sum_{i=1}^n x_i) - 1}$$

As we can see, the above expression is the kernel of

a Beta $\left(\alpha + \sum_{i=1}^n x_i, \beta + n - \sum_{i=1}^n x_i \right)$ PDF.

Thus, the posterior distribution of θ follows,

$$\theta | \text{data} \sim \text{Beta} \left(\alpha + \sum_{i=1}^n x_i, \beta + n - \sum_{i=1}^n x_i \right).$$

Now, the posterior mean,

$$E[\theta | \text{data}] = \frac{\alpha + \sum_{i=1}^n x_i}{\alpha + \sum_{i=1}^n x_i + \beta + n - \sum_{i=1}^n x_i}$$

$$= \frac{\alpha + \sum_{i=1}^n x_i}{\alpha + \beta + n}$$

$$\left[\text{if } x \sim \text{Beta}(a, b) \text{ then } E[x] = \frac{a}{a+b} \right]$$

solution : Q2 :

(1) Given that,

$$f(x) = \begin{cases} cx^3 & \text{if } 0 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

To be a valid density, the integral of $f(x)$ within the range $0 < x < 2$ must be equal to 1.

Now,

$$\int_x f(x) dx = 1$$

$$\Rightarrow \int_0^2 cx^3 dx = 1$$

$$\Rightarrow c \left[\frac{x^4}{4} \right]_{x=0}^{x=2} = 1$$

$$\Rightarrow c [4 - 0] = 1$$

$$\Rightarrow 4c = 1$$

$$\therefore c = \frac{1}{4}$$

so,

$$f(x) = \begin{cases} \frac{1}{4}x^3 & \text{if } 0 < x < 2 \\ 0, & \text{otherwise} \end{cases}$$

solution : Q2 (2) :

To perform inverse transform sampling, the following steps needs to be followed :

i) first, we need to calculate the cumulative distribution function (CDF) $F(x)$ for this given $f(x)$.

ii) then, find the inverse CDF $F^{-1}(u)$, where u is a random sample from Uniform $(0,1)$

iii) generate 10,000 random samples u from Uniform $(0,1)$

iv) finally, apply the inverse CDF to each u to get the corresponding values of x where x has PDF $f(x)$.

In other words, $F^{-1}(u_1), \dots, F^{-1}(u_{10,000}) \sim f(x)$.

Now, lets get the CDF first.

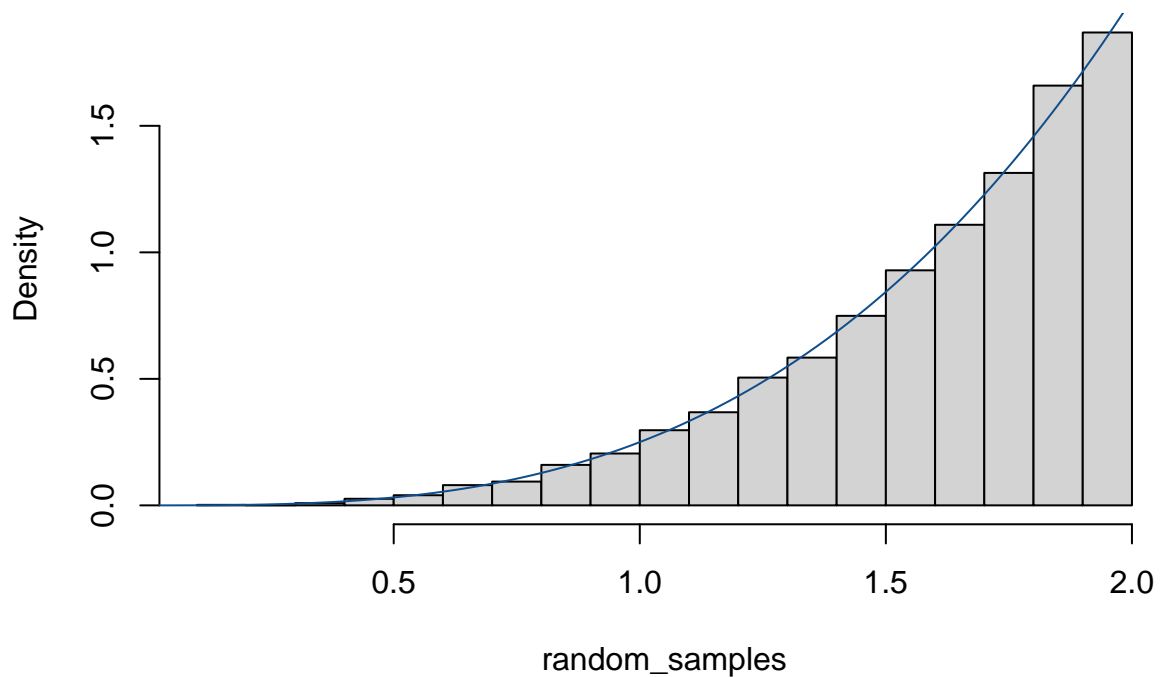
$$\begin{aligned} F(x) &= \int_0^x f(t) dt = \int_0^x \frac{1}{4} t^3 dt \\ &= \int_0^x \frac{1}{4} t^3 dt = \frac{1}{4} \cdot \left[\frac{t^4}{4} \right]_{t=0}^{t=x} \\ &= \frac{1}{16} x^4 \end{aligned}$$

Question 2: Inverse transform Sampling

```
#----- Problem 2 (2): Inverse Transform Sampling
# Defining the inverse CDF:  $2 * u^{(1/4)}$ 
inverse_cdf <- function(u) {
  return(2 * u (1/4))
}

# Generate 10,000 random samples using ITS
set.seed(729)
n_samples <- 10000
u <- runif(n = n_samples, 0, 1)
random_samples <- inverse_cdf(u)
hist(random_samples, probability = T)
curve(expr = .25*(x3), from = 0, to = 2, add = T, col="dodgerblue4")
```

Histogram of random_samples



solution Q3:

given that,

$$f(x, y) = \begin{cases} c(y\sqrt{x} + x\sqrt{y}) & \text{if } 0 < x < 2, 0 < y < 2 \\ 0, & \text{otherwise.} \end{cases}$$

(1) To be a valid density, $f(x, y)$ must integrate to 1 within the range of $0 < x < 2$ and $0 < y < 2$.

now,

$$\int_y \int_x f(x, y) dx dy = 1$$

$$\Rightarrow \int_0^2 \int_0^2 c(y\sqrt{x} + x\sqrt{y}) dx dy = 1$$

$$\Rightarrow c \left[\int_0^2 \left[\frac{2}{3} \cdot y\sqrt{x^3} + \frac{1}{2} x^2\sqrt{y} \right]_{x=0}^{x=2} dy \right] = 1$$

$$\Rightarrow c \left[\int_0^2 \left(\frac{2\sqrt{8}}{3} y + 2\sqrt{y} \right) dy \right] = 1$$

$$\Rightarrow c \left[\frac{2\sqrt{8}}{3} \cdot \frac{1}{2} y^2 + 2 \cdot \frac{2}{3} \sqrt{y^3} \right]_{y=0}^{y=2} = 1$$

$$\Rightarrow c \left[\frac{4\sqrt{8}}{3} + \frac{4\sqrt{8}}{3} \right] = 1$$

$$\Rightarrow \frac{16\sqrt{2}}{3} c = 1$$

$$\text{Thus, } c = \frac{3}{16\sqrt{2}}$$

$$\therefore f(x, y) = \begin{cases} \frac{3}{16\sqrt{2}} (y\sqrt{x} + x\sqrt{y}) & \text{if } 0 < x < 2, 0 < y < 2 \\ 0, & \text{otherwise} \end{cases}$$

(2) The conditional density $f(x|y) = \frac{f(x, y)}{f(y)}$

Now,

$$f(y) = \int_x f(x, y) dx$$

$$= \int_0^2 \frac{3}{16\sqrt{2}} (y\sqrt{x} + x\sqrt{y}) dx$$

$$= \frac{3}{16\sqrt{2}} \cdot \left[\frac{2}{3} y\sqrt{x^3} + \frac{1}{2} x^2\sqrt{y} \right]_{x=0}^{x=2}$$

$$= \frac{3}{16\sqrt{2}} \left[\frac{2\sqrt{8}}{3} y + 2\sqrt{y} \right]$$

$$= \frac{1}{4} y + \frac{3}{8\sqrt{2}} \sqrt{y}$$

$$\therefore f(y) = \frac{1}{4} \left[y + \frac{3}{2\sqrt{2}} \sqrt{y} \right], \quad 0 < y < 2$$

Now,

$$f(x|y) = \frac{f(x, y)}{f(y)}$$

$$= \frac{3}{4\sqrt{2}} \frac{\cancel{\frac{3}{16\sqrt{2}}} (y\sqrt{x} + x\sqrt{y})}{\cancel{\frac{3}{16\sqrt{2}}} \left(y + \frac{3}{2\sqrt{2}} \sqrt{y} \right)}$$

$$0 < x < 2, 0 < y < 2$$

Now, $f(y|x)$ can be written directly using the symmetry or can be calculated as follows. Both results in same expression.

$$f(y|x) = \frac{f(x,y)}{f(x)}$$

$$\therefore f(x) = \int_0^2 f(x,y) dy$$

$$= \int_0^2 \frac{3}{16\sqrt{2}} (y\sqrt{x} + x\sqrt{y}) dy$$

$$= \frac{3}{16\sqrt{2}} \left[\frac{1}{2} y^2 \sqrt{x} + \frac{2}{3} x \sqrt{y^3} \right]_{y=0}^{y=2}$$

$$= \frac{3}{16\sqrt{2}} \left[2\sqrt{x} + \frac{2\sqrt{8}}{3} x \right]$$

$$= \frac{3}{8\sqrt{2}} \sqrt{x} + \frac{1}{4} x$$

$$= \frac{1}{4} \left[x + \frac{3}{2\sqrt{2}} \sqrt{x} \right] ; 0 < x < 2$$

$$\therefore f(y|x) = \frac{f(x,y)}{f(x)} = \frac{\frac{3}{16\sqrt{2}} (y\sqrt{x} + x\sqrt{y})}{\frac{1}{4} \left(x + \frac{3}{2\sqrt{2}} \sqrt{x} \right)}$$

$$= \frac{3}{4\sqrt{2}} \frac{y\sqrt{x} + x\sqrt{y}}{x + \frac{3}{2\sqrt{2}} \sqrt{x}} ; \begin{matrix} 0 < y < 2 \\ 0 < x < 2 \end{matrix}$$

solution : Q3(3) :

As we can see from the full conditional distributions, $F^{-1}(\cdot)$ cannot be easily computed. Thus, we are going to perform rejection sampling method to draw samples within gibbs.

Now, we can rewrite the full conditionals in the following form :

$$\begin{aligned} f(x|y) &= \frac{3}{4\sqrt{2}} \cdot \frac{y\sqrt{x} + x\sqrt{y}}{y + \frac{3}{2\sqrt{2}}\sqrt{y}} \\ &= \frac{3}{4\sqrt{2}} \cdot \frac{1}{y + \frac{3}{2\sqrt{2}}\sqrt{y}} (y\sqrt{x} + x\sqrt{y}) \end{aligned}$$

Since, y is given or known for $f(x|y)$, it can be treated as a constant.

$$\text{let, } c_x = \frac{3}{4\sqrt{2}} \cdot \frac{1}{y + \frac{3}{2\sqrt{2}}\sqrt{y}}$$

$$\therefore f(x|y) = c_x (y\sqrt{x} + x\sqrt{y})$$

Now, let's get the maximum of $f(x|y)$, to get this,

$$\text{set, } \frac{\partial}{\partial x} f(x|y) = 0$$

$$\Rightarrow c_x \left[y \cdot \frac{1}{2\sqrt{x}} + \sqrt{y} \right] = 0$$

$$\Rightarrow c_x \sqrt{y} \left[\frac{\sqrt{y}}{2\sqrt{x}} + 1 \right] = 0$$

$$\Rightarrow \frac{\sqrt{y}}{2\sqrt{x}} = -1 \Rightarrow y = 4x \Rightarrow x = \frac{y}{4}$$

$$\text{since, } 0 < y < 2 \quad \text{so, } \max \{ f(x|y) \} = \frac{2}{4} = \frac{1}{2}.$$

Let's assume an envelope distribution $g(w) \sim \text{Unif}(0,2)$.

then, $m g(w) \geq f(x|y)$; $g(w) = \frac{1}{2}$ [PDF of Uniform]

$$\Rightarrow m \cdot \max \{ g(w) \} \geq \max \{ f(x|y) \}$$

$$\Rightarrow m \cdot \frac{1}{2} \geq \frac{1}{2}$$

$$\text{Thus, } m \geq 1.$$

Next, we will be doing the following steps:

i) Generate z from $\text{uniform}(0,2)$

ii) Calculate the ratio $R = \frac{f_{x|y}(x=z)}{m \cdot g(z)}$

$$R = \frac{f_{x|y}(x=z)}{1 \times \frac{1}{2}} = 2 \cdot f_{x|y}(x=z)$$

iii) if $u < 2 \cdot f_{x|y}(x=z)$ then accept z , otherwise reject where $u \sim \text{unif}(0,1)$. Here, if we accept z , it will be a draw from the full conditional of x .

Now, we can do similar things for $f(y|x)$ as both $f(x|y)$ and $f(y|x)$ are symmetric.

so, generate $z \sim \text{uniform}(0,2)$, then calculate

$$R = \frac{f_{y|x}(y=z)}{m \cdot g(z)} = 2 \cdot f_{y|x}(y=z)$$

then, accept z as a draw from $f_{y|x}(y|x)$

if $u < 2 \cdot f_{y|x}(y=z)$ where $u \sim \text{uniform}(0,1)$

otherwise, reject it.

so far we have,

$$\cancel{f(x|y)} \quad f(x|y) = c_x (y\sqrt{x} + x\sqrt{y})$$

$$\text{and,} \quad f(y|x) = c_y (y\sqrt{x} + x\sqrt{y})$$

where, c_x and c_y are known constant for $f(x|y)$ and $f(y|x)$ respectively. Both function looks similar as x and y are symmetric.

now we can perform gibbs sampling by the following steps:

we will use the order $x \rightarrow y$

step 1: initialize (x^0, y^0)

step 2: sample $x' \sim f(x|y=y^0)$ $y' \sim f(y|x=x')$

using rejection algorithm.

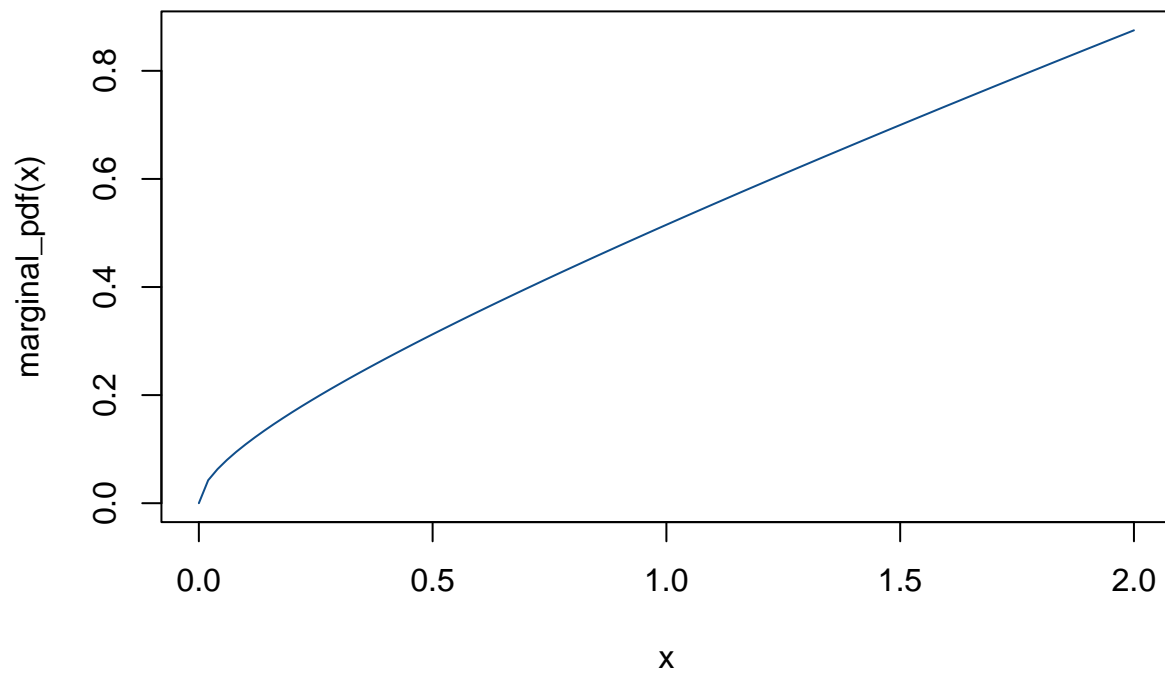
in general: $x^{(i)} \sim f(x|y=y^{(i-1)})$; $y^{(i)} \sim f(y|x=x^{(i)})$

step 3: Repeat step 2

Question 3:

The marginal density X and Y are symmetric function, so their density looks similar.

```
marginal_pdf <- function(v) {  
  .25 * (v + (3/(2*sqrt(2)))) * sqrt(v) )  
}  
curve(marginal_pdf, from = 0, to = 2, col="dodgerblue4")
```



Gibbs Sampler:

```
set.seed(729)
conditional_distribution <- function(x, y) {
  c_x <- (3 / (4 * sqrt(2))) * (y + (3 / (2 * sqrt(2))) * sqrt(y))^-1
  return(c_x*(y * sqrt(x) + x * sqrt(y)))
}

rejection_sampling <- function(y) {

  generate = TRUE

  while(generate) {

    z <- runif(1, min = 0,max = 2)
    u <- runif(1, min = 0, max = 1)

    if(u < conditional_distribution(x = z, y = y))
      generate = FALSE

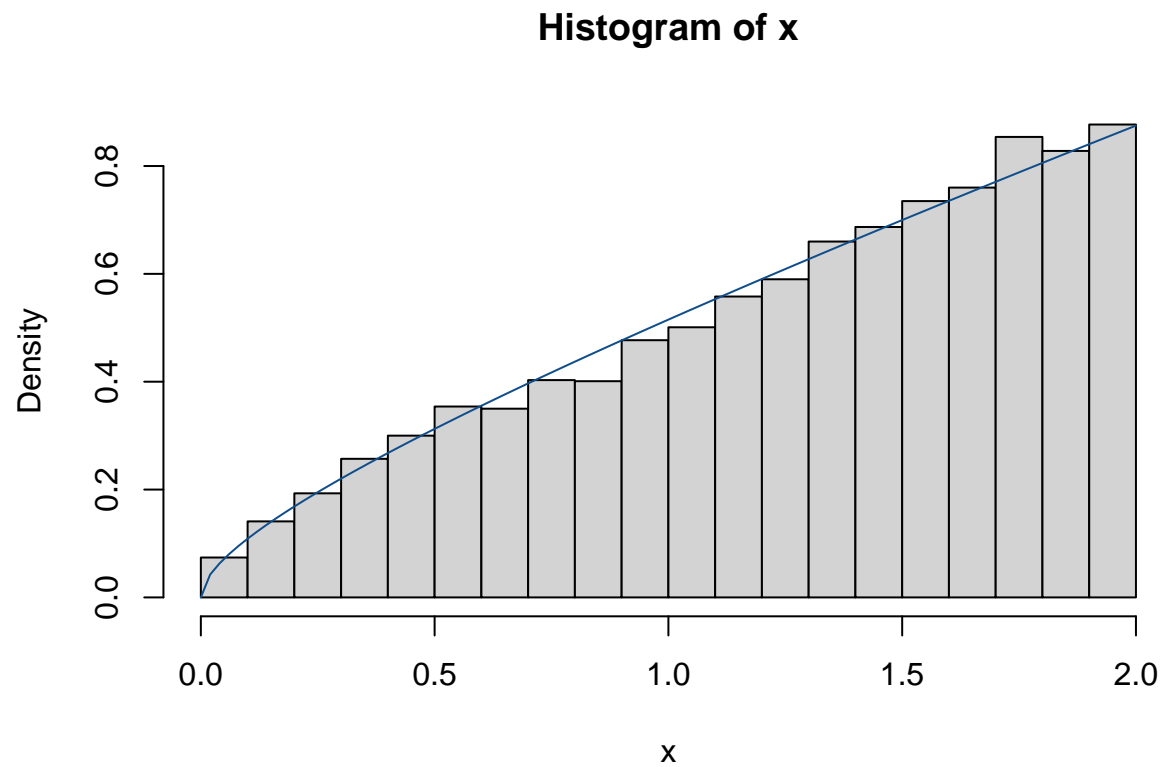
  }
  return(z)
}

# Gibbs sampling
x = y = 1

for(i in 2:10000) {
  x[i] = rejection_sampling(y[i-1])
  y[i] = rejection_sampling(x[i])
}
```

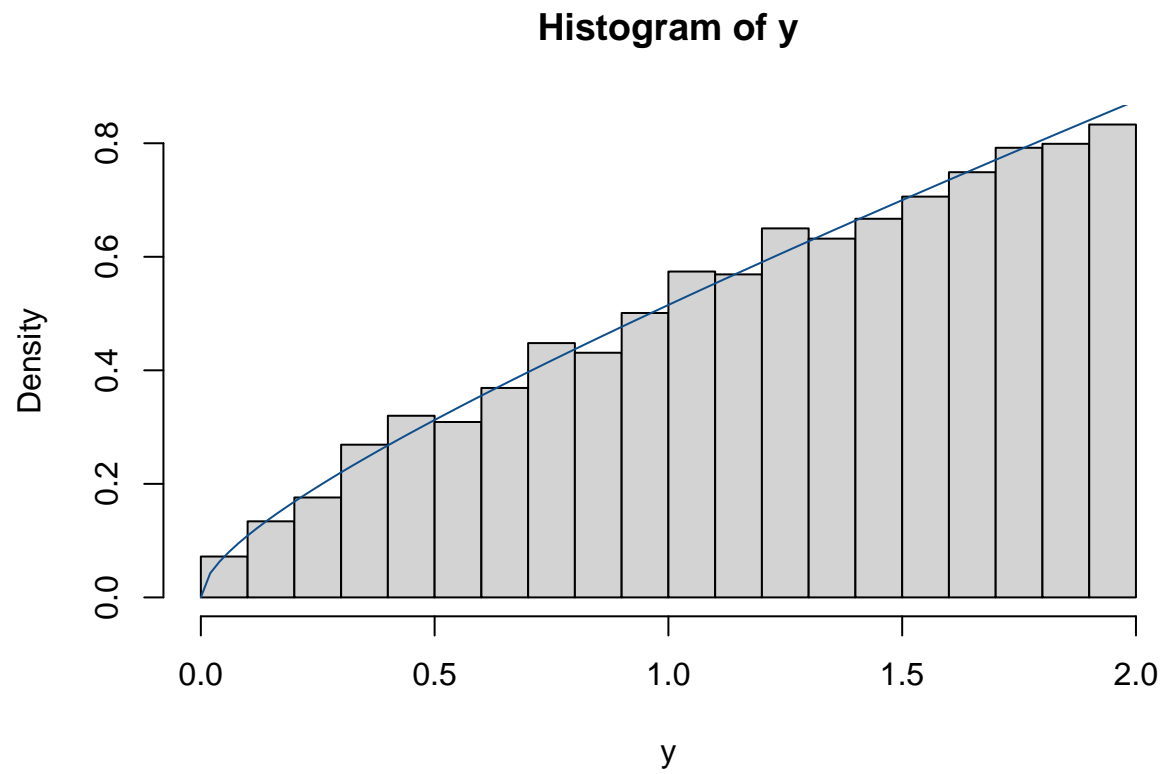
As we can see the pdf of marginal X and the histogram of X using the Gibbs sampling has matched. Thus we can say, the algorithm works well.

```
hist(x, probability = T)
curve(marginal_pdf, from = 0, to = 2, add = T, col="dodgerblue4")
```



As we can see the pdf of marginal Y and the histogram of Y using the Gibbs sampling has matched. Thus we can say, the algorithm works well.

```
hist(y, probability = T)
curve(marginal_pdf, from = 0, to = 2, add = T, col="dodgerblue4")
```



Solution : Q4 -

Given that,

$$x_1, \dots, x_n \stackrel{\text{iid}}{\sim} N(\mu, \sigma^2)$$

$$\text{and } f(\mu) \cdot f(\sigma^2) \propto \frac{1}{\sigma^2}$$

We need to find the conditional posterior of σ^2 .

Let's calculate the likelihood first.

$$L(\mu, \sigma^2 | \underline{x}) = \prod_{i=1}^n \cancel{f(x_i)} \cdot f(x_i | \mu, \sigma^2)$$

$$= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2} (x_i - \mu)^2\right\}$$

$$\propto (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right\}$$

[ignoring constant term.]

We know that,

~~Let's~~

posterior \propto likelihood \times prior.

$$\therefore f(\sigma^2 | \mu, \underline{x}) \propto L(\mu, \sigma^2 | \underline{x}) \cdot f(\mu) \cdot f(\sigma^2)$$

$$\propto (\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right\} \cdot \frac{1}{\sigma^2}$$

$$f(\sigma^2 | \mu, \underline{x}) \propto (\sigma^2)^{-n/2} \cdot \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right\} \cdot (\sigma^2)^{-1}$$

$$\propto (\sigma^2)^{-n/2-1} \cdot \exp \left\{ -\frac{1}{\sigma^2} \cdot \frac{\sum_{i=1}^n (x_i - \mu)^2}{2} \right\}$$

As we can see, the above expression is the kernel of the inverse-gamma PDF.

Thus,

$$\sigma^2 | \mu, \underline{x} \sim \text{IG} \left(\alpha = \frac{n}{2}, \beta = \frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2 \right)$$

$$\sigma^2 | \mu, \underline{x} \sim \text{IG} \left(\alpha = \frac{n}{2}, \beta = \frac{1}{2} \sum_{i=1}^n (x_i - \mu)^2 \right)$$

[if $y \sim \text{IG}(\alpha, \beta)$ then,

$$f(y) = \frac{\beta^\alpha}{\Gamma(\alpha)} \cdot \exp \left\{ -\frac{\beta}{y} \right\} \cdot y^{-\alpha-1}]$$

Now, where, $n = 5$; $\mu = 2$.

$$\therefore \sum_{i=1}^5 (x_i - \mu)^2 = 3.25$$

$$\therefore \sigma^2 | \mu=2, \underline{x} = (1.5, 2.6, 1.0, 3.0, 2.8) \sim \text{IG} \left(\frac{5}{2}, \frac{13}{8} \right)$$

```
x <- c(1.5, 2.6, 1.0, 3.0, 2.8)
n <- length(x)
ss <- sum((x - 2)^2)
n_samples <- 10000
ig_sample <- 1/rgamma(n_samples, shape = n/2, rate = ss/2)
```

