Homework Assignment 3

Jack, 2022/3/20

Due on Mar 27, 2022 at 11:59 pm

1. The Pareto(a, b) distribution has cdf

$$F(x) = 1 - \left(\frac{b}{x}\right)^a, \quad x \geq b > 0, a > 0$$

   (a) Derive the probability inverse transformation $F^{-1}(U)$ and use the inverse transform method to simulate a random sample with size 1000 from the Pareto $(2, 2)$ distribution.

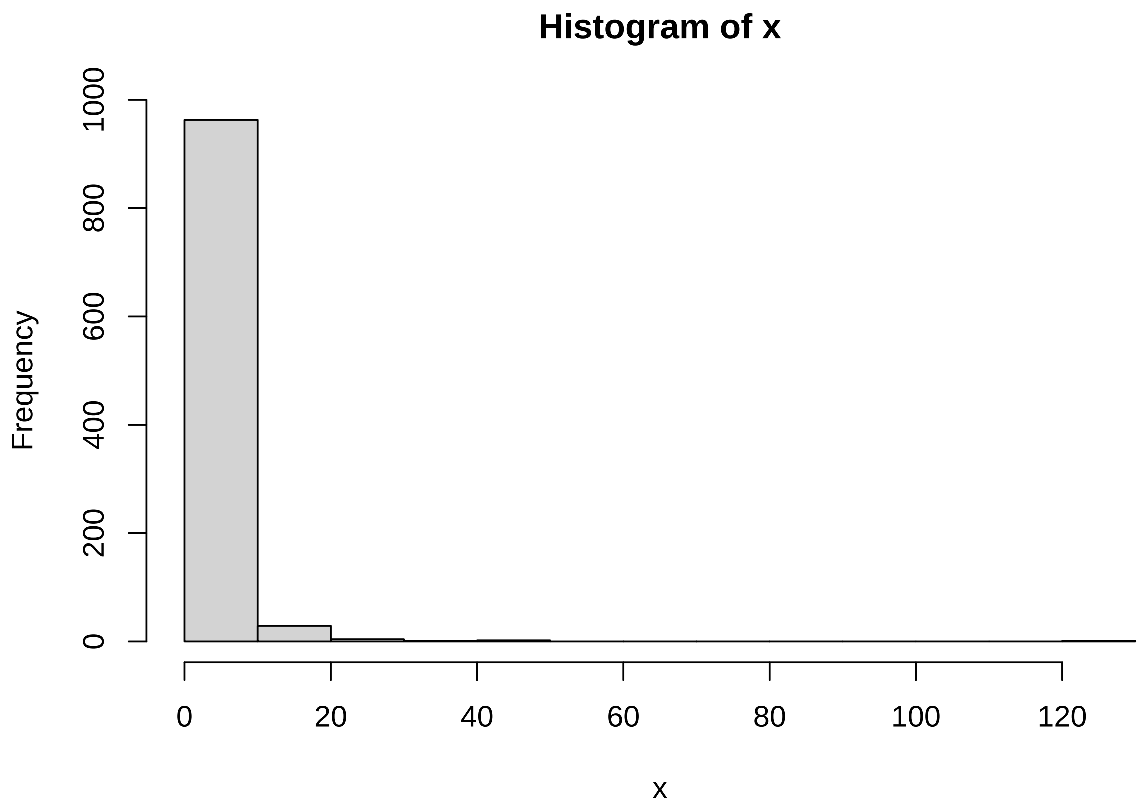
$U = F_x(x) = 1 - \left(\frac{b}{x}\right)^a$

$1 - U = \left(\frac{b}{x}\right)^a$

$(1 - U)^{1/a} = \left(\frac{b}{x}\right)$

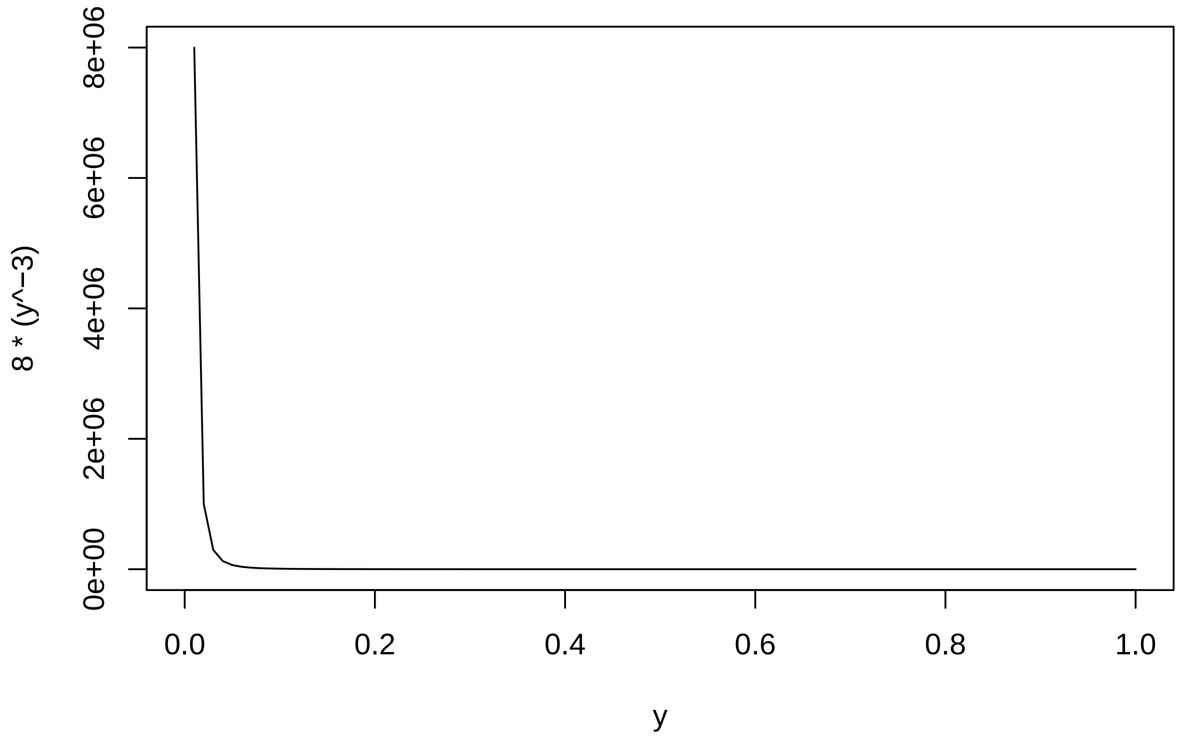
$x = \left(\frac{b}{1 - U^{1/a}}\right)$

```r
n <- 1000
u <- runif(n)
x <- 2/((1-u)^(1/2))
hist(x)
```



(b) Graph the density histogram of the sample with the Pareto $(2, 2)$ density superimposed for comparison.

```
y <- seq(0, 1, .01)
plot(y, 8*(y^-3), type = 'l')
```



2. A discrete random variable X has probability mass function

x	0	1	2	3	4
$p(x)$	0.1	0.2	0.2	0.2	0.3

(a) Use the inverse transform method to generate a random sample of size 1000 from the distribution of X. (Hint: search *if else if* function)

```
pmf = c(0.1, 0.2, 0.2, 0.2, 0.3)
cdf <- cumsum(pmf)
cdf
```

```
## [1] 0.1 0.3 0.5 0.7 1.0
```

```
n <- 1000
x <- numeric(n)
k <- 0
while (k<n) {
  u <- runif(1)
  k <- k+1
  if (u>cdf[1] & u<=cdf[2]){
    x[k] = 1
  } else if (u>cdf[2] & u<=cdf[3]){
    x[k] = 2
  } else if (u>cdf[3] & u<=cdf[4]){
```

```
    x[k] = 3
  } else if (u>cdf[4] & u<=cdf[5]){
    x[k] = 4
  } else { x[k] <- 0 }
}
mean(x)
```

```
## [1] 2.335
```

```
(b) Construct a relative frequency table and compare the empirical with the theoretical probabilities
```

We can find that the mean of the empirical and theoretical are Very close to each other.

```
pmf = c(0.1, 0.2, 0.2, 0.2, 0.3)
#cdf <- cumsum(pmf)
#cdf
n <- 1000
x = numeric(n)
x_0 <- rep(0, times=n*pmf[1])
x_1 <- rep(1, times=n*pmf[2])
x_2 <- rep(2, times=n*pmf[3])
x_3 <- rep(3, times=n*pmf[4])
x_4 <- rep(4, times=n*pmf[5])
x <- c(x_0, x_1, x_2, x_3, x_4)
mean(x)
```

```
## [1] 2.4
```

```
(c) Repeat b) by using $R$ function 'sample' to generate a random sample of $X$ with sample size 1000, a
```

```
n <- 1000
result <- sample(c(0,1,2,3,4), size = n, replace = TRUE, prob=c(0.1,0.2,0.2,0.2,0.3))
result <- data.frame(result)
result %>%
  group_by(result) %>%
  summarise(count = n(), freq = count/n) %>%
  mutate(prob = c(0.1,0.2,0.2,0.2,0.3))
```

```
## # A tibble: 5 x 4
##   result count  freq  prob
##    <dbl> <int> <dbl> <dbl>
## 1      0   100 0.1     0.1
## 2      1   212 0.212   0.2
## 3      2   221 0.221   0.2
## 4      3   185 0.185   0.2
## 5      4   282 0.282   0.3
```

3. Consider the integration $\theta = \int_A g(x)dx$, where

$$g(x) = xe^{-\frac{x^2}{2}}$$

and $A \in \{x : 1 < x < \infty\}$, the importance function

$$f(x) = e^{-(x-1)},$$

for $1 < x < \infty$, compute the Monte Carlo estimator $\hat{\theta}$ and $\mathrm{Var}(\hat{\theta})$ using the importance sampling method.

We should care that the domain of f(x) and g(x) are the same, so we do not need to screen the interval.

```
n <- 10000
g <- function(x){
  res <- x*exp((-x^2)/2)
  return(res)
}
x <- rexp(n,1)*exp(1)
fg <- g(x) / exp(-(x-1))
theta_hat <-mean(fg)
se <- sd(fg)
round(rbind(theta_hat, se), 4)
```

```
##              [,1]
## theta_hat 0.3246
## se        0.2953
```

4. (a) Find two importance functions f_1 and f_2 that are supported on $(1, \infty)$ and are 'close' to

$$g(x) = \frac{x^2}{\sqrt{2\pi}} e^{-x^2/2}, \quad x > 1.$$

Which of your two importance functions should produce the smaller variance in estimating

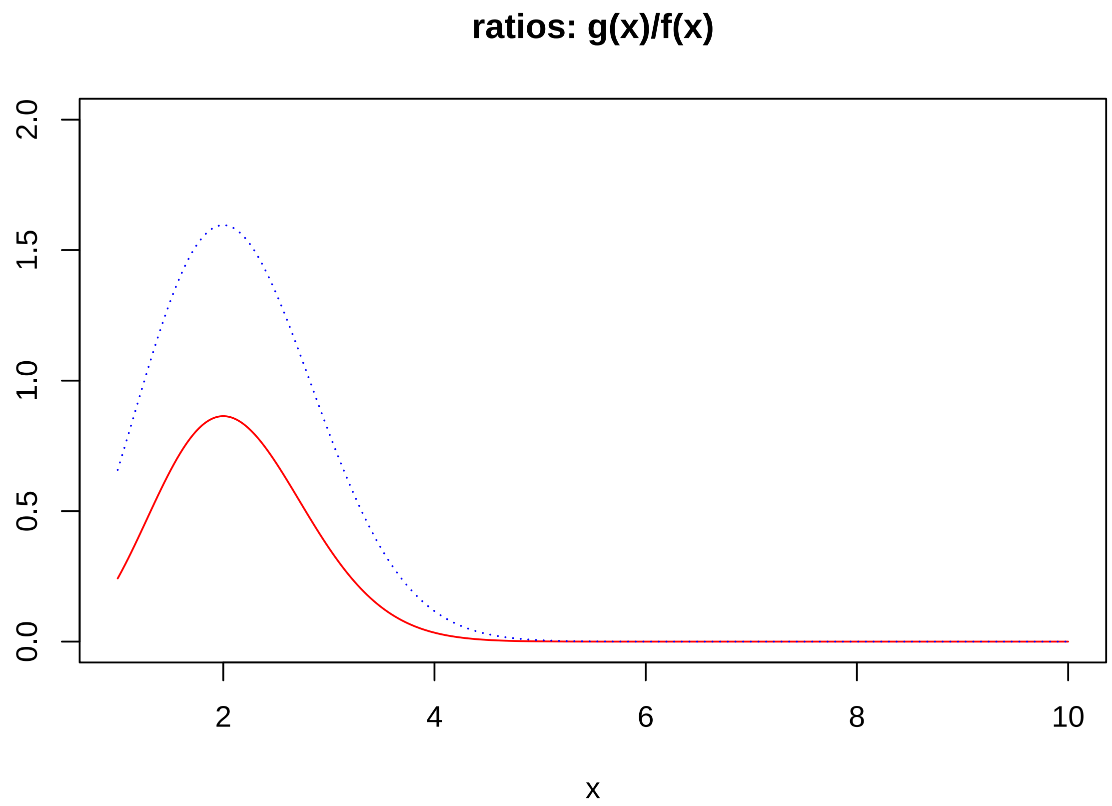
$$\int_1^\infty \frac{x^2}{\sqrt{2\pi}} e^{-x^2/2} dx$$

by importance sampling? Explain.

Ans. We can find that the function g(x) is standard normal distribution. Since the domain of it is from 1 to ∞, for f_1, we can just select the function $1/x^2$. For f_2, we select e^{-x} which is the probability density function (pdf) of a case ($\lambda = 1$) of exponential distribution.

```
x <- seq(1,10,0.01)
g <- x^2/sqrt(2*pi)* exp((-x^2/2))
f_1 <- 1 / (x^2)
f_2 <- exp(-x)

plot(x, g/f_1, type = "l", ylim = c(0,2),ylab = "", col="red",main = 'ratios: g(x)/f(x)')
lines(x, g/f_2, lty = 3,col="blue")
```



From the graph, we can find that $g(x)/f_1(x)$ fluctuates less which means that more closer to a line, so the $f(x) = 1/(x^2)$ should produce the smaller variance in estimating.

```
(b) Obtain the Monte Carlo estimates of the previous integral using those two importance functions $f_{
```

```
n <- 10000
theta_hat <- se <- numeric(2)
g <- function(x){
  res <- (x^2/sqrt(2*pi))*exp((-x^2)/2)*(x>1)
  return(res)
}
# f_1
u <- runif(n)
x <- 1 / (1-u) # inverse transform method
fg <- g(x) / x^(-2)

theta_hat[1] <- mean(fg)
se[1] <- sd(fg)

# f_2
x <- rexp(n, 1)
fg <- g(x) / exp(-x)
theta_hat[2] <- mean(fg)
se[2] <- sd(fg)
round(rbind(theta_hat, se), 4)
```

```
##              [,1]   [,2]
## theta_hat 0.3943 0.3991
## se        0.3065 0.5874
```

Actually, we can find than $f_1(x)$ as importance function can get the smaller variance.

5. A certain type of electronic component has a lifetime Y (in hours) with probability density function given by.

$$f(y \mid \theta) = \begin{cases} \left(\frac{1}{\theta^2}\right) y e^{-y/\theta}, & y > 0 \\ 0, & \text{otherwise.} \end{cases}$$

   That is, Y has a gamma distribution with parameters $\alpha = 2$ and θ. Let $\hat{\theta}$ denote the MLE of θ. Suppose that n such components, tested independently, had lifetimes of $Y_1, Y_2, \ldots, Y_n$ hours. Find the MLE of θ.

Step 1: $L(\theta) = \prod_{i=1}^n f(y|\theta_i) = \theta^{-2n} \prod_{i=1}^n y_i e^{y_i/\theta}$

Step 2: $ln(L(\theta)) = -2n \ln\theta + \ln \prod_{i=1}^n y_i - \frac{1}{\theta} \sum_{i=1}^n y_i$

Step 3: $0 = \frac{\partial}{\partial\theta} = -\frac{2n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^n y_i = -\frac{2n}{\theta} + n\frac{\bar{y}}{\theta^2}$

As result, $\theta = \frac{\sum_{i=1}^n y_i}{2n} = \frac{\bar{y}}{2}$. To conclude, the MLE of θ is $\frac{\bar{y}}{2}$.