Assignment 2

Ehsan Eslmai Shafigh

2023-04-26

Assignment 2

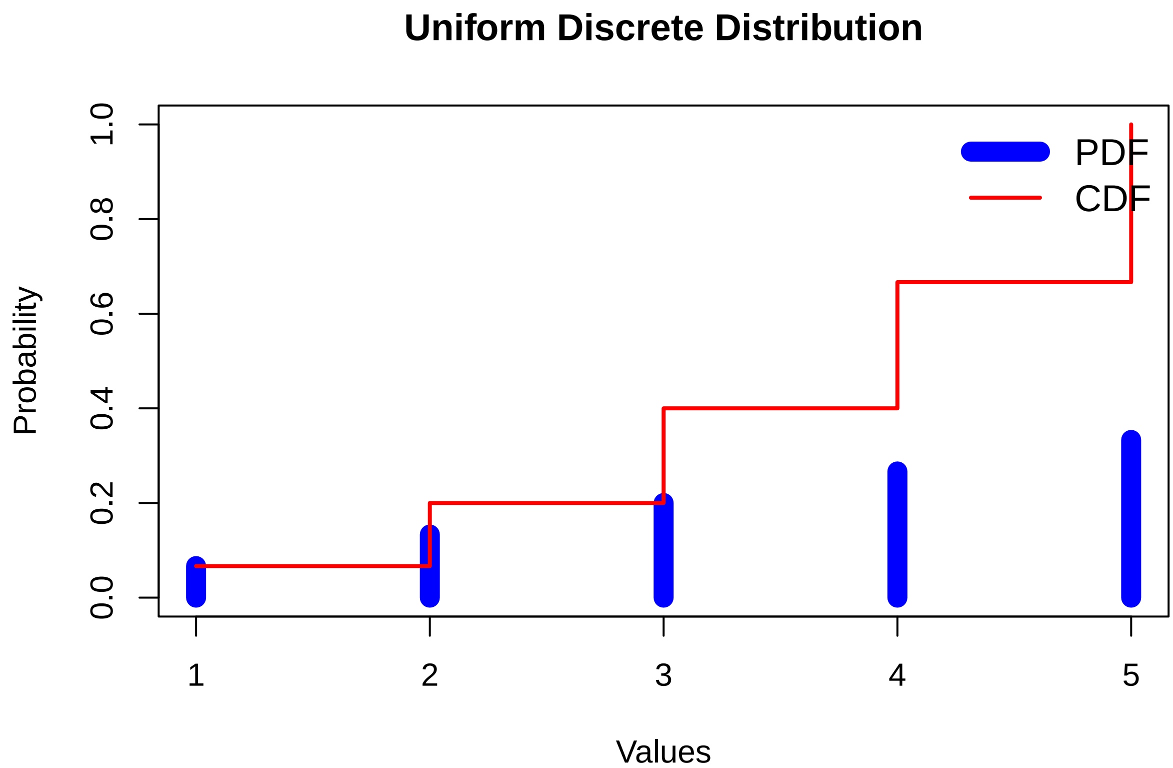
Exercise 1: Discrete Random Variables

```
x = 1:5
min_x = 1
max_x = 6

pdf = function(k) k/15
cdf = function(k) k*(k + 1) /30
```

1)

```
plot(x, pdf(x), type = "h", col = "blue", lwd = 10, ylim = c(0, max(pdf(x), cdf(x))),
     ylab = "Probability", xlab = "Values", main = "Uniform Discrete Distribution")
lines(x, cdf(x), type = "s", col = "red", lwd = 2)
legend("topright", c("PDF", "CDF"), lty = c(1, 1), col = c("blue", "red"),
       lwd = c(10, 2), bty = "n", cex = 1.2)
```



2)

```
mean_ = sum(pdf(x) * x)
variance_ = sum(pdf(x) * x^2 - (pdf(x) * x)^2)

print(mean_)
```

3)

```
## [1] 3.666667
```

```
print(variance_)
```

```
## [1] 10.64889
```

4) first we define the function:

```
func = function(k){k * (6 - k)}
```

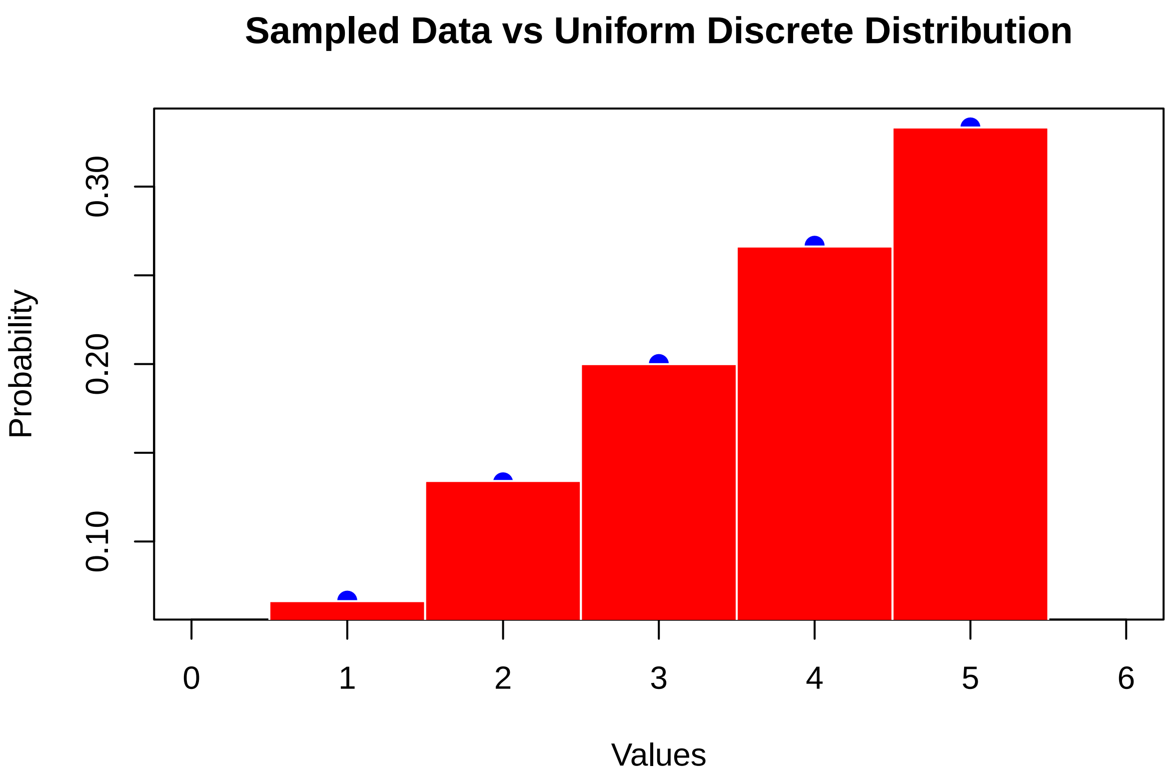
```
expected_value = sum(func(x) * pdf(x))
expected_value
```

```
## [1] 7
```

```
samples = sample(x, size = 1000000, replace = TRUE, prob = pdf(x))
```

5)

```
plot(x, pdf(x), type = "h", col = "blue", lwd = 10, xlim = c(min_x - 1, max_x),
     ylab = "Probability", xlab = "Values", main = "Sampled Data vs Uniform Discrete Distribution")
hist(samples, breaks = seq(min_x - 0.5, max_x + 0.5, by = 1), col = "red", border = "white" , freq = FA
```



6)

Exercise 2: Continuous random variable

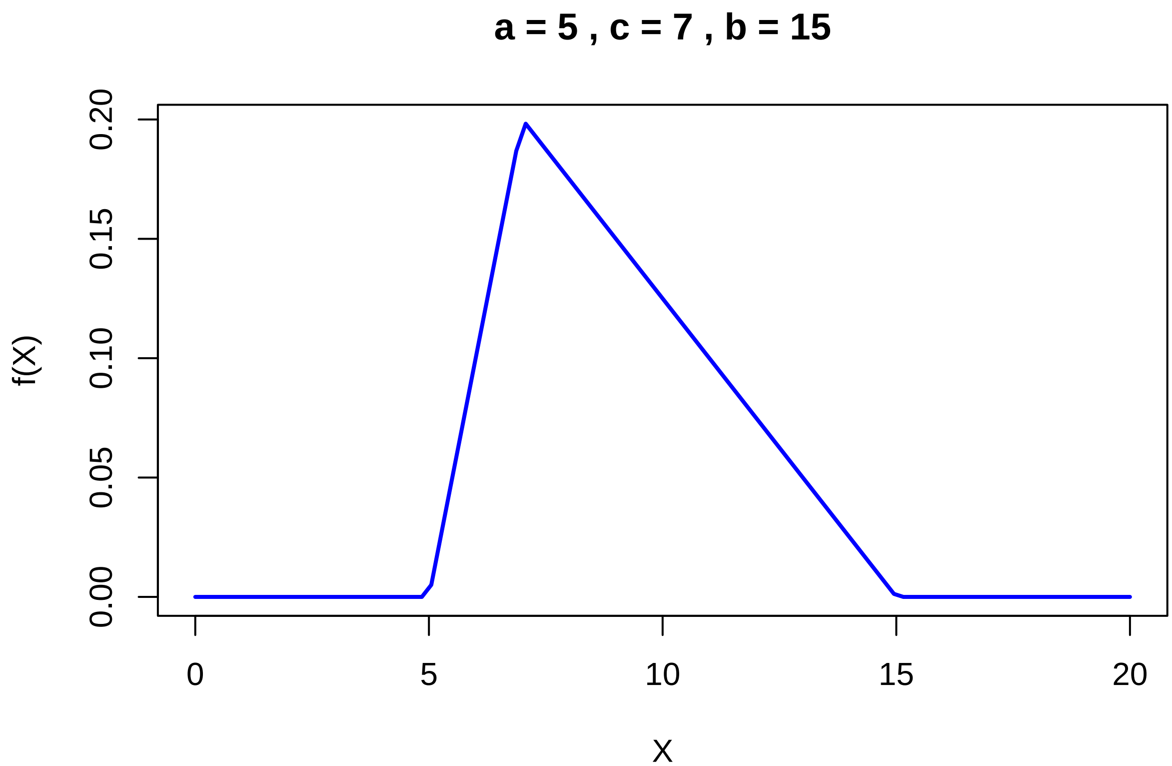
a) we first define the function f(X) and assign some values to a,b and c parameters:

```
f = function(x,a,b,c){ifelse(x<a|x>b, 0, ifelse(x<c, 2*(x-a)/((b-a)*(c-a)) , 2 * (b-x)/((b-a)*(b-c))))}

a = 5
b = 15
c = 7
```

next we plot the function:

```
curve(f(x,a,b,c) , from = 0 , to = 20 , n = 100 , ylab = 'f(X)' , main = 'a = 5 , c = 7 , b = 15' ,col =
```



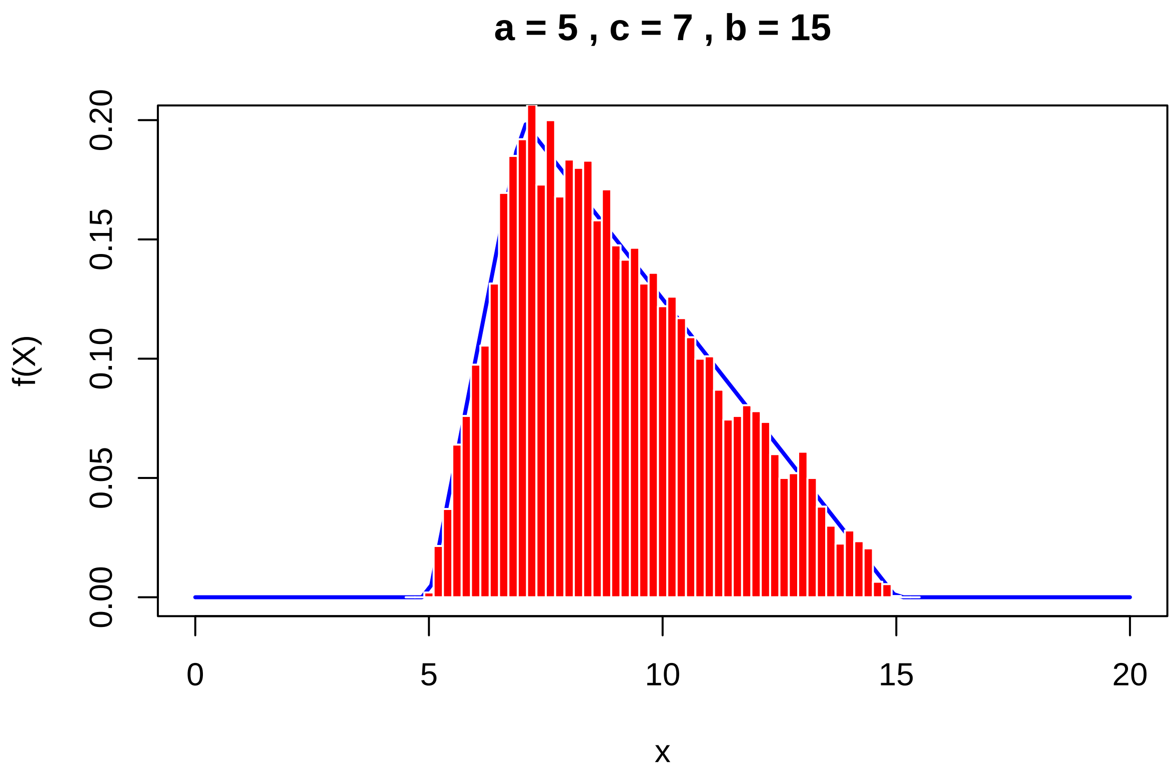
b) We can use the inverse transform sampling method to generate random numbers from the above distribution. The inverse of the cumulative distribution function is found to be:

```
X = function(u, a , b , c){ifelse(u < (c - a)/(b - a) , sqrt(u * (b - a) * (c - a)) + a , -sqrt((b - a)
```

c) Now we just need to generate some numbers between 0 and 1 and use the above transformation:

```
u = runif(10000)
Xs = X(u , a,b,c)

curve(f(x,a,b,c) , from = 0 , to = 20 , n = 100 , ylab = 'f(X)' , main = 'a = 5 , c = 7 , b = 15' ,col =
hist(Xs, breaks = seq(a - 0.5, b + 0.5, by = 0.2), col = "red", border = "white" , freq = FALSE ,add = '
```

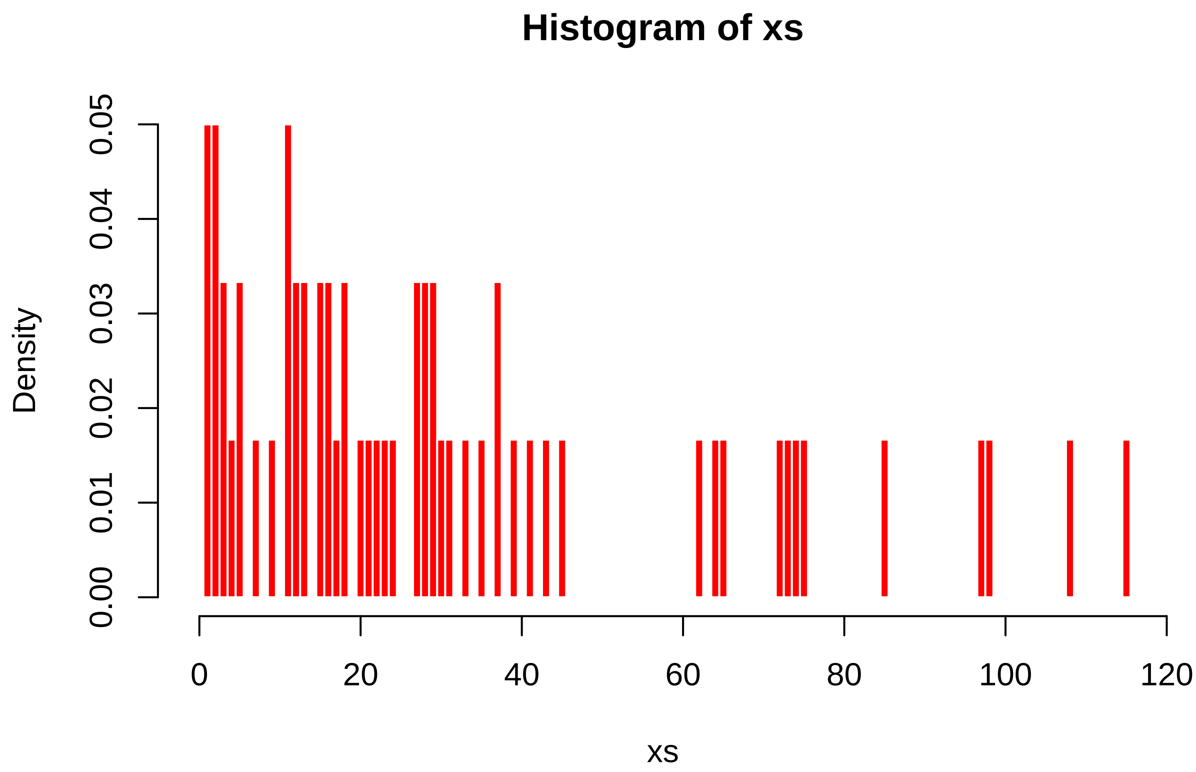


Exercise 3

A) As we know the waiting time follows an exponential distribution, with a lambda $= 1/30$:

```
lambda = 1/30
xs = rexp(60 , rate = 1/30)
breaks = seq(-0.5 , max(xs) + 1 + 0.5, by = 1.0)

hist(xs, breaks = breaks , col = "red", border = "white" , freq = FALSE )
```




```
count = sum(xs<12)
prob = count/length(xs)
prob
```

B)

```
## [1] 0.2666667
```

```
mean(xs)
```

C)

```
## [1] 31.33761
```

As we see the mean of the samples is almost close to the theoretical value 1/lambda = 30

```
count = sum(xs>60)
prob = count/length(xs)
prob
```

D)

```
## [1] 0.2
```

Exercise 4: multiple choices exams

We define the following events:

A: a correct answer is given by the student
B: the student knows the answer

The Bayes Theorem says:

P(B|A) * P(A) = P(A|B) * P(B)

P(A): The probability that a correct answer is given by the student
P(B): The probability that the student knows the answer
P(B'): The probability that the student does not know the answer
P(B|A): The probability that the student knows the answer, when a correct answer is given
P(A|B): The probability that a correct answer is given, when the student knows the answer
P(A|B'): The probability that a correct answer is given, when the student does not know the answer

We want to know P(B|A). P(A) itself should be calculated as:

P(A) = P(A|B) * P(B) + P(A|B') * P(B')

Now the above probabilities can be correctly calculated from the information provided by the question:

P(A|B) = 1, P(A|B') = 0.2, P(B)= 0.7, P(B') = 1 - P(B) = 0.3

```
PAB = 1
PAB_ = 0.2
PB = 0.7
PB_ = 0.3

PA = PAB * PB + PAB_ * PB_

PBA = PAB * PB / PA

PBA
```

```
## [1] 0.9210526
```

So:

P(B|A) = 0.92

Exercise 5: Waiting time

a) If the guy arrives between 10:50 and 11:00, or between 11:20 and 11:30 , he will wait at most for 10 minutes, so the probability can be computed as:

```
10/60 + 10/60
```

```
## [1] 0.3333333
```

b) If the guy arrives between 11:00 and 11:15 or 11:30 and 11:45, he will wait at least for 15 minutes:

```
15/60 + 15/60
```

```
## [1] 0.5
```

c) Let's assume that the arrival time is x. If x is less than 11, he will wait for 11- x, if he arrives between 11 and 11.50 , will wait for 11.50 - x, and if he arrives after 11.50, will wait for 12-x. We generate 1000 random numbers between 10.75 and 11.75, and based on the above mentioned conditions, calculate their waiting time, and we average over them to calculate the average waiting time.

```
ts = runif(100000 , min =  10.75 , max = 11.75)


ts[ts < 11] = 11 - ts[ts < 11]
ts[ts > 11 & ts < 11.50 ] = 11.50 - ts[ts > 11 & ts < 11.50]
ts[ts > 11.50] = 12 - ts[ts > 11.50]

average_waitingtime = mean(ts) * 60
average_waitingtime
```

```
## [1] 15.02198
```

So the average waiting time is around 15 minutes using this MC method.

Exercise 6: stock investment

Given the mean and standard deviation values, the normal distribution is determined. Then we calculate the return rate in question, and derive the corresponding probability:

```
rate = 800/(200 * 85) * 100

1 - pnorm(rate , mean = 10 , sd = 12)
```

```
## [1] 0.6704574
```