Lab Report: Lab4 - Computational Statistics

Dhanush Kumar Reddy Narayana Reddy (dhana004), Udaya Shanker Mohanan Nair (udamo524)

2025-02-18

Introduction

Implementation of 2 Assignment questions of Computational Statistics Lab 4 .

Contributions

Member: Dhanush Kumar Reddy Narayana Reddy, Liu Id: dhana004, Contribution: Report writing and coding of question 2.

Member: Udaya Shanker Mohanan Nair, Liu Id: udamo524, Contribution: Report writing and coding of question 1.

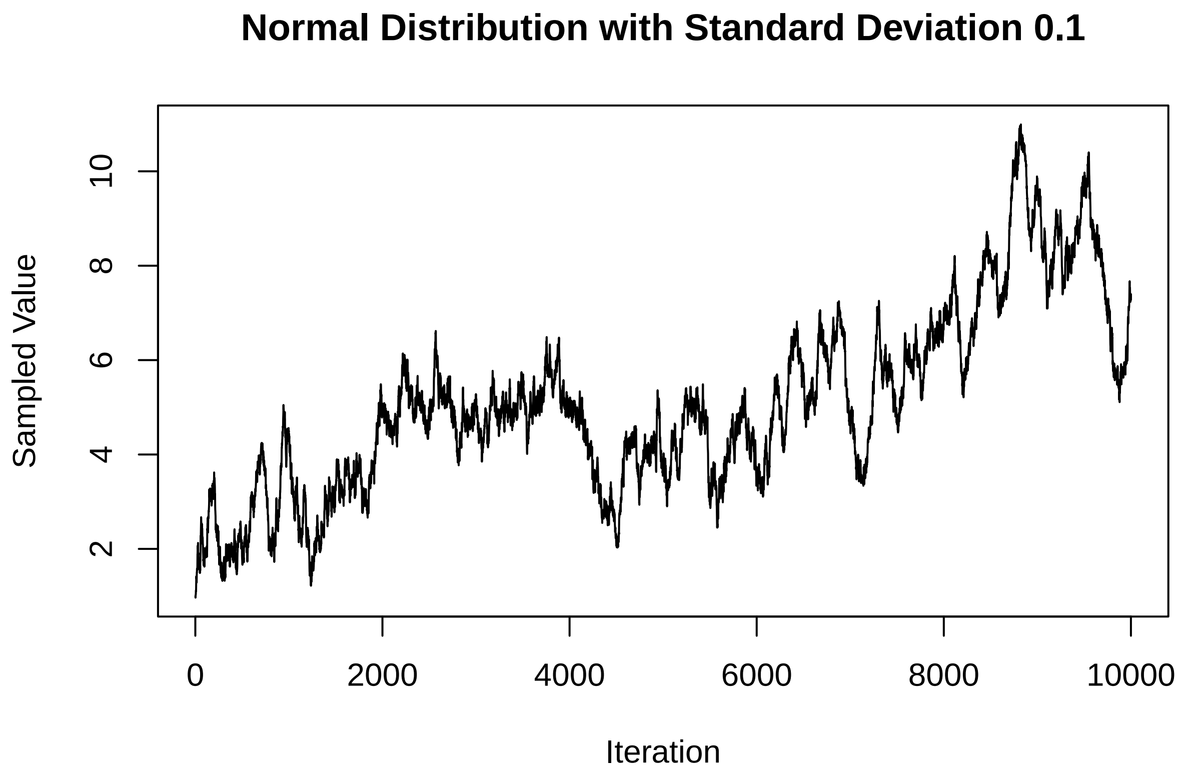
Question 1

Given below a target distribution,

$$f(x) = 120x^5 e^{-x}, \quad x > 0.$$

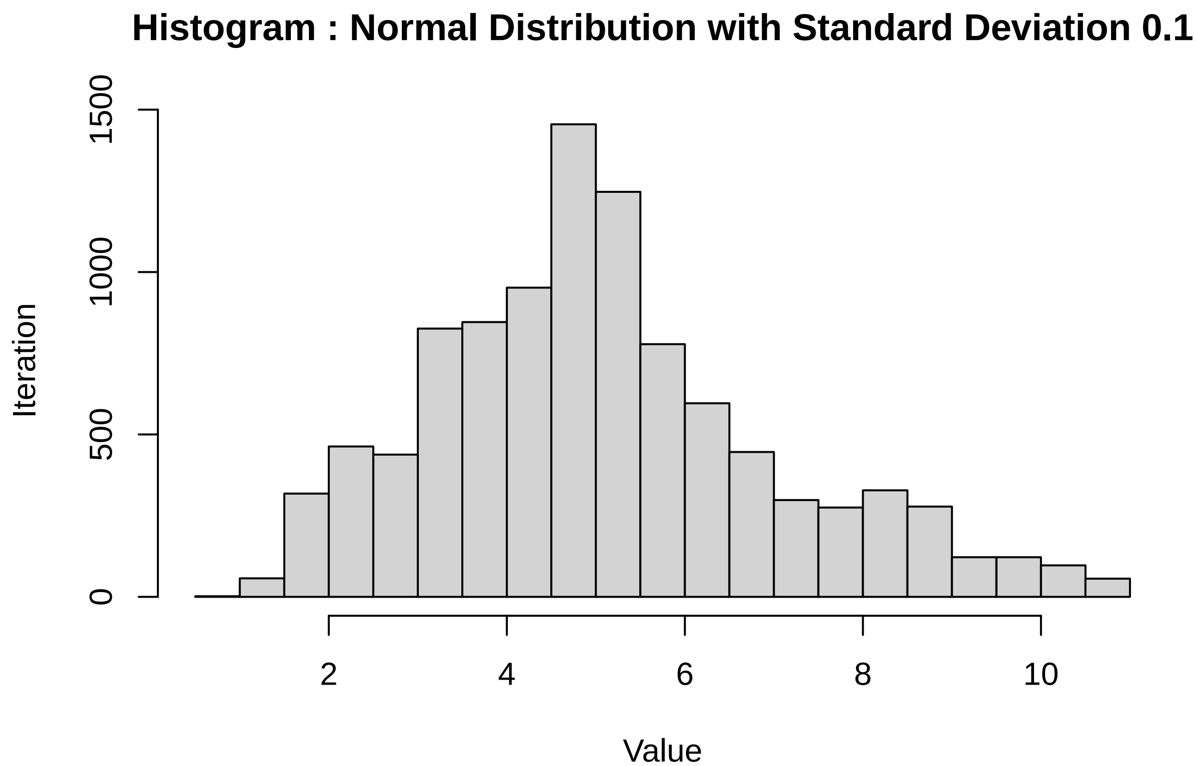
Part A

In this part we are asked to use Metropolis-Hastings Algorithm to generate 10000 samples using a normal distribution.



Convergence :Values appears to have a upward motion initially and then stabilizes after few iterations. After 3000-4000 iterations, the values oscillates around a stable range(4-10), which indicates that it is more likely converged to that taget distribution.

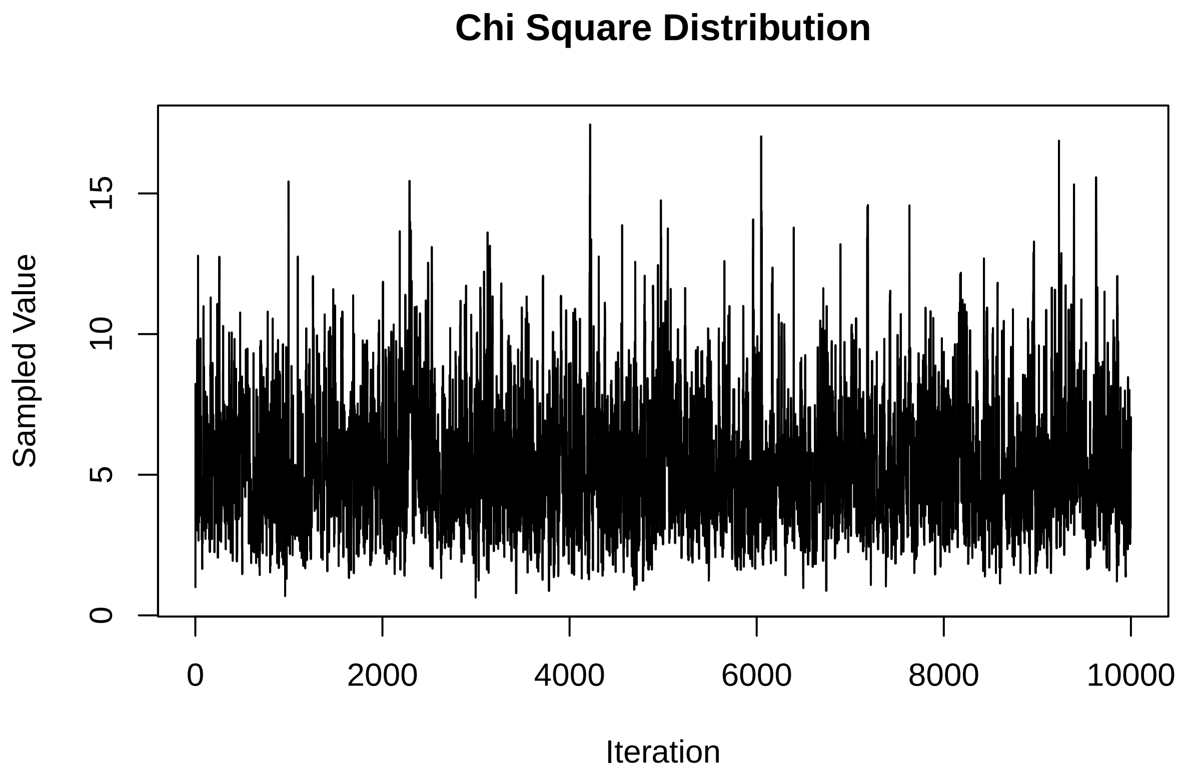
Burn-in Period : Initial few iterations up-to around 3000-4000 shows variations in values,which is considered to be burn-in period after which values reached a stable distribution.

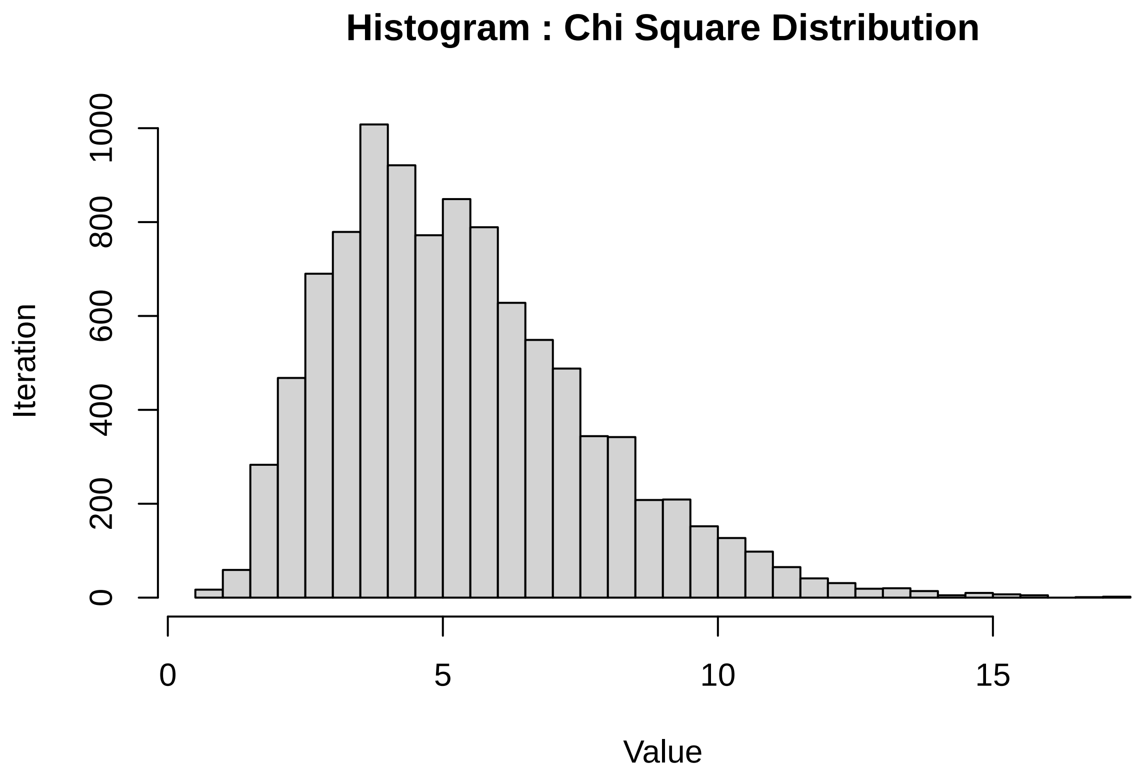



```
## Acceptance Rate of the First Distribution Sample : 9859
```

Part B

In this part we are asked to use Metropolis-Hastings Algorithm to generate 10000 samples using a chi square distribution.

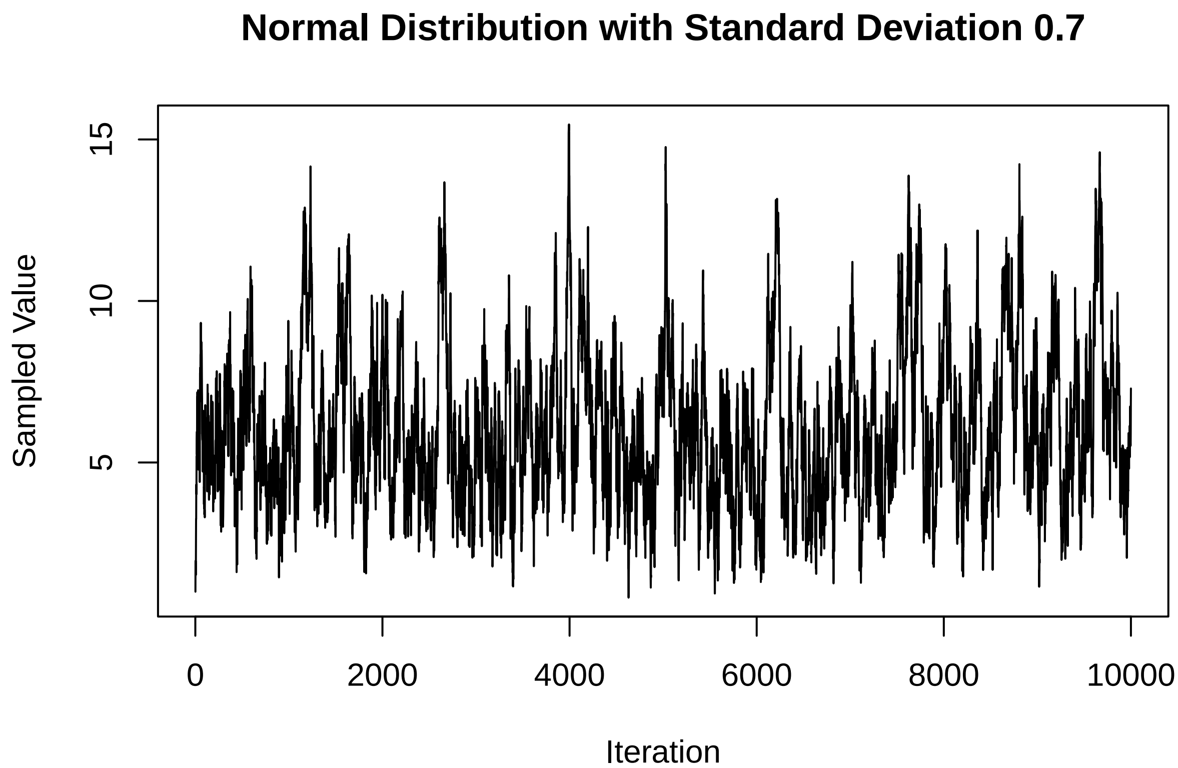

Chi Square Distribution
Sampled Value
0
5
10
15
0
2000
4000
6000
8000
10000
Iteration

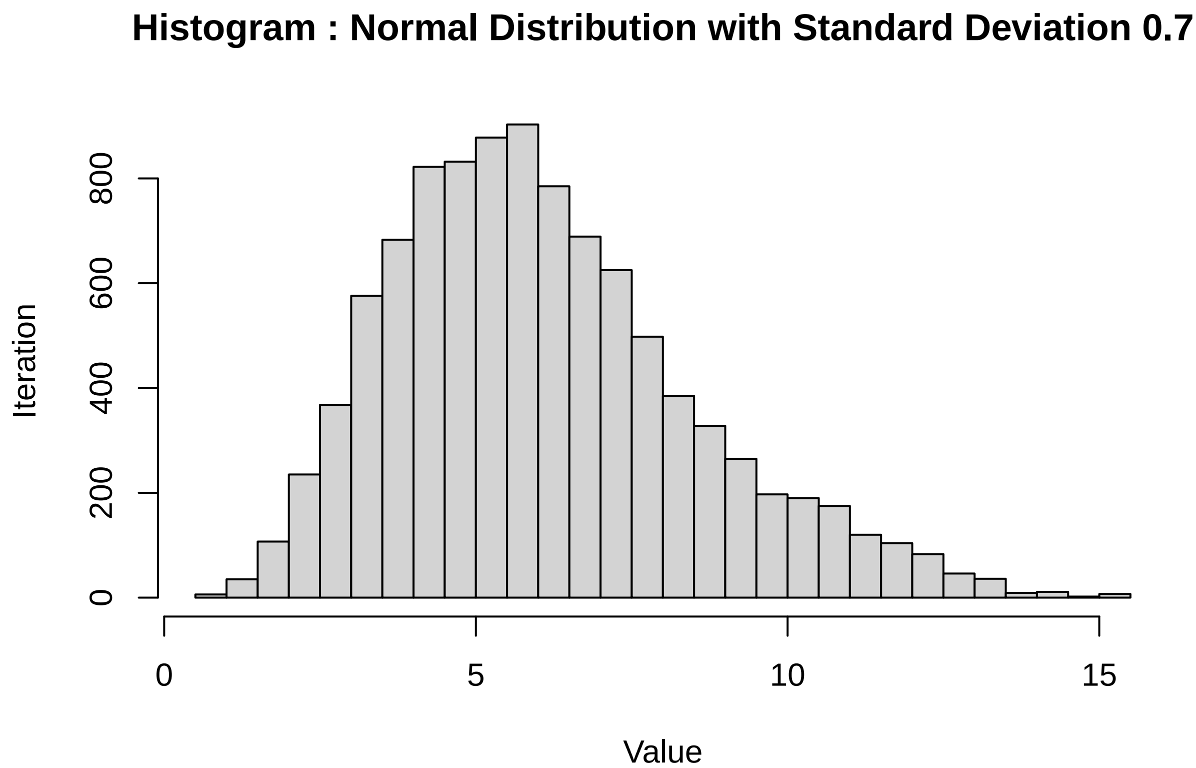



```
## Acceptance Rate: 6085
```

Part C

In this part we choose to use Metropolis-Hastings Algorithm to generate 10000 samples using a normal distribution with another Standard Deviation different from part A. Here in this case we opted 0.4 in for Standard Deviation.


Normal Distribution with Standard Deviation 0.7
Sampled Value
5
10
15
0
2000
4000
6000
8000
10000
Iteration




```
## Acceptance Rate: 9039
```

Part D

Report of the above three generated distributions

Distribution	Acceptance_Rate	E_X
Normal (SD = 0.1)	9859	5.078
Chi-Square (df = floor(X_t) + 1)	6085	5.345
Normal (SD = 0.7)	9039	6.072

Among all the above three distributions, the Normal Distribution with standard deviation 0.7 gives a better option as it is having almost equivalent acceptance rate in comparison with normal distribution(Standard Deviation 0.1) and greatest E(X) of all the three.

Part E

In this part we calculated the E(X) of all the three distributions,

```
## First Dirtibution- Normal Distribution with Standard Deviation 0.1 =  5.077624
```

```
## Second Dirtibution- Chi- Square Distribution =  5.345107
```

```
## First Dirtibution- Normal Distribution with Standard Deviation 0.7 =  6.072416
```

Part F

From this given PDF it is understood that it is a gamma distribution

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x \in (0, \infty)$$

From the above given standard format,

$$f(x) = 120x^5 e^{-x}, \quad x > 0.$$

is a gamma distribution with $\alpha = 6$ (shape) and $\lambda = 1$ (rate).

Mean of gamma distribution, $E(x) = \frac{\alpha}{\lambda} = \frac{6}{1} = 6$

Distribution	Acceptance_Rate	E_X	theoretical E(X)
Normal (SD = 0.1)	9859	5.078	6
Chi-Square (df = floor(X_t) + 1)	6085	5.345	6
Normal (SD = 0.7)	9039	6.072	6

Question 2

Appendix

Question 1

```
fx <- function(x){
  if(x>0){
    res <- 120 * x^5 * exp(-x)
  }
  else{
    res <- 0
  }
  return(res)
}
set.seed(101)
n <- 10000

first_distribution <- numeric(n)
#start value is set to 1
first_distribution[1] <- 1
first_accepted_rate <- 0
for (i in 2:n) {
  value <- rnorm(1, mean=first_distribution[i-1], sd=0.1)
  check_value <- fx(value)/fx(first_distribution[i-1])
  if (runif(1) < check_value) {
    first_distribution[i] <- value
    first_accepted_rate <- first_accepted_rate + 1
  } else {
    first_distribution[i] <- first_distribution[i-1]
```

```
  }
}
compare_df_a <- data.frame(
  Distribution = "Normal (SD = 0.1)",
  Acceptance_Rate = first_accepted_rate,
  E_X = mean(first_distribution)
)

plot(first_distribution, type='l', main='Normal Distribution with Standard Deviation 0.1',
     xlab='Iteration', ylab='Sampled Value')
hist(first_distribution, breaks=30, main='Histogram : Normal Distribution with Standard Deviation 0.1',
     xlab='Value', ylab='Iteration')


cat("Acceptance Rate of the First Distribution Sample :", first_accepted_rate)

second_distribution <- numeric(n)
second_distribution[1] <- 1
second_acceptance_rate <- 0
for (i in 2:n) {
  value <- rchisq(1,df=floor(second_distribution[i-1])+1)
  check_value <- fx(value)/fx(second_distribution[i-1])
  if (runif(1) < check_value) {
    second_distribution[i] <- value
    second_acceptance_rate <- second_acceptance_rate + 1
  } else {
    second_distribution[i] <- second_distribution[i-1]
  }
}
compare_df_b <- data.frame(
  Distribution = "Chi-Square (df = floor(X_t) + 1)",
  Acceptance_Rate = second_acceptance_rate,
  E_X = mean(second_distribution)
)
plot(second_distribution, type='l', main='Chi Square Distribution',
     xlab='Iteration', ylab='Sampled Value')

hist(second_distribution, breaks=30, main='Histogram : Chi Square Distribution',
     xlab='Value', ylab='Iteration')


cat("Acceptance Rate:", second_acceptance_rate)

third_distribution <- numeric(n)
#start value is set to 1
third_distribution[1] <- 1
third_accepted_rate <- 0
for (i in 2:n) {
  value <- rnorm(1, mean=third_distribution[i-1], sd=0.7)
  check_value <- fx(value)/fx(third_distribution[i-1])
  if (runif(1) < check_value) {
    third_distribution[i] <- value
    third_accepted_rate <- third_accepted_rate + 1
```

```r
  } else {
    third_distribution[i] <- third_distribution[i-1]
  }
}
compare_df_c <- data.frame(
  Distribution = "Normal (SD = 0.7)",
  Acceptance_Rate = third_accepted_rate,
  E_X = mean(third_distribution)
)

plot(third_distribution, type='l', main='Normal Distribution with Standard Deviation 0.7',
     xlab='Iteration', ylab='Sampled Value')

hist(third_distribution, breaks=30, main='Histogram : Normal Distribution with Standard Deviation 0.7',
     xlab='Value', ylab='Iteration')

cat("Acceptance Rate:", third_accepted_rate)

compare_df <- rbind(compare_df_a,compare_df_b,compare_df_c)
library(pander)
pander(compare_df)

cat("First Dirtibution- Normal Distribution with Standard Deviation 0.1 = ",mean(first_distribution))
cat("Second Dirtibution- Chi- Square Distribution = ",mean(second_distribution))
cat("First Dirtibution- Normal Distribution with Standard Deviation 0.7 = ",mean(third_distribution))

compare_df$`theoretical E(X)` <- 6

pander(compare_df)
```

Question 2