Lab 2 Probability Distributions

Ashish Chokhani

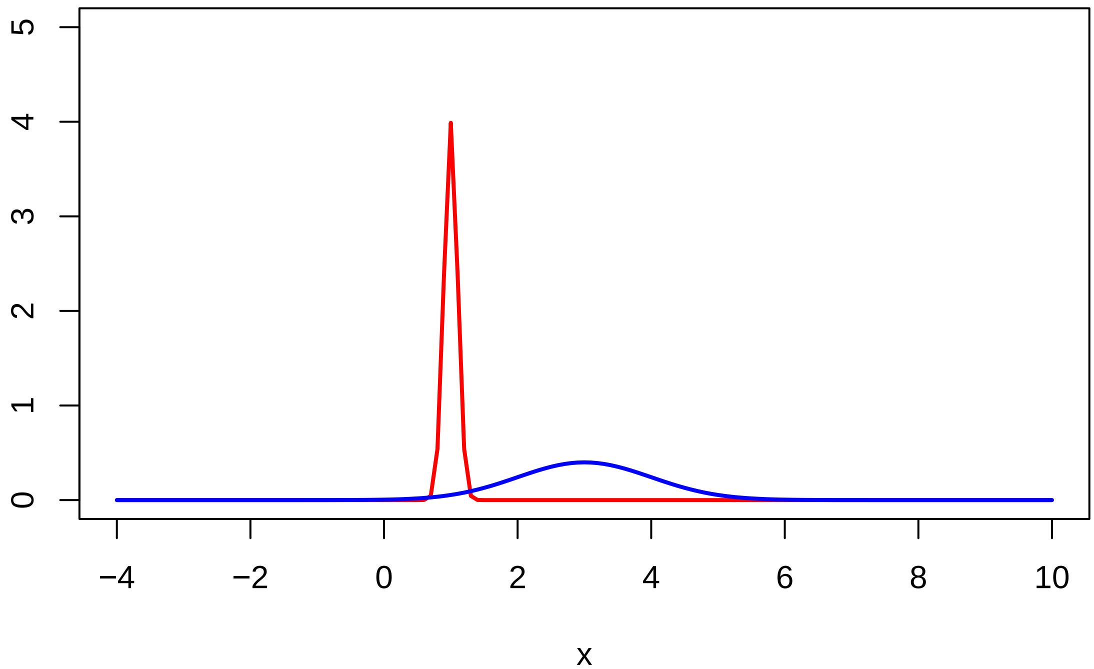
21/1/2025

Contents

Lab 1 Lab Manual Exercise

copy and paste your work by following each example from the lab manual for this exercise

```
rm(list = setdiff(ls(), lsf.str()))

# Plot Normal Distributions with
#-----------------------------------------
# Same standard deviation, different mean
#-----------------------------------------
# Mean 1, sd 1
# Grid of X-axis values
x <- seq(-4, 10, 0.1)

plot(x, dnorm(x, mean = 1, sd = 0.1), type = "l",
     ylim = c(0, 5), ylab = "", lwd = 2, col = "red")
# Mean 3, sd 1
lines(x, dnorm(x, mean = 3, sd = 1), col = "blue", lty = 1, lwd = 2)
```




```
# # Function Syntax
#
# function_name <- function(arg_1, arg_2, ...) {
#    Function body
# }
```

```
# Calculate the 60th %ile of the standard normal.
qnorm(0.6,0,1)
```

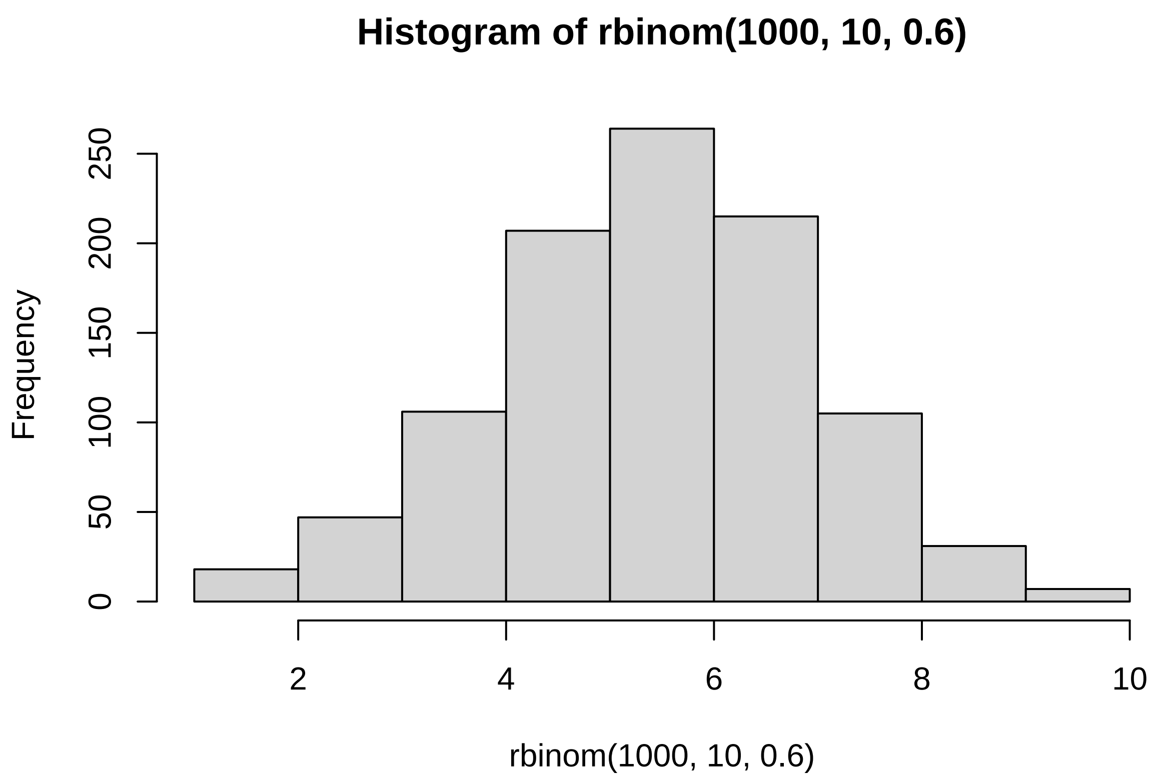
```
## [1] 0.2533471
```

```
# Calculate the probability that a value lies below 0.8 in the standard normal distribution
pnorm(0.8,0,1)
```

```
## [1] 0.7881446
```

```
# Draw 1000 samples of 30 coin tosses with p(heads) = 0.6 # and plot the distribution
# Syntax: rbinom (# observations, # trials per observation, probability of success )
hist(rbinom(1000,30,0.6))
```


```r
# Do the above with only 10 trials per observation
hist(rbinom(1000,10,0.6))
```

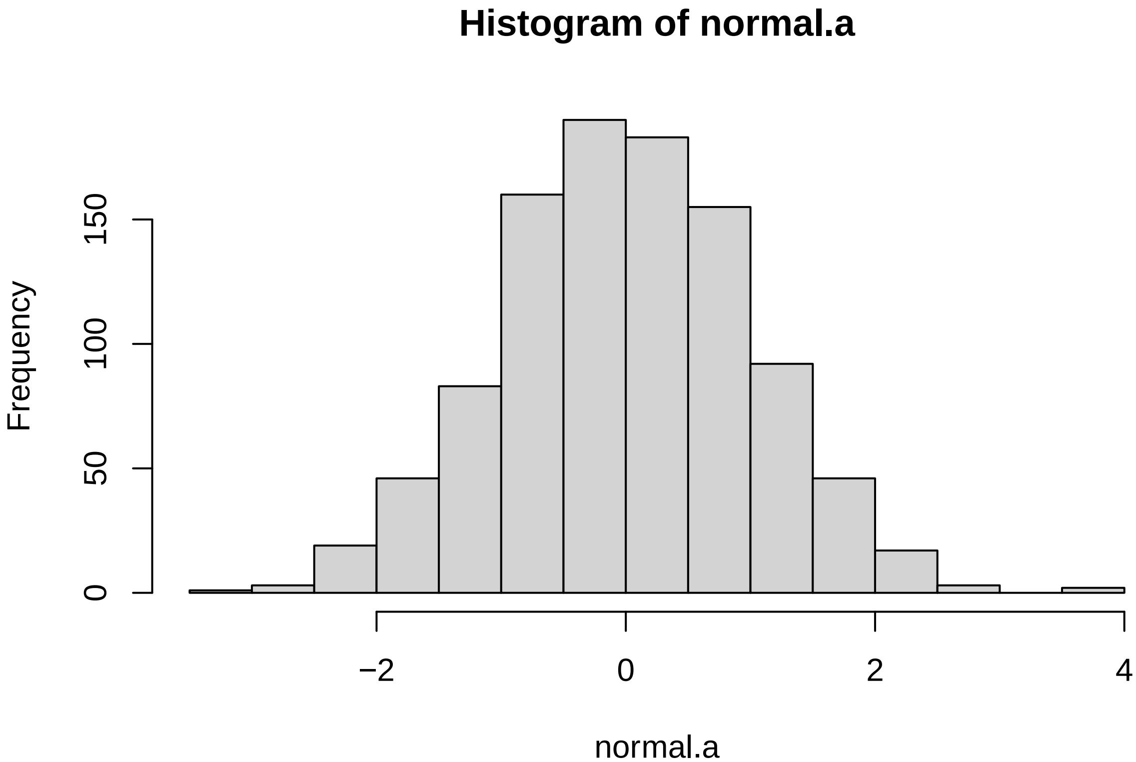



```r
# Do the above with only 5 trials per observation
hist(rbinom(1000,5,0.6))
```

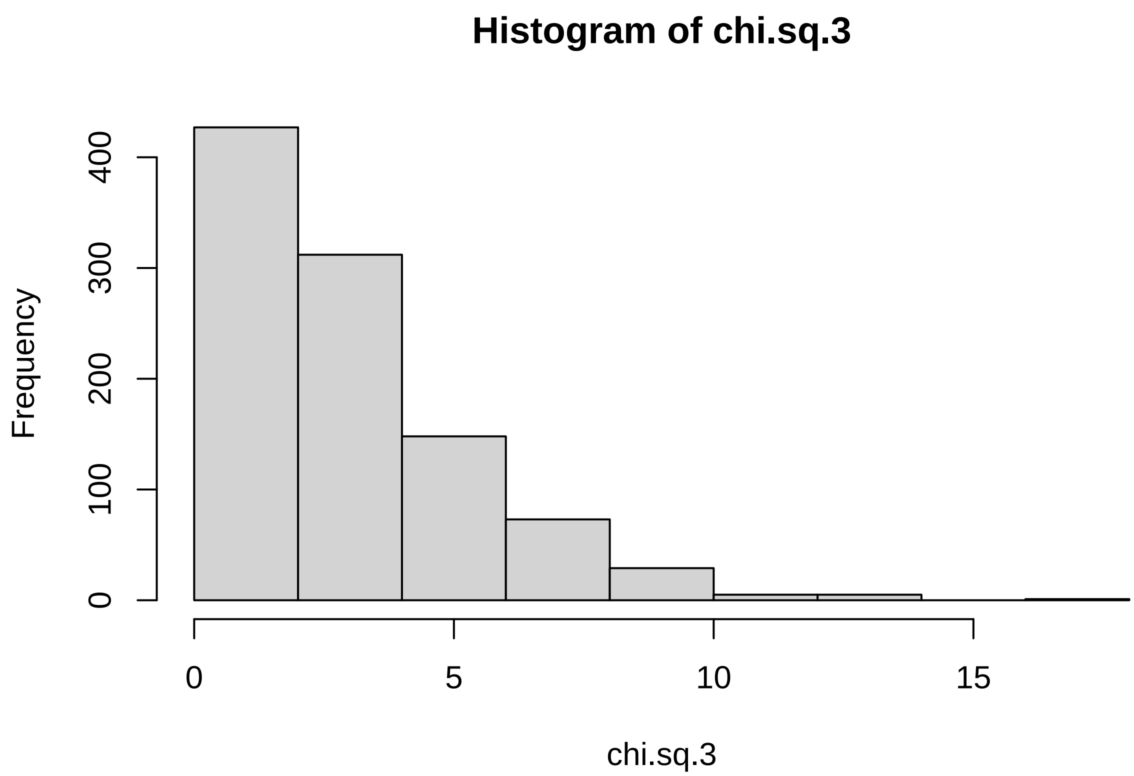

```r
# Do the above with 100 observations and 5 trials per observation
hist(rbinom(100,5,0.6))
```


```r
# Transformations between probability distributions
```

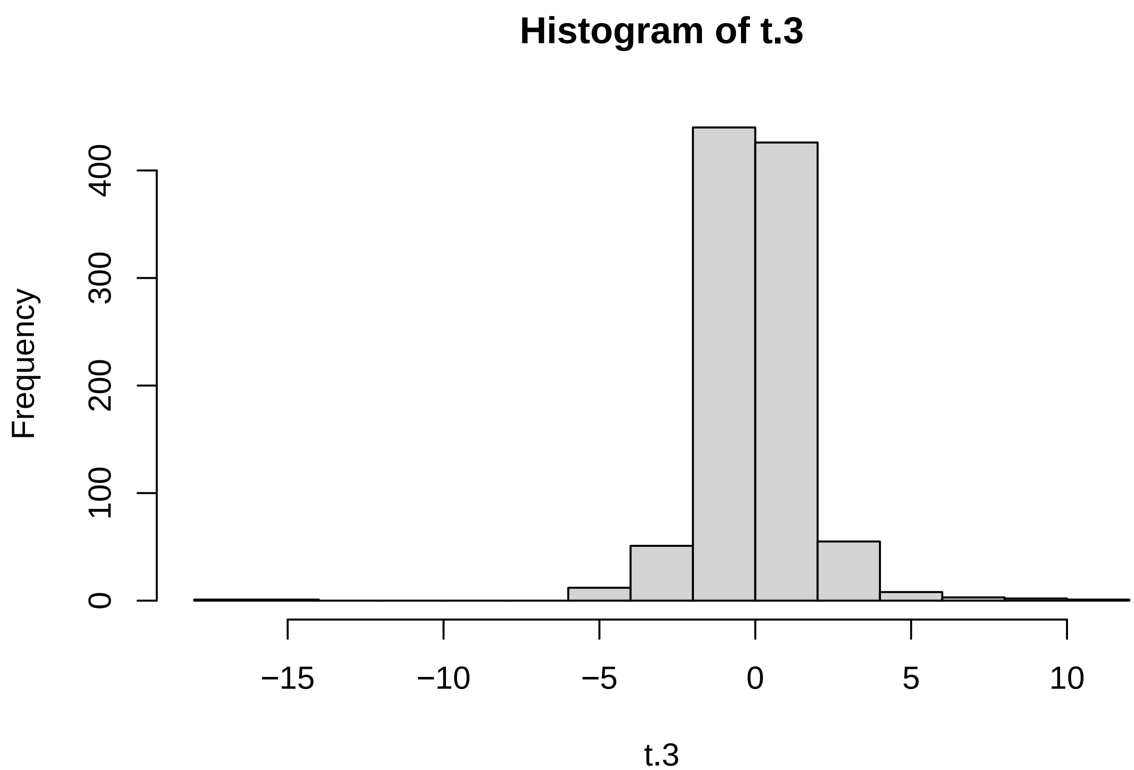
```r
# generate 1000 trials from a normal distribution
normal.a <- rnorm( n=1000, mean=0, sd=1 )
hist( normal.a )
```



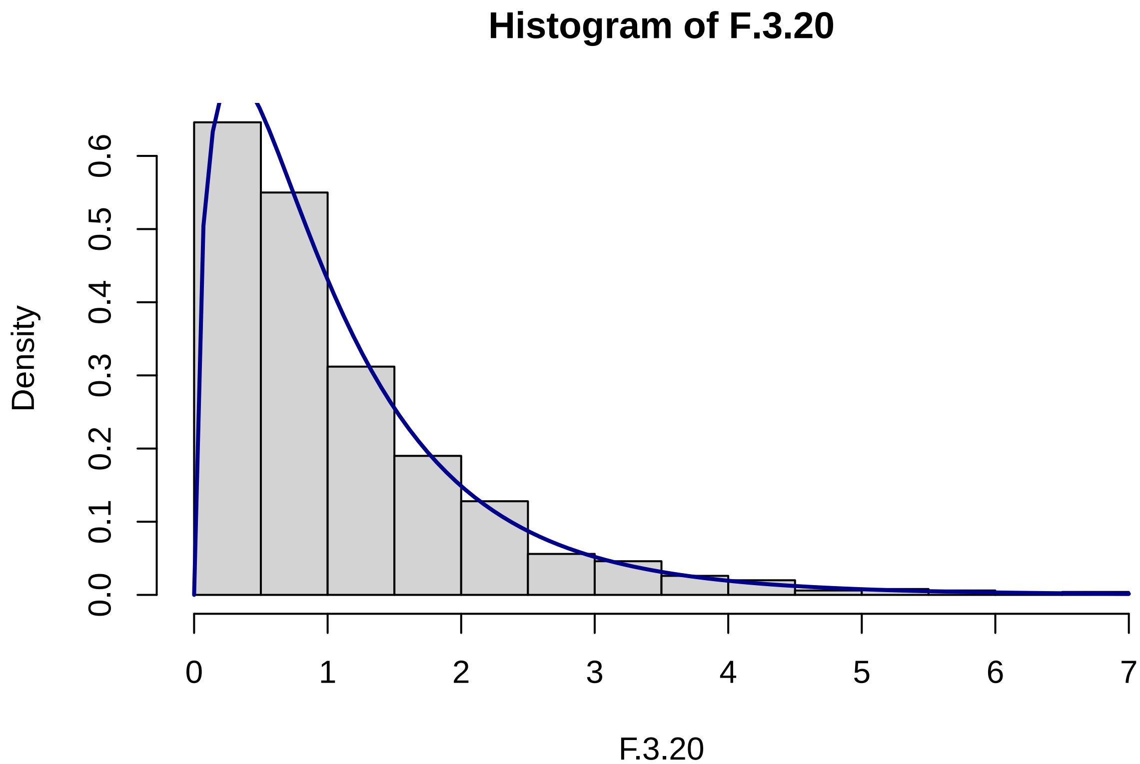

```r
#next, we generate a chi-square distribution with 3 #degrees of freedom:

normal.b <- rnorm( n=1000 )  # another set of normally distributed data
normal.c <- rnorm( n=1000 )  # and another!

# Take the SUM of SQUARES of the above 3 normally distributed variables a, b, and c
chi.sq.3 <- (normal.a)^2 + (normal.b)^2 + (normal.c)^2

# and the resulting chi.sq.3 variable should contain 1000 observations that follow a chi-square distrib

hist(chi.sq.3)
```




```
## Now how do we get to a t-distribution from Normal and chi-sq distributions?
# First, take a scaled chi-sq by dividing it by the degrees of freedom
scaled.chi.sq.3 <- chi.sq.3 / 3
# Then take a normally distributed variable and divide them by the square root of the scaled chi-sq var

normal.d <- rnorm( n=1000 )                    # yet another #set of normally distributed data
t.3 <- normal.d / sqrt( scaled.chi.sq.3 )  # divide by #square root of scaled chi-square to get t
hist (t.3)
```




```
## To get to an F distribution, take the ratio between two scaled chi-sq distributions.
# F distribution with 3 and 20 degrees of freedom:
# first take two chi-sq variables, with 3 dof and 20 dof respectively, and take the ratio:

chi.sq.20 <- rchisq( 1000, 20)                     # generate chi square data with df = 20...
scaled.chi.sq.20 <- chi.sq.20 / 20                 # scale #the chi square variable...
F.3.20 <-  scaled.chi.sq.3  / scaled.chi.sq.20 # take the ratio of the two chi squares...
hist( F.3.20, freq = FALSE)                                  # ... and draw a picture
curve(df(x, 3, 20),
      col="darkblue", lwd=2, add=TRUE, yaxt="n")
```

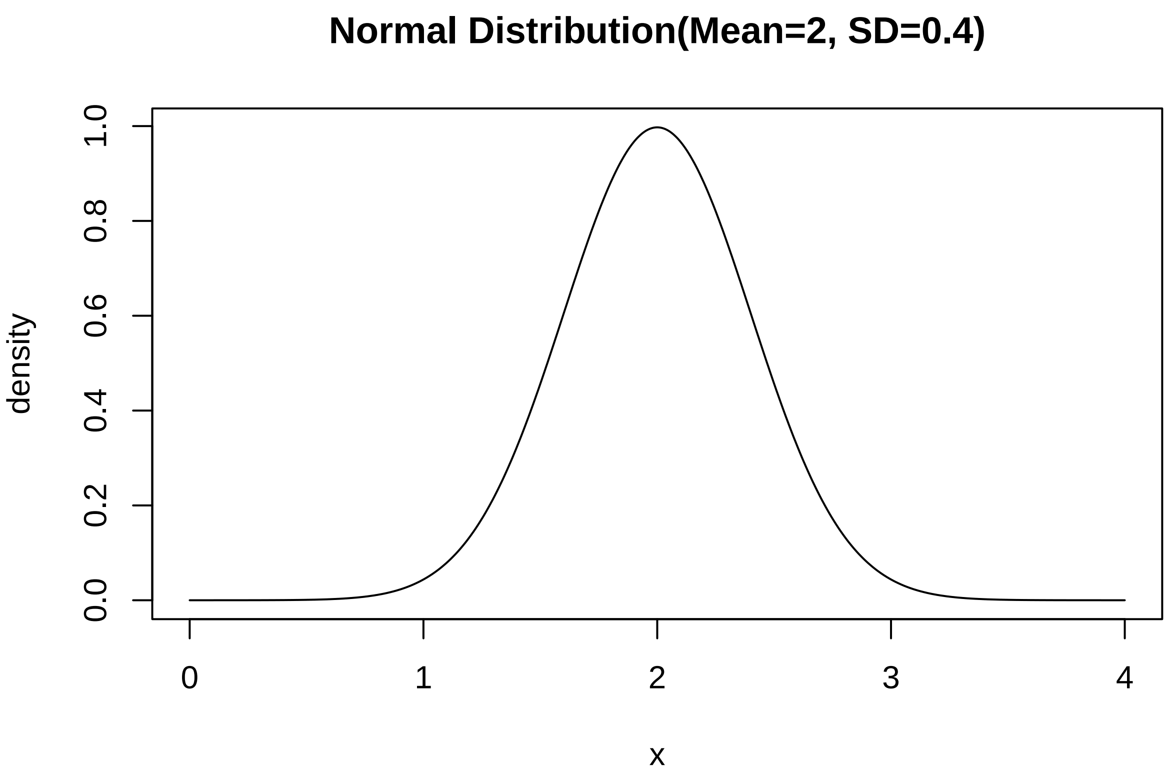



```
## The curve above confirms this looks similar if you use the R built-in function df (just like dnorm, 
```

Lab 1 Generalization exercises

use the code from above to attempt to solve the extra things we ask you do for this assignment

```
# Q1 Plot a normal distribution with mean = 2, s.d. = 0.4

# Ans 1

mean_val <- 2
sd_val <- 0.4
x <- seq(mean_val - 5 * sd_val, mean_val + 5 * sd_val, length.out = 1000)
values <- dnorm(x, mean = mean_val, sd = sd_val)
plot(x, values, type = "l", main = "Normal Distribution(Mean=2, SD=0.4)",
xlab = "x", ylab = "density", )
```




```r
# Q2 Calculate the 85th %ile of the above distribution.

# Ans 2:
percentile_85 <- qnorm(0.85, mean = mean_val, sd = sd_val)
cat("85th %ile of the above distribution :",percentile_85)
```

```
## 85th %ile of the above distribution : 2.414573
```

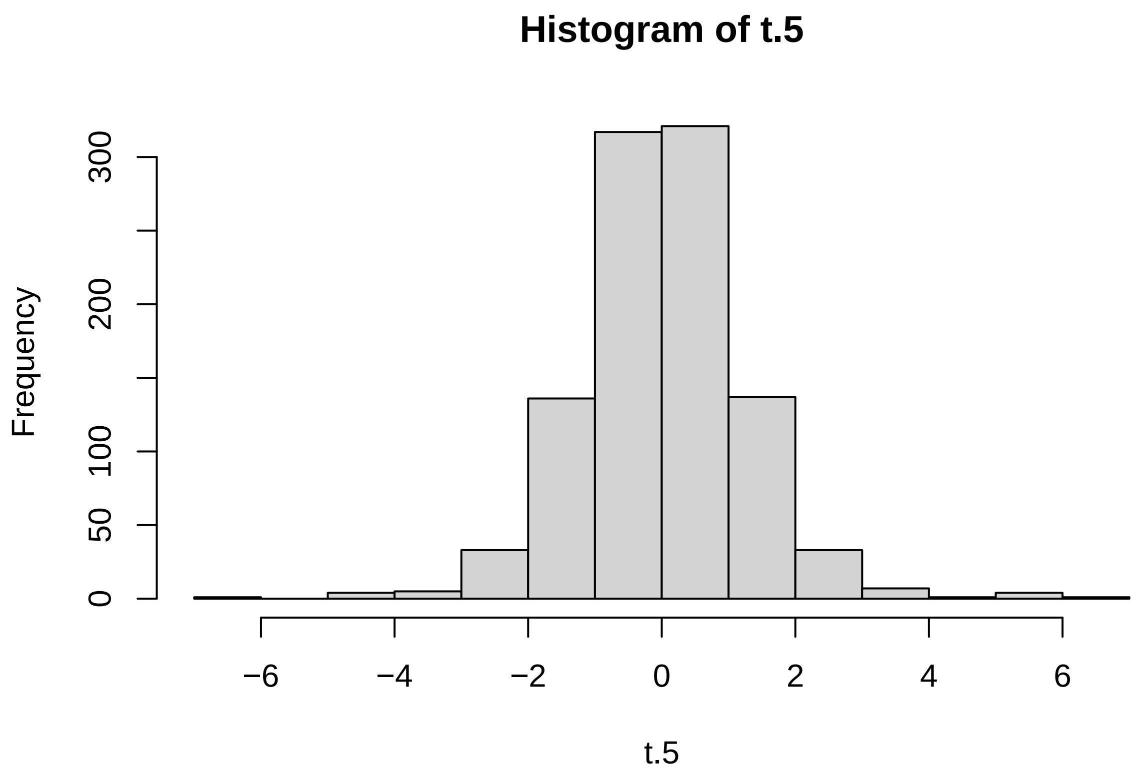
```r
cat("\n")
```

```r
# Q3 Calculate the probability that a value lies in between 1 and 2 given the above distribution

# Ans 3:
probability_between_1_and_2 <- pnorm(2, mean = mean_val, sd = sd_val) - pnorm(1,
mean = mean_val, sd = sd_val)
cat("Probability that a value lies in between 1 and 2 given the above distribution :",probability_betwe
```

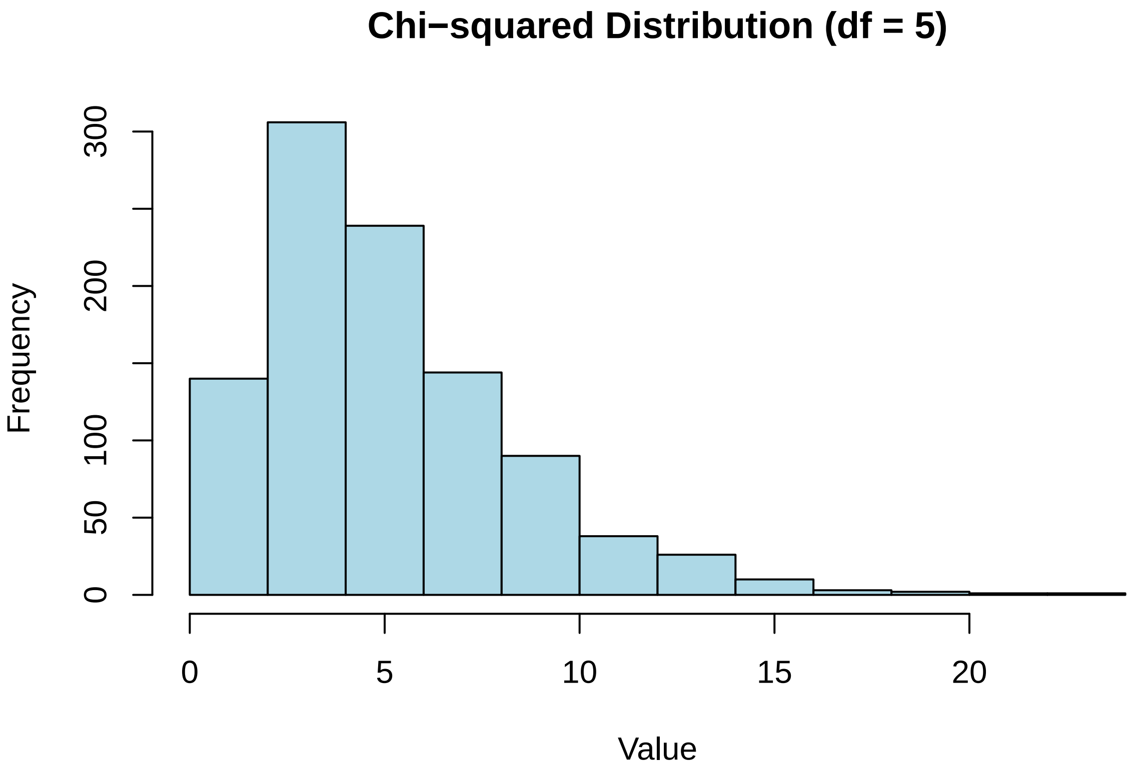
```
## Probability that a value lies in between 1 and 2 given the above distribution : 0.4937903
```

```r
# Q4 Plot a simulated t-distribution with 5 degrees of freedom.

# Ans 4:

normal.1 <- rnorm( n=1000, mean=0, sd=1)
normal.2 <- rnorm( n=1000 )
normal.3 <- rnorm( n=1000 )
normal.4 <- rnorm( n=1000 )
normal.5 <- rnorm( n=1000 )
chi.sq.5 <- (normal.1)^2 + (normal.2)^2 + (normal.3)^2 + (normal.4)^2 + (normal.5)^2
scaled.chi.sq.5 <- chi.sq.5 / 5
normal.d <- rnorm( n=1000 )
t.5 <- normal.d / sqrt( scaled.chi.sq.5 )
```

```r
hist (t.5)
```




```r
# Q5 Plot a chi-sq distribution with 5 degrees of freedom.

# Ans 5.

normal.1 <- rnorm(n=1000)
normal.2 <- rnorm(n=1000)
normal.3 <- rnorm(n=1000)
normal.4 <- rnorm(n=1000)
normal.5 <- rnorm(n=1000)

# Calculate the chi-squared variable
chi.sq.5 <- (normal.1)^2 + (normal.2)^2 + (normal.3)^2 + (normal.4)^2 + (normal.5)^2

# Plot the histogram
hist(chi.sq.5, main="Chi-squared Distribution (df = 5)", xlab="Value", ylab="Frequency", col="lightblue
```



Lab 1

Written answer question

Write your answer here.