Lab 1 Probability Distributions

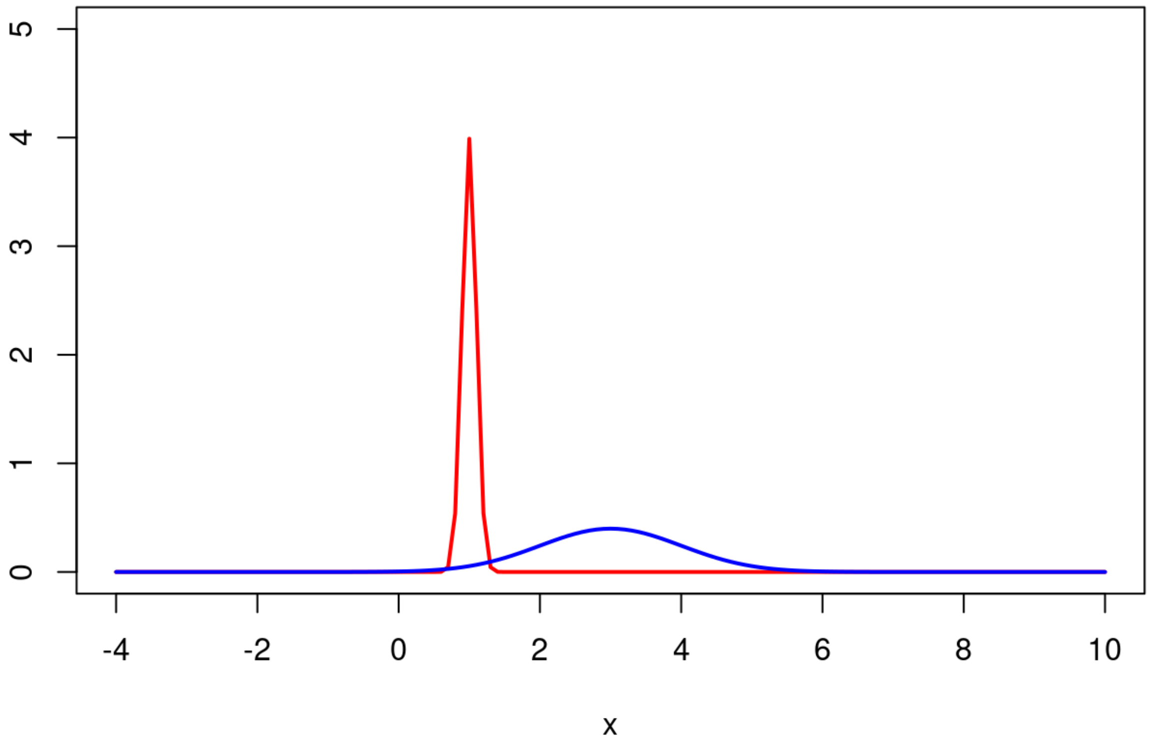
Srujana Vanka

18/01/2024

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
# }
```

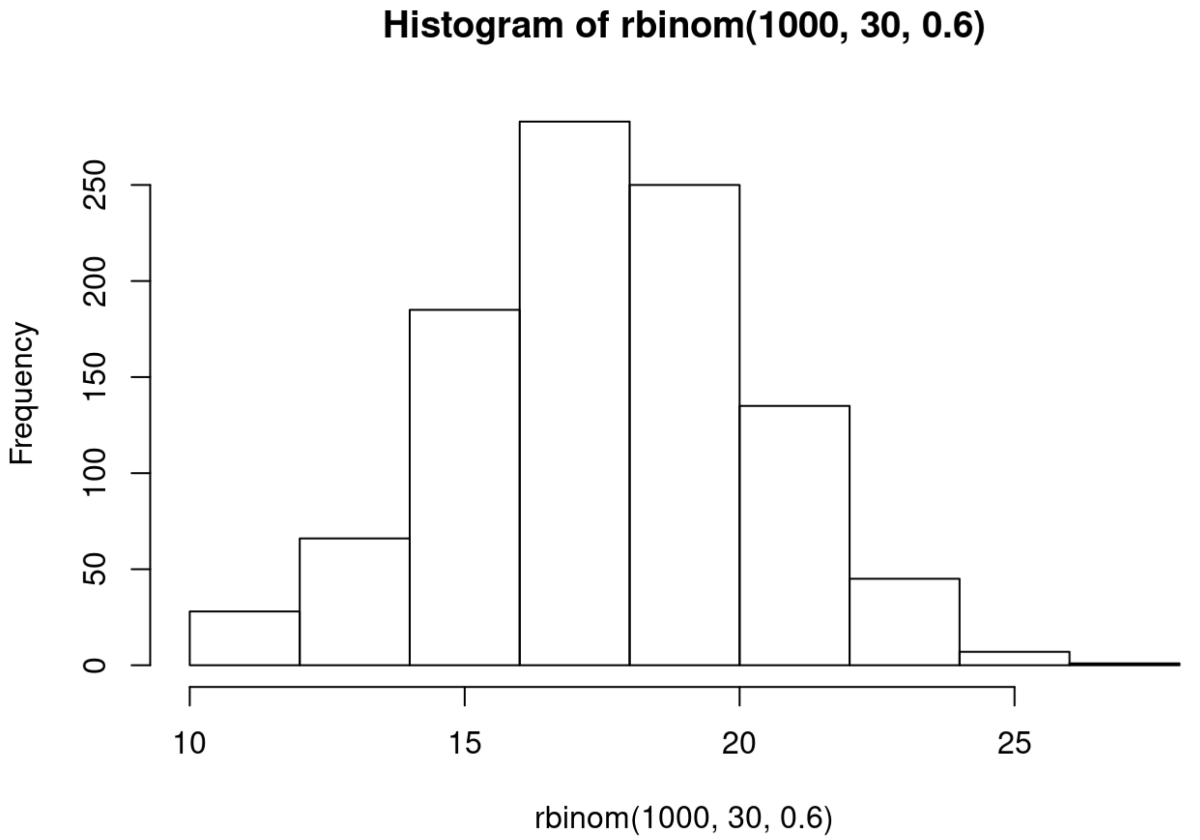
```
# Calculate the 60th %ile of the standard normal.
qnorm(0.6,0,1)
```

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

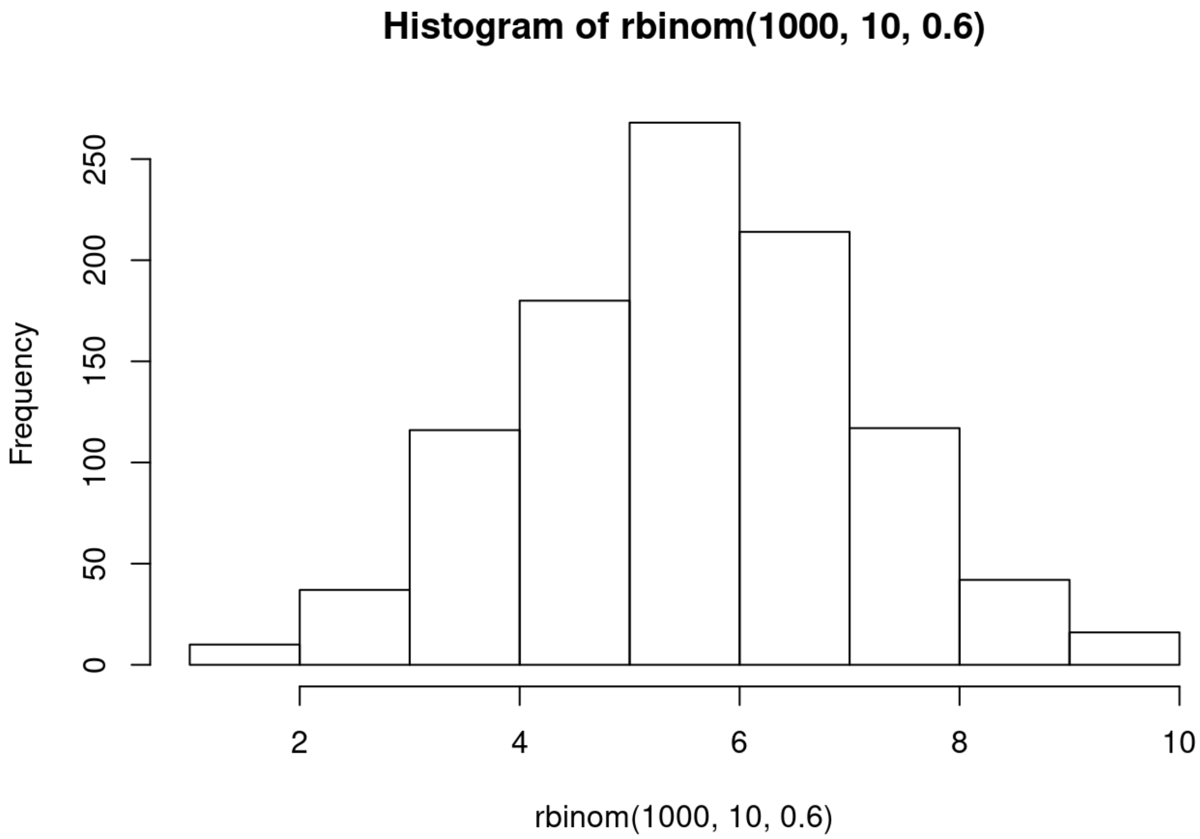
```
# Calculate the probability that a value lies below 0.8 in the standard normal d
istribution
pnorm(0.8,0,1)
```

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

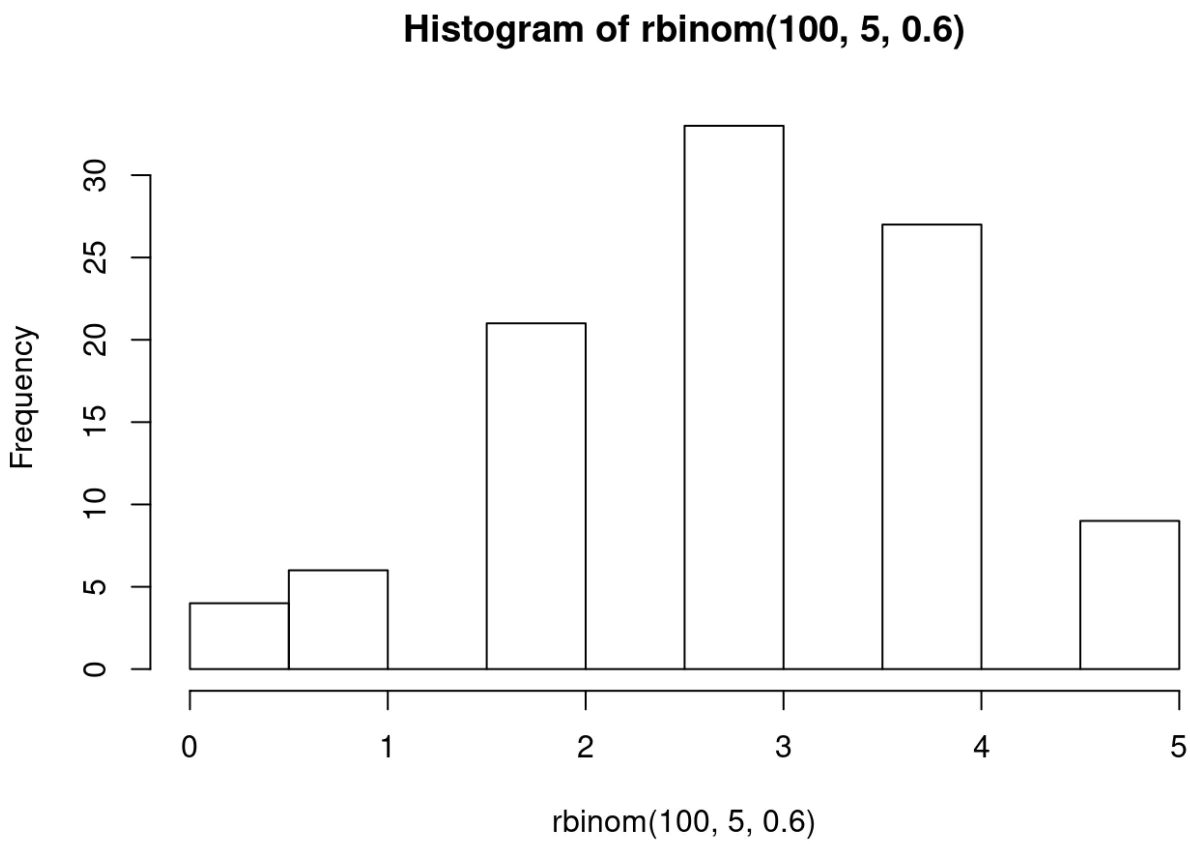
```
# Draw 1000 samples of 30 coin tosses with p(heads) = 0.6 # and plot the distrib
ution
# Syntax: rbinom (# observations, # trials per observation, probability of succe
ss )
hist(rbinom(1000,30,0.6))
```



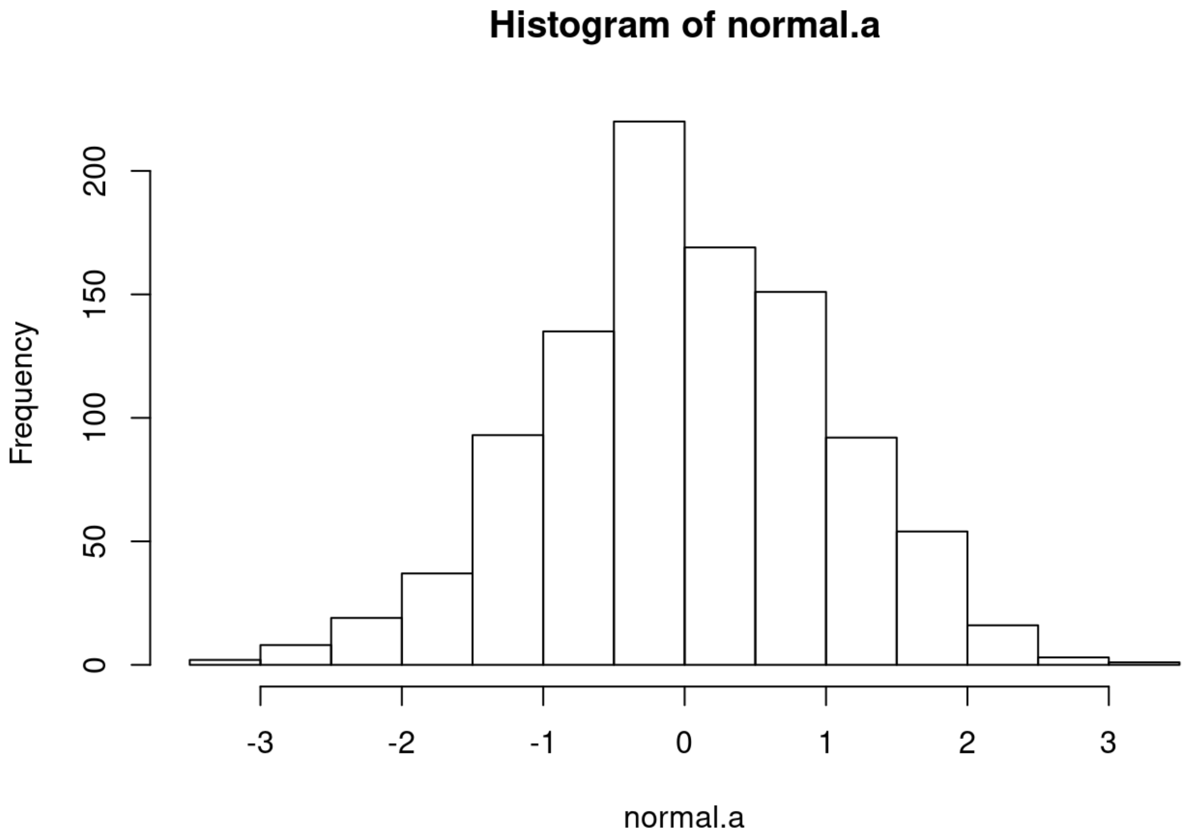

```
# Do the above with only 10 trials per observation
hist(rbinom(1000,10,0.6))
```



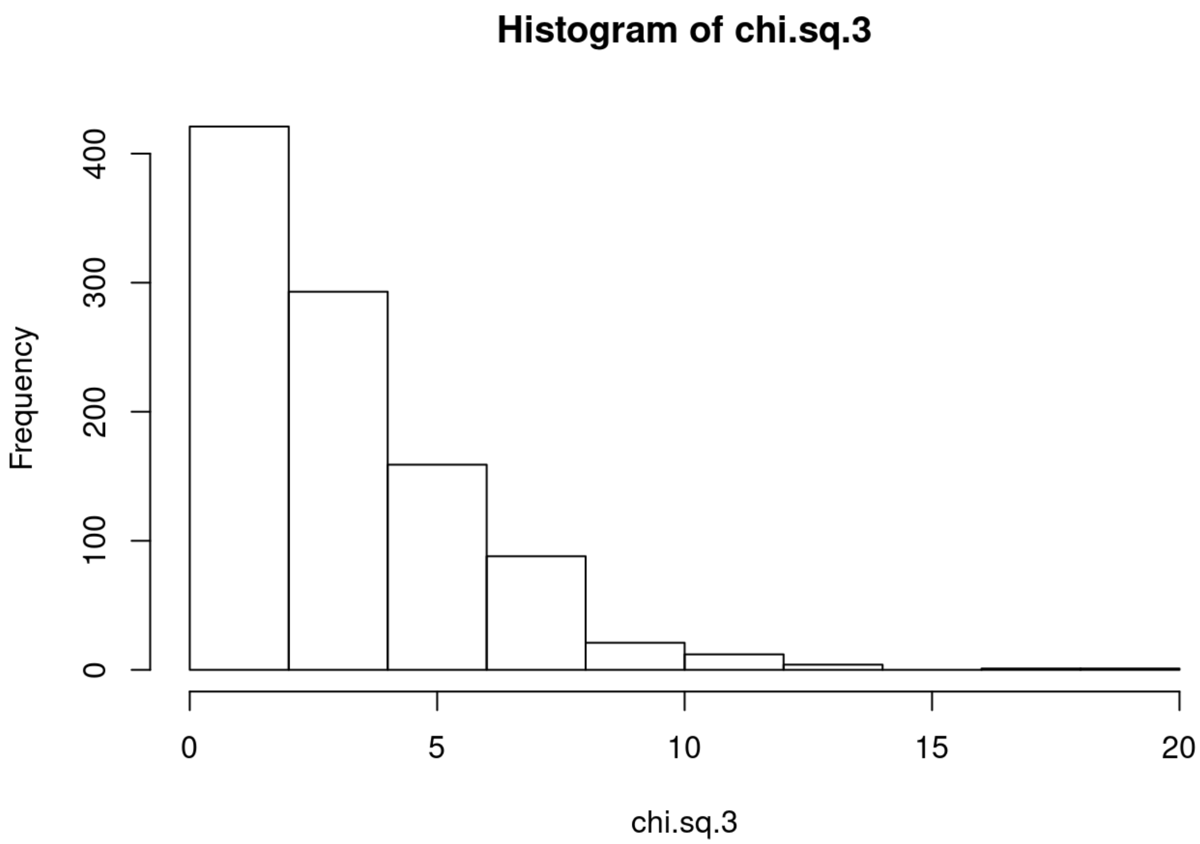
```
# Do the above with 100 observations and 5 trials per observation
hist(rbinom(100,5,0.6))
```



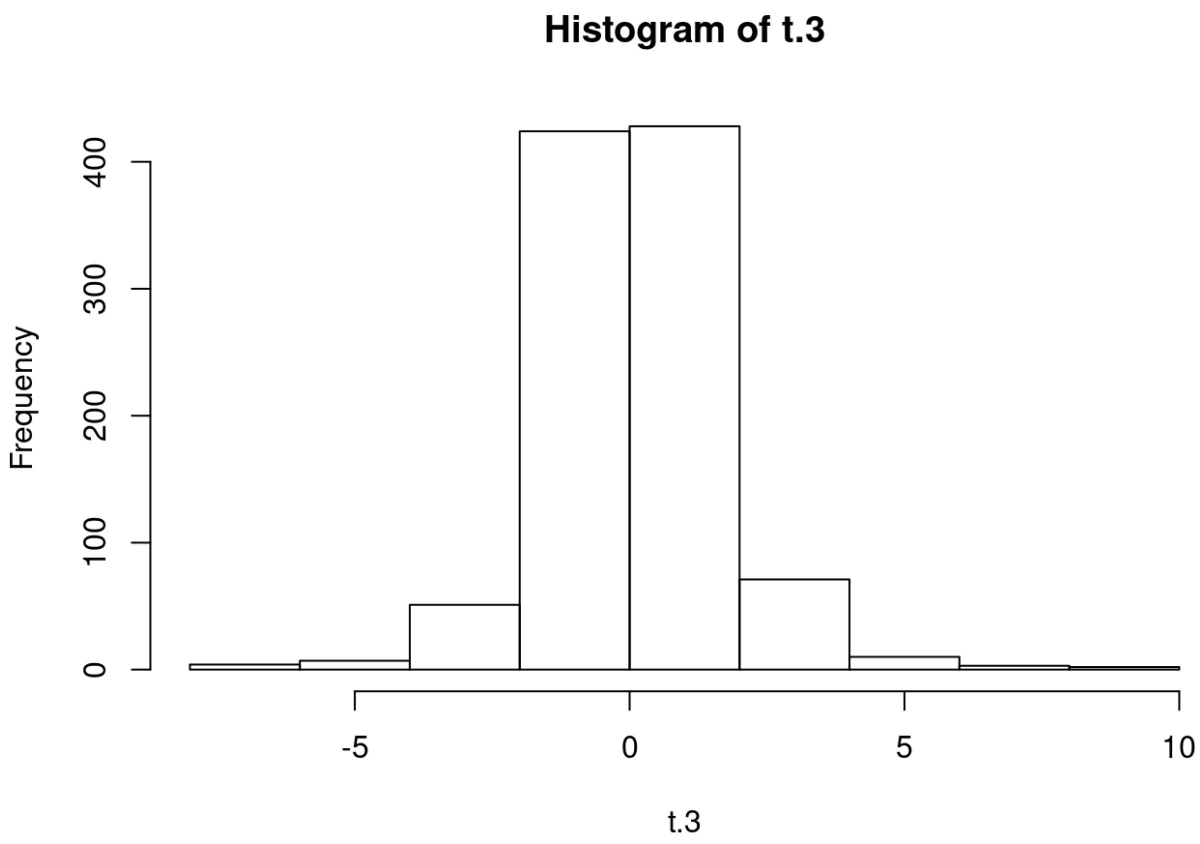

```
# Transformations between probability distributions

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



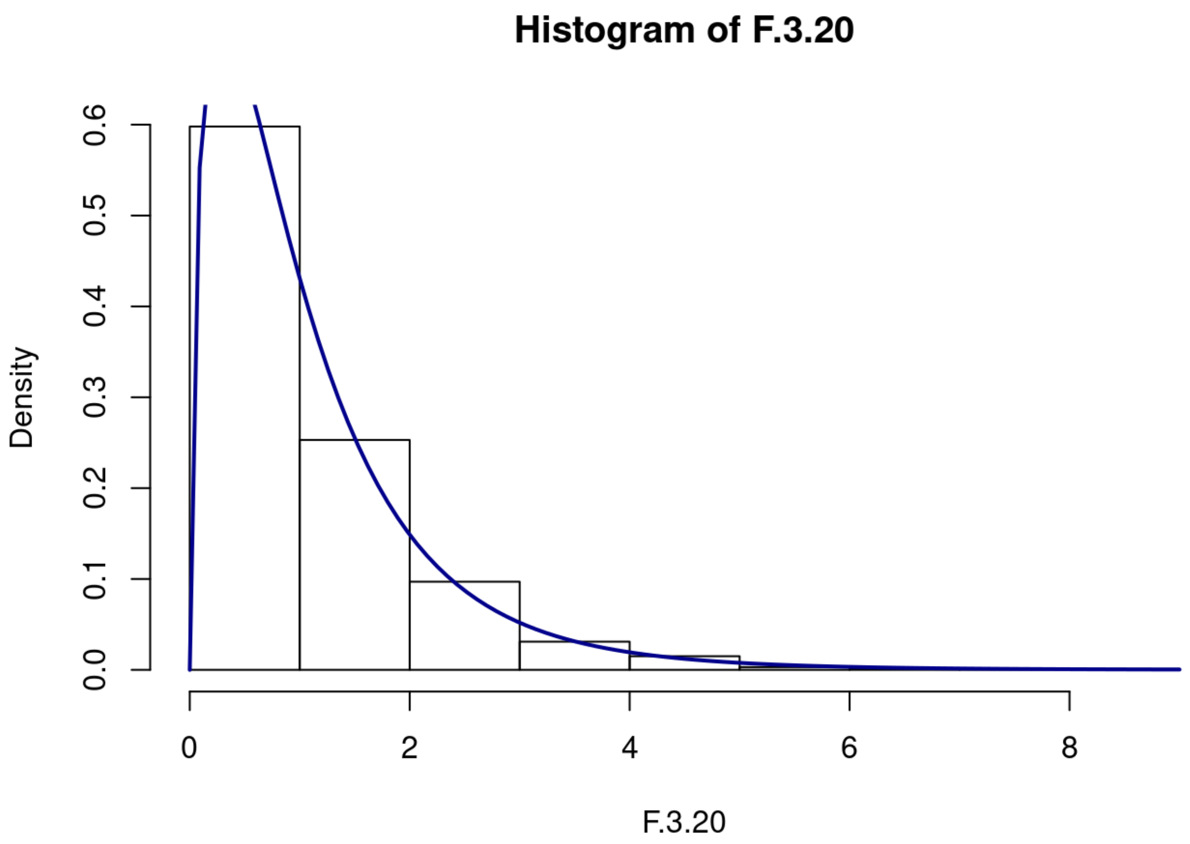

```
#next, we generate a chi-square distribution with 3 #degrees of freedom:

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, an
d 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 foll
ow a chi-square distribution with 3 degrees of freedom. You can use the hist() f
unction to have a look at these observations yourself

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 o
f the scaled chi-sq variable to get a t-distribution with the same degrees of fr
eedom

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




```
## To get to an F distribution, take the ratio between two scaled chi-sq distrib
utions.
# 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 d
f = 20...
scaled.chi.sq.20 <- chi.sq.20 / 20              # scale #the chi square variabl
e...
F.3.20 <-  scaled.chi.sq.3  / scaled.chi.sq.20 # take the ratio of the two chi s
quares...
hist( F.3.20, freq = FALSE)                                  # ... and draw a pic
ture
curve(df(x, 3, 20),
      col="darkblue", lwd=2, add=TRUE, yaxt="n")
```

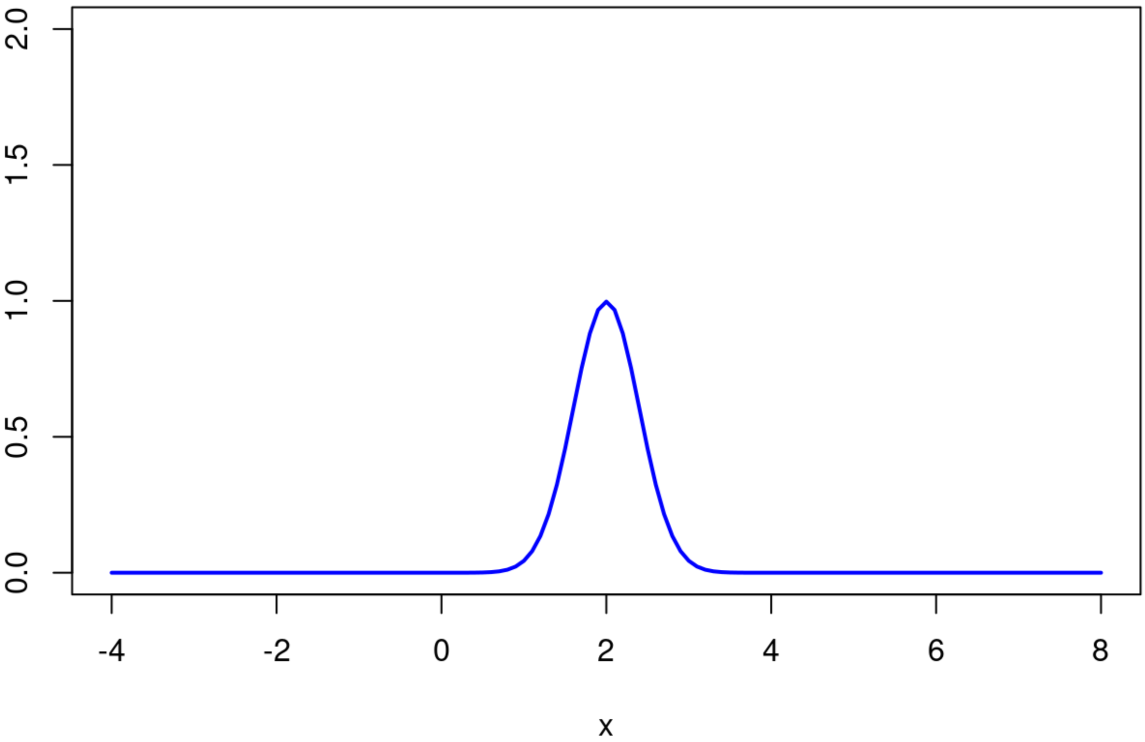



```
## The curve above confirms this looks similar if you use the R built-in functio
n df (just like dnorm, but for the F distribution)
```

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
x <- seq(-4, 8, 0.1)
plot(x, dnorm(x, mean = 2, sd = 0.4), type = "l", ylim = c(0, 2), ylab = "", lwd
= 2, col = "blue")
```



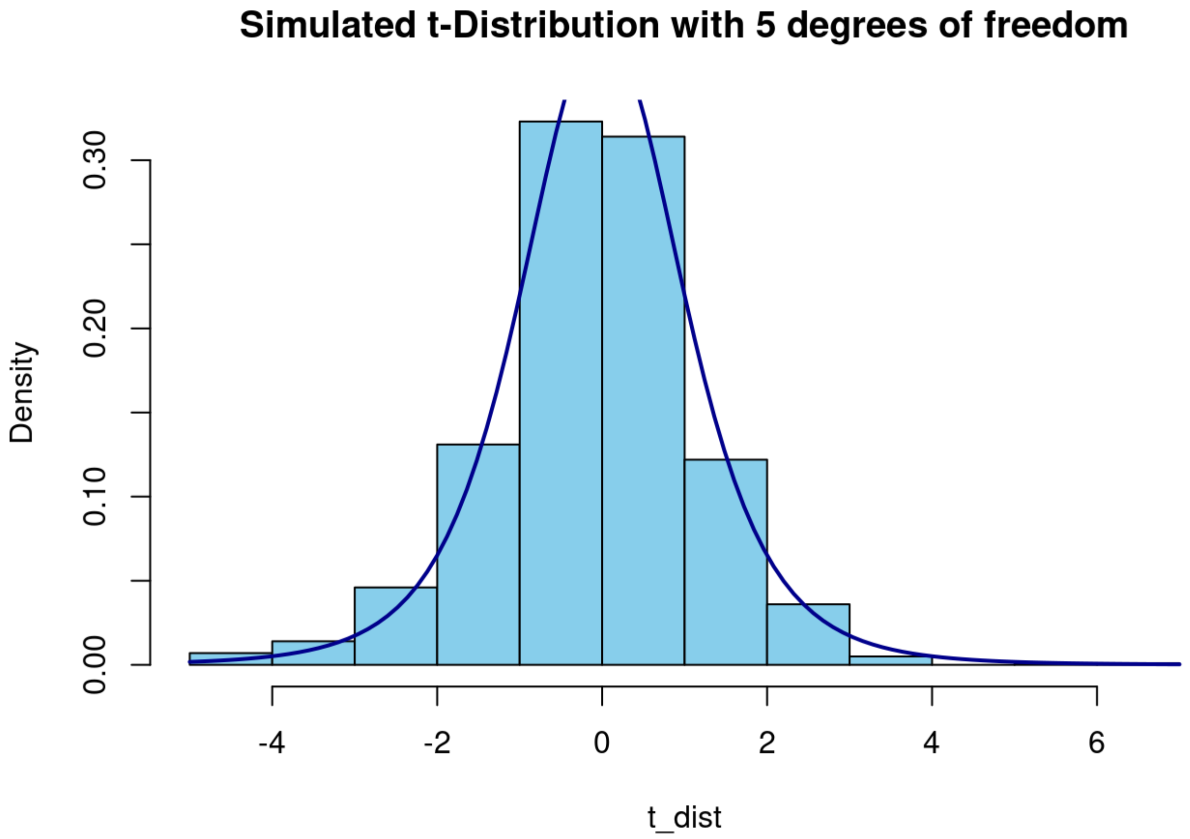

```
# Q2 Calculate the 85th %ile of the above distribution.
qnorm(0.85, mean = 2, sd = 0.4)
```

```
## [1] 2.414573
```

```
# Q3 Calculate the probability that a value lies in between 1 and 2 given the ab
ove distribution
pnorm(2, mean = 2, sd = 0.4) - pnorm(1, mean = 2, sd = 0.4)
```

```
## [1] 0.4937903
```

```
# Q4 Plot a simulated t-distribution with 5 degrees of freedom.
t_dist <- rt(1000, df = 5)
hist(t_dist, main = "Simulated t-Distribution with 5 degrees of freedom", col =
"skyblue", freq = FALSE)
curve(dt(x, df = 5), col = "darkblue", lwd = 2, add = TRUE)
```



Lab 1 Written answer question

Write your answer here.