Name: Shreeya Chatterji
Class: CO16
Roll Number: 102103447

Lab Assignment 7

(1) Use the rt(n, df) function in r to investigate the t-distribution for n = 100 and df = n – 1 and plot the histogram for the same.

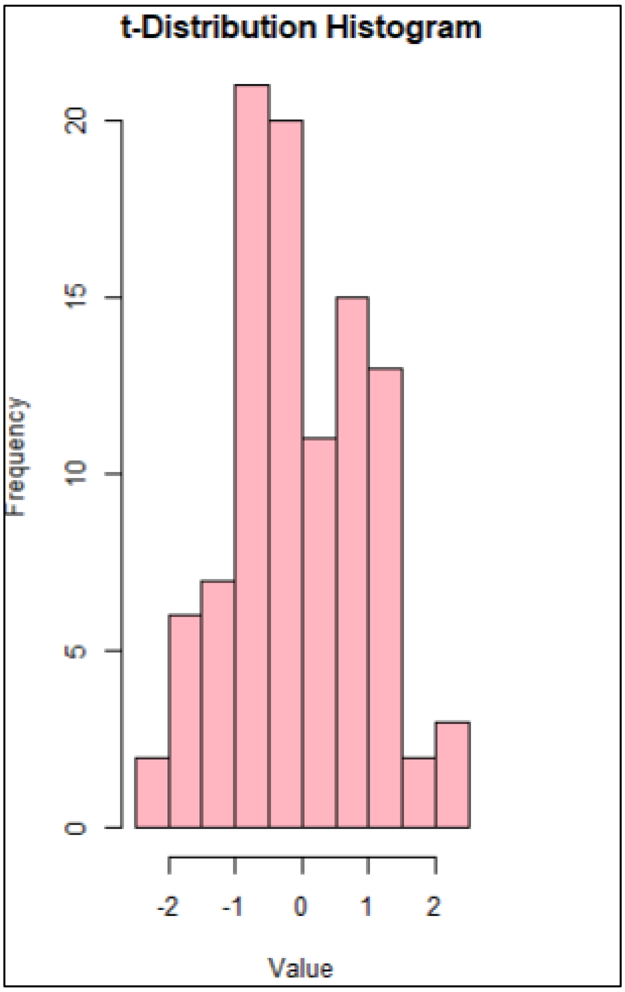
CODE:

```
#Q1
#set the parameters
n<-100
df <- n-1
#Generate random samples from the t-distribution
t_samples<- rt(n,df)
print(t_samples)
#Plot a histogram of the generated data
hist(t_samples, main="t-Distribution Histogram", xlab="Value", ylab="Frequency", col="lightpink", border="black")
```

OUTPUT:

```
> print(t_samples)
  [1] -1.18328119 -1.62892725 -0.56922901  2.02288948 -0.75502740 -1.40174475 -0.89563425  0.68410212 -1.30526593 -0.41272695
 [11]  0.60016551  0.90420299  1.33707917 -0.18523767 -0.11973913 -0.99701605  0.73378431  1.14162339 -0.64106190 -0.15946542
 [21] -1.72714018 -0.31625811 -0.23374461 -0.84797308  1.06427916  1.33954166 -2.21625003 -1.18274174 -0.74691751  1.40175445
 [31]  0.48462040 -1.19997373 -0.67437638  0.68826238  1.22086465  0.92269730  0.37172796 -0.91564275  1.31255406  0.11602252
 [41] -0.84129933 -0.95657087  1.20536715 -0.08974044 -0.87177898 -0.89755059  1.16349435  0.12467832 -0.47688665 -0.21959332
 [51] -0.60341159  0.66904195 -0.08357976  1.13589916 -0.25357310 -0.25467470  1.56921583  0.24415316 -0.16940100  0.72510363
 [61]  2.37628492  0.87315639  0.68695871 -0.46216589 -1.24839879 -0.15623366  0.81687045  0.64436969 -0.60987448  0.10859402
 [71] -0.76806765 -0.44917497 -1.98102457 -0.20230668  1.13611965  1.53144067 -0.07005447 -2.11316017 -0.18708096 -1.93978855
 [81]  0.03746273  0.19372676 -1.33573655  0.77222963 -1.52654588 -1.58245873 -0.45718903 -0.50526442  1.30472358  1.14560076
 [91]  2.14851087  0.72765552  0.07161478  0.53186855  0.24538654 -0.67722358 -0.56004831 -0.89304853  0.46773297 -0.52616187
```

PLOT:



Name: Shreeya Chatterji
Class: CO16
Roll Number: 102103447

(2) Use the rchisq(n, df) function in r to investigate the chi-square distribution with n = 100 and df = 2, 10, 25.

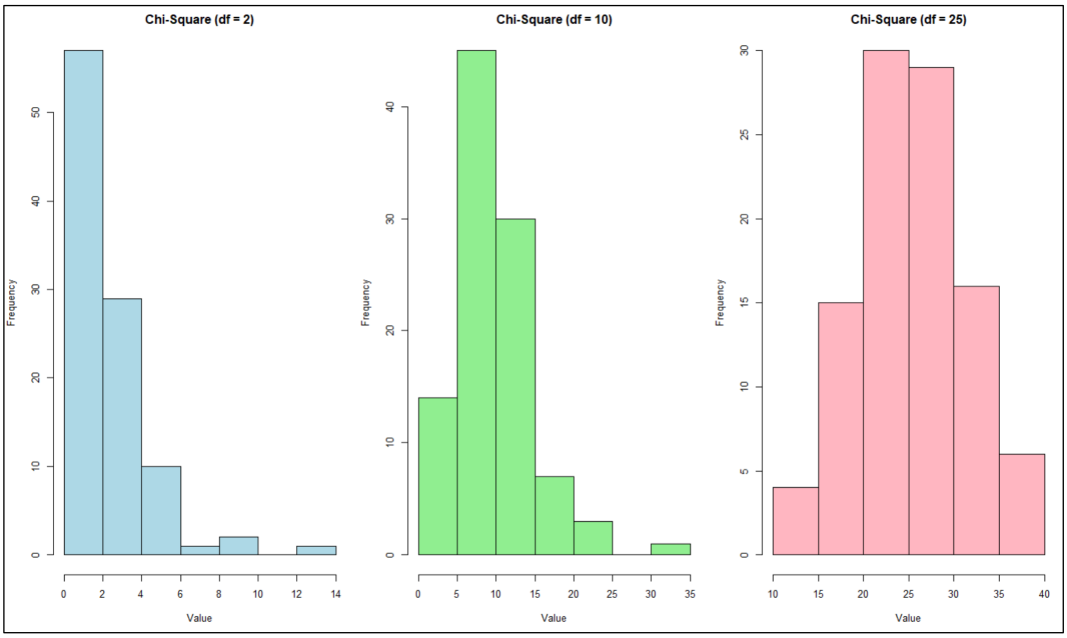
CODE:

```r
#Q2
#Set the parameters
n<-100
dfs<-c(2,10,25)#degrees of freedom

s1=rchisq(n,dfs[1])
s2=rchisq(n,dfs[2])
s3=rchisq(n,dfs[3])

mean(s1)
var(s1)

mean(s2)
var(s2)

mean(s3)
var(s3)
#Generate random samples from the chi-square distribution for each df
#chi_squared_samples<-lapply(dfs, function(df) rchisq(n,df))

#create the histograms for each set of samples
par(mfrow=c(1,3))#Arrange plots in a row

hist(s1, main = "Chi-Square (df = 2)", xlab = "Value", ylab = "Frequency", col = "lightblue")
hist(s2, main = "Chi-Square (df = 10)", xlab = "Value", ylab = "Frequency", col = "lightgreen")
hist(s3, main = "Chi-Square (df = 25)", xlab = "Value", ylab = "Frequency", col = "lightpink")
```

OUTPUT:

```
> mean(s1)
[1] 1.891241
> var(s1)
[1] 3.919405
>
> mean(s2)
[1] 9.916624
> var(s2)
[1] 18.76067
>
> mean(s3)
[1] 24.13628
> var(s3)
[1] 48.91953
> #Generate random s
```

PLOTS:



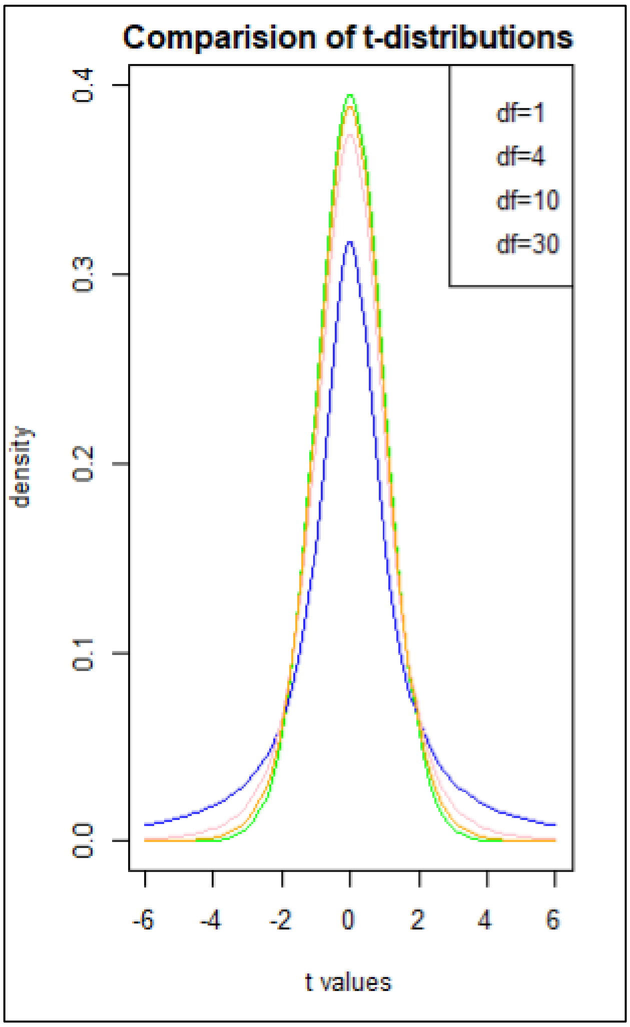
Name: Shreeya Chatterji
Class: CO16
Roll Number: 102103447

(3) Generate a vector of 100 values between -6 and 6. Use the dt() function in r to find the values of a t-distribution given a random variable x and degrees of freedom 1,4,10,30. Using these values plot the density function for students t-distribution with degrees of freedom 30. Also shows a comparison of probability density functions having different degrees of freedom (1,4,10,30).

CODE:

```r
#Q3
#To generate a vector of 100 values between -6 and 6
x<-seq(-6,6,length.out=100)
#Degrees of freedom
df<-c(1,4,10,30)
colors<-c("blue","pink","orange","green")
#Calculate the t-distribution values
t_dist_df1<-dt(x,df[1])
t_dist_df2<-dt(x,df[2])
t_dist_df3<-dt(x,df[3])
t_dist_df4<-dt(x,df[4])
#Plot the comaprision of t-distributions
plot(x, t_dist_df4, type="l",xlab="t values", ylab="density", main="Comparision of t-distributions", col= colors[4])

#Add lines for other degrees of freedom
lines(x,t_dist_df1, col=colors[1])
lines(x,t_dist_df2, col=colors[2])
lines(x,t_dist_df3, col=colors[3])

#Add a legend to the plot
legend("topright", legend=c("df=1","df=4","df=10","df=30"))
```

OUTPUT/PLOT:



Name: Shreeya Chatterji
Class: CO16
Roll Number: 102103447

(4) Write a r-code

(i) To find the 95th percentile of the F-distribution with (10, 20) degrees of freedom.

(ii) To calculate the area under the curve for the interval [0, 1.5] and the interval [1.5, +∞) of a F-curve with v1 = 10 and v2 = 20 (USE pf()).

(iii) To calculate the quantile for a given area (= probability) under the curve for a F-curve with v1 = 10 and v2 = 20 that corresponds to q = 0.25, 0.5, 0.75 and 0.999. (use the qf())

(iv) To generate 1000 random values from the F-distribution with v1 = 10 and v2 = 20 (use rf())and plot a histogram.

CODE:

```
#Q4
v1<-10
v2<-20

#(i) find the 95th percentile of the f-distribution

percentile_95<-qf(0.95,df1=v1,df2=v2)
cat("95th percentile of the F-distribution :",percentile_95,"\n")

#(ii)calculate the area under the curve for the given intervals
area_interval_1<-pf(1.5,df1 =v1,df2=v2,lower.tail = TRUE) #[0,1.5]
area_interval_2<-1-area_interval_1
cat("Area under the curve for [0,1.5]:",area_interval_1,"\n")
cat("Area under the curve for [1.5,+inf]:",area_interval_2,"\n")

#(iii) calculate the quantiles for diff probab
quantile_25<-qf(0.25,df1=v1,df2=v2)
quantile_50<-qf(0.5,df1=v1,df2=v2)
quantile_75<-qf(0.75,df1=v1,df2=v2)
quantile_999<-qf(0.999,df1=v1,df2=v2)

cat("Quantile for p = 0.25",quantile_25,"\n")
cat("Quantile for p = 0.25",quantile_50,"\n")
cat("Quantile for p = 0.75",quantile_75,"\n")
cat("Quantile for p = 0.999",quantile_999,"\n")

#(iv) generate random values and plot a histogram
#set.seed(123) for reproducability

x<-rf(1000,df1=v1,df2=v2)
hist(x,breaks='scott',freq=FALSE,xlim=c(0,3),ylim=c(0,1),xlab="", col="orange")
```

Name: Shreeya Chatterji
Class: CO16
Roll Number: 102103447

OUTPUT:

```
> #(i) find the 95th percentile of the f-distribution
>
> percentile_95<-qf(0.95,df1=v1,df2=v2)
> cat("95th percentile of the F-distribution :",percentile_95,"\n")
95th percentile of the F-distribution : 2.347878
>
> #(ii)calculate the area under the curve for the given intervals
> area_interval_1<-pf(1.5,df1 =v1,df2=v2,lower.tail = TRUE) #[0,1.5]
> area_interval_2<-1-area_interval_1
> cat("Area under the curve for [0,1.5]:",area_interval_1,"\n")
Area under the curve for [0,1.5]: 0.7890535
> cat("Area under the curve for [1.5,+inf]:",area_interval_2,"\n")
Area under the curve for [1.5,+inf]: 0.2109465
>
> #(iii) calculate the quantiles for diff probab
> quantile_25<-qf(0.25,df1=v1,df2=v2)
> quantile_50<-qf(0.5,df1=v1,df2=v2)
> quantile_75<-qf(0.75,df1=v1,df2=v2)
> quantile_999<-qf(0.999,df1=v1,df2=v2)
>
> cat("Quantile for p = 0.25",quantile_25,"\n")
Quantile for p = 0.25 0.6563936
> cat("Quantile for p = 0.25",quantile_50,"\n")
Quantile for p = 0.25 0.9662639
> cat("Quantile for p = 0.75",quantile_75,"\n")
Quantile for p = 0.75 1.399487
> cat("Quantile for p = 0.999",quantile_999,"\n")
Quantile for p = 0.999 5.075246
>
> #(iv) generate random values and plot a histogram
> #set.seed(123) for reproducability
>
> x<-rf(1000,df1=v1,df2=v2)
> hist(x,breaks='scott',freq=FALSE,xlim=c(0,3),ylim=c(0,1),xlab="", col="orange")
```

PLOT: