

Problem-1

Samar Moussa/NIU:1661270 _ Jamia Begum/NIU: 1676891

2022-12-30

```
#exercise:1. Remember one of the examples presented in class:
```

```
#The first row is chest circumference (in inches) of five subjects.
```

```
#Let us call this X.
```

```
#The second row is the respective total volumes of air that can be breathed in  
#and out in one minute (in liters) for the same five subjects. Let us call this Y.
```

```
#X: 39, 29, 60, 40, 32
```

```
#Y: 11, 5, 20, 8, 6
```

```
#Download the package combinat in order to use the function permn().
```

```
#Perform an exact test for  $H_0 : \rho = 0$  against  $H_1 : \rho > 0$ .
```

```
#Use Pearson and Spearman correlation coefficient as well
```

*#Answer: suppose, ##The null hypothesis is, $H_0: \rho = 0$ (i.e X and Y are not correlated) ##The
Alternative hypothesis is, $H_1: \rho > 0$ (i.e X and Y are positively correlated)*

```
x= c(39,29,60,40,32)
```

```
y= c(11,5,20,8,6)
```

```
sttrue1= cor(x,y,method= "pearson") #taking statistic as the Pearson correlation coefficient  
sttrue1
```

```
## [1] 0.9777792
```

```
sttrue2= cor(x,y,method= "spearman") #taking statistic as the spearman correlation coefficient  
sttrue2
```

```
## [1] 0.9
```

```
n=length(y)
```

```
nr=fact(n) #number of rearrangements to be examined
```

```
st1=numeric(nr)
```

```
st2=numeric(nr)
```

```
cnt1=0 #for initiating counting
```

```
cnt2=0
```

```
d=permn(y) #Permuting randomly the Y column leaving the X fixed
```

```
for (i in 1:nr)
```

```
{ st1[i]<-cor(d[[i]],x,method= "pearson")
```

```
if (st1[i] >=sttrue1)cnt1=cnt1+1 #comparing the true statistic and the evaluated statistic
```

```
st2[i]<-cor(d[[i]],x,method= "spearman")
```

```

    if (st2[i] >=sttrue2)cnt2=cnt2+1 #comparing the true statistic and the evaluated statistic
  }
p_value1<-cnt1/nr #pvalue(using pearson)
p_value1

```

```
## [1] 0.025
```

```

p_value2<-cnt2/nr #pvalue(using spearman)
p_value2

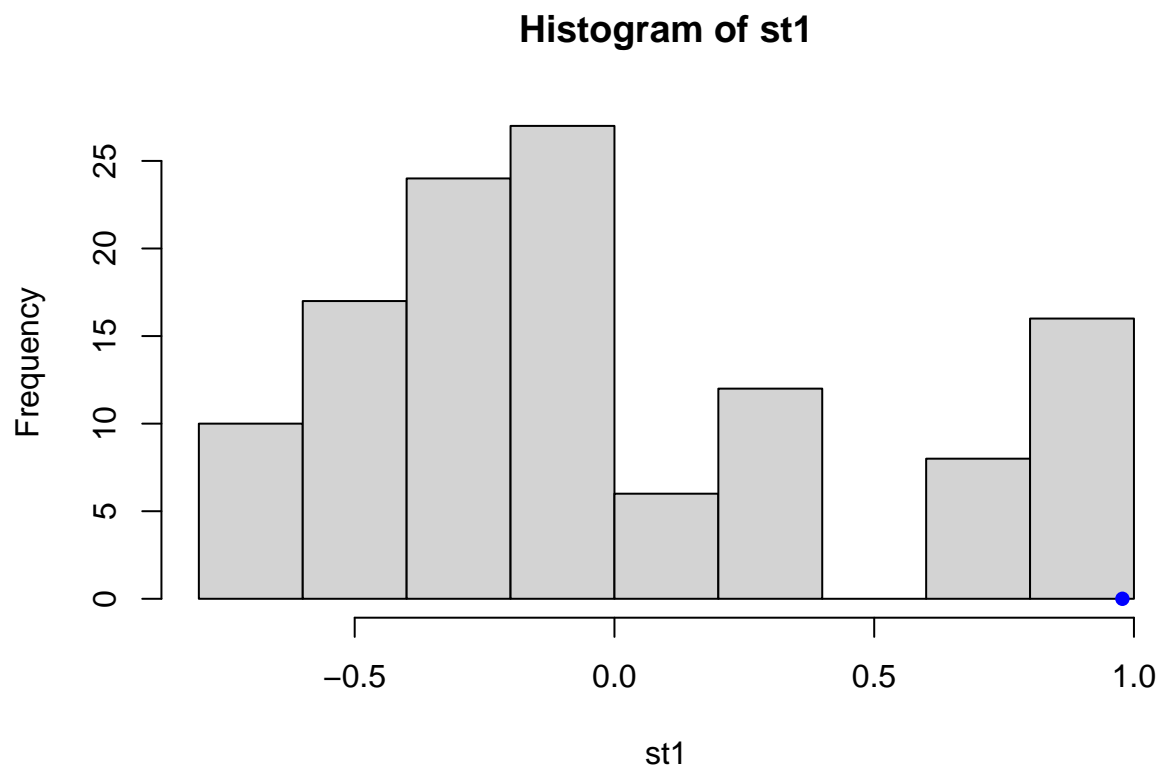
```

```
## [1] 0.04166667
```

```

hist(st1)
points(sttrue1,0,pch = 16,
      col = "blue")

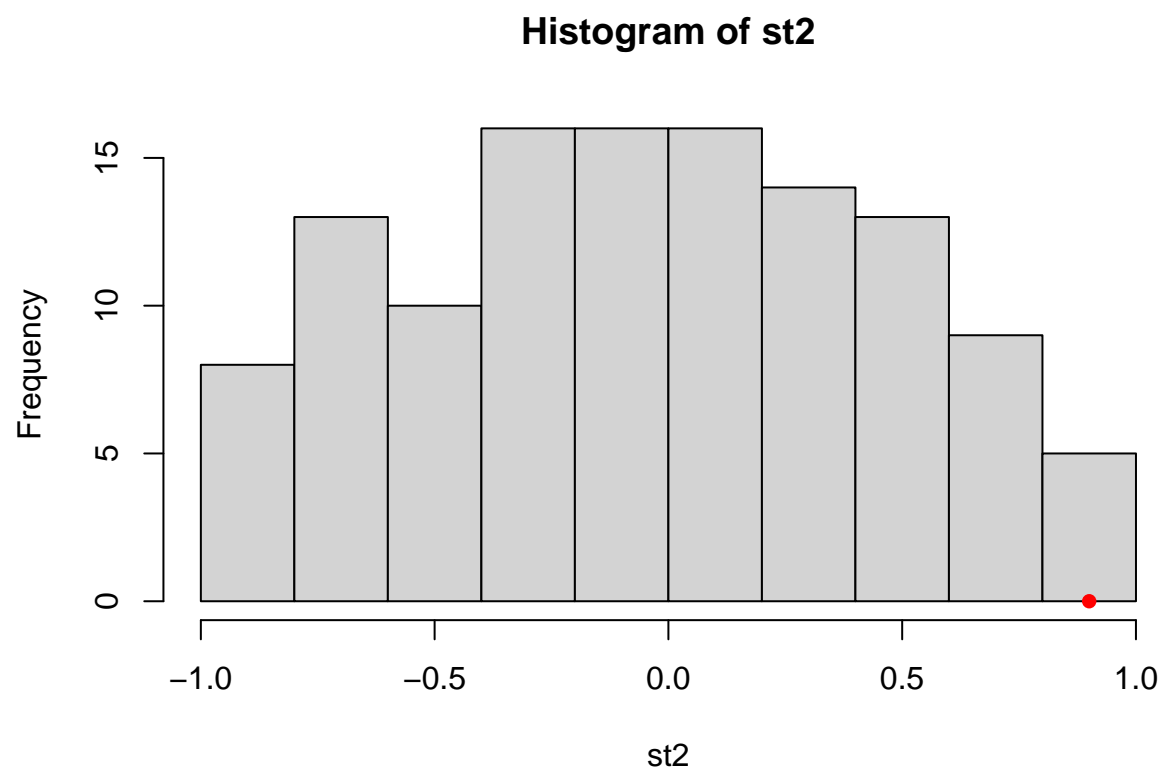
```



```

hist(st2)
points(sttrue2,0,pch = 16,
      col = "red")

```



#in both cases, since the p-value is lower than 0.05, we have lower evidence to support the null hypothesis
#therefore, the null hypothesis is rejected, so we can conclude that
#Chest circumference and volume of air are positively correlated

Problem-2

```
#install.packages("R.utils")  
library(R.utils)
```

```
## Warning: package 'R.utils' was built under R version 4.2.2
```

```
## Loading required package: R.oo
```

```
## Loading required package: R.methodsS3
```

```
## R.methodsS3 v1.8.2 (2022-06-13 22:00:14 UTC) successfully loaded. See ?R.methodsS3 for help.
```

```
## R.oo v1.25.0 (2022-06-12 02:20:02 UTC) successfully loaded. See ?R.oo for help.
```

```
##
```

```
## Attaching package: 'R.oo'
```

```
## The following object is masked from 'package:R.methodsS3':
```

```
##
```

```
##      throw
```

```
## The following objects are masked from 'package:methods':
```

```
##
```

```
##      getClasses, getMethods
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      attach, detach, load, save
```

```
## R.utils v2.12.2 (2022-11-11 22:00:03 UTC) successfully loaded. See ?R.utils for help.
```

```
##
```

```
## Attaching package: 'R.utils'
```

```
## The following object is masked from 'package:utils':
```

```
##
```

```
##      timestamp
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      cat, commandArgs, getOption, isOpen, nullfile, parse, warnings
```

```

Standard= c(2.5, 3.4, 2.9, 4.1, 5.3, 3.4, 1.9, 3.3, 1.8 )
Additive=c(3.5, 6.3, 4.2, 4.3, 3.8, 5.7, 4.4, 0, 0)
diff=Additive - Standard
sttrue= mean(diff)
n=length(diff)
nf=2^n-1
st=numeric(nf)
stat = numeric(n)

for (i in 0:nf)
  { rearr=intToBin(i)
  nn=nchar(rearr)
  u=strrep("0", n-nn)
  rearr=paste(u,rearr,sep="")
  for (j in 1:n){stat[j]=ifelse(substr(rearr,j,j)==0, diff[j], -diff[j])}
  st[i]=mean(stat)
}
pvalue <-length(st[st<=sttrue])/(2^n) # p-value
pvalue

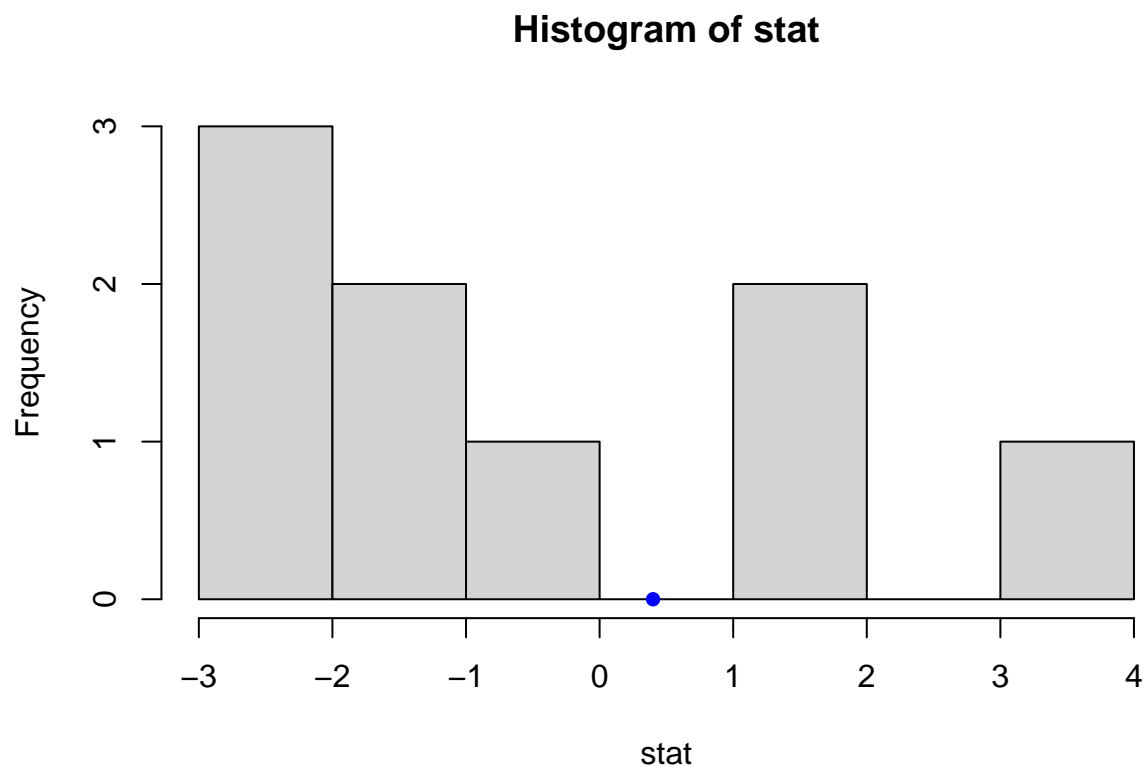
```

```
## [1] 0.7167969
```

```

hist(stat)
points(sttrue,0,pch = 16,
       col = "blue")

```



```
#since the p-value is larger,we have enough evidence to support the  
#null hypothesis therefore,the null hypothesis is not rejected,so we can conclude  
# that the additive is not effective.
```

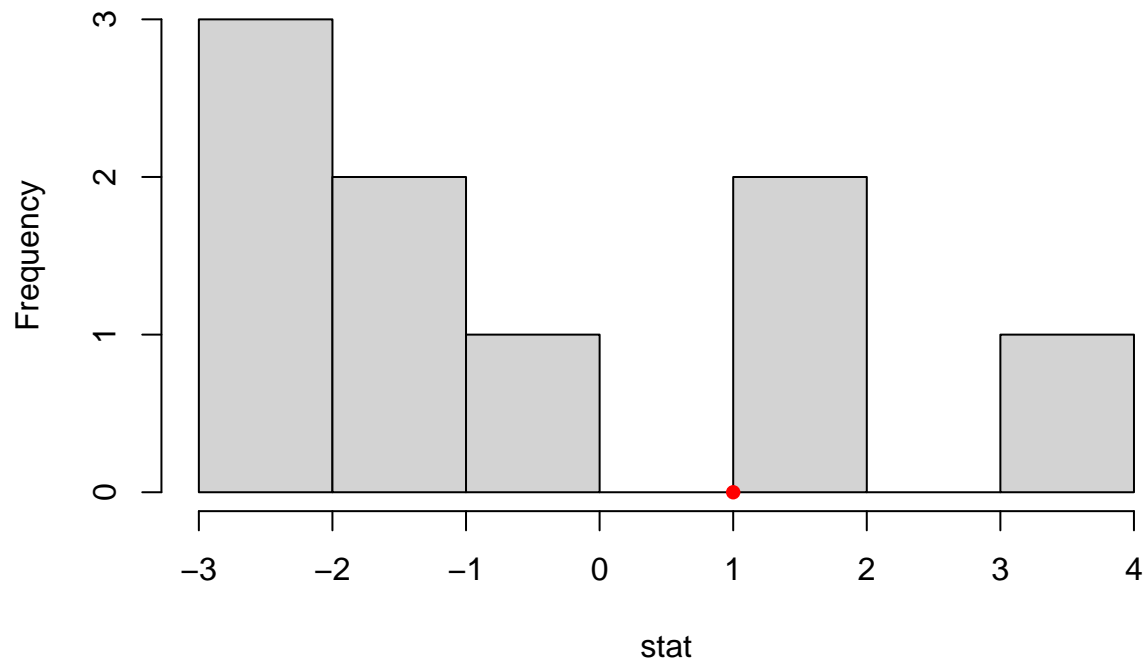
```
#(b) Explore several statistics in order to perform the permutation test.  
#What is the final conclusion?  
#we are going to use several statistic like pearson correlation coefficient and  
#spearman correlation coefficient
```

```
Standard= c(2.5, 3.4, 2.9, 4.1, 5.3, 3.4, 1.9, 3.3, 1.8 )  
Additive=c(3.5, 6.3, 4.2, 4.3, 3.8, 5.7, 4.4, 0, 0)  
diff=Additive - Standard  
n=length(diff)  
sttrue1= median(diff)  
n=length(diff)  
nf=2^n-1  
st1=numeric(nf)  
stat = numeric(n)  
  
for (i in 0:nf)  
{ rearr=intToBin(i)  
nn=nchar(rearr)  
u=strrep("0", n-nn)  
rearr=paste(u,rearr,sep="")  
for (j in 1:n){stat[j]=ifelse(substr(rearr,j,j)==0, diff[j], -diff[j])}  
st1[i]=median(stat)  
}  
pvalue1 <-length(st1[st1<=sttrue1])/(2^n) # p-value  
pvalue1
```

```
## [1] 0.7714844
```

```
hist(stat)  
points(sttrue1,0,pch = 16,  
       col = "red")
```

Histogram of stat



*#in this case,using different statistic the p-value is also larger,we
#have enough evidence to support the null hypothesis
#therefore,the null hypothesis is not rejected,so we can conclude that
#there does not exist a significant difference before and after the additive
#that is the additive is not effective*

prpbem-3

2023-01-15

#Problem-3: Assume that during a three-hour period spent outside, a person recorded the temperature, the amount of time they mowed the grass, and their water consumption. The experiment was conducted on 7 randomly selected days during the summer.

#(a) Fit the Water Consumption as a linear function of the Temperature (T) and the Time mowing the grass (TMG) by means of a multiple regression model. Hint: Use function #lm in R.

```
library(combinat)
```

```
##
```

```
## Attaching package: 'combinat'
```

```
## The following object is masked from 'package:utils':
```

```
##
```

```
##      combn
```

```
df=data.frame(T=c(75,83,85,85,92,97,99),W=c(16,20,25,27,32,48,48),
              TMG=c(1.85,1.25,1.5,1.75,1.15,1.75,1.6))
df
```

```
##      T  W  TMG
## 1 75 16 1.85
## 2 83 20 1.25
## 3 85 25 1.50
## 4 85 27 1.75
## 5 92 32 1.15
## 6 97 48 1.75
## 7 99 48 1.60
```

```
model<-lm(W~ T+TMG,data=df)
an<-anova(model)
an
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: W
```

```
##           Df Sum Sq Mean Sq F value    Pr(>F)
## T           1  905.53   905.53  584.316 1.737e-05 ***
## TMG          1   65.13    65.13   42.029 0.002918 **
## Residuals    4    6.20     1.55
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



```

FTtrue<-an$"F value"[1]      #True F value of T before permutation test
FTMGtrue<-an$"F value"[2]    #gives the significance between TMG and W
FTtrue

```

```
## [1] 584.3158
```

```
FTMGtrue
```

```
## [1] 42.02869
```

```

#F-value of the variable TR without permutation test is 584.3158
#F-value for the variable TMG without permutation test is 42.02869

```

```

#(b) Test the significance of the variables T and TMG using an
#exact permutation test. Compare the results with those obtained
#with the multiple regression model.

```

```

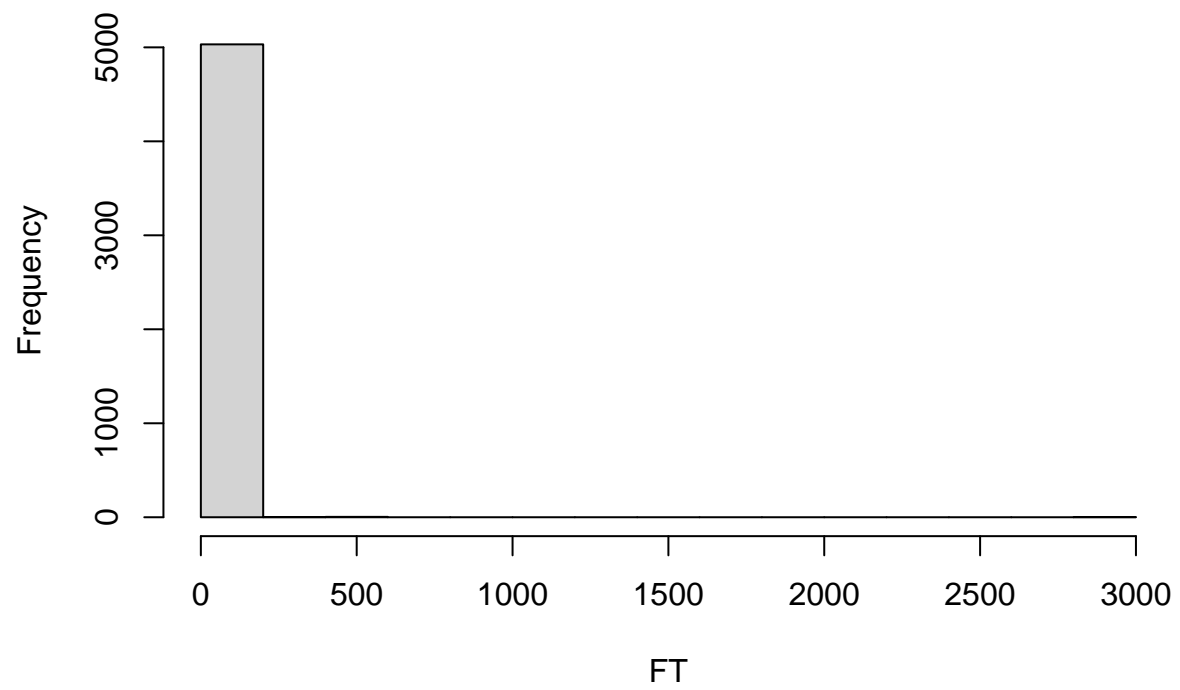
#exact test
T=c(75,83,85,85,92,97,99)
W=c(16,20,25,27,32,48,48)
TMG=c(1.85,1.25,1.5,1.75,1.15,1.75,1.6)
df=data.frame(T=c(75,83,85,85,92,97,99),W=c(16,20,25,27,32,48,48),
              TMG=c(1.85,1.25,1.5,1.75,1.15,1.75,1.6))
n = length(T)
nr=fact(n) #number of rearrangements to be examined
FT=numeric(nr)
FTMG=numeric(nr)

for (i in 1:nr){
  newW<- permn(W)
  newW
  model1<- lm(newW[[i]]~ T+TMG, data=df)
  an1<-anova(model1)
  FT[i]<-an1$"F value"[1]
  FTMG[i]<-an1$"F value"[2]
}

hist(FT, main = " Histogram of the FT")

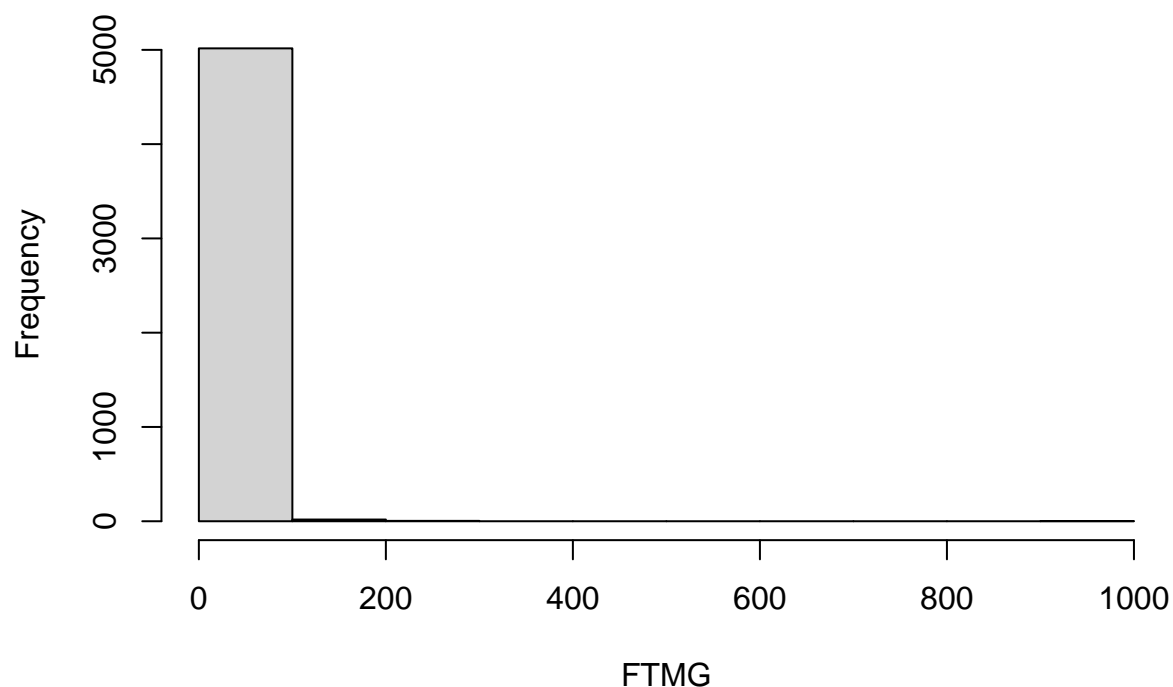
```

Histogram of the FT



```
hist(FTMG, main = " Histogram of the FTMG")
```

Histogram of the FTMG



```
an1
```

```
## Analysis of Variance Table
##
## Response: newW[[i]]
##           Df Sum Sq Mean Sq F value    Pr(>F)
## T           1 815.03   815.03  115.124 0.0004276 ***
## TMG          1 133.51   133.51   18.858 0.0122283 *
## Residuals    4  28.32     7.08
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# F value of T after permutation test
FT1<-an1$"F value"[1]    #gives the significance between T and W
# F value of TMG after permutation test
FTMG1<-an1$"F value"[2]  #gives the significance between TMG and W
FT1
```

```
## [1] 115.124
```

```
FTMG1
```

```
## [1] 42.02869
```

*#the f-values are different after doing exact test than the obtained values from
#the multiple linear regression model*