

HW Mar8

Jinfei(Fiona) Xue

3/8/2019

1

```
df1<-read.table("maybe_uniform.txt",header = F)

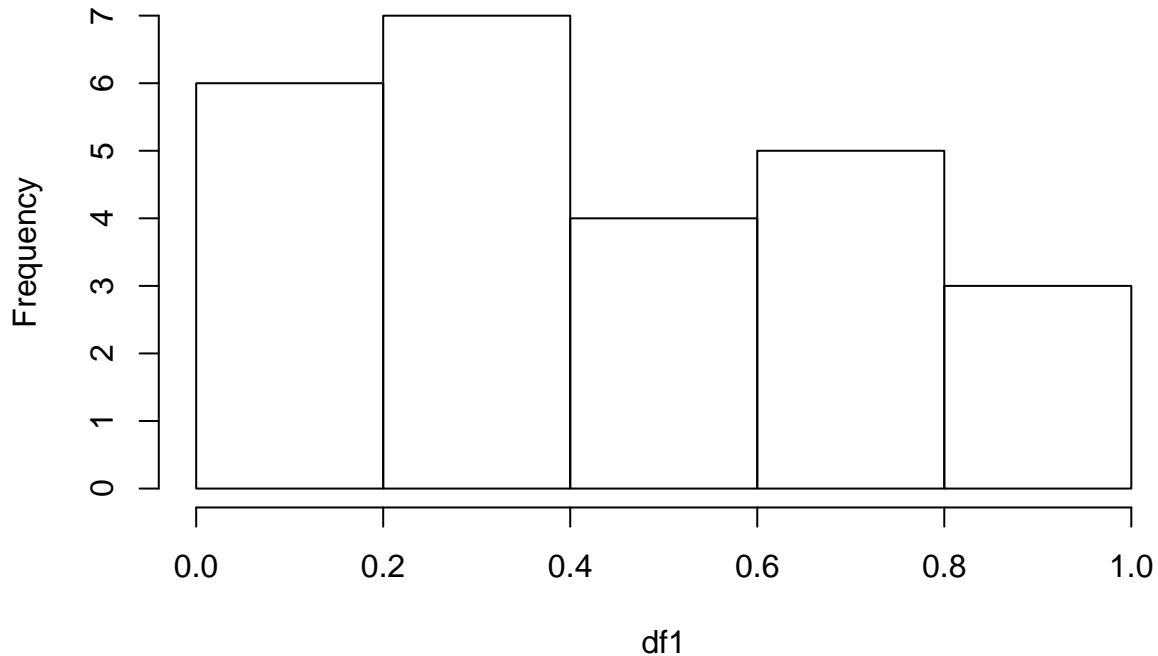
## Warning in read.table("maybe_uniform.txt", header = F): incomplete final
## line found by readTableHeader on 'maybe_uniform.txt'

df1<-as.matrix(df1)
df1<-as.vector(df1)
library(stats)
ks.test(df1,"punif")

##
## One-sample Kolmogorov-Smirnov test
##
## data: df1
## D = 0.18, p-value = 0.3501
## alternative hypothesis: two-sided

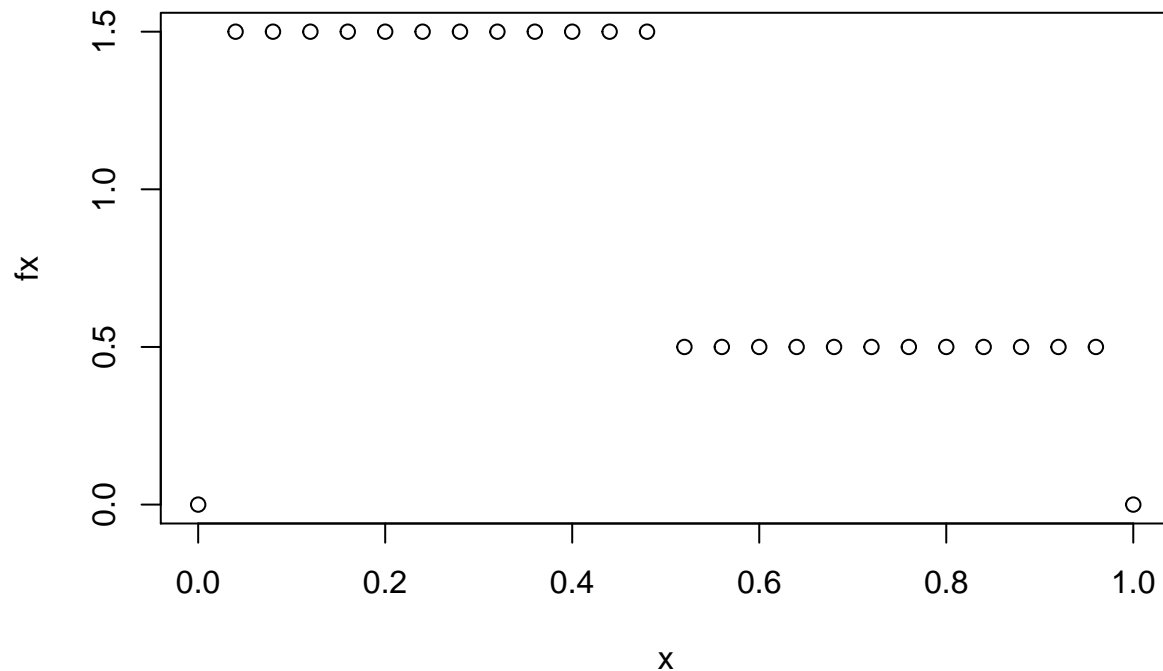
hist(df1)
```

Histogram of df1



We cannot reject the hypothesis that it conforms a uniform distribution.

```
x<-seq(0, 1, by=1/25)
fx <- ifelse(x > 0 & x <=0.5, 3/2,
  ifelse(x > 0.5 & x < 1, 0.5, 0))
plot(x,fx)
```



```
ks.test(fx,df1)
```

```
## Warning in ks.test(fx, df1): cannot compute exact p-value with ties
```

```
##
## Two-sample Kolmogorov-Smirnov test
##
## data: fx and df1
## D = 0.56308, p-value = 0.0006182
## alternative hypothesis: two-sided
```

The D statistic for this model is 0.56, which is greater than 0.18 as of uniform distribution. Therefore the data fit uniform distribution better.

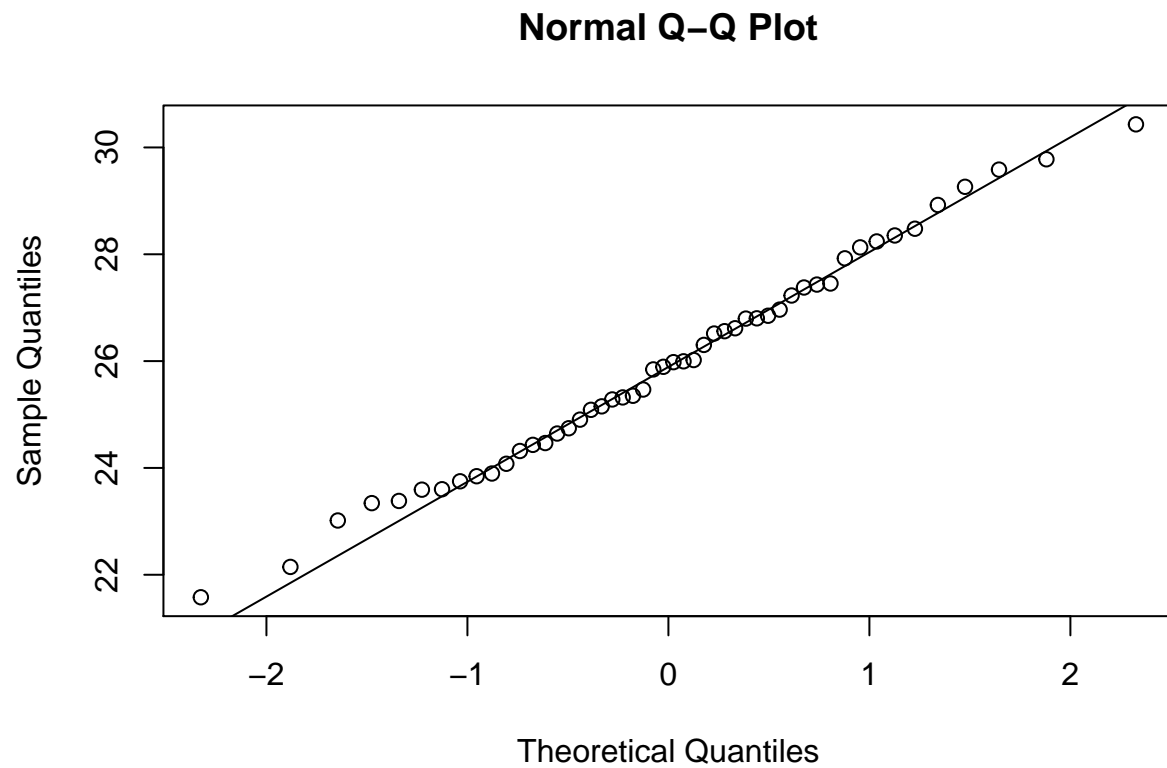
2

```
df2<-read.table("maybe_normal.txt")
df2<-as.vector(as.matrix(df2))
ks.test(df2,"pnorm",26,2)
```

```
##
## One-sample Kolmogorov-Smirnov test
##
## data: df2
## D = 0.06722, p-value = 0.9663
## alternative hypothesis: two-sided
```

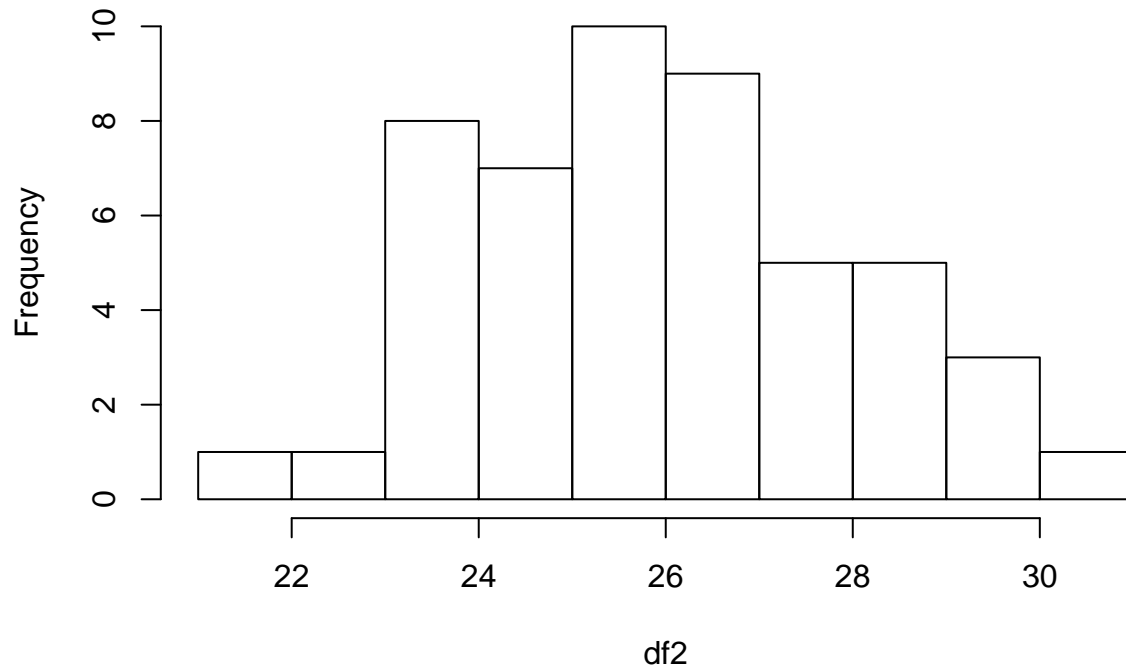
We may have evidence of rejecting that this sample comes from a $N(26,4)$ distribution. We should check qq plot and histogram for further convince.

```
qqnorm(df2)
qqline(df2)
```



```
hist(df2)
```

Histogram of df2



3

```
df3<-as.numeric(as.vector(as.matrix(read.table("maybe_same_1.txt"))))

## Warning in read.table("maybe_same_1.txt"): incomplete final line found by
## readTableHeader on 'maybe_same_1.txt'

## Warning: NAs introduced by coercion
df4<-as.numeric(as.vector(as.matrix(read.table("maybe_same_2.txt"))))

## Warning in read.table("maybe_same_2.txt"): incomplete final line found by
## readTableHeader on 'maybe_same_2.txt'

## Warning in read.table("maybe_same_2.txt"): NAs introduced by coercion
ks.test(df3,df4)

##
## Two-sample Kolmogorov-Smirnov test
##
## data: df3 and df4
## D = 0.2375, p-value = 0.6555
## alternative hypothesis: two-sided
ks.test(df3+2,df4)

## Warning in ks.test(df3 + 2, df4): cannot compute exact p-value with ties
##
## Two-sample Kolmogorov-Smirnov test
```

```
##
## data:  df3 + 2 and df4
## D = 0.8125, p-value = 7.274e-05
## alternative hypothesis: two-sided
```

4

```
df5<-readRDS("norm_sample.Rdata")
ks.test(df5,"pnorm",0,1)

##
## One-sample Kolmogorov-Smirnov test
##
## data:  df5
## D = 0.17724, p-value = 0.3683
## alternative hypothesis: two-sided
Norm<-ecdf(df5)
```

5

```
df6<-read.table("fijiquakes.dat",header = T)
x<-df6$mag
Fn <- ecdf(x)
library(Hmisc)
tot<-sum( (x<=4.9) & (x>4.3))
binconf(tot,length(x),method="wilson",0.05)

## PointEst      Lower      Upper
##      0.526 0.4950118 0.5567892

df7<-read.table("faithful.dat",skip = 25)
waiting<-df7$waiting
avg<-mean(waiting)
var<-var(waiting)
n<-length(waiting)
L<-round(avg-qnorm(0.95)*sqrt(var/n),3)
U<-round(avg+qnorm(0.95)*sqrt(var/n),3)
print(paste("the 90% CI for mean waiting time is:[",L,"",U,""])))

## [1] "the 90% CI for mean waiting time is:[ 69.541 , 72.253 ]"
median(waiting)

## [1] 76
```