

# MA677HW3

Qixuan Zhang

03/08/2019

1.

```
Df1=read.table("maybe_uniform.txt")
```

```
## Warning in read.table("maybe_uniform.txt"): incomplete final line found by
```

```
## readTableHeader on 'maybe_uniform.txt'
```

```
test1=c(Df1$V1,Df1$V2,Df1$V3,Df1$V4,Df1$V5)
```

```
ks.test(test1,"punif")
```

```
##
```

```
## One-sample Kolmogorov-Smirnov test
```

```
##
```

```
## data: test1
```

```
## D = 0.18, p-value = 0.3501
```

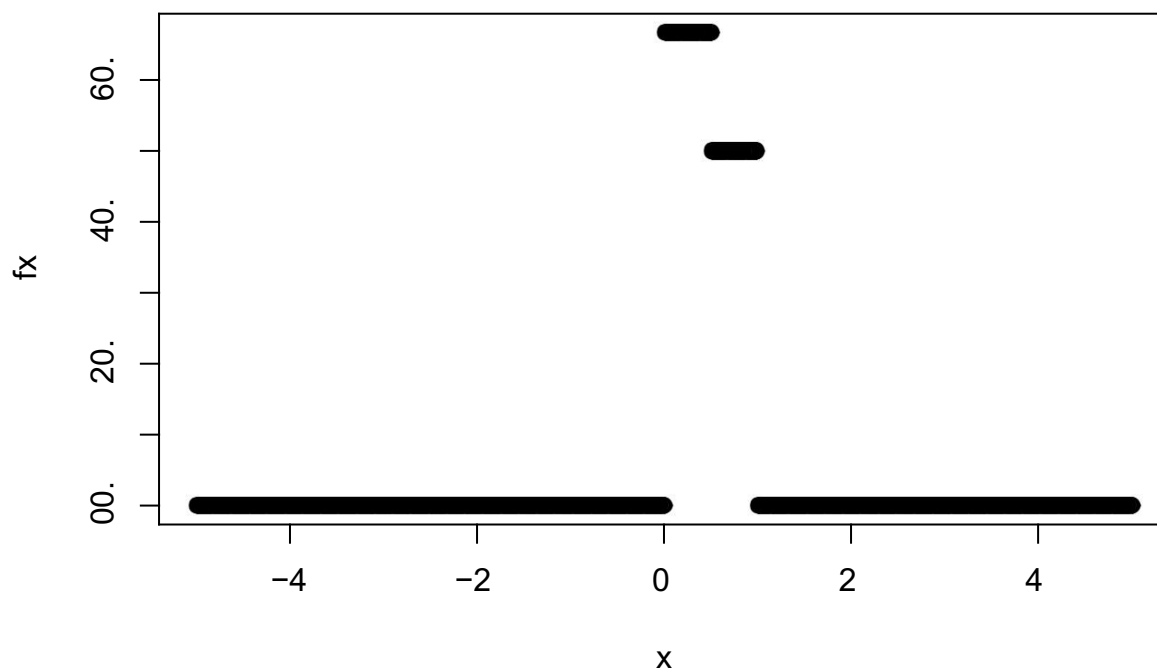
```
## alternative hypothesis: two-sided
```

```
x<-seq(-5, 5, by=0.01)
```

```
fx <- ifelse(x > 0 & x <=0.5, 2/3,
```

```
ifelse(x > 0.5 & x < 1, 0.5, 0))
```

```
plot(x,fx)
```



```
ks.test(fx,"punif")
```

```
## Warning in ks.test(fx, "punif"): ties should not be present for the
```

```
## Kolmogorov-Smirnov test
```

```
##
```

```
## One-sample Kolmogorov-Smirnov test
```

```
##
## data:  fx
## D = 0.9011, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

According to result of ks-test, these data is distributed as a Uniform distribution on  $[0,1]$ . Since in ks-test, the D value of fx is less than that of data, the model below is better than the Uniform

2.

```
Df2=read.table("maybe_normal.txt")
```

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

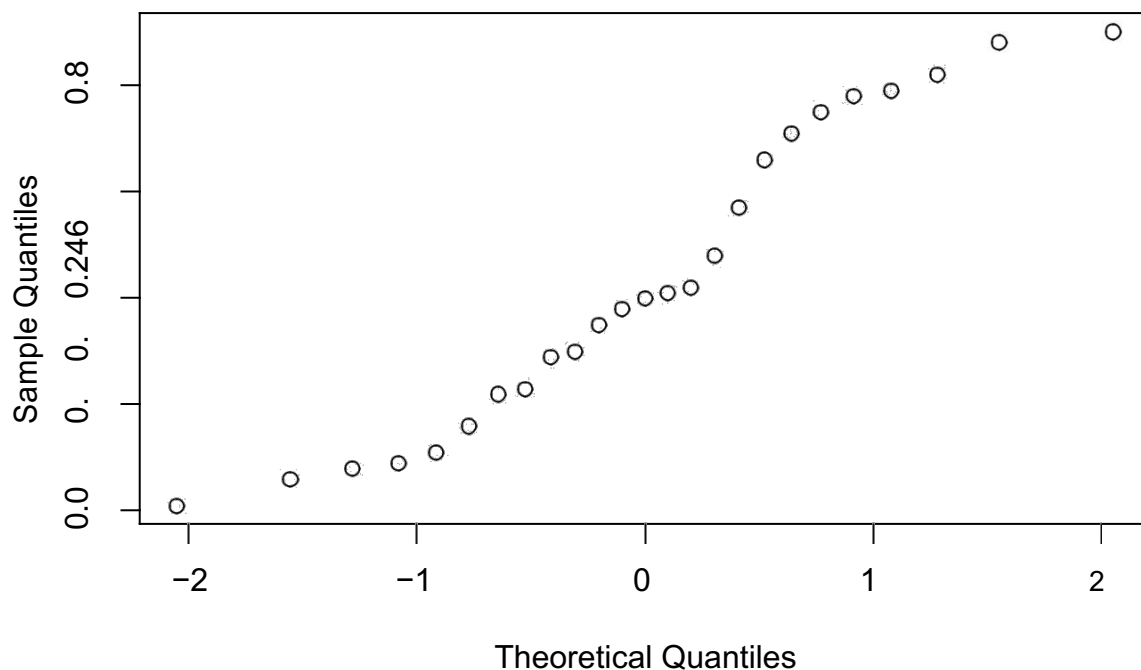
```
d=c(Df2$V1,Df2$V2,Df2$V3,Df2$V4,Df2$V5)
test2=(c(Df2$V1,Df2$V2,Df2$V3,Df2$V4,Df2$V5) -
rep(26,50))/2 ks.test(test2,"pnorm")
```

```
## Warning in ks.test(test2, "pnorm"): ties should not be present for the
## Kolmogorov-Smirnov test
```

```
##
## One-sample Kolmogorov-Smirnov test
##
## data:  test2
## D = 1, p-value < 2.2e-16
## alternative hypothesis: two-sided
```

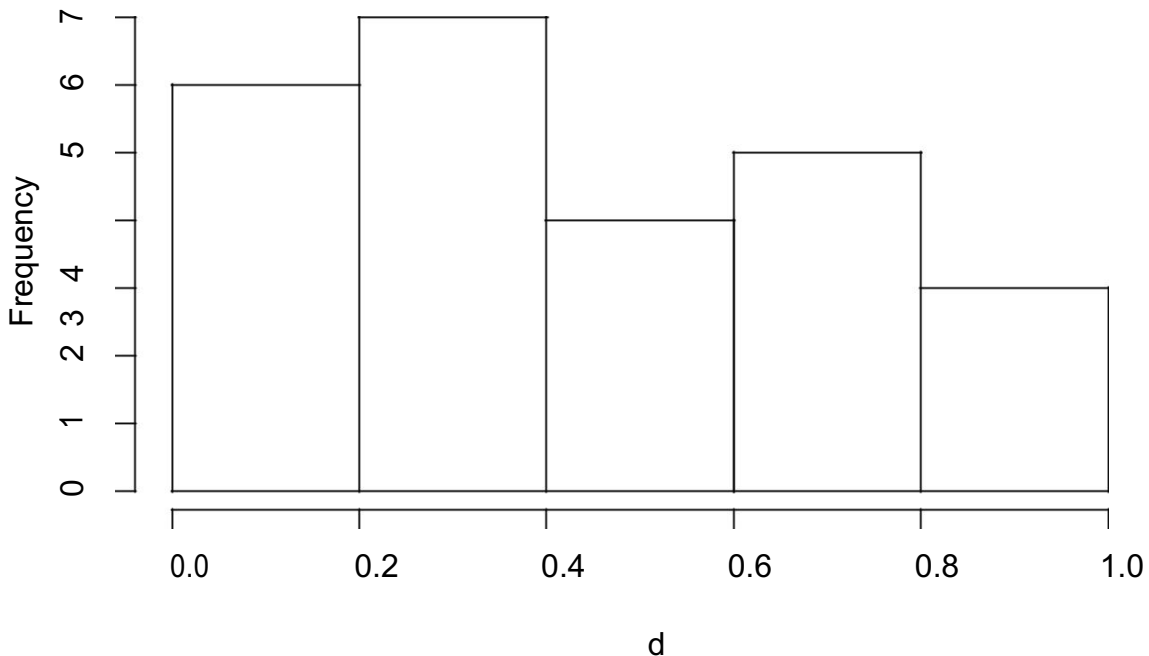
```
qqnorm(d)
```

**Normal Q-Q Plot**



```
hist(d)
```

## Histogram of d



ing to Ks-test and qqnorm, the data generally follow normal disatribution.

3.

```
Df31=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'
```

```
x=c(Df31$V1,Df31$V2,Df31$V3,Df31$V4,Df31$V5)
```

```
Df32=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'
```

```
y=c(Df32$V1,Df32$V2,Df32$V3,Df32$V4,Df32$V5)
```

```
ks.test(x,y)
```

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

```
## Two-sample Kolmogorov-Smirnov test
##
```

```
## data: x and y
```

```
## D = 0.15, p-value = 0.9639
```

```
## alternative hypothesis: two-sided
```

```
ks.test(x+2,y)
```

```
## Warning in ks.test(x + 2, y): cannot compute exact p-value
with ties ##
```

```
## Two-sample Kolmogorov-Smirnov test
##
```

```
## data: x + 2 and y
```

```
## D = 0.85, p-value = 2.129e-07
## alternative hypothesis: two-sided
```

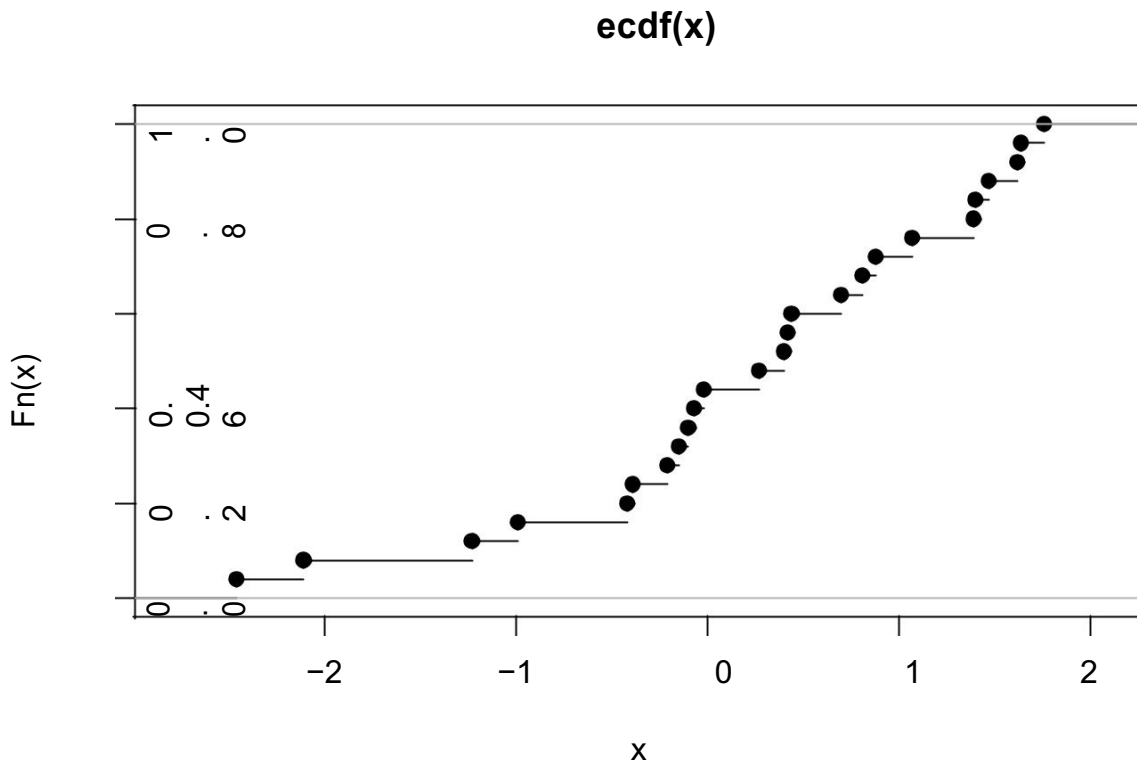
Since p-value of `ks.test(X,Y)` is much larger than 0.05. Then we fail to reject the null hypotheses, thus X and Y are from the same distribution. Since p-value of `ks.test(X+2,Y)` is much smaller than 0.05. Then we reject the null hypotheses, therefore X+2 and Y are not from the same distribution

4.

```
Df4=readRDS("norm_sample.Rdata")
Df4_ecdf=ecdf(Df4)
summary(Df4_ecdf)
```

```
## Empirical CDF: 25 unique values with summary
##   Min. 1st Qu. Median   Mean 3rd Qu.    Max.
## -2.4600 -0.2100  0.4000  0.2448  1.0700  1.7600
```

```
plot.ecdf(Df4)
```



```
us=rnorm(25)
ks.test(Df4,us)
```

```
##
## Two-sample Kolmogorov-Smirnov test
##
## data: Df4 and us
## D = 0.32, p-value = 0.1558
## alternative hypothesis: two-sided
```

5.