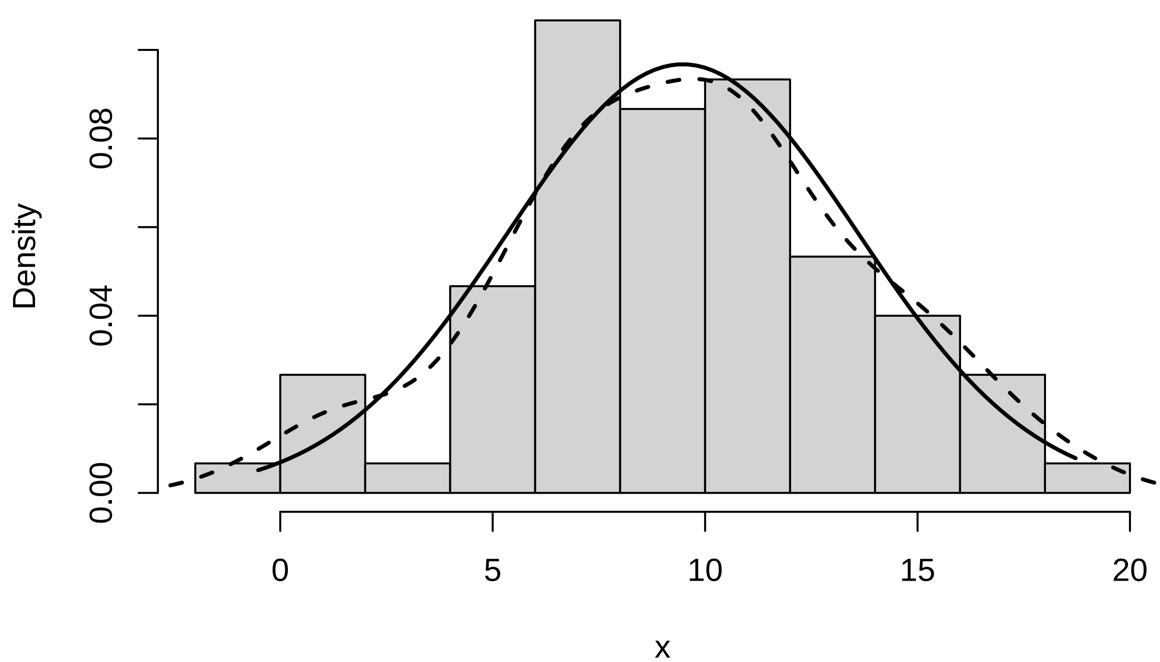
Test of Fit for Two Simulated Populations

jdt

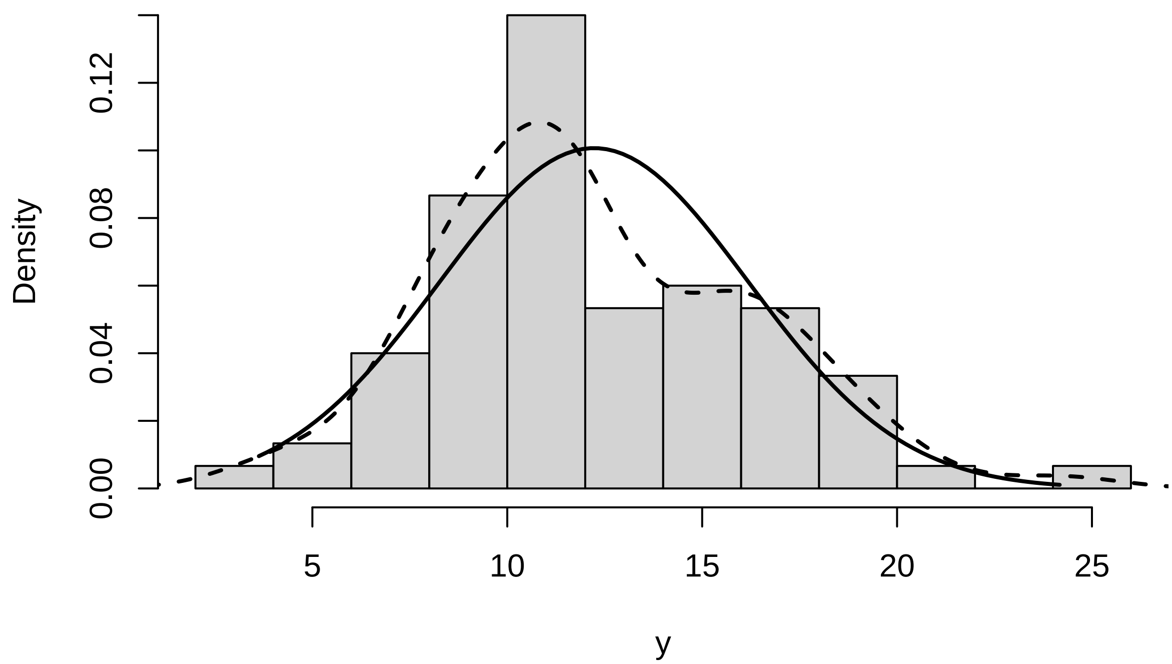
1/28/2023

Analysis with Simulated normal data

```
set.seed(4382)
x = rnorm(75, mean=10, sd = 4)    #x ~ N(10,4)
y = rnorm(75, mean=12, sd = 4)    #y ~ N(12,4)
cert = data.frame(x,y)
#hist(x)
#hist(y)

with(cert, hist(x, main="", freq=FALSE))
with(cert, lines(density(x), main="X", lty=2, lwd=2))
xvals = with(cert, seq(from=min(x), to=max(x), length=100))
with(cert, lines(xvals, dnorm(xvals, mean(x), sd(x)), lwd=2))
```



```r
with(cert, hist(y, main="", freq=FALSE))
with(cert, lines(density(y), main="", lty=2, lwd=2))
xvals = with(cert, seq(from=min(y), to=max(y), length=100))
with(cert, lines(xvals, dnorm(xvals, mean(y), sd(y)), lwd=2))
```



Descriptive statistics

```
##         min       Q1   median       Q3      max     mean       sd  n missing
##  -0.5192659 6.986026 9.614023 12.06477 18.72084 9.474439 4.123197 75       0

## [1] 9.501816

##      2.5%     12.5%     22.5%     32.5%     42.5%     52.5%     62.5%     72.5%
##  1.324368  5.019723  6.458432  7.629620  8.766079  9.682428 10.746588 11.550225
##     82.5%     92.5%
## 13.339813 15.489256

##
##  One Sample t-test
##
## data:  x
## t = -5.3046, df = 74, p-value = 1.131e-06
## alternative hypothesis: true mean is not equal to 12
## 90 percent confidence interval:
##   8.681387 10.267492
```

```
## sample estimates:
## mean of x
##  9.474439
```

```
##       min       Q1   median       Q3      max     mean       sd  n missing
##  3.627661 9.407387 11.58455 15.48663 24.17142 12.22575 3.962322 75       0
```

```
## [1] 12.16128
```

```
##      2.5%     12.5%     22.5%     32.5%     42.5%     52.5%     62.5%     72.5%
##  5.118783  8.113084  9.120299 10.451277 11.123059 11.636086 12.265536 14.784435
##     82.5%     92.5%
## 16.232745 18.159077
```

```
##
##  One Sample t-test
##
## data:  y
## t = 0.49341, df = 74, p-value = 0.6232
## alternative hypothesis: true mean is not equal to 12
## 90 percent confidence interval:
##  11.46364 12.98786
## sample estimates:
## mean of x
##  12.22575
```

Use X and Y

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  x and y
## W = 1797, p-value = 0.0001361
## alternative hypothesis: true location shift is not equal to 0
```

```
##
##  Welch Two Sample t-test
##
## data:  x and y
## t = -4.1667, df = 147.77, p-value = 5.238e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -4.056183 -1.446440
## sample estimates:
## mean of x mean of y
##  9.474439 12.225751
```

```
##
##  Exact two-sample Kolmogorov-Smirnov test
##
## data:  x and y
## D = 0.28, p-value = 0.005377
## alternative hypothesis: two-sided
```

Univariate Test of Fit for Normality

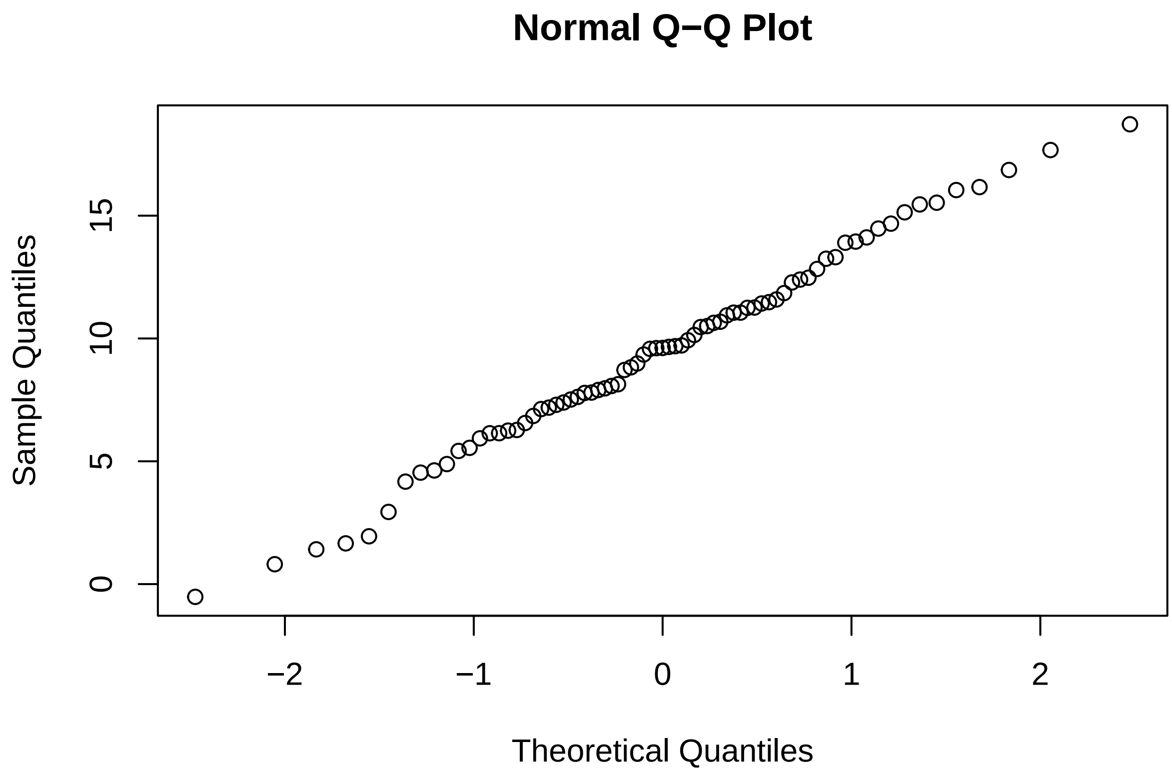
```
##
##  Anderson-Darling normality test
##
## data:  x
## A = 0.14458, p-value = 0.9678
```

```
##
##  Cramer-von Mises normality test
##
## data:  x
## W = 0.0189, p-value = 0.9762
```

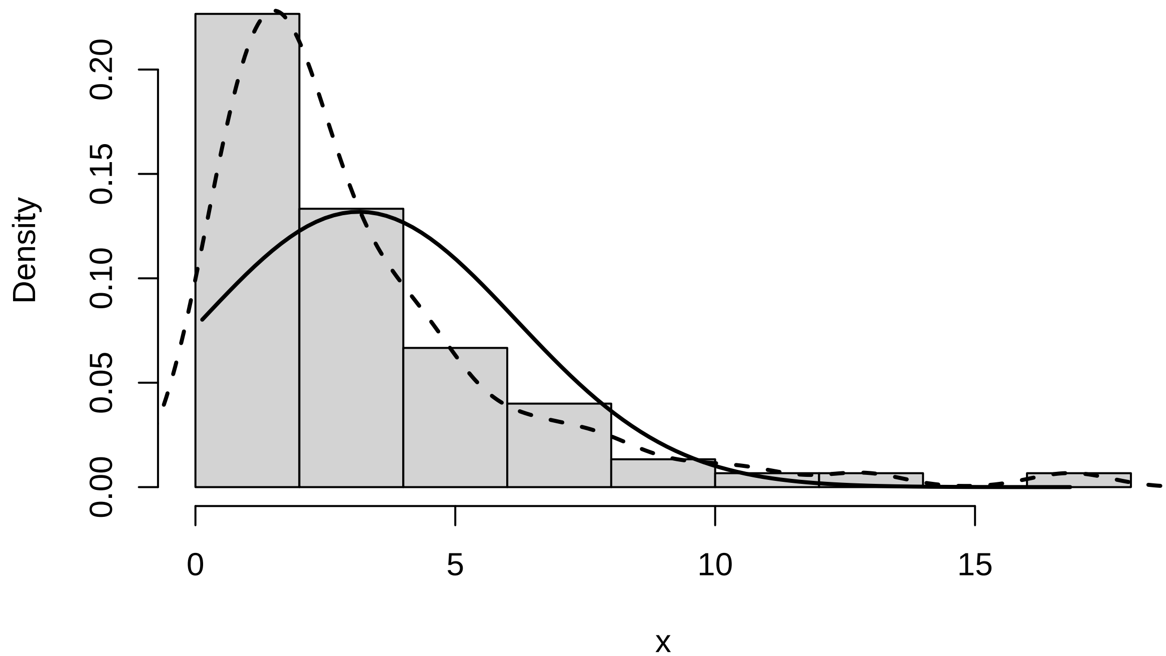
```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  x
## D = 0.043325, p-value = 0.979
```

```
##
##  Pearson chi-square normality test
##
## data:  x
## P = 2.6, p-value = 0.9781
```

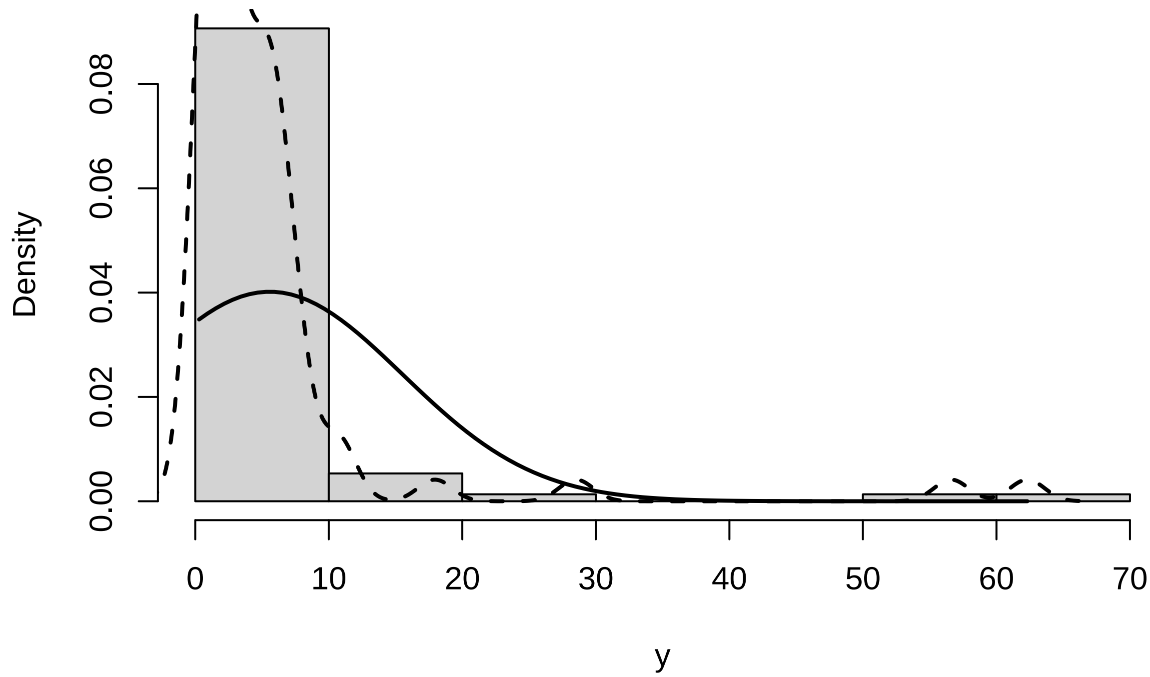
```
##
##  Shapiro-Francia normality test
##
## data:  x
## W = 0.9948, p-value = 0.9718
```



Analysis with Simulated lognormal data

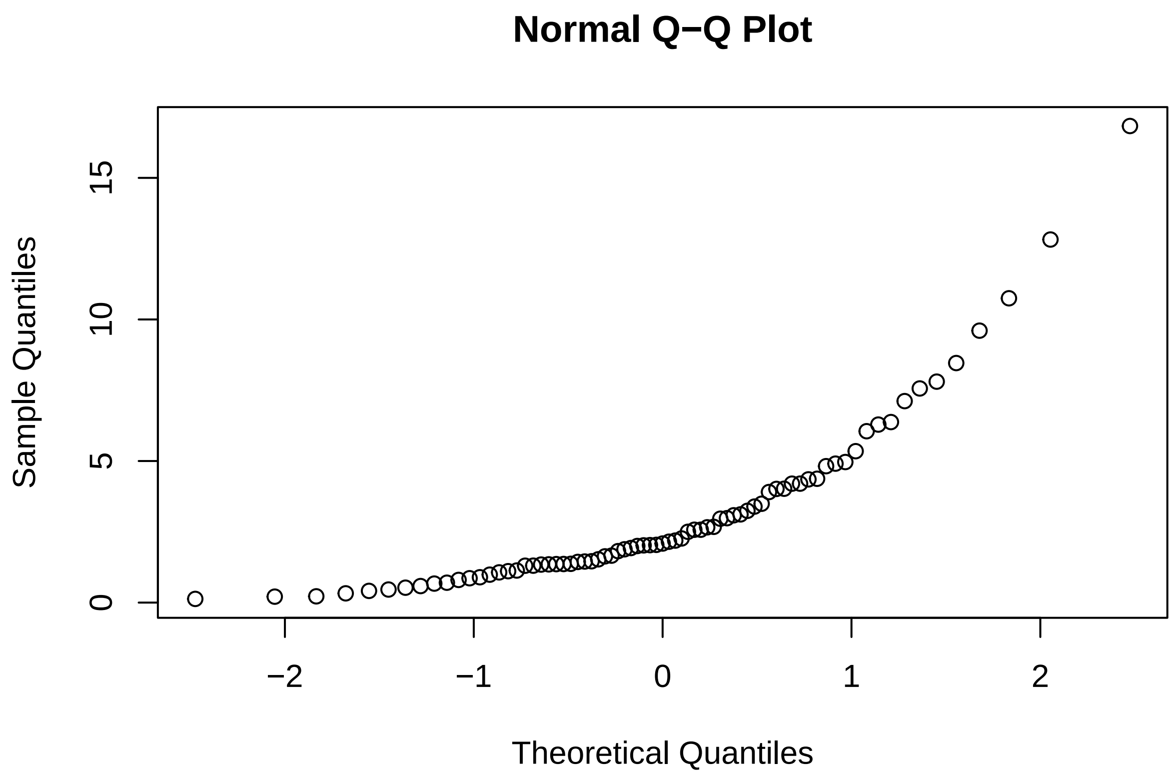


Density
0.00
0.05
0.10
0.15
0.20
0
5
10
15
x




```
##
##  Anderson-Darling normality test
##
## data:  x
## A = 4.4905, p-value = 3.01e-11

##
##  Cramer-von Mises normality test
##
## data:  x
## W = 0.79159, p-value = 1.482e-08

##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  x
## D = 0.17534, p-value = 5.435e-06

##
##  Pearson chi-square normality test
##
## data:  x
## P = 40.68, p-value = 5.719e-06

##
```

```
##  Shapiro-Francia normality test
##
## data:  x
## W = 0.77879, p-value = 3.964e-08
```



Use X and Y

```r
library("coin")
wilcox.test(x,y)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  x and y
## W = 2346, p-value = 0.07985
## alternative hypothesis: true location shift is not equal to 0
```

```r
t.test(x,y)
```

```
##
##  Welch Two Sample t-test
##
## data:  x and y
## t = -2.0249, df = 87.611, p-value = 0.04592
## alternative hypothesis: true difference in means is not equal to 0
```

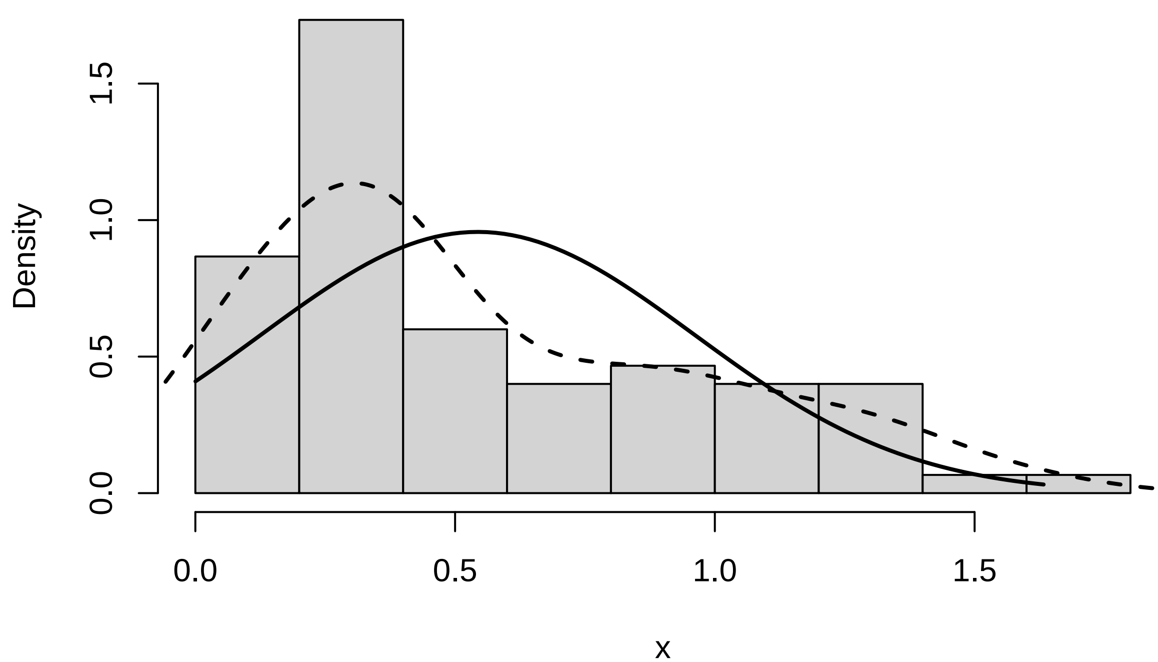
```
## 95 percent confidence interval:
##  -4.8102732 -0.0449811
## sample estimates:
## mean of x mean of y
##  3.148147  5.575774
```

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

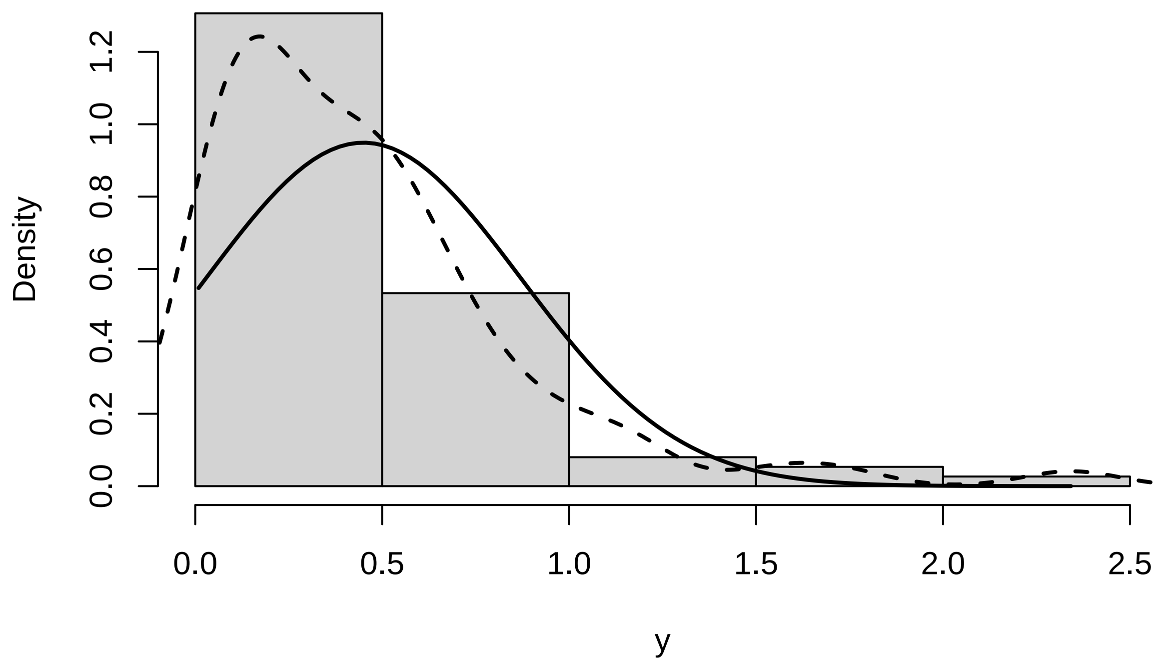
```
##
##  Exact two-sample Kolmogorov-Smirnov test
##
## data:  x and y
## D = 0.21333, p-value = 0.06565
## alternative hypothesis: two-sided
```

Analysis with Simulated exponential data

```r
x = rexp(75, rate = 2)
y = rexp(75, rate = 2.5)

cert = data.frame(x,y)    #redefined data.frame with new x,y

with(cert, hist(x, main="", freq=FALSE))
with(cert, lines(density(x), main="X", lty=2, lwd=2))
xvals = with(cert, seq(from=min(x), to=max(x), length=100))
with(cert, lines(xvals, dnorm(xvals, mean(x), sd(x)), lwd=2))
```



```r
with(cert, hist(y, main="", freq=FALSE))
with(cert, lines(density(y), main="", lty=2, lwd=2))
xvals = with(cert, seq(from=min(y), to=max(y), length=100))
with(cert, lines(xvals, dnorm(xvals, mean(y), sd(y)), lwd=2))
```




```r
library("nortest")
ad.test(x)
```

```
##
##  Anderson-Darling normality test
##
## data:  x
## A = 2.26, p-value = 8.757e-06
```

```r
cvm.test(x)
```

```
##
##  Cramer-von Mises normality test
##
## data:  x
## W = 0.41519, p-value = 1.74e-05
```

```r
lillie.test(x)
```

```
##
##  Lilliefors (Kolmogorov-Smirnov) normality test
##
## data:  x
## D = 0.1733, p-value = 7.586e-06
```

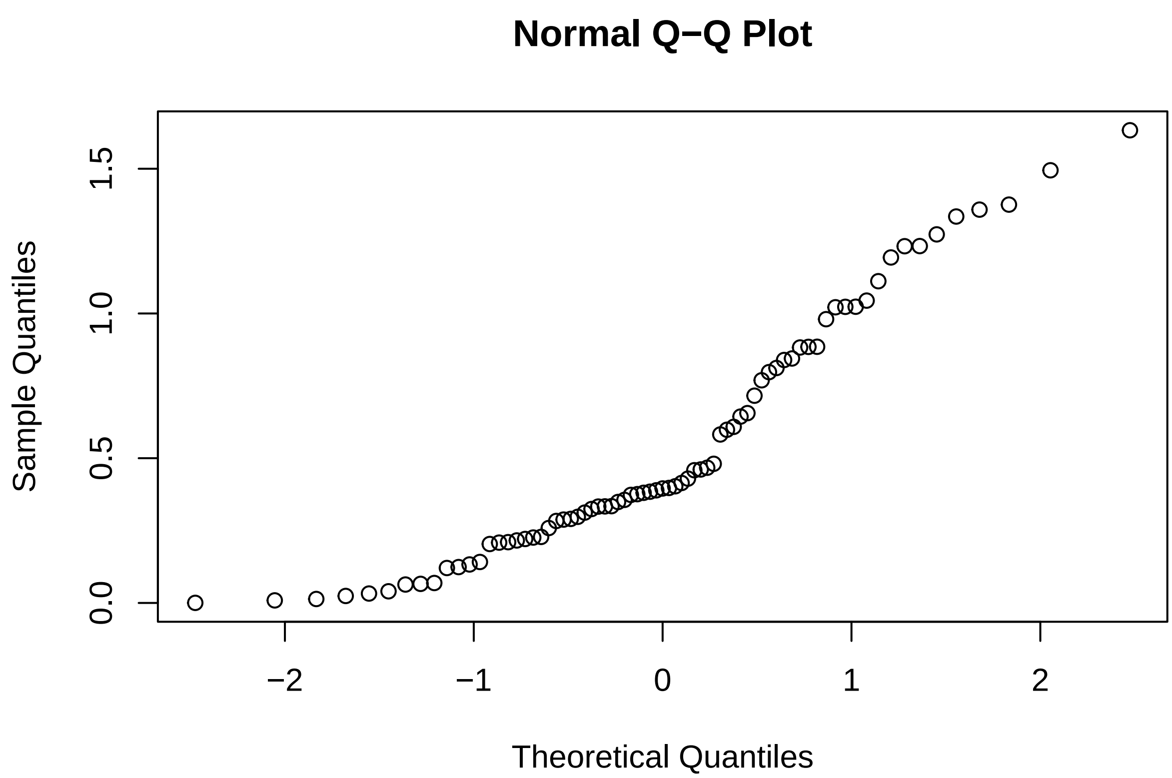
```r
pearson.test(x)
```

```
##
##  Pearson chi-square normality test
##
## data:  x
## P = 25.96, p-value = 0.002074
```

```r
sf.test(x)
```

```
##
##  Shapiro-Francia normality test
##
## data:  x
## W = 0.92087, p-value = 0.0003698
```

```r
qqnorm(x)
```



Use X and Y

```r
library("coin")
wilcox.test(x,y)
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  x and y
## W = 3210, p-value = 0.1356
## alternative hypothesis: true location shift is not equal to 0
```

```r
t.test(x,y)
```

```
##
##  Welch Two Sample t-test
##
## data:  x and y
## t = 1.3714, df = 147.99, p-value = 0.1723
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.04134771  0.22886430
## sample estimates:
## mean of x mean of y
## 0.5435884 0.4498301
```

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

```
##
##  Exact two-sample Kolmogorov-Smirnov test
##
## data:  x and y
## D = 0.16, p-value = 0.2937
## alternative hypothesis: two-sided
```