hw3

```
#Question 1
mu <- read.table("maybe_uniform.txt", header = FALSE)
```
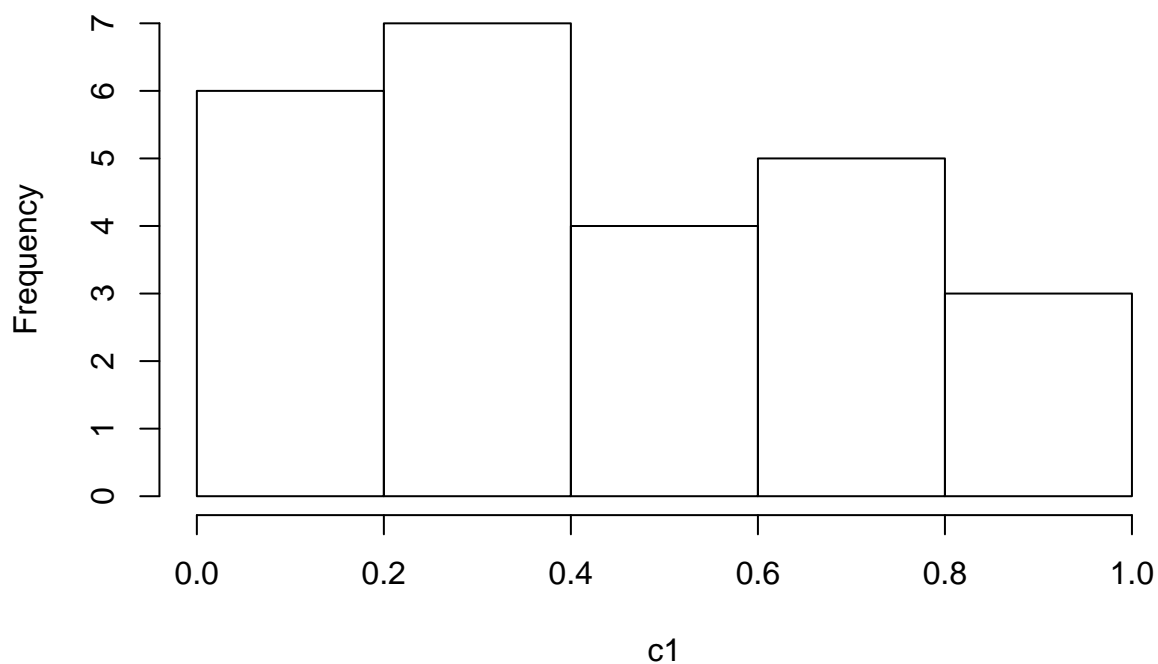
```
## Warning in read.table("maybe_uniform.txt", header = FALSE): incomplete
## final line found by readTableHeader on 'maybe_uniform.txt'
```

```
c1 <- c(mu$V1,mu$V2,mu$V3,mu$V4,mu$V5)
hist(c1)
```
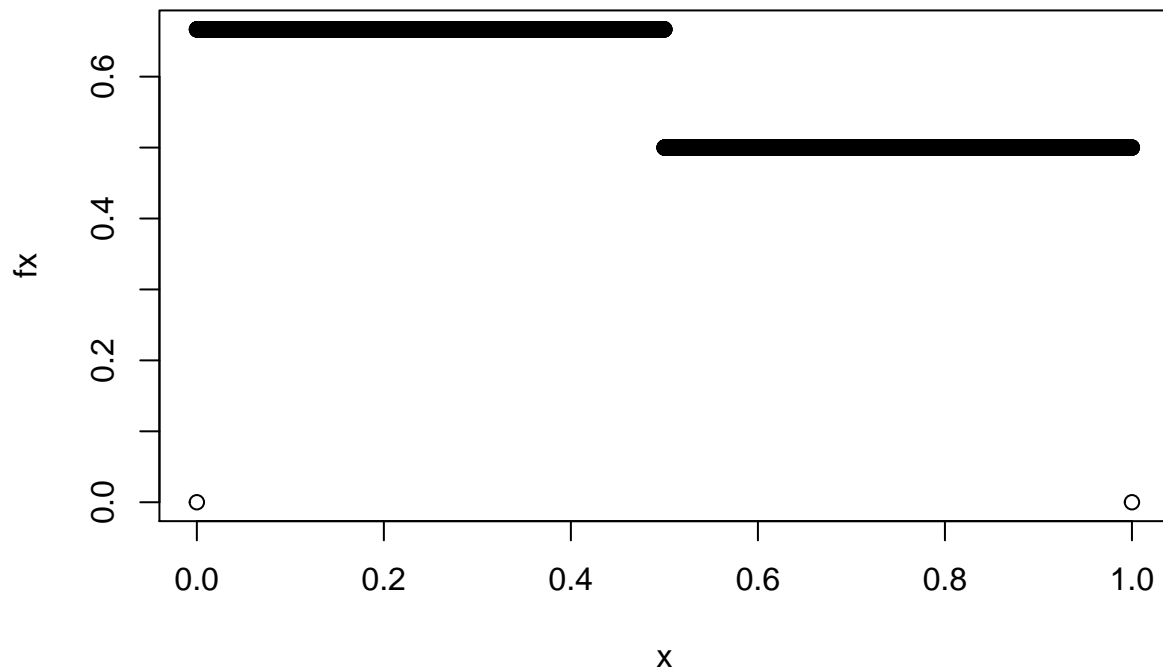
Histogram of c1


```
ks.test(c1,"punif")
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  c1
## D = 0.18, p-value = 0.3501
## alternative hypothesis: two-sided
```

```
x <- seq(0, 1, by = 0.0001)
fx <- ifelse(x >0 & x <= 1/2, 2/3,
ifelse(x >1/2 & x < 1, 1/2, 0))

plot(x, fx)
```

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

```
## Warning in ks.test(mu, x): p-value will be approximate in the presence of
## ties
```

```
##
##  Two-sample Kolmogorov-Smirnov test
##
## data:  mu and x
## D = 0.18004, p-value = 0.394
## alternative hypothesis: two-sided
```

According to the ks.test result, the p-value is larger than 0.05 which means data is uniformly distributed.
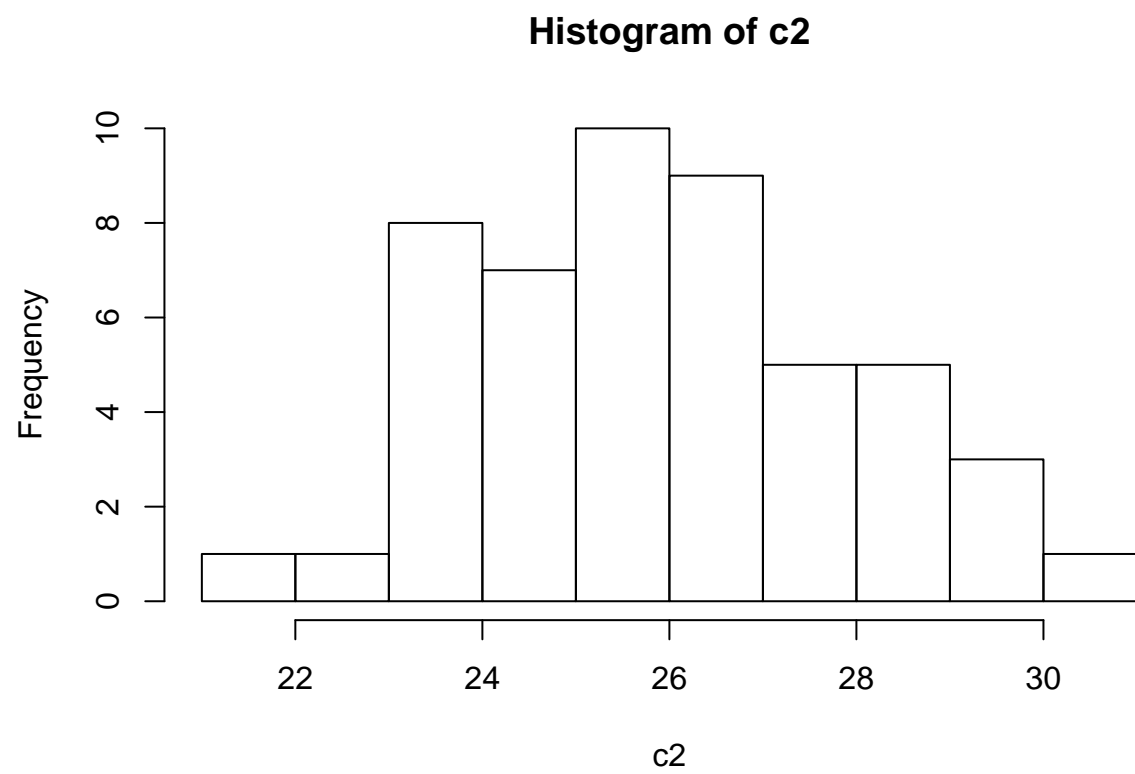Since D-value for fx is also bigger, which means Uniform would be better.

```
#Question 2
table2 <- read.table("maybe_normal.txt")
c2 <- c(table2$V1,table2$V2,table2$V3,table2$V4,table2$V5)

set.seed(2000)
n1 <- rnorm(50, mean = 26, sd = 2)
qqplot(c2,n1)
```
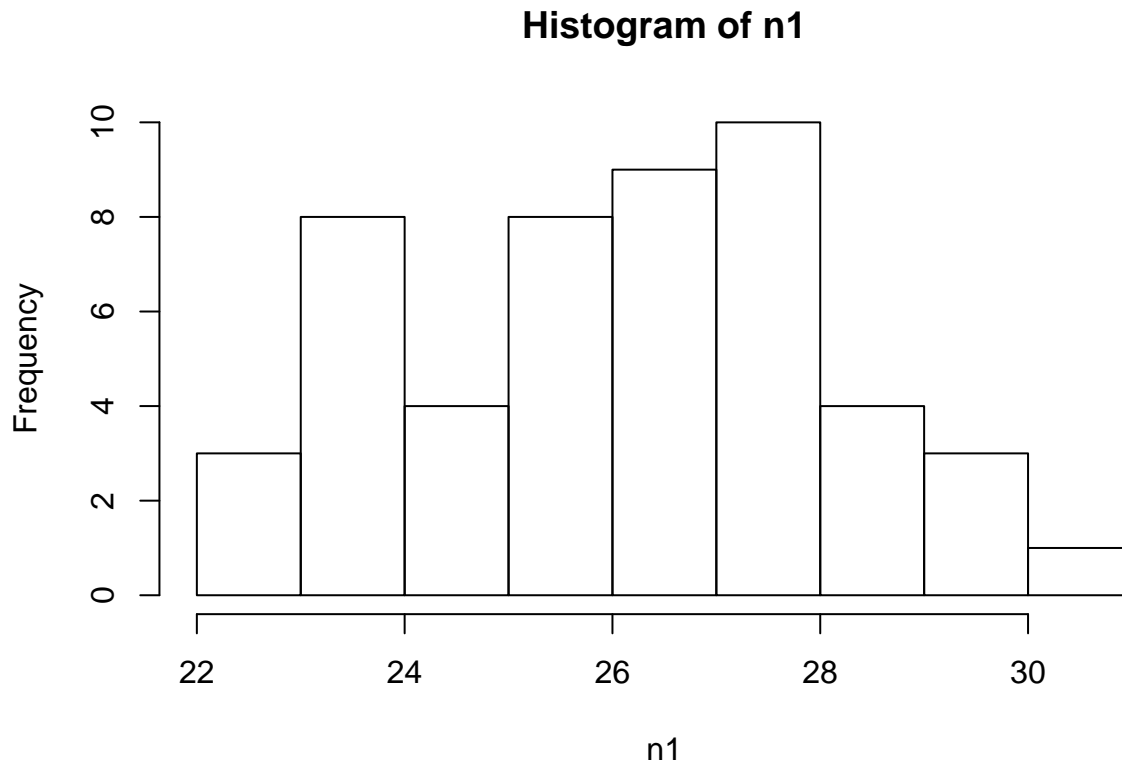
```r
hist(c2)
```

Histogram of c2


```r
hist(n1)
```
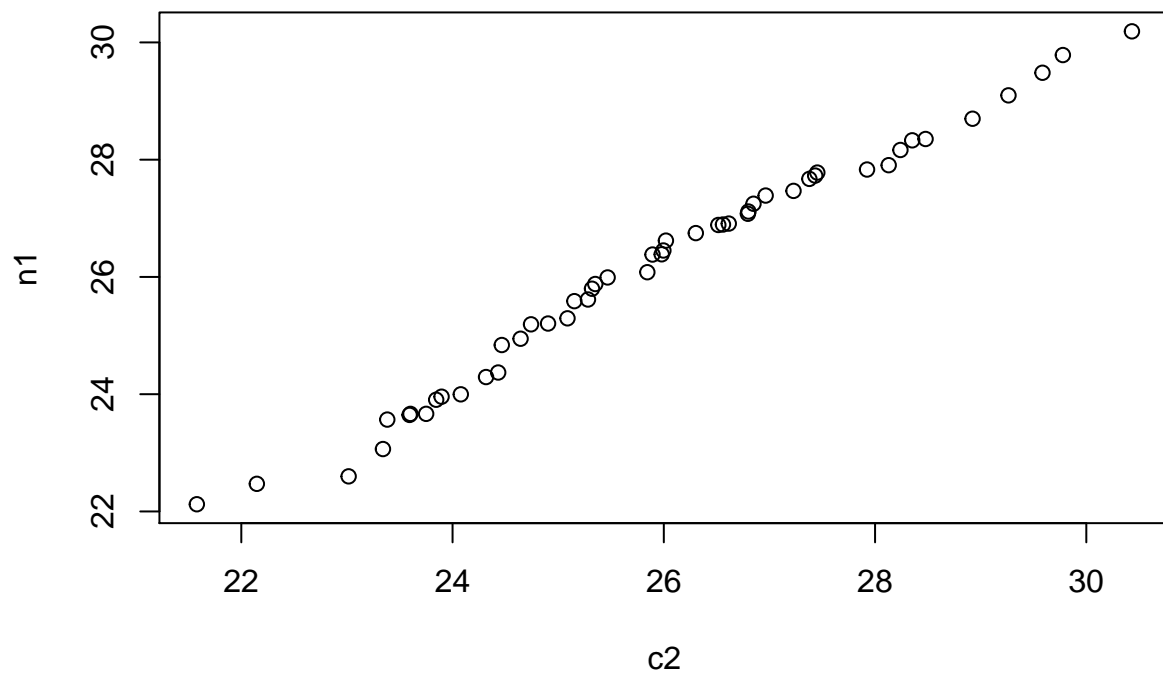
Histogram of n1


```
ks.test(c2,n1)
```

```
##
##  Two-sample Kolmogorov-Smirnov test
##
## data:  c2 and n1
## D = 0.12, p-value = 0.8693
## alternative hypothesis: two-sided
```

The data might generate by normal distribution, base on the result.

```
#Question 3
```

```
same1 <- 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'
```

```
same2 <- 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'
```

```
s1 <- c(same1$V1,same1$V2,same1$V3,same1$V4,same1$V5)
s2 <- c(same2$V1,same2$V2,same2$V3,same2$V4)
ks.test(s1, s2)
```

```
## Warning in ks.test(s1, s2): cannot compute exact p-value with ties
```

```
##
```

```
##  Two-sample Kolmogorov-Smirnov test
##
## data:  s1 and s2
## D = 0.25, p-value = 0.491
## alternative hypothesis: two-sided
ks.test(s1+2,s2)

## Warning in ks.test(s1 + 2, s2): cannot compute exact p-value with ties

##
##  Two-sample Kolmogorov-Smirnov test
##
## data:  s1 + 2 and s2
## D = 0.65, p-value = 0.0001673
## alternative hypothesis: two-sided
```
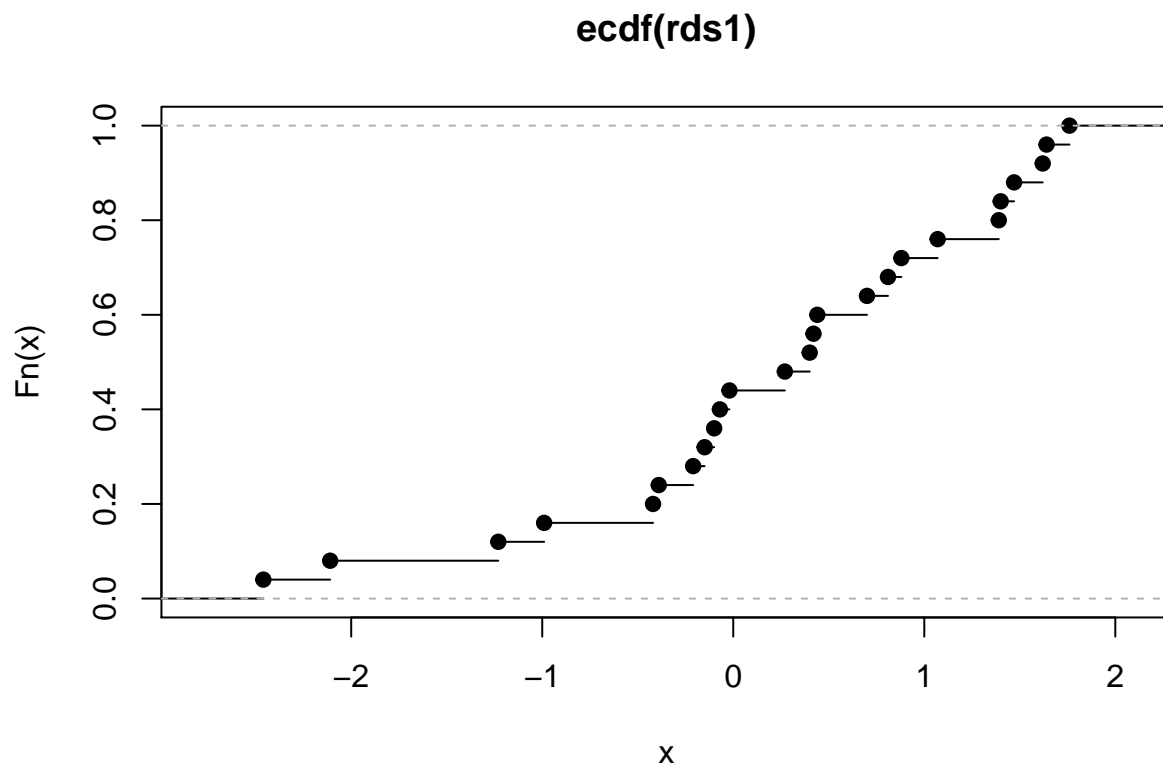
For ks.test(s1,s2), p-value is bigger than 0.05, fail to reject null hypothesis. s1,s2 from same distribution. For ks.test(s1+2,s2), p-value is very small, reject null hypothesis: s1+2,s2 not from same distribution.

```
rds1 <- readRDS("norm_sample.Rdata")
ecdf1 <- ecdf(rds1)
plot(ecdf1)
```

ecdf(rds1)


```
a <- rnorm(25,0,1)
ks.test(rds1,a)

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

```
##
## data:  rds1 and a
## D = 0.28, p-value = 0.285
## alternative hypothesis: two-sided
```

```r
dfns <- as.data.frame(rds1)
dfns$dif <- abs(ecdf1(rds1) - pnorm(rds1))
dstat <- max(dfns$dif)

dstat
```

```
## [1] 0.1372427
```

```r
ks.test(rds1,"pnorm")
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  rds1
## D = 0.17724, p-value = 0.3683
## alternative hypothesis: two-sided
```

```r
# Question 5
fijiquakes <- read.table("fijiquakes(1).dat", header = TRUE)
faithful <- read.table("faithful(1).dat", header = TRUE, skip = 20)

library(sfsmisc)
```
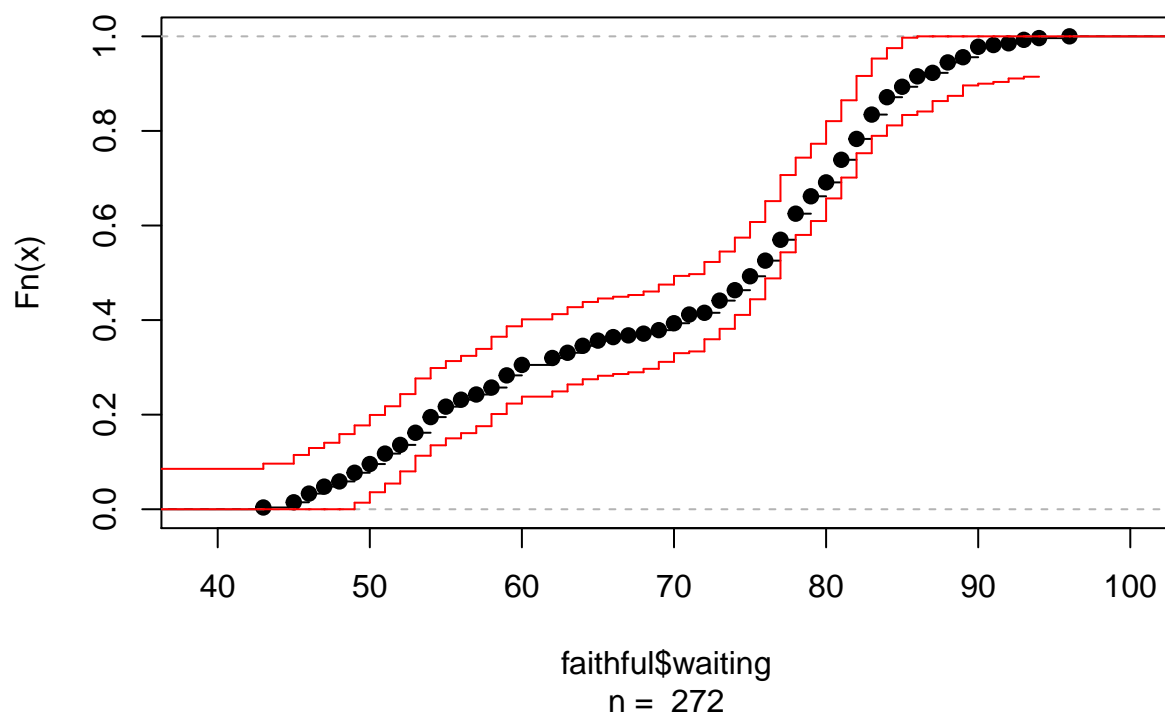
```
## Warning: package 'sfsmisc' was built under R version 3.5.2
```

```r
ecdf.ksCI(fijiquakes$mag)
```

ecdf(fijiquakes$mag) + 95% K.S. bands

fijiquakes$mag
n = 1000
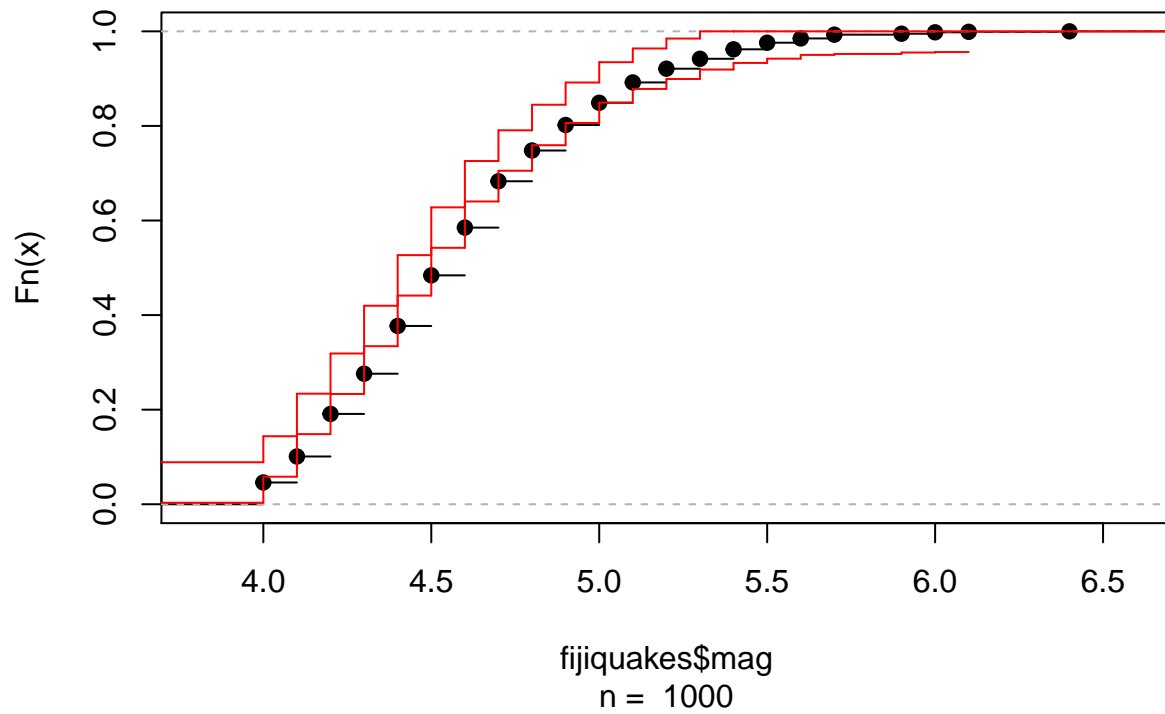
```
ecdf.ksCI(faithful$waiting)
```

ecdf(faithful$waiting) + 95% K.S. bands

faithful$waiting
n = 272

```r
amount <- sum( (fijiquakes$mag <= 4.9) & (fijiquakes$mag > 4.3))
library(Hmisc)
```

```
## Warning: package 'Hmisc' was built under R version 3.5.2
```

```
## Loading required package: lattice
```

```
## Loading required package: survival
```

```
## Loading required package: Formula
```

```
## Warning: package 'Formula' was built under R version 3.5.2
```

```
## Loading required package: ggplot2
```

```
##
## Attaching package: 'Hmisc'
```

```
## The following object is masked from 'package:sfsmisc':
##
##     errbar
```

```
## The following objects are masked from 'package:base':
##
##     format.pval, units
```

```r
binconf(amount, alpha = 0.05, length(fijiquakes$mag), method=c("wilson","exact","asymptotic","all"))
```

```
##  PointEst    Lower     Upper
##     0.526 0.4950118 0.5567892
```

```r
mean(faithful$waiting)
```

[1] 70.89706

```r
median(faithful$waiting)
```

[1] 76