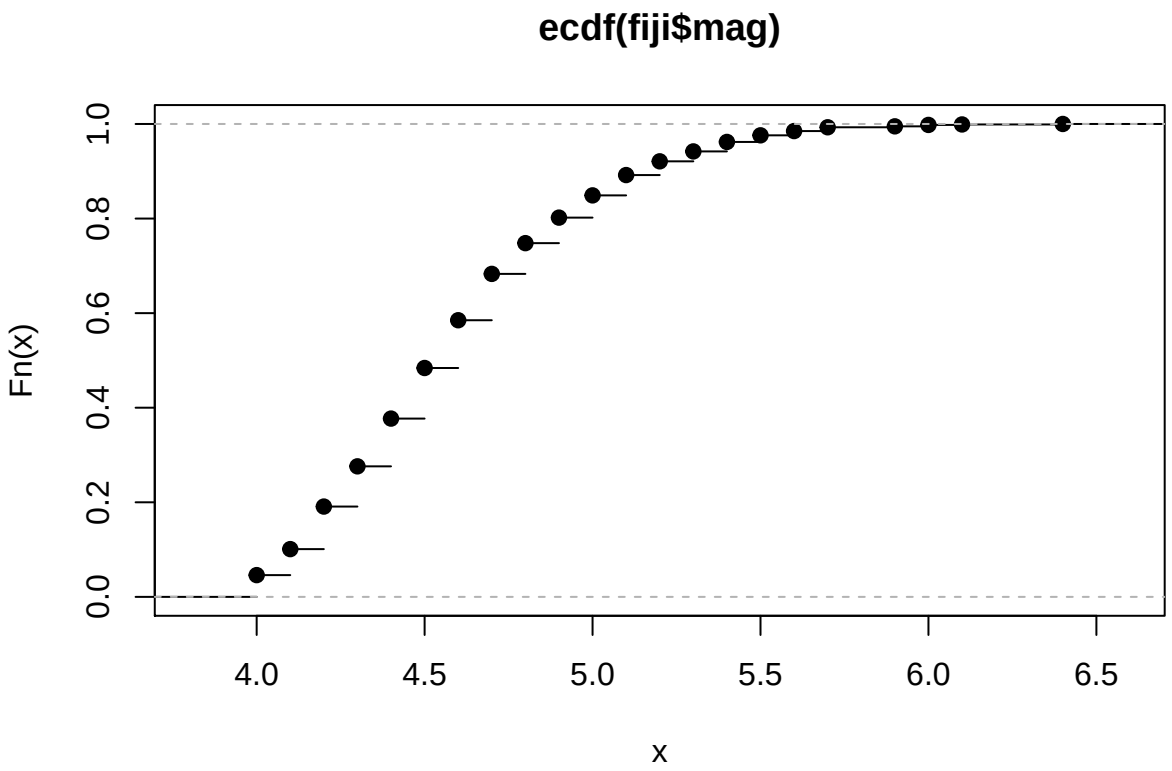
ecdf&KS probs

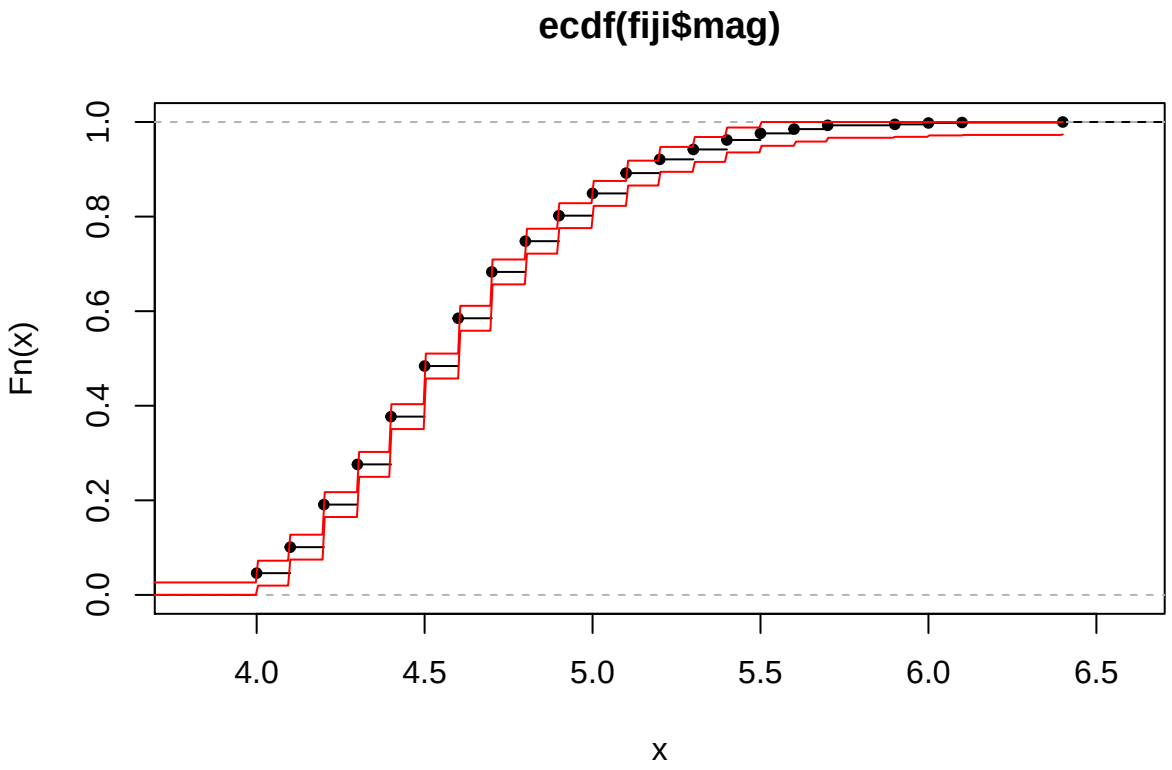
Yu Du

Fiji earthquakes

```
fiji<-read.csv("fijiquakes.csv", head=TRUE, sep="\t")
cdf_fiji <- ecdf(fiji$mag)
plot.ecdf(cdf_fiji)
```




```
n<-length(fiji$mag)
x <- seq(0, max(fiji$mag), length=n)
alpha<-0.5
fx <- cdf_fiji(x)
#Confidence band:
epsilon <- sqrt((1/(2*n))*log(2/alpha))
L <- pmax(fx - epsilon, 0)
U <- pmin(fx + epsilon, 1)
plot.ecdf(cdf_fiji, pch=20)
lines(x, L, type="l", lty=1, col="red")
lines(x, U, type="l", lty=1, col="red")
```




```
#F(4.9)-F(4.3):
a<-4.3
b<-4.9
cdf_fiji(b)-cdf_fiji(a)
```

```
## [1] 0.526
```

```
library(Hmisc)
```

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

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

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

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

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

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

```
sum<-sum((x<=4.9)&(x>4.3))
binconf(sum,length(x),method="wilson",alpha)
```

```
##  PointEst      Lower      Upper
##     0.093 0.08698902 0.09938113
```

Old Faithful Geyser

```
geyser<-read.csv("geysers.csv", head=TRUE)
```

```r
#Mean waiting time:
mean<-mean(geyser$waiting)
mean
```

```
## [1] 70.89706
```

```r
n<-length(geyser$waiting)
#Standard error:
se<-sd(geyser$waiting)/sqrt(n)
z<-1.645
mean-z*se
```

```
## [1] 69.54106
```

```r
mean+z*se
```

```
## [1] 72.25306
```

90% confidence Interval:(70,72)

```r
#Median:
median(geyser$waiting)
```

```
## [1] 76
```

KS problem Part 1.

```r
library(dgof)
```

```
##
## Attaching package: 'dgof'
```

```
## The following object is masked from 'package:stats':
##
##     ks.test
```

```r
#random sample for 25 values
sample<-c(0.42, 0.06, 0.88, 0.40,0.90, 0.38, 0.78, 0.71, 0.57, 0.66,
          0.48, 0.35, 0.16, 0.22, 0.08, 0.11, 0.29, 0.79, 0.75, 0.82, 0.30,
          0.23, 0.01, 0.41, 0.09)
k=ks.test(sample,"punif",0,1)
print(k)
```

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

The null hypothesis is not rejected. The table is form a random sample from the uniform distribution on the interval [0,1].

Part 2.

```r
#Create cdf:
f<-function(x)
  {if(0<x&&x<=1/2) {
  return(3/2*x) }
  else if (1/2<x&&x<1){
   return(1/2*(x+1))
```

```
  }
  else{return(0)}}
```

```
ks.test(sample,f)
```

```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  sample
## D = 0.4, p-value = 0.0004018
## alternative hypothesis: two-sided
```

The null hypothesis is rejected. The 25 values are not a random sample from a continuous distribution with pdf.