# Empiracle Assignment

*Dave Anderson*

*March 8, 2019*

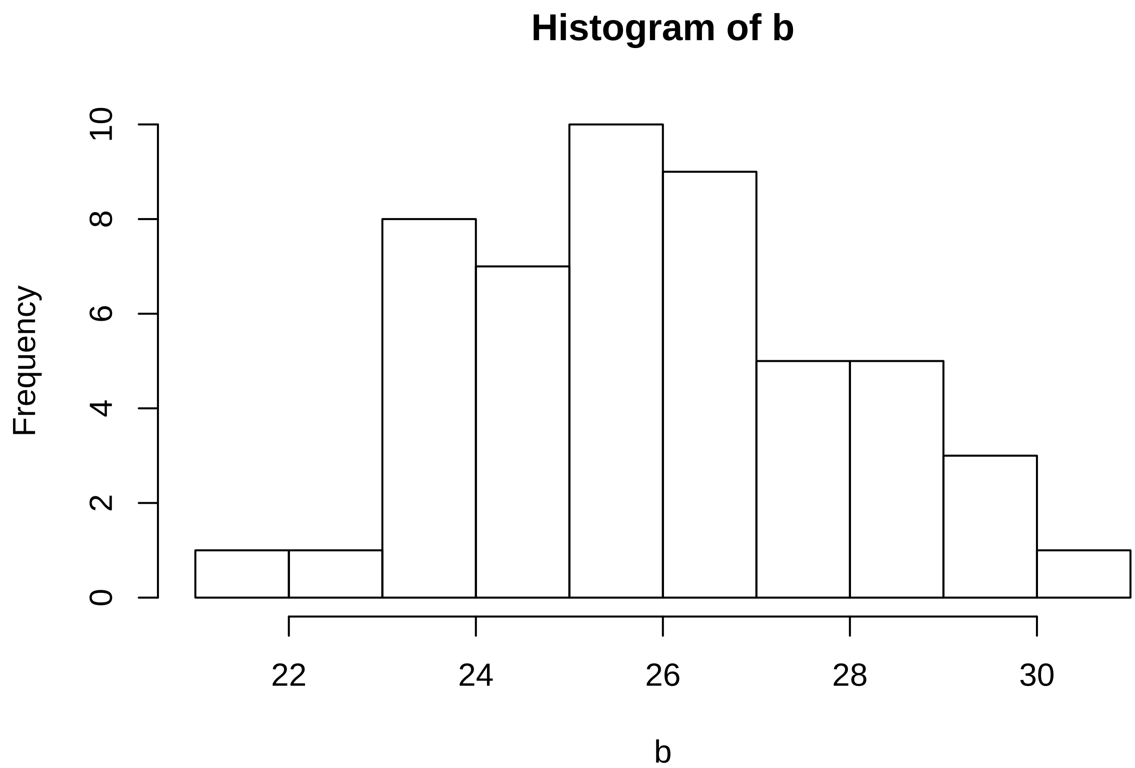
## 1)

```r
a <- 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)


ks.test(a,"punif",0,1)
```

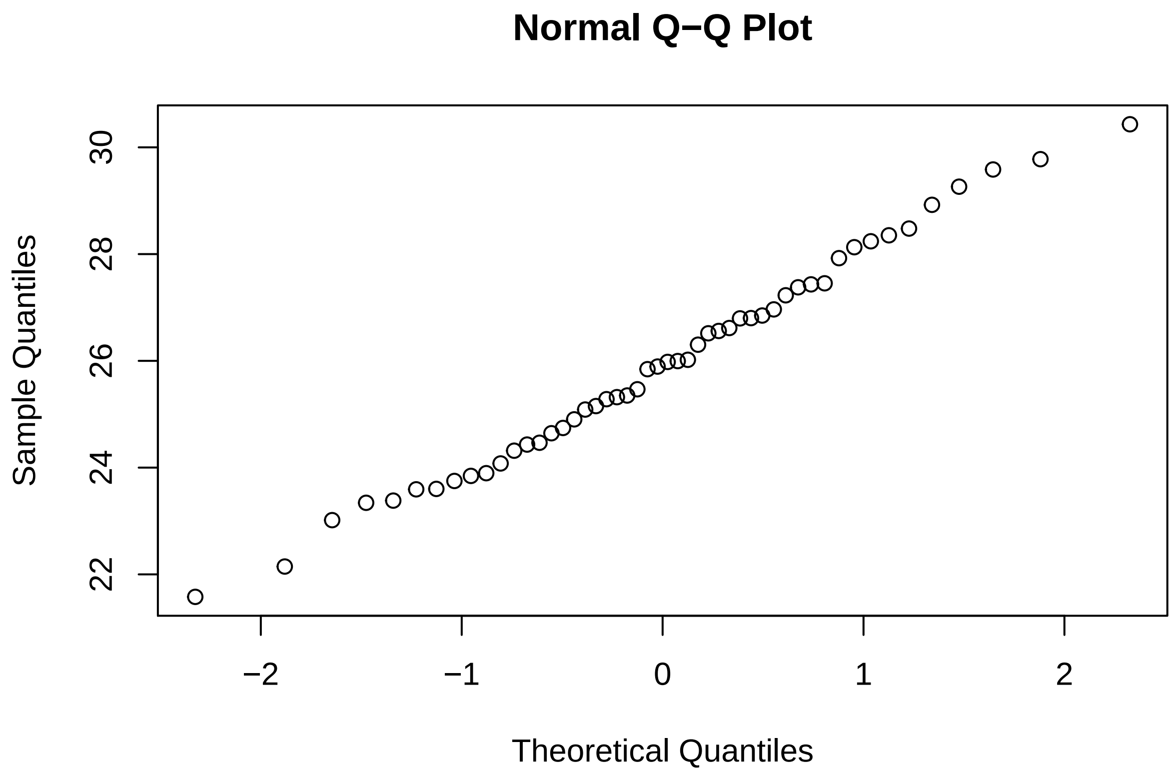
```
##
##  One-sample Kolmogorov-Smirnov test
##
## data:  a
## D = 0.18, p-value = 0.3501
## alternative hypothesis: two-sided
```

## 2)

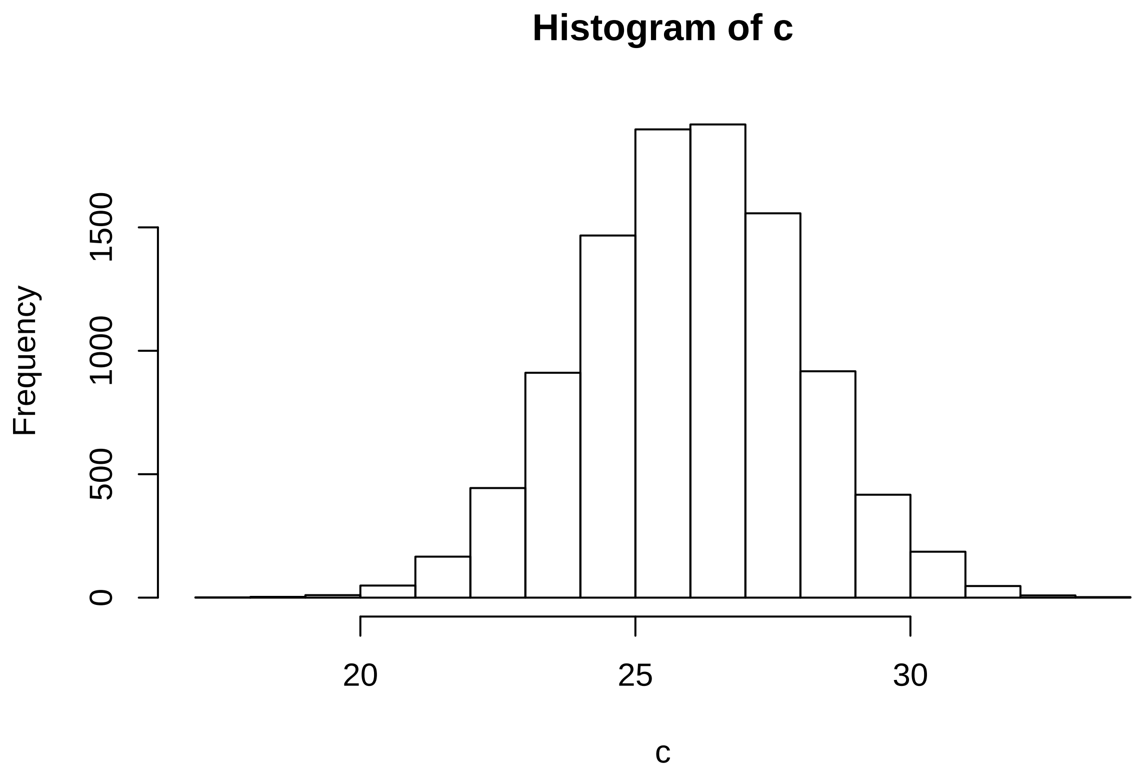
```r
b <- c(25.088, 26.615, 25.468, 27.453, 23.845,
25.996, 26.516, 28.240, 25.980, 30.432,
26.560, 25.844, 26.964, 23.382, 25.282,
24.432, 23.593, 24.644, 26.849, 26.801,
26.303, 23.016, 27.378, 25.351, 23.601,
24.317, 29.778, 29.585, 22.147, 28.352,
29.263, 27.924, 21.579, 25.320, 28.129,
28.478, 23.896, 26.020, 23.750, 24.904,
24.078, 27.228, 27.433, 23.341, 28.923,
24.466, 25.153, 25.893, 26.796, 24.743)

hist(b)
```




```r
qqnorm(b)
```




```
c <- rnorm(10000,26,2)


hist(c)
```




```r
ks.test(b,pnorm,26,2)
```

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

```r
ks.test(b,c)
```

```
##
##  Two-sample Kolmogorov-Smirnov test
##
## data:  b and c
## D = 0.0715, p-value = 0.9611
## alternative hypothesis: two-sided
```

*The data is normally distributed with a mean of 26 and variance of 4. The histogram looks normal, qqplot looks good. I ran a KS test on both the normal distribution, and a simulated dataset. Interestingly, I accidently included the mean as 28 the first time, and my tests had extremely low p-values. Shows how accurate it is!*

## 3)

```r
d <- c(0.61, 0.29, 0.06, 0.59,-1.73,
-0.74, 0.51,-0.56,0.39, 1.64,
```

```r
0.05,-0.06, 0.64,-0.82, 0.31,
1.77, 1.09, -1.28, 2.36, 1.31,
1.05,-0.32,0.40, 1.06,-2.47)

e <- c(2.20, 1.66, 1.38, 0.20,
0.36, 0.00, 0.96, 1.56,
0.44, 1.50,-0.30, 0.66,
2.31, 3.29, -0.27,-0.37,
0.38, 0.70, 0.52,-0.71)

ks.test(d,e)
```

```
##
##  Two-sample Kolmogorov-Smirnov test
##
## data:  d and e
## D = 0.23, p-value = 0.5286
## alternative hypothesis: two-sided
```

```r
ks.test(d+2,e)
```

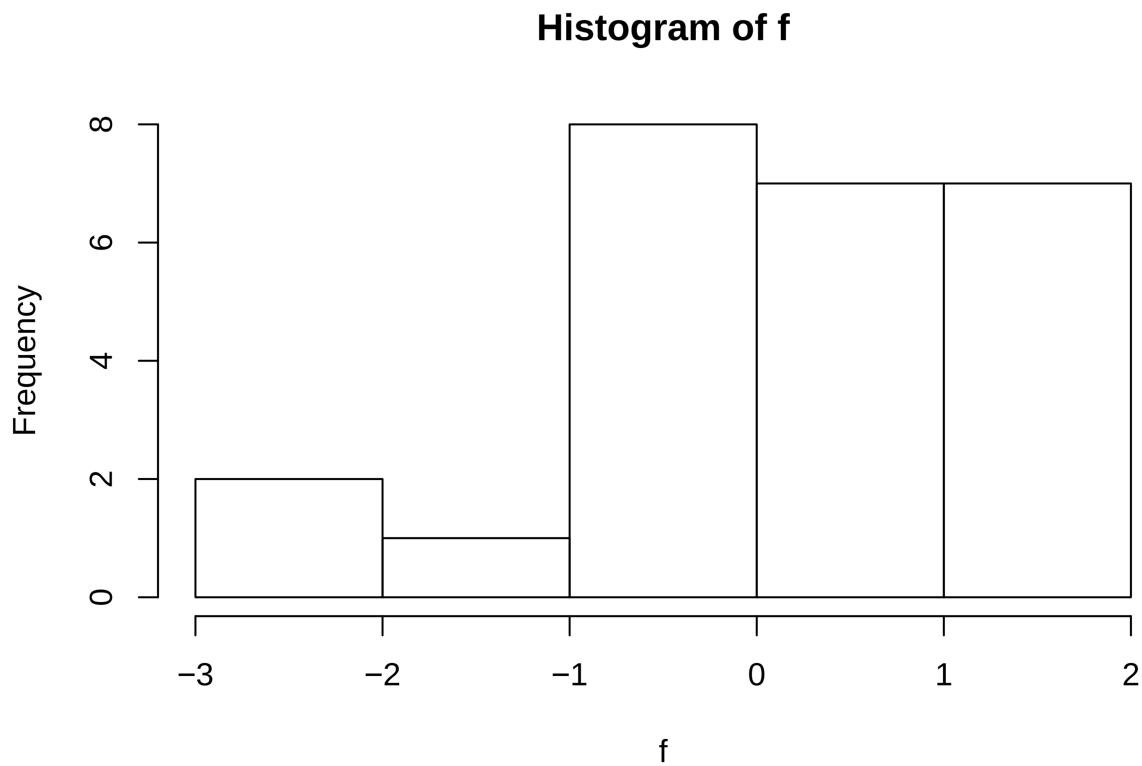
```
## Warning in ks.test(d + 2, e): cannot compute exact p-value with ties
```

```
##
##  Two-sample Kolmogorov-Smirnov test
##
## data:  d + 2 and e
## D = 0.61, p-value = 0.0005127
## alternative hypothesis: two-sided
```

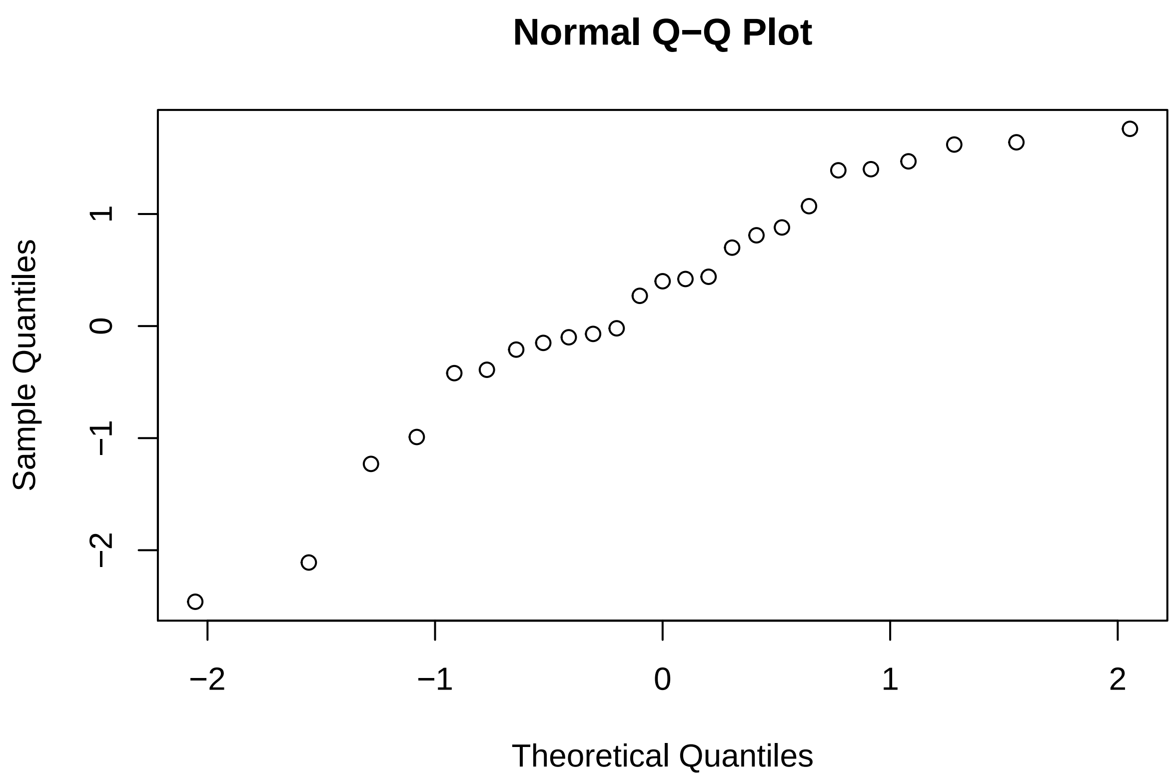
*It seems like X and Y have the same distribution, and X+2 does not have the same distirbution as Y.*

# 4)

```r
f <- readRDS("norm_sample.Rdata")
g <- rnorm(1000,0,1)
hist(f)
```




```r
qqnorm(f)
```




```
ks.test(f,"pnorm",0,1)
```

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

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

```
##
##  Two-sample Kolmogorov-Smirnov test
##
## data:  f and g
## D = 0.175, p-value = 0.4439
## alternative hypothesis: two-sided
```

*Yes, this data does follow the standard normal distribution, but not extremely well.*

```
cdf<- ecdf(f)
```

## 5)

```
fiji <- read.table("fijiquakes(1).dat",header = TRUE)
faith <- read.table("faithful(1).dat",header = TRUE)
```

```r
ecdf_fiji <- ecdf(fiji$mag)
ecdf_faith <- ecdf(faith$waiting)

e_faith <- sqrt(log(2/.1)/(2*272))
e_fiji <- sqrt(log(2/.05)/(2*1000))


answ <- ecdf_fiji(4.9)-ecdf_fiji(4.3)
answ + e_fiji
```

```
## [1] 0.5689469
```

```r
answ - e_fiji
```

```
## [1] 0.4830531
```

*F(4.9) - F(4.3) 95% confidence interval: [0.48, 0.57]

```r
avg <- mean(faith$waiting)

answ2 <- ecdf_faith(mean(avg))

answ2 + e_faith
```

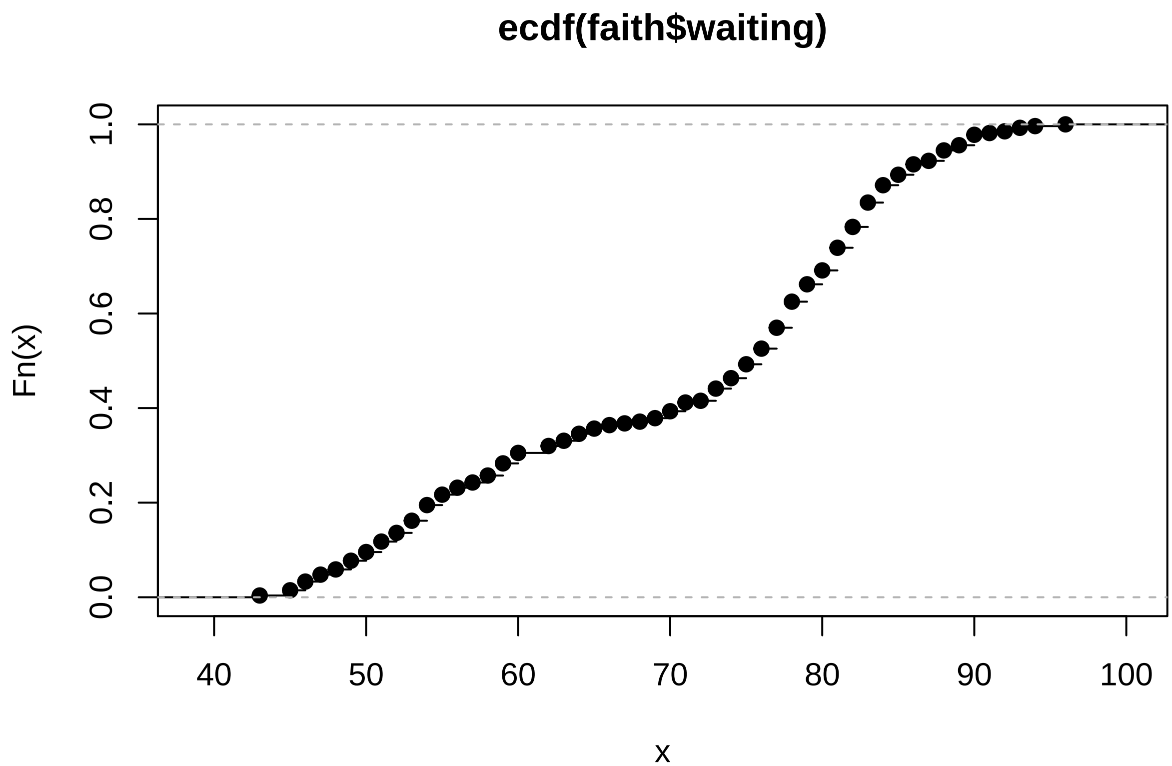
```
## [1] 0.4675906
```

```r
answ2 - e_faith
```

```
## [1] 0.3191741
```

*The confidence band for the mean waitig time, F(70.9) is[0.32,0.46].*

```r
plot(ecdf_faith)
```




```
quantile(ecdf_faith,probs = .5)
```

```
## 50%
##  76
```

*The median wait time based on our ecdf is 76.*