

# MA677 Homework

*Andrew Zhang*

*3/7/2019*

## Problem 1

```
maybe_unif <- data.frame(value = c(0.42, 0.06, 0.88, 0.40, 0.90, 0.38, 0.78, 0.71, 0.57, 0.66, 0.48, 0.1))
chisq.test(maybe_unif)
```

```
## Warning in chisq.test(maybe_unif): Chi-squared approximation may be
## incorrect

##
## Chi-squared test for given probabilities
##
## data:  maybe_unif
## X-squared = 4.4714, df = 24, p-value = 1
```

The p-value is extremely large so we cannot reject the null hypothesis and cannot conclude this is a uniform distribution.

It seems like the model given may be better as it contains uniform values amongst the varying ranges.

## Problem 2

```
maybe_norm <- read.delim("/Users/andrewzhang/Downloads/maybe_normal.txt", header = FALSE, sep = " ")
maybe_norm <- stack(maybe_norm)
maybe_norm$ind <- NULL

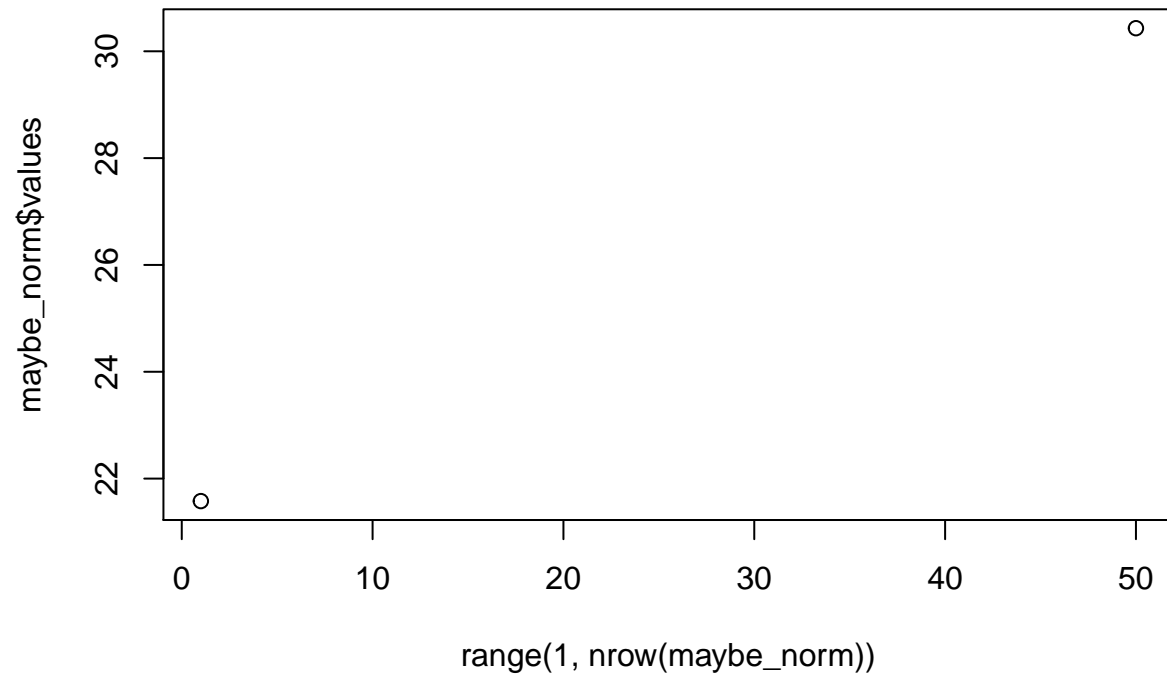
mean(maybe_norm$values)
```

```
## [1] 25.94258
```

```
sd(maybe_norm$values)
```

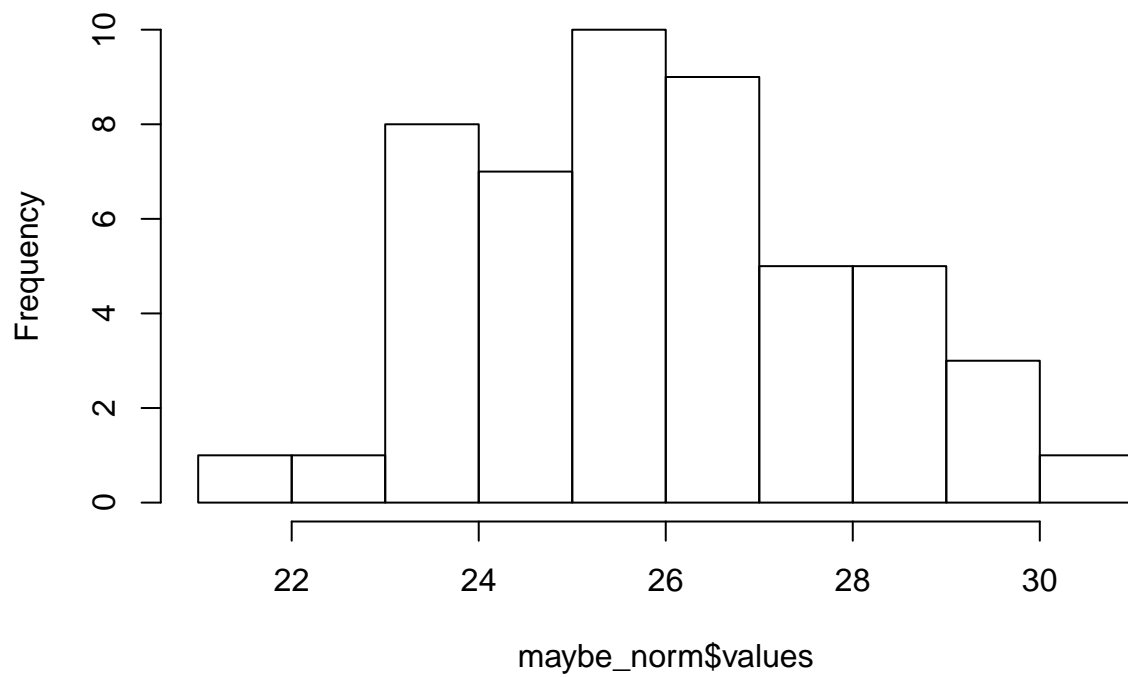
```
## [1] 2.042374
```

```
qqplot(range(1, nrow(maybe_norm)), maybe_norm$values)
```



```
hist(maybe_norm$values)
```

**Histogram of maybe\_norm\$values**



### Problem 3

```
maybe_same1 <- data.frame(value = c(0.61, 0.29, 0.06, 0.59, -1.73, -0.74, 0.51, -0.56, -0.39, 1.64, 0.0))
maybe_same2 <- data.frame(value = c(2.20, 1.66, 1.38, 0.20, 0.36, 0.00, 0.96, 1.56, 0.44, 1.50, -0.30))

ks.test(maybe_same1$value, maybe_same2$value)

##
## Two-sample Kolmogorov-Smirnov test
##
## data: maybe_same1$value and maybe_same2$value
## D = 0.27, p-value = 0.3357
## alternative hypothesis: two-sided
thresh1 <- 1.224 * sqrt((nrow(maybe_same1) + nrow(maybe_same2))/(nrow(maybe_same1)*nrow(maybe_same2)))
thresh1

## [1] 0.3672
```

Since D from the `ks.test` is not bigger than the critical value, we cannot reject the null hypothesis and cannot conclude these two samples are from the same distribution.

### Problem 4

```
norm_samp <- readRDS("/Users/andrewzhang/Downloads/norm_sample.Rdata")
ecdf(scale(norm_samp))

## Empirical CDF
## Call: ecdf(scale(norm_samp))
## x[1:25] = -2.4176, -2.1047, -1.3182, ..., 1.247, 1.3543
ks.test(scale(norm_samp), norm_samp)

##
## Two-sample Kolmogorov-Smirnov test
##
## data: scale(norm_samp) and norm_samp
## D = 0.24, p-value = 0.4755
## alternative hypothesis: two-sided
thresh2 <- 1.224 * sqrt((length(scale(norm_samp)) + length(norm_samp))/(length(scale(norm_samp))*length(norm_samp)))
thresh2

## [1] 0.3461995
```

Here we can see that D is less than the critical value, so we cannot reject the null hypothesis and cannot conclude these samples are of the same distribution.

### Problem 5

```
faith <- read.table("/Users/andrewzhang/Downloads/faithful(1).dat", fill = TRUE)
faith <- faith[16:nrow(faith),]
```

```

faith$V2 <- as.numeric(faith$V2)
faith[4:10] <- NULL
fiji <- read.table("/Users/andrewzhang/Downloads/fijiquakes(1).dat", fill = TRUE, header = TRUE)

ecdf(faith$V2)

## Empirical CDF
## Call: ecdf(faith$V2)
## x[1:126] =      3,      4,      5, ...,    127,    128

ecdf(fiji$mag)

## Empirical CDF
## Call: ecdf(fiji$mag)
## x[1:22] =      4,      4.1,      4.2, ...,      6.1,      6.4

CI1 <- mean(faith$V2) - 1.645 * (sd(faith$V2)/length(faith$V2))
CI2 <- mean(faith$V2) + 1.645 * (sd(faith$V2)/length(faith$V2))
CI <- c(CI1, CI2)
CI

## [1] 65.20303 65.65727

mag49 <- fiji[fiji$mag == 4.9, ]
mag43 <- fiji[fiji$mag == 4.3, ]
magCI1 <- mean(mag49$depth) - mean(mag43$depth) - 1.96 * (sqrt(var(mag49$depth) + var(mag43$depth))/nrow(mag49))
magCI2 <- mean(mag49$depth) - mean(mag43$depth) + 1.96 * (sqrt(var(mag49$depth) + var(mag43$depth))/nrow(mag49))
CI_mag <- c(magCI1, magCI2)
CI_mag

## [1] -120.71944 -98.50976

```