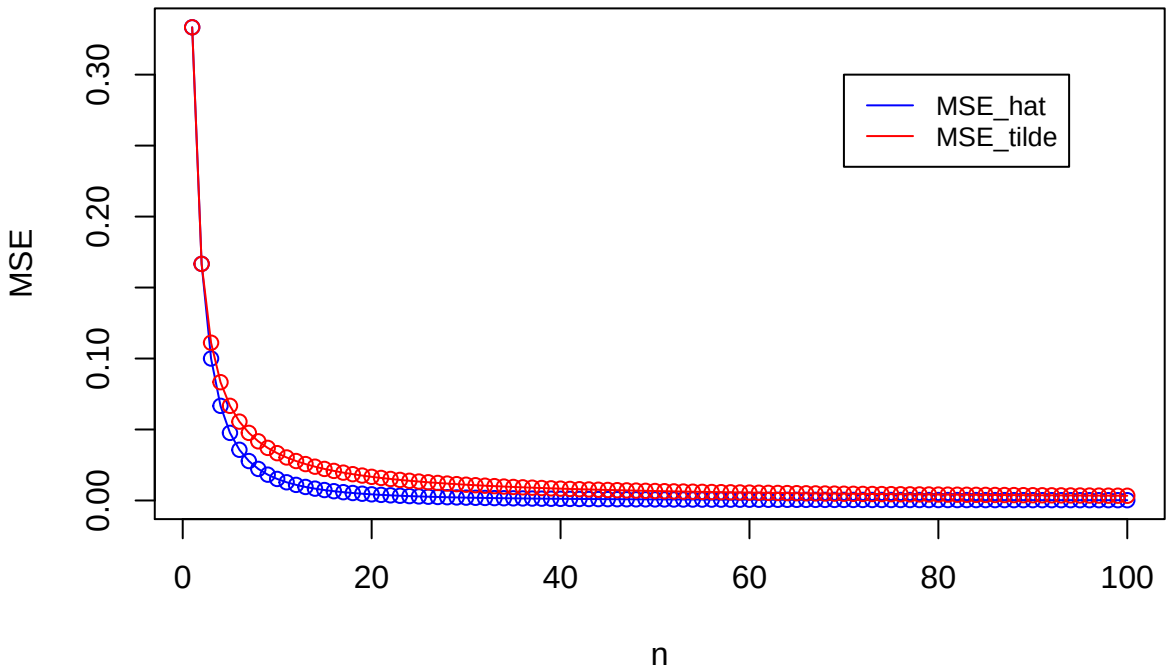
HW02 Report

Ninh (Vincent) DO

February 6, 2019

Problem 1d

```
theta = 1
n = 1 : 100
MSE_hat = 2 * theta ^ 2 / ((n + 1) * (n + 2))
MSE_tilde = theta ^ 2 / (3 * n)
plot(n, MSE_hat, type = "o", col = "blue", xlab = "n", ylab = "MSE")
lines(n, MSE_tilde, type = "o", col = "red")
legend(70, 0.3, legend=c("MSE_hat", "MSE_tilde"),
       col=c("blue", "red"), lty=c(1, 1), cex=0.8)
```



When n is small, prefer theta_hat.

Problem 6

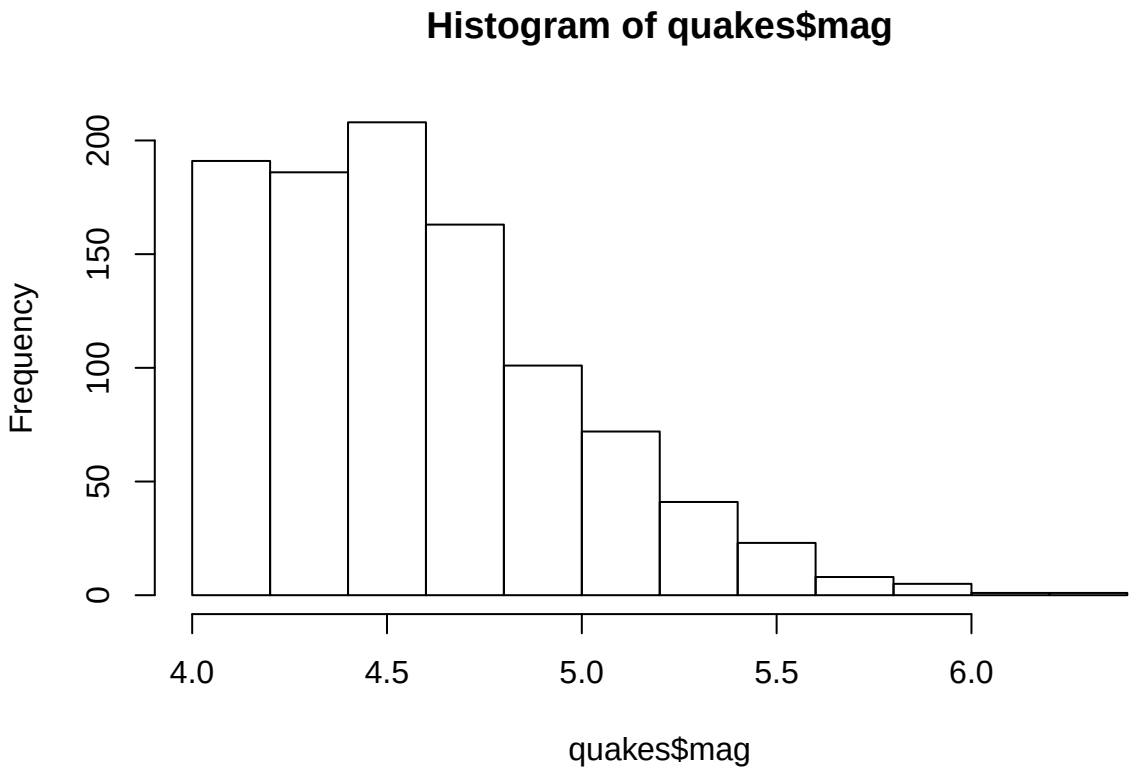
Working with data on the magnitudes of earthquakes near Fiji

```
quakes = read.table(file = "fijiquakes.dat", header = TRUE)
head(quakes)
```

```
##   Obs.    lat   long depth mag stations
## 1    1 -20.42 181.62   562 4.8       41
## 2    2 -20.62 181.03   650 4.2       15
## 3    3 -26.00 184.10    42 5.4       43
## 4    4 -17.97 181.66   626 4.1       19
```

```
## 5    5 -20.42 181.96   649 4.0       11
## 6    6 -19.68 184.31   195 4.0       12
```

```
hist(quakes$mag)
```



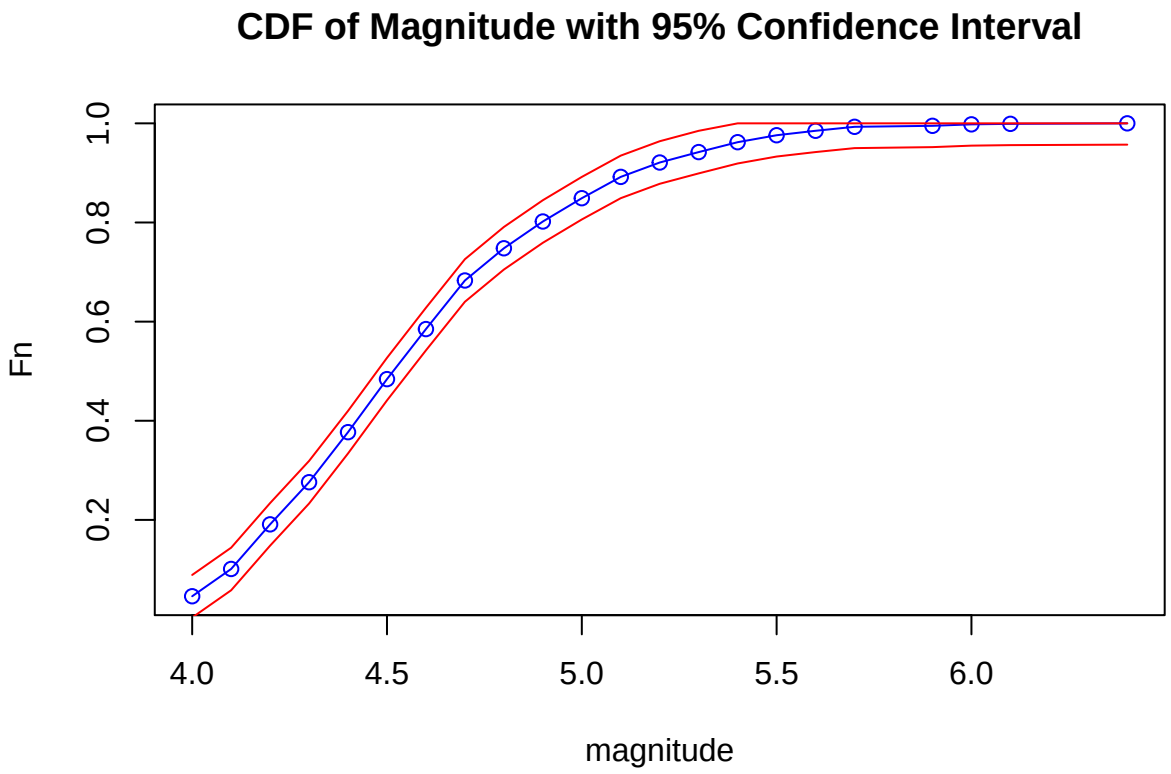

```
n = length(quakes$mag)
mag_cdf_func = ecdf(quakes$mag)
mag_vec = levels(factor(quakes$mag))
mag_cdf = mag_cdf_func(mag_vec)
mag_cdf_lower = pmax(mag_cdf - (1 / (2 * n) * log(2 / 0.05)) ^ 0.5, 0)
mag_cdf_lower
```

```
##  [1] 0.003053059 0.058053059 0.148053059 0.233053059 0.334053059
##  [6] 0.441053059 0.542053059 0.640053059 0.705053059 0.759053059
## [11] 0.806053059 0.849053059 0.878053059 0.899053059 0.919053059
## [16] 0.933053059 0.942053059 0.950053059 0.952053059 0.955053059
## [21] 0.956053059 0.957053059
```

```
mag_cdf_upper = pmin(mag_cdf + (1 / (2 * n) * log(2 / 0.05)) ^ 0.5, 1)
mag_cdf_upper
```

```
##  [1] 0.08894694 0.14394694 0.23394694 0.31894694 0.41994694 0.52694694
##  [7] 0.62794694 0.72594694 0.79094694 0.84494694 0.89194694 0.93494694
## [13] 0.96394694 0.98494694 1.00000000 1.00000000 1.00000000 1.00000000
## [19] 1.00000000 1.00000000 1.00000000 1.00000000
```

```
plot(mag_vec, mag_cdf, type = "o", col = "blue", xlab = "magnitude", ylab = "Fn", main = "CDF of Magnitu
lines(mag_vec, mag_cdf_lower, type = "l", col = "red")
lines(mag_vec, mag_cdf_upper, type = "l", col = "red")
```



Problem 7

Working with data on cloud seeding

```
clouds = read.table(file = "clouds.dat", header = TRUE)
head(clouds)
```

```
##   Unseeded Seeded
## 1   1202.6 2745.6
## 2    830.1 1697.8
## 3    372.4 1656.0
## 4    345.5  978.0
## 5    321.2  703.4
## 6    244.3  489.1
```

```
n = length(clouds$Seeded)
theta = mean(clouds$Seeded) - mean(clouds$Unseeded)
theta
```

```
## [1] 277.3962
```

```
se_theta = (var(clouds$Seeded) / n + var(clouds$Unseeded) / n) ^ 0.5
se_theta
```

```
## [1] 138.8199
```

```
# 95% confidence interval
cf_95 = c(theta - 1.96 * se_theta, theta + 1.96 * se_theta)
cf_95
```

```
## [1]   5.309116 549.483192
```