# lecture 4

## Jung-Jin Lee

## 1/28/2020

```r
ht <- read.table(file = "heights.txt", header = T, sep = " ")
dim(ht)
```

```
## [1] 1375    2
```

```r
pnorm(12, mean = 13, sd = 1)
```

```
## [1] 0.1586553
```

```r
pnorm(13, 13, 1)
```

```
## [1] 0.5
```

```r
1- pnorm(14, mean = 13, sd = 1)
```

```
## [1] 0.1586553
```

```r
pnorm(14, mean = 13, sd = 1, lower.tail = F)
```

```
## [1] 0.1586553
```

```r
pnorm(14, mean = 13, sd = 1) - pnorm(11, mean = 13, sd = 1)
```

```
## [1] 0.8185946
```

```r
1 - pnorm(15, mean = 13, sd = 1)
```

```
## [1] 0.02275013
```

```r
pnorm(15, mean = 13, sd = 1, lower.tail = F)
```

```
## [1] 0.02275013
```

```r
pnorm(-1, mean = 0, sd = 1)
```

```
## [1] 0.1586553
```

```r
pnorm(-1, 0, 1)
```

```
## [1] 0.1586553
```

```r
pnorm(-1)
```

```
## [1] 0.1586553
```

```r
# guess work
pnorm(1)
```

```
## [1] 0.8413447
```

```r
pnorm(1.5)
```

```
## [1] 0.9331928
```

```r
pnorm(1.3)
```

```
## [1] 0.9031995
```

```r
pnorm(1.27)
```

```
## [1] 0.8979577
```

```r
qnorm(0.9, mean = 0, sd = 1)
```

```
## [1] 1.281552
```

```r
pnorm(1.28152)
```

```
## [1] 0.8999945
```

```r
qnorm(0.5, mean = 11, sd = 3)
```

```
## [1] 11
```

```r
# need t_5_0.2
qt(0.8, df = 5)
```

```
## [1] 0.9195438
```

```r
# need t_4_0.025
qt(1-0.025, df = 4)
```

```
## [1] 2.776445
```

```r
# need z_0.025
qnorm(1-0.025)
```

```
## [1] 1.959964
```

```r
mean(iris$Sepal.Width)
```

```
## [1] 3.057333
```

```r
sd(iris$Sepal.Width)
```

```
## [1] 0.4358663
```

```r
sample1 <- c(12, 13, 13, 14, 13); sample2 <- c(10, 11, 16, 13, 15)

alpha <- 0.05
xbar <- mean(sample1)
sdev <- sd(sample1)
n <- length(sample1)
tval <- qt(1-alpha/2, df = n-1)
CI95 <- xbar + tval*sdev/sqrt(n)*c(-1,1)
```
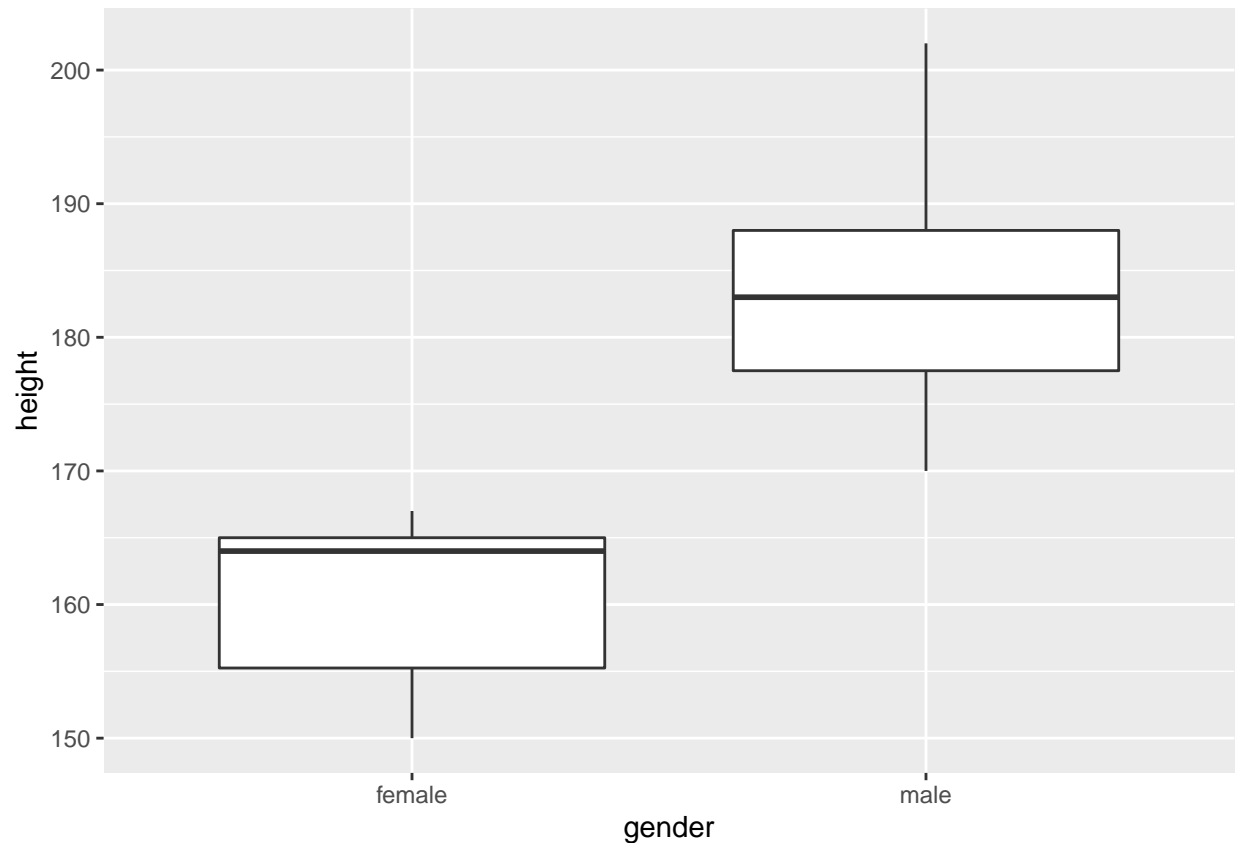
```r
t.test(ht$Mheight)
```

```
##
##  One Sample t-test
##
## data:  ht$Mheight
## t = 983.32, df = 1374, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  62.32821 62.57739
## sample estimates:
## mean of x
##   62.4528
```

```r
st <- starwars
```

```r
d <- st %>%
  filter(species == "Human", !is.na(height))
```

```r
ggplot(d, aes(gender, height)) +
  geom_boxplot()
```

```
female <- d %>% filter(gender == "female") %>%
  arrange(desc(height))

male <- d %>% filter(gender == "male")
t.test(male$height)
```

```
##
##  One Sample t-test
##
## data:  male$height
## t = 106.79, df = 22, p-value < 2.2e-16
## alternative hypothesis: true mean is not equal to 0
## 95 percent confidence interval:
##  178.8068 185.8889
## sample estimates:
## mean of x
##  182.3478
```