MATH 392 Problem Set 6

EJ Arce

19 March 2018

```r
# Load data
x <- c(47, 126, 285, 318, 142, 55, 231,
       102, 164, 85, 242, 62, 289, 290)
```

1.1

$$X \sim Unif(0, \theta)$$

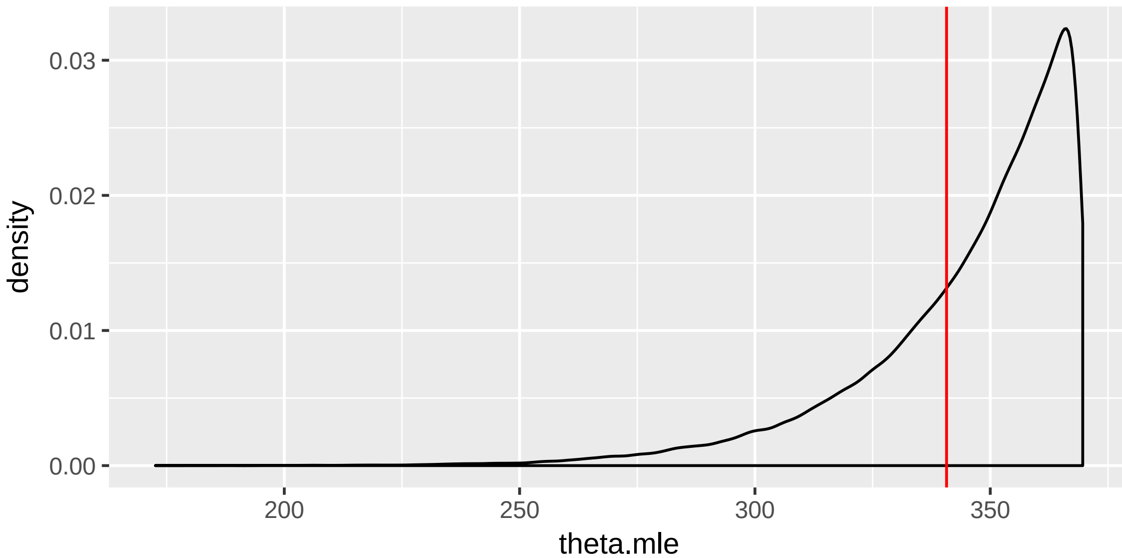
$$H_0 : \theta = 345$$

$$H_A : \theta > 345$$

$$\hat{\theta}_{MLE,corr} = \frac{n+1}{n} X_{max}$$

```r
# Calculate test statistic
xmax <- max(x)
n <- length(x)
theta.test.mle <- (n+1)*xmax/n
```

```r
# Simulate sampling distribution of thetas from null distribution
nsim <- 10^5
max.null <- 345
theta.mle <- rep(NA,nsim)
for(i in 1:nsim){
  sample <- runif(n,0,max.null)
  theta.mle[i] <- (n+1)*max(sample)/n
}
null.samp.mle <- data.frame(theta.mle)
ggplot(null.samp.mle, aes(x=theta.mle)) +
  geom_density() +
  geom_vline(xintercept = theta.test.mle, col = "red")
```



Now, calculate the proportion of simulations that produced estimates more extreme than the test statistic:

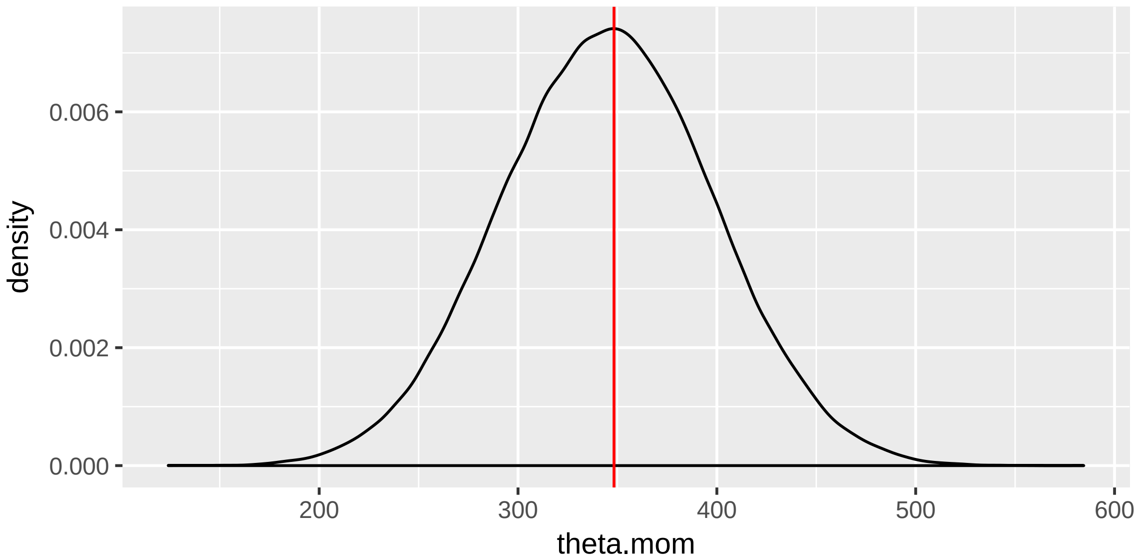
```
(sum(null.samp.mle$theta.mle > theta.test.mle) + 1)/(length(null.samp.mle$theta.mle) + 1)
```

```
## [1] 0.6810232
```

Using a reasonable α level, the simulation results are not statistically significant enough to reject the null hypothesis.

1.2

```
theta.test.mom <- 2*mean(x)
nsim <- 10^5
max.null <- 345
theta.mom <- rep(NA,nsim)
for(i in 1:nsim){
  sample <- runif(n,0,max.null)
  theta.mom[i] <- 2*mean(sample)
}
null.samp.mom <- data.frame(theta.mom)
ggplot(null.samp.mom, aes(x=theta.mom)) +
  geom_density() +
  geom_vline(xintercept = theta.test.mom, col = "red")
```




```
(sum(null.samp.mom$theta.mom > theta.test.mom) + 1)/(length(null.samp.mom$theta.mom) + 1)
```

```
## [1] 0.4769752
```

1.3

The two types of error committed when concluding hypothesis tests are

- Type I: rejecting H_0 when it is actually true
- Type II:retaining H_0 when H_A is actually true

In the case of the German tank problem, committing a type I error would mean that the Western allies believe that there are more German tanks than there actually are. A type II error would mean that the Western allies believe the Germans have 345 tanks, when in fact they have more. Thus, a type II error is more consequential than a type I error.

1.4

Suppose $\alpha = .05$.

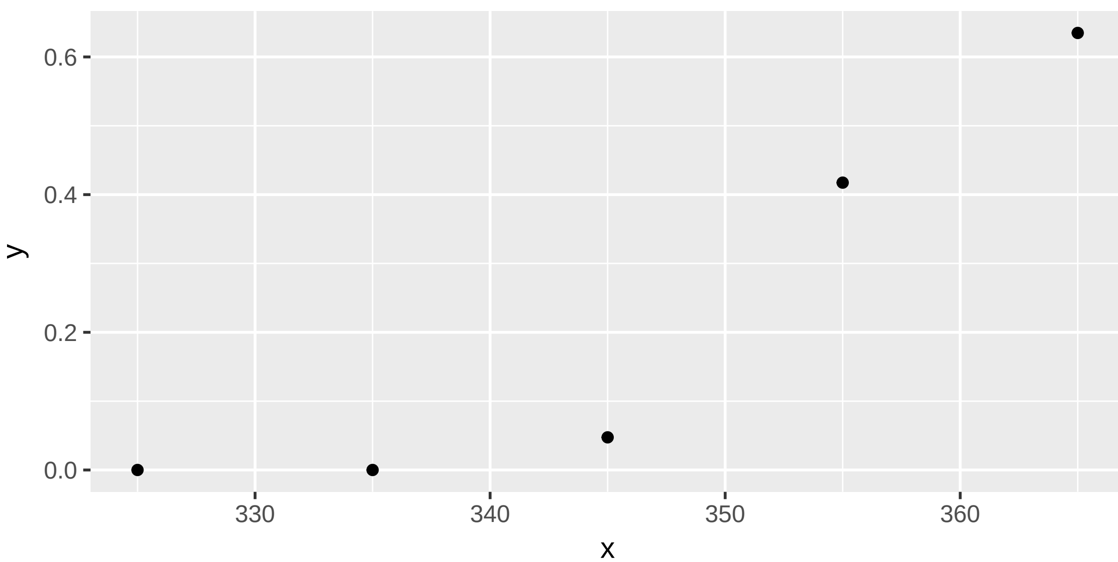
```
n <- 17
nsim <- 10^5
theta1 <- 325
theta2 <- 335
theta3 <- 345
theta4 <- 355
theta5 <- 365
```

```
# Power curve for theta parameters
f.mle <- function(theta.a){
  nsim <- 10^5
  theta.0 <- 345
  test <- (n+1)*(max(x))/n
  thetas.0 <- rep(NA,nsim)
  thetas.a <- rep(NA,nsim)
  for(i in 1:nsim){
    sample.0 <- runif(n,0,theta.0)
```

```
    sample.a <- runif(n,0,theta.a)
    thetas.0[i] <- (n+1)*(max(sample.0))/n
    thetas.a[i] <- (n+1)*(max(sample.a))/n
  }
  sampling.0 <- data.frame(thetas.0)
  sampling.a <- data.frame(thetas.a)
  alpha.crit <- quantile(sampling.0$thetas.0,.95)
  power <- sum(sampling.a$thetas.a > alpha.crit)/length(sampling.a$thetas.a)
  power
}
f.mle(theta5)
```

```
## [1] 0.63534
```

```
d <- data.frame(x = c(theta1,theta2,theta3,theta4,theta5),
                y = c(f.mle(theta1),f.mle(theta2),
                      f.mle(theta3),f.mle(theta4),f.mle(theta5)))
ggplot(d,aes(x=x,y=y)) + geom_point()
```



As expected, as the true alternative hypothesis increases, the power increases.