MA677 Assignment1

Xijia Luo

February 9, 2021

Part 1

In this part, we are discussing how the authors reached their conclusion that the critical value should be between 69 and 73 people cured. Firstly, they set $\alpha(p)$ as the probability of type 1 error, defined to be the probability that we reject the null hypothesis, where this probability is calculated under the assumption that the null hypothesis is true.

Then they set $\beta(p)$ as the probability of type 2 error, defined to be the probability that we accept the null hypothesis, where this probability is calculated under the assumption that the null hypothesis is false.

In this case, n=100, $\alpha(0.6)$,and $\beta(0.8)$ to calculate critical value. Here is the formula and code for the calculation.

$$\alpha(p) = \sum_{m \leq k \leq n} b(n,p,k) = \sum_{60 \leq k \leq 100} b(100,0.6,k)$$

```
n=100
p1=0.6
m1=n*p1
m=rep(0,n-m1)
p=rep(0,n-m1)
for(i in 0:n-m1){
  m[i]=i+m1
  p[i]=pbinom(100,n,p1)-pbinom(m[i]-1,n,p1)
}
a=data.frame(cbind(m,p))
head(a)
```

```
##    m         p
## 1 61 0.4620753
## 2 62 0.3821877
## 3 63 0.3068098
## 4 64 0.2386107
## 5 65 0.1794694
## 6 66 0.1303365
```

$$\beta(p) = 1 - \alpha(p) = \sum_{k \leq m} b(n,p,k) = \sum_{k \leq 80} b(100,0.8,k)$$

```
p2=0.8
m2=n*p2
m=rep(0,n-m2)
p=rep(0,n-m2)
```

```r
for(i in 0:n-m2){
  m[i]=m2-i
  p[i]=pbinom(m[i]-1,n,prob=p2)
}
b=data.frame(cbind(m,p))
head(b)
```

```
##    m          p
## 1 79 0.34596678
## 2 78 0.26106725
## 3 77 0.18908721
## 4 76 0.13135322
## 5 75 0.08747538
## 6 74 0.05583272
```

Critical value with $p<0.05$

```r
cva=a[which(a$p<0.05),1]
min(cva)
```

```
## [1] 69
```

```r
cvb=b[which(b$p<0.05),1]
max(cvb)
```

```
## [1] 73
```

```r
intersect(cva,cvb)
```

```
## [1] 69 70 71 72 73
```

In conclusion, when p is increasing, m will also increasing. As a result, type 1 error will be less likely and type 2 error will be more likely. Based on our work above, we can easily find that the critical value should be between 69 and 73 people cured with bottom and top at heights 0.05 and 0.95.

Part 2

In this section, we will replicate and explain the Figure 3.7 about the power curve.

```r
# X-axis: p value
pmin=0.4
p=seq(from=pmin,to=1,length=n)

# Y-axis: type 1 error rate
m=69
alpha1=rep(0,n)
for(i in 1:n){
  alpha1[i]=pbinom(n,n,p[i])-pbinom(m-1,n,prob=p[i])
}

m=73
alpha2=rep(0,n)
for(i in 1:n){
  alpha2[i]=pbinom(n,n,p[i])-pbinom(m-1,n,prob=p[i])
}

c=data.frame(cbind(p,alpha1,alpha2))
```

Plot

```r
d=ggplot(data=c,mapping=aes(x=p))+
  geom_line(mapping=aes(y=alpha1))+
  geom_line(mapping=aes(y=alpha2))+
  labs(x="p",y="a(p)",caption="Figure 3.7: The power curve.")

#Add a box
d+
  geom_segment(mapping=aes(x=0.6,xend=0.8,y=0.05,yend=0.05))+
  geom_segment(mapping=aes(x=0.6,xend=0.8,y=0.95,yend=0.95))+
  geom_segment(mapping=aes(x=0.6,xend=0.6,y=0.05,yend=0.95))+
  geom_segment(mapping=aes(x=0.8,xend=0.8,y=0.05,yend=0.95))
```

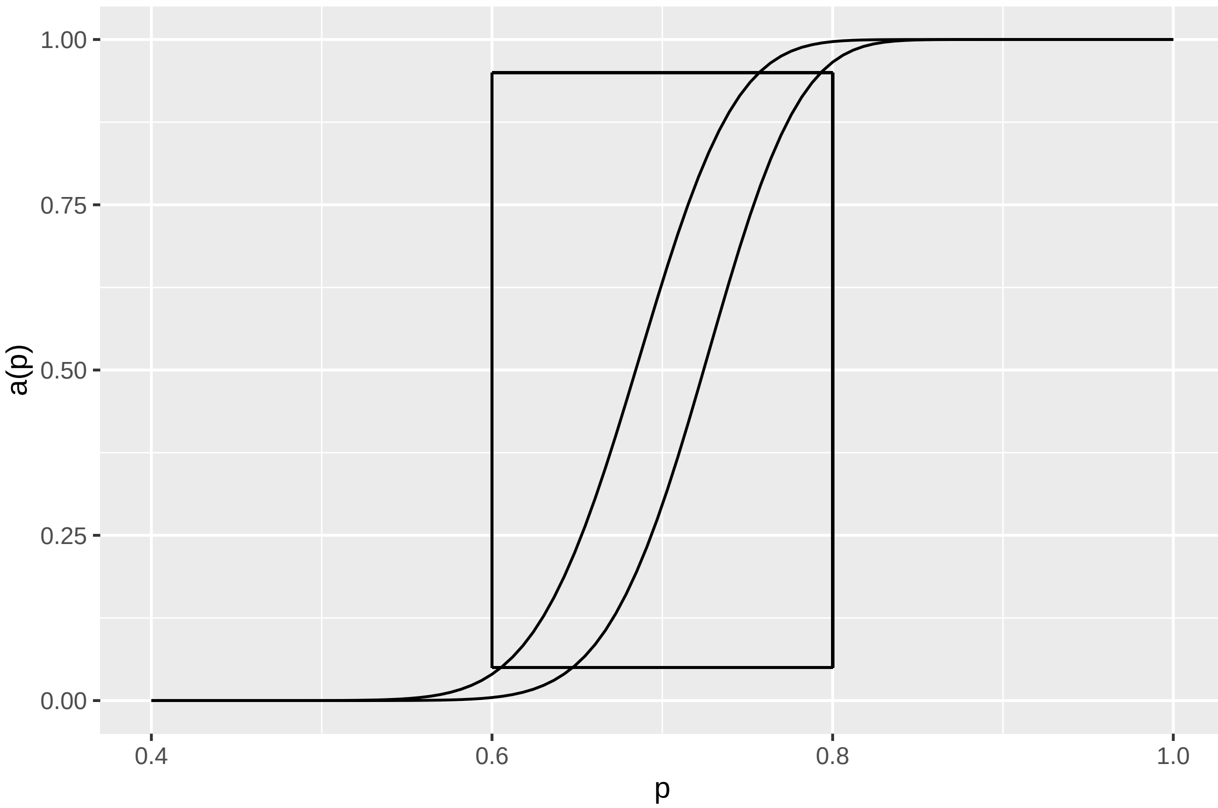

Figure 3.7: The power curve.

In conclusion, this is the PowerCurve for n=100, p ranging from 0.1 to 1. The left curve is when m = 69 and left one is when m = 73. AS we discuss before, when m is increasing, $\alpha(p)$ is also increasing, thats why the curve move to the right. Then is the box with length from 0.6 to 0.8 and height from 0.05 to 0.95. m value is satisfied only when the curve go through the top and bottom of the box at the same time. The two curve represent the limitation of m value satisfying the condition.