Assignment 1

Yuxin Zeng

2021/2/6

Critical Value

$\alpha(p)$ is the probability of type 1 error (to reject the null hypothesis when in fact it's true). Remember our null hypothesis is the additive has no appreciable effect. In this case, $p = 0.6$ so $m = 60$ people are cured, and the total number of people is $n = 100$. So the formula can be written as

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

.

```r
#a: Probability of type 1 error when p=0.6
m1=rep(0,times=40)
P1=rep(0,times=40)

for(i in 0:40){
  m1[i]=i+60
  P1[i]=pbinom(q=100,size=100,prob=0.6)-pbinom(q=m1[i]-1,size=100,prob=0.6)
}
a=data.frame(cbind(m1,P1))
```

$\beta(p)$ is the probability of type 2 error (to accept the null hypothesis when in fact it's false). In other words, the additive does have appreciable effect. In this case, $p = 0.8$ so $m = 80$ people are cured, and the total number of people is $n = 100$. So the formula can be written as

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

.

```r
#b: Probability of type 2 error when p=0.8
m2=rep(0,times=20)
P2=rep(0,times=20)

for(i in 0:20){
  m2[i]=80-i
  P2[i]=pbinom(q=m2[i]-1,size=100,prob=0.8)
}
b=data.frame(cbind(m2,P2))
```

We can see from both the formula and data that, increasing m makes a type 1 error less likely while a type 2 error more likely. So we want to choose some appropriate critical numbers to make the probabilities of each undesirable case less than 0.05.

```r
#Find critical value: both type of errors should less than 0.05
m_a=a[which(a$P1<0.05),1]
min(m_a)
```

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

```r
m_b=b[which(b$P2<0.05),1]
max(m_b)
```

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

```r
intersect(m_a,m_b)
```

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

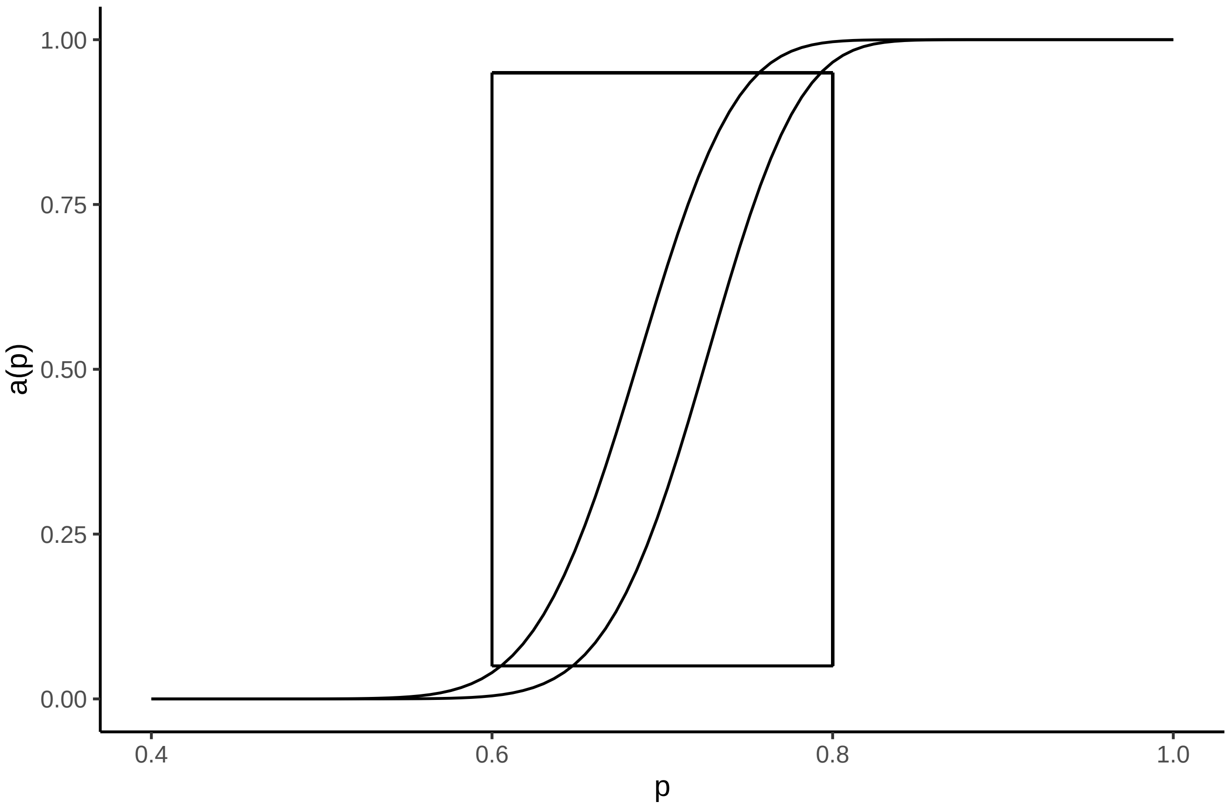
Therefore I reached authors' conclusion that the critical value should be between 69 and 73 people cured. (`min(m_a)` shows that m=69 is the smallest value for m that thwarts a type 1 error, while `max(m_b)` shows that m=73 is the largest which thwarts a type 2.)

Figure 3.7

For $n = 100$ and $m = 69, 73$, plot the function $\alpha(p)$, for p ranging from 0.4 to 1. (The left line represents the case where $m = 69$ and the right line represents the case $m = 73$. As m increases, the graph of $\alpha(p)$ moves to the right. To replicate the figure in textbook I didn't add a legend.)

```r
#X-axis is the probability that new aspirin is effective
p=seq(from=0.4,to=1,length=100)

#Y-axis is the probability of type 1 error
#When m=69
alpha1=rep(0,times=length((p)))
for(i in 1:100){
  alpha1[i]=pbinom(q=100,size=100,prob=p[i])-pbinom(q=68,size=100,prob=p[i])
}

#When m=73
alpha2=rep(0,times=length((p)))
for(i in 1:100){
  alpha2[i]=pbinom(q=100,size=100,prob=p[i])-pbinom(q=72,size=100,prob=p[i])
}

#Now data is ready
dt=data.frame(cbind(p,alpha1,alpha2))

#Plot
#The initial figure
f=ggplot(data=dt,mapping=aes(x=p))+
  theme_bw()+
  theme(panel.grid=element_blank(),panel.border=element_blank(),axis.line=element_line(colour="black"))
  geom_line(mapping=aes(y=alpha1))+
  geom_line(mapping=aes(y=alpha2))+
  labs(x="p",y="a(p)",title="Figure 3.7: The power curve.")+
  theme(plot.title=element_text(hjust=0.5,size=10))

#Add a box
f+
  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.



I included in the graph a box from 0.6 to 0.8, with bottom and top at heights 0.05 and 0.95. Then a value for m satisfies our requirements if and only if the graph of $\alpha(p)$ enters the box from the bottom, and leaves from the top (left bottom is the type 1 and right top is the type 2 criterion)