311assignment

Zhe Yu

2021/2/9

Replicate Figure 3.7

```
#calculate two alpha
p=seq(from=0.4,to=1,length=100)
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])
}
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])
}

#create the data frame
dt=data.frame(p,alpha1,alpha2)

#replicate the plot using ggplot
ggplot(data=dt,mapping=aes(x=p))+
  theme_bw()+
  geom_segment(mapping=aes(x=0.6,xend=0.8,y=0.05,yend=0.05),color="grey")+##add box
  geom_segment(mapping=aes(x=0.6,xend=0.8,y=0.95,yend=0.95),color="grey")+
  geom_segment(mapping=aes(x=0.6,xend=0.6,y=0.05,yend=0.95),color="grey")+
  geom_segment(mapping=aes(x=0.8,xend=0.8,y=0.05,yend=0.95),color="grey")+
  geom_line(mapping=aes(y=alpha1))+
  geom_line(mapping=aes(y=alpha2))+
  labs(x=" ",y=" ",caption="Figure 3.7: The power curve.")+
  theme(plot.caption=element_text(hjust=0.5,size=10))+

  theme(panel.grid=element_blank(),panel.border=element_blank(),axis.line=element_line(colour=
"black"))
```
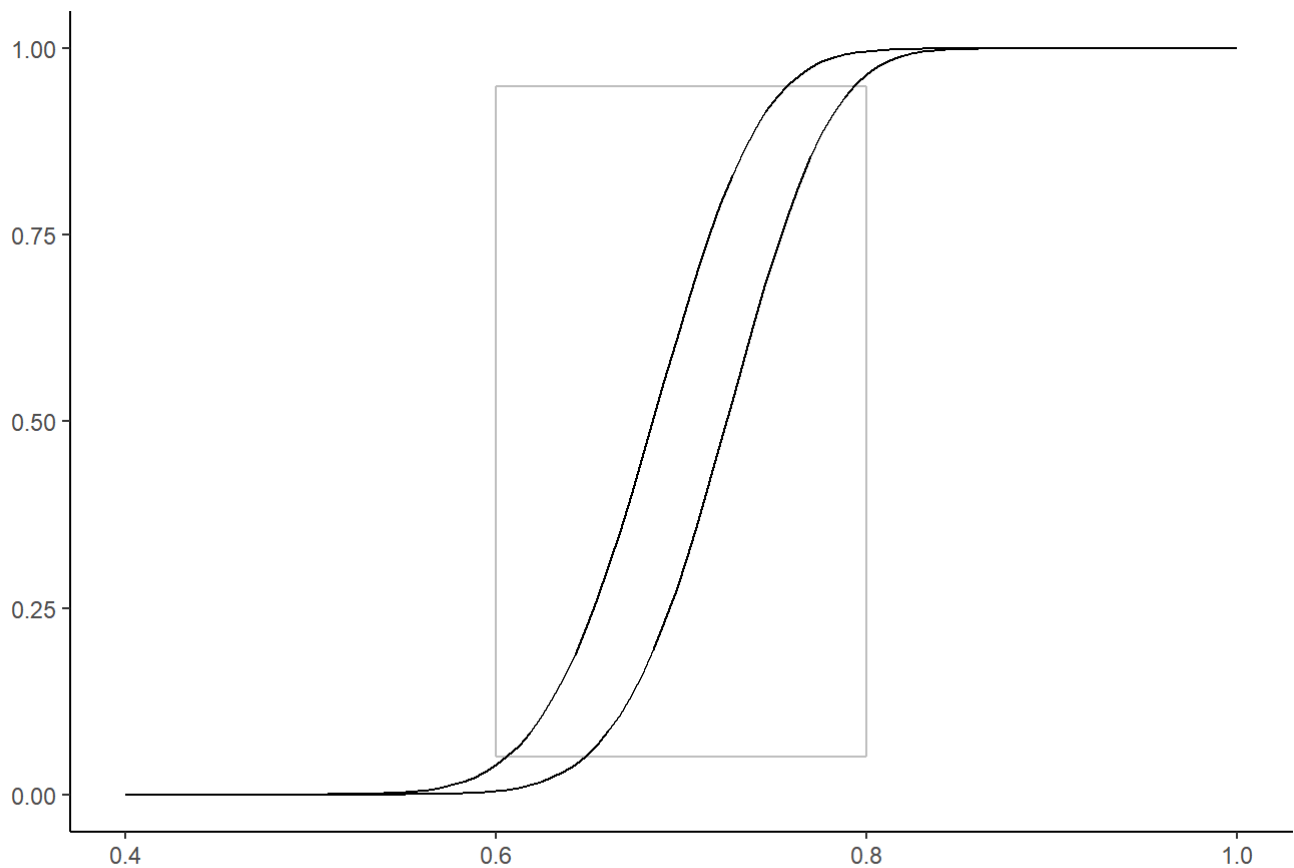
Figure 3.7: The power curve.

Explanation

$\alpha(p)$ is the probability of type 1 error. $\beta(p)$ is the probability of type 2 error. Then choosing m well below np = .8n will increase $\alpha(.8)$, since now $\alpha(.8)$ is all but the lower tail of a binomial distribution. Indeed, if we put $\beta(.8) = 1 - \alpha(.8)$, then $\beta(.8)$ gives us the probability of a type 2 error, and so decreasing m makes a type 2 error less likely. So we should make the probabilities of each type error less than 0.05.

```
#Probability of type 1 error when p=0.6
m.a=rep(0,times=40)
p.a=rep(0,times=40)
for(i in 0:40){
  m.a[i]=i+60
  p.a[i]=pbinom(q=100,size=100,prob=0.6)-pbinom(q=m.a[i]-1,size=100,prob=0.6)
}
a=data.frame(m.a,p.a)
```

```
#Probability of type 2 error when p=0.8
m.b=rep(0,times=20)
p.b=rep(0,times=20)
for(i in 0:20){
  m.b[i]=80-i
  p.b[i]=pbinom(q=m.b[i]-1,size=100,prob=0.8)
}
b=data.frame(m.b,p.b)
```

```r
#both type of errors should less than 0.05
lower <- min(a[which(a$p.a<0.05),1])
upper <- max(b[which(b$p.b<0.05),1])
cat("So the interval is","(",lower,",",upper,")")
```

```
## So the interval is ( 69 , 73 )
```

```r
#both type of errors should less than 0.05
lower <- min(a[which(a$p.a<0.05),1])
upper <- max(b[which(b$p.b<0.05),1])
cat("So the interval is","(",lower,",",upper,")")
```