

MA677 Assignment 1

Yuelin Jiang

2/9/2021

Assignment: Due – Wednesday, February 10

In Chapter 2 of G&S, Example 3.11 uses the binomial distribution to demonstrate the logic of hypothesis testing with a simple example.

The intuitive simplicity of the example and the availability of tools in R for calculations and plotting distributions encourages you to explore hypothesis testing by reproducing the results presented in the book.

In the final paragraph of the example on page 102, the authors write, "A few experiments have shown us that ...". They report the results but don't show you their experiments. Similarly, they show you Figure 3.7, but not how it was produced.

Your assignment is to fill in these gaps. Produce an explanation of the example showing how the authors reached their conclusion that the critical value should be between 69 and 73 people cured. Replicate and explain Figure 3.7.

Reproducing the graph in Figure 3.7

```
n = 100
# make a function to calculate alpha(p) point
alpha <- function(m, n, p){
  sum(dbinom(m:n, n, p, log=FALSE))
}

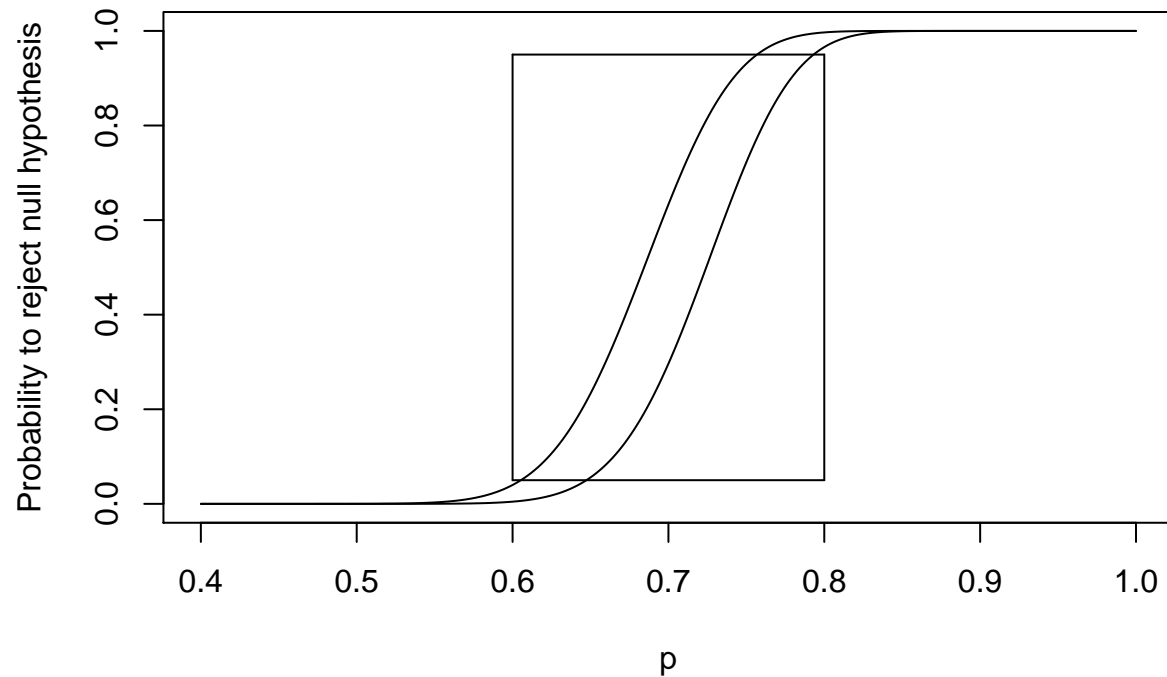
# make iterations for a for loop and x axis
ite <- seq(0.4, 1, by = 0.002)

# function to get a vector of alpha(p) when p takes from 0.4 to 1 given m
fy <- function(m){
  y <- NULL
  for (i in ite) {
    al <- alpha(m, n, i)
    y <- append(y, al)
  }
  y
}

# plotting two lines when m= 69 and 73
plot(ite, fy(69), type = "l", xlab = "p", ylab = "Probability to reject null hypothesis")
lines(ite, fy(73), type = "l")

# plot significant levels from 0.05 to 0.95
segments(0.6, 0.05, 0.6, 0.95)
segments(0.6, 0.95, 0.8, 0.95)
```

```
segments(0.8, 0.05, 0.8, 0.95)
segments(0.6, 0.05, 0.8, 0.05)
```



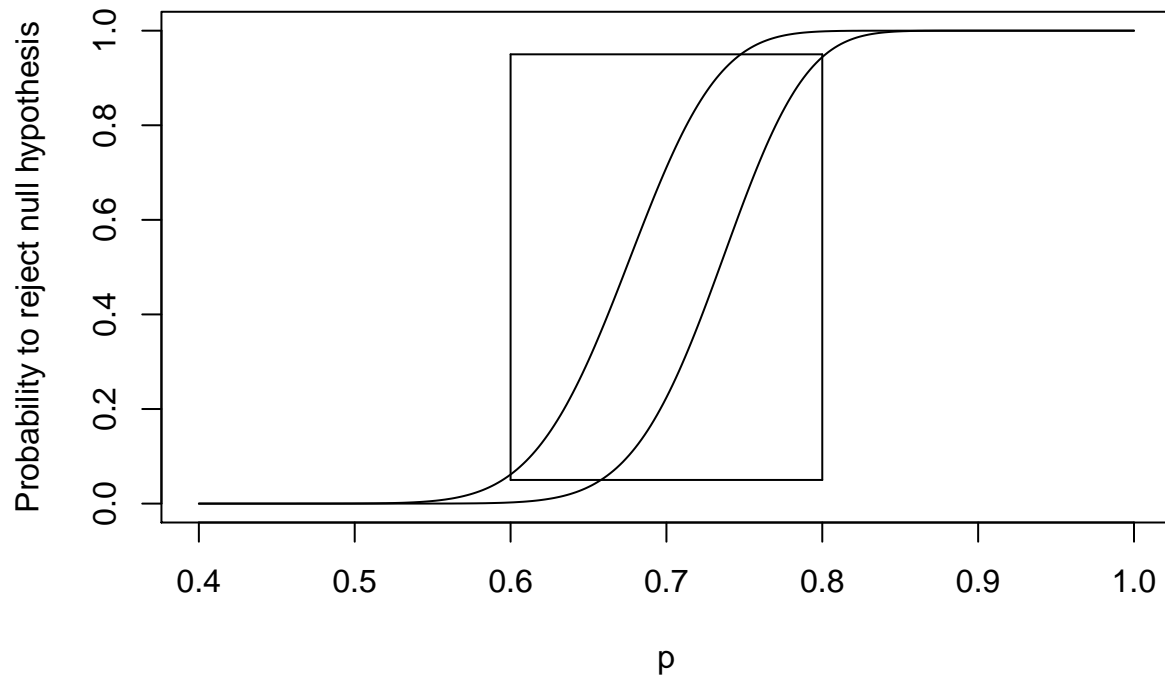
Explaining Critical Value

We conclude critical value to be between 69 and 73 inclusive. Graphically we have to choose critical values m which make our alpha function enters from the bottom of the box and exits from the top of the box. Within the box is our 95% significant range that we will reject the null hypothesis. Any m that make an alpha function enters or exits outside the box increases our risks of Type I and Type II errors. When $p=0.6$ we want the probability to reject the null hypothesis to be low to avoid Type I error, and when $p=0.8$ we want the probability to reject the null hypothesis to be high to avoid Type II error.

For example, when $m = 68$ and $m = 74$ are drawn, as shown below:

```
# plotting critical values that give Type I and Type II errors
plot(ite, fy(68), type = "l", xlab = "p", ylab = "Probability to reject null hypothesis")
lines(ite, fy(74), type = "l")
```

```
# plot significant levels from 0.05 to 0.95
segments(0.6, 0.05, 0.6, 0.95)
segments(0.6, 0.95, 0.8, 0.95)
segments(0.8, 0.05, 0.8, 0.95)
segments(0.6, 0.05, 0.8, 0.05)
```



When $m = 68$, alpha function—the probability to reject null hypothesis at $p=0.6$ is higher than 0.05. That's not a good sign. We have higher risk of committing a Type I error. When $m = 74$, our alpha function—the probability to accept null hypothesis is lower than 0.95. We have higher risk of committing a Type II error. So we conclude critical values have to be between 69 to 73.