

# MAST20005/MAST90058: Week 7 Lab

**Goals:** (i) Basic hypothesis testing for proportions; (ii) Type I and type II error; (iii) Power.

**Data:** 20 students counted the number of yellow lollies and the total number of lollies in a 48.1 gram packet. The data file, `candies.txt`, can be obtained from the shared folder in the computer labs or from the LMS.

## 1 Type I and Type II errors

Let  $X \sim \text{Bi}(1, p)$  and let  $X_1, \dots, X_{10}$  be a random sample of size 10. Consider a test of  $H_0: p = 0.5$  against  $H_1: p = 0.25$ . Let  $Y = \sum_{i=1}^{10} X_i$ . Define the critical region as  $y < 3.5$ .

1. Find the value of  $\alpha$ , the probability of a Type I error. Do not use a normal approximation (use the function `pbinom()`).

```
pbinom(3.5, 10, 0.5)

## [1] 0.171875

pbinom(3, 10, 0.5) # same answer as above! (why?)

## [1] 0.171875
```

2. Find the value of  $\beta$ , the probability of a Type II error. Do not use a normal approximation.

```
1 - pbinom(3.5, 10, 0.25)

## [1] 0.2241249
```

3. Next, carry out a simulation to estimate the Type I error of the test. Simulate 200 observations on  $Y$  when  $p = 0.5$  and then find the proportion of cases when  $H_0$  was rejected. Is this close to  $\alpha$ ?

```
T <- rbinom(200, 10, 0.5) # simulate under H0
alpha <- sum(T < 3.5) / length(T) # estimate
alpha

## [1] 0.19
```

4. Simulate 200 observations on  $Y$  when  $p = 0.25$ . Find the proportion of cases when  $H_0$  was not rejected. Is this close to  $\beta$ ?

```
T1 <- rbinom(200, 10, 0.25) # simulate under H1
beta <- sum(T1 > 3.5) / length(T1) # estimate
beta

## [1] 0.27
```

5. Estimates from simulations are affected by sampling error. The last two questions may be answered more rigorously by computing confidence intervals for your simulation-based estimates in the usual way (note that estimated probabilities for  $\alpha$  and  $\beta$  are just sample proportions):

```
alpha + c(-1, 1) * 1.96 * sqrt(alpha * (1 - alpha) / 200) # CI for alpha
## [1] 0.1356299 0.2443701

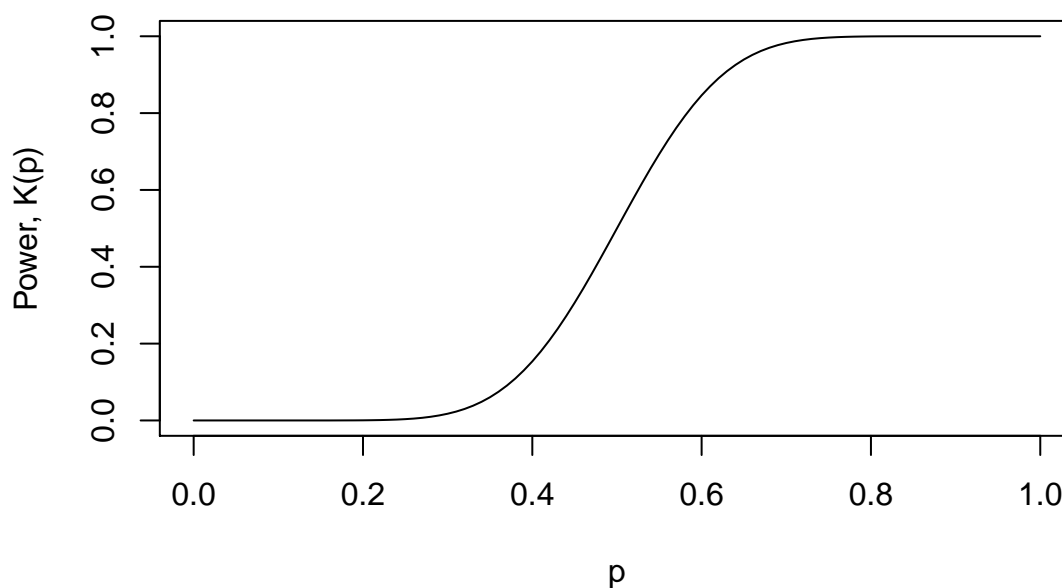
beta + c(-1, 1) * 1.96 * sqrt(beta * (1 - beta) / 200) # CI for beta
## [1] 0.2084704 0.3315296
```

## 2 Power

Let  $p$  be the probability that a tennis player's first serve is good. The player takes lessons to increase  $p$ . After the lessons she wishes to test the null hypothesis  $H_0: p = 0.4$  against the alternative  $H_1: p > 0.4$ . Let  $y$  be the number out of  $n = 25$  serves that are good, and let the critical region be defined by  $y \geq 13$ .

1. Let the power function be  $K(p) = \Pr(Y \geq 13 \mid p)$ . Graph this function for  $0 < p < 1$ .

```
K1 <- function(p)
  1 - pbinom(12, 25, p)
p <- seq(0, 1, 0.01)
K <- K1(p)
plot(p, K, type = "l", ylab = "Power, K(p)")
```



2. Find the value of  $\alpha = K(0.4)$ .

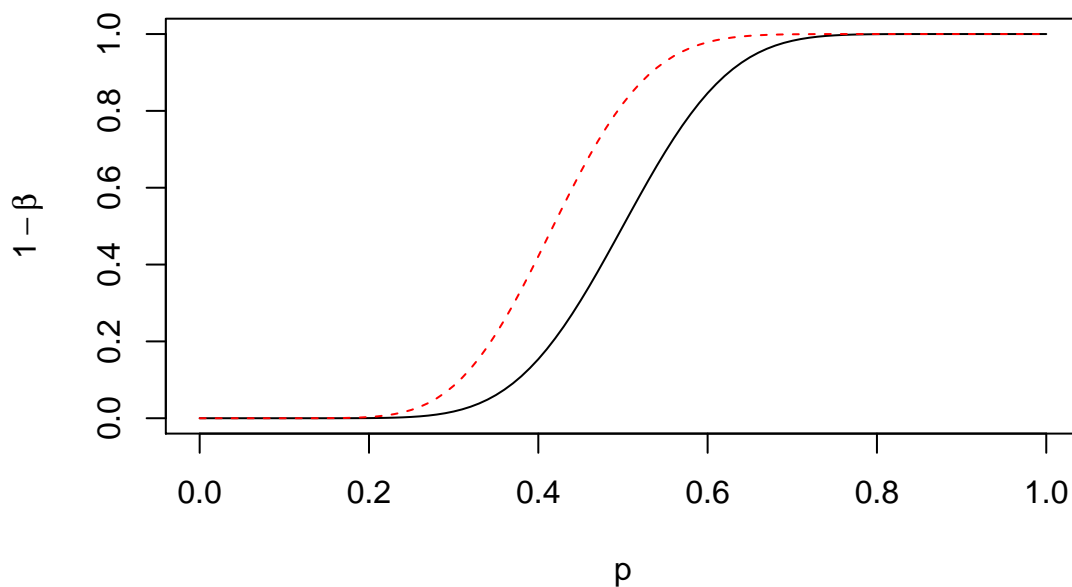
```
K1(0.4)
## [1] 0.1537678
```

3. Find the value of  $\beta$  when  $p = 0.6$ , ( $\beta = 1 - K(0.6)$ )

```
1 - K1(0.6)
## [1] 0.1537678
```

4. What happens to power when the sample size increases? Suppose the player carries out  $n = 30$  serves:

```
K2 <- function(p)
  1 - pbinom(12, 30, p)
curve(K1, from = 0, to = 1, xlab = "p", ylab = expression(1 - beta))
curve(K2, from = 0, to = 1, add = TRUE, col = 2, lty = 2)
```



### 3 Lollies data

Let  $p$  be the proportion of yellow lollies in a packet of mixed colours. It is claimed that  $p = 0.2$ .

1. Define a test statistic and an approximate critical region with a significance level of  $\alpha = 0.05$  to test  $H_0: p = 0.2$  against  $H_1: p \neq 0.2$ .

Reject  $H_0$  if:

$$|z| = \frac{|\hat{p} - 0.2|}{\sqrt{0.2 \times 0.8/n}} > 1.96.$$

2. To perform the test, each of 20 students counted the number of yellow lollies and the total number of lollies in a 48.1 gram packet. The results were:

$y$	$n$	$y$	$n$	$y$	$n$	$y$	$n$
8	56	5	54	10	57	11	57
13	55	14	56	8	59	6	54
12	58	15	57	10	54	7	58
13	56	11	54	11	55	12	58
14	57	13	55	12	56	14	58

```
data <- read.table("candies.txt", header = TRUE) # load the data
```

3. If each student made a test of  $H_0: p = 0.2$  at the 5% level of significance, what proportion of students rejected the null hypothesis?

```
y <- data[, 1]
n <- data[, 2]
p <- y / n
z <- abs(p - 0.2) / sqrt(0.2 * 0.8 / n)
sum(z > 1.96) / length(z) # proportion

## [1] 0.05

which(z > 1.96) # this shows *which* students rejected the null

## [1] 6
```

4. If the null hypothesis were true, what proportion of students do you expect to reject the null hypothesis at the 5% level of significance?

Approximately  $1/20 = 0.05$ .

5. For each of the 20 ratios in part 3, a 95% confidence interval can be constructed. What proportion of these intervals contains  $p = 0.2$ ?

```
b1 <- p - 1.96 * sqrt(p * (1 - p) / n)
b2 <- p + 1.96 * sqrt(p * (1 - p) / n)
sum((b1 <= 0.2) & (0.2 <= b2)) / length(p)

## [1] 0.9
```

6. If the 20 results are pooled do we reject  $H_0: p = 0.2$ ?

```

x <- sum(y)
N <- sum(n)
prop.test(x, N, p = 0.2, alternative = "two.sided")

##
## 1-sample proportions test with continuity correction
##
## data:  x out of N, null probability 0.2
## X-squared = 0.15619, df = 1, p-value = 0.6927
## alternative hypothesis: true p is not equal to 0.2
## 95 percent confidence interval:
##  0.1723170 0.2194814
## sample estimates:
##           p
## 0.1948399

```

We cannot reject the null hypothesis. Therefore, from the data we do not have enough evidence to reject the claim that 20% of lollies in the packet are yellow.

## 4 Comparing two populations

Invadopodia are actin-rich protrusions of the plasma membrane that are associated with degradation of the extracellular matrix in cancer invasiveness and metastasis. Dr Sloan treated one sample of cells using a drug which may reduce the development of invadopodia, while another sample of cells received a neutral treatment. Using a microscope she counted the number of cells developing invadopodia in each tissue portion. In the treatment group she found that 25 out of 351 cells developed invadopodia, while in the neutral group 50 out of 389 cells developed invadopodia.

Carry out a two-sample z-test at the  $\alpha = 0.05$  level of significance to determine whether the treatment is effective in reducing the number of invadopodia. The following tests the null hypothesis  $H_0: p_1 = p_2$  versus the two-sided alternative  $H_1: p_1 \neq p_2$ :

```

x <- c( 25,  50)      # successes
n <- c(351, 389)      # sample sizes
prop.test(x, n, alternative = "two.sided")

##
## 2-sample test for equality of proportions with continuity
## correction
##
## data:  x out of n
## X-squared = 6.0393, df = 1, p-value = 0.01399
## alternative hypothesis: two.sided
## 95 percent confidence interval:
## -0.10279971 -0.01181956
## sample estimates:
##      prop 1      prop 2
## 0.07122507 0.12853470

```

The output above shows the value of the chi-square statistic defined as  $Z^2$ . The observed value of the  $Z$ -statistic is  $|z_{obs}| = \sqrt{6.0393} = 2.4575$ . When  $\alpha = 0.05$ , the rejection region for this test is  $|z| > 1.96$ . Therefore, we reject  $H_0$  and conclude that the treatment is effective in reducing invadopodia development.

## Exercises

1. In Section 1, the simulations (questions 3 and 4) did not give answers that were particularly close to the true values (questions 1 and 2). Furthermore, the confidence intervals for the simulations (question 5) were quite wide, indicating that the simulation estimates are not very precise.
  - (a) Explain why the simulations shown here are not very useful.
  - (b) Improve the simulations and repeat them, showing that they indeed can give accurate and precise estimates.
2.
  - (a) Do question 3 from the tutorial problems.
  - (b) Draw a power curve for a significance level of 0.05, for all possible values of  $p$ .
3. Do question 4 from the tutorial problems.
4. A political party commissions an election poll, asking participants whether they will vote for them rather than a rival party. They will survey 900 people. Let  $p$  the true proportion of people who intend to vote for this party. Upon receiving the responses, they will carry out a hypothesis test of  $H_0: p = 0.5$  against  $H_1: p \geq 0.5$ , using a significance level of 0.1.
  - (a) What is the test statistic and critical region?
  - (b) What is the power when  $p = 0.52$ ?
  - (c) Draw a power curve for  $0.5 \leq p \leq 0.6$ .
  - (d) The poll was run and 465 people responded in favour of the party. Carry out the test and state a conclusion.
  - (e) What can the party do to get a more conclusive result?