Stats. Inference Homework #2

Naeem Chowdhury

2/13/2020

9.66 Exploring Type II Errors

Refer to the web app from Activity 2 at the end of this section, now assuming that we are using the two-sided test $H_0 : p = 0.33$ against $H_a : p \neq 0.33$.

a. Explain the effect of increasing the sample size on the probability of a Type II error when the true $p = 0.50$.

Solution a.

As sample size increases, the probability of a Type II error decreases. This is because as sample size increases, value of the null hypothesis p_0 falls increasingly more standard deviations from the mean.

b. Use the app to find the sample size needed to achieve a power of at least 90% when truly $p = 0.5$.

Solution b.

A sample size of 84 is suffcient to achieve a power of at least 90%.

c. For a fixed sample size, do you think the probability of a Type II error will increase or decrease when the true p is 0.40 instead of 0.50? Check your answer with the app.

Solution c.

I think that for a fixed sample size, the probability of a Type II error will increase when the true p is 0.40 instead of 0.50. This is because the center of the H_a distribution will be closer to the value p_0 of the null hypothesis.

Indeed, it is true!

d. Does the power increase or decrease when the significance level is 0.10 instead of 0.05? Check your answer with the app.

Solution d.

The power increases when the significance level is 0.10 instead of 0.05. This is because by increasing our significance level, we are decreasing the radius around p_0 under which we compute the area under the H_a distribution. This decreases the Type II error probability, increasing the Power.

Problem 2.

Presume you've developed a skin cancer treatment and you were granted permission to test it out on patients. You would like to test if its accuracy differs from the golden standard method which has 28% cure rate. In particular, you'd want the ability to correctly detect a difference of 10%.

a. Formulate the hypotheses for the one-sample proportion test.

Solution a.

Let p be the proportion of patients cured under the skin cancer treatment that I developed.

Then the hypotheses are...

$$H_0 : p = 0.28.$$

$$H_a : p \neq -0.28.$$

b. Interpret the statement: "At a 0.05 significance level, the significance test will have 0.77 power when detecting a difference of 10%."

Solution b.

With a confidence level of 95%, we can detect a 10% difference between our drug and the leading drug 77% of the time.

c. Presume we witness your treatment's results for 100 patients. Obtain the power of one-sample proportion test at $\alpha = 0.05$ significance level when detecting a difference of 10%.

```
n <- 100
p0 <- 0.28
pA <- 0.38
alpha <- 0.05

reject <- qnorm(1-alpha/2, mean=p0, sd= sqrt(p0*(1-p0)/n))
# Calculating the Type II error, the probability that we fail to observe a difference when there actual
type_II <- pnorm(reject, mean = pA, sd = sqrt(p0*(1-p0)/n)) - pnorm(-reject, mean = pA, sd = sqrt(p0*(1
# Power = 1 - type_II
paste("The power of the test is ", 1 - type_II, sep = "")
```

```
## [1] "The power of the test is 0.605347428379385"
```

d. Presume we witness your treatment's results for 100 patients. Obtain the α significance level needed for your test to have the power of 0.85 when detecting a difference of 10%.

```
n <- 100
p0 <- 0.28
pA <- 0.38
alpha <- 0.235

# Quantiles
reject <- qnorm(1-alpha/2, mean=p0, sd= sqrt(p0*(1-p0)/n))

type_II <- pnorm(reject, mean = pA, sd = sqrt(p0*(1-p0)/n)) - pnorm(-reject, mean = pA, sd = sqrt(p0*(1

paste("The power of the test is ", 1 - type_II, " with alpha level of ", alpha, sep = "")
```

```
## [1] "The power of the test is 0.850737054080076 with alpha level of 0.235"
```

e. What # of patients is needed for your test to have the power of 0.85 when detecting a difference of 10%?

Not Sure how to do this one.

f. In this case, do you think it is more important for your test to have lower significance level or higher power? Explain.

g. Provided that you don't have the resources to recruit more than $n = 100$ patients, what can you do in order to increase the power of your test for detecting 10% difference.