Oct 17, 2022

Theoretical Hypothesis Testing

Week 3: Power analysis

Scenario

There are two possible decisions about the null hypothesis:

- Reject
- Do not reject

There are two possible situations of the null hypothesis:

- True null
- False null

Scenario

	Null hypothesis is TRUE	**Null hypothesis is FALSE**
Reject null hypothesis	Type I Error (False positive)	Correct outcome! (True positive)
Fail to reject null hypothesis	Correct outcome! (True negative)	Type II Error (False negative)

Type I error

Type I error is the rejection of the true null hypothesis. (α-type error)

Type II error

Type II error is the failure to reject the false null hypothesis. (β-type error)

Power of the test

The ability of a test to reject the false null hypothesis. $(1 - \beta)$

Neyman-Pearson: A powerful test

A powerful test is defined as that it needs to control of the Type I error probability according to the significance level and its power must be high as possible.

Power analysis

- To show the confidence of the study.
- **To calculate the optimal sample size for a specific power value.**
- Compare the performance of the competitor tests.
- Making sure of performance of a newly developed test.

Elements of power analysis

- **Effect size**
- Sample size
- Significance
- Statistical power (power of the test)

Effect Size

The magnitude of a result presented in the population. It is calculated using a specific statistical measure:

- **Pearson's correlation coefficient** for the relationship between variables
- **Cohen's d** for the difference between samples.

Cohen's d

- Standardized effect size for measuring the difference between sample means.
- It can be used to measure the effect size in *t-test* and *ANOVA*.

$$d = \frac{\bar{x}_A - \bar{x}_B}{\sqrt{\frac{s_A^2 + s_B^2}{2}}}$$

- It can be generalized for k samples $(k > 2)$.
- In the interpretation of the significant differences between populations, it can be used to report the size of the difference.

How to calculate these metrics?

- using ready-to-use function in R
- conducting a simulation study

Application with ready-to-use R functions

- You can use {pwr} package in R for power analysis:

```
install.packages(“pwr”)

library(pwr)

# pwr.t.test() is be used for power analysis of one-sample t-test

# pwr.norm.test() for Z-test with known variance

# pwr.t2n.test() for two-sample t-test
```

Application with ready-to-use R functions

```
library(pwr)

pwr.t.test(n = 100,                          # sample size
           d = 0.5,                          # cohen's d effect size
           sig.level = 0.05,                 # significance level (Type I error)
           type = "one.sample",              # type of t-test: one, two, paired
           alternative = "two.sided")        # alternative hypothesis
```

Application with ready-to-use R functions

```
library(pwr)

pwr.t.test(power = 0.95,                    # power of the test
           d = 0.5,                         # cohen's d effect size
           sig.level = 0.05,                # significance level (Type I error)
           type = "one.sample",             # type of t-test: one, two, paired
           alternative = "two.sided")       # alternative hypothesis
```

Application with ready-to-use R functions

In the pwr.t.test() function calculates the argument that you did not specify.

Application by using simulation

It is be possible to generate random observation from any distribution that you have the distribution function by **inverse transformation method**.

Algorithm of inverse transformation method:

1. Generate $U \sim U(0,1)$

2. Let $X = F_X^{-1}(U)$

Then $X \sim F_X$

Application by using simulation

Thanks to ready-to-use functions in R, we can generate random samples follow any distribution.

```
pval <- numeric(10000)

for(i in 1:10000){

   sample1 <- rnorm(100, 2.5, 4)  # generating random sample from normal dist. (n = 100, mu =
   2.5, sd = 4)

   sample2 <- rnorm(100, 2, 4)

   pval[i] <- t.test(sample1, sample2)$p.value

}

mean(pval < 0.05)
```