Module - 4

ALY 6010 Probability Theory and Statistics

RAHUL AVINASH JADHAV

Northeastern University

College of Professional Studies, Northeastern University, Boston, MA 02115

Contact: jadhav.ra@northeastern.edu

Submitted to Professor: Prof. Roy Wada

Date of submission: 12/05/2021

Introduction

In this assignment we are going to perform Unpaired and paired Two sample t test using the dataset “small sample PSID (4 variables)”. The dataset consists of 60 observations with 4 variables such as Id, year, Gender and Wage2.

We will also use the parameter confidence level and see what difference does confidence level do to the test if it is changed from 0.1 to 0.001.

Analysis

Table 1: Welch Two Sample t-test on Wages based on Gender with alternative as Two sided and count as 40.04(Unpaired testing).

	Significance	estimate	mean Wage of Female	mean wage of Male	statistic	p.value	conf.low	conf.high
1	0.1	1369.847	3484.452	2114.605	4.8	0	598.16	2141.53
2	0.001	1369.847	3484.452	2114.605	4.8	0	356.65	2383.04

1. **Significance 0.1 and Significance 0.001 with confidence level = 0.99 and 0.999 respectively**

   H0 = null hypothesis = true difference in means is equal to 0
   H1 = Alternative hypothesis = true difference in means is not equal to 0

   In both the cases of significance level, the P value is lesser than or equal to significance level, the null hypothesis cannot be rejected.

2. **Findings**
   Change in Significance level makes a huge difference in the confidence level. As you can see when the significance level was 0.1, the Confidence low level was 598.16 and the confidence high level was 2141.53, and when the significance level was 0.001, the confidence low level was 356.65 and the confidence high level was 2383.04.

Table 2: Two Sample t-test on Wages based on years with alternative as Two sided and count as 29(paired testing).

	Significance	mean_of_differences	statistic	p.value	conf.low	conf.high
1	0.1	-35.09	-0.49	0.62	-230.66	160.47
2	0.001	-35.09	-0.49	0.62	-294.73	224.54

1. **Significance 0.1 and Significance 0.001 with confidence level = 0.99 and 0.999 respectively**

   H0 = null hypothesis = true difference in means is equal to 0
   H1 = Alternative hypothesis = true difference in means is not equal to 0

   In both the cases of significance level, the P value is greater than or equal to significance level, the null hypothesis is not rejected.

2. **Findings**
   Change in Significance level makes a huge difference in the confidence level. As you can see when the significance level was 0.1, the Confidence low level was -230.66 and the confidence high level was 160.47, and when the significance level was 0.001, the confidence low level was -294.73 and the confidence high level was 224.54.

Summary

- From this assignment we can see that if the significance level is changed then there is huge difference in the confidence low and confidence high level.
- We also learned how to perform paired and unpaired t test.

Bibliography

- Kabacoff, R. (2011). *R in action: Data analysis and graphics with R*. Manning.

- *S.3.2 Hypothesis Testing (P-Value Approach) | STAT ONLINE*. (n.d.). PennState: Statistics Online Courses. https://online.stat.psu.edu/statprogram/reviews/statistical-concepts/hypothesis-testing/p-value-approach