Statistics on further studies of students in IIT H

MA4240 Applied Statistics

Shreyas Wankhede, Pradeep Mundlik, Akshitha Kola, Dhatri Reddy, Mouktika Cherukupalli, Avinash Malothu, Sumeeth Kumar, Saathvika Marri

IIT Hyderabad

April 27, 2023

Outline

Introduction

This is the introduction slide.

Data visualization

hypothesis Testing

case 1:comparing CGPA of students who are willing to pursue higher studies with students who dont want to pursue

We assume our null hypothesis to be that average CGPA of students willing to go for higher studies is greater than those who dont want to.
Let $\alpha = 0.05$.

$\bar{x_1}$=sample mean of CGPA of people willing to go for higher studies

$\bar{x_2}$=sample mean of CGPA of people who dont want higher studies

s_1^2=sample standard deviation of CGPA of people willing for higher studies

s_2^2=sample standard deviation of CGPA of people who dont want
For Hypothesis Testing we make the following statements-

$$H_0 = \mu_1 - \mu_2 \geq 0$$
$$H_a = \mu 1 - \mu 2 < 0$$

case 1: case 1 continued

information:

$$\bar{x_1} = 8.4447 \tag{1}$$
$$\bar{x_2} = 8.3919 \tag{2}$$
$$s_1^2 = 0.7694 \tag{3}$$
$$s_2^2 = 0.5672 \tag{4}$$
$$n_1 = 88 \tag{5}$$
$$n_2 = 26 \tag{6}$$

since $\frac{s_1^2}{s_2^2} < 4$,we can assume the population variances would be equal.
the degrees of freedom,

$$dof = n_1 + n_2 - 2 = 112$$

case 1 continued

the pooled varience will be:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = 0.5317 \tag{7}$$

the test statistic is given by:

$$t = \frac{\bar{x_1} - \bar{x_2} - 0}{s_p \sqrt{\frac{n_2 + n_1}{n_1 n_2}}} = 0.3244 \tag{8}$$

Using the rejection region approach, we reject H_0 if $t_{0.05,112} \geq -t$ where $t_{0.05,112} = -1.6586$. we have enough statistical evidence to reject null hypothesis since observed t is lesser than 1.6586

Case 4: Hypothesized proportion testing if there is enough evidence that the proportions of people opting for masters,MBA,Phd are not all equal

Sample data :

Masters	MBA	PhD	Total
50	24	13	87

Let P_{Ms}, P_{MBA}, P_{PhD} denote proportions of students willing to pursue Masters, MBA, PhD for higher studies

$H_0 : P_{Ms} = P_{MBA} = P_{PhD} = \frac{1}{3}$ $H_a :$ atleast one $P \neq \frac{1}{3}$ $\alpha = 0.05$

Also,

$$E = \frac{1}{3}x87 = 29 \tag{9}$$

and,

$$\chi^2 = \sum \frac{(O-E)^2}{E} \tag{10}$$

Case 4: Case 4 continued

$$\begin{aligned}\chi^2 &= \frac{(50-29)^2}{29} + \frac{(24-29)^2}{29} + \frac{(13-29)^2}{29} \\ &= 15.2 + 0.862 + 8.827 \\ &= 24.889\end{aligned} \tag{11}$$

At $df = 3 - 1 = 2$, p value $= 0.0001$
p value $< \alpha = 0.05$
Hence there is enough evidence that population proportions are not all equal.