Hypothesis testing

Test 1: sample test to test the mean age of males going to High school.

This sample testing is to test whether the statement is true that the mean age of males in 22 or older Years old going to High school or not, Assume the confidence level to be 95%, significant level, α = 0.05, Let the population mean be μ.

So, H0: μ = 21

$$z = \frac{\overline{x} - \mu}{s/\sqrt{n}}$$

H1: μ > 21

α = 0.05

test statistic value, z = 6.073614

After using Rstudio :

Critical value, c.v = z0.05 =   1.644854

Sample size, n = 76

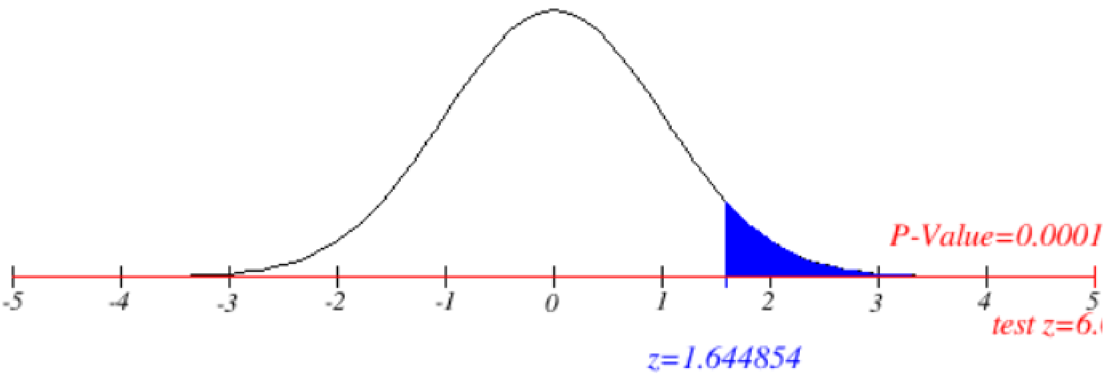
Sample Standard deviation, σ = 6.459102

Sample mean, $\overline{X}$ = 25.5

Decision making:

Since the test statistic value (6.073614) is greater than the critical value (1.644854)

which falls within the critical region. Hence, we reject the null hypothesis.



Hypothesis testing

Test 2: sample test to test the mean age of females going to Elementary school.

This sample testing is to test whether the statement is true that the mean age of females in 25 or older Years old going to Elementary school or not, Assume the confidence level to be 98%, significant level, α = 0.02, Let the population mean be μ.

So, H0: μ = 24

$$z = \frac{\overline{x} - \mu}{s / \sqrt{n}}$$

H1: μ > 24

α = 0.02

test statistic value, z = 3.020303

After using Rstudio :

P value, P-v = 0.9987374

Sample size, n = 24

Sample Standard deviation, σ = 6.150144

Sample mean, $\overline{X}$ = 27.79167

Decision making:

Since the p-value = 0.9987374 is smaller than the significant value

α = 0.02, Hence, the null hypothesis is rejected.

Correlation Analysis

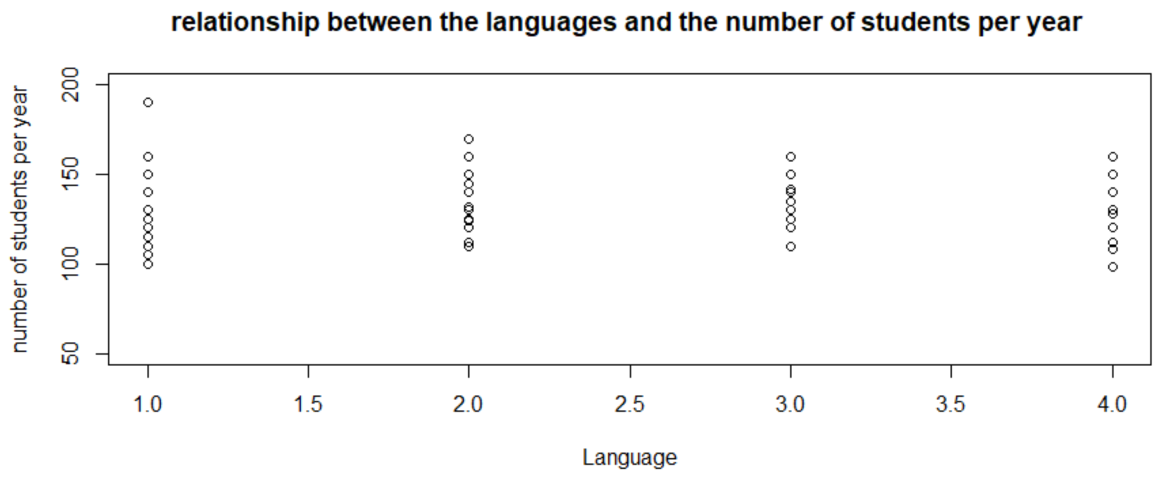
Test 3: Correlation Analysis to explain the relationship between the languages and the number of students per year in years 2000 - 2021.

This test sample is to measure the relationship between the languages and the number of students per year in years 2000 – 2021,

Assume the confidence level to be 99%, significant level, $\alpha = 0.01$.

H0: $\rho = 0$

H1: $\rho \neq 0$



$\alpha = 0.01$ with 2 tail test.

After using Rstudio :

Correlation coefficient r = 0.1104422

Degree of freedom = 98

Sample size = 100

Critical value

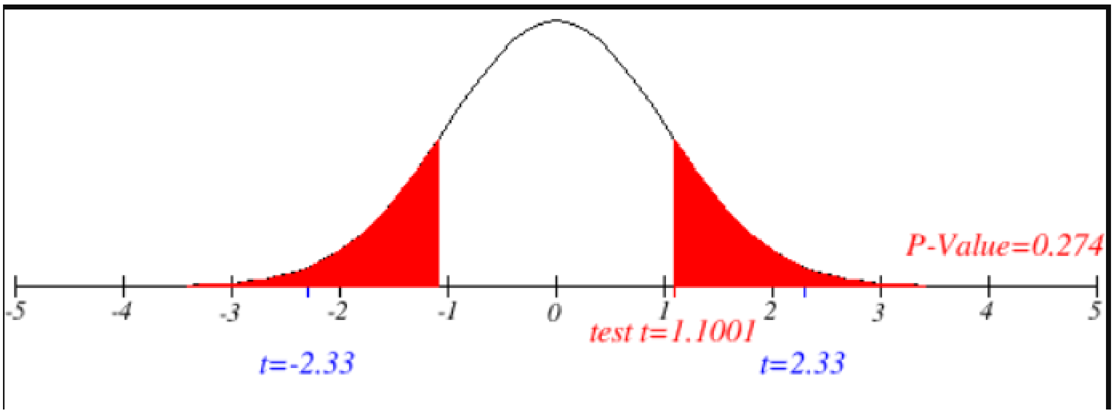
-t = -2.33

t = 2.33

$$t = \frac{r}{\sqrt{\frac{1-r^2}{n-2}}}$$

test statistic:

t = 1.1001



Hence , the critical shown in the provided graph of normal distribution is

-1 and 1 not 2.33 , - 2.33

Decision making:

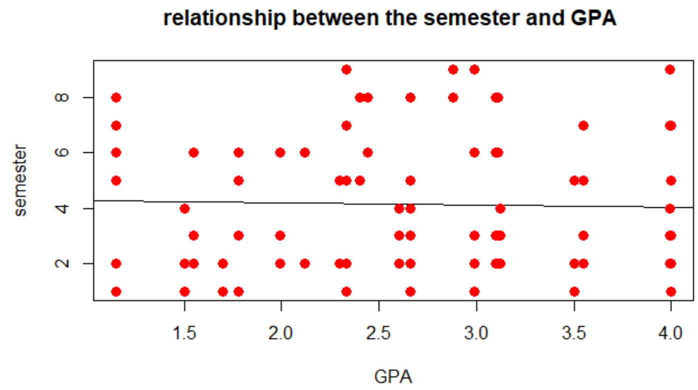
Since t = 1.1001 is located between -t and t , Hence we failed to reject the null hypothesis.

Regression analysis

Test 4: Regression analysis to explain the relationship between the semester and GPA that the students had during their studies. Assume that the confidence level = 95%, significant level, α = 0.05.

H0: β1 = 0

H1: β1 ≠ 0



The r-squared value of regression, R^2 = 0.0008262 so that's only 0.0008262%.

Sample size = 100

Degree of freedom = 98

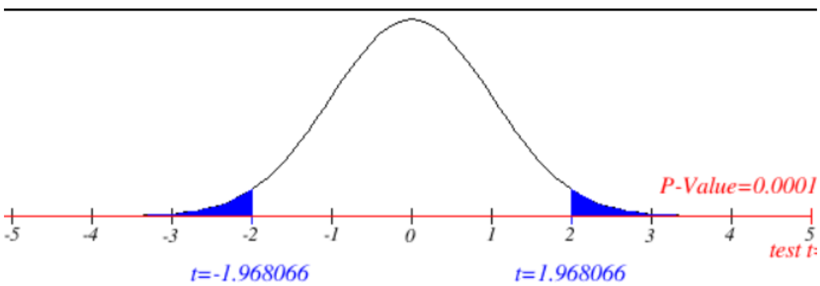
α = 0.05.

critical value:

t = -1.968066

t = 1.968066

ŷ = 2.68788 -0.01067



Decision making:

As you can see in the provided normal distribution the critical value is -2, 2 instead of 1.968066, -1.968066

Test statistic t = -0.285

Since the Test statistic t = -0.285 is greater than t = -1.968066 and

t = 1.968066 so we fail to reject the hypothesis.

Chi Square test of independence

Test 5:Chi Square test of independence to determine whether their number of classes affects the sleeping hours for students. Assume the confidence level = 95%, significant level, α = 0.05.

H0: the number of classes (independent)

H1: the number of classes (independent)

After using Rstudio:

contingency table

	1	2	3	4	5	6	7	8
1	1	2	0	2	3	4	0	0
2	4	4	3	1	3	3	1	1
3	0	3	1	1	0	3	2	0
4	1	2	0	6	1	2	2	1
5	0	1	1	3	2	1	1	2
6	0	2	1	3	2	2	1	0
7	2	3	1	3	0	0	5	1
8	1	1	3	0	0	0	1	0

number of classes 1 – 8

sleeping hours

1	2	0	2	3	4	0	0
4	4	3	1	3	3	1	1
0	3	1	1	0	3	2	0
1	2	0	6	1	2	2	1
0	1	1	3	2	1	1	2
0	2	1	3	2	2	1	0
2	3	1	3	0	0	5	1
1	1	3	0	0	0	1	0

Contingency table that used for test statistic:

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
1	1	2	2	7	0	0
2	4	7	1	6	1	1
3	0	4	1	3	2	0
4	1	2	6	3	2	1
5	0	2	3	3	1	2
6	0	3	3	4	1	0
7	2	4	3	0	5	1
8	1	4	0	0	1	0

$x^2 = 45.569$

α = 0.05

```
Pearson's Chi-squared test

data:  tbl2
X-squared = 45.569, df = 35, p-value = 0.1089
```

The degree of freedom is 35

Critical value = 7.814

p-value = 0

$$\chi^2 = \sum_{\text{all cells}} \frac{[o_{ij} - e_{ij}]^2}{e_{ij}}$$

Observed count

Expected count

Decision making:

Since the test statistic, χ 2=45.569 is more than Critical value = 7.814, falls within the critical region also p-value = 0 is less than α = 0.05 Hence, we reject the null hypothesis.