



EE-495
Special Topics In Computer Engineering

Case Study No. 2
Case Study on Chi Square Test

Haneen Alamoudi - 1708436

Submission Date: 19 March 2021 (WEEK 9)

Contents

Introduction.....	3
Output	4
Code.....	4
Conclusion	5

Introduction

Pearson's chi-squared test is a statistical test applied to sets of categorical data to evaluate how likely it is that any observed difference between the sets arose by chance. The main objective of this case study to conclude whether gender and the choice of pets is related to each other starting with the data at table 1. The case study significance factor is 0.05.

	Dog	Cat	Bird	Total
Men	207	282	241	730
Women	234	242	232	708
Total	441	524	473	1438

Null Hypothesis: There is no relation between gender and the choice of pets

Alternate Hypothesis: There is a significant relation between gender and the choice of pets

Theoretical Solution (chi-square):

If the calculated value of chi-square is less or equal to the tabular value -> H0 holds true, table 2 shows the expected values that was calculated using the equation $\frac{\text{row total} * \text{column total}}{\text{ground total}}$.

	Dog	Cat	Bird	Total
Men	223.873	266.008	240.118	730
Women	217.126	257.992	232.882	708
Total	441	524	473	1438

And the chi-square table is shown in the table below, $\sum \frac{(\text{Observed_value} - \text{calculated_value})^2}{\text{Calculated_value}} = 4.5422$. So, when comparing the chi-square value from the critical value of 5.991 H0 is accepted and there is no relation between gender and the choice of pets.

Output

The screenshot shows the Spyder Python IDE interface. The left pane displays a Python script for a chi-squared contingency test. The script includes a header with author information and a date, followed by the test implementation using SciPy. The right pane shows the Variable explorer with the following data:

Name	Type	Size	Value
alpha	float	1	0.05
data	list	2	[[207, 282, 241], [234, 242, 232]]
dof	int	1	2
expected	Array of float64	(2, 3)	[[223.87343533 266.00834492 240.11821975] [217.12656467 257.99165508 ...]]
p	float64	1	0.1031971404730939
stat	float64	1	4.542228269825232

The bottom pane shows the console output, which includes a `TypeError: cannot unpack non-iterable function object` and the execution of the script, resulting in the p-value and the decision to reject the null hypothesis.

Code

```
"""
Created on Thu Mar 18 20:55:07 2021
@author: Haneen Alamoudi
ID: 1708436
"""

from scipy.stats import chi2_contingency

data = [[207, 282, 241], [234, 242, 232]]
stat, p, dof, expected = chi2_contingency(data)

alpha = 0.05
print("p value is " + str(p));
if p <= alpha:
    print('Dependent (reject H0)')
else:
    print('Independent (reject H0)')
```

Conclusion

Based on the calculation done using the Python code, $p = 0.1031971404730939$. since P-value is greater than 0.05, we cannot reject the null hypothesis because it is not significant. And conclude that there is no relation between gender and the choice of pets.