Assignment\_\_3: Pravin Dattatray Shinde batch :5th aug 2024 syeda saymreen

Chi-Square test

--Association Between Device Type and Customer Satisfaction

Background:

Mizzare Corporation has collected data on customer satisfaction levels for two types of smart home devices: Smart Thermostats and Smart Lights. They want to determine if there's a significant association between the type of device purchased and the customer's satisfaction level.

Data Provided:

The data is summarized in a contingency table showing the counts of customers in each satisfaction level for both types of devices:

| Satisfaction Level | Smart Thermostat | Smart Light | Total |

|---------------------|------------------|-------------|------- |

| Very Satisfied | 50 | 70 | 120 |

| Satisfied | 80 | 100 | 180 |

| Neutral | 60 | 90 | 150 |

| Unsatisfied | 30 | 50 | 80 |

| Very Unsatisfied | 20 | 50 | 70 |

| \*\*Total\*\* | \*\*240\*\* | \*\*360\*\* | 600 |

Steps:

1. State the Hypotheses:

- Null Hypothesis (H₀)\*\*: There is no significant association between the type of device purchased and customer satisfaction levels. (i.e., satisfaction levels are independent of device type).

- Alternative Hypothesis (H₁)\*\*: There is a significant association between the type of device purchased and customer satisfaction levels.

2. Compute the Chi-Square Statistic:

To compute the Chi-square statistic, we calculate the expected frequencies for each cell using the formula:

\[

E = \frac{{(\text{{row total}}) \times (\text{{column total}})}}{{\text{{grand total}}}}

\]

The observed and expected frequencies are used to calculate the Chi-square statistic:

\[

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

\]

Where:

- \( O \) is the observed frequency.

- \( E \) is the expected frequency.

3. Determine the Critical Value:

Using the significance level \( \alpha = 0.05 \) and degrees of freedom \( df = (r-1)(c-1) = 4 \), the critical value is obtained from the Chi-square distribution table.

4. Make a Decision:

Compare the Chi-square statistic with the critical value. If the statistic is greater than the critical value, reject the null hypothesis. Otherwise, fail to reject it.

Python Code Implementation:

```python

import numpy as np

import scipy.stats as stats

Observed frequencies from the table

observed = np.array([

[50, 70],

[80, 100],

[60, 90],

[30, 50],

[20, 50]

])

Row totals and column totals

row\_totals = np.array([120, 180, 150, 80, 70])

col\_totals = np.array([240, 360])

grand\_total = 600

-Expected frequencies

expected = np.outer(row\_totals, col\_totals) / grand\_total

-Calculate the chi-square statistic

chi\_square\_stat = ((observed - expected) \*\* 2 / expected).sum()

-Degrees of freedom

df = (observed.shape[0] - 1) \* (observed.shape[1] - 1)

-Critical value from chi-square distribution with df=4 at alpha=0.05

alpha = 0.05

critical\_value = stats.chi2.ppf(1 - alpha, df)

-P-value

p\_value = 1 - stats.chi2.cdf(chi\_square\_stat, df)

chi\_square\_stat, critical\_value, p\_value

```

-5. Results:

- Chi-square statistic\*\*: 5.64

- Critical value\*\*: 9.49

- P-value: 0.23

6. Conclusion:

Since the Chi-square statistic (5.64) is less than the critical value (9.49), and the p-value (0.23) is greater than the significance level (0.05), we \*\*fail to reject the null hypothesis\*\*.

There is not enough evidence to suggest a significant association between the type of smart home device purchased and customer satisfaction levels. Therefore, customer satisfaction appears to be independent of the device type.