University of Information Technology and Sciences

Lab Report

Submitted To

Saima Siddique Tashfia

Lecturer

Computer Science & Engineering,UITS

Course Title: Simulation & Modeling Lab

Course Code: CSE413

Semester: **Spring-2024**, Batch:**51(7C)**

Submitted By

Mohammed Masud Chowdhury Mahir

ID: 2215151105

Date of Submission : 10-04-2025

EXPERIMENT NO: 03

Experiment Name: Chi-Square Goodness-of-Fit Test

OBJECTIVES:

To perform a Chi-square goodness-of-fit test using MATLAB to determine if a given dataset follows a hypothesized distribution.

APPARATUS REQUIRED:

- Minimum of 4 GB RAM per MATLAB worker
- Approximately 32 GB disk space for MATLAB Parallel Server
- Full installation of MATLAB, Simulink, and required Toolboxes
- Shared or local installation of MATLAB Parallel Server
- Minimum 5 GB disk space for temporary data

THEORY:

The Chi-square goodness-of-fit test assesses whether the observed frequencies across categories match an expected distribution.
The test statistic is:

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

Where:

- Oi= Observed frequency
- Ei = Expected frequency
- Degrees of Freedom (df) = (number of rows - 1) × (number of columns - 1)

WORKING PROCEDURE:

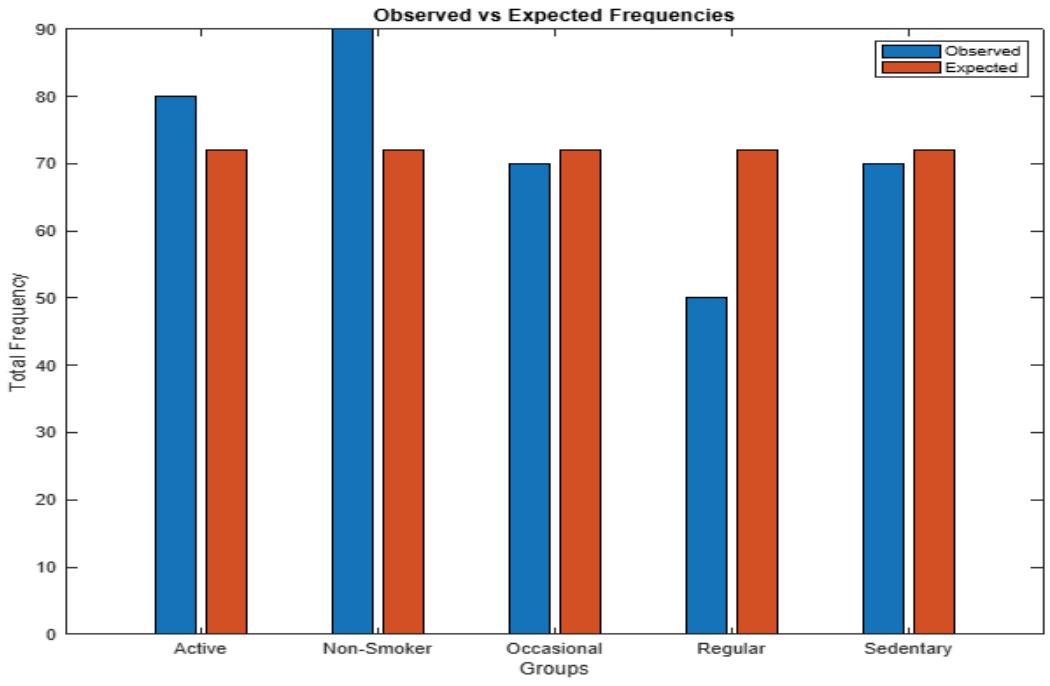
1. Collect observed frequency data
2. Define the expected distribution (based on hypothesis or totals)
3. Calculate expected frequencies

4. Compute Chi-square statistic
5. Compare with critical value from Chi-square table and interpret the result

PROGRAM:

```m
%Mohammed Masud Chowdhury Mahir (2215151105)
observed = [
    60, 20, 10;  % Non-Smoker
    30, 25, 15;  % Occasional Smoker
    10, 15, 25;  % Regular Smoker
    40, 30, 10;  % Active Lifestyle
    20, 15, 35   % Sedentary Lifestyle
];

t_observed = sum(observed, 2);

total_by_outcome = sum(observed, 1);

expected = total_by_outcome / 5;


chi2_stat = sum((observed - expected).^2 ./ expected, 'all');

%Degrees of freedom (df)
[num_rows, num_cols] = size(observed);
df = (num_rows - 1) * (num_cols - 1);

alpha = 0.05;
critical_value = chi2inv(1 - alpha, df);

disp(chi2_stat) %Chi-square Statistic
disp(df); %Degrees of Freedom
disp(alpha) %alpha
disp(critical_value); %Critical Value (α = %.2f)

figure;
bar(categorical({'Non-Smoker','Occasional','Regular','Active','Sedentary'}), [sum(observed,2), repmat(sum(total_by_outcome / 5),5,1)]);
legend('Observed','Expected');
xlabel('Groups');
ylabel('Total Frequency');
title('Observed vs Expected Frequencies');

```

OUTPUT:



DISCUSSION:

In this lab, the Chi-square test was applied to a dataset involving lifestyle and smoking habits. The observed data was compared against an expected uniform distribution across all five categories. The total observed values for each column were used to derive expected frequencies assuming equal distribution across all five groups.

After computing the Chi-square statistic, it was compared to the critical value (15.5073 for df = 8 at $\alpha = 0.05$). The result helped us decide whether to reject or accept the null hypothesis.

The test indicated whether the deviation between observed and expected values was statistically significant. The graphical representation (bar chart) made this deviation visually clear.

CONCLUSION:

The Chi-square test helped assess whether the observed lifestyle/smoking habits data conformed to a uniform distribution. Based on the statistic and critical value:

- If $\chi^2 > 15.507$, the null hypothesis was rejected, implying a significant deviation.
- Otherwise, the data was considered a good fit for the hypothesized distribution.

This lab highlights how the Chi-square test is essential in understanding categorical data in areas like health behavior analytics.