Journal Homepage: -www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

Article DOI:10.21474/IJAR01/19746
DOI URL: http://dx.doi.org/10.21474/IJAR01/19746

RESEARCH ARTICLE

CHI-SQUARE TESTS: A QUICK GUIDE FOR HEALTH RESEARCHERS

Dr.Valarmathi S.[1], Ms. Hemapriya A.S[2] and Dr. Jasmine S. Sundar[3]

1. Research Officer (Statistics), Department of Epidemiology, The Tamilnadu Dr. MGR Medical University, Chennai, Tamilnadu, India.
2. Master of Public Health, Independent Researcher, Chennai, Tamilnadu, India.
3. Assistant Professor, Department of Epidemiology, The Tamilnadu Dr. MGR Medical University, Chennai, Tamilnadu, India.

Manuscript Info

Manuscript History
Received: 27 August 2024
Final Accepted: 29 September 2024
Published: October 2024

Key words:-
Chi-Square Test, Applications, Medical Research

Abstract

Pearson's chi-square (X^2) tests are essential nonparametric statistical tools for analyzing associations among categorical data, making them crucial for research involving non-numeric variables. These tests are widely utilized in various research fields due to their independence from normal distribution assumptions. Chi-square tests are utilized to assess whether there is a significant association between groups, populations, or criteria, and to examine how closely observed data distributions align with expected ones. The three primary types of chi-square tests are: the Goodness-of-Fit test, which checks if the distribution of categorical data in a sample conforms to a predefined distribution; the Test of Independence, which investigates whether there is a relationship between categorical variables within a single sample; and the Test of Homogeneity, which compares the frequency counts of a categorical variable across multiple populations to see if their distributions are similar.For valid and reliable results, it is crucial to consider factors such as random sampling, adequate cell counts, sufficient sample size, and mutually exclusive variables. In healthcare research, chi-square tests are essential for analyzing risk factors, evaluating AYUSH treatments, assessing nursing interventions, and studying health behavior trends. This review underscores their importance in statistical analysis and evidence-based decision-making.

Introduction:-

In the early 20th century, the field of statistics was on the brink of a transformation, and at the heart of this revolution was Karl Pearson. In 1900, Pearson introduced the chi-square test, a pioneering method that would become fundamental in analyzing categorical data. The chi-square test's story begins with Pearson's frustration over the limitations of existing statistical tools, which were inadequate for dealing with the complexities of categorical data. This frustration drove him to create a new test, designed to determine whether the differences between observed and expected frequencies were the result of chance or actual underlying relationships.Pearson's creation quickly found its place in the scientific community. One of the earliest and most fascinating applications was in genetics. Researchers used the chi-square test to study Mendelian inheritance patterns, validating the laws of inheritance proposed by Gregor Mendel. In the 1920s, Sir Ronald A. Fisher, a leading figure in statistics, used the

Corresponding Author:-Dr.Valarmathi S.
Address:-Research Officer (Statistics), Department of Epidemiology, The Tamilnadu Dr. MGR Medical University, Chennai, Tamilnadu, India.

chi-square test to analyze Mendel's experimental data, helping to solidify its importance.As the chi-square test gained popularity, it became a vital tool for health researchers. It enabled them to explore associations between risk factors and health outcomes, providing insights into public health and epidemiology. The test's ability to handle categorical data made it indispensable for analyzing survey results, clinical trials, and epidemiological studies. Yet, with its widespread use, important questions arose: "How can health researchers effectively apply the chi-square test to evaluate new treatments?", "What are the critical assumptions and potential pitfalls when using this test?", and "How do sample size and cell frequencies influence the reliability of the results?" These questions underscore the need for a deep understanding of thethe mechanics of the chi-square test and its correct application.In essence, the chi-square test is a story of innovation and significant impact. From its creation by Karl Pearson to its pivotal role in contemporary health research, the chi-square test demonstrates how a statistical tool can transform scientific investigation and contribute to major advancements in our understanding of the world.

Methods:-

This review article explores the applications and methodologies of Pearson's chi-square tests in healthcare research. The review provides a detailed analysis of the three main chi-square tests: Goodness-of-Fit, Independence, and Homogeneity. Through an analysis in existing literature, identified using keywords such as "Chi-square test," "Pearson's chi-square," "Categorical data analysis," "Goodness-of-Fit test," "Test of Independence," "Test of Homogeneity," "Non-parametric statistics," "Healthcare research," "Clinical trials," "Epidemiology," "Public health," "Statistical methods in medicine," "Contingency tables," "Random sampling," and "SPSS chi-square analysis" in PubMed and Google Scholar, the article highlights the theoretical foundations, assumptions, and practical implications of these tests in various healthcare settings.

What is Chi-Square Test?

The chi-square test is a non-parametric statistical method used to assess the difference or association between observed and expected categorical variables in a dataset(1). Developed by Karl Pearson (1857-1936), it has become a key tool in hypothesis testing, particularly when dealing with categorical data.(2). One of the main advantages of Pearson's chi-square test is its ability to use statistical methods that do not assume a normal distribution. This feature makes the chi-square test a flexible tool for analyzing data across a wide range of studies.The significance of the chi-square value is determined by consulting a chi-square table, taking into account the appropriate degrees of freedom and the chosen level of significance.(2)

The chi-square test serves two primary purposes: determining whether there is an association between two or more groups, populations, or criteria, and evaluating how well the observed data distribution fits the expected distribution(3).

Chi-square tests are particularly useful for analyzing categorical data, such as gender (male or female) or smoking status (smokers vs. non-smokers). However, they are not appropriate for analyzing parametric or continuous data, such as measurements of height or weight(3). By focusing on categorical data, Pearson's chi-square test offers a reliable method for investigating relationships and patterns within datasets, making it an essential tool for statisticians and researchers in various fields.

Importance of Contingency Table

Pearson's chi-square tests come in two main types, both designed to assess whether the observed distribution of categorical data significantly differs from the expected distribution. A frequency distribution details how observations are distributed across different categories.

To visually represent frequency distributions, researchers often use frequency distribution tables. These tables illustrate the count of observations within each category. When examining relationships between two categorical variables, analysts employ contingency tables, a specialized form of frequency distribution table. These tables show the number of observations for each combination of categories.

Example: Examining Treatment Response by Gender

In a clinical trial investigating the efficacy of a new treatment for hypertension, researchers categorize participants by gender to analyze treatment response rates.

	Improved	**Not improved**	**Total**
Male	80	20	100
Female	70	30	100
Total	150	50	200

From the above table,

The Observed Distribution,

$O_{\text{Male, Improved}} = 80$

$O_{\text{Male, Not Improved}} = 20$

$O_{\text{Female, Improved}} = 70$

$O_{\text{Female, Not Improved}} = 30$

The Expected Distribution

$E_{\text{Male, improved}}$ = Total for Male X Total Improved / Total Observations = 75

$E_{\text{Male, Not improved}}$ = Total for Male X Total Not Improved / Total Observations = 25

$E_{\text{Female, improved}}$ = Total for Female X Total Improved / Total Observations = 75

$E_{\text{Female, Not improved}}$ = Total for Female X Total Not Improved / Total Observations = 25

Key Considerations for Effectively Using Pearson's Chi-square Test

When applying Pearson's chi-square test, several key considerations help ensure the validity and reliability of the results:

1. **Random Sampling**: The data should be randomly drawn from the population to avoid bias and ensure that the sample accurately represents the population.

Example: Survey on Dietary Habits - To study the association between dietary habits and heart disease, researchers' randomly select 200 participants from a population. Random sampling ensures that the sample accurately represents the entire population, minimizing bias.

2.**Adequate Cell Counts**: The expected counts in each cell should be adequate, typically no less than 5, and there should be no cells with a count of zero.This helps in maintaining the accuracy of the chi-square approximation(4) (If expected counts are less than 5, methods such as Fisher's Exact Test or Yates' Correction for Continuity can be used to find associations)

Example: Study on Smoking and Lung Cancer - To investigate the relationship between smoking and lung cancer, researchers ensure that each cell in their data table has an expected count of at least 5, with no cells having a count of zero. This prevents inaccuracies in the chi-square approximation.

2. **Sufficient Sample Size**: The sample size should be sufficiently large to minimize the risk of type II errors, where the null hypothesis is incorrectly accepted despite being false. Although there is no strict threshold, a minimum sample size typically ranges from 20 to 50.

Example: Effectiveness of a New Drug - In a clinical trial testing a new drug, researchers use a sample size of 50 participants to ensure the results are reliable and to minimize type II errors.

3. **Mutually Exclusive Variables**: The variables being analyzed must be mutually exclusive, meaning each observation should be counted only once in a specific category and should not appear in multiple categories. This ensures proper categorization and analysis of the data.

Example: Vaccine Efficacy Study - In a study to determine the efficacy of a vaccine, researchers categorize participants as either vaccinated or unvaccinated, ensuring that each participant is only counted in one category.

Types Of Chi-Square Tests

The chi-square test is categorized into three types based on its purpose and conditions of application:

1. Chi-square Goodness-of-Fit test
2. Chi-square Independence test
3. Chi-square Homogeneity test

Table 1:- Chi-Square Tests and Attributes.

Chi –Square test Attributes	Test of Goodness of Fit	Test of Independence	Test of Homogeneity
Sampling type	Sample from population	Single dependent sample	Two (or more) independent samples
Interpretation	Difference from population	Association between variables	Difference in proportions
Null hypothesis	No difference in distribution between sample and population	No association between Variables	No difference in proportion between groups

Table 1 shows the comparison of three different types of chi-square tests(5). These tests are distinguished based on the null hypothesis they address and the type of sampling involved. The interpretation of results can vary accordingly, shaped by these specific aspects of hypothesis testing and sampling methodology.

Table 2:- Sample Analysis Related to Chi-Square Tests.

Test Type	Aim	H_0 Hypothesis	Result
Chi-Square Goodness-of-Fit Test	To determine if the distribution of cancer types among patients in a hospital matches the expected distribution based on national cancer registry data.	The distribution of cancer types among patients in the hospital follows the expected distribution based on national cancer registry data.	The Chi-square test statistic (χ^2) was calculated to be 12.48. With 3 degrees of freedom and a significance level of 0.01, the critical value is 11.34. Since $\chi^2 >$ 11.34, we reject the null hypothesis. Therefore, there is a significant difference between the observed and expected distributions of cancer types among patients in the hospital.
Chi-Square Test of Independence	To investigate whether there is an association between patient age groups and the prevalence of diabetes.	There is no association between patient age groups and the prevalence of diabetes.	The Chi-square test statistic (χ^2) was calculated to be 25.67. With 2 degrees of freedom and a significance level of 0.01, the critical value is 9.21. Since $\chi^2 >$ 9.21, we reject the null hypothesis. Therefore, there is a significant association between patient age groups and the prevalence of diabetes.
Chi-Square Test of Homogeneity	To compare the recovery rates from a specific surgery among patients from three different hospitals.	The recovery rates from the surgery are the same across the three hospitals.	The Chi-square test statistic (χ^2) was calculated to be 10.83. With 2 degrees of freedom and a significance level of 0.01, the critical value is 9.21. Since $\chi^2 >$ 9.21, we reject the null hypothesis. Therefore, there is a significant difference in the recovery rates among the three hospitals.
McNemar Test	To evaluate the effectiveness of a new treatment by comparing the number of patients who improved versus those who did not before and after the treatment.	There is no significant change in the number of patients who improved before and after the treatment.	The McNemar test statistic was calculated to be 7.68 with 1 degree of freedom and a significance level of 0.01. The critical value is 6.63. Since $\chi^2 >$ 6.63, we reject the null hypothesis. Therefore, there is a

			significant change in the number of patients who improved before and after the treatment.

Table 2 presents the statistical analyses that yielded significant findings in healthcare research. The Chi-Square Goodness-of-Fit Test produced a test statistic of 12.48, exceeding the critical value of 11.34 (with 3 degrees of freedom and a significance level of 0.01), indicating a significant difference between observed and expected cancer type distributions. The Chi-Square Test of Independence resulted in a test statistic of 25.67, surpassing the critical value of 9.21 (with 2 degrees of freedom and a significance level of 0.01), revealing a significant association between patient age groups and diabetes prevalence. The Chi-Square Test of Homogeneity produced a test statistic of 10.83, exceeding the critical value of 9.21 (with 2 degrees of freedom and a significance level of 0.01), showing significant differences in recovery rates from surgery across three hospitals. The McNemar Test yielded a test statistic of 7.68, surpassing the critical value of 6.63 (with 1 degree of freedom and a significance level of 0.01), indicating a significant change in patient improvement rates before and after the new treatment. These results highlight the effectiveness of chi-square tests in analyzing healthcare data.

1. Chi-Square Goodness-of-Fit Test

Definition:A nonparametric test used to assess whether a single sample of categorical data aligns with a specified distribution or expectation.

When to Use: To determine if the observed frequencies of categories in your sample correspond to the expected frequencies based on a known distribution.

Example:To determine if the distribution of cancer types among patients in a hospital matches the expected distribution based on national cancer registry data.

Interpretation: Compare the calculated χ^2 value to the critical value determined by the degrees of freedom (df) and the selected significance level (α). If χ^2 is greater than the critical value, reject the null hypothesis (H_0), suggesting a significant difference between the observed and expected distributions.

2. Chi-Square Test of Independence

Definition: A test used to determine whether there is an association between two categorical variables within a sample, by comparing the observed frequencies to the expected frequencies assuming independence, is the Chi-square Test of Independence.

When to Use: To investigate whether two categorical variables are related.

Example: To examine whether there is an association between patient age groups and the prevalence of diabetes.

Interpretation:Compare the calculated χ^2 value to the critical value based on the degrees of freedom (df) and the chosen significance level (α). If χ^2 is greater than the critical value, reject the null hypothesis (H_0), suggesting a significant association between the two variables.

3. Chi-Square Test of Homogeneity

Definition: A statistical method used to compare the distribution of a single categorical variable across two or more different populations.

When to Use: To compare the distributions of a categorical variable across multiple populations.

Example: To compare the recovery rates from a specific surgery among patients from three different hospitals.

Interpretation: Compare the calculated χ^2 to the critical value based on df and α. If $\chi^2 >$ critical value, reject H0, indicating a significant difference in the distributions among the populations.

McNemar Test

Definition:A test used for paired nominal data to determine if there are differences on a dichotomous dependent variable between two related groups.

When to Use: When you have paired data (e.g., pre-test/post-test data) and want to assess whether there is a significant change in proportions. It is particularly useful in before-and-after studies or matched pairs designs.

Example: To evaluate the effectiveness of a new treatment by comparing the number of patients who improved versus those who did not before and after the treatment.

Interpretation: Compare the number of discordant pairs (e.g., patients who improved versus those who did not) and use the McNemar test statistic to determine if there is a significant change.

When to Use a Different Test?

If the expected counts in any cell are less than 5, consider using Fisher's Exact Test or Yates' Correction for Continuity.

Methods For Conducting Chi-Square Tests
Manual Calculation using the formula,

The comprehensive form of all three tests in the Karl Pearson family of chi-square tests—Goodness-of-Fit,

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i},$$

Independence, and Homogeneity—essentially employs the same formula.

Example Scenario: You aim to examine whether there is a relationship between gender (independent variable) and smoking status (dependent variable) within a group of individuals.
Let's assume we have the following observed frequencies:

Step 1: Formulate the Hypotheses
H0: There is no association between gender and smoking status.
H1: There is an association between gender and smoking status.
Step 2: Set up the Contingency Table

	Non-Smoker	**Smoker**	**Total**
Male	200	150	350
Female	250	180	430
Total	450	330	780

Step 3: Calculate Expected Frequencies
Calculate expected frequencies for each cell:
$E_{ij} = n_i * n_j / n$
Where:
n_i and n_j are the row and column totals, respectively.
n is the total sample size
Now, $E_{\text{male non-smokers}} = 350 \times 450/780 = 201.92$
$E_{\text{male smokers}} = 350 \times 330/780 = 148.08$
$E_{\text{female non-smokers}} = 430 \times 450/780 = 248.08$
$E_{\text{female smokers}} = 430 \times 330/780 = 181.92$

Step 4: Compute the Chi-Square Statistic
Now, calculate the chi-square statistic:

$$\chi 2 = \sum (Oij - Eij)^2 / Eij$$

Calculate for each cell and sum them up:
$\chi 2 = (200-201.92)^2/201.92 + (150-148.08)^2/148.08 + (250-248.08)^2/248.08 + (180-11.92)^2/181.92$
$\chi 2 = 1.84/201.92 + 1.16/148.08 + 1.17/248.08 + 1.84/181.92$
$\chi 2 = 0.009+0.008+0.005+0.010$
$\chi 2 = 0.032$
Step 5: Determine Degrees of Freedom (df)
Degrees of freedom df=(2−1)(2−1)=1df = (2-1)(2-1) = 1df=(2−1)(2−1)=1

Step 6: Compare with Critical Value
At a significance level of $\alpha=0.05$, the critical value of chi-square with 1 degree of freedom is approximately 3.84 (from chi-square distribution table).

Step 7: Make a Decision
Since $\chi^2 = 0.032$ is less than 3.84, we do not reject the null hypothesis (H_0). This indicates that there is no significant association between gender and smoking status in this sample.

Figure 1:- How to Perform a Chi-Square Test Using SPSS.

Launch SPSS on your computer

↓ **Input data**

- Enter your data into the SPSS Data Editor.
- Each row should represent an individual case, and columns should represent your variables.
- Create columns for "Gender" and "Smoking Status".

↓ **Define Variables**

- Click on the "Variable View" tab.
- Name your variables (e.g., "Gender" and "Smoking_Status").
- Set the "Type" to "String" or "Numeric".
- Label your variables.

↓ **Navigate to Chi-Square Test**

- Click on Analyze.
- Select Descriptive Statistics > Crosstabs....

↓ **Set-up Crosstabs**

- Move "Gender" to the Row(s) box.
- Move "Smoking_Status" to the Column(s) box.

↓ **Select Statistics**

- Click on the Statistics... button.
- Check the box for Chi-square.

↓ **Display Results**

- Optionally, click on the Cells... button.
- Select additional display options.

↓ **Run the test**

↓

- Click OK to run the Chi-square test.
- SPSS will output the results.
- Click Continue.

Interpret Results

- Find the Chi-square test results in the Output Viewer.
- Look for the Pearson Chi-Square value and p-value.
- Compare the p-value to your significance level.

Example Output Interpretation

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	0.032	1	0.857

In Figure 1, the p-value of 0.857 exceeds the significance level (e.g., 0.05), so we fail to reject the null hypothesis. Therefore, the data does not show a significant association between gender and smoking status.

Table 3:- Implications Of Chi-Square Tests In Different Healthcare Research Settings.

Type of Study	Medical	AYUSH	Nursing	Allied Health
Cohort Studies	Analyze associations between risk factors and outcomes over time. **Example:** Investigating association between smoking status and lung cancer development over 10 years.	Evaluate effectiveness of AYUSH treatments over time. **Example:** Analyzing association between daily yoga practice and stress reduction over a year.	Examine impact of nursing interventions over time. **Example:** Assessing association between wound care management and healing rates in diabetic patients over a year.	Examine how various factors, such as lifestyle behaviors or interventions, impact their health outcomes throughout an extended period. **Example:** Investigating association between physical activity levels and cardiovascular health over a decade.
Case-Control Studies	Assess association between exposure and disease status. **Example:** Comparing prevalence of a genetic mutation in cancer patients vs. healthy individuals.	Assess relationship between AYUSH practices and health outcomes. **Example:** Studying impact of Ayurvedic herbs on digestion in patients with gastrointestinal disorders.	Evaluate relationship between nursing practices and patient outcomes. **Example:** Investigating association between medication administration protocols and medication errors.	Evaluate relationship between exposures and health conditions. **Example:** Comparing prevalence of ergonomic risk factors among individuals with and without musculoskeletal disorders.
Cross-Sectional Studies	Examine associations between variables at a single point in time. **Example**: Studying	Explore prevalence of AYUSH therapies across demographic groups.	Explore nursing workforce characteristics and their impact on	Analyze distribution of health behaviors and outcomes across diverse populations.

	association between dietary habits and prevalence of cardiovascular disease.	**Example:** Analyzing association between age groups and use of homeopathic treatments for allergies.	patient care. **Example:** Analyzing association between nurse staffing levels and patient satisfaction scores.	**Example:** Assessing association between socioeconomic status and adherence to dietary guidelines.
Clinical Trials	Compare treatment outcomes between different intervention groups. **Example:** Testing efficacy of a new antidepressant using response rates in treatment vs. placebo groups.	Compare treatment outcomes between different therapeutic approaches. **Example:** Evaluating efficacy of Siddha medicine for arthritis relief using pain reduction rates.	Assess effectiveness of nursing interventions compared to standard care. **Example:** Evaluating impact of new nursing protocol on reducing patient falls.	Assess effectiveness of interventions on health outcomes. **Example**: Evaluating impact of a new physical therapy technique on improving mobility in stroke patients.

Table 3 highlights the diverse applications of chi-square tests in healthcare research. In cohort studies, they analyze long-term associations, such as the link between smoking and cancer or AYUSH treatments and stress reduction. Case-control studies use chi-square tests to assess exposure-disease relationships, like genetic mutations in cancer patients or the effects of AYUSH practices. Cross-sectional studies examine variables at a single time point, such as dietary habits and disease prevalence. Clinical trials apply chi-square tests to compare outcomes between intervention groups, evaluating the efficacy of new treatments or therapies. These uses demonstrate the flexibility of chi-square tests in various research contexts.

Conclusion:-

This comprehensive article on Pearson's chi-square tests provides a thorough understanding of its applications and methodologies across various healthcare research settings. It explains the different types of chi-square tests—goodness-of-fit, independence, and homogeneity—by outlining their specific applications and providing practical examples for better understanding and interpretation.By emphasizing the importance of random sampling, adequate cell counts, and sufficient sample sizes, it guides researchers in ensuring the validity and reliability of their chi-square analyses. Moreover, the article illustrates how to conduct these tests manually and using SPSS, making it accessible for both beginners and experienced researchers. Overall, this resource serves as a valuable tool for healthcare professionals, statisticians, and researchers aiming to employ chi-square tests effectively in analyzing categorical data, thereby contributing to evidence-based decision-making and advancements in healthcare research.

Disclosure Statement:
The authors declare that they have no competing interests.

Funding:
Nil.

References:-

1. Turney S. Scribbr. 2022 [cited 2024 Jun 21]. Chi-Square (X2) Tests | Types, Formula & Examples. Available from: https://www.scribbr.com/statistics/chi-square-tests/
2. Nihan ST. Karl Pearsons chi-square tests. Educ Res Rev. 2020 Sep 30;15(9):575–80.
3. Rana R, Singhal R. Chi-square Test and its Application in Hypothesis Testing. J Pract Cardiovasc Sci. 2015 Apr;1(1):69.
4. Contingency Tables Involving Small Numbers and the χ2 Test | Journal of the Royal Statistical Society Series B: Statistical Methodology | Oxford Academic [Internet]. [cited 2024 Jun 26]. Available from: https://academic.oup.com/jrsssb/article-abstract/1/2/217/6905373?redirectedFrom=fulltext
5. The Chi-Square Test: Often Used and More Often Misinterpreted - Todd Michael Franke, Timothy Ho, Christina A. Christie, 2012 [Internet]. [cited 2024 Jun 26]. Available from: https://journals.sagepub.com/doi/10.1177/1098214011426594.