

How to Perform a Chi-Square Test of Independence to Determine Relationships Between Variables

Authored by
stats writer

January 22, 2026

RECOMMENDED CITATION

stats writer (2026). *How to Perform a Chi-Square Test of Independence to Determine Relationships Between Variables*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=127066>

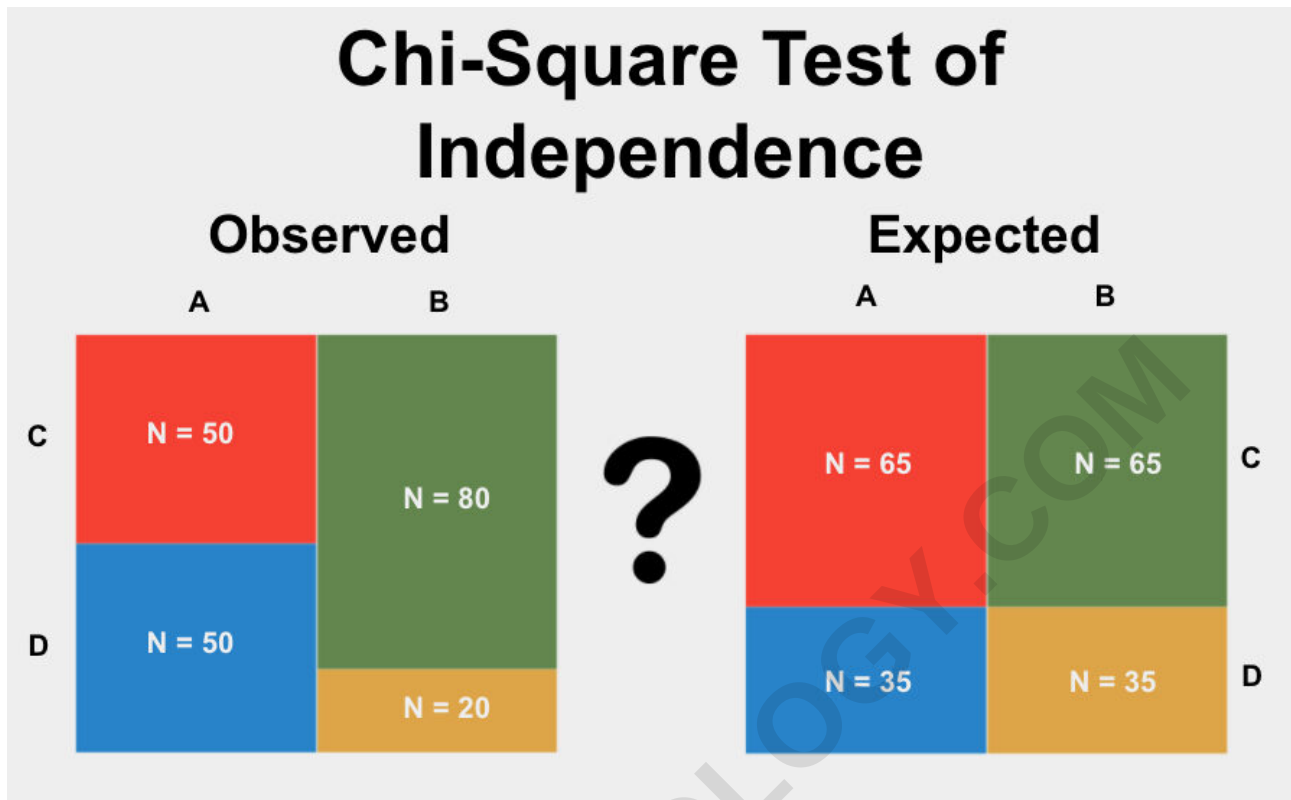
The Chi-Square Test of Independence is a statistical method essential for researchers across various disciplines. It is expertly utilized to determine whether there is a statistically significant relationship between two sets of categorical variables. The power of this test lies in its ability to compare the actual observed counts in a dataset against the theoretical expected counts that would occur if the variables were completely unrelated (or independent). By calculating the Chi-Square statistic, we gain insight into the divergence between the observed and expected results, allowing us to judge if this divergence is large enough to be considered statistically significant. This technique is indispensable in fields like public health, market research, and the social sciences for identifying underlying patterns and making robust, data-driven conclusions about how variables interact.

Understanding the Chi-Square Test of Independence

The Chi-Square Test of Independence (often denoted as χ^2) serves a critical function in bivariate analysis: it assesses the degree to which the proportions of categories across two different group variables differ from one another. Specifically, it tests the null hypothesis (H_0) that there is no association between the two variables, meaning they are statistically independent. If the test results lead us to reject the null hypothesis, we conclude that a significant relationship or dependency exists between the variables, suggesting that knowing the category of one variable helps predict the category of the other.

To properly execute this powerful statistical procedure, a researcher must ensure their data meets specific structural requirements. Primarily, the data must consist of two variables, both of which are nominal or categorical variables, and each variable must possess two or more possible options or levels. Furthermore, the test is based on a contingency table where the frequencies of intersection between the categories are counted. A crucial guideline for the validity of the test is that the expected count within every single cell of this table should ideally exceed a minimum threshold, often cited as 10, though some literature accepts 5. Meeting these prerequisites ensures the asymptotic assumptions of the test statistic hold true.

The underlying mechanism of the test involves comparing the actual counts observed in the study--the observed frequencies--with the counts that would theoretically be present if the two variables were perfectly independent--the expected frequencies. A large discrepancy between these two frequency sets results in a large Chi-Square statistic, suggesting that the observed relationship is unlikely to be due to random chance. This comparison is central to inferential statistics, allowing researchers to draw conclusions about a larger population based on sample data and determine if the observed differences are statistically significant.



The **Chi-Square Test of Independence** is also known by several related terms in statistical literature, including The Chi-Square Test of Homogeneity or simply the Chi-Squared Test of Independence. While technically distinct in their sampling methods--homogeneity tests whether distributions are the same across populations, while independence tests association within a single population--they utilize the same core computational formula.

Essential Assumptions for Valid Chi-Square Results

Like all methods in inferential statistics, the Chi-Square Test of Independence relies on specific underlying assumptions. These assumptions dictate the conditions under which the statistical method is mathematically sound and capable of producing accurate and unbiased results. If the collected data violates these fundamental properties, the conclusions drawn from the test may be unreliable or outright incorrect, potentially leading to erroneous interpretations of the relationship between variables.

Understanding and verifying these assumptions is a mandatory preliminary step before conducting the analysis. Failing to check these conditions is a common source of statistical error in research. By adhering to these requirements, we ensure that the calculated test statistic accurately reflects the true relationship, or lack thereof, within the population being studied. The key assumptions for this test ensure data quality and the applicability of the Chi-Square distribution, which is asymptotic.

The primary assumptions for the Chi-Square Test of Independence involve requirements regarding sampling methods, variable interaction, and data structure, summarized as follows:

The data must originate from a Random Sample.

The observations must exhibit Independence.

The categories within variables must be **Mutually Exclusive**.

The **Expected Cell Counts** must be sufficiently large.

Let us delve into a more detailed explanation of what each of these critical assumptions implies for your data collection and preparation process, ensuring methodological rigor in your study.

The Necessity of a Random Sample

The assumption of a random sample is paramount to valid statistical inference. It requires that every data point, or observation, included in the analysis must have been drawn from the population using a simple random sampling technique. This means that every individual, event, or subject in the population had an equal chance of being selected for the sample. When sampling is not random--for instance, if selection is based on convenience or is systematically biased--the sample may not accurately represent the target population, severely limiting the generalizability of the findings.

Failure to achieve random sampling introduces what statisticians refer to as **sampling bias**. Bias is a systematic error that distorts the results, typically by favoring certain outcomes or characteristics, thereby rendering the sample non-representative. If your groups were not randomly selected or assigned, the observed differences might be attributable to the sampling method itself rather than a true underlying relationship between the variables of interest. Therefore, ensuring that the data points for both groups originate from a representative, simple random sample is foundational to obtaining valid and generalizable results from the Chi-Square test, allowing conclusions to be extrapolated beyond the immediate dataset.

Independence of Observations

The assumption of Independence requires that each observation or data point collected should be independent of all others. In practical terms, this means that the measurement taken from one subject or unit should not influence, or be influenced by, the measurement taken from any other subject or unit. This is critical because the mathematical calculation of the **Chi-Square statistic** assumes that errors and variations are not systematically linked across observations; dependent data violates this core premise.

A common violation of this assumption occurs when **repeated measurements** are taken over time from the same unit of observation, such as tracking the behavior of the same set of customers

monthly, or measuring the performance of the same group of students multiple times (e.g., longitudinal studies). Since measurements from the same source are inherently related (they 'depend' on that single source), they violate the independence rule. When dependence is present, specialized statistical methods designed for paired data or dependent samples (such as the McNemar Test) should be employed instead of the standard Chi-Square Test of Independence.

Requirement for Mutually Exclusive Categories

The data used in a contingency table must be structured so that the categories within each categorical variable are mutually exclusive. This means that any given observation must fall into one, and only one, category for that variable; there can be no overlap. For example, if a variable records a person's response to an opinion survey, the categories might be "Strongly Agree," "Agree," "Neutral," "Disagree," and "Strongly Disagree." A respondent cannot logically belong to both the "Agree" and "Disagree" categories simultaneously.

If the categories overlap, the counts within the contingency table become inflated or ambiguous, violating the fundamental principle of discrete counting upon which the Chi-Square formula is built. Ensuring that the variable groupings are discrete and non-overlapping is essential for accurately calculating both the observed frequencies and the subsequent expected frequencies, thus preserving the integrity of the test statistic and ensuring that the degrees of freedom are calculated correctly.

Defining the Conditions: When to Apply the Chi-Square Test

The Chi-Square Test of Independence is not universally applicable; its usage is restricted to specific data structures and research questions. A researcher must meticulously review the nature of their variables and the structure of their collected data against a set of five key criteria to determine if this test is the appropriate statistical tool. Applying the test when the criteria are not met can lead to flawed conclusions, emphasizing the need for precision in method selection and ensuring the output is truly statistically significant.

You should proceed with the Chi-Square Test of Independence only when your study satisfies all the following methodological and structural conditions simultaneously. These conditions ensure that the underlying mathematical requirements for the Chi-Square distribution are met, guaranteeing the validity of the resulting p-value.

The objective is to test for a **Difference or Association** between variables.

The variables of interest must be **Proportional or Categorical**.

Each categorical variable must have **Two or more Options** (levels).

The samples used for comparison must be **Independent Samples**.

There must be **Sufficient Observations** (expected frequency typically more than 10) in every cell

of the contingency table.

We will now elaborate on each of these criteria to provide clear guidance on recognizing scenarios appropriate for the Chi-Square Test of Independence, helping researchers make sound methodological decisions before proceeding to data analysis.

Testing for Association or Difference

The primary purpose of the Chi-Square Test of Independence is to determine if a relationship or association exists between two variables. When conducting an analysis, it is vital to distinguish between three main objectives: testing for difference, testing for relationship/association, and prediction. This particular test falls squarely into the domain of assessing association. You are seeking to answer: Does the distribution of categories in Variable A depend on the category chosen in Variable B, or are the two distributions independent of each other?

For instance, if you are studying whether socioeconomic status (Variable A, categorized as Low/Medium/High) influences the type of transportation used (Variable B, categorized as Car/Bus/Bike), you are testing for an association. If the test reveals a statistically significant result, it suggests that the proportion of individuals using a Car is not the same across all socioeconomic statuses. If your objective, however, is predicting a continuous outcome (like income) or comparing means, other statistical models, such as ANOVA or regression analysis, would be more appropriate.

Requirement for Categorical Data Types

The Chi-Square Test is strictly designed for use with nominal or categorical variables. A categorical variable is one that classifies observations into distinct, non-ordered groups. Classic examples include demographic characteristics like eye color, geographical region, or response choices like 'Pass'/'Fail' or 'Male'/'Female'. The counts or frequencies derived from these categories form the basis of the contingency table upon which the test is performed.

Proportional variables, which represent the fraction or percentage of a category within a total group (e.g., success rate, market share), are derived directly from these categorical counts and are thus also suitable for testing via Chi-Square, provided the raw counts meet the minimum cell frequency requirements. It is essential to remember that if your variable of interest is continuous--measured on an interval or ratio scale, such as height or measured reaction time--the Chi-Square test is inappropriate. In such a scenario, if you wanted to compare two groups on a continuous outcome, a parametric test like the Independent Samples T-Test would be the correct alternative.

Defining Category Levels: Two or More Options

Each of the two categorical variables being compared must have at least two levels or options. Variables like "Made a Purchase" (Yes/No), "Color Preference" (Black/White/Red/Blue), or "Educational Level" (High School/College/Graduate) all meet this requirement. The test is robust enough to handle contingency tables that are larger than 2x2 (e.g., 2x3, 3x3, or 4x5 tables). The number of options directly affects the degrees of freedom used in the Chi-Square calculation, but the core requirement remains that the variables must be able to classify the data into multiple distinct groups.

If you are working with a scenario where both variables are strictly dichotomous (only two options each, creating a 2x2 table), and you satisfy the cell count requirement of having more than 10 observations in every cell, the Chi-Square Test is suitable. However, for this specific 2x2 design, statisticians might also suggest the **Two-Proportion Z-Test** as an alternative, particularly if the sample sizes are very large, although the Chi-Square test remains the broader, more commonly applied choice.

The Requirement for Independent Samples

The assumption of **independent samples** is crucial for the Chi-Square Test of Independence. This criterion means that the data collected from one sample group must not be related in any way to the data collected from the other sample group. The two samples must be distinct, non-paired, and drawn independently from the population, ensuring that the selection of one observation does not influence the probability of selecting any other observation.

Violations of this assumption commonly occur in designs involving repeated measures, where the same individuals are measured under different conditions or at different time points (e.g., pre-test and post-test data). In such cases, the observations are dependent because the outcome for the first measurement influences the outcome for the second measurement. If your data involves repeated measures or paired observations from a single sample, you must avoid the Chi-Square Test of Independence and instead consider using specific non-parametric tests designed for related samples, such as the McNemar Test, which correctly handles the dependency structure within the data.

The Critical Requirement of Sufficient Cell Counts

The reliability of the Chi-Square approximation to the theoretical Chi-Square distribution heavily depends on having an adequate sample size, specifically quantified through the expected frequencies in the contingency table. The widely accepted rule-of-thumb recommends that the expected frequencies (not the observed ones) in every single cell of your contingency table must be greater than or equal to 10. If this condition is not met, the Chi-Square distribution may not

accurately model the sampling distribution of the test statistic, leading to inflated Type I error rates.

The term "cell" refers simply to the count of observations falling into a unique combination of categories for the two variables. For example, if I analyze survey responses comparing Gender (Male/Female) and Voting Status (Voted/Did Not Vote), the cell for "Male and Voted" must have an expected count of 10 or more. If the cell counts are too low, the shape of the sampling distribution of the test statistic deviates significantly from the theoretical Chi-Square distribution, leading to inaccurate p-values and unreliable hypothesis testing conclusions.

If your data fails this crucial assumption--meaning you have fewer than 10 expected observations in one or more cells--you must opt for an alternative method. For smaller sample sizes, especially in 2x2 tables, we recommend using Fisher's Exact Test, which provides an exact p-value calculation without relying on the Chi-Square approximation. Furthermore, if you have very large datasets (e.g., more than 1000 total observations) where the cell counts are sufficiently large, the G-Test (or likelihood ratio test) is often preferred for its superior statistical properties, though the Chi-Square test remains highly robust.

Applied Example: Analyzing Treatment Efficacy

To illustrate the practical application of the Chi-Square Test of Independence, consider a clinical trial designed to compare the efficacy of two distinct medical interventions, labeled Treatment A and Treatment B, in treating a specific disease. The structure of the variables is as follows:

Primary Grouping Variable: Treatment Type (A or B)

Outcome Variable: Disease Status (**Recovered (Yes)** or **Not Recovered (No)**)

The core research question driving the study is straightforward: Is there a statistically significant association between the type of treatment received and the patient's recovery status? In other words, do the recovery rates differ significantly between the two treatment groups? This design produces a 2x2 contingency table, provided we assume all necessary assumptions regarding sampling and cell counts have been verified.

The formal structure of the hypothesis test begins with the establishment of the null hypothesis (H_0) and the alternative hypothesis (H_a). The null hypothesis states that there is no difference in the proportion of recovery between Treatment A and Treatment B; that is, the variables are independent. Conversely, the alternative hypothesis states that there is a significant association between treatment type and recovery status. Given that both variables are categorical variables with two options each, and assuming our data meets all other critical assumptions, the Chi-Square Test of Independence is the correct analytical choice.

Upon executing the analysis, the procedure yields a **Chi-Square statistic** and an associated

probability value (p-value). The p-value quantifies the probability of observing the data distribution found in your study--or an even more extreme distribution--if the null hypothesis were actually true (meaning recovery rates were truly identical). The standard convention in statistical research is to set a significance level (α) at 0.05. If the calculated p-value is less than or equal to 0.05, we conclude that the results are statistically significant. This allows us to reject the null hypothesis and confidently assert that the observed difference in recovery rates between Treatment A and Treatment B is highly unlikely to be due merely to random sampling variation, indicating a real and meaningful association between the treatment and the clinical outcome.

ARABPSYCHOLOGY.COM