NO.1.

What statistical hypothesis test can be applied to two samples of categorical data with the same categories to determine if the frequency distributions are significantly different?

A Pearson’s chi-square test is a [statistical test](https://www.scribbr.com/statistics/statistical-tests/) for categorical data. It is used to determine whether your data are significantly different from what you expected. There are two types of Pearson’s chi-square tests:

The [chi-square goodness of fit test](https://www.scribbr.com/statistics/chi-square-goodness-of-fit/)is used to test whether the frequency distribution of a categorical variable is different from your expectations.

The [chi-square test of independence](https://www.scribbr.com/statistics/chi-square-test-of-independence/)is used to test whether two categorical variables are related to each other.

Chi-square is often written as Χ2 and is pronounced “kai-square” (rhymes with “eye-square”). It is also called chi-squared.

What is a chi-square test?

Pearson’s chi-square (Χ2) tests, often referred to simply as chi-square tests, are among the most common [nonparametric tests](https://www.scribbr.com/statistics/statistical-tests/#nonparametric). Nonparametric tests are used for data that don’t follow the [assumptions of parametric tests](https://www.scribbr.com/frequently-asked-questions/assumptions-of-statistical-tests/), especially the assumption of a [normal distribution](https://www.scribbr.com/statistics/normal-distribution/).

If you want to [test a hypothesis](https://www.scribbr.com/statistics/hypothesis-testing/) about the distribution of a [categorical variable](https://www.scribbr.com/methodology/types-of-variables/#quantitative-vs-categorical) you’ll need to use a chi-square test or another nonparametric test. Categorical variables can be [nominal](https://www.scribbr.com/statistics/nominal-data/) or [ordinal](https://www.scribbr.com/statistics/ordinal-data/) and represent groupings such as species or nationalities. Because they can only have a few specific values, they can’t have a normal distribution.

Note: Parametric tests can’t test hypotheses about the distribution of a categorical variable, but they can involve a categorical variable as an [independent variable](https://www.scribbr.com/methodology/independent-and-dependent-variables/#independent) (e.g., [ANOVAs](https://www.scribbr.com/statistics/one-way-anova/)).

Test hypotheses about frequency distributions

There are two types of Pearson’s chi-square tests, but they both test whether the observed [frequency distribution](https://www.scribbr.com/statistics/frequency-distributions/)of a categorical variable is significantly different from its expected frequency distribution. A frequency distribution describes how observations are distributed between different groups.

Frequency distributions are often displayed using [frequency distribution tables](https://www.scribbr.com/statistics/frequency-distributions/#how-to). A frequency distribution table shows the number of observations in each group. When there are two categorical variables, you can use a specific type of frequency distribution table called a contingency table to show the number of observations in each combination of groups.

Example: Bird species at a bird feeder

| Frequency of visits by bird species at a bird feeder during a 24-hour period | |
| --- | --- |
| Bird species | Frequency |
| House sparrow | 15 |
| House finch | 12 |
| Black-capped chickadee | 9 |
| Common grackle | 8 |
| European starling | 8 |
| Mourning dove | 6 |

A chi-square test (a [chi-square goodness of fit test](https://www.scribbr.com/statistics/chi-square-goodness-of-fit/)) can test whether these observed frequencies are significantly different from what was expected, such as equal frequencies.

Example: Handedness and nationality

|  |  |  |
| --- | --- | --- |
| Contingency table of the handedness of a sample of Germans and Dutch | | |
|  | Right-handed | Left-handed |
| Germans | 236 | 19 |
| Dutch | 157 | 16 |

A chi-square test (a test of independence) can test whether these observed frequencies are significantly different from the frequencies expected if handedness is unrelated to nationality.

The chi-square formula

Both of Pearson’s chi-square tests use the same formula to calculate the [test statistic](https://www.scribbr.com/statistics/test-statistic/), chi-square (Χ2):

\begin{equation*} X^2=\sum{\frac{(O-E)^2}{E}} \end{equation*}

Where:

Χ2 is the chi-square test statistic

Σ is the summation operator (it means “take the sum of”)

*O* is the observed frequency

*E* is the expected frequency

The larger the difference between the observations and the expectations (*O* − *E*in the equation), the bigger the chi-square will be. To decide whether the difference is big enough to be statistically significant, you compare the chi-square value to a critical value.

When to use a chi-square test

A Pearson’s chi-square test may be an appropriate option for your data if all of the following are true:

You want to test a hypothesis about one or more categorical variables. If one or more of your variables is quantitative, you should use a different [statistical test](https://www.scribbr.com/statistics/statistical-tests/). Alternatively, you could convert the quantitative variable into a categorical variable by separating the observations into intervals.

The sample was randomly selected from the [population](https://www.scribbr.com/methodology/population-vs-sample/).

There are a minimum of five observations expected in each group or combination of groups.

Types of chi-square tests

The two types of Pearson’s chi-square tests are:

[Chi-square goodness of fit test](https://www.scribbr.com/statistics/chi-square-goodness-of-fit/)

[Chi-square test of independence](https://www.scribbr.com/statistics/chi-square-test-of-independence/)

Mathematically, these are actually the same test. However, we often think of them as different tests because they’re used for different purposes.

Chi-square goodness of fit test

You can use a [chi-square goodness of fit test](https://www.scribbr.com/statistics/chi-square-goodness-of-fit/)when you have one categorical variable. It allows you to test whether the frequency distribution of the categorical variable is significantly different from your expectations. Often, but not always, the expectation is that the categories will have equal proportions.

Example: Hypotheses for chi-square goodness of fit test Expectation of equal proportions

[Null hypothesis](https://www.scribbr.com/statistics/null-and-alternative-hypotheses/#definition): The bird species visit the bird feeder in equal proportions.

[Alternative hypothesis](https://www.scribbr.com/statistics/null-and-alternative-hypotheses/#alternative): The bird species visit the bird feeder in different proportions.

Expectation of different proportions

Null hypothesis: The bird species visit the bird feeder in the same proportions as the average over the past five years.

Alternative hypothesis: The bird species visit the bird feeder in different proportions from the average over the past five years.

Chi-square test of independence

You can use a chi-square test of independence when you have two categorical variables. It allows you to test whether the two variables are related to each other. If two variables are independent (unrelated), the probability of belonging to a certain group of one variable isn’t affected by the other [variable](https://www.scribbr.com/methodology/types-of-variables/).

Example: Chi-square test of independence

Null hypothesis: The proportion of people who are left-handed is the same for Americans and Canadians.

Alternative hypothesis: The proportion of people who are left-handed differs between nationalities.

NO.2.

The concept of hypothesis Lesting is statistical tern largely introduced by Fisher, & Ney 7 & Pearson in the early 20th Century, in an attempt to standardise statistical metrics, Explain What you understand by hypothesis testing in brief. Briefly a hypothesis you would like to test in a given machine learning or data science project of your choice.

Land grabbing is the large-scale acquisition of land through buying or leasing of large pieces of land by domestic and transnational companies, governments, and individuals. It is often driven by the desire to access natural resources, such as water, minerals, and fertile land for agriculture.

Or

Land grabbing refers to the acquisition of large tracts of land, often by governments, corporations, or wealthy individuals, without the consent or fair compensation to the local communities or small-scale landholders who traditionally inhabit or use the land. It typically involves the displacement of local people, often indigenous communities or small farmers, from their homes and livelihoods.

Land grabbing can have a number of negative consequences, including:

Displacing local communities from their traditional land and livelihoods

Depriving people of access to food and water

Damaging the environment

Escalating social tensions and conflict

Land grabbing is a complex issue with no easy solutions. However, there are a number of things that can be done to address it, such as:

Strengthening land rights and tenure security

Promoting sustainable agriculture and natural resource management

Increasing transparency and accountability in land deals

Supporting civil society organizations working to combat land grabbing

Here are some examples of land grabbing:

A foreign investment company buys up thousands of hectares of land in a developing country to grow biofuels.

A government leases land to a mining company to extract minerals, without the consent of the local community.

A wealthy individual buys up large tracts of land to build a luxury resort.

Land grabbing is a global problem, but it is particularly prevalent in developing countries with weak land governance systems. It is estimated that between 2006 and 2012, over 200 million hectares of land was grabbed worldwide.

It is important to note that not all land acquisitions are considered land grabbing. For example, if a farmer buys land from another farmer with the farmer's consent, this is not considered land grabbing. However, if a government or company forces farmers to sell their land at a low price, or if they take over land without the consent of the local community, this is considered land grabbing.

Collecting data on land grabbing.

There are a number of different sources of data on land grabbing. Some of the most important sources include:

The Land Matrix Global Observatory: This is a global land monitoring initiative that collects data on large-scale land deals. The Land Matrix database includes information on over 1,000 land deals covering over 100 million hectares of land in 106 countries.

GRAIN: This is an international non-profit organization that works to support small-scale farmers and peasant communities. GRAIN maintains a database of land grabs, which includes information on over 400 recent, large-scale land grabs by foreign investors for the production of food crops.

The Oakland Institute: This is a policy research institute that focuses on issues of food, land, and water justice. The Oakland Institute has published a number of reports on land grabbing, including a report on the global scale of land grabbing and a report on the impact of land grabbing on women and indigenous peoples.

Other sources of data on land grabbing include:

Government reports

Academic studies

Civil society reports

News articles

It is important to note that there is no single, comprehensive database of land grabs. This is because land grabbing is a complex issue with no easy definition. Additionally, land grabbing deals are often opaque and difficult to track.

Despite these challenges, there is a growing body of data on land grabbing. This data provides valuable insights into the scale and scope of land grabbing, as well as its impact on local communities and the environment.

Here is a summary of some of the key findings from the data on land grabbing:

Land grabbing is a global problem, but it is particularly prevalent in developing countries with weak land governance systems.

The most common targets of land grabbing are agricultural land, forests, and mining sites.

Land grabbing is often driven by foreign investors, but domestic investors and governments are also involved.

Land grabbing can have a number of negative consequences, including displacing local communities, depriving people of access to food and water, damaging the environment, and escalating social tensions and conflict.

The data on land grabbing is an important tool for raising awareness of this issue and advocating for policies to protect local communities and their land rights

Here are some ways to explore data on land grabbing:

Use the Land Matrix Global Observatory database. The Land Matrix database is a great resource for exploring data on land grabbing at the global, regional, and national levels. The database includes information on over 1,000 land deals covering over 100 million hectares of land in 106 countries. You can use the database to filter and sort data by a variety of criteria, such as country, region, target sector, and investor type.

Use GRAIN's land grab database. GRAIN's land grab database is a good resource for exploring data on recent, large-scale land grabs by foreign investors for the production of food crops. The database includes information on over 400 land grabs in over 70 countries. You can use the database to filter and sort data by country, region, target crop, and investor type.

Use the Oakland Institute's reports on land grabbing. The Oakland Institute has published a number of reports on land grabbing, which include valuable data on the scale and scope of land grabbing, as well as its impact on local communities and the environment.

Use other sources of data, such as government reports, academic studies, civil society reports, and news articles. There is a growing body of data on land grabbing from a variety of sources. You can use this data to explore different aspects of land grabbing, such as the trends in land grabbing over time, the impact of land grabbing on different sectors of the economy, and the responses of governments and communities to land grabbing.

Once you have collected some data on land grabbing, you can use a variety of tools and methods to explore it. For example, you can use data visualization tools to create charts and graphs that illustrate the data in different ways. You can also use statistical analysis tools to identify patterns and trends in the data.

Here are some examples of questions that you can explore using data on land grabbing:

What are the countries that are most affected by land grabbing?

What are the sectors that are most targeted by land grabbing?

What are the types of investors that are involved in land grabbing?

What are the impacts of land grabbing on local communities and the environment?

What are the responses of governments and communities to land grabbing?

By exploring data on land grabbing, you can gain a deeper understanding of this complex issue and its impacts. You can also use this information to advocate for policies and practices that protect local communities and their land rights.

To analyse data on land grabbing, you can use a variety of statistical methods. Some common methods include:

Descriptive statistics: Descriptive statistics can be used to summarize the data and identify patterns and trends. For example, you could calculate the mean, median, and mode of land acquisition sizes in different countries or regions. You could also create a histogram to visualize the distribution of land acquisition sizes.

Inferential statistics: Inferential statistics can be used to draw conclusions about the population based on a sample of data. For example, you could use a t-test to compare the mean land acquisition sizes in two different countries. You could also use a chi-squared test to test for an association between land grabbing and deforestation.

Regression analysis: Regression analysis can be used to identify the relationship between different variables. For example, you could use a regression model to predict the impact of land grabbing on food security in a given country.

You can also use machine learning algorithms to analyse data on land grabbing. Machine learning algorithms can be used to identify patterns and trends in data that are difficult to detect using traditional statistical methods. For example, you could use a machine learning algorithm to identify the factors that are most likely to lead to land grabbing in a given region.

Once you have analysed the data, you can use the findings to write a report or article, or to create a presentation or infographic. You can also use the findings to advocate for policies and practices that protect local communities and their land rights.

Here are some examples of research questions that you could analyse using data on land grabbing:

What are the factors that are most likely to lead to land grabbing in a given region?

What are the impacts of land grabbing on food security?

What are the impacts of land grabbing on poverty and inequality?

What are the impacts of land grabbing on the environment?

What are the most effective ways to prevent and mitigate the negative impacts of land grabbing?

By analysing data on land grabbing, you can gain a deeper understanding of this complex issue and its impacts. You can also use this information to advocate for policies and practices that protect local communities and their land rights.

Here are some of the ways on how to communicate the data of land grabbing:

Use simple language and clear visuals. Avoid using technical jargon and acronyms that your audience may not understand. Instead, use simple language and clear visuals to explain the data in a way that is easy to understand.

Tell a story with the data. Use the data to tell a story about the impact of land grabbing on local communities and the environment. This will help to make the data more relatable and engaging for your audience.

Focus on the human impact. Land grabbing is a complex issue with a wide range of impacts. However, it is important to focus on the human impact of land grabbing in your communication. This will help to make the issue more personal and relevant for your audience.

Use a variety of communication channels. There are many different ways to communicate data. You can use traditional channels, such as reports and articles, or you can use more creative channels, such as infographics, videos, and social media. Choose the channels that are most likely to reach your target audience.

Here are some examples of how you can communicate the data of land grabbing:

Create an infographic. Infographics are a great way to communicate complex data in a visually appealing and easy-to-understand way. You can use an infographic to illustrate the scale and scope of land grabbing, the impacts of land grabbing on local communities and the environment, or the responses of governments and communities to land grabbing.

Write a blog post or article. You can write a blog post or article to explain the data on land grabbing in more detail. In your blog post or article, you can discuss the key findings from the data, the implications of the data, and what can be done to address the issue of land grabbing.

Create a video. Videos are a powerful way to communicate data and tell stories. You can create a video to tell the story of a community that has been affected by land grabbing, or to explain the complex issue of land grabbing in a way that is easy to understand.

Use social media. Social media is a great way to reach a large audience and share information about land grabbing. You can use social media to share infographics, articles, videos, and other resources about land grabbing.

By communicating the data of land grabbing in a clear and engaging way, you can help to raise awareness of this important issue and advocate for policies and practices that protect local communities and their land rights.