Classical Tests and Contingency Tables
UNIT 4

Dr. Praveen Barapatre

Classical Tests and Contingency Tables

- Classical tests are a collection of statistical procedures used to analyze data and test hypotheses.
- Contingency tables, on the other hand, are a way to organize categorical data.
- They are often used in conjunction with classical tests to investigate relationships between two or more categorical variables.

Contingency Tables

- A contingency table is a grid-like table that shows the frequency distribution of two or more categorical variables.
- It allows you to see how often each combination of categories occurs in your data.

Contingency Table			
	Boy	Girl	Sum
like Snickers	43	30	73
doesn't like Snickers	8	19	27
Sum	51	49	100

Can you be absolutely certain that boys like Snickers more than girls do, from this table?

- No, you can’t. The reason is that you have to ask a massive number of people before you can say anything for certain about these kinds of matters. You should at least ask 1000 people to get a decent impression of who likes “Snickers” more.
- You also have to make sure you ask similar amounts of boys and girls. If you ask 900 boys and 100 girls, that’s not good enough to be sure.

Contingency Table

- A contingency table presents a good overview of probabilities and data from complex experiments. The table is easy to read and extract information from, for further calculations. In addition to that, contingency tables are useful when you’ve been given some values, but not all of them. The missing ones can then be obtained by using the values you’ve already got in the contingency table.

Goodness of fit tests

- Goodness-of-fit tests are a type of classical test used to assess how well a sample data set matches a hypothesized probability distribution.

Purpose:

- Determine if a sample comes from a population with a specific probability distribution (e.g., normal distribution, binomial distribution).
- Assess how well a model fits the observed data.

Functionality:

- Compares observed frequencies (counts) in your data to expected frequencies under the assumed distribution.
- Calculates a test statistic based on these differences.
- Uses the test statistic and a chi-square distribution to determine if the observed discrepancies are likely due to random chance or indicate a significant difference

Anderson-Darling

- The Anderson-Darling test is a specific type of goodness-of-fit test used to assess how well a sample data set follows a

What it Does:

- The Anderson-Darling test evaluates how closely your data resembles a normal distribution.
- It accomplishes this by comparing the observed cumulative distribution function (CDF) of your data with the theoretical CDF of a normal distribution.

<u>How it Works:</u>

- Order your data: Arrange your data points from smallest to largest.
- Calculate weights: The Anderson-Darling test assigns more weight to discrepancies in the tails of the distribution compared to the center. This is unlike the Kolmogorov-Smirnov test, another goodness-of-fit test, which gives equal weight to all data points.
- Compare the CDFs: The test calculates a statistic based on the weighted squared differences between the observed CDF of your data and the expected CDF of a normal distribution.
- Evaluate the statistic: A higher statistic indicates a larger discrepancy between your data and the normal distribution. You can compare this statistic to critical values from tables or software to determine if the observed difference is statistically significant.

- The Anderson-Darling test is more powerful than the Kolmogorov-Smirnov test for detecting deviations from normality, especially in the tails of the distribution.
- It's particularly useful when you suspect non-normality but the deviation might be subtle.
- The test is most effective when the parameters of the normal distribution (mean and standard deviation) are known. However, it can still be used for estimations with some adjustments.

Chi-square test

- The chi-square test, also denoted by χ^2 (chi-squared), is a fundamental statistical test used for analysing categorical data.
- Independence: Whether two categorical variables are related or independent of each other.
- Goodness-of-fit: How well your observed data matches an expected distribution (often used for categorical data).

Chi-Square Test of Independence

This is the most common application of the chi-square test. It's used to determine if there's a statistically significant association between two categorical variables.

Here's how it works:

- You set up a contingency table showing the frequency distribution of both variables.
- You define the null hypothesis (H_0) that the two variables are independent.
- You calculate the expected frequencies for each category under the assumption of independence.
- You compare the observed frequencies with the expected frequencies using the chi-square statistic (χ^2).
- Based on the chi-square statistic and a chi-square distribution table (or software), you determine the p-value, which represents the probability of observing such a discrepancy due to random chance.
- If the p-value is less than a significance level (e.g., 0.05), you reject the null hypothesis and conclude that there's a statistically significant association between the two variables.

Chi-Square Goodness-of-Fit Test

This application of the chi-square test is used to assess how well your observed data fits a specific expected distribution (e. g., uniform Distribution, binomial distribution). Here's the process:

- You define the expected distribution and calculate the expected frequencies for each category.
- You compare the observed frequencies in your data with the expected frequencies using the chi-square statistic (χ^2).
- Similar to the test of independence, you calculate the p-value based on the chi-square statistic and a chi-square distribution.
- A low p-value indicates that the observed data deviates significantly from the expected distribution.

- The chi-square test is most reliable with large samples and sufficient expected frequencies in each category (usually a minimum of 5).
- It's a non-parametric test, meaning it doesn't assume a specific underlying distribution for the data itself (unlike tests like t-tests).
- The chi-square test only tells you whether there's a statistically significant association or deviation, not the nature or direction of the relationship.