The document "Module 04 - The Tests.ipynb - Colab.pdf" provides a comprehensive overview of various statistical tests used in hypothesis testing. These tests help determine if there are significant differences between groups or if observed patterns are likely due to chance.

Here's a breakdown of the tests covered in the document, along with their definitions and use cases:

**1. T-Test**

* Definition: A statistical test used to compare the means of two groups. It determines if the difference between the means is statistically significant or likely due to random variation.
* Use Case:
  + One-Sample t-Test: To test if a sample mean is significantly different from a known or hypothesized population mean.
  + Independent Two-Sample t-Test: To test if there is a significant difference between the means of two independent groups.
  + Paired t-Test: To test if there is a significant difference between the means of two dependent or related groups (not covered in the document).

**2. Levene's Test**

* Definition: A statistical test used to assess the equality of variances across different groups or samples. It helps determine if the assumption of equal variances is met for tests like ANOVA.
* Use Case: To check the assumption of equal variances before performing tests like ANOVA or t-tests that assume equal variances.

**3. Shapiro-Wilk Test**

* Definition: A statistical test used to assess if a dataset follows a normal distribution. It helps determine if the normality assumption is met for tests that require normally distributed data.
* Use Case: To check the normality assumption before performing tests like t-tests or ANOVA that assume normality.

**4. Kolmogorov-Smirnov Test**

* Definition: A non-parametric test used to compare a sample with a reference probability distribution (one-sample K-S test) or to compare two samples (two-sample K-S test). It assesses if the samples come from the same distribution.
* Use Case: To test if a dataset follows a specific distribution or if two datasets come from the same distribution, especially when normality assumptions may not be met.

**5. Chi-Square Test for Independence**

* Definition: A statistical test used to determine if there is a significant association between two categorical variables. It analyzes the frequencies in a contingency table to assess if the observed differences between groups are likely due to chance.
* Use Case: To test relationships between categorical variables, such as the association between gender and political affiliation or between education level and product preference.

**6. ANOVA (Analysis of Variance)**

* Definition: A statistical test used to compare the means of three or more groups. It assesses if there are statistically significant differences between the means of the groups or if the observed differences are due to random variation.
* Use Case: To compare the effectiveness of different treatments, the performance of different groups on a test, or the impact of different factors on a response variable.

**7. Kruskal-Wallis Test**

* Definition: A non-parametric test used to compare the distributions of three or more groups. It is an alternative to ANOVA when the data does not meet the assumptions of normality or equal variances.
* Use Case: To compare groups when the data is ordinal or when the assumptions of ANOVA are violated.

**8. Fisher's Exact Test**

* Definition: A statistical test used to determine if there are non-random associations between two categorical variables, especially when sample sizes are small. It calculates the exact probability of observing the data in a contingency table, assuming no association between the variables.
* Use Case: To analyze contingency tables with small sample sizes where the chi-square test may not be appropriate.

**9. Z-Test**

* Definition: A statistical test used to compare a sample mean to a known population mean when the population standard deviation is known and the data is normally distributed or the sample size is large.
* Use Case: To test hypotheses about population means when the population standard deviation is known.

**10. Welch's t-Test**

* Definition: A variation of the independent t-test used when the variances of the two groups being compared are not assumed to be equal. It adjusts the degrees of freedom and provides a more accurate test in cases of unequal variances.
* Use Case: To compare the means of two independent groups when the assumption of equal variances is violated.

**11. Mann-Whitney U Test**

* Definition: A non-parametric test used to compare the distributions of two independent groups. It assesses if one group tends to have larger or smaller values than the other.
* Use Case: To compare groups when the data is ordinal or when the assumptions of the t-test are violated.