Statistical Tests

Types of Statistical Tests

Statistics are the arrangement of statistical tests which analysts use to make inference from the data given. These tests enables us to make decisions on the basis of observed pattern from data. There is a wide rangeof statistical tests. The choice of which statistical test to utilize relies upon the structure of data, the distribution of the data, and variable type.There are many different types of tests in statistics like t-test,Z-test,chi-square test, anova test ,binomial test, one sample median test etc.

When to perform a statistical test

You can perform statistical tests on data that have been collected in a statistically valid manner – either through an experiment or through observations made using probability sampling methods.

For a statistical test to be valid, your sample size needs to be large enough to approximate the true distribution of the population being studied.

To determine which statistical test to use, you need to know:

- whether your data meets certain assumptions.
- the types of variables that you're dealing with.

Choosing a Statistical test

parametric statistical test:

Parametric tests are used if the data is normally distributed .A **parametric statistical test** makes an assumption about the population parameters and the distributions that the data came from. These types of test includes **t-tests,z-tests** and **anova tests**, which assume data is from normal distribution.

Z-test:

A z-test is a statistical test used to determine whether two population means are different when the variances are known and the sample size is large. In z-test mean of the population is compared.The parameters used are population mean and population standard deviation. Z-test is used to validate a hypothesis that the sample drawn belongs to the same population.

Ho: Sample mean is same as the population mean(Null hypothesis)

Ha: Sample mean is not same as the population mean(Alternate hypothesis)

z = (x − μ) / (σ / √n),

where , x=sample mean, u=population mean, σ / √n = population standard deviation.

If z value is less than critical value accept null hypothesis else reject null hypothesis.

T-test:

In t-test the mean of the two given samples are compared. A t-test is used when the population parameters (mean and standard deviation) are not known.

Paired T-Test:

Tests for the difference between two variables from the same population(pre- and post test score). For example- In a training program performance score of the trainee before and after completion of the program.

Independent T-test:

 The independent t-test which is also called the two sample t-test or student's t-test, is a statistical test that determines whether there is a statistically significant difference between the means in two unrelated groups.For example -comparing boys and girls in a population.

One sample t-test:

 The mean of a single group is compared with a given mean. For example-to check the increase and decrease in sales if the average sales is given.

t = (x1 − x2) / (σ / √n1 + σ / √n2),

where x1 and x2 are mean of sample 1 and sample 2 respectively.

ANOVA Test:

Analysis of variance (ANOVA) is a statistical technique that is used to check if the means of two or more groups are significantly different from each other. ANOVA checks the impact of one or more factors by comparing the means of different samples. If we use a t-test instead of ANOVA test it won't be reliable as number of samples are more than two and it will give error in the result.

The hypothesis being tested in ANOVA is

Ho: All pairs of samples are same i.e. all sample means are equal

Ha: At least one pair of samples is significantly different

In anova test we calculate F value and compare it with critical value

F= ((SSE1 — SSE2)/m)/ SSE2/n-k,

where SSE = residual sum of squares, m = number of restrictions

k = number of independent variables

Non parametric statistical test:

Non parametric tests are used when data is not normally distributed. Non parametric tests include **chi-square test**.

Chi-square test(χ^2 test):

chi-square test is used to compare two categorical variables. Calculating the Chi-Square statistic value and comparing it against a critical value from the Chi-Square distribution allows to assess whether the observed frequency are significantly different from the expected frequency.

The hypothesis being tested for chi-square is-

Ho: Variable x and Variable y are independent

Ha: Variable x and Variable y are not independent.

$$\chi^2 = \sum \frac{(o-e)^2}{e}$$

Chi-square formula