# Hypothesis Testing

**Hypothesis** : Statement about the probability distribution of a random variable that we either accept or reject.

- **Left sided**: Alternative states parameter of interest is less.
- **Right sided**: Alternative states parameter of interest is greater.
- **Two sided**: Either less or greater.

**Acceptance region**: The range of values of test statistic for which we do not reject the null hypothesis.
**Rejection region**: The range of values of test statistic for which we reject the null hypothesis.

**Critical values:** Boundaries of critical and acceptance region.

**Types of error:**

1. Type I error: Null hypothesis is TRUE, but is rejected. Its probability is denoted by $\alpha$ a.k.a significance level.
2. Type II error: Null hypothesis is FALSE, but is accepted. Its probability is denoted by β.

Instead of these errors, it is convenient to work with power.

**Power** = 1 - β

**P-value :** Probability that the test statistic will take on a value at least as extreme as the observed value of statistic when H0 is True. It represents the smallest level of significance that would lead to the rejection of H0.

If P is less than or equal to $\alpha$ we reject the null hypothesis, else we fail to reject.

**Parametric tests:** Parametric tests are those that make assumptions about the parameters of the population distribution from which the sample is drawn.

1. ANOVA
2. T test
3. F test

**Non-parametric tests:** Non-parametric tests are those that do not make assumptions about the parameters of the population distribution from which the sample is drawn.

1. Sign test
2. Mann-Whitney U test (Wilcoxon Rank Sum test)
3. Wilcoxon Signed Rank test
4. Kruskal Wallis H test
5. Chi-Square Goodness of Fit test

1. ANOVA:

- **Purpose**: To compare the means of three or more independent samples to see if at least one sample mean is significantly different from the others.
- **Assumptions**: The data in each sample are normally distributed, the variances are equal across samples (homogeneity of variances), and the samples are independent.

2. T test:

- **Purpose**:
  - One sample: To compare the means of the sample to see if the sample mean of one is significantly different from the given value.
  - Independent samples: To compare the means of two independent samples to see if the sample mean of one is significantly different from the other.
- **Assumptions**: The data in each sample are normally distributed, the variances are equal across samples (homogeneity of variances), and the samples are independent.

3. F test:

- **Purpose:** To compare the variances of two independent samples to see if the sample variance of one is significantly different from the other.
- **Assumptions**: The data in each sample are normally distributed and the samples are independent.

4. Z test:

Similar to T-test, but population standard deviation is known. Generally used for sample size > 30.

5. Sign Test:

- **Purpose:** To determine if the median of a set of paired observations differs significantly from a given value
- **Assumptions:** The paired differences are independent of each other.

6. WIlcoxon SIgned Rank test

- **Purpose:** To test for a significant difference in paired observations.
- **Assumptions:** The paired differences are dependent, come from a continuous distribution, and are symmetrically distributed about their median.

7. Mann-Whitney U test

- **Purpose:** To determine whether there is a statistically significant difference between the medians of two independent samples.
- **Assumptions: T**he observations within each sample are independent.

8. Kruskal Wallis H test

- **Purpose: T**o determine whether there are statistically significant differences between the medians of three or more independent samples.
- **Assumptions:** The observations within each group are independent.

9. Chi-Square Goodness of Fit test

- **Purpose:** To test whether the observed frequencies of categorical data match the expected frequencies from a hypothesized distribution
- **Assumptions:** The observations are independent of each other and the sample size is sufficiently large for the chi-square approximation to be valid.

To calculate a confidence interval around the mean of data that is not normally distributed:

1. Find a distribution that matches the shape of the data and use that distribution to calculate the confidence interval.
2. Perform a transformation on the data to make it fit a normal distribution, and then find the confidence interval for the transformed

data.