Degrees Of Freedom

-Mukesh Kumar

Understanding Degrees of Freedom

- **Degrees of Freedom (df)** is a concept used in statistics to determine the number of **independent values** or quantities that can vary in an analysis without violating any constraints. It's a key component in many statistical tests, such as the Chi-Square test, t-test, and ANOVA.

- **1. Basic Example**

- Imagine you have a list of 5 numbers. If you know the mean of these numbers, and you know 4 of them, you can always calculate the 5th number. So, even though there are 5 numbers, only 4 of them can vary independently. The degrees of freedom in this case would be 4 (which is 5 - 1).

- **2. In the Context of a t-Test**

- **One Sample t-Test**: When you're comparing the mean of a single sample to a known value, the degrees of freedom are calculated as $n - 1$, where n is the number of observations in the sample. This accounts for the fact that once you know the mean of the sample, only $n - 1$ values can vary freely.
- **Two-Sample t-Test**: When comparing the means of two independent samples, the degrees of freedom are typically $n_1 + n_2 - 2$, where n_1 and n_2 are the sample sizes. This is because you're estimating two means, so you lose 2 degrees of freedom.

- **3. Chi-Square Test**

- **Goodness-of-Fit Test**: Here, the degrees of freedom are calculated as $k - 1$, where k is the number of categories. For example, if you have a distribution with 5 categories, the degrees of freedom would be $5 - 1 = 4$.
- **Test of Independence**: For a contingency table, the degrees of freedom are calculated as $(r - 1) \times (c - 1)$, where r is the number of rows and c is the number of columns. For example, in a 3x2 table, the degrees of freedom would be $(3 - 1) \times (2 - 1) = 2$.