Partitioning of variance in Anova or Hypothesis Testing in Anova:

1. Null Hypothesis :

We compare the mean of two or more groups

2. Alternate Hypothesis : Atleast one of the sample mean is not equal

We use F Test:

F =

Example:

Let’s suppose we have three samples

X1 = [1,2,3,4,5], 3

X2 = [6,7,3,2,1], 19/5

X3 = [5,6,3,2,4], 4

Mean is getting compared based on the variance that is present

Variance between samples is when we calculate the variance for each sample and find the difference between variance of different samples

Variance within sample has multiple steps involved

:

: Atleast one sample mean is not equal

One way Anova:

One factor with atleast two levels which are independent.

Q. Doctors want to test a new medication which reduces headaches. They split the participants into 3 groups [15mg, 30mg, 45mg]. Later on the doctor asks the patients to rate the headache from 1 to 10. Are there any differences between the three groups using alpha = 0.05?

Data:

|  |  |  |
| --- | --- | --- |
| 15 mg | 30 mg | 45 mg |
| 9 | 7 | 4 |
| 8 | 6 | 3 |
| 7 | 6 | 2 |
| 8 | 7 | 3 |
| 8 | 8 | 4 |
| 9 | 7 | 3 |
| 8 | 6 | 2 |

1. Defining null and alternate hypothesis:

Null Hypothesis :

Alternate Hypothesis : Not all means are equal

2. Significance value

CI = 95%

3. Degree of freedom:

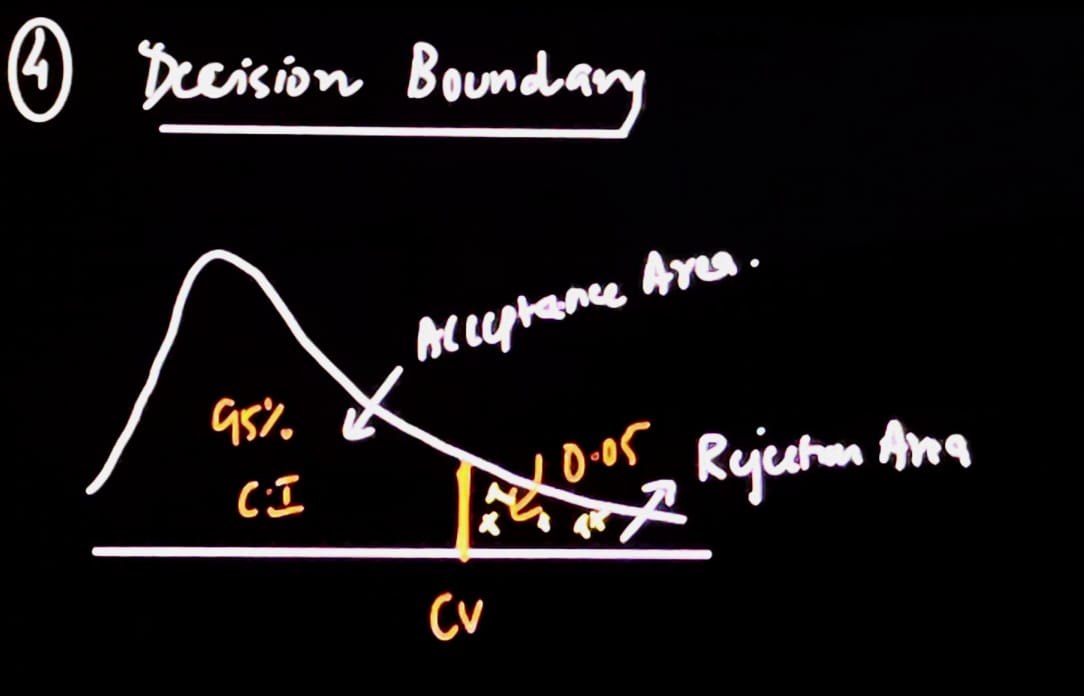
N = 21 [Total Sample size]

N = 7 [Sample size]

a = 3 [No. of categories]

(2,18) = (df1, df2), Used in f table along with significance value to find out the critical value

4. Decision Boundary:



Using f table, we will find the critical value.

Cross referencing, we get the critical value as 3.5546

Decision Rule:  
If f calculated value is greater than critical value, we reject the null hypothesis.

5. F Test Statistics:

F =

Sum of Squares needs to be calculated for between and within samples and total. It helps in finding out the variance.

We also need to calculate the degree of freedom, mean squared and finally the f value.

Data:

|  |  |  |
| --- | --- | --- |
| 15 mg | 30 mg | 45 mg |
| 9 | 7 | 4 |
| 8 | 6 | 3 |
| 7 | 6 | 2 |
| 8 | 7 | 3 |
| 8 | 8 | 4 |
| 9 | 7 | 3 |
| 8 | 6 | 2 |

I. SS Between samples = -

15 mg: 9 + 8 + 7 + 8 + 8 + 9 + 8 = 57

30 mg: 7 + 6 + 6 + 7 + 8 + 7 + 6 = 47

45 mg: 4 + 3 + 2 + 3 + 4 + 3 + 2 = 21

This is for

SS = = 98.67

Through this we are able to calculate the sum of squares, or the error in variance between the samples.

II. SS Within Samples =

= + + + + + + + + + + + + + + + + + + + +

= 853 let’s say

=853 - = 10.29

III. Total SS = 108.96

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | SS | Df | MS = SS / Df | F |
| Between | 98.67 | 2 | 49.34 |  |
| Within | 10.29 | 18 | 0.54 |  |
| Total | 108.96 | 20 |  |  |

F =

F Test = = 49.34 / 0.54 = 86.56

Since F is greater than 3.5546, we reject the null hypothesis.

Conclusion: There are differences.