F- Distribution

F- distribution → chi- square → $(\text{normal})^2$

1. Continous probability distbn: The F- distribution is a continous prob distbn used in statistical hypo testing and analysis of variance (Anova).

2. Degrees of freedom: The F- distribution is defined by two parameters the degree of freedom for the numerator (df1) and the degree of freedom for the denominator (df2).

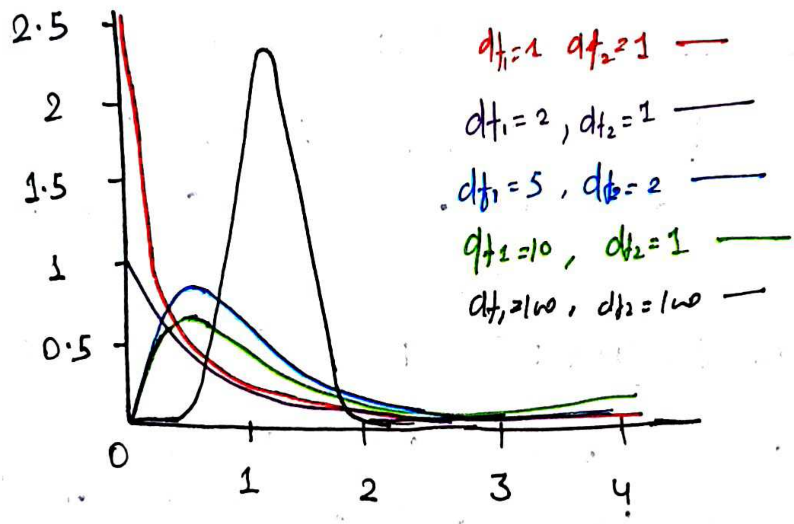
3. Positively skewed and bounded: The shape of the F- distbn is ~~commonly used to test~~ positively skewed, with its left bound at zero. The distbn's shape depends on the values of the degree of freedom.

4. Testing equality of variance: The F- distbn is commonly used to test hypo about the equality of two variances in diff samples or pop.

5. Comparing statistical models: The F- distbn is also used to compare the fit of diff statistical models particularly in the context of ANOVA.

6. F-statistics: The F- statistics is calculated by dividing the ratio of two sample variances or mean square from an ANOVA table.

This value is then compared to critical values from the f- distⁿ to determine statistical significance



$\chi_1^2 \Rightarrow df_1$

$\chi_2^2 \Rightarrow df_2$

$$F = \boxed{\frac{\chi_1^2/df_1}{\chi_2^2/df_2}} \rightarrow \text{f-dist}^n$$

One Way ANOVA test

One way ANOVA (Analysis of variance) is a statistical method used to compare the mean of three or more independent groups to determine if there are any significant diff betn them. It is an extension of the t-test, which is used for comparing the means of two independent groups. The term "one-way" refers to the fact that there is only one independent variable (factor) with multiple levels (groups) in this analysis.

The primary purpose of one way ANOVA is to test null hypo that all the group means are equal. The alternative hypo is that atleast one group mean is significance different from the others.

$\boxed{\text{t-test}} \rightarrow$ Sample $\rightarrow$ gender

Male, Female, Male $\}$ T-test only work with 2 category

* but ANOVA work with more than 2 category.
  Male, female, other

Null hypo = Mean(male) = Mean(female) = mean(other)

Alter Hypo = atleast one group mean is differend or not equal.

Mean(male) = Mean(female) ≠ mean(other)

Example 1

num	category
3	A
1	B
8	C
8	B
6	C
6	A
⋮	⋮

$\longrightarrow$

A	B	C
3	1	8
6	8	6
3	9	10

$\boxed{n=9}$ $\boxed{k=3}$

$n \downarrow$ num $k \downarrow$ category

grand mean $\rightarrow$ $\bar{\bar{x}} = \frac{54}{9}$ ← sum of all num

$\bar{\bar{x}} = 6$

$\boxed{\bar{x}_A = 4 \mid \bar{x}_B = 6 \mid \bar{x}_C = 8}$

$\boxed{\bar{\bar{x}} = 6} \rightarrow$ SST → Sum of Square Total

$(6-3)^2 + (6-6)^2 + (6-3)^2 + (6-1)^2 + (6-8)^2 + (6-9)^2 +$

$\bar{\bar{x}}$ $(6-8)^2 + (6-6)^2 + (6-10)^2 = 76$

$\boxed{SSW} \rightarrow$ Sum of Square Within

$\bar{x}_A \Rightarrow (4-3)^2 + (4-6)^2 + (4-3)^2 + \bar{x}_B \Rightarrow (6-1)^2 + (6-8)^2 +$

$(6-9)^2 + \bar{x}_C \Rightarrow (8-8)^2 + (8-6)^2 + (8-10)^2 = 52$

$\boxed{\text{degree of freedom for SSW} \Rightarrow n-k \Rightarrow 9-3 = 6}$

$\boxed{\text{degree of freedom for SST} \Rightarrow n-1 \Rightarrow 9-1 = 8}$

$\boxed{SSB} \rightarrow$ Sum of Square Between

$$= (\bar{\bar{x}} - \bar{x}_A)^2 + (\bar{\bar{x}} - \bar{x}_B)^2 + (\bar{\bar{x}} - \bar{x}_C)^2$$

$$= \cancel{(6-4)^2} + \cancel{(6-6)^2} + \cancel{(6-8)^2} =$$

A	B	C
3	1	8
6	8	6
3	9	10
3	3	3

Sample

$$\Rightarrow \text{Sample}_A (\bar{\bar{x}} - \bar{x}_A)^2 + \text{Sample}_B (\bar{\bar{x}} - \bar{x}_B)^2 + \text{Sample}_C (\bar{\bar{x}} - \bar{x}_C)^2$$

$$3 \times 4 + 0 + 3 \times 4 = 24$$

degree of freedom for SSB $= k-1 = 3-1 = 2$

quantity	value	df
SSB	24	2
SSW	52	6
SST	76	8

$\boxed{SST = SSW + SSB}$

$\downarrow$

verify

SST

$\swarrow$ Internal variance groups (SSW)

$\searrow$ $\boxed{SSB}$

$$F\text{-ratio} = \frac{\frac{SSB}{df_{ssb}}}{\frac{SSW}{df_{ssw}}} = \frac{\frac{24}{2}}{\frac{52}{6}} = 1.4$$

P_value ⇒ code import scipy.stats as stats

f_statistic = 1.4

df1 = 2

df2 = 6

p-value = stats.sf(f_static, df1, df2)

P_value = 0.31 > α (0.05)

null hypo can't rejected [$\mu_A = \mu_B = \mu_C$].

Steps

- Define the null and alternative hypothesis.
- Calculate the overall mean (grand mean) of all the groups combined and mean of all the groups individually.
- Calculate the "between group" and "within-group" sum of squares (SS)
- Find the betⁿ group and within group degree of freedoms.
- Calculate the "between-group" and "within-group" mean squares (MS) by dividing their respective sum of squares by their degrees of freedom.

- Calculate the F-statistic by dividing the "betn group" mean square by the "within group" mean square

Anova Table

Source of variate	Sum of Square (SS)	Degree of freedom	Mean Square (MS) (SS is divide by d.f) is an estimation of variance to be used in F-ratio	F- ratio
Betn sample or categories (SSB)	$n_1(\bar{x}_1-\bar{\bar{x}})^2+\dots+n_k(\bar{x}_k-\bar{\bar{x}})^2$	$(k-1)$	$\frac{SS\ bet^n}{(k-1)}$	$\frac{MS\ bet^n}{MS\ within}$
Within samples or categories (SSW)	$\Sigma(x_{i1}-\bar{x}_1)^2+\dots+$ $\Sigma(x_{ik}-\bar{x}_k)^2\ i=1,2,3$	$(n-k)$	$\frac{SS\ within}{(n-k)}$	
Total	$\Sigma(x_{ij}-\bar{\bar{x}})^2$ $i,j=1,2,3---$	$(n-1)$		

- Calculate the p- value associated with the calculated F- statistic using the F- distn and the appropriate degrees of freedom. The p-value represent the prob of obtaining an F statistic as extreme or more extreme that the calculated value, assuming the null hypo is true.

- choose α (0.05)

- If p-value is less than or equal to alpha, reject the hypo.

Assumption

1. Independence

2. Normality

3. Homogeniety of variance : pop variance of all categories is same. If not then use welch's ANOVA.

Post-hoc Test

Post hoc test, also known as post-hoc pairwise comparisons or multiple comparison tests, are used in the content of ANOVA when the overall test indicates a significant diff among the group means. These tests are performed after the initial one-way ANOVA to determine which specific groups or pair of groups have significantly diff means.

The main purpose of post hoc test is to control the family-wise error rate (FWER) and adjust the significant level for multiple comparisons to avoid inflated Type I error. There are several post hoc test available, each with diff characteristic and assumption. Some common post hoc test include:

1. Bonferroni Correction:

$$\mu_A = \mu_B = \mu_C$$

P_val ← $\boxed{AB \rightarrow t\text{-test}}$

P_val ← $\boxed{BC \rightarrow t\text{-test}}$

P_val ← $\boxed{CA \rightarrow t\text{-test}}$

2. Tukey's HSD (Honestly Significant Difference) Test:

This test controls the FWER and is used when the sample sizes are equal and the variances are assumed to be equal accross the groups. It is one of the most commonly used post-hoc test.

use code

$$\frac{0.05}{3} \rightarrow \text{category}$$

$\rightarrow \mu_A, \mu_B, \mu_C$

* Mostly Tusky's HSD used

why t-test is not used for more than 3 category's?

1. Increase Type I error :- When you perform multiple comparison using individual t-test the prob of making a type I error (false positive) increase. The more test you perform, the higher the

chance that you will incorrectly reject the null hypo in at least on of the tests, even if the null hypo is true for all groups.

2. Difficulty in interpreting results.

if you have 4 group and you perform 6 pairwise t-test. It can be challenging

3. Inefficiency

Application in Machine Learning

1. Hyperparameter tuning
2. Feature selection
3. Algorithm comparison
4. Model stability assessment