F - Test

Z-test
T-test
Parametric test

μ

$H_0: \mu = \text{value}$
$H_a: \mu \neq$ "

Two-tailed

$>$
$<$

one-tailed

- Find out whether the two independent estimates of pop variance differ significantly

OR

- Two sample drawn (randomly) from the Normal distribution having same variance.

μ, $\bar{x}$

Ratio Test

$$\boxed{F = \frac{\sigma_1^2}{\sigma_2^2} \quad \text{OR} \quad \frac{S_1^2}{S_2^2}}$$

• variance :

$$\boxed{\sigma^2 = \frac{\Sigma (x-\mu)^2}{n}}$$

N

$$F = \frac{\text{Larger estimates of var}}{\text{smaller " " "}}$$

ν : Degree of freedom. $\nu = n - 1$.

$\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1$.

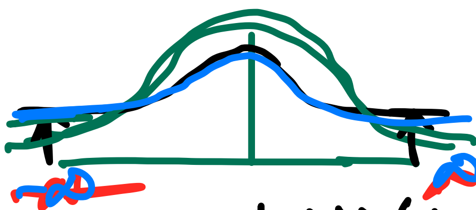
larger variance   smaller var;

$\nu_1 =$ numerator $\nu_2 =$ denominator.

H_0 = accepted.

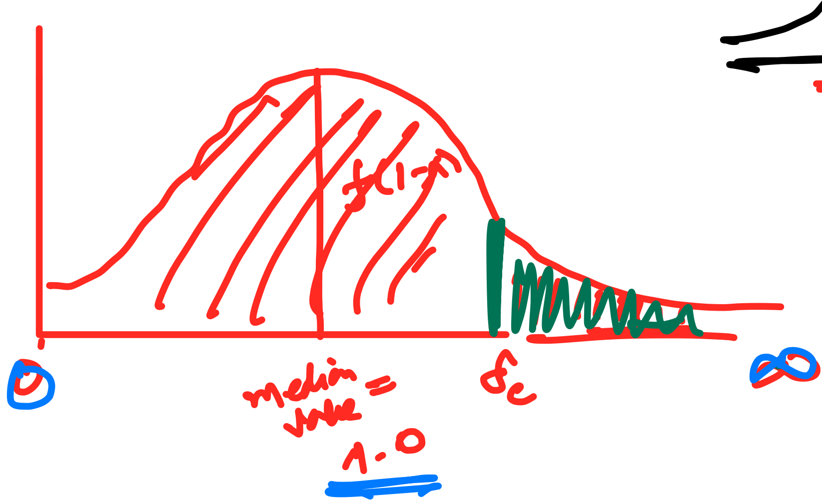
Calculated F value > Tab F-value.

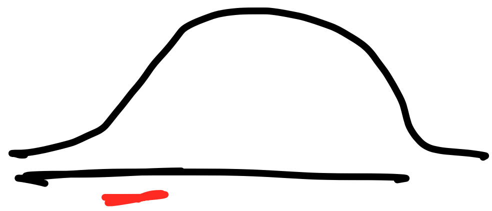
H_0 = reject



$t >>> 60$

F - distribution :



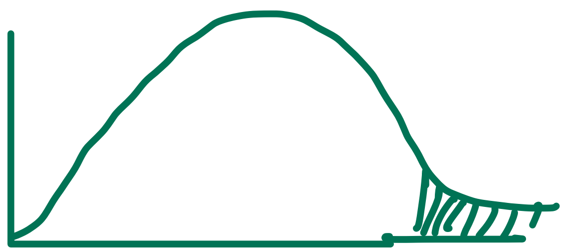


1. Skewed towards right
   range $0 \rightarrow \infty$
   median roughly $= 1$

2. Value of F will always be greater 0

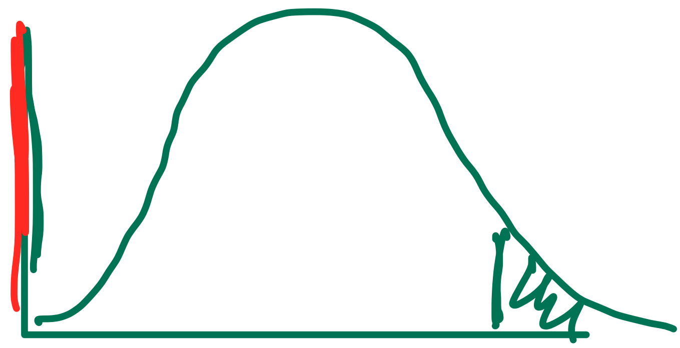
3. Shape of F-dist^n depends
   - d of (Numerator)
   - dof (denom^n)

4.



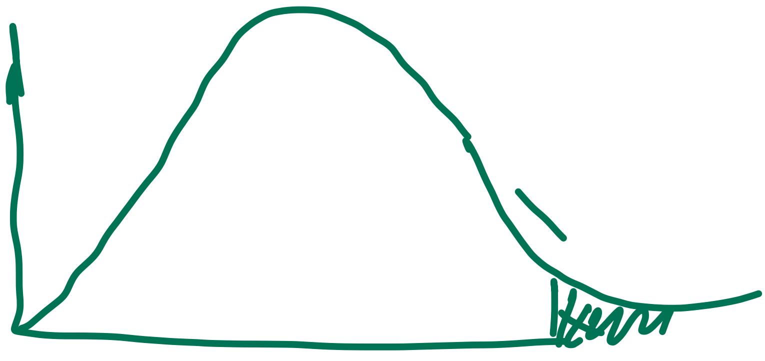
$df_1 = df_2 = 5$

$f_{0.05} = 5.05$



$df_1 = df_2 = 30$

$f_{0.05} = 1.84$



$df_1 = df_2 = 60.$

$f_{0.05} = 1.53$

$H_0: \sigma_1^2 = \sigma_2^2$; $H_a: \sigma_1^2 \neq \sigma_2^2$

Null Hypothesis

$F = \frac{S_1^2}{S_2^2}$ → A, B

$\frac{\sigma_1^2}{\sigma_2^2}$

$\frac{\sum(x-\bar{x})^2}{n-1}$

x_A	$(x_A-\bar{x}_A)$	$(x_A-\bar{x}_A)^2$	x_B	$(x_B-\bar{x}_B)$	
16	-14	196	14	-19	361
17	-13	169	16	-17	289
25	-5	25	24	-9	81
26	-4	16	28	-5	25
32	2	4	32	-1	1
34	4	16	35	2	4
38	8	64	37	4	16
40	10	100	42	9	81
42	12	144	43	10	100
			45	12	144
			47	14	196

$\sum A = 270$

$\bar{x}_A = \frac{270}{9} = 30$

$\sum(x_A-\bar{x}_A)^2 = 734$

$\sum B = \frac{363}{11} = 33$

$\sum(x_B-\bar{x}_B)^2 = 1298$

$S_2^2 = \frac{734}{8} = 91.75$

$v_1 = 10$

$v_2 = 8$

$S_1^2 = \frac{1298}{10} = 129.8$

$F = \frac{129.8}{91.75} = \boxed{1.4147}$ Fcalc

Tabulated F = 3.35

$1.414 < 3.35$

H_0 is accepted

Both the pop have same Variance

$\sigma_1^2 = \sigma_2^2$

Chi square - Test χ^2

Z-Test, F-test, t-test. →

Assumption: Samples are drawn from normally distributed population

parametric :

Non-parametric test :

- No exact info is available regarding pop;
  - binomial, poisson, normal.

χ^2 - test.

χ^2 - Test : magnitude of the descripancy b/w Theoretical & empirical /observed value.

Formula:

Observed freq → O; Expected freq. → E

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

1 2 3
3 4 7
4 6 → 10

Row Total; column total

$$E = \frac{RT \times CT}{N}$$

N → No; of observation.

$$\nu = (r-1)(c-1)$$

Condition:

1. At least 5 observation in each cell

< 5 overestimated → to rejection (high chance)

Yates

2. Independent observation & completely random

3. $N \geq 50$

4. data : original units.

% age. ✗

Applications.

1. Test for independence of attributes:

χ^2 : help

2. χ^2 to check goodness of fit

- actual sample distⁿ matches or coincide → known prob distribution.

3. χ^2. (Yate's correction)

2x2 $\qquad |O-E| - 0.25$

Date: 18/01/23

1842, 1823, 1835, 1846, 1821, 1844, 1839, 1851

1828, 1843, 1850, 1831, 1833, 1840

χ^2 Test for pop; variance

↳ χ^2: parameteric.

$$\chi^2 = \frac{S_s^2}{S_p^2} \times (n-1)$$

S_s^2 → var; of sample

$(n-1)$ → n: sample size.

S_p^2 → var; of pop.

H_0: Quinine = No effect

H_a: " ≠ No "

Treatment	Fever	E	No fever	E	Total
1. Quinine	20	30	480	470	500
2. No "	100	90	1400	1410	1500
Total	120		1880		2000

N = 2000

$$\chi^2 = \sum\left[\frac{(O-E)^2}{E}\right]; \quad E = \frac{RT \times CT}{N}$$

$N = 2000$

$$\frac{5\cancel{00} \times \cancel{120}^{6}}{\cancel{2000}} = 30, \quad \frac{\cancel{500}^{25} \times 18\cancel{80}}{\cancel{2000}_1} = 470$$

$$\frac{15\cancel{00} \times \cancel{120}^{6}}{\cancel{2000}} = 90 \qquad \frac{15\cancel{00} \times 188\cancel{0}}{2\cancel{000}} = 1410$$

- calculation of χ^2

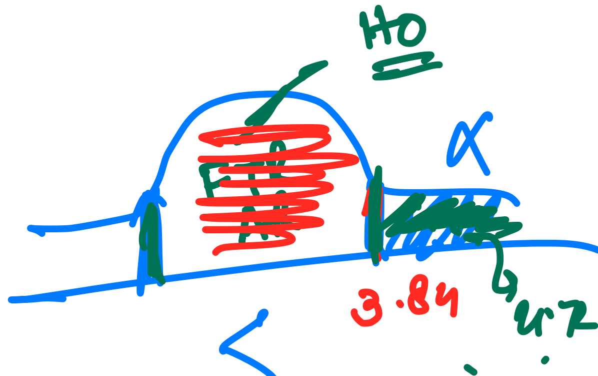
2×2

-0.25

O	E	$\lvert(O-E)\rvert$	$(O-E)^2$	$\frac{(O-E)^2}{E}$
20	30	-10	100	3.33
100	90	10	100	1.11
480	470	10	100	0.21
1400	1410	-10	100	0.07

$\Sigma = 4.72$

$$DOF = \nu = (r-1)(c-1)$$
$$(2-1)(2-1) = 1$$



Fail to reject
Accepted region

5%

$$\chi^2_{0.05} = 3.84$$

$$\chi^2_{cal} = 4.72$$

Ho Rejected

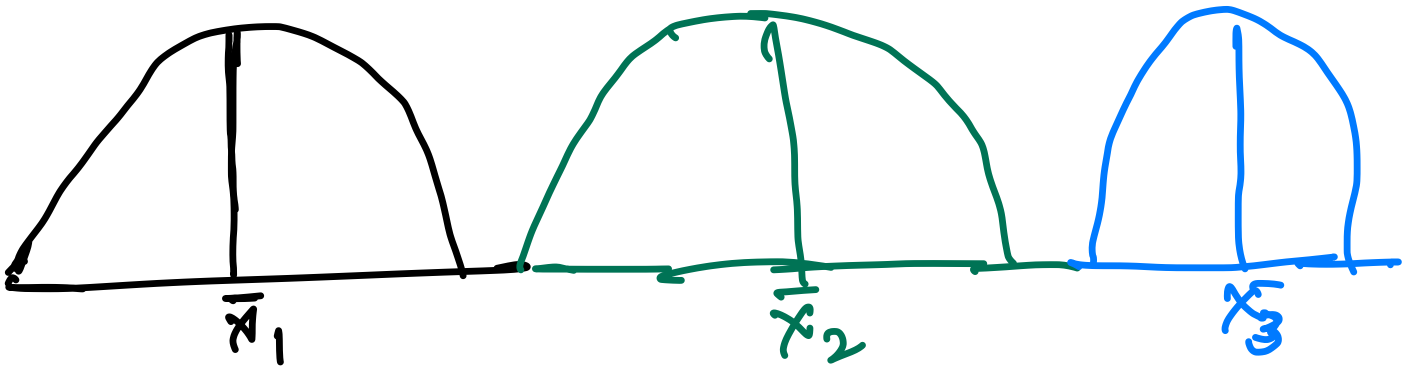
ANOVA (Analysis of Variance)

One way

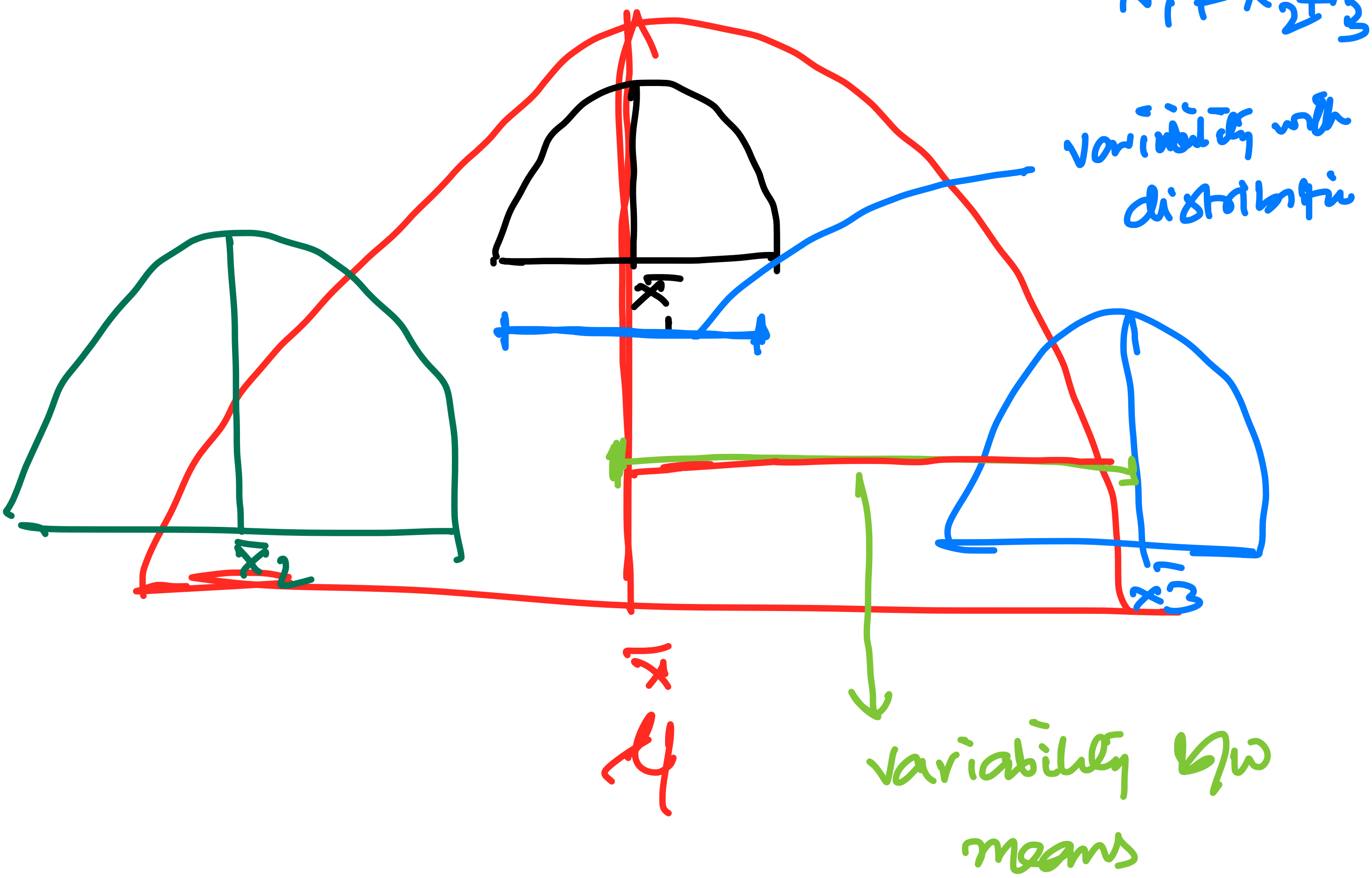
Two way

VET CECOS FAST

PESH
L&B
KHI



$\bar{x}_1 \neq \bar{x}_2 \neq \bar{x}_3$



$$\text{ANOVA} = \frac{\text{variability b/w the means (a)}}{\text{variability within the distri (b)}}$$

Total vari = $\underline{a + b}$

Assumption:

1: Normal distribution.

2: $s_1^2 = s_2^2 = s_3^2 \cdots s_r^2$

3.

$H_0: \mu_1 = \mu_2 = \mu_3 \cdots \neq \mu_r$

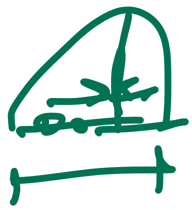
$H_a: \mu_1 \neq \mu_2 \neq \mu_3 \cdots \neq \mu_r$

One-way

Example

A	B	C
2	3	4
4	5	6
6	7	8
12	15	18

$\bar{x}_A = 4$



STEP-I:

$H_0 = \bar{x}_A = \bar{x}_B = \bar{x}_C$

$H_a = \bar{x}_A \neq \bar{x}_B \neq \bar{x}_C$

STEP II :

Calc $\boxed{\text{Variance b/w the Samples.}}$ ✓

(a) mean of each sample. ✓

$\bar{x}_A = 4$ $\bar{x}_B = 5$ $\bar{x}_C = 6$

(b) Grand mean ✓

$\bar{\bar{x}} = 5$

(c) Difference b/w the means of sample & $\bar{\bar{x}}$

$(\bar{x}_A - \bar{\bar{x}})$	$(\bar{x}_A - \bar{\bar{x}})^2$	$(\bar{x}_B - \bar{\bar{x}})$	$(\bar{x}_B - \bar{\bar{x}})^2$	$(\bar{x}_C - \bar{\bar{x}})$	$(\bar{x}_C - \bar{\bar{x}})^2$
-1	1	0	0	1	1
-1	1	0	0	1	1
-1	1	0	0	1	1
	3		0		3

$\sum (3+3) \rightarrow 6$ ✓

III: $\boxed{\text{variance within the sample}}$

(a): $\bar{X}_A = 4$ $\bar{X}_B = 5$ $\bar{X}_C = 6$

(b) deviation of items in sample from mean

$(A-\bar{X}_A)$	$(A-\bar{X}_A)^2$	$(B-\bar{X}_B)$	$(B-\bar{X}_B)^2$	$(C-\bar{X}_C)$	$(C-\bar{X}_C)^2$
-2	4	-2	4	-2	4
0	0	0	0	0	0
2	4	2	4	2	4
	8		8		8

$8 + 8 + 8 = 24$

Ratio of F

Source of variance	Sum of Squares	ν	Mean Sum of squares	F
B/w Sample	SSC 6	$\nu_1 = c-1$ 2	MSC	
within "	SSE = 24	$\nu_2 = n-c$ 6	MSE	

$$MSC = SSC/c-1 = \frac{SSC}{\nu_1} \qquad 6/2 = 3$$

$$MSE = SSE/n-c = \frac{SSE}{\nu_2} \qquad \frac{24}{6} = 4$$

$$F = \frac{MSC}{MSE} = 3/4 = 0.75$$

F Table

5.14

Ho accepted.

Data:

DAYS	A	B	C	D
MON	2	3	4	5
TUE	4	5	6	7
WED	6	7	8	9

Source of Variance	Sum of squares	DOF	Mean of squares	
B/w the Columns	SSC 15	$\nu_1 = c-1 = \boxed{3}$	$MSC = \frac{SSC}{\nu_1} = 5$	MSC/MSR F = 4.7
B/w the rows	SSE 32	$\nu_2 = (r-1) = 2$	$MSE = \frac{SSE}{\nu_2}$ 16	MSE/MSR = 0
Residual or Error	SSR — 0	$\nu_3 = \nu_1 * \nu_2 = 6$	$MSR = \frac{SSR}{\nu_3}$ 0	
	SST 47	$\nu = n-1 = 11$		

ν_2

4.2

STEP-I

5

Days	A	B	C	D	Total
Mon	-3	-2	-1	0	-6
TUE	-1	0	1	2	2
WED	+1	2	3	4	10
	-3	0	3	6	6

T = Grand total 6

$$C.F = \frac{T^2}{N} = \frac{36}{12} = 3$$

II: SSC

$$= \frac{A^2}{n_A} + \frac{B^2}{n_B} + \frac{C^2}{n_C} + \frac{D^2}{n_D} - C.F$$

$$= \cancel{\frac{9}{3}}^{\,3} + 0 + \cancel{\frac{9}{3}}^{\,3} + \cancel{\frac{36}{3}}^{\,12} - 3$$

$$= 15$$

SSE ?

$$\frac{M^2}{n_M} + \frac{Tu^2}{n_{Tu}} + \frac{W^2}{n_W} - C.F$$

$$= \frac{\cancel{36}^{\,9}}{\cancel{4}} + \frac{\cancel{4}^{\,1}}{\cancel{4}} + \frac{\cancel{100}^{\,25}}{\cancel{4}} - 3$$

$$= 32$$

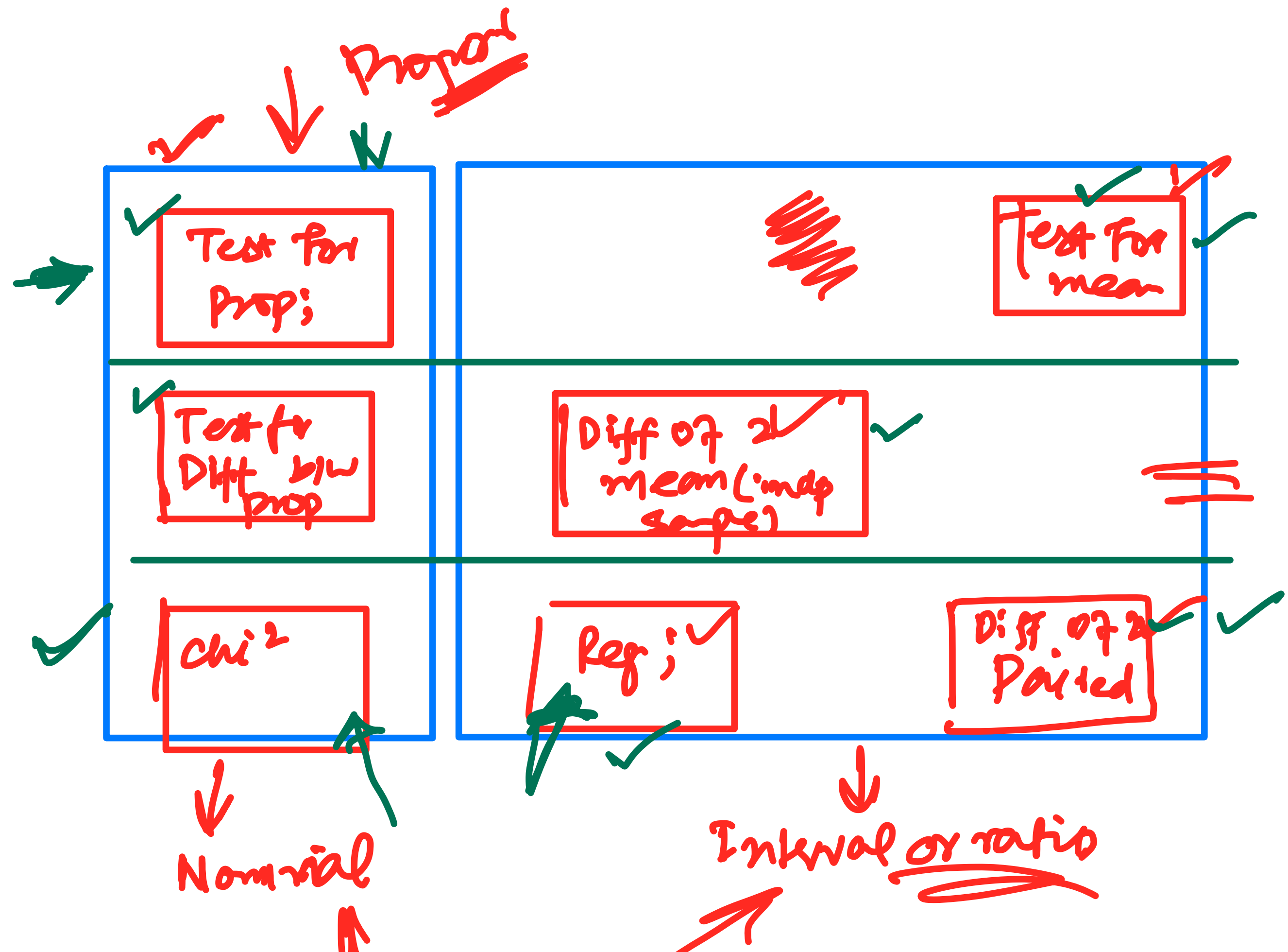
SSR = SST − (SSC + SSE)

SST =

$$(-3)^2 + (-1)^2 + \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ (4)^2 - C.F$$

$$= 47$$

$$SSR = 47 - (15 + 32) = 0$$



1) Data:

2) Samples

3) purpose.

1 sample:

2 Samples.

1 Sample → 2 measure

3) Purpose of Analysis