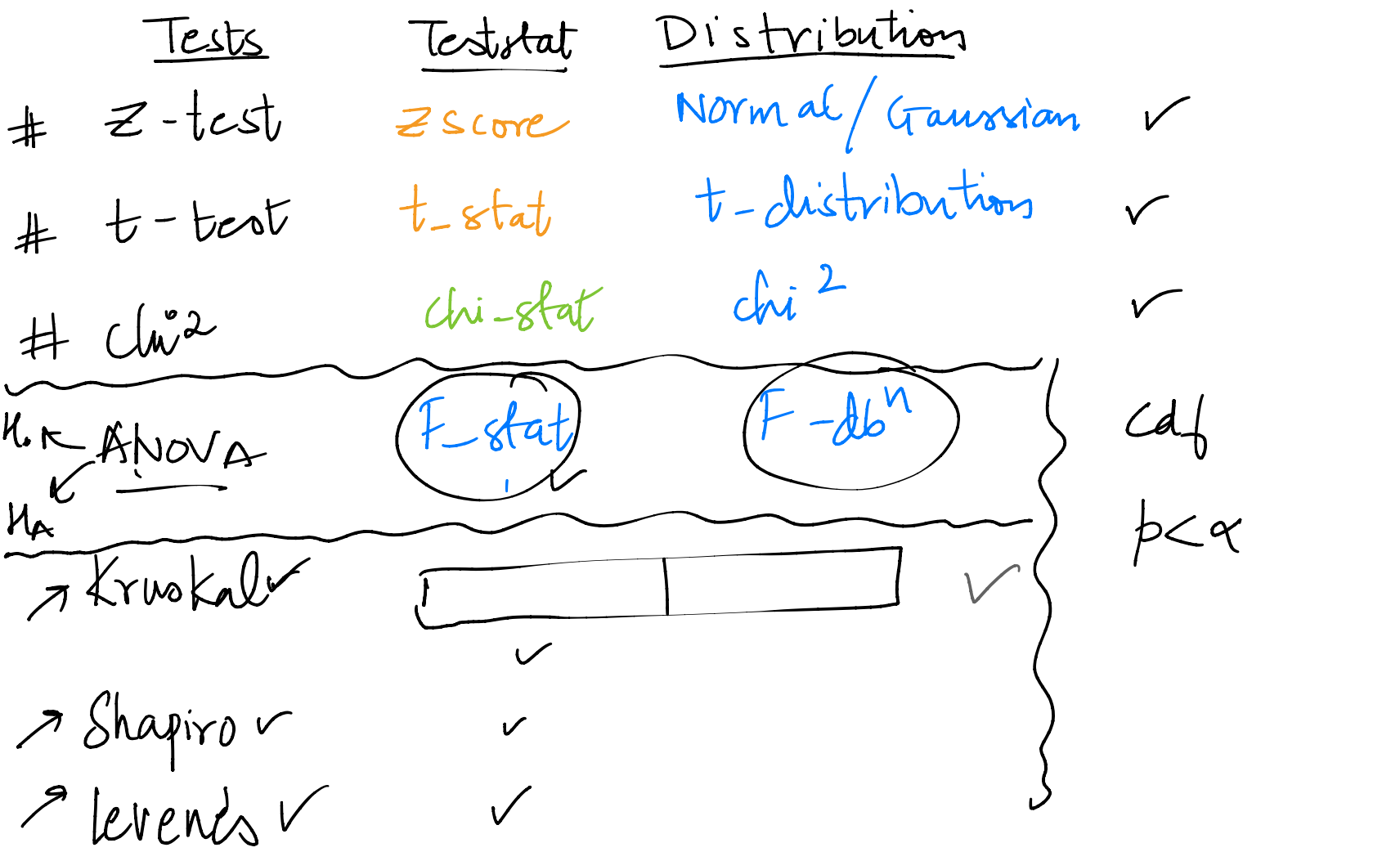
1st July 2023

ANOVA

Analysis of Variance


Tests
Teststat
Distribution
z-test
z score
Normal / Gaussian
t-test
t_stat
t-distributions
chi^2
chi_stat
chi^2
F_stat
F-dbn
cdf
H0
ANOVA
HA
p<α
Kruskal
Shapiro
levene

Aerofit :

which test

- Income v/s Gender → T-test
  (Num) (Cat)
- Product v/s (Gender) → chi^2
  (Cat) (Cat)
- Income v/s Product →
  (Num) (Cat)
  → KP281
  → 481
  → 781

3 Htest

50 cities (Cat)

how many t tests ?

$\frac{50 \times 49}{2 \times 1} = 25 \times 49$

Locations

$C_1 \quad C_2 \quad C_3 \quad - - - - - \quad C_{10}$

# of test $= {}^{10}C_2 = \frac{10 \times 9}{2 \times 1} = 45$

*

$C_1 \quad C_2 \quad C_3 \quad C_4 \quad C_5 \quad = 5\#$

C_1C_3 -

C_1C_4 /

C_1C_5 /

${}^5C_2 = \frac{5 \times 4}{2 \times 1}$

$= 10$

Income v/s Product

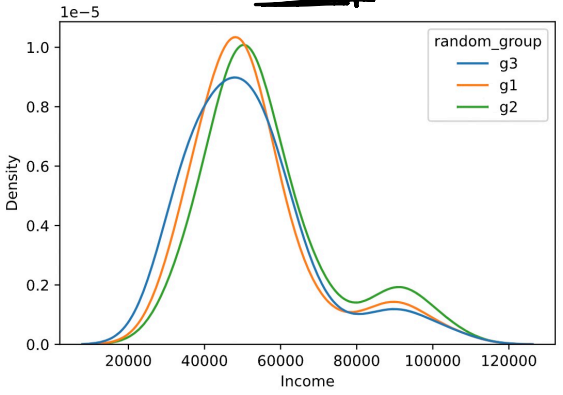
(Num) (Cat) → more than 2 categories

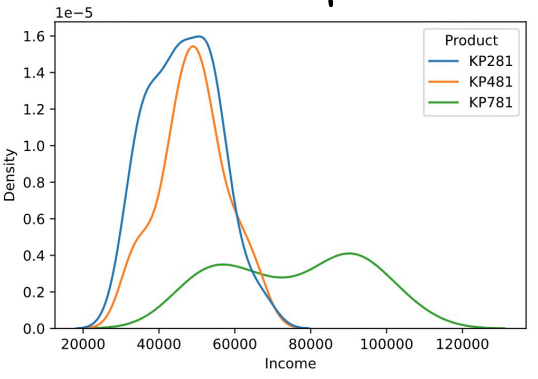
→ ANOVA

Setup -1

$\alpha = 0.05$

Setup -2





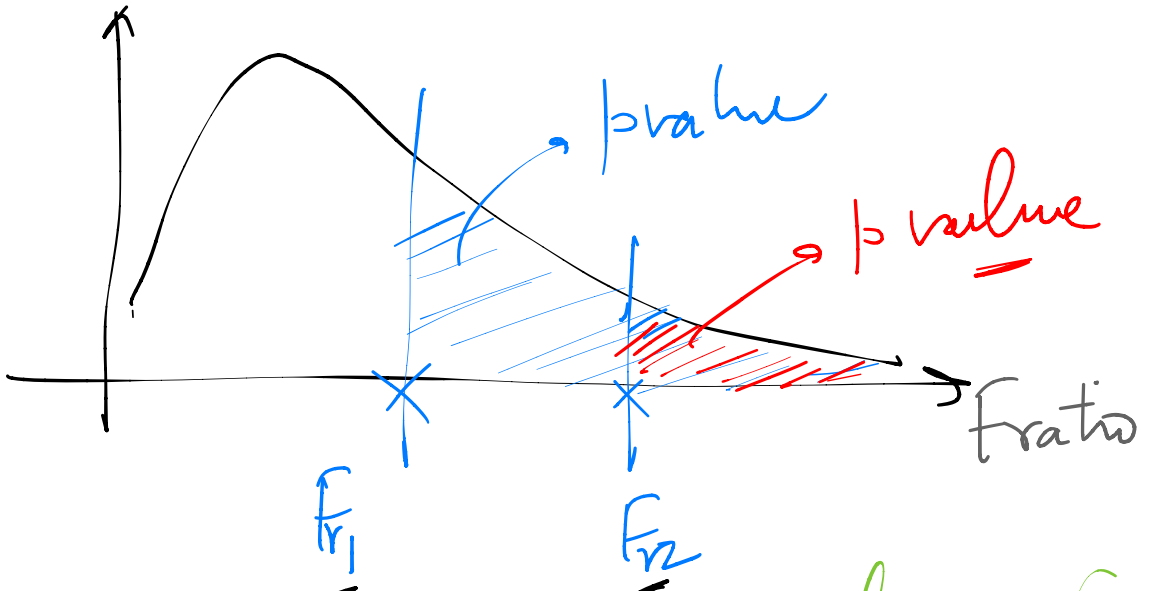
H_0 : all means are equal $\longleftrightarrow$ H_0 :

H_A : Some means are diff. $\longleftrightarrow$ H_A : ~~same~~

(i) Variance b/w groups

(ii) Variance within groups

$$F\text{ratio} = \frac{\text{Var. b/w groups}}{\text{Var. within groups}}$$

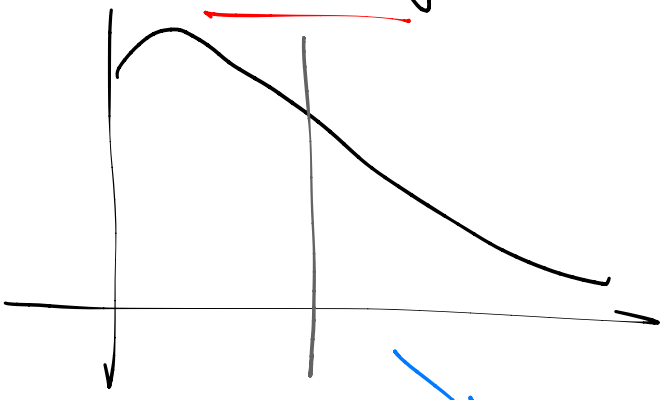


higher Fratio → lower p value

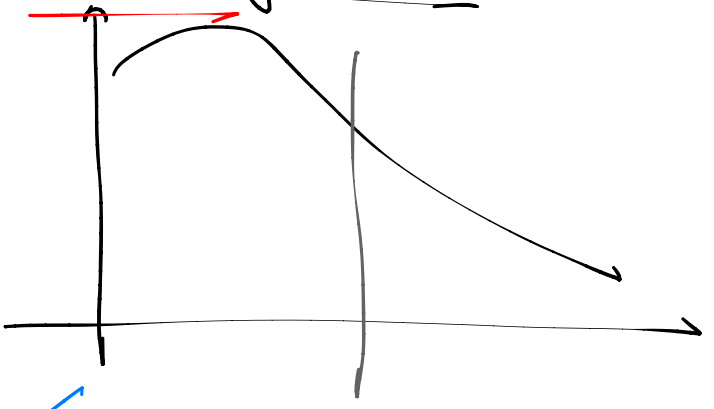
↓

more likely to reject H_0

Sachin — I^st Innings



II^nd Innings



median $P_{50} \leq P_{50}$

$P_1 \leq P_1$

$P_{25} \leq P_{25}$

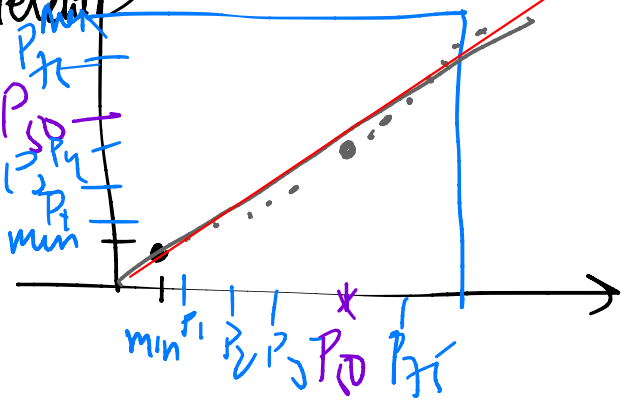
$P_{75} \leq P_{75}$

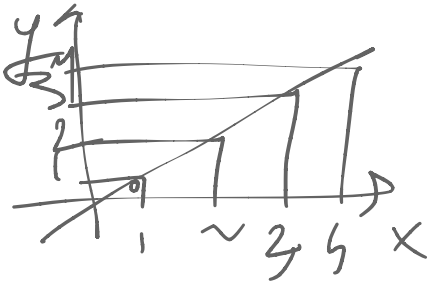
$\min(I) \leq \min(II)$.

$\max(I) \leq \max(II)$

1st

IInd Quantity (Percentiles)





QQ plot

IInd

2nd