low CAL   low FAT   low CARB

$\bar{X}_1, S_1, n_1$ $\bar{X}_2, S_2, n_2$ $\bar{X}_3, S_3, n_3$

Summary Statistics

$\bar{X}$ S n

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

$$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

Inference about μ_1, μ_2, μ_3

1. $H_0: \mu_1 = \mu_2 = \mu_3$

H_a: at least two of them are different from each other

~~$(\mu_1 \neq \mu_2 \neq \mu_3)$~~

$\boxed{\mu_1 \neq \mu_2 = \mu_3}$ Multiple comparison

- diet D   diet E

1 Factor = Diet

3 levels $\longrightarrow$ 5 levels   adding 2 more trts

$\bar{X}_4, S_4, n_4; \bar{X}_5, S_5, n_5.$

$\mu_1 - - - \mu_5.$

- Gender – Female / Male   2 levels

$3 \times 2 = 6 \rightarrow$ # trtments.   N balanced

Diet

Gender	low CAL	low FAT	low CARB
M	N/6	N/6	N/6
F	N/6	N/6	N/6

trt = combination of factor levels.

μ σ^2

$\bar{X}$ S n

$$\chi_n^2 = \sum_{i=1}^{n} X_i^2 \qquad X_i \overset{iid}{\sim} N(0,1)$$

$$t_n = \frac{X}{\sqrt{Y/n}} \qquad X \sim N(0,1) \qquad Y \sim \chi_n^2 \qquad X \perp\!\!\!\perp Y$$

independent

$$F_{n_1,n_2} = \frac{X_1/n_1}{X_2/n_2} \qquad X_1 \sim \chi_{n_1}^2 \qquad X_2 \sim \chi_{n_2}^2 \qquad X_1 \perp\!\!\!\perp X_2$$

Chapter 14 Analysis of Variance (ANOVA)

- One-way ANOVA

ANOVA determines if diff in population means for several groups by comparing VAR <u>BTWN</u> groups to VAR <u>WITHIN</u> groups.

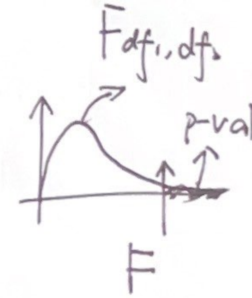
We refer to different levels of the factor in one-way ANOVA as groups & we let $g = \#$ groups.

The predictor is categorical, but the response is quantatitive.

- Test statistic $F = \dfrac{\text{VAR BTWN } g}{\text{VAR WITH } g}$

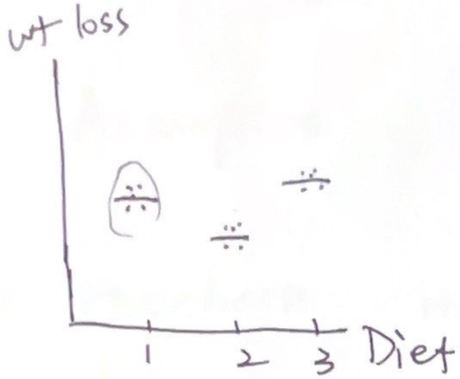
Sig difference if $F >>> 1$.

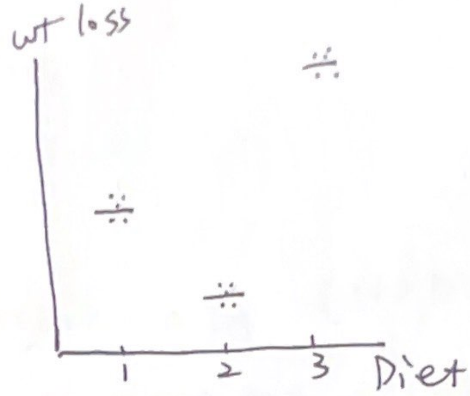
no Sig diff if $F \approx 1$.



conditional

df_1, df_2. $F \mid H_0 \sim F_{df_1, df_2}$ $P(F_{df_1,df_2} \geq F)$



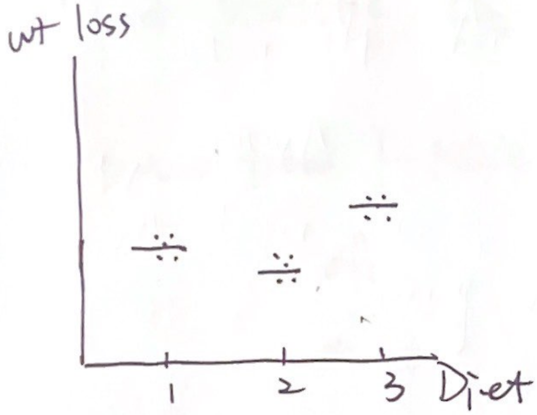


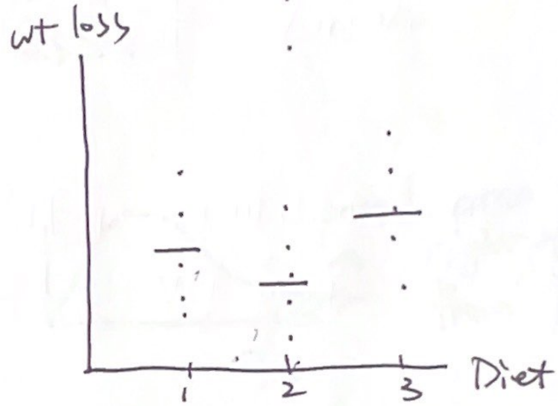
VAR WITHIN groups

VAR BTWN groups

$F = \dfrac{\text{BTWN} \text{ “}\uparrow\text{”}}{\text{WITH} \text{ “}=\text{”}}$

$F_2 > F_1$

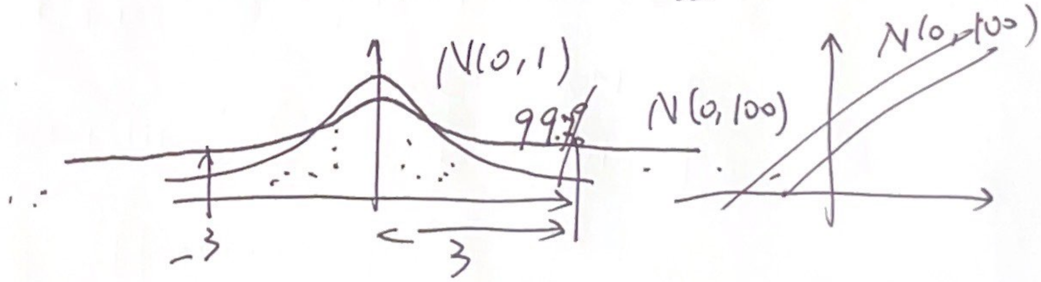




WITHIN
VAR BTWN groups

$F = \dfrac{\text{BTWN} \text{ “}=\text{”}}{\text{WITH} \text{ “}\uparrow\text{”}}$

$F_2 < F_1$



3
-5

- Assumptions

- Hypothesis $H_0: \mu_1 = \mu_2 = \cdots = \mu_g$ (NO DIFF)

  Ha: at least one μ_i's different from at least one other

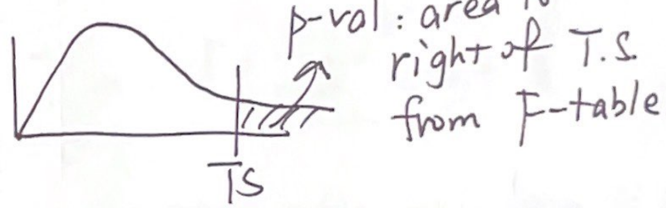
  (some DIFF; not all the same).

- Test statistic $F = \frac{MSG}{MSE}$.

- p-val from F-table

  df_1

  df_2



- Desicion small p-val → Rej Ho

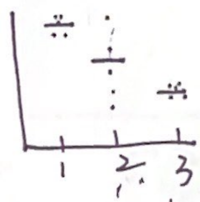
- Conclusion Some diff in population means

  → Follow-up to figure out which ones

• Assumptions:

1. SRS of exp units for each group
   - random assignment to groups
     or random selection from population
2. Original distribution of response variable is normal for each group
   - Make sure no major outliers
     (less important if samples are large)
3. Equal variances for all groups in population (homoscedasticity)
   - Check stdev of all groups are similar

   Rule of thumb: Biggest stdev is NOT more than 2x smallest stdev.

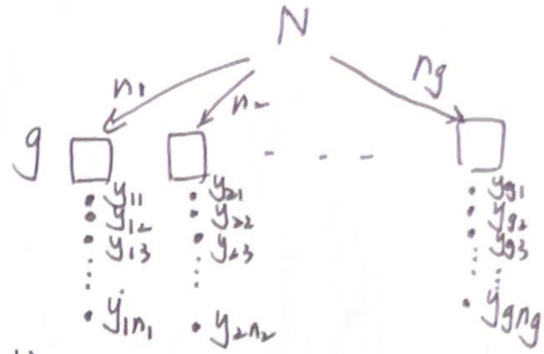


$$F = \frac{\text{VAR BTWN}}{\text{VAR WITH} \uparrow} \downarrow$$

   - If the group sizes are the same, this assumption is NOT that important.

ANOVA is fairly robust.

Symbols & Notation.



→ g = number of groups

→ y_{ij} = observation j ~~in~~ in group i

→ n_i = # of observations in group i

→ $N = \sum_{i=1}^{g} n_i$ = Total # of observations.

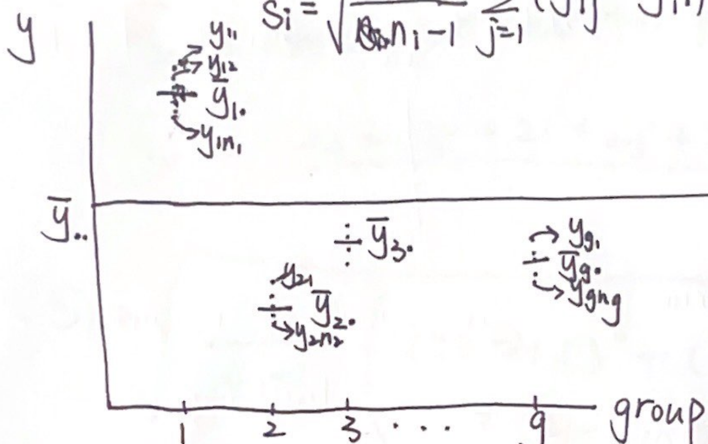
→ $\bar{y}_{i.} = \bar{y}_i$ = average of observations in group i.

→ s_i = standard deviation for group i

→ $\bar{y}_{..} = \bar{y}$ = average of all observations

$$s_i = \sqrt{\frac{1}{n_i - 1}\sum_{j=1}^{n_i}(y_{ij} - \bar{y}_{i.})^2}$$

$y_{i.}$
$\vdots$
y_{in_i}



ANOVA table

(Adj SS) sum of squares → mean squares

Source	df	SS	MS	F	p-val
Group	$g-1$ (2)	SSG (122.17)	$MSG = \frac{SSG}{g-1}$ $\left(\frac{122.17}{2}\right)$	$F = \frac{MSG}{MSE}$	from F-table w/ $df_{num} = g-1$, $df_{den} = N-g$
Error	$N-g$ (9)	SSE (60.75)	$MSE = \frac{SSE}{N-g}$		
Total	$N-1$ (11)	SST (182.92)			

* SSG + SSE = SST

p-val ⟸ $P(F_{df_{num}, df_{den}} \geq F)$

$N=12 \quad n_1=n_2=n_3=4$

1 Factor Diet w/ 3 levels.

3 trt

$y_{23} = 26$ lbs

$y_{32} = 27$ lbs

$\bar{y}_{2.}$ = average wt loss for ppl in group 2.

$$\bar{y}_{..} = \frac{22+18+21+25+24+21+\cdots+32}{\underset{\searrow N}{12}}$$

$$S_1 = \sqrt{\frac{1}{4-1}\left((22-21.5)^2 + (18-21.5)^2 + (21-21.5)^2 + (25-21.5)^2\right)}$$

S_1, S_2, S_3

$H_0: \mu_1=\mu_2=\mu_3 \quad H_a:$ at least . . . .

$$\frac{122.17}{2} = 61.085 \qquad \frac{61.085}{6.75} = 9.05$$

$$\frac{60.75}{9} = 6.75$$

0.7% De: Rej

Con: very strong