One-Way ANOVA

- Sometimes we have more than two groups
- Matt Wynn Understanding the clinician's role in dementia diagnosis
- Diagnosing dementia isn't easy or often clear-cut

CDR – Clinical dementia rating

- WashU School of Medicine
- Assesses performance in six areas: memory, orientation, judgment & problem solving, community affairs, home & hobbies, and personal care
- Composite rating
  - 0 – None
  - 0.5 – Very mild
  - 1 – Mild
  - 2 – Moderate
  - 3 - Severe

Three Groups

0 – No Impairment	0.5 – Very Mild Impairment	1 – Mild Impairment

Participants' Anxiety scores following appointment (20-80)

0 – No Impairment	0.5 – Very Mild Impairment	1 – Mild Impairment
30	35	45
35	40	55
30	45	55
25	45	50
40	55	50

Participants' Anxiety scores following appointment (20-80)

0 – No Impairment	**0.5 – Very Mild Impairment**	**1 – Mild Impairment**
30	35	45
35	40	55
30	45	55
25	45	50
40	55	50
Ybar = 32	*Ybar = 44*	*Ybar = 51*

Restricted model – 0 Group

0 – No Impairment	Prediction	Error	Squared Error
30	42.333	-12.333	152.111
35	42.333	-7.333	53.778
30	42.333	-12.333	152.111
25	42.333	-17.333	300.444
40	42.333	-2.333	5.444
Ybar = 32			$E_R = 663.889$

Full model – 0 Group

0 – No Impairment	Prediction	Error	Squared Error
30	32	-2	4
35	32	3	9
30	32	-2	4
25	32	-7	49
40	32	8	64
Ybar = 32			$E_F = 130$

Restricted Model – 0.5 group

0.5 – Very Mild	Prediction	Error	Squared Error
35	42.333	-7.333	53.778
40	42.333	-2.333	5.444
45	42.333	2.667	7.111
45	42.333	2.667	7.111
55	42.333	12.667	160.444
Ybar = 44			$E_R = 233.889$

Full Model – 0.5 group

0.5 – Very Mild	Prediction	Error	Squared Error
35	44	-9	81
40	44	-4	16
45	44	1	1
45	44	1	1
55	44	11	121
Ybar = 44			$E_F = 220$

Restricted Model – 1 group

1 – Mild	Prediction	Error	Squared Error
45	42.333	2.667	7.111
55	42.333	12.667	160.444
55	42.333	12.667	160.444
50	42.333	7.667	58.778
50	42.333	7.667	58.778
Ybar = 51			$E_R = 445.556$

Full Model – 1 group

1 – Mild	Prediction	Error	Squared Error
45	51	-6	36
55	51	4	16
55	51	4	16
50	51	-1	1
50	51	-1	1
Ybar = 51			$E_F = 70$

Wrapping it up

$E_R = 663.889 + 233.889 + 445.556 = 1343.333$

$E_F = 130 + 220 + 70 = 420$

$$F = \frac{(E_R - E_F) / (df_R - df_F)}{E_F / df_F}$$

$$F = \frac{(1343.333 - 420) / (df_R - df_F)}{418 / df_F}$$

Degrees of freedom (ν)

- Restricted model
  - N – one grand mean (N – 1)
- Full model
  - N – one mean for each group (N – a)

Wrapping it up

$$F = \frac{(1343.333 - 420) / (14 - 12)}{420 / df_F}$$

$$F = \frac{(1343.333 - 420) / (14 - 12)}{420 / 12}$$

$$F = \frac{461.667}{35.000} = 13.190$$

Critical value for $F(2, 12) = 3.885$

A trick

- E_R = Squared deviations from the grand mean
- E_F = Squared deviations from the group mean
- If equal n, $E_R - E_F = n \times \Sigma(\text{Grand Mean} - \text{Group Mean})^2$

$E_R - E_F$

- $E_R - E_F = n\Sigma(\text{Group mean} - \text{Grand mean})^2$
- $5[(32 - 42.333)^2 + (44 - 42.333)^2 + (51 - 42.333)^2]$
- $5[106.778 + 2.778 + 75.111]$
- $E_R - E_F = 923.333$

Components of the ANOVA

ANOVA Formula

$$F = \frac{(E_R - E_F) / (df_R - df_F)}{E_F / df_F}$$

$$F = \frac{923.333 / 2}{420 / 12}$$

ANOVA Formula

$$F = \frac{(E_R - E_F) / (df_R - df_F)}{E_F / df_F}$$

$$F = \frac{923.333 / 2}{420 / 12}$$

$$F = \frac{SS_B / df_{Numerator}}{SS_W / df_{Denominator}}$$

$E_R – E_F$

- Sum of squared deviations from the grand mean – Sum of squared deviations from the group mean
- $E_R – E_F = n\Sigma(\text{Group mean} – \text{Grand mean})^2$
- Sum of squared deviations *between groups*

E_F

- Sum of squared deviations from the group mean
- Errors reflect how much an individual deviates from the group
- Sum of squared deviations *within groups*

ANOVA Formula

$$F = \frac{(E_R - E_F) / (df_R - df_F)}{E_F / df_F}$$

$$F = \frac{923.333 / 2}{420 / 12}$$

$$F = \frac{SS_B / df_{Numerator}}{SS_W / df_{Denominator}}$$

$$F = \frac{MS_B}{MS_W}$$

ANOVA Output

Source	df	SS	MS	F
CDR Condition	2	923.333	461.667	13.190
Error	12	420	35.000	
Total	14	1343.333		

Why use an ANOVA?

Why not just run a number of t-tests?

- Inflates type-1 error rate FWER: the probability that one or more of your "family" of multiple tests is false
-
-

P(making an error) = alpha

P(not making an error) = 1 - alpha

P(not making an error in m tests) = (1 - alpha)^m

P(making at least 1 error in m tests) = 1 - (1 - alpha)^m

Using an ANOVA helps control for inflated FWER by using a single cohesive statistical test, rather than a series of t-tests!

ANOVA allows a test of these various means while maintaining an a priori alpha level

Why is it called the Analysis of Variance if we are interested in means?

$$\sum_i \sum_j (Y_{ij} - \bar{Y}_{..})^2$$

Why is it called the Analysis of Variance if we are interested in means?

$$\sum_i \sum_j (Y_{ij} - \bar{Y}_{..})^2 = (\bar{Y}_{.j} - \bar{Y}_{..})^2$$

Why is it called the Analysis of Variance if we are interested in means?

$$\sum_i \sum_j (Y_{ij} - \bar{Y}_{..})^2 = (\bar{Y}_{.j} - \bar{Y}_{..})^2 + (Y_{ij} - \bar{Y}_{.j})^2$$

Why is it called the Analysis of Variance if we are interested in means?

$$\sum_i \sum_j (Y_{ij} - \bar{Y}_{..})^2 = (\bar{Y}_{.j} - \bar{Y}_{..})^2 + (Y_{ij} - \bar{Y}_{.j})^2$$

$$SS_{\text{Total}} = SS_{\text{Between (Method)}} + SS_{\text{Within (Residual)}}$$

Why is it called the Analysis of Variance if we are interested in means?

$$\sum_i \sum_j (Y_{ij} - \bar{Y}_{..})^2 = (\bar{Y}_{.j} - \bar{Y}_{..})^2 + (Y_{ij} - \bar{Y}_{.j})^2$$

$$\mathrm{SS}_{\text{Total}} = \mathrm{SS}_{\text{Between (Method)}} + \mathrm{SS}_{\text{Within (Residual)}}$$

$$s^2_{\text{Total}} = s^2_{\text{Explained}} + s^2_{\text{Unexplained}}$$

ANOVA test statistic

- The ratio of SS between (signal) with the SS within (noise)

Effect Sizes

ETA SQUARED (η^2)

$$\eta^2 = \frac{SS_{between/effect}}{SS_{total}}$$

Interpretation: Proportion of the variability in the outcome variable that can be explained in terms of the predictor.

$\eta^2 = 0$; there is no relationship at all between the outcome and predictor

$\eta^2 = 1$; the relationship between the outcome and predictor is perfect

See tip in textbook on page 440 about magnitude

Proportionate Reduction in Error (PRE)

- To what extent does the full model reduce the errors made?
- $PRE = \frac{E_R - E_F}{E_R}$

R^2 or η^2

- Always going to be at least somewhat greater than zero
- Even when no actual differences occur

Proportionate Reduction in Error (PRE)

- To what extent does the full model reduce the errors made?
- $$PRE = \frac{E_R - E_F}{E_R} = \frac{923.333}{1343.333} = 0.687$$

scale of 0 - 1

If it's 0, then knowing X does not help predict Y. The full model does not reduce the errors made. If it's 1, then knowing X 100% predicts Y. The full model completely reduces the errors made.

Another alternative

- Cohen's *d* using the extreme groups
- Mostly useful for power

Assumptions for the F test

Typical assumptions

- The following should hold in order to assume your observed F maps on to the population distribution for F
- 1) Normally distributed scores
- 2) Equal variances
- 3) Scores should be independent

What happens when you have ... ?

- 1) Unequal n
- 2) Unequal variance

Different formula for E_F

- E_F = Deviations from the group mean
- Σ(Sum of the squared deviations x Sample size for the group)
- Weight the sum of squares by the group size

When is the F test robust? (Maxwell & Delaney, 2004)

	Equal Sample Sizes	Unequal Sample Sizes
Equal variances		
Unequal variances		

When is the F test robust?

	Equal Sample Sizes	Unequal Sample Sizes
Equal variances	Appropriate	Appropriate
Unequal variances		

When is the F test robust?

	Equal Sample Sizes	Unequal Sample Sizes
Equal variances	Appropriate	Appropriate
Unequal variances	Good, unless there's a very large difference	???

When is the F test robust?

	Equal Sample Sizes	Unequal Sample Sizes
Equal variances	Appropriate	Appropriate
Unequal variances	Good, unless there's a very large difference	
- Large sample with large variances		Conservative

When is the F test robust?

	Equal Sample Sizes	Unequal Sample Sizes
Equal variances	Appropriate	Appropriate
Unequal variances	Good, unless there's a very large difference	
- Large sample with large variances		Conservative
- Large sample with small variances		Liberal