Factorial Within Subjects

Psy 420

Ainsworth

Factorial WS Designs

- Analysis
  - Factorial – deviation and computational
- Power, relative efficiency and sample size
- Effect size
- Missing data
- Specific Comparisons

Example for Deviation Approach

	b_1: Science Fiction			b_2: Mystery			Case Means
	a_1: Month 1	a_2: Month 2	a_3: Month 3	a_1: Month 1	a_2: Month 2	a_3: Month 3	
S_1	1	3	6	5	4	1	$S_1 = 3.333$
S_2	1	4	8	8	8	4	$S_2 = 5.500$
S_3	3	3	6	4	5	3	$S_3 = 4.000$
S_4	5	5	7	3	2	0	$S_4 = 3.667$
S_5	2	4	5	5	6	3	$S_5 = 4.167$
Treatment Means	$a_1b_1 = 2.4$	$a_2b_1 = 3.8$	$a_3b_1 = 6.4$	$a_1b_2 = 5$	$a_2b_2 = 5$	$a_3b_2 = 2.2$	GM = 4.133

Analysis – Deviation

- What effects do we have?
  - A
  - B
  - AB
  - S
  - AS
  - BS
  - ABS
  - T

Analysis – Deviation

- DFs
  - $DF_A = a - 1$
  - $DF_B = b - 1$
  - $DF_{AB} = (a - 1)(b - 1)$
  - $DF_S = (s - 1)$
  - $DF_{AS} = (a - 1)(s - 1)$
  - $DF_{BS} = (b - 1)(s - 1)$
  - $DF_{ABS} = (a - 1)(b - 1)(s - 1)$
  - $DF_T = abs - 1$

Sums of Squares - Deviation

- The total variability can be partitioned into A, B, AB, Subjects and a Separate Error Variability for Each Effect

$$SS_{Total} = SS_A + SS_B + SS_{AB} + SS_S + SS_{AS} + SS_{BS} + SS_{ABS}$$

$$\sum\left(Y_{ijk} - \bar{Y}_{...}\right)^2 = \sum n_j\left(\bar{Y}_{.j.} - \bar{Y}_{...}\right)^2 + \sum n_k\left(\bar{Y}_{..k} - \bar{Y}_{...}\right)^2 +$$

$$+\left[\sum n_{jk}\left(\bar{Y}_{.jk} - \bar{Y}_{...}\right)^2 - \sum n_j\left(\bar{Y}_{.j.} - \bar{Y}_{...}\right)^2 - \sum n_k\left(\bar{Y}_{..k} - \bar{Y}_{...}\right)^2\right]$$

$$+jk\sum\left(\bar{Y}_{i..} - \bar{Y}_{...}\right)^2 + \left[k\sum\left(Y_{ij.} - \bar{Y}_{.j.}\right)^2 - jk\sum\left(\bar{Y}_{i..} - \bar{Y}_{...}\right)^2\right] +$$

$$+\left[j\sum\left(Y_{i.k} - \bar{Y}_{.j.}\right)^2 - jk\sum\left(\bar{Y}_{i..} - \bar{Y}_{...}\right)^2\right] +$$

$$+\left[\sum\left(Y_{ijk} - \bar{Y}_{.jk}\right)^2 - jk\sum\left(\bar{Y}_{i..} - \bar{Y}_{...}\right)^2 - k\sum\left(Y_{ij.} - \bar{Y}_{.j.}\right)^2 - j\sum\left(Y_{i.k} - \bar{Y}_{.j.}\right)^2\right]$$

Analysis – Deviation

- Before we can calculate the sums of squares we need to rearrange the data
- When analyzing the effects of A (Month) you need to **AVERAGE** over the effects of B (Novel) and vice versa
- The data in its original form is only useful for calculating the AB and ABS interactions

For the A and AS effects

- Remake data, averaging over B

	a_1: Month 1	a_2: Month 2	a_3: Month 3	Case Means
S_1	3	3.5	3.5	**S_1 = 3.333**
S_2	4.5	6	6	**S_2 = 5.500**
S_3	3.5	4	4.5	**S_3 = 4.000**
S_4	4	3.5	3.5	**S_4 = 3.667**
S_5	3.5	5	4	**S_5 = 4.167**
Treatment Means	**a_1 = 3.7**	**a_2 = 4.4**	**a_3 = 4.3**	**GM = 4.133**

$$SS_A = \sum n_j \left(\bar{Y}_{.j.} - \bar{Y}_{...}\right)^2 = 10*[(3.7-4.133)^2 + (4.4-4.133)^2 + (4.3-4.133)^2] = 2.867$$

$$SS_S = jk\sum\left(\bar{Y}_{i..} - \bar{Y}_{...}\right)^2 = (3*2)*[(3.333-4.133)^2 + (5.5-4.133)^2 + (4-4.133)^2 +$$
$$+(3.667-4.133)^2 + (4.167-4.133)^2 = 16.467$$

$$SS_{AS} = k\sum\left(Y_{ij.} - \bar{Y}_{.j.}\right)^2 - jk\sum\left(\bar{Y}_{i..} - \bar{Y}_{...}\right)^2 =$$
$$k\sum\left(Y_{ij.} - \bar{Y}_{.j.}\right)^2 = 2*[(3-3.7)^2 + (4.5-3.7)^2 + (3.5-3.7)^2 + (4-3.7)^2 + (3.5-3.7)^2 +$$
$$+(3.5-4.4)^2 + (6-4.4)^2 + (4-4.4)^2 + (3.5-4.4)^2 + (5-4.4)^2 +$$
$$+(3.5-4.3)^2 + (6-4.3)^2 + (4.5-4.3)^2 + (3.5-4.3)^2 + (4-4.3)^2] = 20.6$$

$$SS_{AS} = 20.6 - 16.467 = 4.133$$

For B and BS effects

- Remake data, averaging over A

	b_1: Science Fiction	b_2: Mystery	Case Means
S_1	3.333	3.333	S_1 = 3.333
S_2	4.333	6.667	S_2 = 5.500
S_3	4.000	4.000	S_3 = 4.000
S_4	5.667	1.667	S_4 = 3.667
S_5	3.667	4.667	S_5 = 4.167
Treatment Means	b_1 = 4.200	b_2 = 4.067	GM = 4.133

$$SS_B = \sum n_k \left(\bar{Y}_{..k} - \bar{Y}_{...}\right)^2 = 15*[(4.2-4.133)^2 + (4.067-4.133)^2] = .133$$

$$SS_{BS} = j\sum\left(Y_{i.k} - \bar{Y}_{..k}\right)^2 - jk\sum\left(\bar{Y}_{i..} - \bar{Y}_{...}\right)^2 =$$

$$j\sum\left(Y_{i.k} - \bar{Y}_{..k}\right)^2 = 3*[(3.333-4.2)^2 + (4.333-4.2)^2 + (4-4.2)^2 +$$

$$+(5.667-4.2)^2 + (3.667-4.2)^2 + (3.333-4.067)^2 + (6.667-4.067)^2 +$$

$$+(4-4.067)^2 + (1.667-4.067)^2 + (4.667-4.067)^2] = 50$$

$$SS_S = jk\sum\left(\bar{Y}_{i..} - \bar{Y}_{...}\right)^2 = 16.467$$

$$SS_{BS} = 50 - 16.467 = 33.533$$

For AB and ABS effects

- Use the data in its original form

	b_1: Science Fiction			b_2: Mystery			Case Means
	a_1: Month 1	a_2: Month 2	a_3: Month 3	a_1: Month 1	a_2: Month 2	a_3: Month 3	
S_1	1	3	6	5	4	1	$S_1 = 3.333$
S_2	1	4	8	8	8	4	$S_2 = 5.500$
S_3	3	3	6	4	5	3	$S_3 = 4.000$
S_4	5	5	7	3	2	0	$S_4 = 3.667$
S_5	2	4	5	5	6	3	$S_5 = 4.167$
Treatment Means	$a_1b_1 = 2.4$	$a_2b_1 = 3.8$	$a_3b_1 = 6.4$	$a_1b_2 = 5$	$a_2b_2 = 5$	$a_3b_2 = 2.2$	GM = 4.133

$$SS_{AB} = \sum n_{jk}\left(\bar{Y}_{.jk} - \bar{Y}_{...}\right)^2 - \sum n_j\left(\bar{Y}_{.j.} - \bar{Y}_{...}\right)^2 - \sum n_k\left(\bar{Y}_{..k} - \bar{Y}_{...}\right)^2 =$$

$$\sum n_{jk}\left(\bar{Y}_{.jk} - \bar{Y}_{...}\right)^2 = 5*[(2.4-4.133)^2 + (3.8-4.133)^2 + (6.4-4.133)^2$$
$$+(5-4.133)^2 + (5-4.133)^2 + (2.2-4.133)^2] = 67.467$$

$$\sum n_j\left(\bar{Y}_{.j.} - \bar{Y}_{...}\right)^2 = SS_A = 2.867$$

$$\sum n_k\left(\bar{Y}_{..k} - \bar{Y}_{...}\right)^2 = SS_B = .133$$

$$SS_{AB} = 67.467 - 2.867 - .133 = 64.467$$

$$SS_{ABS} = \left[\sum \left(Y_{ijk} - \overline{Y}_{.jk} \right)^2 - jk \sum \left(\overline{Y}_{i..} - \overline{Y}_{...} \right)^2 - k \sum \left(Y_{ij.} - \overline{Y}_{.j.} \right)^2 - j \sum \left(Y_{i.k} - \overline{Y}_{.j.} \right)^2 \right] =$$

$$\sum \left(Y_{ijk} - \overline{Y}_{.jk} \right)^2 = (1-2.4)^2 + (1-2.4)^2 + (3-2.4)^2 +$$
$$+ (5-2.4)^2 + (2-2.4)^2 + (3-3.8)^2 + (4-3.8)^2 +$$
$$+ (3-3.8)^2 + (5-3.8)^2 + (4-3.8)^2 + (6-6.4)^2 +$$
$$+ (8-6.4)^2 + (6-6.4)^2 + (7-6.4)^2 + (5-6.4)^2 +$$
$$+ (6-5)^2 + (8-5)^2 + (4-5)^2 + (3-5)^2 +$$
$$+ (5-5)^2 + (4-5)^2 + (8-5)^2 + (5-5)^2 +$$
$$+ (2-5)^2 + (6-5)^2 + (1-2.2)^2 + (4-2.2)^2 +$$
$$+ (3-2.2)^2 + (0-2.2)^2 + (3-2.2)^2 = 64$$

$$SS_{ABS} = 64 - 16.467 - 20.6 - 50 = 9.867$$

$$SS_{Total} = (1-4.133)^2 + (1-4.133)^2 + (3-4.133)^2 +$$
$$+(5-4.133)^2 + (2-4.133)^2 + (3-4.133)^2 + (4-4.133)^2 +$$
$$+(3-4.133)^2 + (5-4.133)^2 + (4-4.133)^2 + (6-4.133)^2 +$$
$$+(8-4.133)^2 + (6-4.133)^2 + (7-4.133)^2 + (5-4.133)^2 +$$
$$+(6-4.133)^2 + (8-4.133)^2 + (4-4.133)^2 + (3-4.133)^2 +$$
$$+(5-4.133)^2 + (4-4.133)^2 + (8-4.133)^2 + (5-4.133)^2 +$$
$$+(2-4.133)^2 + (6-4.133)^2 + (1-4.133)^2 + (4-4.133)^2 +$$
$$+(3-4.133)^2 + (0-4.133)^2 + (3-4.133)^2 = 131.467$$

Analysis – Deviation

- DFs
  - $DF_A = a - 1 = 3 - 1 = 2$
  - $DF_B = b - 1 = 2 - 1 = 1$
  - $DF_{AB} = (a - 1)(b - 1) = 2 * 1 = 2$
  - $DF_S = (s - 1) = 5 - 1 = 4$
  - $DF_{AS} = (a - 1)(s - 1) = 2 * 4 = 8$
  - $DF_{BS} = (b - 1)(s - 1) = 1 * 4 = 4$
  - $DF_{ABS} = (a - 1)(b - 1)(s - 1) = 2 * 1 * 4 = 8$
  - $DF_T = abs - 1 = [3*(2)*(5)] - 1 = 30 - 1 = 29$

Source Table - Deviation

Source	SS	df	MS	F
A	2.867	2	1.433	2.774
AS	4.133	8	0.517	
B	0.133	1	0.133	0.016
BS	33.533	4	8.383	
AB	64.467	2	32.233	26.135
ABS	9.867	8	1.233	
S	16.467	4	4.117	
Total	131.467	29		

Analysis – Computational

- Example data re-calculated with totals

	b_1: Science Fiction			b_2: Mystery			Case Totals
	a_1: Month 1	a_2: Month 2	a_3: Month 3	a_1: Month 1	a_2: Month 2	a_3: Month 3	
S_1	1	3	6	5	4	1	S_1=20
S_2	1	4	8	8	8	4	S_2=33
S_3	3	3	6	4	5	3	S_3=24
S_4	5	5	7	3	2	0	S_4=22
S_5	2	4	5	5	6	3	S_5=25
Treatment Totals	$a_1b_1 = 12$	$a_2b_1 = 19$	$a_3b_1 = 32$	$a_1b_2 = 25$	$a_2b_2 = 25$	$a_3b_2 = 11$	T = 124

Analysis – Computational

- Before we can calculate the sums of squares we need to rearrange the data
- When analyzing the effects of A (Month) you need to **SUM** over the effects of B (Novel) and vice versa
- The data in its original form is only useful for calculating the AB and ABS interactions

Analysis – Computational

- For the Effect of A (Month) and A x S

	Month			Case Totals
	a_1: Month 1	a_2: Month 2	a_3: Month 3	
S_1	6	7	7	S_1=20
S_2	9	12	12	S_2=33
S_3	7	8	9	S_3=24
S_4	8	7	7	S_4=22
S_5	7	10	8	S_5=25
Treatment Totals	37	44	43	T = 124

$$SS_A = \frac{\sum\left(\sum a_j\right)^2}{bs} - \frac{T^2}{abs} = \frac{37^2 + 44^2 + 43^2}{2(5)} - \frac{124^2}{2(3)(5)} =$$

$$SS_A = \frac{5154}{10} - \frac{15376}{30} = 515.4 - 512.533 = 2.867$$

$$SS_S = \frac{\sum\left(\sum s_i\right)^2}{ab} - \frac{T^2}{abs} = \frac{20^2 + 33^2 + 24^2 + 22^2 + 25^2}{2(3)} - 512.533 =$$

$$SS_S = \frac{3174}{6} - 512.533 = 529 - 512.533 = 16.467$$

$$SS_{AS} = \frac{\sum\left(\sum a_j s_i\right)^2}{b} - \frac{\sum\left(\sum a_j\right)^2}{bs} - \frac{\sum\left(\sum s_i\right)^2}{ab} + \frac{T^2}{abs} =$$

$$SS_{AS} = \frac{6^2 + 9^2 + 7^2 + 8^2 + 7^2 + 7^2 + 12^2 + 8^2 + 7^2 + 10^2 + 7^2 + 12^2 + 9^2 + 7^2 + 8^2}{2} -$$

$$-515.4 - 529 + 512.533 = \frac{1072}{2} - 515.4 - 529 + 512.533 =$$

$$SS_{AS} = 536 - 515.4 - 529 + 512.533 = 4.133$$

Analysis – Traditional

- For B (Novel) and B x S

	Type of Novel		Case Totals
	b_1: Science Fiction	b_2: Mystery	
S_1	10	10	S_1=20
S_2	13	20	S_2=33
S_3	12	12	S_3=24
S_4	17	5	S_4=22
S_5	11	14	S_5=25
Treatment Totals	63	61	T = 124

$$SS_B = \frac{\sum\left(\sum b_k\right)^2}{as} - \frac{T^2}{abs} = \frac{63^2 + 61^2}{3(5)} - 512.533 =$$

$$SS_B \frac{7690}{15} - 512.533 = 512.667 - 512.533 = .133$$

$$SS_{BS} = \frac{\sum\left(\sum b_k s_i\right)^2}{a} - \frac{\sum\left(\sum b_k\right)^2}{as} - \frac{\sum\left(\sum s_i\right)^2}{ab} + \frac{T^2}{abs} =$$

$$SS_{BS} = \frac{10^2 + 13^2 + 12^2 + 17^2 + 11^2 + 10^2 + 20^2 + 12^2 + 5^2 + 14^2}{3} -$$

$$-512.667 - 529 + 512.533 =$$

$$SS_{BS} = \frac{1688}{3} - 512.667 - 529 + 512.533$$

$$SS_{BS} = 562.667 - 512.667 - 529 + 512.533 = 33.533$$

Analysis – Computational

- Example data re-calculated with totals

	b_1: Science Fiction			b_2: Mystery			Case Totals
	a_1: Month 1	a_2: Month 2	a_3: Month 3	a_1: Month 1	a_2: Month 2	a_3: Month 3	
S_1	1	3	6	5	4	1	S_1=20
S_2	1	4	8	8	8	4	S_2=33
S_3	3	3	6	4	5	3	S_3=24
S_4	5	5	7	3	2	0	S_4=22
S_5	2	4	5	5	6	3	S_5=25
Treatment Totals	a_1b_1 = 12	a_2b_1 = 19	a_3b_1 = 32	a_1b_2 = 25	a_2b_2 = 25	a_3b_2 = 11	T = 124

$$SS_{AB} = \frac{\sum\left(\sum a_j b_k\right)^2}{s} - \frac{\sum\left(\sum a_j\right)^2}{bs} - \frac{\sum\left(\sum b_k\right)^2}{as} + \frac{T^2}{abs} =$$

$$SS_{AB} = \frac{12^2 + 19^2 + 32^2 + 25^2 + 25^2 + 11^2}{5} - 515.4 - 512.667 + 512.533 =$$

$$SS_{AB} = \frac{2900}{5} - 515.4 - 512.667 + 512.533 =$$

$$SS_{AB} = 580 - 515.4 - 512.667 + 512.533 = 64.467$$

$$SS_{ABS} = \sum Y^2 - \frac{\sum\left(\sum a_j b_k\right)^2}{s} - \frac{\sum\left(\sum a_j s_i\right)^2}{b} - \frac{\sum\left(\sum b_k s_i\right)^2}{a} +$$

$$+ \frac{\sum\left(\sum a_j\right)^2}{bs} + \frac{\sum\left(\sum b_k\right)^2}{as} + \frac{\sum\left(\sum s_i\right)^2}{ab} - \frac{T^2}{abs} =$$

$$SS_{ABS} = 644 - 580 - 536 - 562.667 + 515.4 + 512.667 + 529 - 512.533 = 9.867$$

$$SS_{Total} = \sum Y^2 - \frac{T^2}{abs} = 644 - 512.533 = 131.467$$

Analysis – Computational DFs are the same

- DFs
  - $DF_A = a - 1 = 3 - 1 = 2$
  - $DF_B = b - 1 = 2 - 1 = 1$
  - $DF_{AB} = (a - 1)(b - 1) = 2 * 1 = 2$
  - $DF_S = (s - 1) = 5 - 1 = 4$
  - $DF_{AS} = (a - 1)(s - 1) = 2 * 4 = 8$
  - $DF_{BS} = (b - 1)(s - 1) = 1 * 4 = 4$
  - $DF_{ABS} = (a - 1)(b - 1)(s - 1) = 2 * 1 * 4 = 8$
  - $DF_T = abs - 1 = [3*(2)*(5)] - 1 = 30 - 1 = 29$

Source Table – Computational Is also the same

Source	SS	df	MS	F
A	2.867	2	1.433	2.774
AS	4.133	8	0.517	
B	0.133	1	0.133	0.016
BS	33.533	4	8.383	
AB	64.467	2	32.233	26.135
ABS	9.867	8	1.233	
S	16.467	4	4.117	
Total	131.467	29		

Relative Efficiency

- One way of conceptualizing the power of a within subjects design is to calculate how efficient (uses less subjects) the design is compared to a between subjects design
- One way of increasing power is to minimize error, which is what WS designs do, compared to BG designs
- WS designs are more powerful, require less subjects, therefore are more efficient

Relative Efficiency

- Efficiency is not an absolute value in terms of how it's calculated
- In short, efficiency is a measure of how much smaller the WS error term is compared to the BG error term
- Remember that in a one-way WS design:

  $SS_{AS}=SS_{S/A} - SS_{S}$

Relative Efficiency

$$\text{Relative Efficiency} = \frac{MS_{S/A}}{MS_{AS}}\left(\frac{df_{AS}+1}{df_{AS}+3}\right)\left(\frac{df_{S/A}+3}{df_{S/A}+1}\right)$$

- $MS_{S/A}$ can be found by either re-running the analysis as a BG design or for one-way:

  $SS_S + SS_{AS} = SS_{S/A}$ and

  $df_S + df_{AS} = df_{S/A}$

$$MS_{S/A} = \frac{(s-1)MS_S + (a-1)(s-1)MS_{AS}}{as-1}$$

Relative Efficiency

- From our one-way example:

Source	SS	df	MS	F
A	41.20	2	20.60	17.35
S	9.70	4	2.43	2.04
AxS	9.50	8	1.19	
Total	60.40	14		

$$SS_{S/A} = SS_S + SS_{AS} = 9.7 + 9.5 = 19.2$$

$$df_{S/A} = df_S + df_{AS} = 4 + 8 = 12$$

$$MS_{S/A} = 19.2/12 = 1.6$$

Relative Efficiency

$$\text{Relative Efficiency} = \frac{1.6}{1.2}\left(\frac{8+1}{8+3}\right)\left(\frac{12+3}{12+1}\right) = 1.26$$

- From the example, the within subjects design is roughly 26% more efficient than a between subjects design (controlling for degrees of freedom)
- For # of BG subjects (n) to match WS (s) efficiency:

$$n = s(\text{relative efficiency}) = 5(1.26) = 6.3 \approx 7$$

Power and Sample Size

- Estimating sample size is the same from before
- For a one-way within subjects you have and AS interaction but you only estimate sample size based on the main effect
- For factorial designs you need to estimate sample size based on each effect and use the largest estimate
- Realize that if it (PC-Size, formula) estimates 5 subjects, that total, not per cell as with BG designs

Effects Size

- Because of the risk of a true AS interaction we need to estimate lower and upper bound estimates

$$\eta^2 = \eta_L^2 = \frac{SS_A}{SS_{Total}} = \frac{SS_A}{SS_A + SS_S + SS_{AS}}$$

$$partial\ \eta^2 = \frac{SS_A}{SS_A + SS_{error}} = \frac{SS_A}{SS_A + SS_{AS}}$$

$$\hat{\varpi}_L^2 = \frac{df_A(MS_A - MS_{AS})}{df_A(MS_A - MS_{AS}) + as(MS_{AS}) + s(MS_S)}$$

$$partial\ \hat{\varpi}^2 = \frac{df_A(MS_A - MS_{AS})}{df_A(MS_A - MS_{AS}) + as(MS_{AS})}$$

Missing Values

- In repeated measures designs it is often the case that the subjects may not have scores at all levels (e.g. inadmissible score, drop-out, etc.)
- The default in most programs is to delete the case if it doesn't have complete scores at all levels
- If you have a lot of data that's fine
- If you only have a limited cases…

Missing Values

- Estimating values for missing cases in WS designs (one method: takes mean of A_j, mean for the case and the grand mean)

$$Y_{ij}^{*} = \frac{sS_i^{'} + aA_j^{'} - T^{'}}{(a-1)(s-1)}$$

$where:$

$Y_{ij}^{*} =$ predicted score to replace missing score

$s =$ number of cases

$S_i^{'} =$ sum of known values for that case

$a =$ number of levels of A

$A_j^{'} =$ sum of known values for A

$T^{'} =$ sum of all known values

Specific Comparisons

- F-tests for comparisons are exactly the same as for BG

$$F = \frac{n(\sum w_j \bar{Y}_j)^2 / \sum w_j^2}{MS_{error}}$$

- Except that MS_{error} is not just MS_{AS}, it is going to be different for every comparison
- It is recommended to just calculate this through SPSS