Unbalanced Designs & Quasi F-Ratios

ANOVA for unequal n's, pooled variances, & other useful tools

Unequal n's

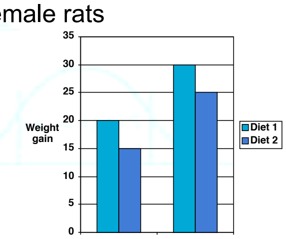
- Focus (so far) on Balanced Designs
  - Equal n's in groups (CR-p and CRF-pq)
  - Observation in every block (RB-p RBF-pq)
- What happens when cell n's are unequal?
  - Induce correlations between the factors
  - SS no longer independent
    - SS_{total} is not clearly partitioned
    - ANOVA assumptions may not hold

Unequal n's

- Example: fake data from a study of the effects of two different diets on weight gain in male and female rats

	Female	Male
Diet 1	20	30
Diet 2	15	25

  - Main effect of Diet
    Diet 1 > Diet 2
  - Main effect of sex
    Male > Female
  - Interaction? No



Unequal n's

- Calculate traditional ANOVA with sums:

	df	SS
Diet	1	5
Sex	1	245
DxS	1	75

* However, based on the means, $SS_{DxS} = 0$

	Female		Male	
Diet 1	20	-	30	= 10
Diet 2	15	-	25	= 10

Unequal n's

- Why does this happen?

	Female	Male
Diet 1	20 n = 8	30 n = 2
Diet 2	15 n = 2	25 n = 8

* Think in terms of orthogonal contrasts
  - With equal n's:

	B1	B2
A1	1 n = 5	1 n = 5
A2	-1 n = 5	-1 n = 5

	B1	B2
A1	1 n = 5	-1 n = 5
A2	1 n = 5	-1 n = 5

Cross-product = (1)(1)/5 + (1)(-1)/5 + (-1)(1)/5 + (-1)(-1)/5 = 0

Unequal n's

- Why does this happen?

	Female	Male
Diet 1	20 n = 8	30 n = 2
Diet 2	15 n = 2	25 n = 8

* Think in terms of orthogonal contrasts
  - With unequal n's:

	B1	B2
A1	1 n = 8	1 n = 2
A2	-1 n = 2	-1 n = 8

	B1	B2
A1	1 n = 8	-1 n = 2
A2	1 n = 2	-1 n = 8

Cross-product = (1)(1)/8 + (1)(-1)/2 + (-1)(1)/2 + (-1)(-1)/8 = 1/4 -1

Unequal n's

- Simpler case: just one cell with different n

	B1	B2
A1	1 n = 4	1 n = 5
A2	-1 n = 5	-1 n = 5

	B1	B2
A1	1 n = 4	-1 n = 5
A2	1 n = 5	-1 n = 5

Cross-product = (1)(1)/4 + (1)(-1)/5 + (-1)(1)/5 + (-1)(-1)/5 = 1/4 - 1/5

Effects are correlated because of unequal n's

Unweighted Means Method

- Most common approach
- Unequal n's are ignored when calculating the marginal means
- For example:
  - $\bar{x}_{jk} = \sum_i x_{jki}/n_{jk}$ (cell)
  - $\bar{x}_{j\cdot} = \sum_k \bar{x}_{ik}/n_{j\cdot}$ (marg)
  - $\bar{x}_{\cdot k} = \sum_j \bar{x}_{ik}/n_{\cdot k}$ (marg)

	B1	B2	
A1	11 9 5 15 10 **10** n = 5	12 18 16 14 **15** n = 4	**12.5**
A2	7 4 10 11 **8** n = 4	9 4 8 **7** n = 3	**7.5**
	9.0	**11.0**	

Unweighted Means Method

- Effect of A is calculated using 12.5 and 7.5
- Effect of B is calculated using 9.0 and 11.0

	B1	B2	
A1	11 9 5 15 10 **10** n = 5	12 18 16 14 **15** n = 4	**12.5**
A2	7 4 10 11 **8** n = 4	9 4 8 **7** n = 3	**7.5**
	9.0	**11.0**	

Unweighted Means Method

- Formally, the hypotheses are:
  - H_0 for A : $\sum \frac{\mu_{jk}}{q} - \sum \frac{\mu_{j'k}}{q} = 0$ or $\mu_{j\cdot} - \mu_{j'\cdot} = 0$
  - H_0 for B: $\sum \frac{\mu_{jk}}{p} - \sum \frac{\mu_{jk'}}{p} = 0$ or $\mu_{\cdot k} - \mu_{\cdot k'} = 0$
  - H_0 for AxB: $(\mu_{jk} - \mu_{j'k}) - (\mu_{jk'} - \mu_{j'k'}) = 0$

In all cases, the comparison is done after the n's have been collapsed into cell means--info about n_{jk} differences are lost

Unweighted Means Method

- In regression terms, the SS for each effect is computed *after* all other effects have been removed from the model
  - Analogous to semi-partial correlation
  - Remove induced correlations before calculating the SS for each effect
  - Reflect "unique" contributions of each effect

Weighted Means Method

- Second most common approach
- Difference in n use to weight the means
- For example:
  - $\overline{x}_{jk} = \sum_i x_{jki}/n_{jk}$ (cell)
  - $\overline{x}_{j\cdot} = \sum_i \sum_k x_{iki}/\sum n_{jk}$ (marg)
  - $\overline{x}_{\cdot k} = \sum_i \sum_j x_{ik}/\sum n_{jk}$ (marg)

	B1	B2	
A1	11 9 5 15 10 **10** n = 5	12 18 16 14 **15** n = 4	**12.22**
A2	7 4 10 11 **8** n = 4	9 4 8 **7** n = 3	**7.57**
	9.11	**11.57**	

Weighted Means Method

- Effect of A is calculated using 12.22 and 7.57 (smaller than unweighted means difference)
- Effect of B is calculated using 9.11 and 11.57 (larger than unweighted means difference)

	B1	B2	
A1	11, 9, 5, 15, 10 **10** n = 5	12, 18, 16, 14 **15** n = 4	**12.22**
A2	7, 4, 10, 11 **8** n = 4	9, 4, 8 **7** n = 3	**7.57**
	9.11	**11.57**	

Weighted Means Method

- Formally, the hypotheses are:
  - H_0 for A : $\sum \frac{n_{jk}\mu_{jk}}{n_{j\cdot}} - \sum \frac{n_{j'k}\mu_{j'k}}{n_{j'\cdot}} = 0$
  - H_0 for B: $\sum \frac{n_{jk}\mu_{jk}}{n_{\cdot k}} - \sum \frac{n_{jk'}\mu_{jk'}}{n_{\cdot k'}} = 0$
  - H_0 for AxB: $(\mu_{jk} - \mu_{j'k}) - (\mu_{jk'} - \mu_{j'k'}) = 0$

Notice the addition of the cell sample sizes for effects of A and B

Interaction looks the same

Weighted Means Method

- In regression terms, the SS for each effect is computed *before* all other effects have been removed from the model
  - Analogous to simple correlation
  - Induced correlations: variance from other effects are picked up by a given effect SS
  - No longer seeing the unique contributions

Weighted v Unweighted

- What do the differences in n's reflect?
  - Differences in reflect relative frequency of the conditions in the population:
    use **weighted means**
  - Differences due to some aspect of the treatment:
    use **weighted means**
  - Otherwise, use unweighted means method
    - the differences carry no *real* meaning

Weighted v Unweighted

- What leads to unequal n's?
  - Patients versus controls
  - Loss of observations due to difficulty of one condition relative to other conditions
  - Loss of observations due to random choice of trial types
  - Loss of observations due to technical difficulties
  - Loss of participants

Weighted v Unweighted

- How do you identify the correct SS for weighted and unweighted models?
  - Different ways to calculate SS's for a model
    - Type I
    - Type II
    - Type III
  - Like regression, these methods are dependent on how you want to look at the contributions of each term in the model

Type I SS

- Hierarchical Decomposition
  - Each term adjusted *only* for the terms that have already been entered in the model
    - Weighted SS for the first term entered
    - Sequential SS for the second term
    - Correct SS for the interaction
  - In SPSS: order matters
    - List variables in specific order

NOTE: run multiple times with each effect as the first variable and combine to get weighted means analysis

Type II SS

- Factor Sequential
  - Each term is adjusted for other effects that *do not include that term* in the model
  - In SPSS: the two effects will be resolved without the contribution of each other.
  - This gives you the sequential SS for your effects (as if each was the 2nd term)

NOTE: this ends up being something in between a weighted and unweighted analysis.

Type III SS

- Unweighted Analysis
  - Each term is adjusted for all relevant terms in the model
    - Reflects unique contribution of each variable
  - Gives you the unweighted SS for each effect and the correct interaction
  - In SPSS: Type III is the default (but be sure to check!)

NOTE: Use this for unweighted analyses

Example

- Going back to the previous example:

	B1	B2
A1	11 9 5 15 10	12 18 16 14
A2	7 4 10 11	9 4 8

Weighted analysis:

1. Run ANOVA with A entered first using Type I SS
2. Run ANOVA with B entered first using Type I SS
3. Combine the results

Example

With A entered first

	Type I SS	df	MS
A	85.17	1	85.17
B	22.43	1	22.43
AxB	34.84	1	34.84
Error	116.00	12	9.67
Total	258.44	15	

With B entered first

	Type I SS	df	MS
B	23.83	1	23.83
A	83.76	1	83.76
AxB	34.84	1	34.84
Error	116.00	12	9.67
Total	258.44	15	

Weighted Analysis

	SS	df	MS
A	85.17	1	85.17
B	23.83	1	23.83
AxB	34.84	1	34.84
Error	116.00	12	9.67
Total	258.44	15	

Example

- Going back to the previous example:

	B1	B2
A1	11 9 5 15 10	12 18 16 14
A2	7 4 10 11	9 4 8

Unweighted analysis:

Run ANOVA with Type III SS

	Type III SS	df	MS
A	96.77	1	96.77
B	15.48	1	15.48
AxB	34.84	1	34.84
Error	116.00	12	9.67
Total	258.44	15	

Example

- Comparison

	Weighted SS	Unweighted SS
A	85.17	96.77
B	23.83	15.48
AxB	34.84	34.84
Error	116.00	116.00
Total	258.44	258.44

Which is appropriate?
Depends on what the unequal n's mean!

Reporting Stats

- **Be consistent**: match descriptive and inferential stats
- Weighted analysis:
  - Report weighted means
  - ANOVA values should be from weighted analysis (using Type I SS repeatedly)
- Unweighted analysis:
  - Report unweighted means
  - ANOVA values should be from unweighted analysis (Type III SS)

Contrasts and Unequal n's

- **Just one minor change in the way you use the equation:**

$$SS_\psi = \frac{\psi^2}{\sum (c^2/n)}$$

Square each weight and dived by the cell n

ANOVA assumptions ≠ n's

- Most studies concerned with homogeneity of variance and normality (e.g., Milligan et al., 1987)
- Homogeneity of variance
  - Simulations paired various sample size patterns with various unequal variances
  - Result: unbalanced ANOVA is *very* sensitive to inhomogeneity
  - Type I error rates can be too high or too low depending on the exact mapping of variance to sample size

ANOVA assumptions ≠ n's

- Normality
  - News is far more promising
  - Unbalanced ANOVA is almost as robust to normality violations as a balanced ANOVA
- Upshot?
  - Worry about homogeneity of variance
  - Do not worry about normality
  - Better yet, try not to have unequal n's!

Interim summary

- Unequal n's happen
- To deal with them…
  - Know why the n's are not equal
  - Understand weighted v unweighted analyses
- Be consistent with your stats
- Be clear in your results sections

Quasi F-Ratios

What to do when you don't have an appropriate error term

Quasi-F ratios

- Sometimes we do not have the error terms we need to assess certain effects in the model (think E[MS])
- We can "create" an F value that will test the effect by pooling the available values
- Pooling produces a "quasi-F" statistic
  - F will have specific degrees of freedom
  - F can be used to assess significance

Quasi-F ratios

- Example
  - CRF-pqr, A, B, and C as random effects

	E(MS)
A	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 + nr\sigma_{\alpha\beta}^2 + nqr\sigma_\alpha^2$
B	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + np\sigma_{\beta\gamma}^2 + nr\sigma_{\alpha\beta}^2 + npr\sigma_\beta^2$
C	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + np\sigma_{\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 + npq\sigma_\gamma^2$
AxB	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nr\sigma_{\alpha\beta}^2$
AxC	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2$
BxC	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + np\sigma_{\beta\gamma}^2$
AxBxC	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2$
Residual	σ_ε^2

- **Which effects can we test?**
- **Which can we not?**

Quasi-F ratios

- Focus on effect of A:
  $E(MS_A) = \sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 + nr\sigma_{\alpha\beta}^2 + nqr\sigma_\alpha^2$
  - Need to isolate $nqr\sigma_\alpha^2$
  - Need to remove $\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 + nr\sigma_{\alpha\beta}^2$

	E(MS)
A	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 + nr\sigma_{\alpha\beta}^2 + nqr\sigma_\alpha^2$
B	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + np\sigma_{\beta\gamma}^2 + nr\sigma_{\alpha\beta}^2 + npr\sigma_\beta^2$
C	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + np\sigma_{\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 + npq\sigma_\gamma^2$
AxB	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nr\sigma_{\alpha\beta}^2$
AxC	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2$
BxC	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + np\sigma_{\beta\gamma}^2$
AxBxC	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2$
Residual	σ_ε^2

Use combinations of other MS values
e.g.,
$E(MS_{AxB}) + E(MS_{AxC}) - E(MS_{AxBxC})$

Quasi-F ratios

- Focus on effect of A:
  $E(MS_A) = \sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 + nr\sigma_{\alpha\beta}^2 + nqr\sigma_\alpha^2$
  - Need to isolate $nqr\sigma_\alpha^2$
  - Need to remove $\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 + nr\sigma_{\alpha\beta}^2$

$$\begin{aligned} E(MS_{AxB}) &= \sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nr\sigma_{\alpha\beta}^2 \\ + E(MS_{AxC}) &= +\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 \\ - E(MS_{AxBxC}) &= -\sigma_\varepsilon^2 - n\sigma_{\alpha\beta\gamma}^2 \\ \hline E(MS_{AxB}) + E(MS_{AxC}) - E(MS_{AxBxC}) &= \sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nr\sigma_{\alpha\beta}^2 + nq\sigma_{\alpha\gamma}^2 \end{aligned}$$

This means that the effect of A can be evaluated using the quasi-F'

$$F' = \frac{MS_A}{MS_{AxB} + MS_{AxC} - MS_{AxBxC}}$$

Quasi-F ratios

- General form for quasi-F'

$$F' = \frac{MS_1}{MS_2 + MS_3 - MS_4}$$

  - Degrees of freedom for the numerator is just the df_1 (df for the effect you are testing)
  - Degrees of freedom for the denominator must also be pooled. Use nearest integer value to:

$$df = \frac{(MS_2 + MS_3 - MS_4)^2}{MS_2^2/df_2 + MS_3^2/df_3 + MS_4^2/df_4}$$

Quasi-F ratios

- Potential problem:
  - Depending on the effect sizes, this formula can yield a negative denominator

$$F' = \frac{MS_1}{MS_2 + MS_3 - MS_4}$$

This can be circumvented by using a variation on the formula

$$F'' = \frac{MS_1 + MS_4}{MS_2 + MS_3}$$

$$F'' = \frac{MS_A + MS_{AxBxC}}{MS_{AxC} + MS_{AxB}}$$

Quasi-F ratios

- Degrees of freedom for F"
  - Numerator

$$\nu_1 = \frac{(MS_1 + MS_4)^2}{MS_1^2/df_1 + MS_4^2/df_4}$$

  - Denominator

$$\nu_2 = \frac{(MS_2 + MS_3)^2}{MS_2^2/df_2 + MS_3^2/df_3}$$

Quasi-F ratios

- What would you do for the effect of B?
- What would you do the effect of C?

	E(MS)
A	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 + nr\sigma_{\alpha\beta}^2 + nqr\sigma_\alpha^2$
B	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + np\sigma_{\beta\gamma}^2 + nr\sigma_{\alpha\beta}^2 + npr\sigma_\beta^2$
C	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + np\sigma_{\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 + npq\sigma_\gamma^2$
AxB	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nr\sigma_{\alpha\beta}^2$
AxC	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2$
BxC	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + np\sigma_{\beta\gamma}^2$
AxBxC	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2$
Residual	σ_ε^2

$$F' = \frac{MS_1}{MS_2 + MS_3 - MS_4}$$

$$F'' = \frac{MS_1 + MS_4}{MS_2 + MS_3}$$

MS_1?

MS_2?

MS_3?

MS_4?

Quasi-F ratios

- Suppose I give you this table, but I tell you A, B, and C are random variables

	df	SS	MS	F
A	2	10.72	5.36	0.42
B	2	656.06	328.03	25.57
C	1	16.00	16.00	1.25
AxB	4	319.78	79.95	6.23
AxC	2	0.17	0.085	0.007
BxC	2	46.16	23.08	1.80
AxBxC	4	1.67	0.42	0.033
residual	18	231.00	12.33	
Total	35	1281.56		

	E(MS)
A	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 + nr\sigma_{\alpha\beta}^2 + nqr\sigma_\alpha^2$
B	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + np\sigma_{\beta\gamma}^2 + nr\sigma_{\alpha\beta}^2 + npr\sigma_\beta^2$
C	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + np\sigma_{\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2 + npq\sigma_\gamma^2$
AxB	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nr\sigma_{\alpha\beta}^2$
AxC	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + nq\sigma_{\alpha\gamma}^2$
BxC	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2 + np\sigma_{\beta\gamma}^2$
AxBxC	$\sigma_\varepsilon^2 + n\sigma_{\alpha\beta\gamma}^2$
Residual	σ_ε^2

What's the problem?

$$F'' = \frac{MS_1 + MS_4}{MS_2 + MS_3}$$

How can you fix it?

$$\nu_1 = \frac{(MS_1 + MS_4)^2}{MS_1^2/df_1 + MS_4^2/df_4} \quad \nu_2 = \frac{(MS_2 + MS_3)^2}{MS_2^2/df_2 + MS_3^2/df_3}$$

Quasi-F ratios

- Effect of A
  - $F''(2, 4) = 0.07$, $p = 0.93$
- Effect of B
  - $F''(2, 6) = 3.19$, $p = 0.13$
- Effect of C
  - $F''(1, 2) = 0.71$, $p = 0.49$

Quasi-F ratios

- Distribution of quasi-F values (F' or F")
  - Not actually a central F
  - Central F is a good approximation of the distribution
- These principles can be used any time you need to figure out an error term, provided you can figure out E(MS) values

Quasi-F ratios & Contrasts

- How do you handle contrasts?
- No single clear approach
  - If you use the F', then the same denominator and df can be used for the contrasts.
  - Common approach: separate tests on subsets of data
- Quasi-F's for procedure
  - Justify ignoring irrelevant factors
  - Proceed with simpler model

Summary so far

- Weighted and unweighted analyses for unequal n's: know when to use them
- Quasi F ratios:
  - F' or F"
  - Pay attention to kinds of effects you have!

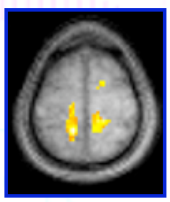
Factorial Design Walk-through

What constitutes a complete analysis?

What are the steps?

- Example: fMRI and spatial learning
  - ∗ All participants were scanned while learning three different environments
    - One from the ground-level perspective
    - One from the aerial perspective
    - One from a "hybrid" perspective (aerial-with-turns)
  - ∗ Want to know the effect of condition and hemisphere in the anterior superior parietal cortex (ROI defined from a previous study)

Data & Predictions



- Data
  - ∗ Extract percent signal change (relative to baseline)
    - For each participant (n = 14)
    - In each condition (p = 3)
    - In each hemisphere (q = 2)
  - ∗ Predictions
    - Ground vs. Aerial (replication)
    - Two alternatives for hybrid condition
      - » If area involved in orientation, hybrid = ground > aerial
      - » If not, ground > hybrid = aerial

The Data

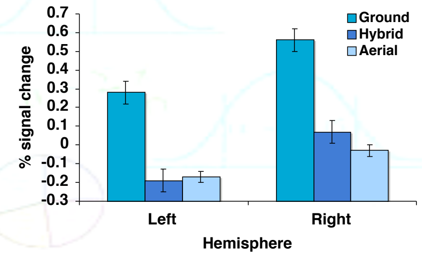
- Look at the data!

	Left	Right	Marginal
Ground	0.28 (0.06)	0.56 (0.08)	0.42 (0.06)
Hybrid	-0.19 (0.06)	0.07 (0.06)	-0.06 (0.04)
Aerial	-0.17 (0.05)	-0.03 (0.02)	-0.10 (0.03)
Marginal	-0.02 (0.04)	0.20 (0.03)	

  - ∗ Main effect of hemisphere?
  - ∗ Main effect of condition?
  - ∗ Interaction? (Let's look graphically)

Sphericity and Contrasts

- Look at the data!



Sphericity and Contrasts

- ANOVA table

Source	SS	df	MS	F	G-G p
BLOCK	.458	13	0.035		
HEMI	1.072	1	1.072	15.36	.002
Error(HEMI)	0.908	13	0.070		
COND	4.658	2	2.329	38.59	<.001
Error(COND)	1.569	26	0.060		
HEMI * COND	0.075	2	0.038	1.215	.313
Error(HEMI*COND)	0.805	26	0.031		

Main Effects

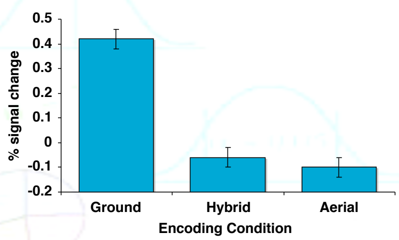
- Step through each one systematically
- Main effect of hemisphere
  - Is sphericity met? NOT RELEVANT!
  - Significant effect p = 0.002
  - Effect size (η_p^2 = 0.54 or η_G^2 = 0.25)
  - Only two levels:
    - Conclude that right superior parietal cortex was more active than left superior parietal cortex

Main Effects

- Main effect of condition
  - ∗ Is sphericity met? G-Gε = 0.74 (no sig. violation)
  - ∗ Significant effect $p < 0.001$ (G-G corrected)
  - ∗ Effect size ($\eta_p^2 = 0.75$ or $\eta_G^2 = 0.59$)
  - ∗ Three levels--how do they differ?
    - Start with the graph
    - Keep in mind the predictions as well

Main Effects

- What contrasts would be interesting?



Main Effects

- Assume sphericity is NOT met here (even though it is)
  - ∗ What data for a given subject is relevant?

sub	L_G	L_H	L_A	R_G	R_H	R_S
1	0.38	-0.62	-0.09	1.21	0.09	0.02
2	0.20	-0.08	-0.26	0.73	0.14	0.11
3	0.36	-0.07	-0.03	0.75	-0.13	0.06
4	0.22	-0.10	-0.02	0.74	-0.01	-0.02
5	0.37	-0.09	-0.23	0.14	0.38	-0.01
6	0.36	-0.06	-0.25	0.28	0.05	-0.16
7	-0.16	-0.76	-0.09	0.82	0.14	0.09
8	0.58	-0.12	-0.03	0.44	0.09	-0.03
9	-0.04	0.04	-0.78	0.35	0.28	0.03
10	0.39	-0.21	-0.09	0.59	0.08	-0.08
11	0.57	-0.16	-0.17	0.72	-0.49	-0.17
12	0.58	-0.07	-0.09	0.72	0.24	-0.10
13	0.14	-0.09	-0.14	0.17	0.04	-0.02
14	-0.03	-0.20	-0.09	0.20	0.06	-0.08

Main Effects

- Assume sphericity is NOT met here (even though it is)
  - ∗ What data for a given subject is relevant?
    - Marginal means for each subject? NO
    - Cell means for each subject? YES
      - » Contrast value would be the same either way
      - » Better estimate of residual error with full set
  - ∗ How do you set up the weights?
    - Weight every cell mean to construct contrast

Ground v Aerial & Hybrid

- Determine the weights first:

		Ground	Hybrid	Aerial
	c	+2	-1	-1
Left	1	+2	-1	-1
Right	1	+2	-1	-1

Ground v Aerial & Hybrid

c	2	-1	-1	2	-1	-1		
sub	L_G	L_H	L_A	R_G	R_H	R_S	ψ	ψ^2
1	0.38	-0.62	-0.09	1.21	0.09	0.02	3.78	14.27
2	0.20	-0.08	-0.26	0.73	0.14	0.11	1.95	3.82
3	0.36	-0.07	-0.03	0.75	-0.13	0.06	2.40	5.78
4	0.22	-0.10	-0.02	0.74	-0.01	-0.02	2.06	4.23
5	0.37	-0.09	-0.23	0.14	0.38	-0.01	0.97	0.93
6	0.36	-0.06	-0.25	0.28	0.05	-0.16	1.71	2.94
7	-0.16	-0.76	-0.09	0.82	0.14	0.09	1.94	3.76
8	0.58	-0.12	-0.03	0.44	0.09	-0.03	2.12	4.51
9	-0.04	0.04	-0.78	0.35	0.28	0.03	1.04	1.09
10	0.39	-0.21	-0.09	0.59	0.08	-0.08	2.25	5.07
11	0.57	-0.16	-0.17	0.72	-0.49	-0.17	3.58	12.84
12	0.58	-0.07	-0.09	0.72	0.24	-0.10	2.63	6.92
13	0.14	-0.09	-0.14	0.17	0.04	-0.02	0.81	0.66
14	-0.03	-0.20	-0.09	0.20	0.06	-0.08	0.65	0.42
						$\sum$	27.90	67.23

Ground v Aerial & Hybrid

$$\psi = 27.90/14 = 1.99$$

$$SS_{\psi} = \frac{14 * 1.99^2}{12} = 4.63$$

$$SS_{res\_1} = \frac{67.23 - (27.90^2/14)}{12} = 0.97$$

$$MS_{res\_1} = 0.97/(14-1) = 0.07$$

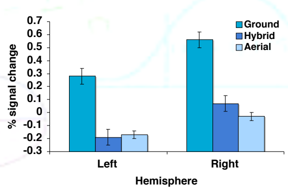
$$F_{\psi} = 4.63/0.07 = 62.22$$

Ground v Aerial & Hybrid

- All other aspects remain the same
  - How much of the effect is accounted for?
    - % effect = SS_{ψ} /SS_{effect} = 4.63/4.66 = 0.99
    - Ground > Aerial & Hybrid
  - Would we need to do more?
    - Not really
    - Only other interesting hypothesis from our prediction is Aerial v Hybrid, but there is no variance left for this contrast

Interactions

- Same procedure applies to contrast-contrast interactions (if it had been significant)
  - Define weights for each variable
  - Example: Aerial v Hybrid x Left v Right



Aerial v Hybrid x Left v Right

- Determine the weights first:

		Ground	Hybrid	Aerial
	c	0	+1	-1
Left	+1	0	+1	-1
Right	-1	0	-1	+1

Aerial v Hybrid x Left v Right

c	0	1	-1	0	-1	1
sub	L_G	L_H	L_A	R_G	R_H	R_S
1	0.38	-0.62	-0.09	1.21	0.09	0.02
2	0.20	-0.08	-0.26	0.73	0.14	0.11
3	0.36	-0.07	-0.03	0.75	-0.13	0.06
4	0.22	-0.10	-0.02	0.74	-0.01	-0.02
5	0.37	-0.09	-0.23	0.14	0.38	-0.01
6	0.36	-0.06	-0.25	0.28	0.05	-0.16
7	-0.16	-0.76	-0.09	0.82	0.14	0.09
8	0.58	-0.12	-0.03	0.44	0.09	-0.03
9	-0.04	0.04	-0.78	0.35	0.28	0.03
10	0.39	-0.21	-0.09	0.59	0.08	-0.08
11	0.57	-0.16	-0.17	0.72	-0.49	-0.17
12	0.58	-0.07	-0.09	0.72	0.24	-0.10
13	0.14	-0.09	-0.14	0.17	0.04	-0.02
14	-0.03	-0.20	-0.09	0.20	0.06	-0.08

- $\sum c^2$?
- n for SS_ψ?
- n for SS_{res_i}?

Factorial Summary

- Keep the big picture in mind
- Deal with effects separately
- Contrasts & sphericity
  - Use all of the subject data at the level it was entered into the ANOVA
  - Be VERY careful about:
    - $\sum c^2$
    - Correct number of observations
  - All of this is easy in a spreadsheet or Matlab