HUDM 5123 - Linear Models and Experimental Design
Notes 06 - Review and Linear Contrasts

1 Review

1.1 Simple and Multiple Linear Least Squares Regression

1. Sample mean

$$\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{Y_1 + Y_2 + \cdots + Y_n}{n}.$$

2. Sample variance

$$s_Y^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(Y_i - \bar{Y}\right)^2.$$

3. Sample covariance

$$s_{XY} = \frac{1}{n-1}\sum_{i=1}^{n}\left[\left(X_i - \bar{X}\right)\left(Y_i - \bar{Y}\right)\right].$$

4. Sample correlation

$$r_{XY} = \frac{s_{XY}}{\sqrt{s_X^2 s_Y^2}}.$$

5. Multiple regression equation and interpretation

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i + \cdots + \beta_p X_i + \epsilon_i.$$

6. R^2 equation and meaning

$$R^2 = \frac{\text{TSS} - \text{RSS}}{\text{TSS}} = \frac{\text{RegSS}}{\text{TSS}},$$

7. Sample residual variance and standard error

$$s^2 = \frac{1}{n-2}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2 \qquad s = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}$$

1.2 OLS Regression Diagnostics

1. Four assumptions about the error term (linearity, constant variance, normality, independence) and how to use data to diagnose violations

$$\epsilon_i \overset{iid}{\sim} N\left(0, \sigma_\epsilon^2\right),$$

2. No multicollinearity assumption and variance inflation factors

$$\text{VIF} = \frac{1}{1 - R_j^2},$$

3. Leverage, discrepancy, and influence

1.3 One-way ANOVA

1. Checking assumptions for categorical predictor (group variance ratios and qq plots, for example)
2. Dummy coding vs deviation coding and interpretation of coefficients
3. Incremental F test

$$F_0 = \frac{(RSS_R - RSS_F)/(df_R - df_F)}{RSS_F/df_F}$$

4. Specification of full and reduced models
5. ANOVA table

Source	Sum of Squares	df	Mean Square	F
Regression	RegSS	$df_R - df_F$	$\frac{RegSS}{df_R - df_F}$	$\frac{RegMS}{RMS}$
Residuals	RSS	df_F	$\frac{RSS}{df_F}$	

6. ANOVA omnibus test null and alternative hypotheses for testing the effect of factor A on the outcome (assume factor A has a levels)

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$$
$$H_1 : \text{at least one } \mu_j \neq \mu_k$$

1.4 ANCOVA

1. Baseline imbalance measures (standardized mean difference d and variance ratio r

$$d = \frac{\bar{X}_T - \bar{X}_C}{s_{\text{pooled}}}, \text{ where } s_{\text{pooled}} = \sqrt{\frac{(N_T - 1)s_T^2 + (N_C - 1)s_C^2}{N_T + N_C - 2}}$$
$$r = \frac{s_T^2}{s_C^2}.$$

2. Full and reduced models for ANCOVA vs ANOVA
3. Meaning of 'controlling for' a covariate
4. Explain and predict changes when going from ANOVA to ANCOVA (residual SS, residual df, treatment group SS, F statistics, p-value
5. Treatment by covariate interaction full and reduced models and interpretation
6. Relationships between adjusted and unadjusted means (Fig. 9.5 from Maxwell and Delaney)

1.5 Categorical Predictors and Interactions

1. Two-factor notation including cell means, marginal means, and grand mean

	C_1	C_2	$\cdots$	C_c	
R_1	μ_{11}	μ_{12}	$\cdots$	μ_{1c}	$\mu_{1\bullet}$
R_2	μ_{21}	μ_{22}	$\cdots$	μ_{2c}	$\mu_{2\bullet}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
R_r	μ_{r1}	μ_{r2}	$\cdots$	μ_{rc}	$\mu_{r\bullet}$
	$\mu_{\bullet 1}$	$\mu_{\bullet 2}$	$\cdots$	$\mu_{\bullet c}$	$\mu_{\bullet\bullet}$

where

$$\mu_{j\bullet} = \frac{1}{c}\sum_{k=1}^{c}\mu_{jk} \qquad \mu_{\bullet k} = \frac{1}{c}\sum_{j=1}^{r}\mu_{jk} \qquad \mu_{\bullet\bullet} = \frac{1}{rc}\sum_{j=1}^{r}\sum_{k=1}^{c}\mu_{jk}$$

2. Definition of main effects, simple effects and interaction effects
3. Interaction plot interpretation
4. Dummy coding vs deviation coding and interpretation of coefficients
5. Full and reduced models for testing (a) interaction and (b) main effects
6. Testing decision tree (main effects vs simple effects)

2 Linear Contrasts

The incremental F test used for the one-way ANOVA is called the *omnibus F test* because it tests the null hypothesis that the treatment means are *all* equal. For example, for the one-way ANOVA with five groups, $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$. If this null hypothesis is rejected, we simply conclude the group means are not all equal. We do not learn which group means or how many pairs of group means differ. This is dissatisfying because researchers typically want to know specifically how the group means differed. As an example, consider the following experiment with one factor that has five levels at different dosages of a drug.

conventional		experimental		
drug1	drug2	drug3	drug4	placebo
μ_1	μ_2	μ_3	μ_4	μ_5

Linear contrast comparisons allow us to test individual groups (or weighted linear combinations of them) against one another. Some individual comparisons that might be of interest:

1. the third drug vs. the placebo,
2. the average of conventional drugs vs. the average of new drugs,
3. drug2 vs. drug 4,
4. the average effect of all drugs vs. placebo.

Comparisons that focus on the difference between only two means are called *pairwise comparisons*. The first and third examples above focus on pairwise comparisons, while the others, which involve averages of group means, are called *complex comparisons*. A *contrast*, denoted by ψ, is formed as a linear combination of the group means, μ_j, with coefficients c_j.

$$\psi = c_1\mu_1 + c_2\mu_2 + \cdots + c_j\mu_j = \sum c_j\mu_j$$

Here are the contrasts for the four examples given above:

1. $\psi_1 = 0 * \mu_1 + 0 * \mu_2 + 1 * \mu_3 + 0 * \mu_4 + -1 * \mu_5 = \mu_3 - \mu_5,$
2. $\psi_2 = 0.5 * \mu_1 + 0.5 * \mu_2 + -0.5 * \mu_3 + -0.5 * \mu_4 + 0 * \mu_5 = \frac{\mu_1 + \mu_2}{2} - \frac{\mu_3 + \mu_4}{2},$
3. $\psi_3 = 0 * \mu_1 + 1 * \mu_2 + 0 * \mu_3 + -1 * \mu_4 + 0 * \mu_5 = \mu_2 - \mu_4,$
4. $\psi_4 = .25 * \mu_1 + .25 * \mu_2 + .25 * \mu_3 + .25 * \mu_4 + -1 * \mu_5 = \frac{\mu_1 + \mu_2 + \mu_3 + \mu_4}{4} - \mu_5.$

The null hypothesis for all the contrasts above is the same: $H_0 : \psi = 0.$ The only restriction placed on a contrast is that the coefficients sum to zero: $\sum c_j = 0.$ Also, although not required, it is good practice to standardize all contrasts so that the sum of absolute values of c_j's is 2.

2.1 Hypothesis Testing for Contrasts

The t statistic may be used to test the null hypothesis that an ANOVA contrast is equal to zero.

$$t = \frac{\hat{\psi}}{SE_{\psi}}, \text{ where } SE_{\psi} = \sqrt{MS_{\text{error}} \sum \frac{c_j^2}{n_j}}.$$

Note that MS_{error}, pronounced 'mean squared error', is the residual variance, s^2, from section 1.1 above. The formula for the $(1-\alpha)\%$ confidence interval for a contrast is

$$\hat{\psi} - t^* SE_{\psi} \leq \psi \leq \hat{\psi} + t^* SE_{\psi},$$

where, t^* is the upper $(1-\alpha)/2$ critical value from the t distribution with df_F degrees of freedom.

3 Testing Contrasts for the One Factor Design

24 participants with mild hypertension were randomly assigned to one of four treatment groups. The combination group combines all aspects of the other three treatments. The scores in the table are the systolic blood pressure readings for each subject 2 weeks after termination of treatment. Test the pairwise comparison between drug therapy and biofeedback. Test the complex comparison between the combined treatment and the average of the other treatments.

Systolic Blood Pressure Data				
	Drug Therapy	Biofeedback	Diet	Combination
	84	81	98	91
	95	84	95	78
	93	92	86	85
	104	101	87	80
		80	94	81
		108		
Mean	94.0	91.0	92.0	83.0
Var	67.3	132.0	27.5	26.5

The one-way ANOVA for the treatment factor is not significant ($F(3, 16) = 1.66$; $p = .22$). Suppose, instead, that the research questions were more specific. First, the researchers hypothesized that there would be a difference between the efficacy of drug therapy and biofeedback. Second, they hypothesized that the combination level of the factor would be more effective than the other three treatments when each taken alone. These two hypotheses may be formulated as linear contrasts; the first is a pairwise comparison and the second is a complex comparison.

1. pairwise comparison between drug therapy and biofeedback

$$\psi_1 = 1\mu_1 - 1\mu_2 + 0\mu_3 + 0\mu_4$$

2. complex comparison between the combined and the average of the other treatments

$$\psi_2 = 1/3\mu_1 - 1/3\mu_2 + 1/3\mu_3 - 1\mu_4$$

See accompanying R code for the tests.

4 Recap for Two-Way ANOVA

In the analysis of a two-way ANOVA design, the first step is to create an interaction plot of the marginal means. Next, after checking assumptions, run the ANOVA and assess the significance of the two-way interaction via the p-value. Combine that evidence with what you see graphically to make a decision about how you should proceed. If there is no graphical evidence of an interaction (i.e., trace plots of marginal means appear to have similar, if not identical, slope) and the test for interaction is non-significant, proceed by examining the ANOVA test results for the main effects of the two factors. Then follow up significant main effects results with pairwise comparisons that average over the levels of the other factor. If, on the other hand, there is graphical evidence of an interaction and the test for interaction is significant, proceed by examining simple main effects. Pick one factor to condition on and then look for omnibus test results of the other factor at each level of the first. Follow up any that are significant with simple pairwise comparisons.

If there is ambiguity in the sense that the plot and the statistical significance results appear to disagree (e.g., the plot seems to show strongly non-parallel slopes across groups but the test for interaction is non-significant) then you will need to use your best judgment as to how to proceed by examining other factors. Was the sample size very small so that a statistical test for interaction would likely be underpowered? If so, you might prefer to look at simple effects. Is the sample size very large so that even though the plot doesn't appear to show much in the way of different slopes across groups, the test for interaction is still highly significant? If so, you might favor a main effects approach over a simple effects approach. Is the interaction of a non-crossing type such that the main effects cannot cancel out? If so, you might prefer to look at main effects even in the face of a significant interaction. Etc. No matter which way you decide to parse the data, you should explain to the reader when you write up your results what the factors were that motivated your decision.

5 Two-Way ANOVA with Significant Interaction

This example, from Maxwell & Delaney, p. 310, has to do with a cognitive neuroscience study. Theoretical considerations imply that amnesic patients will have a deficit in explicit memory but not on implicit memory. Huntington's disease patients, on the other hand, are expected to have the opposite; they will have no deficit in explicit memory, but will have a deficit in implicit memory. Fifteen participants diagnosed with Amnesia, fifteen participants with Hungington's diesease, and a control group of fifteen participants with no

neurodegenerative disease were recruited for the study. Each participant was then randomly assigned (5 per group) to one of three tasks: (1) artificial grammar task, which consists of classifying letter sequences as either following or not following grammatical rules; (2) classification learning task, which consists of classifying hypothetical patients as either having or not having a certain disease based on symptoms probabilistically related to the disease; and (3) recognition memory task, which consists of recognizing particular stimuli as stimuli that have previously been presented during the task. There are five observations per cell of the design.

The key point to understand before proceeding to the analysis is that these tasks have been selected to map onto the theoretical differences between the three types of research participants. In particular, the first two tasks are known to reflect implicit memory processes, whereas the third task is known to reflect explicit memory processes. Thus, if the theory is correct, we would expect to see relatively higher scores on the first two tasks for the amnesic group but relatively higher scores on the third task for the Huntington group.

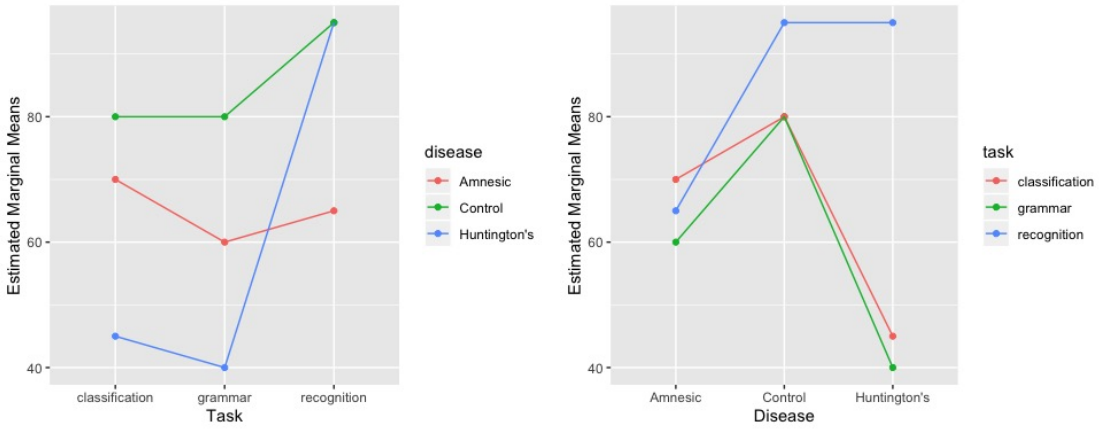

Figure 1: Interaction plots for the cognitive neuroscience data

The two-way ANOVA test of interaction between disease group and task type is significant ($F(4, 36) = 7.92; p = .0001$), suggesting that the task type has a differential effect on participants' recall ability depending on which disease group they are in. The interaction plot certainly appears to corroborate this two-way interaction when considering, in particular, the recognition task performance for the Huntington's group relative to the other two. Thus, we might decide to follow up the two-way interaction by testing simple main effects. To do so, we must decide which factor to condition on. On the one hand, we could explore the effect of disease category at each level of the task factor. On the other, we could explore the effect of task type at each level of disease category. Suppose we decide to explore the effect of task at each level of disease; that is, we will condition on the levels of the disease factor.

Based on the three simple omnibus tests of the main effect of task at each level disease, we find a significant task effect for the Huntington's group ($F(2, 36) = 29.3; p < .0001$) and non-significant results for the simple main effect of task for the amnesic group ($F(2, 36) = 0.8; p = .46$) and the control group ($F(2, 36) = 2.4; p = .11$). Following up on the significant simple main effect for the Huntington's group, simple pairwise comparisons of task levels for the Huntington's group reveal significant differences in mean score for grammar vs

recognition ($t = -6.92; p < .0001$) and classification vs recognition ($t = -6.29; p < .0001$) and no significant difference for classification vs grammar ($t = 0.63; p = .53$). Thus, we find significant evidence of a task effect such that those with Huntington's perform significantly worse on the first two tasks than the third. See accompanying R code.