Contrasts for Comparison of Means

If the analysis of variance produces a significant F-statistic, this signals the need for further analyses. The most common post ANOVA analysis is the comparison of means. There are several different analytic tools available for this type of analysis, but the most basic is ***contrast of means***. Contrasts are of the following basic form:

$$\sum_{i=1}^{v} c_i \cdot \mu_i, \text{ where } \sum_{i=1}^{v} c_i = 0 \text{ and at least one } c_i \neq 0.$$

Implementing contrasts using SAS Proc GLM is a relatively straightforward process. A SAS Proc GLM contrast statement has the following form:

contrast "label" effect values;

A contrast is initiated with the **contrast** statement. The **"label"** appears in either single or double quotation marks and is a self specified label which identifies the contrast to the user (e.g., "mean 1 vs. mean 2"). The **effect** indicates to SAS the name of the effect or treatment factor that appears in the model (e.g., laboratory). The **values** represent the coefficients of the contrasts (e.g., the c_i's).

Using the EPA laboratory data as an example, suppose that a person was interested in comparing the mean for the first laboratory with the mean for the second laboratory. The full SAS code for this analysis is:

```
title "Analysis of EPA Laboratory Data";
title2 "Using a Completely Randomized Design";
title3 "With a One-Way Treatment Structure";
data epa;
input laboratory ddt @@;
cards;
1 1005 1 1015 1 1033 1 1028 1 1023 1 1043
2  995 2 1008 2  976 2 1014 2 1011 2  982
3  950 3  975 3  988 3 1015 3 1008 3  994
;
run;
proc print data = epa;
run;
proc glm data = epa;
  class laboratory;
  model ddt = laboratory;
  contrast "Lab 1 vs. Lab 2" laboratory 1 -1  0;
run;
```

It might also be of interest to compare the first laboratory with the third and the second laboratory with the third. This could be accomplished by adding the following contrast statements to the above code:

```
  contrast "Lab 1 vs. Lab 3" laboratory 1  0 -1;
  contrast "Lab 2 vs. Lab 3" laboratory 0  1 -1;
```

11

Simultaneous Inference for Two or More Contrasts

It is also sometimes of interest to assess two or more contrasts simultaneously. Remembering that all SAS statements end with a semicolon, simultaneous contrasts can be created by replacing the semicolons of a set of contrasts with comma, except for the last in the set, which must have a semicolon. For example, suppose that it is of interest to simultaneously assess the contrast of the means for laboratory 1 and laboratory 3, and the contrast of the means for laboratory 2 and laboratory 3. The full SAS code with the individual contrasts and the simultaneous contrasts would appear as follows:

```
title "Analysis of EPA Laboratory Data";
title2 "Using a Completely Randomized Design";
title3 "With a One-Way Treatment Structure";
data epa;
input laboratory ddt @@;
cards;
1 1005 1 1015 1 1033 1 1028 1 1023 1 1043
2  995 2 1008 2  976 2 1014 2 1011 2  982
3  950 3  975 3  988 3 1015 3 1008 3  994
;
run;
proc print data = epa;
run;
proc glm data = epa;
  class laboratory;
  model ddt = laboratory;
  contrast "Lab 1 vs. Lab 2" laboratory 1 -1  0;
  contrast "Lab 1 vs. Lab 3" laboratory 1  0 -1;
  contrast "Lab 2 vs. Lab 3" laboratory 0  1 -1;
  contrast "Lab 1 vs. Lab 3 and Lab 2 vs. Lab 3" laboratory 1  0 -1,
                                                 laboratory 0  1 -1;
run;
```

Inference Using Orthogonal Polynomial Contrasts

As discussed in class, orthogonal polynomial contrasts are used to assess trends (linear, quadratic, etc.) in the response means when the treatment (factor) levels are quantitative. Many books on analysis of variance provide the contrast coefficients for equally spaced quantitative treatment levels up to $v = 6$ or 7 levels. If the treatment levels are not equally spaced a rather arduous set of mathematical calculations are needed to obtain the contrast coefficients. SAS provides a simple alternative. An *interactive matrix language* procedure in SAS, Proc IML, allows for the quick computation of the contrast coefficients for any number of treatment levels and any spacing among the treatment levels.

Example
Kuehl(2000) reports the results of an experiment which was conducted to test the effects of nirtorgen fertilizer on lettuce production. Five rates of ammonium nitrate were applied to four replicate plots in a completely randomized design. The response was the number of lettuce heads harvested from each plot and appear in the following table:

Treatment	Heads of Lettuce per Plot			
0	104	114	90	140
50	134	130	144	174
100	146	142	152	156
150	147	160	160	163
200	131	148	154	163

For these data, the orthogonal polynomials for the linear, quadratic, cubic and quartic contrasts can be found in Table A.2 on page 702 of the text. These values are:

Trend	c_1	c_2	c_3	c_4	c_5
Linear	-2	-1	0	1	2
Quadratic	2	-1	-2	-1	2
Cubic	-1	2	0	-2	1
Quartic	1	-4	6	-4	1

The SAS Proc GLM run for analyzing the above data, including the orthogonal polynomials is:

```
options ls = 80;
Title "Analysis of the Lettuce Heads Data";
title2 "Using Orthogonal Polynomial Contrasts";
data Lettuce;
input Heads Fertilizer @@;
cards;
104    0  114    0   90    0  140    0
134   50  130   50  144   50  174   50
146  100  142  100  152  100  156  100
147  150  160  150  160  150  163  150
131  200  148  200  154  200  163  200
;
proc glm data = Lettuce;
  class Fertilizer;
  model Heads = Fertilizer;
  contrast "Linear    " Fertilizer -2 -1  0  1  2;
  contrast "Quadratic" Fertilizer  2 -1 -2 -1  2;
  contrast "Qubic     " Fertilizer -1  2  0 -2  1;
  contrast "Quintic  " Fertilizer  1 -4  6 -4  1;
run;
```

Although this example only involves equally spaced treatment levels (0, 50, 100, 150 and 200), it is still serves the purpose well.

If you did not have your book available, or the treatment levels were not equally spaced or the number of treatment levels exceeded the maximum of 7, then an alternative approach would have to be devised.

The SAS Proc IML code for determining the orthogonal polynomial coefficients is:

```
OPTIONS LS=72;

 * LINEAR IS COL #2, QUAD IS COL #3, CUBIC IS COL # 4, ETC.;
 * CONTRASTS ARE VERTICAL NOT HORIZONTAL;

PROC iml;
 * Levels represents a vector of treatment levels;
   Levels = {0 50 100 150 200};

 * Contrast represents the matrix of orthogonal polynomial coefficients;
   Contrasts = ORPOL(Levels);

PRINT Levels Contrasts;
RUN;
```

You will notice that the coefficients generated by Proc IML do not appear to be the same as those provided in the text. However, these coefficients are simply a constant multiple of those provided in the text. For example, if you multiply each coefficient in the linear contrast by 3.162275, you will get the same linear contrast as provided in the text.

Example

A researcher in agronomy was interested in grain yield at five (5) different planting densities. An experiment was constructed in which the five planting densities (50%, 75%, 100%, 150% and 200% of normal) were randomly assigned to 15 equal sized plots (completely randomized design). The yield in bushels per plot appear in the following table:

	Planting Density (% of normal)				
	50%	75%	100%	150%	200%
	12.5	16.2	18.2	18.2	14.8
	11.8	16.0	20.8	19.5	16.9
	12.9	17.0	18.6	17.5	16.3
Means	12.4	16.4	19.2	18.4	16.0

The SAS Proc IML code for determining the orthogonal polynomial coefficients is:

```
OPTIONS LS=72;

 * LINEAR IS COL #2, QUAD IS COL #3, CUBIC IS COL # 4, ETC.;
 * CONTRASTS ARE VERTICAL NOT HORIZONTAL;

PROC iml;
 * Levels represents a vector of treatment levels;
   Levels = {50 75 100 150 200};

 * Contrast represents the matrix of orthogonal polynomial coefficients;
   Contrasts = ORPOL(Levels);

PRINT Levels Contrasts;
RUN;
```

The SAS code for the analysis of the data presented above appears below. Incorporate the appropriate coefficients for the linear and quadratic orthogonal polynomials.

```
options pageno = 1 pagesize = 45 linesize = 100;
data density;
input yield density @@;
cards;
12.5 50   16.2 75   18.2 100   18.2 150   14.8 200
11.8 50   16.0 75   20.8 100   19.5 150   16.9 200
12.9 50   17.0 75   18.6 100   17.5 150   16.3 200
;
proc print data = density;
run;
proc plot data = density;
  plot yield*density;
run;
proc glm data = density;
class density;
model yield = density;
contrast "Linear Trend"    density -0.539796 -0.332182 -0.124568  0.290659 0.705887;
contrast "Quadratic Trend" density  0.544488 -0.085746 -0.467316 -0.484466 0.493040;
run;
```

15