Contrasts

Contrasts (*Q*) allow you to ask more direct questions and therefore give you less ambiguous results than multiple comparison procedures.

Single df comparisons can be stated in terms of hypotheses. For example, in the switchgrass variety trial the comparison of Upland vs Lowland genotypes is:

$$(\mu_1 + \mu_3 + \mu_4 + \mu_5 + \mu_6 + \mu_7 + \mu_{10})/7 = (\mu_2 + \mu_8 + \mu_9)/3$$

or

$$3\mu_1 + 3\mu_3 + 3\mu_4 + 3\mu_5 + 3\mu_6 + 3\mu_7 + 3\mu_{10} - 7\mu_2 - 7\mu_8 - 7\mu_9 = 0$$

Contrast coefficients (c_i) associated with treatment means are based on hypothesis.

Contrasts

A contrast is defined as:

$$Q = \sum c_i Y_i$$

where

$$\sum c_i = 0$$

Contrasts

Equations

$$SS(Q) = MS(Q) = \frac{(\sum c_i Y_i)^2}{r \sum c_i^2}$$ Using treatment totals

$$SS(Q) = MS(Q) = \frac{r(\sum c_i \bar{Y}_i)^2}{\sum c_i^2}$$ Using treatment means

Contrasts

Switchgrass Example

$$SS(Q) = MS(Q) = \frac{r(\sum c_i \bar{Y}_i)^2}{\sum c_i^2}$$

$$MS_Q = \frac{6(3(12.6+14.2+12.7+9.4+6.7+13.9+9.1)-7(8.3+9.6+9.4))^2}{(3)^2+(-7)^2+(3)^2+(3)^2+(3)^2+(3)^2+(3)^2-(7)^2-(7)^2+(3)^2}$$

$$MS_Q = \frac{6(-44.3898)^2}{210} = 56.299$$

$$F = \frac{56.299}{8.651} = 6.51$$ P > F = 0.014

∴ Reject the null hypothesis and conclude that the mean yields of lowland and upland varieties in the trial differ.

Contrasts
Orthogonality

A pair of contrasts is orthogonal when:

$$\sum c_{1i}c_{2i} = 0$$

Contrasts
Orthogonality Example

Contrast	Treatment		
	1	2	3
1 vs. 2 + 3	2	-1	-1
2 vs. 3	0	1	-1
c_1c_2	0	-1	1

$$\sum c_{1i}c_{2i} = 2(0) + 1(-1) - 1(-1) = 0$$

Contrasts

Sets of contrasts may be tested simultaneously:

The hypothesis $\mu_1 = \mu_2 = \mu_3$

requires 2 df to test because

if $\mu_1 = \mu_2$ and $\mu_2 = \mu_3$ then $\mu_1 = \mu_3$

Contrasts

Rules for assigning coefficients:

<u>Groups compared are of equal size, assign</u>:

- + 1 to one group
- - 1 group to the other
- 0's to any treatment levels not in the comparison

Example: to compare the mean of treatments 1 & 2 to that of 4 & 5 the contrast coefficients are: 1 1 0 -1 -1

Contrasts

Rules for assigning coefficients:

Groups are of unequal size, assign:

- group 1 coefficients equal to the number of treatment levels in group 2
- group 2 negative coefficients equal to the number of treatment levels in group 1
- 0's to any treatment levels not in the comparison

Example: to compare the mean of treatments 1 & 2 to that of 3 & 4 & 5 the contrast coefficients are: 3 3 -2 -2 -2

Contrasts

Cover Crop Example

Source	**df**	**Treatment level / C_i**				
Treatment	4	clover	vetch	oat	wheat	rye
legume vs grass	1					
clover vs vetch	1					
oat vs wheat	1					
oat vs rye	1					
Error	15					
Total	19					

Contrasts

Cover Crop Example

Source	**df**	**Treatment level / C_i**				
Treatment	4	clover	vetch	oat	wheat	rye
legume vs grass	1	3	3	-2	-2	-2
clover vs vetch	1	1	-1	0	0	0
oat vs wheat	1	0	0	1	-1	0
oat vs rye	1	0	0	1	0	-1
Error	15					
Total	19					

If the complete set of contrasts is orthogonal the SS for the four contrasts will sum to the Treatment SS. Would that happen here?

Contrasts

Seed Treatment Example

Three acid treatments were applied to rice seed to determine effects on early seedling growth.

Treatment	Replications					Mean
Control	4.23	4.38	4.10	3.99	4.25	4.19
HCl	3.85	3.78	3.91	3.94	3.86	3.87
Propionic	3.75	3.65	3.82	3.69	3.73	3.73
Butyric	3.66	3.67	3.62	3.54	3.71	3.64

Data source: https://psfaculty.plantsciences.ucdavis.edu/agr205/Lectures/2011_Lectures/L4_Contrasts.doc

Contrasts

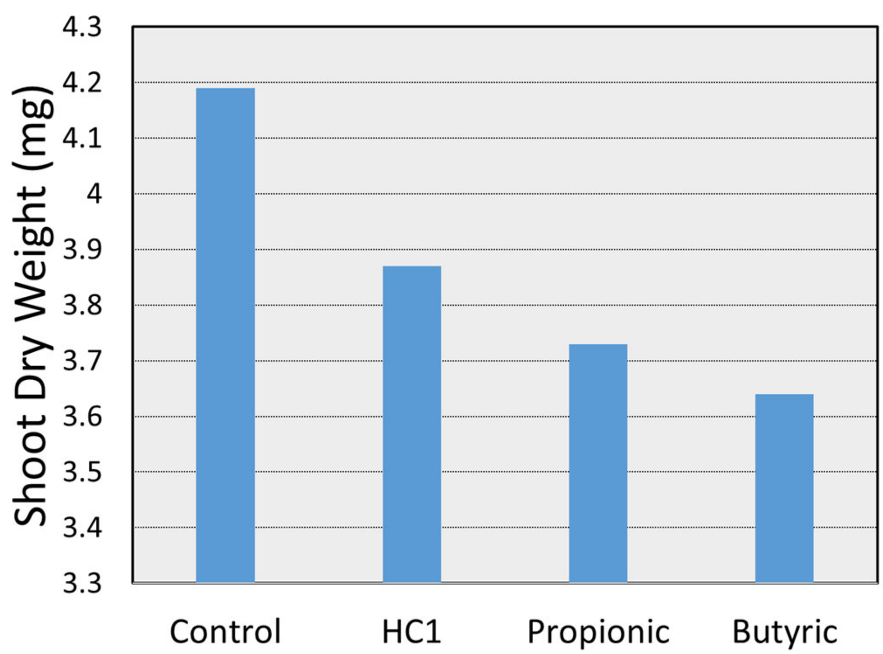
Seed Treatment Example

```
proc glm;
  class treatment;
  model ShootWt = treatment / ss3;
  means treatment;
  contrast 'Control vs acid' treatment 1 -3 1 1;
  contrast 'Inorganic vs organic' treatment 1 0 -2 1;
  contrast 'Between organics' treatment 1 0 0 -1;
run;
```

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Treatment	3	0.873695	0.291232	33.87	<.0001
Control vs acid	1	0.741482	0.741482	86.24	<.0001
Inorganic vs organic	1	0.112853	0.112853	13.13	0.0023
Between organics	1	0.01936	0.01936	2.25	0.1529
Error	16	0.13756	0.008598		
Corrected Total	19	1.011255			

* Note: this set of contrasts is orthogonal so contrast SS sum to treatment SS.
** Order of contrast coefficients required given in class level information.

Seed Treatment Example



Contrasts

Homework Example

	Treatment								
Comparison	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
	Oat	Oat	Oat	Wheat	Wheat	Wheat	Rye	Rye	Rye
		Vetch	Pea		Vetch	Pea		Vetch	Pea
Pure vs Mix									
Oat vs Wheat									
Oat vs Rye									
Wheat vs Rye									
Pea vs Vetch									
Oat Mix vs Wheat Mix									
Oat Mix vs Rye Mix									
Wheat Mix vs Rye Mix									

Contrasts vs. Estimates

For the contrast Pure vs Mix:

$$(\mu_1 + \mu_4 + \mu_7)/3 = (\mu_2 + \mu_3 + \mu_5 + \mu_6 + \mu_8 + \mu_9)/6$$

We normally write the coefficients as integers:

$$2(\mu_1 + \mu_4 + \mu_7) - 1(\mu_2 + \mu_3 + \mu_5 + \mu_6 + \mu_8 + \mu_9) = 0$$

The SAS code is:

```
contrast 'Pure vs Mix' trt 2 -1 -1 2 -1
-1 2 -1 -1;
```

Contrasts vs. Estimates

We are not really comparing *means* - we are comparing *effects* based on the linear additive model:

$0 = 2(\mu_1 + \mu_4 + \mu_7) - 1(\mu_2 + \mu_3 + \mu_5 + \mu_6 + \mu_8 + \mu_9)$

$= 2(\mu + T_1) + 2(\mu + T_4) + 2(\mu + T_7) - (\mu + T_2) - (\mu + T_3) - (\mu + T_5) - (\mu + T_6) - (\mu + T_8) - (\mu + T_9)$

$= 2T_1 + 2T_4 + 2T_7 - T_2 - T_3 - T_5 - T_6 - T_8 - T_9$

Contrasts vs. Estimates

<u>For the ESTIMATE statement the coefficients matter!</u>

If want to estimate the linear combination of:

$(\mu_1 + \mu_4 + \mu_7)/3 = (\mu_2 + \mu_3 + \mu_5 + \mu_6 + \mu_8 + \mu_9)/6$

you have to use .333 and .166 as the coefficients for the ESTIMATE statement.

In SAS you can use the DIVISOR option with the ESTIMATE statement to divide coefficients by a constant.

```
estimate ‘Pure vs Mix' trt -2 1 1 -2 1 1-2 1
1 / divisor=6;
```

Contrasts vs. Estimates
Switchgrass Example

To estimate the mean difference between Upland and Lowland genotypes:

$(\mu_1 + \mu_3 + \mu_4 + \mu_5 + \mu_6 + \mu_7 + \mu_{10})/7 - (\mu_2 + \mu_8 + \mu_9)/3$

or

$(0.14)\mu_1 - (.33)\mu_2 + (0.14)\mu_3 + (0.14)\mu_4 + (0.14)\mu_5 + (0.14)\mu_6 + (0.14)\mu_7 - (.33)\mu_8 - (.33)\mu_9 + (0.14)\mu_{10}$

The correct coefficients for calculating the estimate of the mean difference are: 0.14 -0.33 0.14 0.14 0.14 0.14 0.14 -0.33 -0.33 0.14

Contrasts vs. Estimates
Switchgrass Example

$$\bar{d} = (12.6 + 14.2 + 12.7 + 9.4 + 6.7 + 13.9 + 9.1)/7 - (8.3 + 9.6 + 9.4)/3$$

$$\bar{d} = 11.22 - 9.11 = 2.11$$

You can use a *t*-test to find the probability that this difference would occur:

$$t = \frac{Q}{s_Q} = \frac{Q}{s\sqrt{(\sum c_i^2 / r}} = \frac{2.11}{2.94\sqrt{.476/6}} = 2.55$$

Where s is the standard deviation (RMSE) from the ANOVA.

The *P* of a difference of 2.11 occurring by chance is 0.014
$\therefore$ conclude that the means are different.

* Note that $(2.55)^2 = 6.5$, the F-value from the contrast and that the SEQ equals the SED when comparing 2 means.

Estimates

Seed Treatment Example

```
estimate 'Control vs acid' treatment 1 -3 1 1 / divisor=3;
estimate 'Inorganic vs organic' treatment 1 0 -2 1 / divisor=2;
estimate 'Between organics' treatment 1 0 0 -1;
```

Parameter	Estimate	Standard Error	t Value	Pr > |t|
Control vs acid	-0.44466667	0.04788180	-9.29	<.0001
Inorganic vs organic	-0.18400000	0.05078632	-3.62	0.0023
Between organics	-0.08800000	0.05864299	-1.50	0.1529

Estimates calculate the mean difference for each contrast and test the significance of the difference using a t-test.

The coefficients used to calculate the estimate must sum to one for each mean in the comparison. For the first estimate statement above they are: .333 -1 .333 .333 or using a divisor: 1/3 -3/3 1/3 1/3.