Analysis of Variance
Part 6: Contrasts & Linear and Quadratic Trends

STAT 705: Regression and Analysis of Variance

Contrasts

- A linear combination of the t population means has the form: $c_1\mu_1 + c_2\mu_2 + . . . + c_t\mu_t$
- Examples involving three means:
  - Sum of the means: $\mu_1 + \mu_2 + \mu_3$
    - » $c_1 =1$, $c_2 = 1$, $c_3 = 1$
  - Average of the means: $(\mu_1 + \mu_2 + \mu_3)/3$
    - » $c_1 =1/3$, $c_2 = 1/3$, $c_3 = 1/3$
- A contrast is a linear combination whose coefficients (the c's) sum to 0

Examples of Contrasts

- Difference of means 1 and 2
  - $\mu_1 - \mu_2 = (1)\mu_1 + (-1)\mu_2 + (0)\mu_3$
  - $c_1 = 1, c_2 = -1, c_3 = 0 \Rightarrow c_1 + c_2 + c_3 = 0$
- Mean 1 minus the average of means 2 and 3
  - $\mu_1 - (\mu_2 + \mu_3)/2 = \mu_1 - 0.5\mu_2 - 0.5\,\mu_3$
  - $c_1 = 1, c_2 = -0.5, c_3 = -0.5 \Rightarrow c_1 + c_2 + c_3 = 0$
  - This might be appropriate if treatment 1 was a reference (control) treatment and treatments 2 and 3 were new treatments

Example

- An experiment was conducted to determine the effects of types of preservatives (A, B) in different amounts (100, 400) on prevention of bacteria growth. A control treatment was also included. The treatments are
  1. (A, 100)
  2. (A, 400)
  3. (B, 100)
  4. (B, 400)
  5. Control (no preservative)
- Let μ_1, μ_2, μ_3, μ_4, and μ_5 denote the means of the populations of responses for these treatments

Contrasts for the Example

In order, the treatments are

(A, 100), (A, 400), (B, 100), (B, 400), Control

Here are some contrasts that may be of interest and their coefficients (c_1, c_2, c_3, c_4, c_5)

- (A, 100) vs. Control ⇨ (1, 0, 0, 0, -1)
- Average of A's vs. Control ⇨ (.5, .5, 0, 0, -1)
- Average of 100's vs. average of 400's ⇨ (.5, .5, -.5, -.5, 0)
- Average of treatments vs. control ⇨ (.25, .25, .25, .25, -1)

Hypothesis Tests for Contrasts

- Hypotheses are H_0: population contrast = 0

  vs. H_a: population contrast $\neq 0$

- Test statistic: $t = \dfrac{\text{estimated contrast}}{\text{standard error of etimated contrast}}$

- Use the t distribution with df = dfE

- For the population contrast $c_1\mu_1 + c_2\mu_2 + . . . + c_t\mu_t$

  The estimated contrast is $c_1\overline{Y}_{1\cdot} + c_2\overline{Y}_{2\cdot} + ... + c_t\overline{Y}_{t\cdot}$

  The standard error is $SE = \sqrt{MSE}\sqrt{\dfrac{c_1^2}{n_1} + \dfrac{c_2^2}{n_2} + ... + \dfrac{c_t^2}{n_t}}$

Example

- Suppose there are 3 treatments with observations as shown. The ANOVA table is also shown.

Treatment	1	2	3
Data	4	1	2
	5	2	3
	5	3	4
	8		5
Means	5.5	2.0	3.5

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	21.64	310.82	5.41	0.0327
Error	8	16.00	2.00		
Corrected Total	10	37.64			

- We want to test whether the mean of treatment 1 is statistically different than the average of treatments 2 and 3

Example, continued

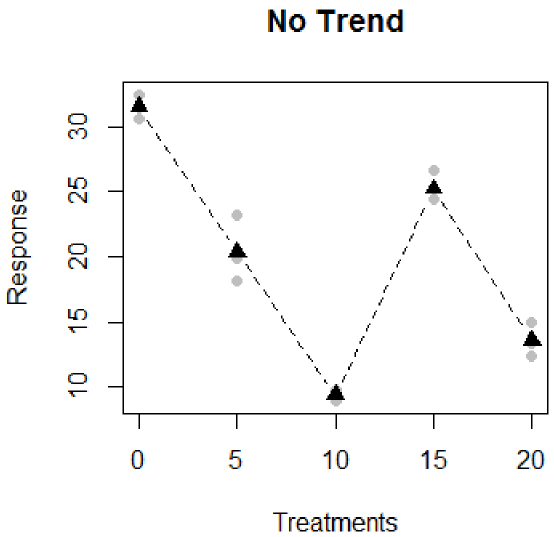
- Estimated contrast = (1) 5.5 – (.5) 2.0 – (.5) 3.5 = 2.75
- Standard error = $\sqrt{2.00}\sqrt{\frac{1^2}{4}+\frac{.5^2}{3}+\frac{.5^2}{4}} = 0.89$
- Test statistic = 2.75 / 0.89 = 3.09
- Critical value = 2.306 (from t table with df = 8)
- Decision: Reject H_0, since 3.09 > 2.306
- The contrast is significantly from 0, i.e., the mean of treatment 1 is significantly different from the average of the means of treatments 2 and 3.

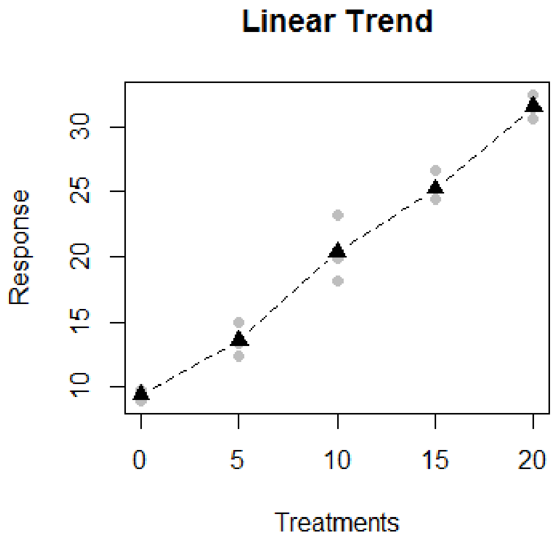
Linear and Quadratic Trends

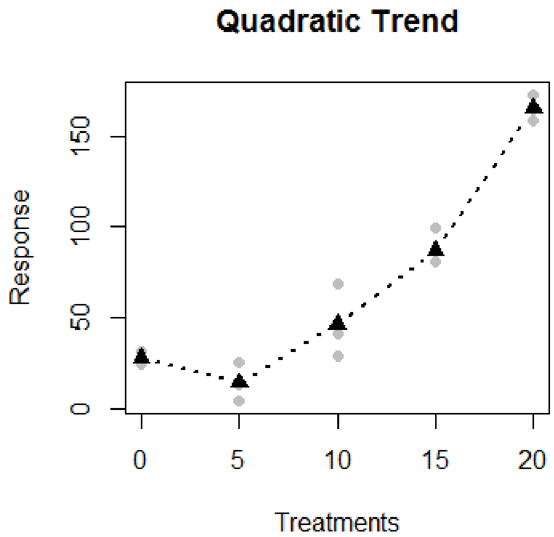
- For the next few slides, we take a closer look at some special contrasts
- The techniques that will be presented apply only when
  1) the treatments are quantitative (i.e., numeric), and
  2) the levels for the treatments are evenly spaced
- Examples
  - Wooden beams are tested at different amount of pressure (100 psi, 200 psi, 300 psi, 400 psi) to see how strong they are.
  - Different amounts of nitrogen (0%, 5%, 10%) are applied to plots of wheat to see how this affects yields.
  - Frozen food products are stored at different temperatures (-5°F, -10°F, -15°F) to see how long they last before spoiling.

Plot the Means

- Since the treatment levels are quantitative, we can plot them against their means
- We look for trends in the means
- Examples







Contrasts to Test for Trends

- We can use contrasts to test for trends in the treatment means (but only when the treatments are quantitative)
- If there are 3 treatments, we can test for linear and quadratic trends
- If we have a 4^{th} treatment, we can also test for cubic trends and so forth depending on how many quantitative treatments we have
- We will focus just on testing for linear and quadratic trends regardless of the number of treatments we have

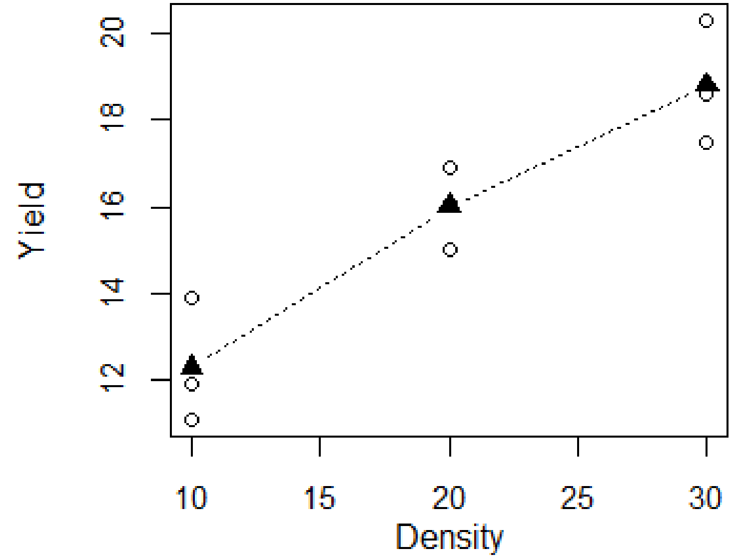
Contrast Coefficients for Trends

- When there are three treatments, contrast coefficients are
  - for linear trends: -1, 0, 1
  - for quadratic trends: 1, -2, 1
- When there are four treatments, contrast coefficients are
  - for linear trends: -3, -1, 1, 3
  - for quadratic trends: 1, -1, -1, 1
- Additional coefficients (for up to 8 treatments) are given in a separate file on the course website

Example

Beans were planted in 3 different densities: 10 plants per plot, 20 plants per plot, and 30 plants per plot. Yield was measured on each plot. The data and graph below show that the yield increases linearly with density. We will test to see if this trend is significant.

Density	**10**	**20**	**30**
Yield	11.9	16.1	20.3
	11.1	15.0	18.6
	13.9	16.9	17.5
Means	12.3	16.0	18.8



Example: Linear Contrast

- Use SAS to get the ANOVA table
- MSE = 1.66, dfE = 6

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	63.78	31.89	19.21	0.0025
Error	6	9.96	1.66		
Corrected Total	8	73.74			

- Treatment means are 12.3, 16.0, 18.8
- For 3 treatments, linear contrast coefficients are -1, 0, 1
- Estimated contrast = (-1)(12.3) + (0)(16.0) + (1)(18.8) = 6.5
- Standard error of estimate $= \sqrt{1.66}\sqrt{\frac{(-1)^2}{3} + \frac{0}{3} + \frac{1^2}{3}} = 1.052$
- Test statistic = 6.5 / 1.052 = 6.18
- Use t table (with df = dfE = 6 & α=0.05) to get critical value = 2.447
- 6.18 > 2.447, so we reject the null hypothesis that the contrast equals 0
- At α=0.05, we conclude that there <u>is</u> a significant linear trend

KANSAS STATE UNIVERSITY

Example: Quadratic Contrast

- We use the same ANOVA table, so MSE = 1.66 and dfE = 6
- The treatment means are still 12.3, 16.0, 18.8
- For 3 treatments, quadratic contrast coefficients are 1, -2, 1
- Estimated contrast = (1)(12.3) + (-2)(16.0) + (1)(18.8) = -0.9
- Standard error of estimate $= \sqrt{1.66}\sqrt{\frac{1^2}{3}+\frac{(-2)^2}{3}+\frac{1^2}{3}} = 1.822$
- Test statistic = -0.9 / 1.822 = -0.49
- The critical value does not change (it is 2.447)
- |-.49| is not greater than 2.447, so we do not reject the null hypothesis that the contrast equals 0
- At α=0.05, we conclude that there is <u>not</u> a significant quadratic trend

Other Polynomial Contrasts

- Contrasts can be used to compare treatment means, regardless of whether the treatments are quantitative or qualitative
- We use contrasts to test for trends ONLY when the treatments are numeric
- We have examined only linear and quadratic trends, with treatment levels that are evenly spaced
  - If there are more than 3 treatments, then higher-degree contrasts (cubic, 4^{th} degree, etc.) can be tested
  - If the treatment levels are not evenly spaced, then the contrast coefficients will be different
  - See the supplemental file on the course website

What You Should Know

- Contrasts in general
  - What they are
  - How to define them
  - How to test whether a contrast is equal to 0
- Polynomial contrasts
  - When are they used?
  - Know the coefficients for linear and quadratic trends when there are three evenly spaced treatments
- You are <u>not</u> responsible for knowing
  - Polynomial contrasts when the treatments are not evenly spaced
  - Cubic, 4th degree, or higher polynomials