STAT 500

ANOVA - Contrasts

ANOVA: Contrasts

- Donut example
  - We rejected the null hypothesis of an equal mean amount of oil absorbed for the four cooking oils.
  - Conclusion: At least some of the means for the four cooking oils are different.
  - Question: Which ones and by how much?

ANOVA: Comparisons and Contrasts

- Additional Analyses
  - Inference for a single population mean
  - Linear combinations of means - contrasts as special case
  - Pairwise comparisons

ANOVA: Single Population Mean

- $100(1-\alpha)\%$ confidence interval for a single group mean

$$\bar{Y}_{i.} \pm t_{N-r,1-\alpha/2}\sqrt{\frac{MS_{error}}{n_i}}$$

- Note: MS_{error} is the estimate of the population variance σ^2.

- Note: df for t distribution is $N-r$.

- Valid for a single population mean - not used for comparison between means.

ANOVA: Contrasts

Contrast:
Linear combination of the population means

$$\gamma = \sum_i c_i \mu_i$$

where

$$\sum_{i=1}^{r} c_i = 0$$

Contrast Examples

- Ex. Difference between group 1 mean and mean of groups 2 and 3.

$$\gamma = \mu_1 - \frac{\mu_2 + \mu_3}{2}$$

($c_1 = 1$, $c_2 = -0.5$, $c_3 = -0.5$, and $c_4 = 0$)

- Ex. Difference between two group means

$$\gamma = \mu_i - \mu_k$$

($c_i = 1, c_k = -1,$ all other c's = 0)

Estimation: Contrasts

- Point estimate: $\hat{\gamma} = \sum_i c_i \bar{Y}_{i.}$
- Standard error assuming σ^2 is known:

$$\sigma_{\hat{\gamma}} = \sqrt{\sigma^2 \sum_i (c_i^2/n_i)}$$

- Standard error when σ^2 is NOT known:

$$S_{\hat{\gamma}} = \sqrt{MS_{error} \sum_i (c_i^2/n_i)}$$

- $100(1-\alpha)\%$ confidence intervals:

$$\hat{\gamma} \pm t_{N-r,1-\alpha/2} S_{\hat{\gamma}}$$

Hypothesis Test: Contrasts

- Test $H_o : \gamma = \sum_i c_i \mu_i = 0$
  - Using t distribution

$$t = \frac{\hat{\gamma} - 0}{S_{\hat{\gamma}}} \text{ has } N - r \text{ d.f.}$$

  - Using F distribution

$$F = \frac{SS_\gamma}{MS_{error}} \text{ has } (1, N - r) \text{ d.f.}$$

where $SS_\gamma = \frac{\hat{\gamma}^2}{(\sum_i c_i^2 / n_i)}$

Orthogonal Contrasts

- Two contrasts $\gamma_1 = \sum_i c_i\mu_i$ and $\gamma_2 = \sum_i b_i\mu_i$ are *orthogonal* if

$$\sum_i b_i c_i / n_i = 0$$

- Example: $\gamma_1 = \mu_1 - \frac{\mu_2+\mu_3}{2}$ and $\gamma_2 = \mu_2 - \mu_3$
  - $c_1 = 1, c_2 = -0.5, c_3 = -0.5$ and $b_1 = 0, b_2 = 1, b_3 = -1$
  - $\sum_i b_i c_i / n_i = 0(1)/n_1 + 1(-0.5)/n_2 + -1(-0.5)/n_3 = 0$

Orthogonal Contrasts

If γ_1 and γ_2 are orthogonal contrasts, then

- they represent statistically unrelated pieces of information. In another word, one contrast conveys no information about the other.
  - Estimates $\hat{\gamma}_1$ and $\hat{\gamma}_2$ are *uncorrelated*.
  - Hypothesis tests for γ_1 and γ_2 are independent - results of one test do not affect results of other.
  - Confidence intervals for γ_1 and γ_2 are independent - results of one do not affect results of other.

Orthogonal Contrasts

- A set of contrasts are orthogonal if all pairs are orthogonal.
- For r means, there are at most $r-1$ mutually orthogonal contrasts in a set.
- For r means, there are many possible sets of $r-1$ mutually orthogonal contrasts.
- SS_{model} can be decomposed by $r-1$ mutually orthogonal contrasts.
- Example: Let $r = 4$ as in the donut example, and let $\gamma_1, \gamma_2, \gamma_3$ be mutually orthogonal contrasts. Then

$$SS_{\text{model}} = SS_{\gamma_1} + SS_{\gamma_2} + SS_{\gamma_3}$$

Orthogonal Polynomial Contrasts

Analyze trends for quantitative factors or ordered treatments

Assume equal spacing of levels and equal sample sizes

For a factor with three equally spaced levels:

Trend	μ_1	μ_2	μ_3
Linear	-1	0	1
Quadratic	-1	2	-1

$$SS_{\text{model}} = SS_{\text{linear}} + SS_{\text{quad}}$$

Orthogonal Polynomial Contrasts (cont.)

For a factor with five equally spaced levels:

Trend	μ_1	μ_2	μ_3	μ_4	μ_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

$$SS_{\text{model}} = SS_{\text{linear}} + SS_{\text{quad}} + SS_{\text{cubic}} + SS_{\text{quartic}}$$

Orthogonal Contrasts

Why are orthogonal contrasts useful?

- F test from the ANOVA table
  - Tests whether all groups have the same mean
  - We don't always care about the F-test.
    Contrasts focus attention on specific questions.
  - Researcher must specify the questions
- Independence of test results means we can interpret tests for contrasts individually.
- Motivate partitioning of SS into "interesting" and "everything else" parts

Orthogonal Contrasts

Why are orthogonal contrasts useful?

- Researchers specify one question "Did type1 oil has a difference mean from the other three types?", answered by contrast with c=(-1,1/3, 1/3, 1/3).
- Does this contrast explain all differences among means? $SS_{model} = SS_1 + SS_2 + SS_3 = SS_1+$ rest

Source	SS	d.f.
Model	SS_{model}	3
Type1vsOthers	SS_c	1
rest	$SS_{model} - SS_c$	3-1
Error	SS_{error}	df_{error}