Analysis of Variance (ANOVA)

Lecture 14

Objectives

- Define terms.
- Explain why many pair-wise comparisons are inappropriate.
- Recognize when an ANOVA or Kruskal-Wallis is appropriate.
- Understand the rationale behind an ANOVA.
- Perform the calculations necessary for an ANOVA and Kruskal-Wallis.
- Understand the necessity of multiple comparison tests.
- Be able to conduct parametric and nonparametric multiple comparison tests.
- Interpret the results of the tests.

Overview

- If we want to compare the means of two groups, we use a t-test (assuming we can meet the assumptions).
- What if we have to make comparisons among more than 2 groups? -- You can't use a t-test.

Overview

- It might seem that you could compare sample means two-by-two until all comparisons are made.
- With three groups (A, B, and C) you might think you could compare:
  - A with B
  - A with C
  - B with C
- You can NOT do this

Overview

- Each additional t-test increases the probability of committing a Type I error.
  - Multiple comparison inflation of Type I error.

Example

- If we assume α=0.05 (there is a 5% chance of committing a type I error if two means are compared then:

Example

Groups	Comparisons	Prob. 1 Type I
2	1	0.05
3	3	0.14
4	6	0.26
10	45	0.90

Example

- Multiple comparison inflation of Type I error.
- Bonferroni inequality
  - Go to previous slide and demonstrate

Overview

- We could decrease α (say to 0.01) to decrease our chances of committing a Type 1 error, in doing so that would increase β and our corresponding chance that we would commit a type II error.

Overview

- We need to use an Analysis of Variance (ANOVA) or Kruskal-Wallis.
  - With an ANOVA or Kruskal-Wallis, we can make comparisons between more than 2 groups.

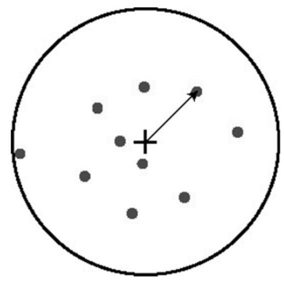
ANOVA

- The many different varieties or flavors of ANOVAs.
  - To compare means from more than 2 groups you use a One-way ANOVA (also called a single classification or single factor ANOVA).

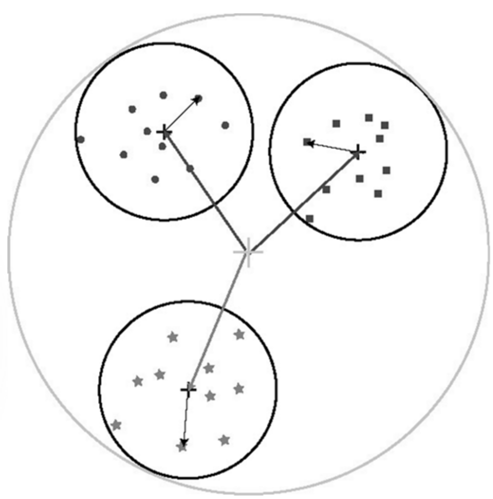
Rationale

- Samples give rise to two sources of variability
  - The variability around each mean within a sample
  - The variability among the samples due to difference among the means of the populations from which the samples are drawn.
  - $\text{Variability}_{total} = \text{Variability}_{within} + \text{Variability}_{among}$

Rationale



Rationale



Rationale

Assumptions

- Random sampling (very important)
- Independent samples
- Equal variances
- Normally distributed populations

- Best if sample sizes are nearly equal (maximizes the power of the test – not an actual assumption).
- Fairly robust

Hypotheses

- H0: Samples are drawn from populations with equal means. (μ_1=μ_2=μ_3= … =μ_i)
- H1: Samples are drawn from populations with different means.
  - (Normally no statistical hypothesis because more than 1 deviation is possible).

Example

- Example 18

Multiple Comparison Tests

- An ANOVA or Kruskal-Wallis test will tell you if the samples differ significantly, but they will not tell you which of the samples differs significantly from the others.

Multiple Comparison Tests

- To determine which pairs of samples are responsible for generating the differences detected by the one-way ANOVA or Kruskal-Wallis you need to use a comparison among means (or magnitudes) test.

Multiple Comparison Tests

- Test for comparisons among means are typically only used when a significant difference has been detected using an ANOVA or a Kruskal-Wallis.

Multiple Comparison Tests

- In these tests, the significance level (α) represents the probability of committing at least one Type I error while comparing all of the means – not the probability of committing a Type I error for any single comparison.

Multiple Comparison Tests

- As a consequence the significance level for any one test is less than the overall error rate – resulting in comparisons that conservative (not very sensitive to real differences in means – low power).

Multiple Comparison Tests

- There are large variety of multiple comparison tests. Major tests include:
  - Duncan's – Uncommonly encountered
  - Newman-Keuls' – Fairly commonly encountered
  - Tukey's – Fairly commonly encountered
  - Scheffe's – Commonly encountered

Assumptions

- Random samples
- Independent samples
- Equal variances (parametric version)
- Normal distributions (parametric version)

- Equal sample sizes are desirable for maximum power and robustness.

Assumptions

- How the tests respond to deviations from these assumptions is generally not known, but the Tukey's test is known to be remarkably robust.
  - Tukey's LSD test
  - Tukey's HSD test

Kruskal-Wallis test (Nonparametric ANOVA)

- Advantages
  - Data types that can be utilized (ratio, interval, ordinal, and “nominal”)
  - No assumption of normality.
  - No assumption of equal variances.

Assumptions

- Random samples.
- Independent samples.
- 3 or more samples.
  - 2 is an option for a One-Way ANOVA it is not for a Kruskal-Wallis test. (use a Mann-Whitney U for 2 samples).
- There needs to be a minimum of 5 observations per sample.

Example

- Example 19

Ambiguous Results

- Occasionally multiple comparison tests will yield somewhat ambiguous results such as:

$$X_1 = X_2$$

$$X_2 = X_3 = X_4$$

$$X_1 \neq X_3 = X_4$$

Ambiguous Results

- The four samples appear to be drawn from two different populations: sample 1 and sample 2 were drawn from one population and sample 2, sample 3, and sample 4 were drawn from another population.
  - It is impossible for sample 2 to have come both populations.

Ambiguous Results

- Therefore all we can state is:

$$\mu_1 \neq \mu_3 = \mu_4$$

- Repeating the analysis with larger sample sizes may yield a more decisive results as the test would have more power