

Analysis of Variance (ANOVA)

Need to know

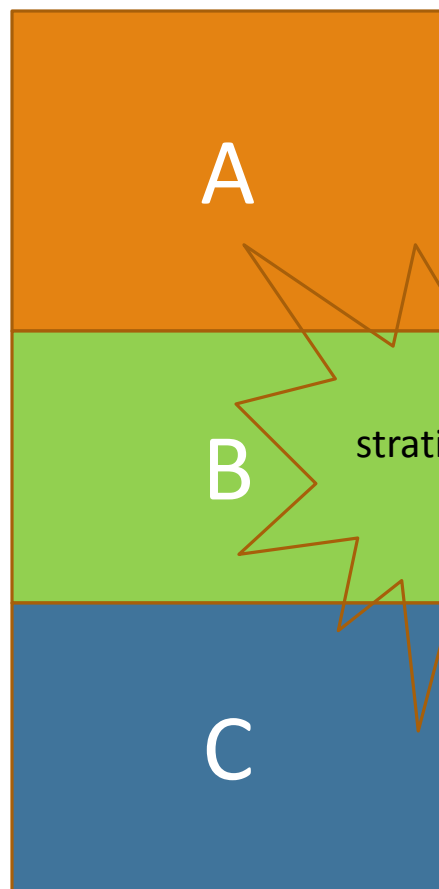
Concept

Nature of data

Computing

Interpreting

Concept



4, 7, 10, 3

Mean & Variance
of (A)

6, 5, 9, 8, 4

Mean & Variance
of (B)

5, 7, 7, 9, 3

Mean & Variance
of (C)

Mean &
Variance of
(A+B+C)

Mean & Variance of
Mean A, Mean B, Mean C

When I talk about between groups variability, what am I talking about?

Exercise

How many kinds of MEAN do you notice? (2 or 3 or 4 or 5)

Where does INDIVIDUAL variation lie? (Pink or Yellow or Gray)

Where does GROUP-WISE variation lie? (Pink or Yellow or Gray)

Can we test GROUP-WISE variation by two-sample t test? (Yes or No)

Nature of data

Parameters and Measurement scale

Factor (Background characteristics/Experience/Stratification)

Outcome (change of differentiation by the factor focused in analysis)

Measurement scales of Factor and Outcome

The basic ANOVA situation

Two variables: 1 Categorical, 1 Quantitative

Main Question: Do the (means of) the quantitative variables depend on which group (given by categorical variable) the individual is in?

If categorical variable has only 2 values:

- 2-sample t-test

ANOVA allows for 3 or more groups

Why ANOVA?

If we have to compare a continuous measure in more than two groups

- Test sample dataset SPSS Cars.sav

Two-sample t-tests are problematic

- Increasing the risk of a Type I error
- At .05 level of significance, with 100 comparisons, 5 will show a difference when none exists (experiment-wise error)
- So the more t-tests you run, the greater the risk of a type I error (rejecting the null when there is no difference)

ANOVA allows us to see if there are differences between means with an **OMNIBUS** test

Variance – why do scores vary?

A representation of the spread of scores

What contributes to differences in scores?

- Individual differences
- Which group you are in

What does Sum Square (SS) between represent?

What does Mean Square (MS) (either within or between) represent?

Variance to compare Means

We are applying the variance concept to means

- How do means of different groups compare to the overall mean

Do the means vary so greatly from each other that they exceed individual differences within the groups?

Between/Within Groups

Variance can be separated into two major components

- **Within groups** – variability or differences in particular groups (individual differences)
- **Between groups** - differences depending what group one is in or what treatment is received

Bottom Line

INDIVIDUAL VARIANCE

GROUP VARIANCE

Computing variances

Fundamental Concepts

You are able to compare MULTIPLE means

Between-group variance reflects differences in the way the groups were treated

Within-group variance reflects individual differences

Null hypothesis: no difference in means

Alternative hypothesis: difference in means

Sum of Squares

We are comparing “variance estimates”

- $\text{Variance} = \text{SS}/\text{df}$

The charge is to partition the variance into between and within group variance

Critical factors:

- BETWEEN GROUP VARIANCE
- WITHIN GROUP VARIANCE

How does the between group variance compare with the within group variance?

Designed Experiments of Interest

One-factor completely randomized designs

$$\text{Total SS} = \text{Treatment SS} + \text{Error SS}$$

$$\text{SS(Total)} = \text{SST} + \text{SSE}$$

Randomized Block Designs

$$\text{Total SS} = \text{Treatment SS} + \text{Block SS} + \text{Error SS}$$

$$\text{SS(Total)} = \text{SST} + \text{SSB} + \text{SSE}$$

Word check

When I talk about between groups variability, what am I talking about?

What does SS between represent?

What does MS (either within or between) represent?

What does the F ratio represent?