UNIVERSITY OF TWENTE.

Analysis of (Co-)Variance
Technical Lecture

Overview

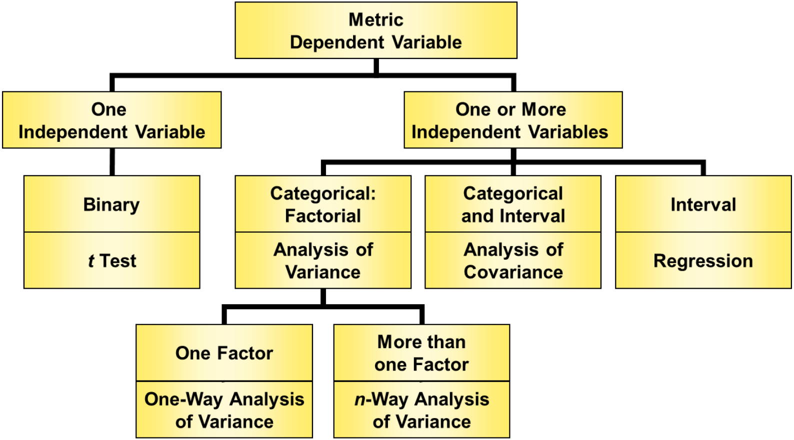
Introduction: Analysis of Variance (ANOVA)

- ANOVA is used as a test of means for two or more populations. Typically, the null hypothesis is that all population means are equal.
- ANOVA must have a metric dependent variable, i.e., interval or ratio scaled.
- There must also be one or more independent variables (factors) that are all categorical (non-metric). Each possible outcome of a factor represents a level of that factor or, alternatively, a category.

Introduction

- One-way ANOVA involves only one categorical variable, or a single factor.
- If two or more factors are involved, the analysis is termed n-way ANOVA.
- If the set of independent variables consists of both categorical and metric variables, the technique is called analysis of covariance (ANCOVA). In this case, the categorical independent variables are still referred to as factors, whereas the metric independent variables are referred to as covariates.

Relationship among t-Test, ANOVA, ANCOVA, and Regression



One-Way ANOVA

Potential Applications: One-Way ANOVA

Researchers are often interested in examining the differences in the mean values of the dependent variable for several categories of a single independent variable or factor.

For example:

- Does the number of defective products differ among different machines?
- Does the number of work-related accidents differ across different working environments?
- Do different checklists for trouble shooting affect the reparation time?

Example: One-Way ANOVA

Has the study programme an effect on the beer consumption?

Observation	Beer consumption	Study
Student1	0	Theology
Student2	14	Theology
Student3	16	Theology
Student4	20	Business
Student5	33	Business
Student6	37	Business
Student7	29	Engineering
Student8	31	Engineering
Student9	45	Engineering

Average consumption: 25

Statistics Associated with One-Way ANOVA

SS_Y:

- Total variation in Y.

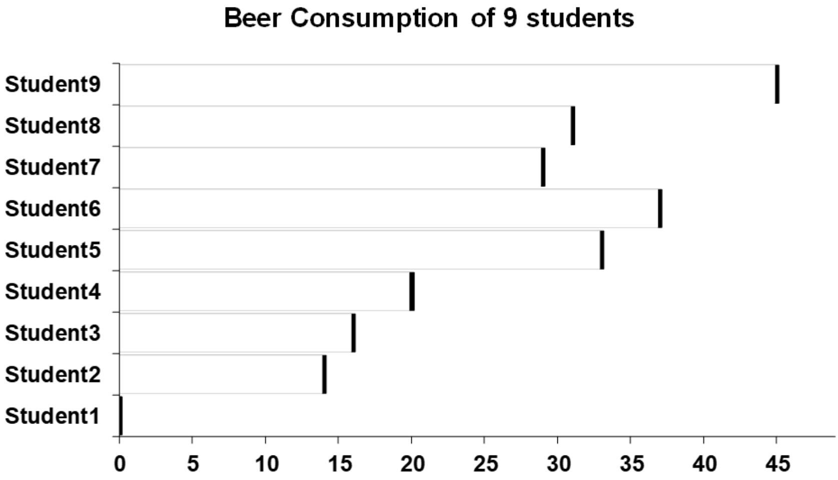
SS_X:

- The variation in Y related to the variation in the means of the categories of X.
- This represents variation between the categories of X, or the portion of the sum of squares in Y related to X.

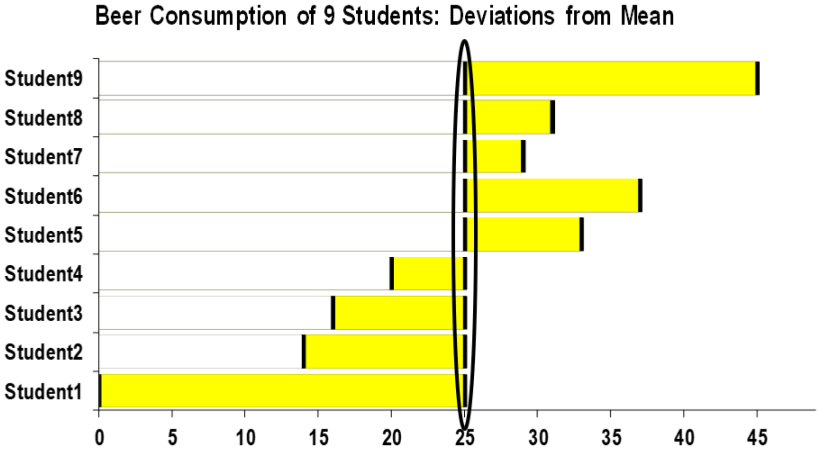
SS_{error}/SS_{within}:

- Variation in Y due to the variation within each of the categories of X.
- This variation is not accounted for by X.

Example: One-Way ANOVA



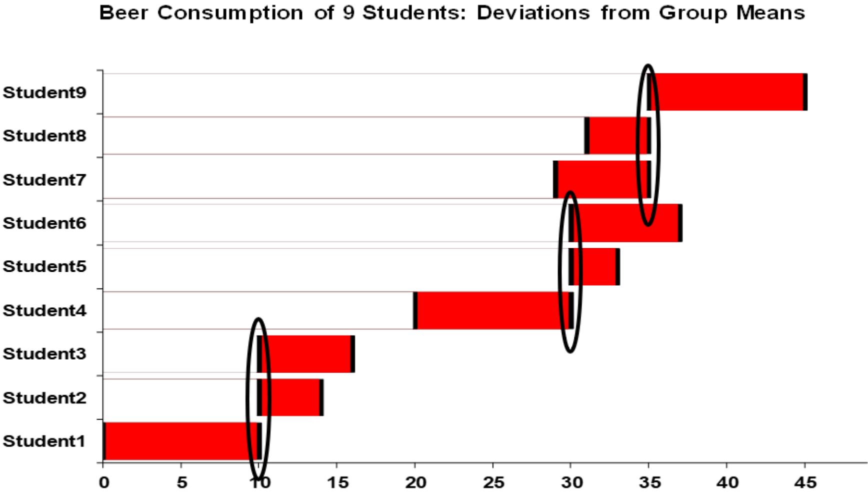
Example: One-Way ANOVA



Example: Calculating SS_Y

Observation	Consumption	Deviation from mean	Squared deviation
Student1	0	-25	625
Student2	14	-11	121
Student3	16	-9	81
Student4	20	-5	25
Student5	33	8	64
Student6	37	12	144
Student7	29	4	16
Student8	31	6	36
Student9	45	20	400
			$SS_Y = 1512$

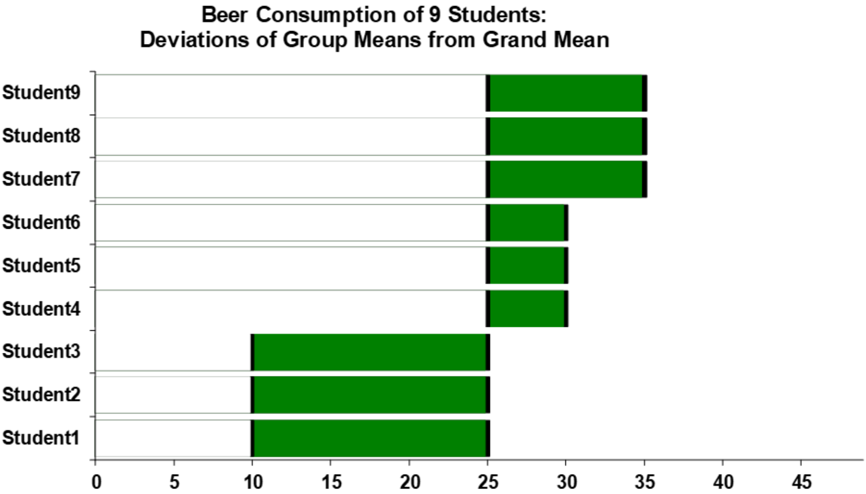
Example: One-Way ANOVA



Calculating SS_{error}

Observation	Consumption	Deviation from group mean	Squared deviation
Student1	0	-10	100
Student2	14	4	16
Student3	16	6	36
Student4	20	-10	100
Student5	33	3	9
Student6	37	7	49
Student7	29	-6	36
Student8	31	-4	16
Student9	45	10	100
			$SS_{error} = 462$

Example: One-Way ANOVA



Calculating SS_X

Observation	Group mean	Deviation from mean	Squared deviation
Student1	10	-15	225
Student2	10	-15	225
Student3	10	-15	225
Student4	30	5	25
Student5	30	5	25
Student6	30	5	25
Student7	35	10	100
Student8	35	10	100
Student9	35	10	100
			$SS_X = 1050$

Test Significance

We can use an F-test to statistically assess whether the mean of the dependent variable differs across groups:
$\mathrm{H}_0 : \mu_1 = \mu_2 = ... = \mu_c$.

For our example: $\mathrm{H}_0 : \overline{\mathrm{Beer}}_{\mathrm{Theo}} = \overline{\mathrm{Beer}}_{\mathrm{Busi}} = \overline{\mathrm{Beer}}_{\mathrm{Engi}}$

F-statistic:

$$F = \frac{SS_X/(c-1)}{SS_{error}/(N-c)} = \frac{1050/2}{462/6} \approx 6.81 > \underbrace{5.14}_{F(2{,}6)_{0.95}} \Rightarrow \text{significant}$$

Factors that Affect Significance

Is the difference in sample means large enough to conclude that the population means do in fact differ among two or more populations?

The answer depends on:

- The size of the difference between group means (the variability of group means).
- The sample sizes in each group. Larger sample sizes give more reliable information and even small differences in means may be significant if the sample sizes are large enough.
- The variances of the dependent variable in each group. For the same absolute difference in means, the difference is more significant if in each group the dependent variable values tightly cluster about their respective means.

Assumptions in ANOVA

The salient assumptions in analysis of variance can be summarized as follows:

- The dependent variable is metric.
- Ordinarily, the categories of the independent variable are assumed to be fixed, i.e., fixed-effects model.
- The error term is normally distributed, with a zero mean and a constant variance $\Rightarrow$ the dependent variable is normally distributed with a constant variance across groups.
- The error is independent of any of the categories of X.
- Observations are mutually independent, i.e., no repeated measurement. Otherwise, the F-test can be seriously distorted.

Dependent Variable is Normally Distributed Across Groups

This assumption is needed for statistical inferences.

Consider P-P, Q-Q plots, and statistical tests for normality (same as in regression analysis).

For large sample sizes $(N > 200)$ a violation of the assumption can be neglected (Central Limit Theorem).

Equal Variance Across Groups

This assumption is important, because it strongly affects the F-test.

It is assessed by the Levene's test:

- H_0: The variance of the dependent variable is the same across groups
- H_1: The variance of the dependent variable differs across groups

In case of a significant Levene's test:

- When groups have equal size, this is not harmful.
- Otherwise use Welch test instead of F-test.

Quantify the Effects

eta^2:

$$\eta^2 = \frac{SS_X}{SS_Y} = \frac{1050}{1512} \approx 0.694$$

- Measures the strength of the effect of X (independent variable or factor) on Y (dependent variable).
- The value of η^2 varies between 0 and 1.
- How much of the variance in the dependent variable is explained by the independent variable (like R^2 in regression analysis).

Quantify the Effects

Partial eta^2:

$$\text{Partial } \eta^2 = \frac{SS_X}{SS_X + SS_{error}} = \frac{1050}{1050 + 462} \approx 0.694$$

- In case of an one-way ANOVA, the partial η^2 and the η^2 are equal but they differ if more than one independent variable is included.
- The value of partial η^2 varies between 0 and 1.
- How much of the variance can be explained when controlling for the other factors, i.e., the percent of partial variance (variance of the dependent variable after accounting for the other variables in the model) that can be explained.

Quantify the Effects

The most commonly used effect size measure in ANOVA is omega squared, ω^2. The relative contribution of a factor X is calculated as follows:

$$\omega_X^2 = \frac{SS_X - (df_X \cdot MS_{error})}{SS_{total} + MS_{error}} = \frac{1050 - (2 * 462/(9-3))}{1512 + 462/(9-3)} \approx 0.56$$

It is a less biased estimate for the population variance explained.

In general, effect size measures are typically only considered for statistically significant effects.

Multiple Comparisons

If the null hypothesis of equal means is rejected, we can only conclude that not all of the group means are equal. We may wish to examine differences among specific means.

This can be done by

- specifying appropriate **a priori contrasts**, or
- performing **post-hoc tests**

to determine which of the means are statistically different.

Multiple Comparisons: A priori contrasts

- A priori contrasts are determined before conducting the analysis, based on the researcher's theoretical framework. The contrasts selected are orthogonal (they are independent in a statistical sense).
- A priori contrasts do not control for the family-wise error rate (Think of using corrections such as Bonferroni).

Multiple Comparisons: Post-Hoc Tests

To investigate which group means differ significantly, post-hoc test can be applied.

Post-hoc tests usually control for the family-wise error rate.

For our example:
$\overline{\text{Beer}}_{\text{Theo}} \leftrightarrow \overline{\text{Beer}}_{\text{Busi}}$
$\overline{\text{Beer}}_{\text{Theo}} \leftrightarrow \overline{\text{Beer}}_{\text{Engi}}$
$\overline{\text{Beer}}_{\text{Engi}} \leftrightarrow \overline{\text{Beer}}_{\text{Busi}}$

Cannot be applied for ANCOVA.

Various post-hoc tests...

- least significant difference (most powerful),
- Duncan's multiple range test,
- Student-Newman-Keuls,
- Tukey's alternate procedure,
- honestly significant difference,
- modified least significant difference, and
- Scheffé's test (most conservative).

N-Way ANOVA

N-Way ANOVA

One is often concerned with the effect of more than one factor simultaneously.
For example:

- How does the production error rate differ across different machines and different manuals?
- How does the weight loss differ across different diets and different exercise programs?

N-Way ANOVA: Decomposing the Total Variation

Consider the simple case of two factors X_1 and X_2 having categories c_1 and c_2. The total variation in this case is partitioned as follows:

$$\begin{aligned} SS_Y = & SS \text{ due to } X_1 + SS \text{ due to } X_2 \\ & + SS \text{ due to interaction of } X_1 \text{ and } X_2 + SS_{error} \end{aligned}$$

N-Way ANOVA: Testing Significance

The significance of the overall effect may be tested by an $F-$test, as follows:

$$F = \frac{(SS_{X_1} + SS_{X_2} + SS_{X_1 X_2})/df_n}{SS_{error}/df_d} = \frac{MS_{X_1, X_2, X_1 \cdot X_2}}{MS_{error}},$$

where

- df_n: degrees of freedom for the numerator: $(c_1 - 1) + (c_2 - 1) + (c_1 - 1)(c_2 - 1) = c_1 c_2 - 1$
- df_d: degrees of freedom for denominator: $N - c_1 c_2$
- MS: mean squares

N-Way ANOVA: Testing Significance

If the overall effect is significant, the next step is to examine the significance of the interaction effect. Under the null hypothesis of no interaction, the appropriate F-test is:

$$F = \frac{SS_{X_1 X_2}/df_n}{SS_{error}/df_d} = \frac{MS_{X_1 \cdot X_2}}{MS_{error}},$$

where

- df_n: degrees of freedom for the numerator: $(c_1 - 1)(c_2 - 1)$
- df_d: degrees of freedom for denominator: $N - c_1 c_2$
- MS: mean squares

N-Way ANOVA: Testing Significance

In the next step, the significance of the main effect of each factor is tested.

Main effect, is the effect of one independent variable averaged across the levels of any other independent variable.

For X_1 it is tested as follows:

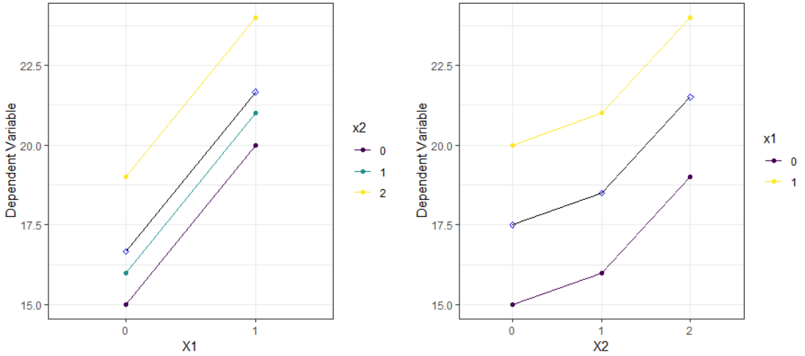
$$F = \frac{SS_{X_1}/df_n}{SS_{error}/df_d} = \frac{MS_{X_1}}{MS_{error}},$$

where

- $df_n = c_1 - 1$
- $df_d = N - c_1 c_2$

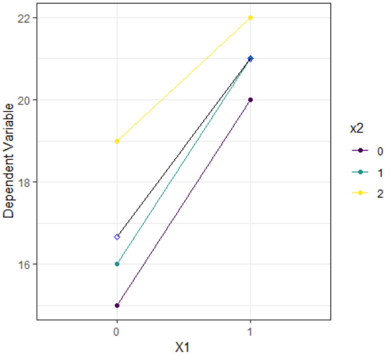
Issues in Interpretation: Patterns of Interactions

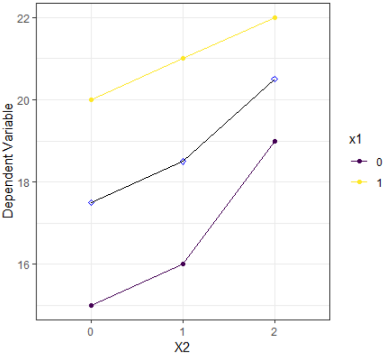
No Interaction:



Issues in Interpretation: Patterns of Interactions

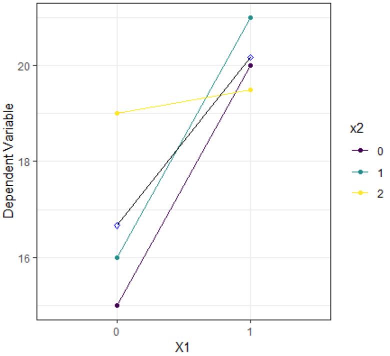
Ordinal Interaction:

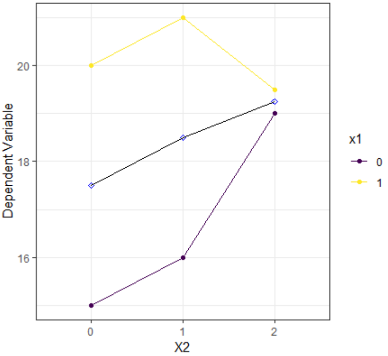




Issues in Interpretation: Patterns of Interactions

Disordinal Interaction:





ANCOVA

Analysis of Covariances

When examining the differences in the mean values of the dependent variable related to the effect of the controlled independent variables (in an experiment), it is often necessary to take into account the influence of independent variables which are not controlled for in the experiment (and to use 'statistical control' instead!).

For example:
In determining how different price levels will affect a household's cereal consumption, it may be necessary to control for the household's size.

Analysis of Covariances

ANCOVA is mainly used for two purposes:

- In quasi-experimental (observational) designs, to remove the effects of variables which modify the relationship of the categorical independents to the interval dependent.
- In experimental designs, to control for factors which cannot be randomized but which can be measured on an interval scale.

Goal:

- Reducing the error term in the model.
- Procedures of statistical control (e.g. "what if" analysis).

Repeated Measures ANOVA

Repeated Measures ANOVA

- One way of controlling the differences between subjects is by observing each subject under each experimental condition.
- Since repeated measurements are obtained from each respondent, this design is referred to as within-subjects design or repeated measures ANOVA.
- Repeated measures ANOVA may be thought of as an extension of the paired-samples t-test to the case of more than two related samples.

Potential Application

- Four commercials are rated by the same group of people. Were the commercials rated differently?
- Do colors affect the stress level?

Repeated Measures ANOVA

In the case of a single factor with repeated measures, the total variation, with $nc - 1$ degrees of freedom, may be split into between-people variation and within-people variation.

$$SS_{total} = SS_{between\ people} + SS_{within\ people}$$

The between-people variation, which is related to the differences between the means of people, has $n - 1$ degrees of freedom. The within-people variation has $n(c - 1)$ degrees of freedom.

Repeated Measures ANOVA

The within-people variation may, in turn, be divided into two different sources of variation. One source is related to the differences between treatment means, and the second consists of residual or error variation:

$$SS_{within\ people} = SS_X + SS_{error}$$

The degrees of freedom corresponding to the treatment variation (SS_X) are c - 1, and those corresponding to residual variation are (c - 1) (n -1).

A test of the null hypothesis of equal means may now be constructed in the usual way:

$$F = \frac{SS_X/(c-1)}{SS_{error}/((n-1)(c-1))} = \frac{MS_X}{MS_{error}}$$

Appendix

Family-wise error rate

All Pairwise Comparisons Alpha = 0.05		
Groups	Comparisons	Experimentwise Error Rate
2	1	0.05
3	3	0.142625
4	6	0.264908109
5	10	0.401263061
6	15	0.53670877
7	21	0.659438374
8	28	0.762173115
9	36	0.842220785
10	45	0.900559743
11	55	0.940461445
12	66	0.966134464
13	78	0.981700416
14	91	0.990606054
15	105	0.995418807
		From StatisticsByJim.com

Patterns of Interactions (ANCOVA)