Analysis of Research Design (2)

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Greta Dohrwardt & Alisa Hirn

EFMD EQUIS ACCREDITED
AACSB ACCREDITED
ASSOCIATION OF MBAS ACCREDITED

29.05.2024

Agenda

The idea of Inferential Statistics

- Significance testing
- Confidence Intervall
- Effect Size

General Linear Model

- Two-Variable Linear Model
- General Case
- Dummy Variables

Experimental Analysis

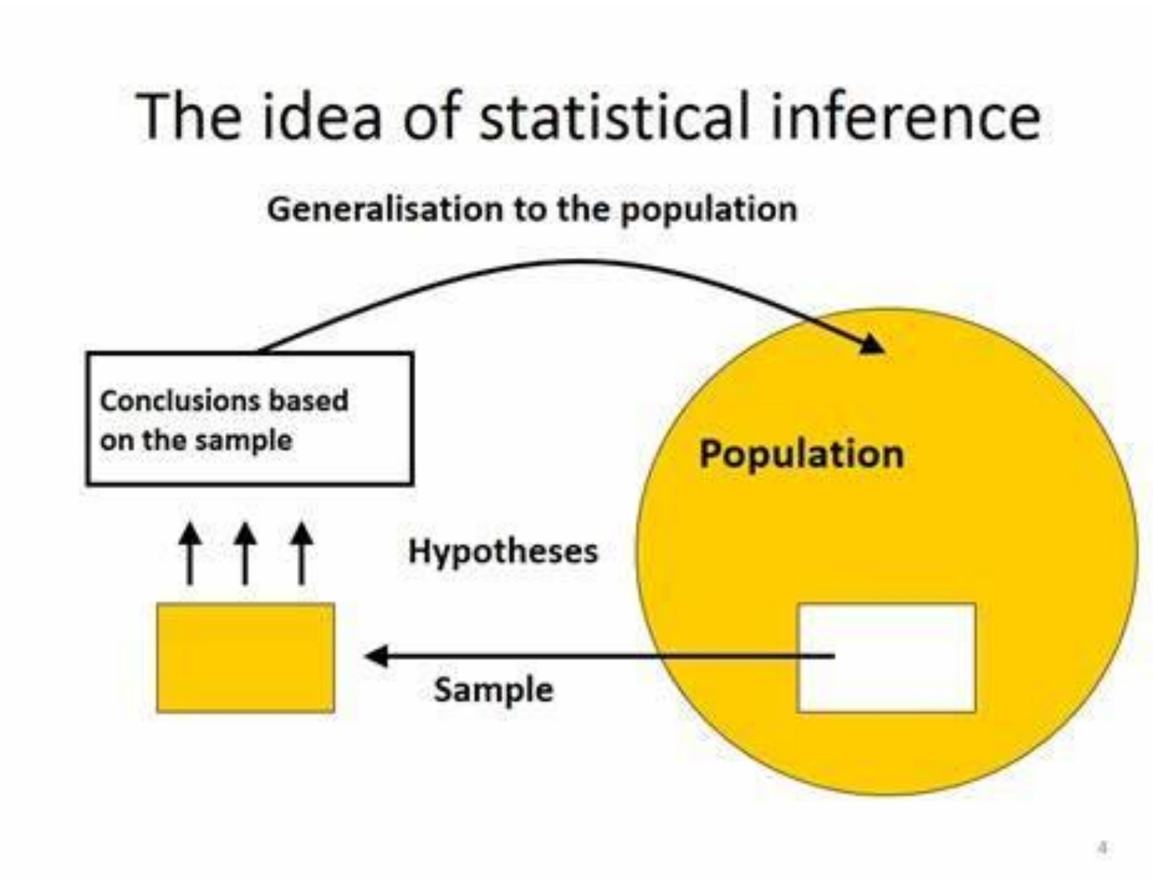
- The t-Test
- Factorial Design Analysis
- Randomized Block Analysis
- Analysis of Covariance

Inferential Statistics

- **Descriptive statistics:** simply what is going on in data
- **Inferential statistics:** make inferences form our data to general conditions

→ **General Linear Model** (t-test, ANOVA, ANCOVA regression analysis) and **multivariate analysis** (factor analysis, multidimensional analysis, cluster analysis,…)

→ **Dummy variable** (proxy variable) for modelling two separate lines (each treatment group) with single equation



Inferential Statistics

- How do we decide whether to make inferences?
  Not every result as important as any other result

→ **Significance testing**, **confidence intervals** and **effect size** interpret precision and magnitude of results

→ significance level is the level of risk we are willing to accept as the price of our inference from sample population

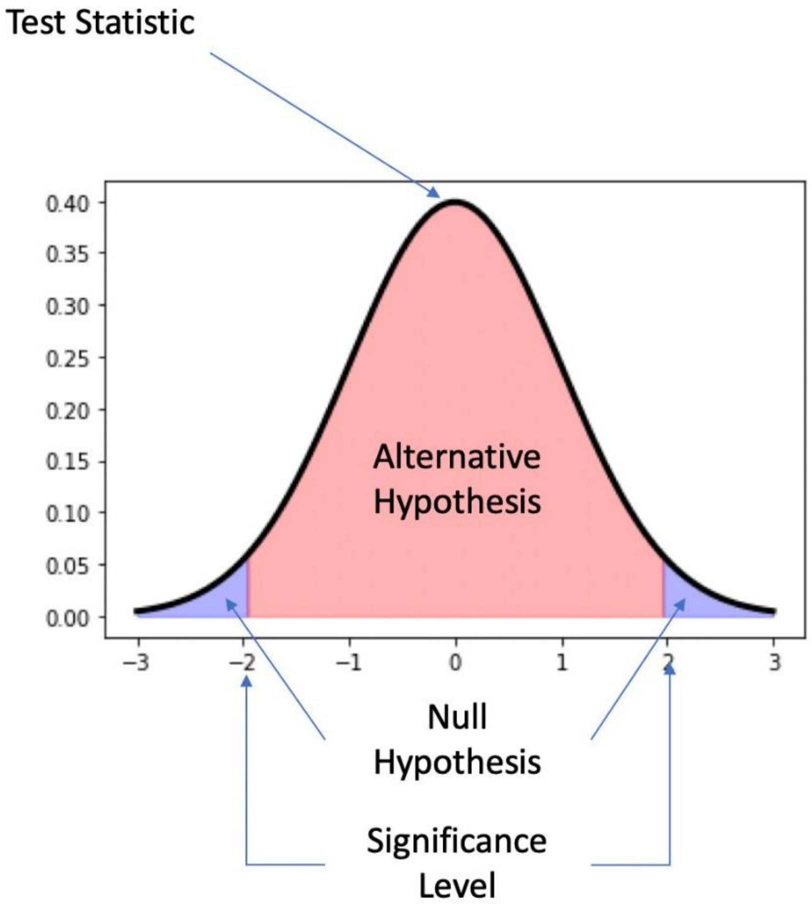
P-value = how likely you are to have found a particular set of observations if the null hypothesis were true

Significance testing

H_A : People are happier in summer (There is an effect)
H_0 : People are equally happy in summer and winter (There is no effect)

- **if *p-value* is 0.05** → 5% of the time you would see a test statistic at least as extreme as the one you found if the null hypothesis was true
- The smaller the p-value, the less likely it is that the observed data can occur under the null hypothesis.

Statistical significance is another way of saying that the *p-value* of a statistical test is small enough to reject the null hypothesis of the test



Significance testing

In February 2014, George Cobb, Professor Emeritus of Mathematics and Statistics at Mount Holyoke College, posed these questions to an ASA discussion forum:

Q:*Why do so many colleges and grad schools teach p=0.05?*
A: *Because that's still what the scientific community and journal editors use.*
Q: *Why do so many people still use p = 0.05?*
A: *Because that's what they were taught in college or grad school*.

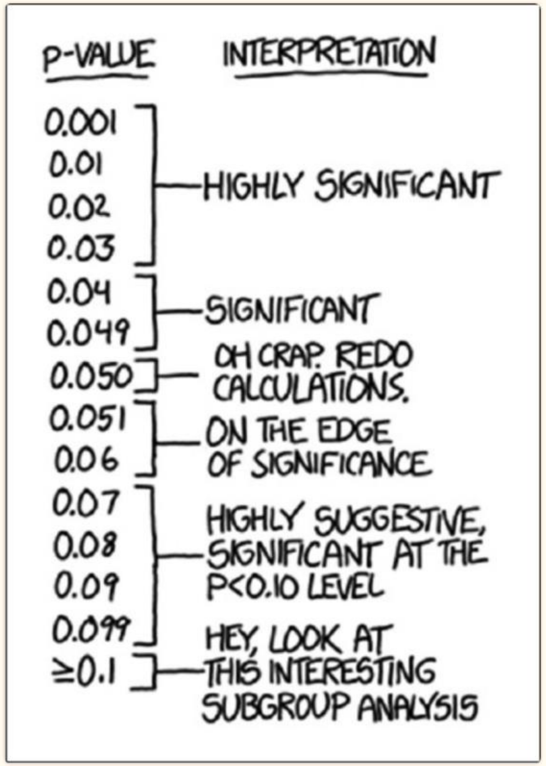


Figure 1

"If all else fails, use 'significant at *P* > 0.05 level' and hope no one notices." (http://xkcd.com/1478/, Randall Munroe, Creative Commons Attribution-NonCommercial 2.5 License)

Significance testing

American Statistical Association statement on P-values

1. P-values can indicate how incompatible the data are with a specified statistical model
2. P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
3. Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.
4. Proper inference requires full reporting and transparency.
5. A p-value, or statistical significance, does not measure the size of an effect or the importance of a result.
6. By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis.

ASA News
AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics®

732 North Washington Street, Alexandria, VA 22314 • (703) 684-1221 • Toll Free: (888) 231-3473 • www.amstat.org • www.twitter.com/AmstatNews

AMERICAN STATISTICAL ASSOCIATION RELEASES STATEMENT ON STATISTICAL SIGNIFICANCE AND *P*-VALUES

Provides Principles to Improve the Conduct and Interpretation of Quantitative Science

March 7, 2016

Significance testing

So.. Should we get rid of p-value?

- The p-value can only tell you **whether or not the null hypothesis is supported**. It cannot tell you whether your alternative hypothesis is true, or why.

→ Statistical significance

- supplement them in very sensible ways with estimates of the precision and importance of our results → **confidence intervals** an **effect sizes**

→ Practical significance

Confidence intervall

This is the range of values you expect your estimate to fall between if you redo your test, within a certain level of confidence

- the range that encompasses plus or minus two standard errors

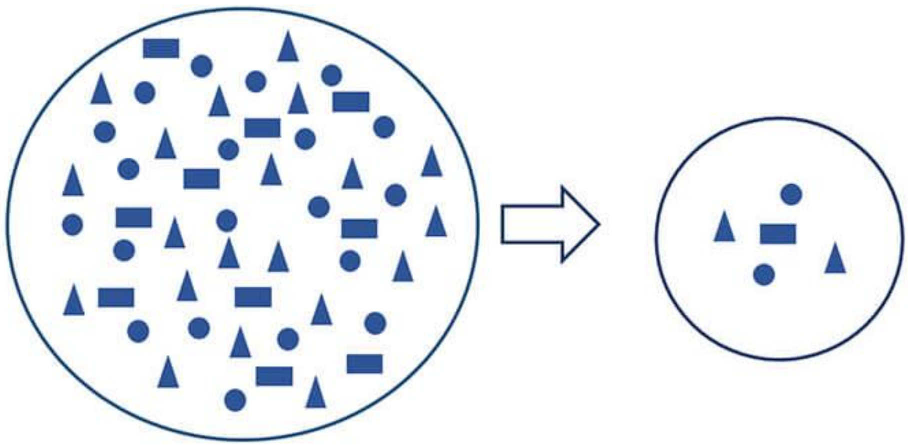
→ confident that f. ex. 95 percent of the time our estimate will be in our interval defined by two times the standard error.

- **p-value + confidence interval** → probability of our results and how it fits our level of risk cutoff criterion for testing hypothesis

Confidence intervall

- error, reflecting inevitable inaccuracy that occurs when observing only one population
→ **Using the standard error** to report on how big the range is; how precise estimate
- This is the **range of values you expect your estimate to fall between** if you redo your test, within a certain level of confidence
- f. ex. confident that 95 percent of the time our estimate will be in our **interval defined by two times the standard error.**

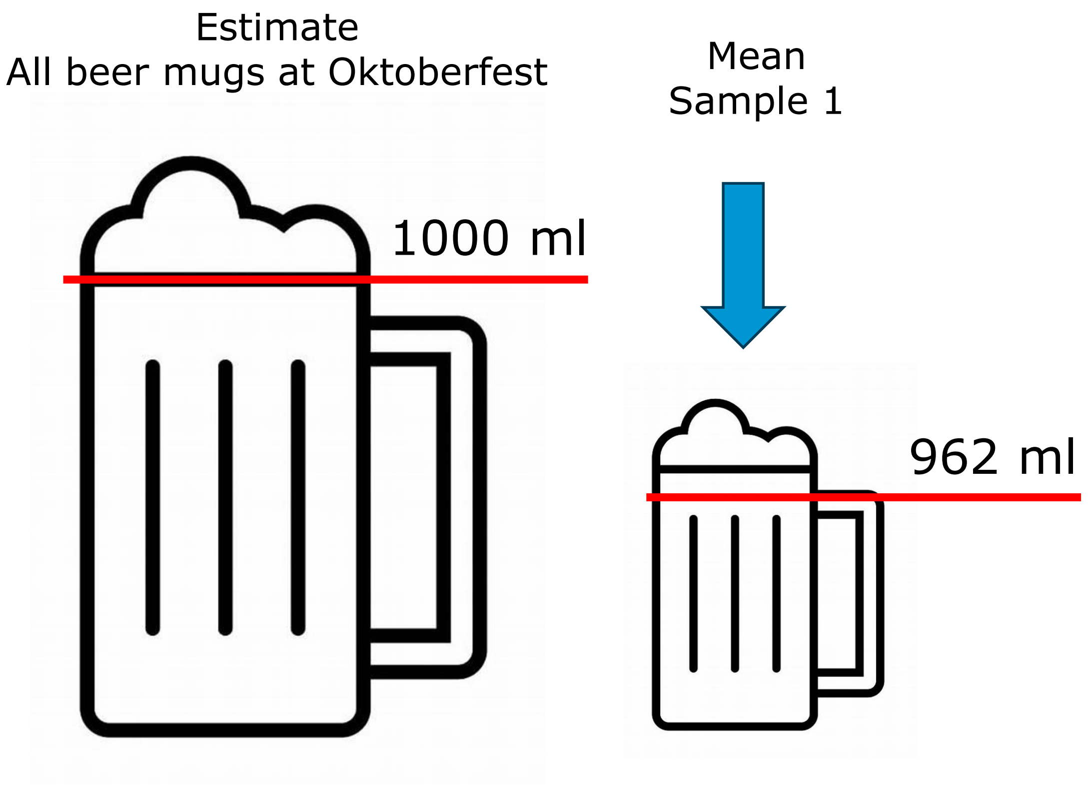
Unknown parameter of the population is estimated based on the sample
→ f. ex. variance, mean, etc.



Population **Sample**

Confidence intervall

Is poured too little beer into the mugs at Oktoberfest systematically?

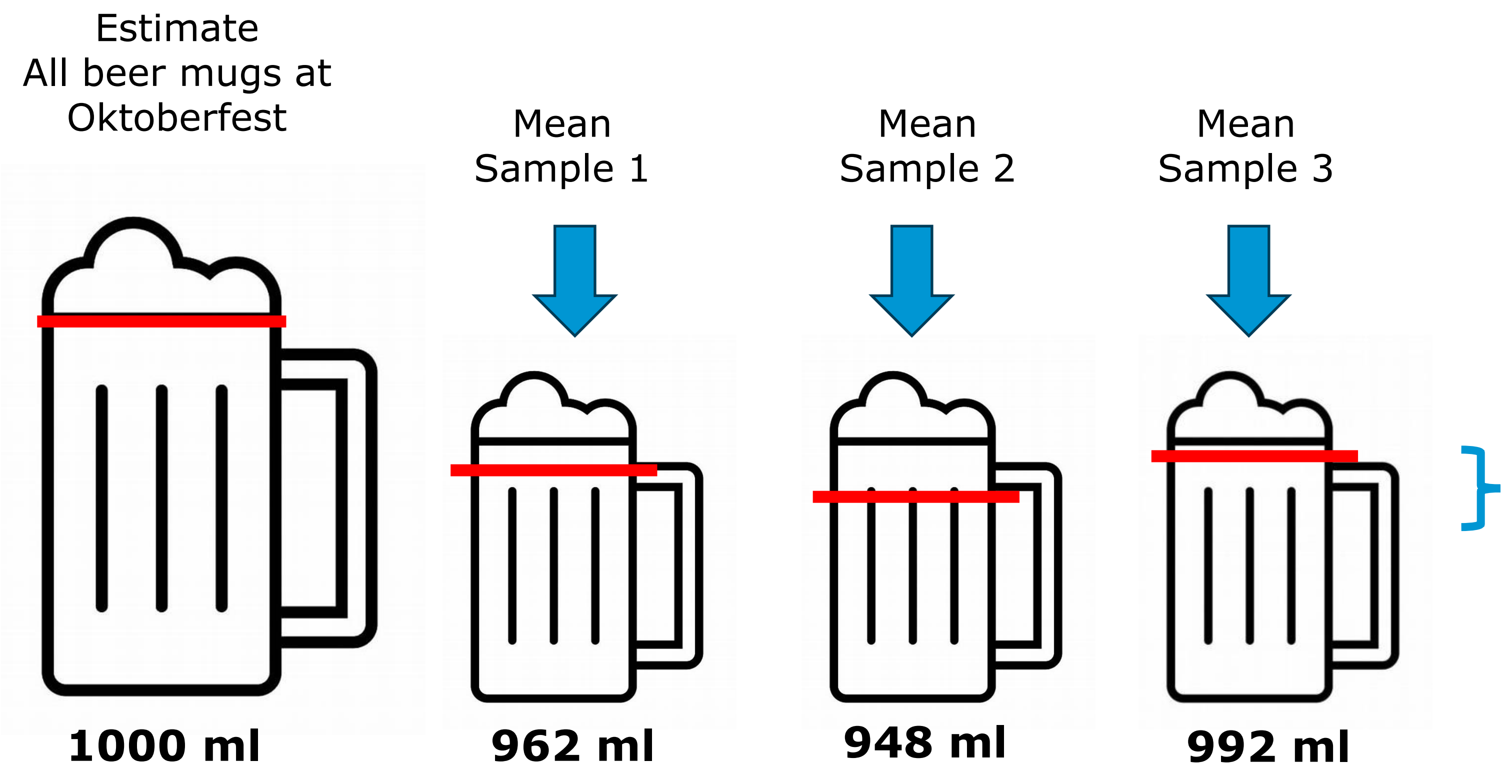


- How accurate is this estimate of 962ml?
- What is the most likely range of the true average?
- Is it possible that the total/true average is also 950ml? Is it even possible that in reality the average is 1000ml, but we were just unlucky in this sample?

What range is the true value with high probability?

Confidence intervall

Is poured too little beer into the mugs at Oktoberfest systematically?



$$CI = \bar{x} \pm z\frac{s}{\sqrt{n}}$$

CI = confidence interval

$\bar{x}$ = sample mean

z = confidence level value

s = sample standard deviation

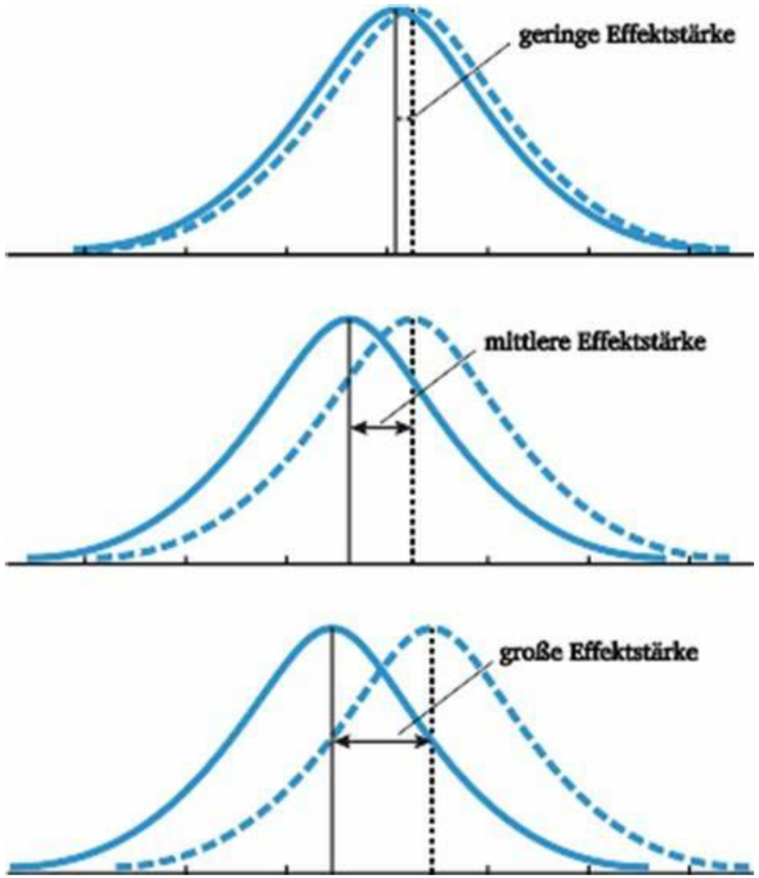
n = sample size

confidence level (probability of error)
range in which true parameter lies with 95% certainty; α = 0.05 or 5%

Effect size

effect size is a way of quantifying the difference between two groups

→ "How big is the result?" Practical relevance/significance



- Effect sizes complement statistical hypothesis testing, and play an important role in power analyses, sample size planning, and in meta-analyses
- **magnitude** and **direction** of the difference between two groups or the relationship between two variables

Effect size

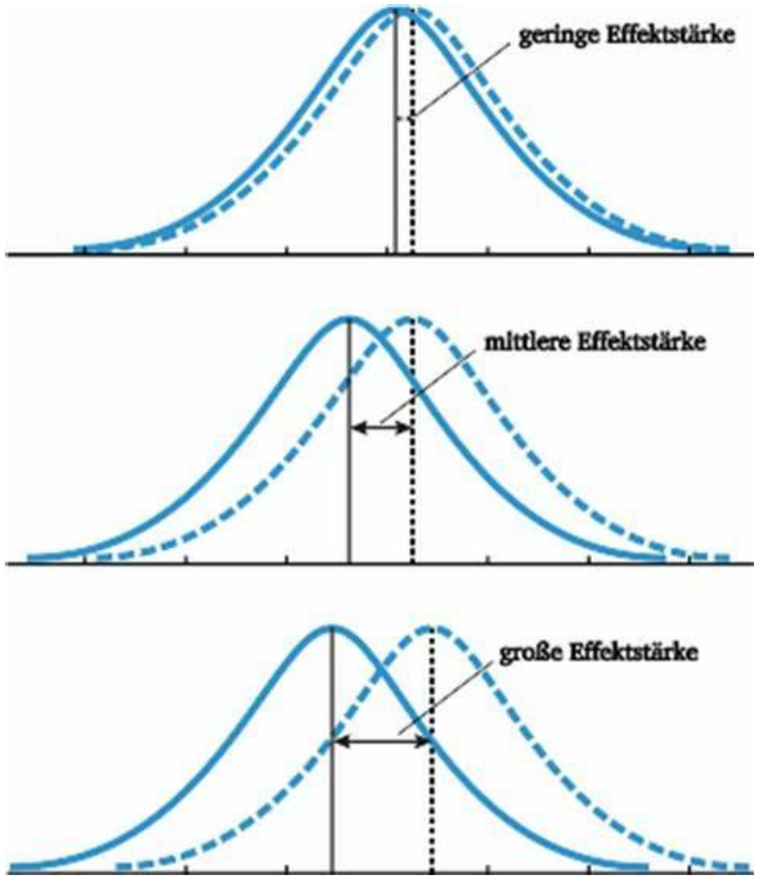
f. ex. **Cohen's d = (x1 – x2) / s**

- X1 = sample mean group 1
- X2 = sample mean group 2
- s = standard deviation of the population

→ How far the **signal-to-noise ratio deviates from zero**

→ If H0 (no relationship) is true, signal-to-noise ratio is zero, and so is the effect

→ Since you can always diminish the noise level by increasing the sample size, your estimate effect and the p-value associated with it reflect the sample size



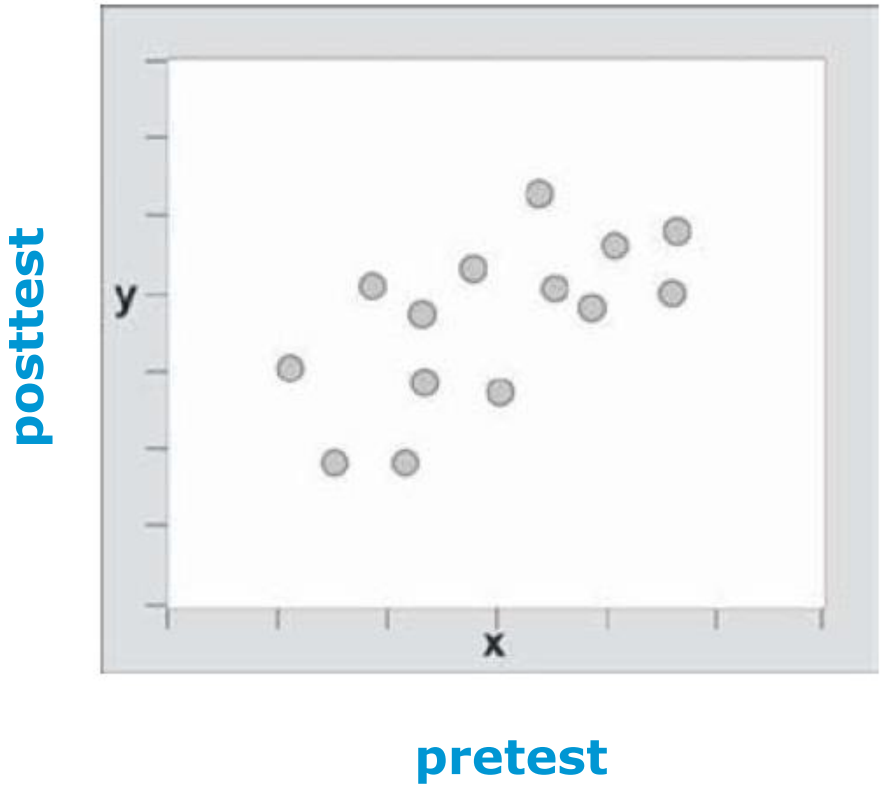
General Linear Model

- system of equations used as mathematical framework for most of the statistical analyses used in applied social research
- Foundation for t-test, ANOVA, ANCOVA, regression analysis, multivariate analysis,…

Why General Linear Model?

- Linear: fitting a line
- Model: equation that summarizes the line that you fit

→ **aim: describing general patterns**



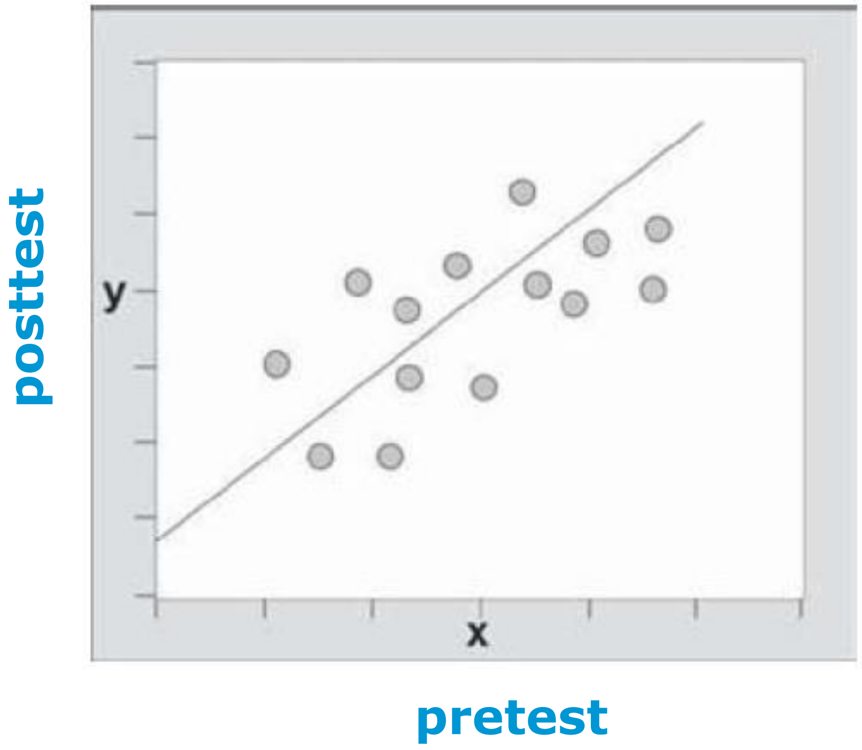
General Linear Model: Two-variable Linear Model

<u>**Two-variable-Linear Model**</u>

Each dot = pretest + posttest score → positive relationship

→ **What is happening? How to summarize the data?**

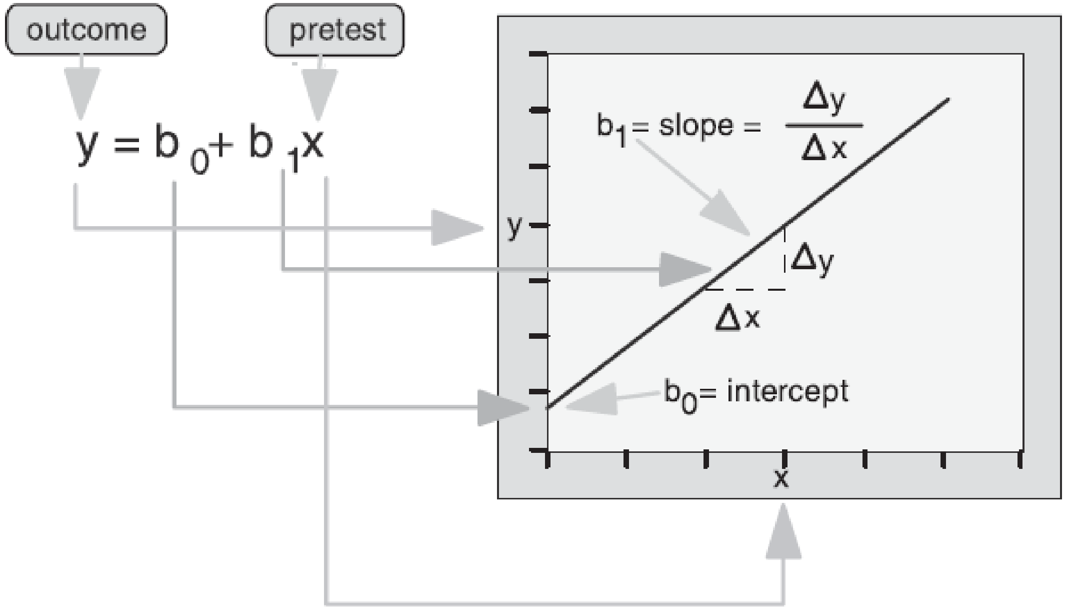
Regression line
describes relationship between two variables, just like any deskriptive statistic, f.ex. mean



General Linear Model: Two-variable Linear Model

- Y = y-axis variable, outcome or posttest
- X = x-axis variable, pretest
- b0 = intercept, value of y when x=0
- b1 = slope (change of y for a change in x of one unit)

Why do we need slope?
→ describes the way this line fits to bivariate plot



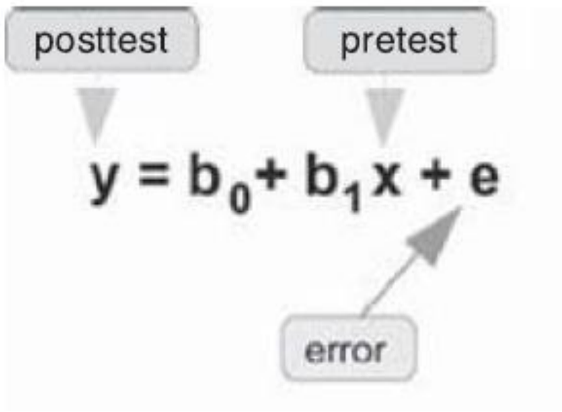
General Linear Model: Two-variable Linear Model

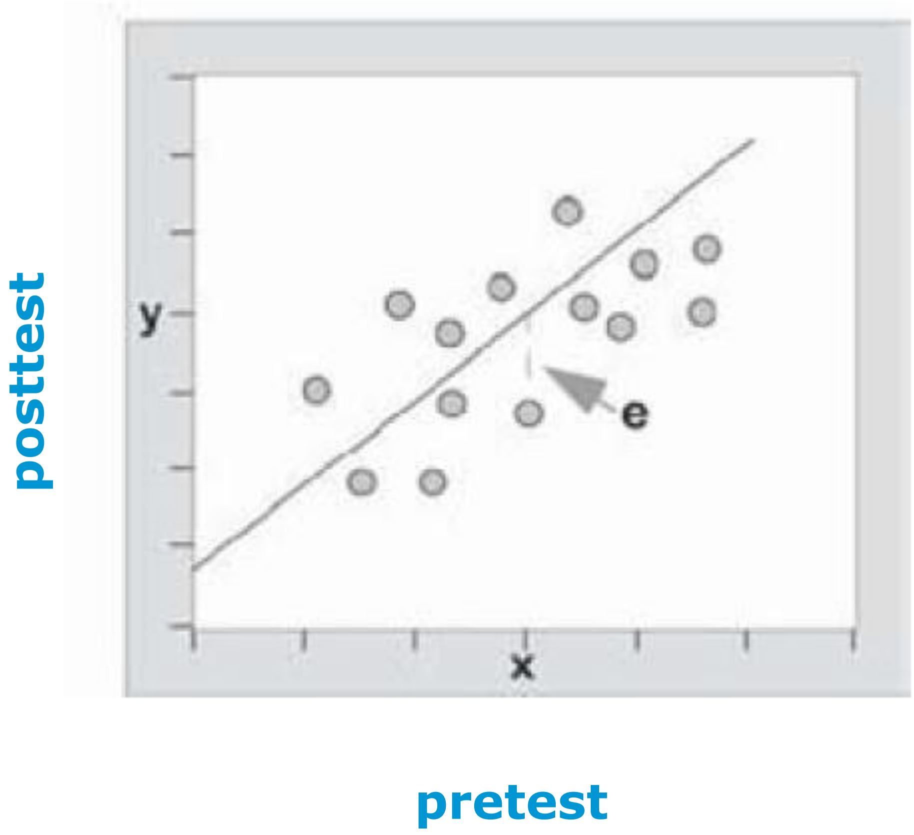
- Equation does not perfectly fit cloud of points

→ One more compontent?

Error term

Captures degree to which the line is in error in describing each point





General Linear Model: General Case

- Difference: each of four terms can represent a **set of variables**

$$y = b_0 + bx + e$$

Set of outcome variables

Set of intercepts
(value of each y when each x = 0)

Set of coefficients
(one each for each x)

Set of pre-program variable / covariates
(adjust in your study)

General Linear Model: General Case

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

$$y = b_0 + bx + e$$

- Difference: each of four terms can represent a **set of variables**
- Summarize a variety of research outcomes:
  - Experimental or quasi-experimental study: represent program/treatment with dummy-coded variables (Z for dummy-coded x)
  - Multiple outcome variables
  - Multiple pretests
  - → For each x-value you estimate a b-value that represents x-y-relationship
  - → Test relationships between variables or differences between groups

Major problem: which equation summarizes the data best?

→ **Model specification** = process of stating the equation that believe best summarizes the data for a study

Gerneral Linear Model: Dummy Variables

Dummy variable (Z)

Numerical variable used in regression analysis to represent subgroup of the sample in your study

- Treat a nominal-variable like an interval-level variable
- Like switches, turn paramenters „on and off"

→ Enables to **use as single regression equation to represent multiple groups**

Gerneral Linear Model: Dummy Variables

Example:

Posttest-only two-group randomized experiment

Eaquation for each group separatly:

$$y_i = \beta_0 + \beta_1 Z_i + e_i$$

y_i = outcome score for the i^{th} unit
β_0 = coefficient for the *intercept*
β_1 = coefficient for the *slope*
Z_i = 1 if i^{th} unit is in the treatment group
0 if i^{th} unit is in the control group
e_i = residual for the i^{th} unit

→ Predicted value for control group is β0

→ Predicted value for treatment group is β0 + β1

Control group
0

Treatment group
1

Gerneral Linear Model: Dummy Variables

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Example:

Posttest-only two-group randomized experiment

Eaquation for each group separatly:

$$y_i = \beta_0 + \beta_1 Z_i + e_i$$

First, determine effect for each group:

For Control group ($Z_i = 0$):

$$y_C = \beta_0 + \beta_1(0) + 0$$

$$y_C = \beta_0$$

For treatment group ($Z_i = 1$):

$$y_T = \beta_0 + \beta_1(1) + 0$$

$$y_T = \beta_0 + \beta_1$$

e_i averages to 0 across the group

→ Predicted value for control group is β0

→ Predicted value for treatment group is β0 + β1

Control group
0

Treatment group
1

Gerneral Linear Model: Dummy Variables

Example:
Posttest-only two-group randomized experiment

How to determine the difference between the two groups?

→ Difference between equations of two groups

treatment	control
$y_T = \beta_0 + \beta_1$	$y_C = \beta_0$

$$y_T - y_C = (\beta_0 + \beta_1) - \beta_0$$

$$y_T - y_C = \beta_0 + \beta_1 - \beta_0$$

$$y_T - y_C = \beta_1$$

Control group
0

Treatment group
1

Gerneral Linear Model: Dummy Variables

Example:
Posttest-only two-group randomized experiment

How to determine the difference between the two groups?
→ Difference between equations of two groups

<u>Take away:</u>

- Create seperate equations for each group by substituting dummy variable
- Find difference between two groups by difference between their equations

Control group
0

Treatment group
1

What is one of the simplest inferential tests when you want to compare the average performance of two groups on a single measure to see whether there is a difference…?

Understanding the t-Test

- **Purpose**: To analyze the differences between two groups in a posttest-only randomized experimental design.

- **Key Requirements**:
  - Analysis has two groups
  - Uses a post-only measure
  - Has two distributions (measures), each with an average and variation
  - Assesses treatment effect = statistical (non-chance) difference between the groups

Why Use a t-Test? The t-Test assesses whether the means of two groups are *statistically* different from each other.

- **Significance**: Determines if observed differences are due to the treatment and not by chance.

Why “t“-Test?

Remember the formula for the straight line: y=mx+b?

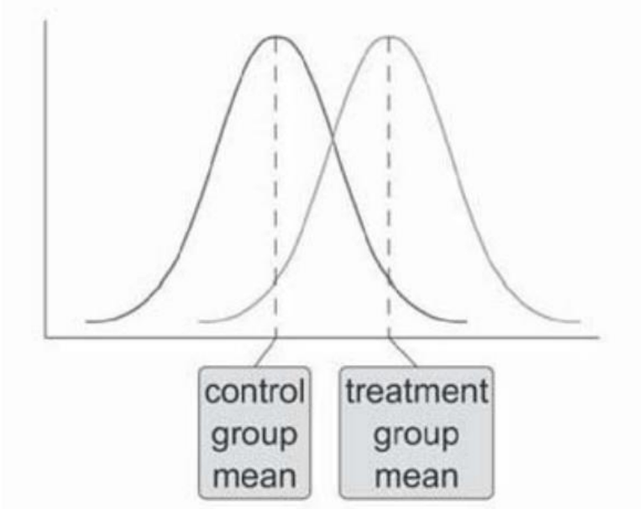
Using the name t-Test is like calling that formula the y-formula.

The statician who invented this analysis first wrote out the formula, he used the letter “t“ by to symbolize the value that describes the **difference** between the groups

Idealized distribution for treated and control group posttest values

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

„Is there a differnce between the groups?“: Each group can be represented by a bell-shaped curve describing the distribution on a single variable.



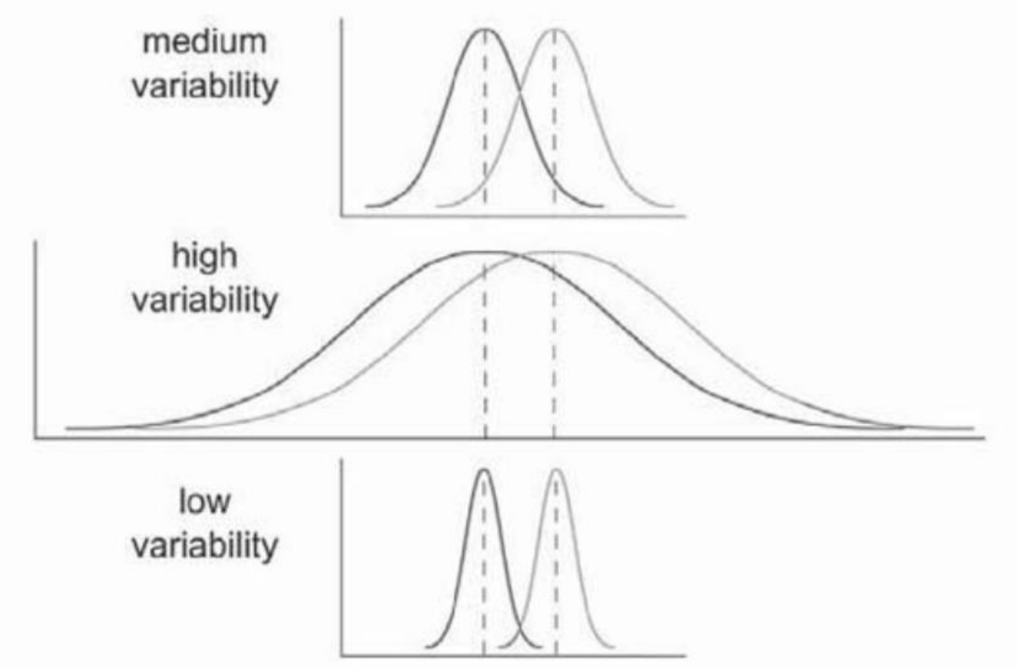
- **Dotted line:** the Distribution for the treated groups in a study
- **Solid line:** Control groups in a study

→ **Idealized/smoothed distribution**

- Figure indicates where the control and treatment group means are located.
- The t-Test asseses whether the means of two groups are statistically different from each other.

Three Scenarios for Differences Between Means

- **Moderate Variability**: Scores within each group show moderate overlap.
- **High Variability**: High overlap, making group differences less striking despite identical mean differences.
- **Low Variability**: High-variability case with little overlap between curves.



Conclusion: Differences must be judged relative to the variability of scores. The t-Test quantifies this relationship.

T-Test Summary

1. Control group and treatment group posttest scores.
2. Calculation of t-value.
3. Interpretation of results.

☐ **Key Takeaway**: t-Test provides a clear method to assess treatment effects in experimental designs.

- **Summary**:
  - t-Test evaluates differences between two group means.
  - Takes into account the variability within groups.
  - Essential for determining statistical significance in experiments.

Statistical Analysis of a t-Test: Formula

- t-Test formula is a ratio
  - **Numerator (top part of ratio)**: Difference between the two means (averages).
  - **Denominator (bottom part of ratio)**: Measure of variability or dispersion of the scores.

- Formula is an example of the signal-to-noise metaphor in research.
  - **Signal**: Difference between the means (effect of the treatment).
  - **Noise**: Variability in the data (makes it harder to detect the effect).

☐ The ratio that is computed is called a t-value – describes differences between the groups relative to variability of the scores in the groups

Formula

FIGURE 14–7a Formula for the *t*-test

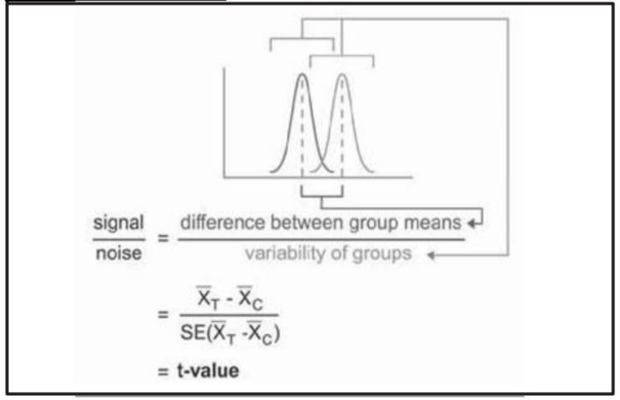


FIGURE 14–7b Formula for the standard error of the difference between the means

$$SE(\bar{X}_T - \bar{X}_C) = \sqrt{\frac{var_T}{n_T} + \frac{var_C}{n_C}}$$

FIGURE 14–7c Formula for the *t*-test

$$t = \frac{\bar{X}_T - \bar{X}_C}{\sqrt{\frac{var_T}{n_T} + \frac{var_C}{n_C}}}$$

- **Figure 14-7a:** shows the formula for the t-test and how the numerator and denominator are related to the distribution.
- **Figure 14-7b:** specific formula for standard error of the differnce.
- **Figure 14-7c:** Final formula for the t-test.

t-Value

- **t-Value**: The computed ratio that describes the difference between the groups relative to the variability of the scores.

- The formula's numerator and denominator relate to the distributions.
  - **Numerator**: Difference between means.
  - **Denominator**: Standard error of the difference, calculated as:

$$\text{Standard Error} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- s_1^2 and s_2^2 are the variances of the groups.
- n_1 and n_2 are the sample sizes.

Interpreting the t-Value

- **Sign of t-Value**:
  - Positive if the first mean is larger.
  - Negative if the first mean is smaller.

- **Significance Testing**:
  1. Calculate the t-value.
  2. Look up the p-value associated with the t-value in a table of significance (many statistical programs automatically provide the p-value).
  3. Look up the p-value associated with the t-value in a table of significance (many statistical programs automatically provide the p-value). → To test whether the t-ration is large enough to say that the difference between the groups is not likely to have been a chance finding.
  4. Set a risk level (called the alpha level) at .05. → Means that 5 times out of 100, you would find a statistically significant difference between the means even if there were none.
  5. Determine degrees of freedom (df): $df = (n_1 + n_2 - 2)$
  6. Use df, t-value, and alpha to find the p-value in a significance table.
  7. Compare the p-value with alpha to determine if the t-value is significant.

Conclusion: If significant, the difference between group means is unlikely due to chance

Testing the Significance

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Methods to Estimate Treatment Effect

- **Three Approaches**:
  - **Independent t-Test**: As described.
  - **One-Way ANOVA**: Between two independent groups.
  - **Regression Analysis**: Regress posttest values onto a dummy-coded treatment variable. (most general)

→ **Note**: All three methods yield identical results.

Regression formula for t-test or two-group one-way analysis of variance (ANOVA)

- It is identical to the formula to introduce dummy variables.

- Essentially this formula is the equation for a straight line with a random error term thrown in.

$$y_i = \beta_0 + \beta_1 Z_i + e_i$$

where:

y_i = outcome score for the i^{th} unit

β_0 = coefficient for the *intercept*

β_1 = coefficient for the *slope*

Z_i = 1 if i^{th} unit is in the treatment group

0 if i^{th} unit is in the control group

e_i = residual for the i^{th} unit

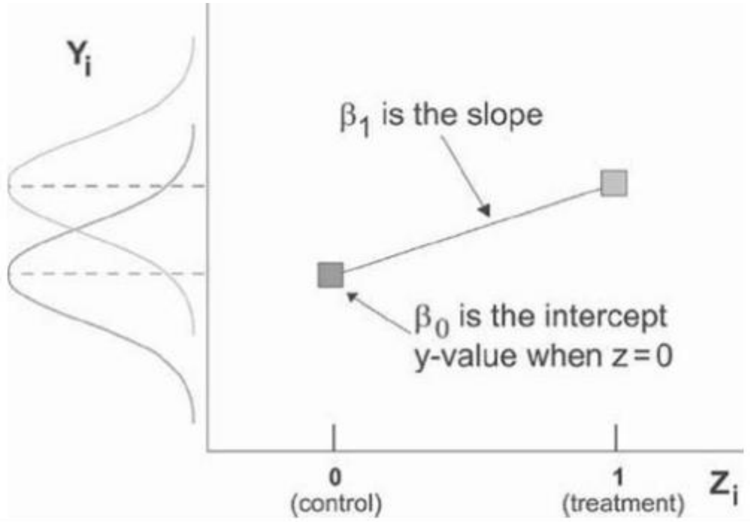
Elements of the Equation in Graphic Form

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

- Equation for a Straight Line: $y=mx+b$
- In statistical terms:

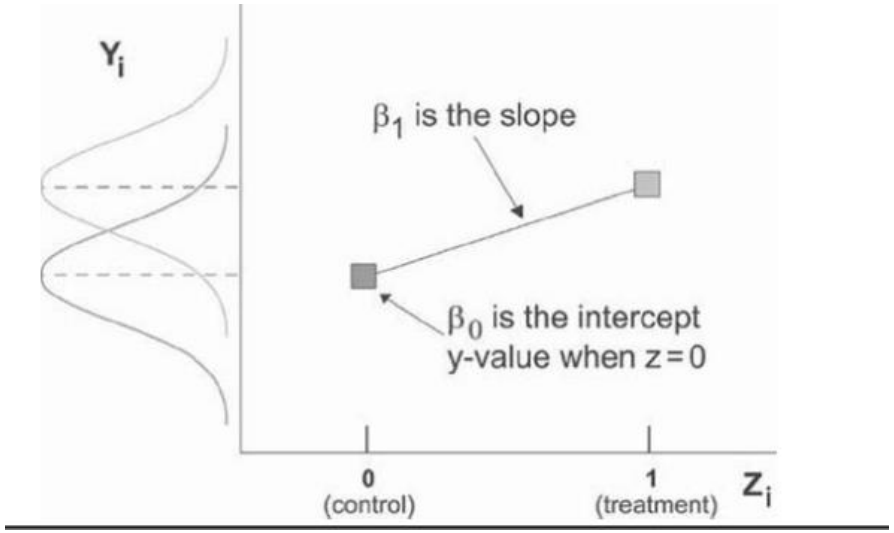
$$y_i = \beta_0 + \beta_1 Z_i + \epsilon_i$$

- y_i: Posttest score
- β_0: Intercept (control group mean)
- β_1: Slope (difference between means)
- Z_i: Dummy variable (0 for control, 1 for treatment)



- **Figure Explanation**: Graph shows posttest on vertical axis.
  - Horizontal axis: Dummy variable Z (0 for control, 1 for treatment).
  - The slope of the line indicates the difference in posttest means.
- **Conclusion**: Slope (difference in means) adjusted for variability provides the treatment effect.

Elements of the Equation in Graphic Form



$y_i = \beta_0 + \beta_1 Z_i + \epsilon_i$

- y_i: Posttest score (the outcome measured after treatment)
- β_0: Intercept (average score for the control group)
- β_1: Slope (difference in average scores between the treatment and control groups)
- Z_i: Dummy variable (0 for control group, 1 for treatment group)
- ϵ_i: Error term (captures variability not explained by the model)

Figure Explanation:

- The graph shows posttest scores on the vertical axis.
- **Horizontal axis:** Dummy variable Z (0 for control, 1 for treatment).
- **The slope of the line indicates the difference in posttest means:** The line's slope shows the difference in average posttest scores between the groups.

Conslusion

- The slope ($\beta 1$) indicates the treatment effect (difference in means) after adjusting for variability.
- A significant slope means the treatment likely had a real effect.

Summary

The t-test, on-way ANOVA, and regression analysis all yield the *same* results in this case

The regression-analysis method utilizes a dummy variable (Z) for treatment

Regression analysis is the most general model of three

Factorial Design Analysis

- **Definition**: Factorial design involves experiments with more than one independent variable (factor).
- **Purpose**: To study the interaction effects between factors and their individual effects on the dependent variable.
- **Example**: 2x2 factorial design includes two factors, each with two levels.

- **Main Effects**: The effect of each independent variable on the dependent variable.
- **Interaction Effects**: How the independent variables interact to influence the dependent variable.
- **Design Notation**: Described using numbers to indicate levels of factors (e.g., 2x3 factorial design).

- **Setup**: Two factors (e.g., Type of Instruction and Gender), each with two levels (e.g., Traditional vs. Experimental instruction, Male vs. Female).
- **Groups**: Four groups resulting from the combination of levels (Traditional-Male, Traditional-Female, Experimental-Male, Experimental-Female).
- **Analysis**: Evaluate the main effect of each factor and their interaction effect.

Regression model for a simple 2x2 factorial design

Equation:

$$y_i = \beta_0 + \beta_1 Z_{1i} + \beta_2 Z_{2i} + \beta_3 Z_{1i} Z_{2i} + \epsilon_i$$

Explanation:

- y_i: Outcome score for the i-th unit
- β_0: Coefficient for the intercept
- β_1: Mean difference on factor 1
- β_2: Mean difference on factor 2
- β_3: Interaction of factor 1 and factor 2
- Z_{1i}: Dummy variable for factor 1 (0 = 1 hour per week, 1 = 4 hours per week)
- Z_{2i}: Dummy variable for factor 2 (0 = in class, 1 = pull-out)
- ϵ_i: Residual for the i-th unit

Key Points:

- This equation models the impact of two factors and their interaction on the outcome.
- Each coefficient (β) represents the effect size of the corresponding variable.
- Dummy variables (Z) allow us to include categorical factors in the regression model.

Factorial Design

Analyzing Main Effects

- Main effects represent the independent effect of one factor on the dependent variable, averaged over the levels of the other factor.
- **Calculation**: Compare the means of the levels of one factor, ignoring the other factor.
- **Example**: Calculate the average performance of all students receiving traditional instruction versus experimental instruction.

Analyzing Interaction Effects

- **Definition**: Interaction effects occur when the effect of one factor depends on the level of another factor.
- **Identification**: Determine if the differences between levels of one factor vary across levels of the other factor.
- **Graphical Representation**: Interaction plots can visually display interaction effects by plotting means for each combination of factor levels.

Statistical Analysis of Factorial Designs

- **ANOVA**: Analysis of Variance (ANOVA) is used to assess main and interaction effects.
- **Steps**:
    - Compute sum of squares for main effects, interaction effects, and error.
    - Calculate mean squares by dividing sum of squares by respective degrees of freedom.
    - F-ratios: Compare mean squares of effects to mean squares of error.
- **Significance Testing**: Determine if observed effects are statistically significant.

Summary

- Factorial designs allow for the examination of multiple factors and their interactions.
- Main effects and interaction effects provide comprehensive insights into the factors' influences.
- ANOVA is the primary tool for analyzing factorial designs.

Advantages:

- Efficiency: Examines multiple factors simultaneously.
- Interaction Insights: Reveals interaction effects not identifiable in single-factor designs.

Disadvantages:

- Complexity: More factors increase the complexity of the design and analysis.
- Larger Sample Size: Requires more participants to maintain power.

Randomizes Block Design

- **Definition**: A randomized block design is an experimental setup that groups subjects into blocks based on a certain characteristic before randomly assigning treatments within each block.
- **Purpose**: To control for variability among subjects and increase the precision of the experiment by reducing the impact of confounding variables.
- **Example**: Blocking by age, gender, or pre-existing conditions to ensure balanced groups.

Key Features

- **Blocking**: Subjects are divided into homogeneous blocks to control for specific variables.
- **Randomization**: Within each block, subjects are randomly assigned to different treatment groups.
- **Improved Precision**: By controlling for block variables, this design increases the accuracy and reliability of the treatment effect estimation.

Statistical Analysis

ANOVA for Randomized Block Design: Analysis of Variance (ANOVA) is used to compare the means across treatment groups while accounting for block effects.

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij}$$

- Y_{ij}: Response variable
- μ: Overall mean
- τ_i: Effect of the i-th treatment
- β_j: Effect of the j-th block
- ϵ_{ij}: Random error term

- **Sum of Squares:** Partitioning the total variability into components due to treatments, blocks, and error.

Randomizes Block Design

- RD can also be presented in regression analysis notation
- Figure shows a model for a case where there are four blocks or homogeneous subgroups
- Note: a number of dummy variables are used to specify this model
- Z1: treatment Group
- Dummy variables Z2,Z3,Z4 indicates blocks 2-4

$$y_i = \beta_0 + \beta_1 Z_{1i} + \beta_2 Z_{2i} + \beta_3 Z_{3i} + \beta_4 Z_{4i} + e_i$$

where:

y_i = outcome score for the i^{th} unit

β_0 = coefficient for the intercept

β_1 = mean difference for treatment

β_2 = blocking coefficient for block 2

β_3 = blocking coefficient for block 3

β_4 = blocking coefficient for block 4

Z_{1i} = dummy variable for treatment (0 = control, 1 = treatment)

Z_{2i} = 1 if block 2, 0 otherwise

Z_{3i} = 1 if block 3, 0 otherwise

Z_{4i} = 1 if block 4, 0 otherwise

e_i = residual for the i^{th} unit

Steps in Randomized Block Design

1. **Identify Blocking Variable**: Choose a characteristic that is expected to affect the outcome.
2. **Form Blocks**: Group subjects into blocks based on the chosen variable.
3. **Random Assignment**: Randomly assign subjects within each block to different treatment groups.
4. **Conduct Experiment**: Apply treatments and measure outcomes.
5. **Analyze Data**: Use appropriate statistical methods to assess treatment effects within and across blocks.

Scenario: Testing a new drug's effect on blood pressure with blocks based on age groups (young, middle-aged, elderly).

Steps:

1. Form blocks by age group.
2. Randomly assign subjects within each age group to the drug or placebo.
3. Measure blood pressure changes.
4. Analyze the results using ANOVA to determine the drug's effect while accounting for age differences.

Example and Conslusion

Advantages:

- Controls for variability within blocks.
- Increases precision of treatment effect estimation.
- Reduces experimental error.

Disadvantages:

- More complex to design and analyze.
- Requires careful selection of blocking variables.

Practical Considerations

- **Selecting Blocking Variables**: Choose variables that significantly influence the outcome.
- **Sample Size**: Ensure each block has sufficient sample size for reliable analysis.
- **Implementation**: Proper randomization within blocks is crucial for validity.

Summary:

- Randomized block design enhances experimental precision by controlling for specific variables.
- ANOVA is used for analyzing data, accounting for block effects.
- Effective for reducing experimental error and increasing the reliability of results.

Analysis of Covariance

- **Definition**: ANCOVA combines ANOVA and regression to evaluate whether population means differ when controlling for covariates.

- **Purpose**: To increase statistical power by reducing error variance.

- **Application**: Commonly used to adjust for pre-existing differences between groups.

Key Components

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

- **Dependent Variable**: The outcome being measured.
- **Independent Variable(s)**: The factors or treatments being tested.
- **Covariate(s)**: Continuous variables that are controlled for, reducing the impact of extraneous variables.

$$Y_{ij} = \mu + \tau_i + \beta(X_{ij} - \bar{X}) + \epsilon_{ij}$$

- Y_{ij}: Dependent variable
- μ: Overall mean
- τ_i: Treatment effect
- β: Regression coefficient for the covariate
- X_{ij}: Covariate value
- $\bar{X}$: Mean of the covariate
- ϵ_{ij}: Random error term

Steps in Performing ANCOVA

WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

1. **Collect Data**: Gather data for dependent variable, independent variable(s), and covariate(s).

1. **Check Assumptions**:
   - Linearity: Relationship between covariate and dependent variable should be linear.
   - Homogeneity of Regression Slopes: Slopes of regression lines should be similar across groups.

2. **Conduct ANCOVA**:
   - Adjust the dependent variable for the covariate.
   - Compare adjusted means of the groups.

3. **Interpret Results**: Determine if the adjusted group means differ significantly.

Statistical Analysis in ANCOVA

- **Adjusting for Covariates**: The covariate adjusts the dependent variable to account for pre-existing differences.
- **F-Test**: Used to test the significance of the main effects and interaction effects.
- **Sum of Squares**:
  - **Total Sum of Squares (SST)**: Total variation in the data.
  - **Sum of Squares for Regression (SSR)**: Variation explained by the model.
  - **Sum of Squares for Error (SSE)**: Variation not explained by the model.

Example:

- **Scenario**: Evaluating a new educational program's effect on student performance, controlling for initial test scores.
- **Steps**:
  - Measure initial test scores (covariate).
  - Apply the program (independent variable).
  - Measure final test scores (dependent variable).
  - Use ANCOVA to adjust final scores based on initial scores.

Summary

- ANCOVA adjusts for covariates to compare group means more accurately.
- Enhances the ability to detect significant differences by controlling for extraneous variables.
- Essential tool for improving experimental precision.

Advantages:

- Controls for pre-existing differences.
- Increases statistical power by reducing error variance.
- Provides a clearer understanding of the treatment effect.

Disadvantages:

- More complex analysis.
- Assumptions need to be met for valid results.

Practical Considerations

- **Choosing Covariates**: Select covariates that are related to the dependent variable but not influenced by the treatment.
- **Sample Size**: Ensure adequate sample size for reliable analysis.
- **Software**: Most statistical software packages (e.g., SPSS, SAS, R) can perform ANCOVA.