UNDERSTANDING RESEARCH RESULTS: STATISTICAL INFERENCE

PSYC214

Chapter 13

Assoc. Prof. Dr. Şenel Hüsnü Raman

Samples and Populations

Inferential statistics are used to determine whether the results match what would happen if the experiments were repeatedly conducted with multiple samples.

- Help researchers make conclusions on the basis of sample data

Inferential Statistics

These allow researchers to make inferences about the true difference in population based on a sample of data.

They give the probability that the difference between means reflects random error rather than a real difference.

Null and Research Hypotheses

The **null hypothesis** is that the population means are equal and that the observed difference is due to random error.

The **research hypothesis** is that the population means are not equal.

Statistical significance indicates that there is a low probability that the difference between the obtained sample means was due to random error.

Probability and Sampling Distributions

Probability is the likelihood of the occurrence of some event or outcome.

- Used in statistical inference
- **Alpha level:** The probability required for significance

Sampling distributions are based on the assumption that the null hypothesis is true.

Sample size refers to the total number of observations.

Group Differences: The *t* and *F* Tests 1

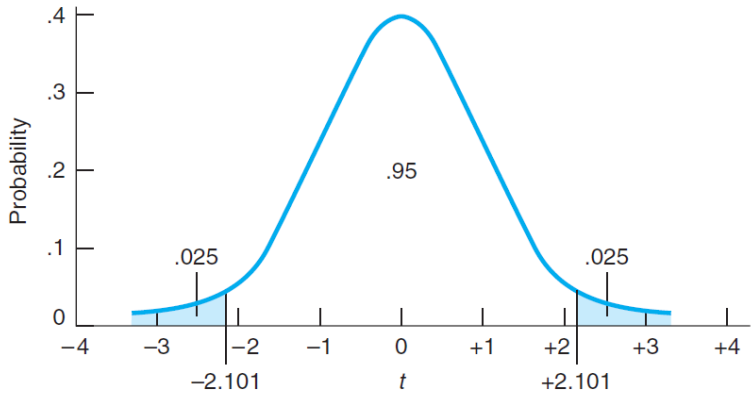
The ***t* test** examines whether two groups are significantly different from each other.

The *t* value is a ratio of two aspects of data:

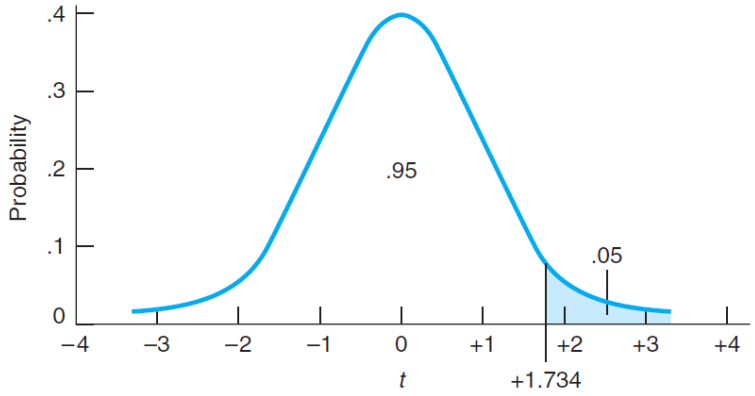
- Difference between the group means
- Variability within groups

$$t = \frac{\text{group difference}}{\text{within-group variability}}$$

Figure 13.1 — Sampling Distributions of *t* Values with 18 Degrees of Freedom



Critical Value for Two-Tailed Test with .05 Significance Level



Critical Value for One-Tailed Test with .05 Significance Level

Group Differences: The *t* and *F* Tests 2

Degrees of freedom (***df***) represents the number of scores free to vary once the means are known.

One-tailed versus two-tailed tests

- One-tailed tests: Critical *t* chosen when research hypothesis specifies a direction of difference between the groups.
- Two-tailed tests: Critical *t* chosen when research hypothesis does not specify a predicted direction of difference.

Group Differences: The *t* and *F* Tests

The ***F* test** or **analysis of variance** is used in the following situations:

- There are more than two levels of an independent variable.
- Factorial design with two or more independent variables has been used.

Systematic variance is the deviation of the group means from the grand mean.

Error variance is the deviation of the individual scores in each group from their respective group means.

The larger the *F* ratio is, the more likely it is that the results are significant.

Group Differences: The *t* and *F* Tests

Calculating effect size:

$$\text{effect size } r = \sqrt{\frac{t^2}{t^2 + df}}$$

Cohen's *d:* Effect size estimate used when comparing two means

$$d = \frac{M_1 - M_2}{\sqrt{\frac{(SD_1^2 + SD_2^2)}{2}}}$$

Group Differences: The *t* and *F* Tests

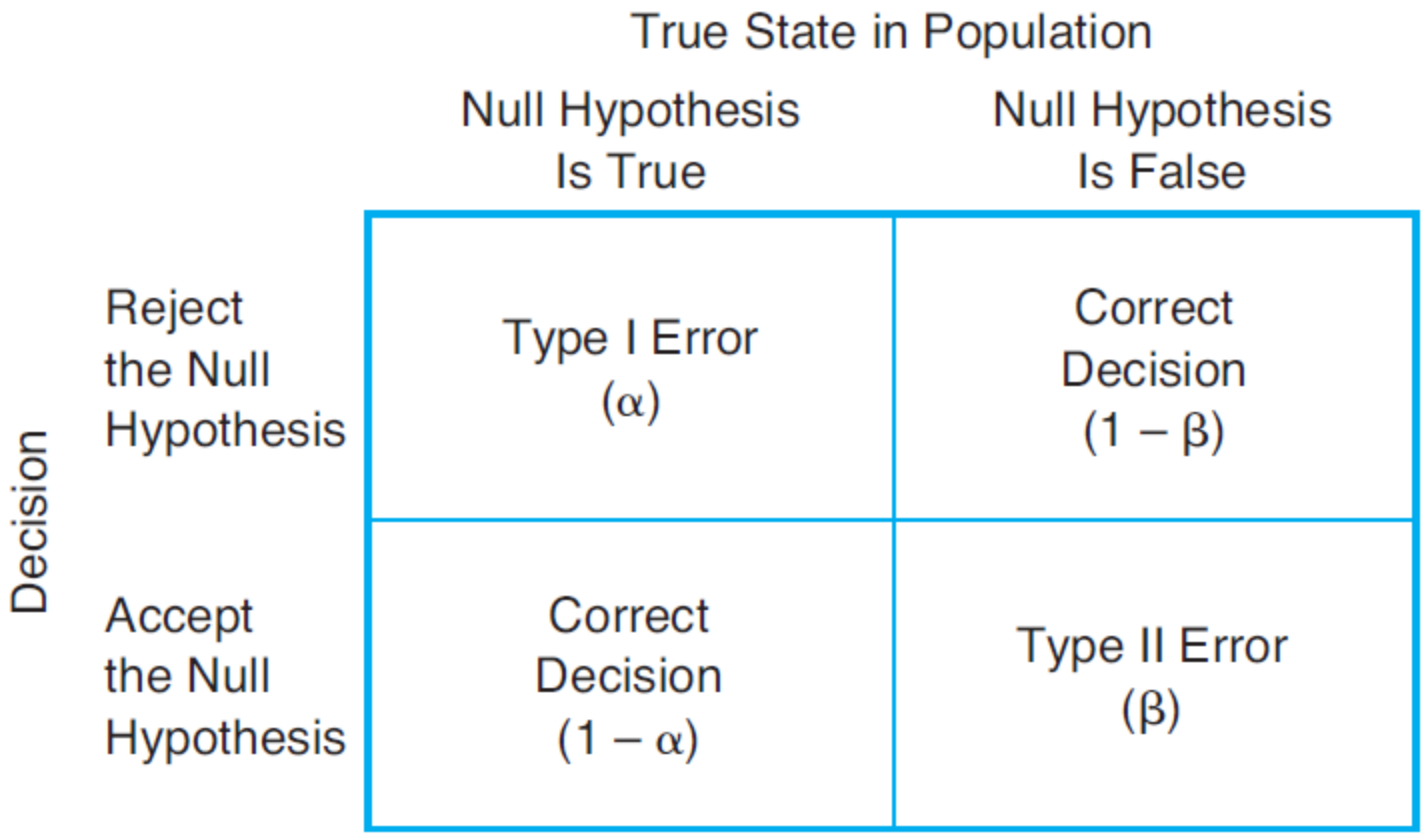
A **confidence interval** is an interval of values within which there is a given level of confidence where the population value lies.

- Represented in bar graphs as a vertical I-shaped line bounded by upper and lower limits.

Statistical significance tests

- The goal is to help testers decide if the obtained results are reliable.
- The chosen significance level indicates how confident you wish to be when making the decision.
- Significant results are most likely when a large sample size is used.

Figure 13.3 — Decision Matrix for Type I and Type II Errors



Correct Decisions

Occur in two instances:

1. When we reject the null hypothesis and research hypothesis is true in the population
2. When we accept the null hypothesis, and the null hypothesis is true in the population.

Type I Errors

A **Type I error** occurs when we reject the null hypothesis when it is actually true.

The probability of making a Type I error is based on significance level chosen.

- The higher the significance or alpha level, the greater the probability of a Type I error.

Type II Errors

Type II errors occur when we accept the null hypothesis when in the population the research hypothesis is true.

The probability of this type of error is related to three factors:

1. Significance (alpha) level
2. Sample size
3. Effect size

Figure 13.4 — Decision Matrix for a Juror

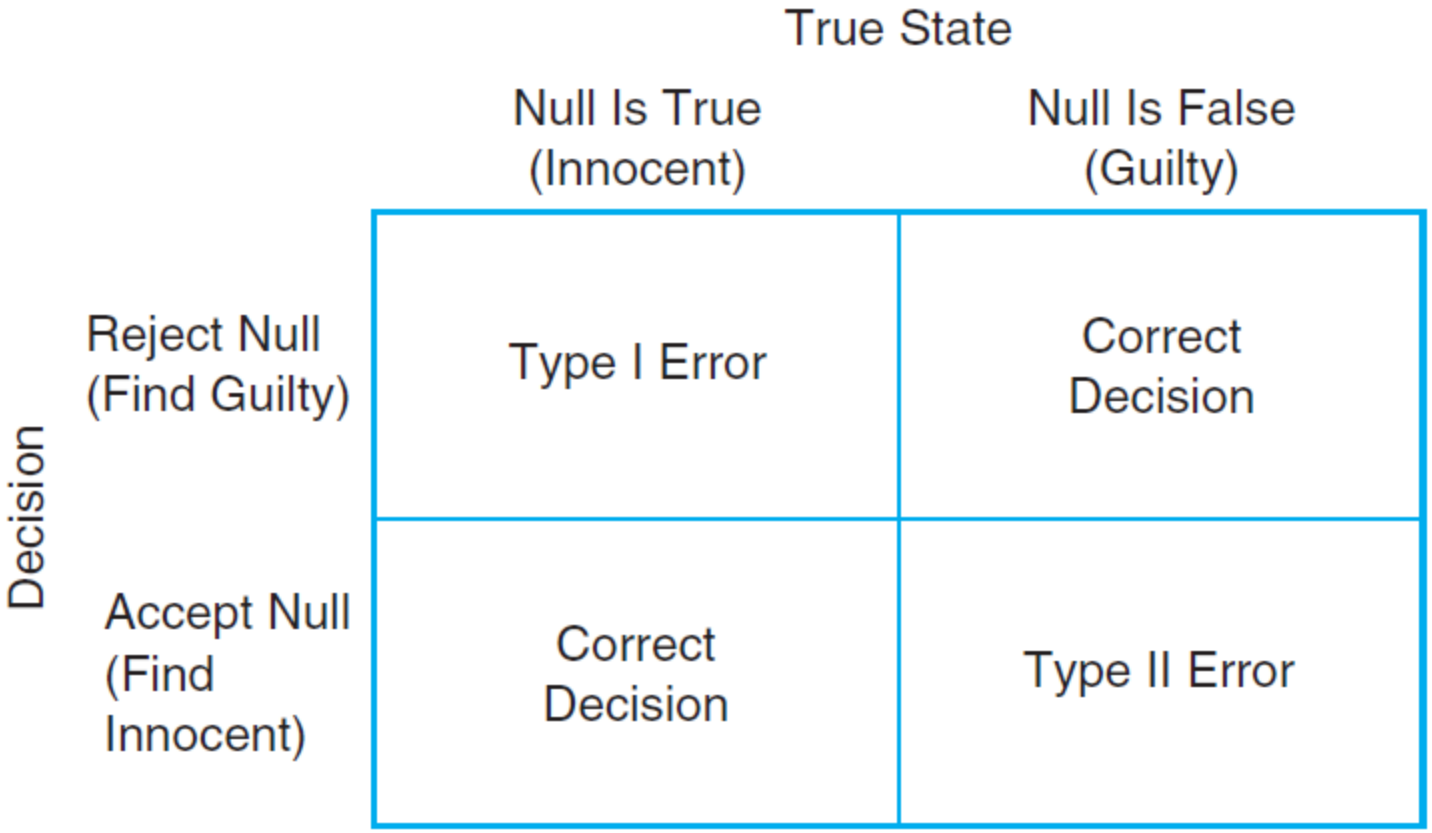
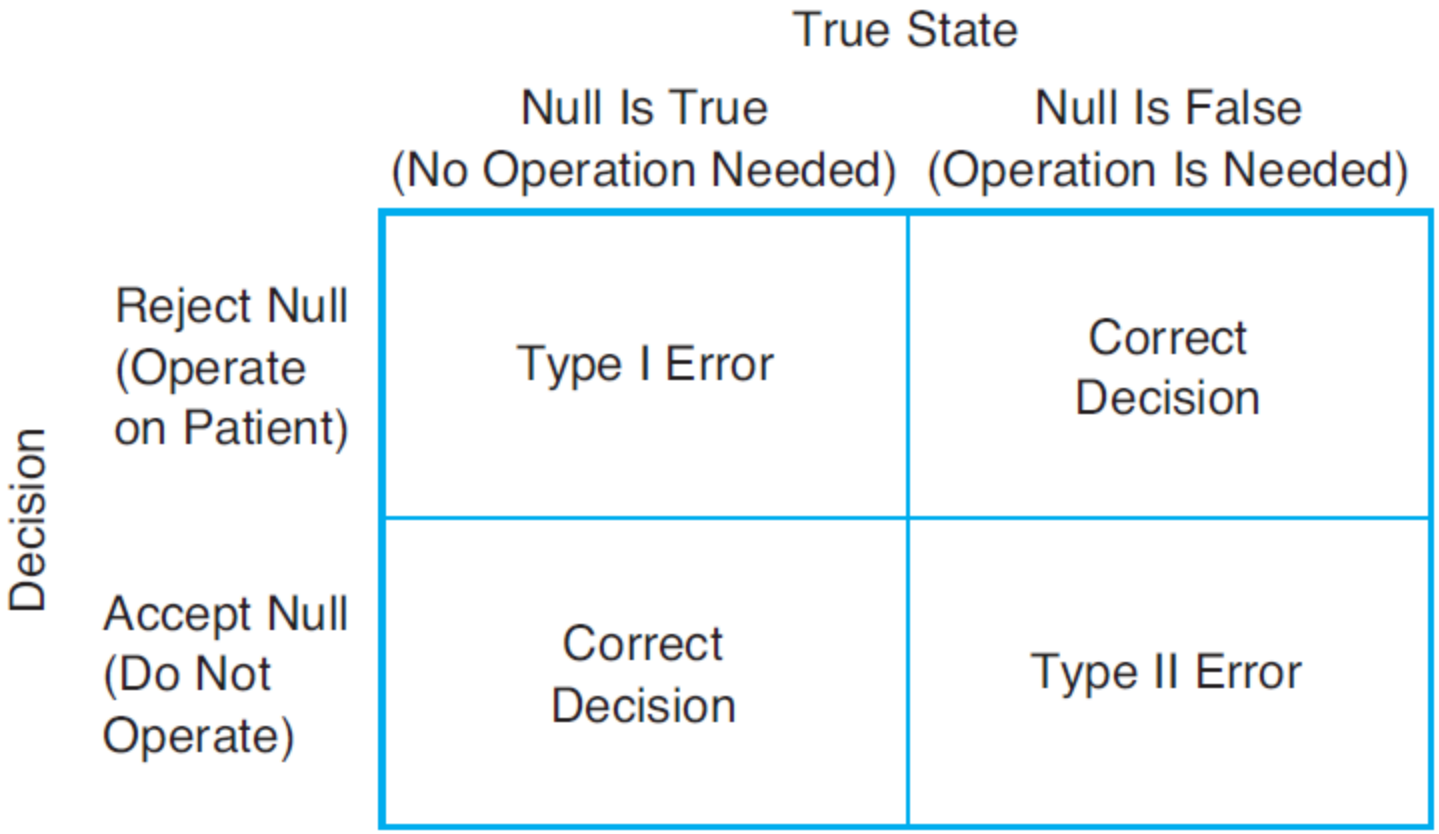


Figure 13.5 — Decision Matrix for a Doctor



Choosing a Significance Level

Researchers have traditionally used a .05 or a .01 significance level in the decision to reject the null hypothesis.

The chosen level specifies the probability of a Type I error if the null hypothesis is rejected.

The level chosen and the consequences of a Type I or a Type II error are determined by the use of the results.

Interpreting Nonsignificant Results

The results of a single study can be nonsignificant even when a relationship between variables in the population exist.

A meaningful result can be overlooked when the significance level is very low.

Sample sizes should be large enough to find a real effect.

Evidence of non-related variables should come from multiple studies.

Choosing a Sample Size: Power Analysis

Power of a statistical test: Determines optimal sample size based on probability of correctly rejecting the null hypothesis

$$\text{Power} = 1 - p \text{ (Type II error)}$$

Effect sizes and desired power

- Smaller effect sizes require larger samples to be significant at the .05 level.
- Higher desired power demands a greater sample size.
- Researchers usually use a power between .70 and .90 to determine sample size.

The Importance of Replications

Scientists attach little importance to the results of a single study.

Detailed understanding requires numerous studies examining the same variables.

Researchers look at the results of studies that replicate previous investigations.

Significance of a Pearson Correlation Coefficient

The Pearson *r* correlation coefficient is used to describe the strength of the relationship between two variables when both variables have interval or ratio scale properties.

A statistical significance test helps to:

- Decide the rejection of a null hypothesis
- Conclude that the true population correlation is greater than 0.00

Computer Analysis of Data

Statistical analysis software packages, such as the following, make it easy to calculate statistics for any data set:

- SPSS
- SAS
- SYSTAT
- R
- Microsoft Excel

Steps in Computer Analysis

Input data into rows and columns.

- Rows represent cases or each participant's data.
- Columns contain a participant's score for a specific variable.

Provide instructions for the statistical analysis.

Interpret the output.

Selecting the Appropriate Significance Test

Variables maybe have a variety of scale properties:

- Nominal
- Ordinal
- Interval
- Ratio

Nominal scale properties have two or more discrete values.

Interval or ratio scale properties have many values.

Research Studying Two Variables

Independent Variable (IV)	Dependent Variable (DV)	Statistical test
Nominal *Right-handed; left-handed*	Nominal *Vegetarian—yes/no*	Chi-square
Nominal (2 groups) *Right-handed; left-handed*	Interval/ratio *Grade point average*	t test
Nominal (3 groups) *Study time (low, medium, high)*	Interval/ratio *Test score*	One-way analysis of variance
Interval/ratio *Optimism score*	Interval/ratio *Sick days last year*	Pearson correlation

Research with Multiple Independent Variables

IV	DV	Statistical test
Nominal (2 or more variables)	Interval/ratio	Analysis of variance (factorial design)
Interval/ratio (2 or more variables)	Interval/ratio	Multiple regression