

PSY104  
ANALYSIS OF  
VARIANCE (ANOVA)

Statistics in Social Sciences

# Comparison of More Than Two Groups

- As we saw in the previous lesson, the method we used to compare two groups was **the t-test**.
- So, if our study has three groups (groups A, B, and C), how do we compare these three groups?
- What if we perform three t-tests? 1,2,3, and so on...
- Can we compare A with B, then A with C, and then B with C??? By using the t-test with an  $\alpha=0.05$ ?
- The answer is YES :)
- But...

# Type 1 Error & Performing Multiple T-tests

- Let's say that when performing a t-test, our alpha criterion is 0.05, meaning we have a 95% probability and a 5% chance of making a Type I error.
- If we perform 3 t-tests and take  $\alpha=0.05$ , we have a  $0.95 \times 0.95 \times 0.95 = .857$  chance of not making a Type I error.
- Our probability of making a Type I error is  $1 - .857 = 0.143$ , or **14.3%**. This value, which was previously 5%, has increased to 14.3%, so performing three separate t-tests would not be the correct approach.
- If we had compared 10 groups using t-tests,  $1 - (0.95)^{10} = 0.4012$ , meaning we would have a 40% probability of making a Type I error.

# ANOVA / F-test

- When applying the t-test, we were testing whether the means of the two groups were equal.
- If the p-value was greater than alpha (0.05), we could not reject the null hypothesis (that the means of the two groups are equal).
- If the p-value was less than alpha, we rejected the null hypothesis and accepted the alternative hypothesis (that the means of the two groups are different).
- Similarly, in ANOVA, we test the null hypothesis that the means of more than two groups (e.g., 3 groups;  $\text{GROUP1}=\text{GROUP2}=\text{GROUP3}$ ) are equal.
  - That is, if the p-value is greater than 0.05, we accept the null hypothesis; if the p-value is less than 0.05, we reject the null hypothesis and say that there is a difference between the three groups.
- While the t-value is used in the t-test to decide whether the p-value is significant, the **F-value/ratio** obtained from the **F-test** is used in ANOVA.

# What is F Test?

- The F test is the statistical test used in ANOVA to determine whether there are significant differences among group means.
- It does not compare means directly. Instead, it compares **variances**.
- More specifically, it compares:
  - Variability **between groups**- grup ortalamalarının genel ortalamadan ne kadar uzaklaştığı
  - Variability **within groups**- aynı gruptaki bireylerin ortalamadan ne kadar uzaklaştığı
- Ortalamalar birbirine yakınsa- Varyans (?)
- Ortalamalar birbirinden uzaksa- Varyans (?)

$$F = \frac{\text{Gruplar Arası Varyans}}{\text{Grup İçi Varyans}}$$

# Assumptions of ANOVA

1. **Level of Measurement (Ölçüm Düzeyi):**The dependent variable should be measured at the **interval** or **ratio** level (continuous). *Bağımlı değişken aralık veya oran düzeyinde (sürekli) olmalıdır.* (Örn: Sınav puanı, boy uzunluğu vb.). **Bağımsız değişken ise kategorik (en az 3 grup) olmalıdır.**
2. **Independence of Observations (Gözlemlerin Bağımsızlığı):** There is no relationship between the observations in each group or between the groups themselves. *Her bir gruptaki gözlemler birbirinden ve diğer gruplardan tamamen bağımsız olmalıdır.* –Random assignment!!!
3. **Normality (Normallik):** The dependent variable should be normally distributed for each category of the independent variable. *Bağımlı değişken, bağımsız değişkenin her bir grubu için normal dağılım göstermelidir.*– Checking Normality: Skewness & Kurtosis, Shapiro-Wilk test, Kolmogorov-Smirnov test.
4. **Homogeneity of Variances (Varyansların Homojenliği):**The variance of data in each group should be approximately equal. *Gruplardaki verilerin varyansları (dağılımları) birbirine eşit (benzer) olmalıdır.*
  - Test: Use Levene's Test.  $p > 0.05$  olmalı!! If violated ( $p < 0.05$ ): Use Welch ANOVA or Brown-Forsythe test.

# Assumption of Homogeneity of Variance

- As in the t-test, ANOVA also assumes that the variances of the groups are homogeneous, meaning they are close to each other.
- As in the t-test, we can check the homogeneity of variance here using Levene's test.
- **BE CAREFUL!!:** In Levene's test, our null hypothesis ( $H_0$ ) is that the variances of each group are equal. If the p-value is greater than 0.05, we cannot reject this hypothesis and can say that variance homogeneity is satisfied.
- If the p-value obtained from Levene's test is less than 0.05, we reject this hypothesis and say that the assumption of variance homogeneity is not satisfied.
- In this case, we cannot use the F and p values given in the ANOVA table. Instead, we use the F value obtained through the Brown–Forsythe F (1974) and Welch's F (1951) statistics provided in SPSS and the corresponding p value.

## Descriptives

Plant Growth								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Organic	48	3.3750	2.43752	.35183	2.6672	4.0828	1.00	7.00
Inorganic	48	3.7083	2.47505	.35724	2.9897	4.4270	2.00	8.00
Mixed	48	4.4167	2.71194	.39143	3.6292	5.2041	2.00	9.00
Total	144	3.8333	2.56387	.21366	3.4110	4.2557	1.00	9.00

## Test of Homogeneity of Variances

		Levene Statistic	df1	df2	Sig.
Plant Growth	Based on Mean	.414	2	141	.662
	Based on Median	.196	2	141	.822
	Based on Median and with adjusted df	.196	2	138.137	.822
	Based on trimmed mean	.400	2	141	.671

## ANOVA

Plant Growth					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	27.167	2	13.583	2.098	.126
Within Groups	912.833	141	6.474		
Total	940.000	143			

# Assumption of Normality

- However, it is important to note that the dependent variable must satisfy the normality condition for each group.
- Normality can be checked visually as well as numerically using numerical data. Additionally, two tests have been developed to test for normality.
- Visually, using histograms and P-P plots,
- Numerically, using skewness and kurtosis values,
- Using tests such as the Kolmogorov–Smirnov test and Shapiro–Wilk tests.
- Normality and outlier checks can be performed using the
- **SPSS>Analyze>Descriptive Stat.>Explore**

- <https://www.youtube.com/watch?v=j3m-05WpnHw>

# Violation of Normality

- As with the t-test, it is possible to follow a ranking.
- Although ANOVA is resistant to non-normal data, results may be biased when normality is not ensured.
- When normality is not ensured: increase the sample size, transform the data and perform ANOVA on the transformed data, or use non-parametric tests (Kruskal Wallis).

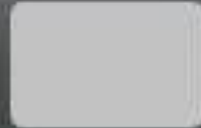
# One-Way ANOVA

- One-way ANOVA is a [statistical procedure](#) used to compare the variance in the group means within a sample, considering **only one independent variable**.
- One-way ANOVA is applied when a study involves a single independent variable with two or more groups and a continuous dependent variable.
- The primary goal of conducting a one-way ANOVA is to compare three or more categorical groups to determine if there is a difference between them.
- Example: There is **one independent variable (factor)**.
- The independent variable has **three or more levels**.
- The dependent variable is **continuous** (interval or ratio scale).

**Example:** Suppose we want to examine whether teaching method affects exam performance.

- Independent variable: Teaching Methods (Traditional/Online/Hybrid)
- Dependent variable: Exam score
- If the F-test is significant, we **reject** the null hypothesis and conclude that not all means are equal.
- However, we must conduct **post-hoc tests** (such as Tukey or Bonferroni) to determine which specific groups differ.

## One-Way ANOVA



**Compare means of independent groups**



**Preferable when there are three or more groups of data**



**One independent variable and dependent variable**



**The independent variable should have three or more levels**



**Uses F-distribution for calculation**



# Two-Way ANOVA

- Two-way ANOVA is a method used to **compare the differences between the means of groups categorized based on two independent** variables.
- Two-way ANOVA is conducted to understand the effects of the independent variables on the dependent variables, how two-factor variables affect the dependent variable.

## When do we use Two-Way ANOVA?

Two-Way ANOVA is used when:

- There are **two independent variables**.
- The dependent variable is **continuous**.
- To examine both main effects and interaction effects.

# Two-Way ANOVA

- Two-way ANOVA is a method used to **compare the differences between the means of groups categorized based on two independent** variables.
- Two-way ANOVA is conducted to understand the effects of the independent variables on the dependent variables, how two-factor variables affect the dependent variable.
- Independent variables: Teaching method (Traditional, Online, Hybrid) & Gender (Male, Female)
- Dependent variable: Exam score
- **What does Two-Way ANOVA test?**

Two-Way ANOVA examines three effects:

- The main effect of the first independent variable (e.g., teaching method).
- The main effect of the second independent variable (e.g., gender).
- The **interaction effect** between the two independent variables.

# Two-Way ANOVA

- **What does Two-Way ANOVA test?**

Two-Way ANOVA examines three effects:

- The main effect of the first independent variable (e.g., teaching method).
- The main effect of the second independent variable (e.g., gender).
- The **interaction effect** between the two independent variables.

## **What is an interaction effect?**

- An interaction occurs when the effect of one independent variable depends on the level of the other independent variable.
- For example: If the online method works better for females but not for males, then there is an interaction between teaching method and gender.
- Interaction effects are important because they show conditional relationships rather than simple average differences. – it means that **it changes depending on context.**

# ANOVA

- The F value obtained from an ANOVA tells us whether there is a difference between groups that can be attributed to the experimental effect.
- If the F value is statistically significant, the first conclusion we draw is that there is a significant difference among the three groups.
- It could be that the means of all three groups differ from one another.
- Another scenario could be that the means of Group 1 and Group 2 are equal, but Group 3 has a different mean from both of them. There are many possible patterns like these that could lead to a significant F value.
- The F value only indicates that there is a significant difference among the groups; it does not provide any additional information about where the difference lies.
- **GRUPLAR ARASI FARKLARI NASIL ANLAYACAĞIZ? ☺**

# ANOVA

- GRUPLAR ARASI FARKLARI NASIL ANLAYACAĞIZ? ☺

**The answer is - POST-HOC TESTS!!!**

- All available groups are compared with each other in a pairwise manner.
- This method uses a stricter decision criterion to ensure that the familywise error rate does not exceed .05.
- Dependent V. : Depression (Beck Depression Inventory Score- BDI Score), Independent V. : Medicine Administered (3 Level)

Dependent Variable: BDI Score  
Tukey HSD

(I) Medicine administered	(J) Medicine administered	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
None	Placebo	6.680*	1.891	.003	1.74	11.62
	Homeopathic	6.200*	1.891	.008	1.26	11.14
	Pharmaceutical	14.840*	1.891	<.001	9.90	19.78
Placebo	None	-6.680*	1.891	.003		
	Homeopathic			.994		
	Pharmaceutical	9.160*	1.891	<.001		

**P > .05? MEAN DIFFERENCE NOT "SIGNIFICANT"**

\* The mean difference is significant at the 0.05 level.

# ANOVA

- Most important - the **sample sizes** because these affect which assumptions we'll need for our ANOVA.
- **Eta-squared** is an effect size measure: it is a single, standardized number that expresses *how different* several sample means are (that is, how far they lie apart). – *Bir fark var ama bu fark ne kadar büyük?» sorusuna cevap verir.*

BDI Score

Medicine administered	N	Min	Max	Mean	SD	Skewness	Kurtosis
None	25	<b>EQUAL SAMPLE SIZES</b>		41.72	6.328	-.364	-.407
Placebo	25			35.04	7.056	.413	.085
Homeopathic	25			35.52	6.634	-.738	.431
Pharmaceutical	25			26.88	6.704		
Total	100	15	53	34.79	8.451		

ANOVA Effect Sizes<sup>a</sup>

		Point Estimate	95% Confidence Interval	
			Lower	Upper
BDI Score	Eta-squared → " $\eta^2 = 0.39$ " ←	.393	.229	.501
	Epsilon-squared	.374	.205	.485
	Omega-squared Fixed-effect	.372	.204	.483
	Omega-squared Random-effect	.165	.079	.237

<sup>a</sup> Eta-squared and Epsilon-squared are estimated based on the fixed-effect model.

# ANOVA

- **Eta-squared** is an effect size measure: it is a single, standardized number that expresses *how different* several sample means are (that is, how far they lie apart). – *Bir fark var ama bu fark ne kadar büyük?» sorusuna cevap verir.*
- $\eta^2 = 0.01$  indicates a **small** effect;
- $\eta^2 = 0.06$  indicates a **medium** effect;
- $\eta^2 = 0.14$  indicates a **large** effect.
- IF,  $\eta^2 = 0.39$  is a huge effect: our 3 medicines resulted in dramatically different mean depression scores.

ANOVA Effect Sizes<sup>a</sup>

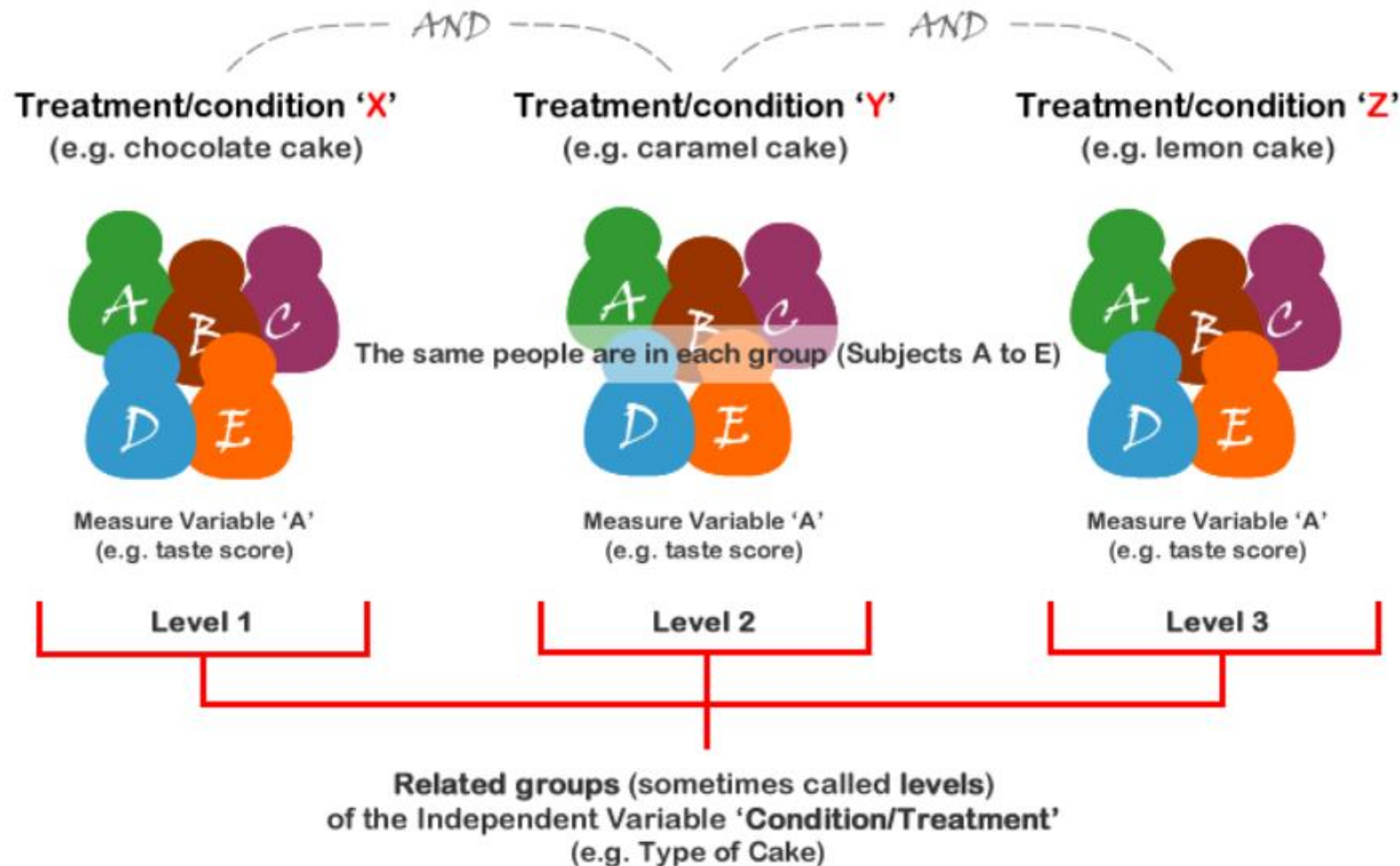
		Point Estimate	95% Confidence Interval	
			Lower	Upper
BDI Score	Eta-squared → “ $\eta^2 = 0.39$ ” ←	.393	.229	.501
	Epsilon-squared	.374	.205	.485
	Omega-squared Fixed-effect	.372	.204	.483
	Omega-squared Random-effect	.165	.079	.237

a. Eta-squared and Epsilon-squared are estimated based on the fixed-effect model.

# Repeated Measures ANOVA

- The repeated measures analysis of variance (ANOVA) is used to test the difference in our dependent variable between three or more groups of observations in which all participants participate in all groups or levels.
- Our grouping variable (3 or more) is our independent variable.
- In other words, **we use the repeated measures ANOVA when we have a research question with a continuous dependent variable and a categorical independent variable with three or more categories in which **the same participants** are in each category.**
- **Example ? ☺**

# Repeated Measures ANOVA



*Any Questions??* 😊

Asst. Prof. Dr. Sena DOĞRUYOL

Statistics in Social Sciences