



One- and Two-way ANOVA

Gadise Bekele, M.D., M.P.H.

Department of Epidemiology and Biostatistics

December, 2022

Outline

- ANOVA(One vs Two Way)
- Assumption
- Test on SPSS
- Interpretation

ANalysis Of VAriance (ANOVA)

- is a set of statistical methods used to assess the mean differences across two or more groups
- One-way ANOVA : where we have a continuous variable paired with an ordinal or nominal variable with more than two categories,
- It splits the total variance into two groups: between variance and within variance.



-
- The between variance measures the variation between groups, whereas the within variance measures the variation within groups.
 - Whenever the between variation is considerably larger than the within variation, we can say that there are differences within groups.
 - The formula for an ANOVA analysis or f-test is **between-group variance/within-group variance**



One-way ANOVA

- What it does: it will tell you whether there are significant differences in the mean scores on the dependent variable across the three or more groups.
- Post-hoc tests can then be used to find out where these differences lie.
- Assumptions: 6 assumptions
- Non-parametric alternative: Kruskal-Wallis Test



■ Two different types of one-way ANOVA:

- **Between-groups ANOVA**, which is used when you have different subjects or cases in each of your groups (this is referred to as an independent groups design); and
- **Repeated-measures analysis of variance**, which is used when you are measuring the same subjects under different conditions (or measured at different points in time) (this is also referred to as a within-subjects design).



Assumptions

- Level of measurement
- Random sampling
- Independence of observations
- Normal distribution
- Homogeneity of variance



Checking Normality

PracticedataANOVA.sav [DataSet2] - IBM SPSS Statistics Data Editor

Age	Sex	ward_admitted
1	58	Male
2	70	Male
3	64	Female
4	55	Male
5	50	Female
6	48	Male
7	70	Male
8	61	Female
9	35	Female
10	28	Female
11	34	Female
12	82	Male
13	48	Male
14	28	Female
15	38	Male
16	19	Female
17	35	Female
18	27	Female
19	30	Female
20	51	Female
21	27	Male
22	35	Female
23	43	Male

Explore: Plots

Boxplots: Factor levels together, Dependents together, None

Descriptive: Stem-and-leaf, Histogram

Normality plots with tests

Spread vs Level with Levene Test: None, Power estimation, Transformed Power: Natural log, Untransformed

Display: Both

Continue Cancel Help

*Output1 [Document1] - IBM SPSS Statistics Viewer

95% Confidence Interval for Mean	Lower Bound	Upper Bound	48.78	52.99
5% Trimmed Mean			50.92	
Median			52.00	
Variance			297.289	
Std. Deviation			17.242	
Minimum			18	
Maximum			90	
Range			72	
Interquartile Range			30	
Skewness			-.093	.151
Kurtosis			-1.048	.301

Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
_102_Age_in_Year	.083	260	.000	.966	260	.000

a. Lilliefors Significance Correction

SPSS Steps to run One-Way ANOVA

- Step 1: Go to Analyze → Compare Means → One-Way ANOVA.
- Step 2: Put your continuous variable (age) as **Dependent Variable** and your categorical variable (ward_admitted) as **Factor**
- Step 3: Click on the **option** box and mark the Descriptive and Homogeneity of Variance test options → click continue → click on **Post Hoc** box and select a test which can give you the result of comparison between the groups. Since we run ANOVA when the homogeneity of variance assumption is met, select a test under Equal variance assumed (Eg. LSD) → Click ok → check the output





6: Visible: 3 of 3 Variables

	Age	Sex	ward_admitted	var											
1	58	Male	Obstetrics/Gyne...												
2	70	Male	Medical Ward												
3	64	Female	Medical Ward												
4	55	Male	Medical Ward												
5	50	Female	Surgical Ward												
6	48	Male	Medical Ward												
7	70	Male	Medical Ward												
8	61	Female	Medical Ward												
9	35	Female	Obstetrics/Gyne...												
10	28	Female	Obstetrics/Gyne...												
11	34	Female	Obstetrics/Gyne...												
12	82	Male	Medical Ward												
13	48	Male	Medical Ward												
14	28	Female	Obstetrics/Gyne...												
15	38	Male	Surgical Ward												
16	19	Female	Obstetrics/Gyne...												
17	35	Female	Obstetrics/Gyne...												
18	27	Female	Obstetrics/Gyne...												
19	30	Female	Obstetrics/Gyne...												
20	51	Female	Surgical Ward												
21	27	Male	Surgical Ward												
22	35	Female	Surgical Ward												
23	43	Male	Surgical Ward												

One-Way ANOVA

Dependent List:

Factor:

Buttons: Contrasts..., Post Hoc..., Options..., Bootstrap..., OK, Paste, Reset, Cancel, Help





Visible: 3 of 3 Variables

	Age	Sex	ward_admitted	var											
1	58	1	3												
2	70	1	1												
3	64	2	1												
4	55	1	1												
5	50	2	2												
6	48	1	1												
7	70	1	1												
8	61	2	1												
9	35	2	3												
10	28	2	3												
11	34	2	3												
12	82	1	1												
13	48	1	1												
14	28	2	3												
15	38	1	2												
16	19	2	3												
17	35	2	3												
18	27	2	3												
19	30	2	3												
20	51	2	2												
21	27	1	2												
22	35	2	2												
23	43	1	2												

One-Way ANOVA: Options

Statistics

- Descriptive
- Fixed and random effects
- Homogeneity of variance test
- Brown-Forsythe
- Welch
- Means plot

Missing Values

- Exclude cases analysis by analysis
- Exclude cases listwise

Continue Cancel Help



Visible: 3 of 3 Variables

	Age	Sex	ward_admitted	var											
1	58	1	3												
2	70	1	1												
3	64	2	1												
4	55	1	1												
5	50	2	2												
6	48	1	1												
7	70	1	1												
8	61	2	1												
9	35	2	3												
10	28	2	3												
11	34	2	3												
12	82	1	1												
13	48	1	1												
14	28	2	3												
15	38	1	2												
16	19	2	3												
17	35	2	3												
18	27	2	3												
19	30	2	3												
20	51	2	2												
21	27	1	2												
22	35	2	2												
23	43	1	2												

One-Way ANOVA: Post Hoc Multiple Comparisons

Equal Variances Assumed

- LSD
- Bonferroni
- Sidak
- Scheffe
- R-E-G-W F
- R-E-G-W Q
- S-N-K
- Tukey
- Tukey's-b
- Duncan
- Hochberg's GT2
- Gabriel
- Waller-Duncan
- Type I/Type II Error Ratio: 100
- Dunnett
- Control Category: Last

Test

- 2-sided
- < Control
- > Control

Equal Variances Not Assumed

- Tamhane's T2
- Dunnett's T3
- Games-Howell
- Dunnett's C

Significance level: 0.05

Continue Cancel Help





- Output
 - Log
 - Explore
 - Title
 - Notes
 - Active Dataset
 - Case Processing
 - Descriptives
 - Tests of Normality
 - _102_Age_in_Year
 - Title
 - Normal Q-Q Plot
 - Detrended Normal Q-Q Plot
 - Log
 - Oneway
 - Title
 - Notes
 - Active Dataset
 - Descriptives
 - Test of Homogeneity of Variances
 - ANOVA
 - Post Hoc Tests
 - Title
 - Multiple Comparisons

[DataSet2] C:\Users\User\Desktop\ARCFPNC\Downloads\PracticedataANOVA.sav

Descriptives

_102_Age_in_Year

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Medical Ward	175	56.65	14.889	1.126	54.42	58.87	18	90
Surgical Ward	53	43.92	16.448	2.259	39.39	48.46	19	80
Obstetrics/Gynecology Ward	32	30.94	10.236	1.809	27.25	34.63	19	62
Total	260	50.89	17.242	1.069	48.78	52.99	18	90

Test of Homogeneity of Variances

_102_Age_in_Year

Levene Statistic	df1	df2	Sig.
5.981	2	257	.003

ANOVA

_102_Age_in_Year

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	21108.158	2	10554.079	48.531	.000
Within Groups	55889.607	257	217.469		
Total	76997.765	259			



Output

- Log
- Explore
 - Title
 - Notes
 - Active Dataset
 - Case Processing
 - Descriptives
 - Tests of Normality
 - _102_Age_in_Year
 - Title
 - Normal Q-Q Plot
 - Detrended Normal Q-Q Plot
- Log
- Oneway
 - Title
 - Notes
 - Active Dataset
 - Descriptives
 - Test of Homogeneity of Variances
 - ANOVA
 - Post Hoc Tests
 - Title
 - Multiple Comparisons

Statistic	df1	df2	Sig.
	5.981	2	.003

ANOVA

_102_Age_in_Year

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	21108.158	2	10554.079	48.531	.000
Within Groups	55889.607	257	217.469		
Total	76997.765	259			

Post Hoc Tests

Multiple Comparisons

Dependent Variable: _102_Age_in_Year
LSD

(I)	(J)	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Medical Ward	Surgical Ward	12.721*	2.312	.000	8.17	17.27
	Obstetrics/Gynecology Ward	25.708*	2.835	.000	20.12	31.29
Surgical Ward	Medical Ward	-12.721*	2.312	.000	-17.27	-8.17
	Obstetrics/Gynecology Ward	12.987*	3.301	.000	6.49	19.49
Obstetrics/Gynecology Ward	Medical Ward	-25.708*	2.835	.000	-31.29	-20.12
	Surgical Ward	-12.987*	3.301	.000	-19.49	-6.49

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

- Check the **Test of homogeneity of variance table** for the p-value of Leven's statistics(if p-value >0.05 , then assumption is met)
- Check the **ANOVA** table for the p-value of the independent group ANOVA test(If p-value < 0.05 , it means there is statically significant difference in the mean of the three or more groups we are comparing). If there is significant difference between the groups, then check the Post Hoc table to identify where the difference is.
- Check the **Post Hoc Tests** table to identify between which variables there is statistically significant difference
- Check the **Descriptive table** to get the mean with CI values of the variables. Using these values we assess which variables has a higher or lower mean values



Interpretation

- There is a statistically significant difference in the mean age of patients between the three inpatient wards. Being admitted to the medical ward is associated with a relatively older age(56.65), followed by surgical ward(43.92) and then Oby/gyn ward(30.94)



Two-way between groups ANOVA

- Looks at the individual and joint effect of two independent variables on one dependent variable
- The advantage of using a two-way design is that we can test the 'main effect' for each independent variable and also explore the possibility of an 'interaction effect'.
- An interaction effect occurs when the effect of one independent variable on the dependent variable depends on the level of a second independent variable.



-
- What you need: Three variables;
 - two categorical independent variables and
 - one continuous dependent variable

 - What it does: it allows you to simultaneously test for the effect of each of your independent variables on the dependent variable and also identifies any interaction effect.



Procedure for two-way ANOVA

- 1. From the menu at the top of the screen, click on Analyze, then click on General Linear Model, then on univariate
- 2. Click on your dependent, continuous variable and move these into the box labelled **Dependent variable**
- 3. Click on your two independent, categorical variables and move these into the box labelled **Fixed Factors**
- 4. Click on the **Option** button.
 - Click on Descriptive Statistics, Estimates of effect size and Homogeneity tests
 - Click on Continue
- 5. Click on the **Post Hoc** button
 - From the Factors listed on the left-hand side, choose the independent variables(s) you are interested in (these variables should have three or more levels or groups)
 - Click on the arrow button to move it into the Post Hoc Tests for section
 - Choose the test you wish to use
 - Click on Continue



■ 6. Click on the **Plots** button

- In the Horizontal box, put the independent variable that has the most groups
- In the box labelled Separate Lines, put the other independent variable
- Click on **Add**
- In the section labelled Plots, you should now see your two variables listed

■ 7. Click on Continue and then Ok



	Age	Sex	ward_admitted	Length_of_stay2	var								
1	58	1	3	1									
2	70	1											
3	64	2											
4	55	1											
5	50	2											
6	48	1											
7	70	1											
8	61	2											
9	35	2											
10	28	2											
11	34	2											
12	82	1											
13	48	1											
14	28	2											
15	38	1											
16	19	2											
17	35	2											
18	27	2											
19	30	2											
20	51	2	2	1									
21	27	1	2	2									
22	35	2	2	3									
23	43	1	2	1									

Univariate

Dependent Variable:

Fixed Factor(s):

Random Factor(s):

Covariate(s):

WLS Weight:

Model...
Contrasts...
Plots...
Post_Hoc...
Save...
Options...
Bootstrap...

OK Paste Reset Cancel Help





1 : Length_of_stay2 1

Visible: 4 of 4 Variables

	Age	Sex	var									
1	58	1										
2	70	1										
3	64	2										
4	55	1										
5	50	2										
6	48	1										
7	70	1										
8	61	2										
9	35	2										
10	28	2										
11	34	2										
12	82	1										
13	48	1										
14	28	2										
15	38	1										
16	19	2										
17	35	2										
18	27	2										
19	30	2										
20	51	2										
21	27	1										
22	35	2	2				3					
23	43	1	2				1					

Univariate: Options

Estimated Marginal Means

Factor(s) and Factor Interactions:
 (OVERALL)
 Sex
 Length_of_stay2
 Sex*Length_of_stay2

Display Means for:
 (OVERALL)
 Sex
 Sex*Length_of_stay2
 Length_of_stay2

Compare main effects

Confidence interval adjustment:
 LSD(none)

Display

Descriptive statistics
 Estimates of effect size
 Observed power
 Parameter estimates
 Contrast coefficient matrix

Homogeneity tests
 Spread vs. level plot
 Residual plot
 Lack of fit
 General estimable function

Significance level: .05 Confidence intervals are 95.0 %

Continue Cancel Help

Data View Variable View





1: Length_of_stay2 1 Visible: 4 of 4 Variables

	Age	Sex	ward_admitted	Length_of_stay2	var									
1	58	1												
2	70	1												
3	64	2												
4	55	1												
5	50	2												
6	48	1												
7	70	1												
8	61	2												
9	35	2												
10	28	2												
11	34	2												
12	82	1												
13	48	1												
14	28	2												
15	38	1												
16	19	2												
17	35	2												
18	27	2												
19	30	2												
20	51	2												
21	27	1	2	2										
22	35	2	2	3										
23	43	1	2	1										

Univariate: Post Hoc Multiple Comparisons for Observed Means

Factor(s):
Sex
Length_of_stay2

Post Hoc Tests for:
Length_of_stay2

Equal Variances Assumed

- LSD
- S-N-K
- Waller-Duncan
- Bonferroni
- Tukey
- Type I/Type II Error Ratio: 100
- Sidak
- Tukey's-b
- Dunnett
- Scheffe
- Duncan
- Control Category: Last
- R-E-G-W-F
- Hochberg's GT2
- Test: 2-sided < Control > Control
- R-E-G-W-Q
- Gabriel

Equal Variances Not Assumed

- Tamhane's T2
- Dunnett's T3
- Games-Howell
- Dunnett's C

Continue Cancel Help

Data View Variable View





1 : Length_of_stay2 1 Visible: 4 of 4 Variables

	Age	Sex	ward_admitted	Length_of_stay2	var									
1	58	1	3	1										
2	70	1												
3	64	2												
4	55	1												
5	50	2												
6	48	1												
7	70	1												
8	61	2												
9	35	2												
10	28	2												
11	34	2												
12	82	1												
13	48	1												
14	28	2												
15	38	1												
16	19	2												
17	35	2												
18	27	2												
19	30	2												
20	51	2	2	1										
21	27	1	2	2										
22	35	2	2	3										
23	43	1	2	1										

Univariate

Univariate: Profile Plots

Factors:
Sex
Length_of_stay2

Horizontal Axis:
Length_of_stay2

Separate Lines:
Sex

Separate Plots:

Plots:
Length_of_stay2*Sex

Buttons: Add, Change, Remove, Continue, Cancel, Help, OK, Paste, Reset, Cancel, Help

Data View Variable View



Output

*Output1.spv [Document1] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Direct Marketing Graphs Utilities Add-ons Window Help

Length_of_stay2

Log

Univariate Analysis

Title

Notes

Active Dataset

Between-Subjects

Descriptive Statistics

Levene's Test of Equality of Variances

Tests of Between-Subjects Effects

Estimated Marginal Means

Title

1. Grand Means

2. _101_Sex

Title

Estimate

Pairwise

Univariate

3. _101_Sex

4. Length_of_stay2

Title

Estimate

Pairwise

Univariate

Post Hoc Tests

Title

Length_of_stay2

Title

Multiple Comparisons

Homogeneity of Variance

Title

Profile Plots

Title

Length_of_stay2

Descriptive Statistics

Dependent Variable: _102_Age_in_Years

101 Sex	Length of stay2	Mean	Std. Deviation	N
Male	1-6 days	52.86	16.400	65
	7-13 days	58.08	14.460	65
	>= 14 days	44.14	15.115	7
	Total	54.89	15.742	137
Female	1-6 days	43.43	16.305	74
	7-13 days	55.58	19.332	36
	>= 14 days	38.15	11.539	13
	Total	46.43	17.800	123
Total	1-6 days	47.84	16.960	139
	7-13 days	57.19	16.312	101
	>= 14 days	40.25	12.839	20
	Total	50.89	17.242	260

Levene's Test of Equality of Error Variances^a

Dependent Variable: _102_Age_in_Years

F	df1	df2	Sig.
1.884	5	254	.098

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

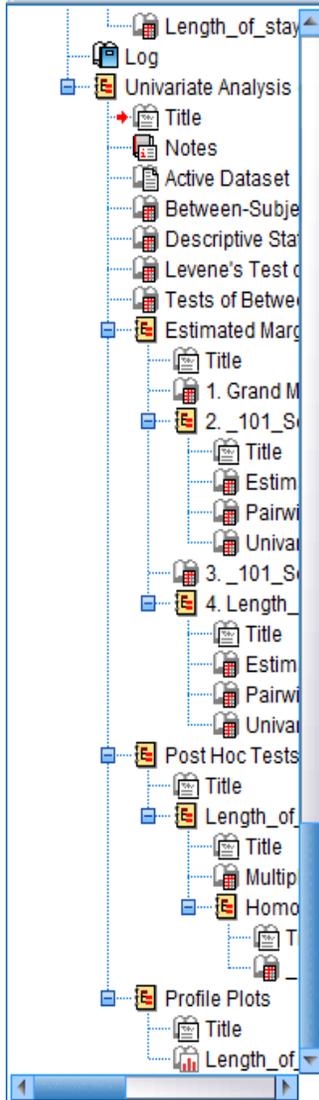
a. Design: Intercept + Sex + Length_of_stay2 + Sex * Length_of_stay2

Open output document

IBM SPSS Statistics Processor is ready

H: 3.95, W: 4.89 in





Levene's Test of Equality of Error Variances^a

Dependent Variable: _102_Age_in_Years

F	df1	df2	Sig.
1.884	5	254	.098

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Sex + Length_of_stay2 + Sex * Length_of_stay2

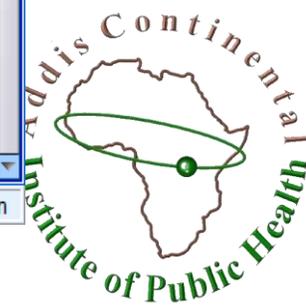
Tests of Between-Subjects Effects

Dependent Variable: _102_Age_in_Years

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	10945.935 ^a	5	2189.187	8.418	.000	.142
Intercept	292661.696	1	292661.696	1125.420	.000	.816
Sex	1099.330	1	1099.330	4.227	.041	.016
Length_of_stay2	6038.463	2	3019.231	11.610	.000	.084
Sex * Length_of_stay2	669.344	2	334.672	1.287	.278	.010
Error	66051.831	254	260.047			
Total	750303.000	260				
Corrected Total	76997.765	259				

a. R Squared = .142 (Adjusted R Squared = .125)

Estimated Marginal Means





- Length_of_stay2
- Log
- Univariate Analysis
 - Title
 - Notes
 - Active Dataset
 - Between-Subject
 - Descriptive Statistics
 - Levene's Test of Homogeneity of Variance
 - Tests of Between-Subjects Equality
 - Estimated Marginal Means
 - Title
 - 1. Grand Means
 - 2. _102_Age_in_Years
 - Title
 - Estimate
 - Pairwise Comparisons
 - Univariate
 - 3. _102_Age_in_Years
 - 4. Length_of_stay2
 - Title
 - Estimate
 - Pairwise Comparisons
 - Univariate
 - Post Hoc Tests
 - Title
 - Length_of_stay2
 - Title
 - Multiple Comparisons
 - Homogeneous Subsets
 - Profile Plots
 - Title
 - Length_of_stay2

The F tests the effect of Length_of_stay2. This test is based on the linearly independent pairwise comparisons among the estimated marginal means.

Post Hoc Tests

Length_of_stay2

Multiple Comparisons

Dependent Variable: _102_Age_in_Years
Tukey HSD

(I) Length of stay2	(J) Length of stay2	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1-6 days	7-13 days	-9.35*	2.108	.000	-14.32	-4.38
	>= 14 days	7.59	3.857	.122	-1.50	16.68
7-13 days	1-6 days	9.35*	2.108	.000	4.38	14.32
	>= 14 days	16.94*	3.947	.000	7.63	26.24
>= 14 days	1-6 days	-7.59	3.857	.122	-16.68	1.50
	7-13 days	-16.94*	3.947	.000	-26.24	-7.63

Based on observed means.
The error term is Mean Square(Error) = 260.047.
*. The mean difference is significant at the .05 level.

Homogeneous Subsets

_102_Age_in_Years

Tukey HSD^{a,b,c}

Length_of_stay2	N	Subset	
		1	2



INTERPRETATION

- **Descriptive statistics:** These provide the Mean scores, Std deviations and N for each subgroup.
- **Levene's Test of Equality of Error Variances:** This test provides a test of one of the assumptions underlying analysis of variance. The value you are most interested in is the Sig. level. You want this to be greater than .05 and therefore not significant.

- **Tests of Between Subjects Effects**

- Interaction effect
- Main effect
- Effect size

Size	Eta squared (% of variance explained)	Cohen's d (standard deviation units)
Small	.01 or 1%	.2
Medium	.06 or 6%	.5
Large	.138 or 13.8%	.8

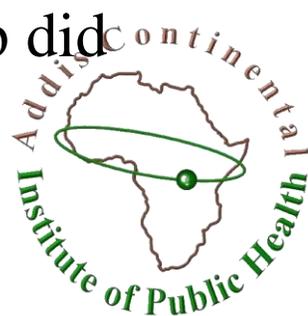
- **Post-hoc tests**

- Multiple comparisons



PRESENTING THE RESULTS FROM TWO-WAY ANOVA

- A two-way between-groups analysis of variance was conducted to explore impact of length of stay and sex on mean age of patients. Subjects were divided into three groups according to their length of stay (Group 1: 1-6 days; Group 2: 7-13 days; Group 3: 14 days and above). The interaction effect between sex and length of stay group was not statistically significant, $P = 0.29$. There was a statistically significant main effect for both sex ($P = 0.04$) and length of stay ($P = 0.0001$). The effect size was small (partial eta squared = .02) for sex and medium (partial eta squared = .08) for length of stay.
- Post-hoc comparisons using the Tukey HSD test indicated that the mean age for the 7-13 days group ($M = 57.19$, $SD = 16.31$) was significantly different from the 1-6 days group ($M = 47.84$, $SD = 16.96$) and ≥ 14 days group ($M = 40.25$, $SD = 12.84$). The mean age of 1-6 days group and ≥ 14 days group did not differ significantly.



▣ Thank You.

