

community project

encouraging academics to share statistics support resources

All stcp resources are released under a Creative Commons licence

stcp-gilchristsamuels-9

The following resources are associated:

One-Way ANOVA – Additional Material worksheet and Normality Testing worksheet

One-Way Analysis of Variance (ANOVA)

Research question type: Differences between several groups of measurements

What kind of variables: Continuous (scale/interval/ratio)

Common Applications: Comparing the means of different groups in scientific or medical experiments when treatments, processes, materials or products are being compared

Example: Grocery Bags¹

A paper manufacturer makes grocery bags. They are interested in increasing the tensile strength of their product. It is thought that strength is a function of the hardwood concentration in the pulp.

An investigation is carried out to compare four levels of hardwood concentration: 5%, 10%, 15% and 20%. Six test specimens were made at each level and all 24 specimens were then tested in random order. The results are shown below:

Hardwood concentration (%)	Tensile strength (PSI)						Mean	Standard deviation	Median
	7	8	15	11	9	10			
5	7	8	15	11	9	10	10.00	2.83	9.50
10	12	17	13	18	19	15	15.67	2.81	16.00
15	14	18	19	17	16	18	17.00	1.79	17.50
20	19	25	22	23	18	20	21.17	2.64	21.00
All							15.96	4.72	

Research question: Are there differences in mean tensile strengths between the different hardwood concentrations?

1. Source: Montgomery, D. and Runger, G. (2011) *Applied Statistics and Probability for Engineers*, 5th ed., Hoboken, NJ: Wiley.



In the analysis of variance we compare the variability **between** the groups (how far the means are apart) to the variability **within** the groups (how much natural variation there is in the measurements). This is why it is called analysis of variance, often abbreviated to **ANOVA**.

Hypotheses

The null hypothesis is:

H₀: There is **no difference in mean** tensile strength between the four hardwood concentrations

The alternative hypothesis is:

H₁: There **is a difference in mean** tensile strength between the four hardwood concentrations

Steps in SPSS

Create two numerical variables called *Concentration* and *Strength*. In the example on the right, *Concentration* has codes 1 for 5% hardwood concentration, 2 for 10%, etc. and *Strength* is the tensile strengths in PSI. These codes can be explained using the Values field in the Variable View.

	Concentration	Strength
1	1	7
2	1	8
3	1	15
4	1	11
5	1	9
6	1	10
7	2	12
8	2	17
9	2	13
10	2	18
11	2	19
12	2	15
13	3	14
14	3	18

Testing ANOVA assumptions

ANOVA has three assumptions:

1. The observations are random samples from normal distributions.
2. The residuals for the whole data set are normally distributed. This theoretically follows from the first assumption but it is worth testing separately with small samples.
3. The groups have equal variances.

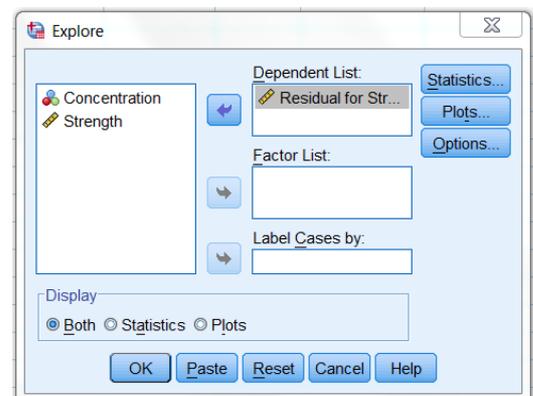
To test the first assumption, select Analyze – Descriptive Statistics – Explore..., select *Strength* on the Dependent List and *Concentration* on the Factor List. Under Plots... select Normality plots with tests. The Shapiro-Wilk normality test for each group is negative ($p > 0.05$, see right), indicating that we may assume that the data is normally distributed (the Kolmogorov-Smirnov test should not be used for small sample sizes).

		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Concentration	Statistic	df	Sig.	Statistic	df	Sig.
Strength	5%	.195	6	.200 [*]	.922	6	.523
	10%	.183	6	.200 [*]	.940	6	.660
	15%	.212	6	.200 [*]	.933	6	.607
	20%	.171	6	.200 [*]	.966	6	.863

a. Lilliefors Significance Correction
*. This is a lower bound of the true significance.

Note: For more thorough normality checking, use histograms and boxplots for large samples and QQ plots for small samples – see the normality testing worksheet.

To test the second assumption we first need to create the residuals:



- Select Analyze – General linear model – Univariate
- Add *Strength* as the Dependent Variable and *Concentration* as the Fixed Factor
- Select Save... and choose Unstandardised residuals

Now we can test the residuals for normality:

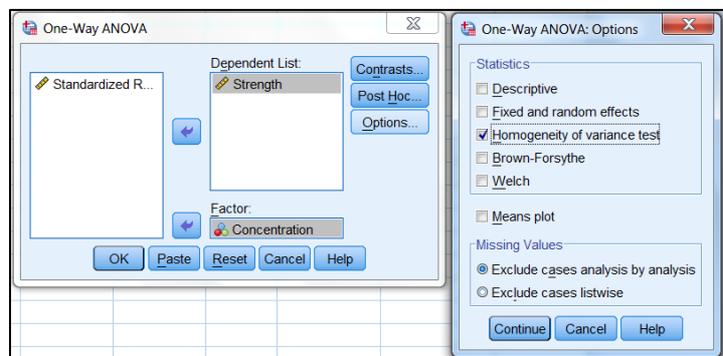
- Select Analyze – Descriptive Statistics – Explore...
- Select *Standardized Residual for Strength* on the Dependent List and leave the Factor List blank
- Under Plots... select Normality plots with tests

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Residual for Strength	.095	24	.200 [*]	.966	24	.576

a. Lilliefors Significance Correction
 *. This is a lower bound of the true significance.

This again gives a non-significant result for the Shapiro-Wilk test (see right), indicating that we may accept the second assumption.

To test the third assumption we need to use a Levene’s test for equality of variance which is contained within the ANOVA analysis in SPSS:



- Select Analyze – Compare Means – One-Way ANOVA
- Select *Strength* on the Dependent List and *Concentration* on the Factor list
- Select Options... then select Homogeneity of variance (Levene's) test in the Statistics list

This produces the output shown on the right. This is not significant at 95% so we may accept assumption 3.

Test of Homogeneity of Variances			
Strength			
Levene Statistic	df1	df2	Sig.
.665	3	20	.583

Note: for robust use of ANOVA see the additional advice sheet.

Running the ANOVA analysis

Now that we have shown that all the assumptions of ANOVA can be accepted we can run the ANOVA analysis. However, the required output table is already provided with the Levene’s test we already carried out.

The ANOVA table gives an F statistic of 19.61 and a p-value of < 0.001 (the value in the *Sig.* column of 0.000 has been rounded to 3 decimal places and should not be quoted in this format). Therefore we reject the null hypothesis with 99.9% confidence (note: subtract the p-value threshold from 1 and multiply by 100 to obtain the confidence level) and conclude that there is very

ANOVA					
Strength					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	382.792	3	127.597	19.605	.000
Within Groups	130.167	20	6.508		
Total	512.958	23			



strong evidence that the mean tensile strengths of the different groups are unequal. However, we do not know which pairs of group mean differences are significantly different, if any. This can be explored by performing **post hoc tests**.

Post hoc tests

Post hoc test are 'after the event' tests used to establish significant differences between pairs of means after an ANOVA has been performed and was found to be statistically significant. In effect they are variations on independent samples t-tests comparing the pairs of means.

The significance level needs to be adjusted in order to reduce the possibility of Type I errors. This leads to many choices of types of post hoc test. Field (2009, p. 375) recommends the following:

- For equal group sizes and similar variances, use Tukey or, for guaranteed control over Type I errors (more conservative), use Bonferroni
- For slightly different group sizes, use Gabriel
- For very different group sizes, use Hochberg's GT2

For example we had equal group sizes so we shall use Tukey and Bonferroni:

- Select Analyze – Compare Means – One-Way ANOVA
- Select *Strength* on the Dependent List and *Concentration* on the Factor list
- Select the Post Hoc... button in the One-way ANOVA Dialog box
- Select Tukey and Bonferroni

The SPSS output indicates statistically significant differences in mean strength between:

- 5% and 10% ($p < 0.01$ for both methods)
- 5% and 15% ($p < 0.001$ for both methods)
- 5% and 20% ($p < 0.001$ for both methods)
- 10% and 20% ($p < 0.01$ for both methods)
- 15% and 20% ($p < 0.05$ for Tukey only)

However, there was no statistically significant difference between 10% and 15% for either method or between 15% and 20% with Bonferroni.

Note: Four groups gives six pairs with each listed twice for both methods

Reference

Field, A. (2009) *Discovering Statistics using SPSS (And sex and drugs and rock 'n' roll)*, 3rd ed., London: SAGE.

