

# One-way ANOVA

---

The **ONE-WAY ANOVA** procedure compares means between two or more groups. It is used to compare the effect of multiple levels (treatments) of a single factor, either discrete or continuous, when there are multiple observations at each level. The null hypothesis is that the means of the measurement variable are the same for the different groups of data.

## Assumptions

The results can be considered reliable if a) observations within each group are independent random samples and approximately normally distributed, b) populations variances are equal and c) the data are continuous. If the assumptions are not met, consider using non-parametric Kruskal-Wallis test.

## How To

- ✓ If observations for each level are in different columns – run the **STATISTICS->ANALYSIS OF VARIANCE (ANOVA)->ONE-WAY ANOVA (UNSTACKED)** command.
- ✓ For stacked data run the **STATISTICS->ANALYSIS OF VARIANCE (ANOVA)->ONE-WAY ANOVA (WITH GROUP VARIABLE)** command, select a **RESPONSE** variable and a **FACTOR** variable. Factor variable is a categorical variable with numeric or text values.
- ✓ LE version includes only **ONE-WAY ANOVA (UNSTACKED, W/O POST-HOC TESTS)** command, and it is similar to the “ANOVA - Single Factor” command from the Analysis Toolpak package for Microsoft Excel and does not include post-hoc comparisons.

## Data Layout

The data for one-way ANOVA can be arranged in two ways, as shown below.

**Samples for each factor level (group) are in different columns**

High NonOECD	High OECD	Low	Lower middle	Upper middle
3.300	4.400	100.600	17.300	16.400
10.200	4.200	43.900	39.600	26.400
13.800	4.900	89.300	42.700	175.500
6.300	5.600	102.700	28.500	14.600
15.700	8.900	88.700	100.500	36.800
10.700	4.000	42.200	53.200	5.500
4.900	3.900	145.500	109.200	18.400
3.900	3.800	154.600	75.700	7.300
101.500	2.800	83.300	22.900	50.400
10.200	4.600	126.000	17.400	15.300
9.100	4.200	54.800	14.600	12.800
5.400	4.700	70.100	84.100	13.500
8.800	3.700	78.500	24.200	48.800

Run the “**ONE-WAY ANOVA (UNSTACKED)**” command.

**Factor levels are defined by values of the factor variable**

Mortality rate	Income
100.6	Low
16.4	Upper middle
26.4	Upper middle
3.3	High NonOECD
175.5	Upper middle
10.2	High NonOECD
14.6	Upper middle
17.3	Lower middle
4.4	High OECD
4.2	High OECD
36.8	Upper middle
13.8	High NonOECD
6.3	High NonOECD
43.9	Low
48.8	High NonOECD

Run the **ONE-WAY ANOVA (WITH GROUP VARIABLE)** COMMAND.

---

## Results

Report includes analysis of variance summary table and post-hoc comparisons.

### ANALYSIS OF VARIANCE TABLE

The basic idea of ANOVA is to split total variation of the observations into two pieces - the variation within groups (*error variation*) and the variation between groups (*treatment variation*) and then test the significance of these components contribution to the total variation.

**SOURCE OF VARIATION** - the source of variation (term in the model).

**SS (SUM OF SQUARES)** - the sum of squares for the term.

**DF (DEGREES OF FREEDOM)** - the number of the degrees of freedom for the corresponding model term.

$$df_{Total} = N - 1$$

$$df_{Treatment} = \text{Factor levels (treatments)} - 1$$

$$df_{Error} = df_{Total} - df_{Treatment}$$

**MS (MEAN SQUARE)** - the estimate of the variation accounted for by this term.

$$MS = SS/DF$$

**F** - the F-test statistic, under the null hypothesis is distributed as  $F_{df\ treatment, df\ error}$ .

$$F = \frac{MS_{Treatments}}{MS_{Error}}$$

**P-LEVEL** - the significance level of the F-test. If p-level is less than the significance level  $\alpha$  – the null hypothesis is rejected, and we can conclude that not all of the group means are equal.

### POST-HOC ANALYSIS (MULTIPLE COMPARISON PROCEDURES)

While significant F-test can tell us that the group means are not all equal, we do not know exactly which means are significantly different from which other ones. With a comparison procedure we compare the means of each two groups. The **SIGNIFICANT** column values show if means difference is significant at the  $\alpha$  alpha level and we should reject the null hypothesis  $H_0$ .

#### **SCHEFFE CONTRASTS AMONG PAIRS OF MEANS**

Scheffe's test is most popular of the post hoc procedures, the most flexible, and the most conservative. Scheffe test corrects alpha for all pair-wise comparisons of means. The test statistic is defined as

$$q = \sqrt{\frac{(\text{Mean}_i - \text{Mean}_j)^2}{MS_{within} (1/N_i + 1/N_j)}}$$

The test statistic is calculated for each pair of means and the null hypothesis is rejected if  $q$  is greater than the critical value  $q_{crit}$ , as previously defined for the original ANOVA analysis

#### **TUKEY TEST FOR DIFFERENCES BETWEEN MEANS**

Tukey's HSD (honestly significant difference) or Tukey A test is based on a studentized range distribution.

The test statistic is defined as  $q = (Mean_i - Mean_j) \sqrt{\frac{MSE}{N}}$ .

Tukey test requires equal sample sizes per group, but can be adapted to unequal sample sizes as well. The simplest adaptation uses the harmonic mean of group sizes as  $N$ .

#### **TUKEY B OR TUKEY WSD (WHOLLY SIGNIFICANT DIFFERENCE) TEST**

Tukey's B (WSD) test is also based on a studentized range distribution. Alpha for Tukey B test is the average of the Newman-Keuls alpha and the Tukey HSD alpha.

#### **NEWMAN-KEULS TEST**

The Newman-Keuls test is a *stepwise* multiple range test, based on a studentized range distribution. The test statistic is identical to Tukey test statistic but Newman-Keuls test uses different critical values for different pairs of mean comparisons - the greater the rank difference between pairs of means, the greater the critical value. The test is more powerful but less conservative than Tukey's tests.

#### **BONFERRONI TEST FOR DIFFERENCES BETWEEN MEANS**

The Bonferroni test is based on the idea to divide the familywise error rate  $\alpha$  among tests and test each individual hypothesis at the statistical significance level of  $1/n$  times what it would be if only one hypothesis were tested, i.e. at the significance level of  $\alpha/n$ .

#### **FISHER LEAST SIGNIFICANT DIFFERENCE (LSD) TEST**

The Fisher LSD test is based on the idea that if an omnibus test is conducted and is significant, the null hypothesis is incorrect. The test statistic is defined as

$$LSD = t_{\alpha} \sqrt{\frac{2 MSE}{N}},$$

where  $t_{\alpha}$  is the critical value of the t-distribution with the df associated with  $MSE$ , the denominator of the F statistic.

## **References**

Design and Analysis: A Researcher's Handbook. 3rd edition. Geoffrey Keppel. Englewood Cliffs, NJ: Prentice-Hall, 1991.

Experimental Design: Procedures for the Behavioral Sciences – 3<sup>rd</sup> Edition (1995). Roger E. Kirk Pacific Grove, CA: Brooks/Cole, 1995.

Handbook of Parametric and Nonparametric Statistical Procedures (3<sup>rd</sup> ed.). Sheskin, David J.. Boca Raton, FL, 1989.