APPENDIX B - ANALYSIS OF VARIANCE (ANOVA / MANOVA)

Description

ANOVA/MANOVA is a powerful technique for analyzing experimental data involving quantitative measurements.

Purpose

The purpose of analysis of variance is to test differences in means (for groups or variables) for statistical significance. This is accomplished by analyzing the variance, that is, by partitioning the total variance into the component that is due to true random error (i.e., within group Sum of the Squares) and the components that are due to differences between means. These latter variance components are then tested for statistical significance, and, if significant, we reject the null hypothesis of no differences between means, and accept the alternative hypothesis that the means (in the population) are different from each other.

Key Assumptions underlying ANOVA:

- The various effects are additive.
- The "experimental" (residual) errors must be independent of the main effects, interactions, and each other.
- The "experimental" errors must have a common variance.
- The "experimental" errors must be randomly and normally distributed.
- Variances are Homogenous (and co-variances for multivariate designs).

Benefits

- ANOVA is more statistically powerful (i.e., we need fewer observations to find a significant effect) than the simple t test.
- ANOVA allows us to detect interaction effects between variables, and, therefore, to test more complex hypotheses about reality. At the heart of ANOVA is the fact that variances can be divided up and partitioned. For example, when several sources of variation are acting simultaneously on a set of observations, the variance of the observations is the sum of the variances of the independent sources. The variance is computed as the sum of squared deviations from the overall mean, divided by **n-1** (sample size minus one). Thus, given a certain **n**, the variance is a function of the sums of (deviation) squares, or **SS** for short.
- ANOVA is a much more flexible and powerful technique that can be applied to much more complex research issues.

Process Flow

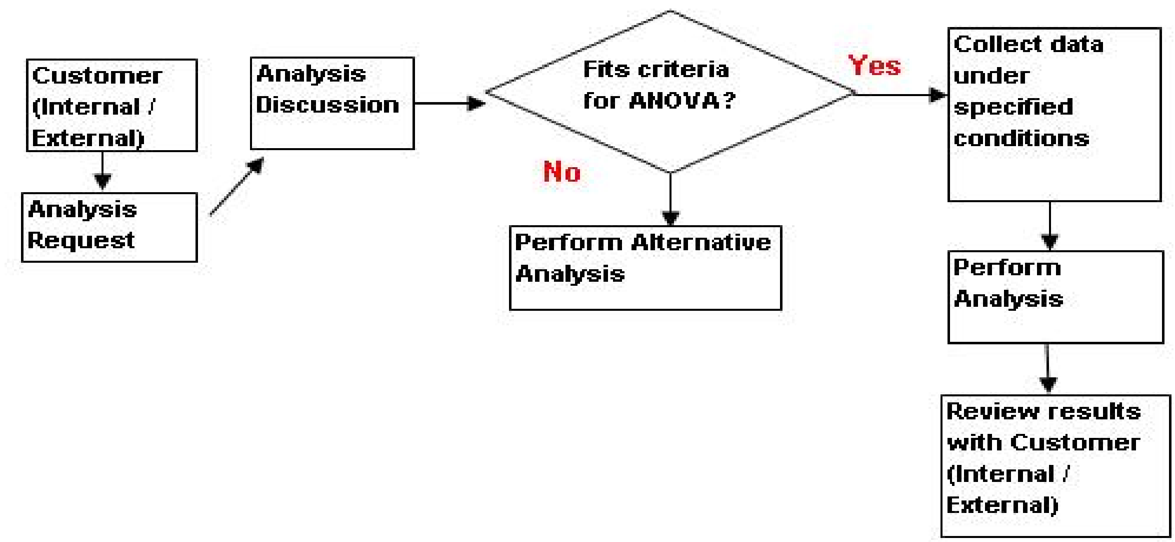
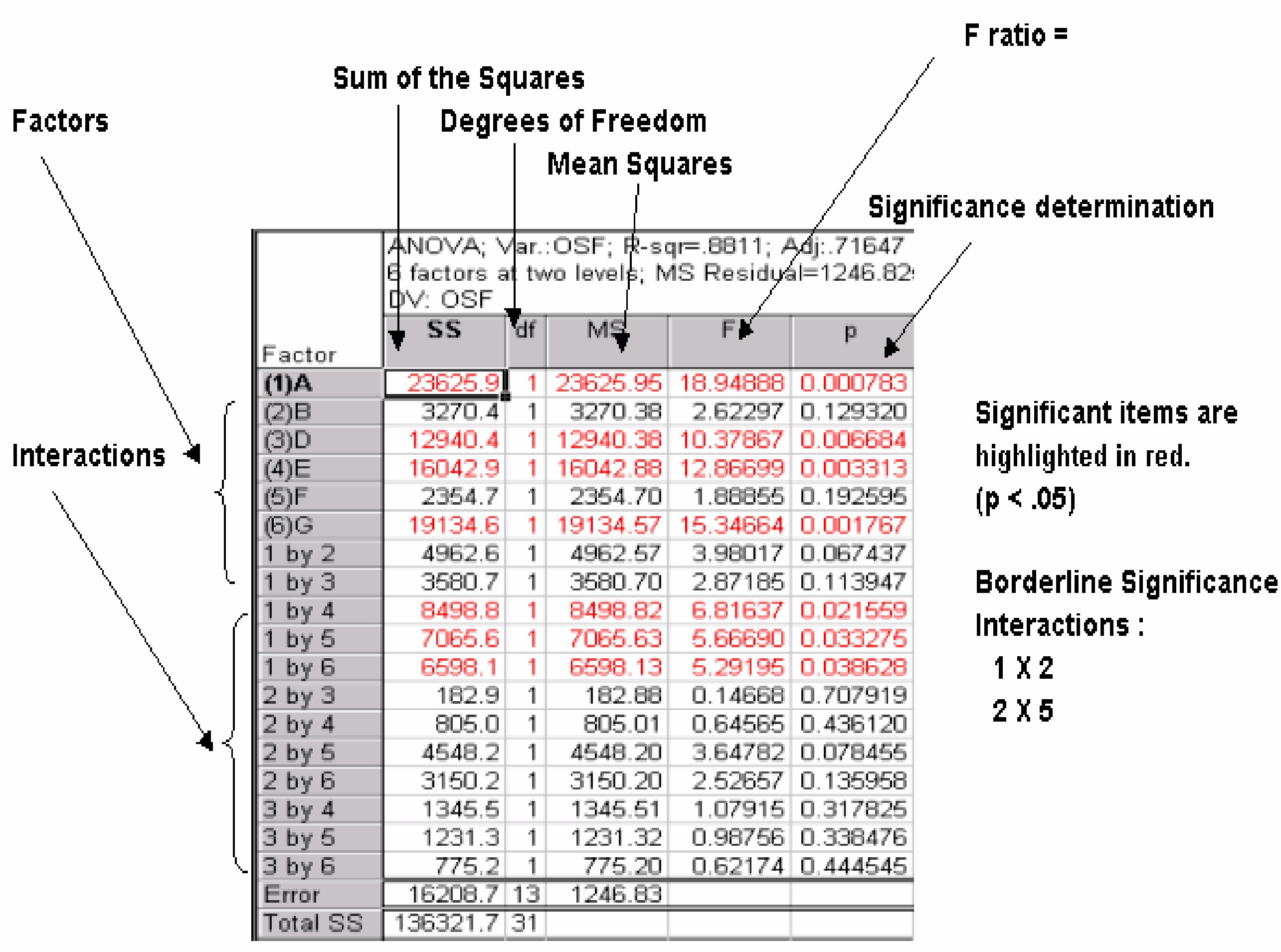


Figure B.1. ANOVA Process Flow

Table 1 below is a 6-factor (2-level) table from a DOE that was conducted. It shows how to interpret the table.



ANOVA; Var.:OSF; R-sqr=.8811; Adj:.71647
6 factors at two levels; MS Residual=1246.82
DV: OSF

Factor	SS	df	MS	F	p
(1)A	23625.9	1	23625.95	18.94888	0.000783
(2)B	3270.4	1	3270.38	2.62297	0.129320
(3)D	12940.4	1	12940.38	10.37867	0.006684
(4)E	16042.9	1	16042.88	12.86699	0.003313
(5)F	2354.7	1	2354.70	1.88855	0.192595
(6)G	19134.6	1	19134.57	15.34664	0.001767
1 by 2	4962.6	1	4962.57	3.98017	0.067437
1 by 3	3580.7	1	3580.70	2.87185	0.113947
1 by 4	8498.8	1	8498.82	6.81637	0.021559
1 by 5	7065.6	1	7065.63	5.66690	0.033275
1 by 6	6598.1	1	6598.13	5.29195	0.038628
2 by 3	182.9	1	182.88	0.14668	0.707919
2 by 4	805.0	1	805.01	0.64565	0.436120
2 by 5	4548.2	1	4548.20	3.64782	0.078455
2 by 6	3150.2	1	3150.20	2.52657	0.135958
3 by 4	1345.5	1	1345.51	1.07915	0.317825
3 by 5	1231.3	1	1231.32	0.98756	0.338476
3 by 6	775.2	1	775.20	0.62174	0.444545
Error	16208.7	13	1246.83		
Total SS	136321.7	31			

Table B.1. Example of ANOVA Results

Factors: A processing variable whose level may change the response variable (output measurement); a method, material, machine, person, environment, or measurement variable changed during the experiment in an attempt to cause change in a response variable; independent variable.

Interaction: If the effect of one factor is different at different levels of another factor, the two factors are said to interact or to have interaction.

Interaction Effect: The effect on the output/dependent response variable caused by the combination of two or more factors, independent of their individual effects.

The Total Sum of the Squares: Total sum of the squares of the deviations from the sum of all the observations divided by the total number of observations (Grand Average) + between-treatment sum of squares divided by the between-treatment degrees of freedom + within-treatment (residual) sum of squares. It is a measure of the overall variation in the measured data.

Degrees of Freedom: Number of tests conducted - number of treatments.

The Mean Square: Individual factor Sum of Squares divided by the Degrees of Freedom.

F Ratio: Mean Square divided by Mean Square Error.

F Test: A means for determining the statistical significance of a factor by comparing calculated **F** values to those contained in an **F** table.

Response: The numerical result of a trial based on a given treatment combination.

Significance Level: The level (usually alpha = .05) of risk that the experimenter is willing to accept for making an incorrect decision. For ANOVA , the derived level of significance is contrasted against the alpha level. If the **p** value is less than alpha, the factor or interaction is said to be significant. As such, it is highlighted.

Treatment: The set of levels of all factors included in a trial in an experiment is called a treatment or treatment combination.

Example: Consider the synchronizer example in the Design of Experiments tool write-up. The experiment was conducted and the following eventual results were obtained. The three factors and the interactive means were tested for statistical significance. Factor 1, factor 3, and the interactive effect of 1 * 2 were found to be significant. As such, we reject the null hypothesis of no differences between means and accept the alternative hypothesis that the means (in the population) are different from each other and hence impact the response variable output. All differences between the highlighted means obtained in the different experimental conditions can be sufficiently explained by the simple additive model for those three variables.

	ANOVA; Var.:**Results**; R-sqr=.97537; Adj:.94253 2**(3-0) design; MS Residual=.0000008 DV: **Results**				
Factor	SS	df	MS	F	p
(1)**Oil Level**	0.000012	1	0.000012	15.00000	0.030466
(2)**Oil Type**	0.000008	1	0.000008	9.60000	0.053363
(3)**Lot Number**	0.000061	1	0.000061	72.60000	0.003396
1 by 2	0.000018	1	0.000018	21.60000	0.018783
Error	0.000002	3	0.000001		
Total SS	0.000102	7			

Table B.2. Additive Model for Three Variables

General Comments

ANOVA is an integral part of the more crucial comparative analysis studies undertaken in any industry.

References

Box, G., Hunter, W., and Hunter, J. *Statistics for Experimenters: An Introduction to Design, Data Analysis and Model Building.* John Wiley & Sons, 1978.

Duncan, A. *Quality Control and Industrial Statistics*. Irwin, Inc., 1986.

Montgomery, D. *Design and Analysis of Experiments,* 3rd edition. John Wiley & Sons, 1991.