

One-Way ANOVA . One-way analysis of variance (ANOVA) is used for experimental data in which there is a continuous response variable and a single independent classification variable. The total variation in the response variable is explained as the sum of the variation due to the effects of the classification variable and the variation due to random error. This dialog is appropriate if you have a continuous dependent variable and a single classification variable.

One-Way ANOVA Output

The output of the analysis includes information about the levels of the independent variable followed by the standard ANOVA table, which includes the Degrees of Freedom, Sum of Squares and the Mean Square for the model and for random error. "R-Square" is the proportion of variability accounted for by the independent variables. "C.V." is the coefficient of variation. "Root MSE" is the square root of the Mean Square for Error. The row of the ANOVA table representing the independent variable is printed again. The F-statistic is the ratio of the Mean Square values, and "Pr > F" is the associated p -value. Small p -values indicate that the independent variable is significant in explaining the variation in the dependent variable, that is, that the means of the levels of the classification variable are not all the same. Additional output may be produced if you specify tests or means comparisons

PROC ANOVA: Introduction

The ANOVA procedure is one of several procedures available in SAS/STAT software for analysis of variance. The ANOVA procedure is designed to handle balanced data (that is, data with equal numbers of observations for every combination of the classification factors), whereas the GLM procedure can analyze both balanced and unbalanced data. Because PROC ANOVA takes into account the special structure of a balanced design, it is faster and uses less storage than PROC GLM for balanced data.

Use PROC ANOVA for the analysis of balanced data only, with the exceptions of one-way analysis of variance. **Caution:** If you use PROC ANOVA for analysis of unbalanced data, you must assume responsibility for the validity of the results.

PROC ANOVA: Syntax

The following statements are available in PROC ANOVA.

```
PROC ANOVA ;  
    CLASS variables ;  
    MODEL dependents=effects ;  
    BY variables ;  
    MEANS effects ;
```

The PROC ANOVA, CLASS, and MODEL statements are required, and they must precede the first RUN statement. The CLASS statement must precede the MODEL statement. If you use BY statement, it must precede the first RUN statement. The MEANS statement must follow the MODEL statement.

CLASS *variables ;*

The CLASS statement names the classification variables to be used in the model. Typical class variables are TREATMENT, SEX, RACE, GROUP, and REPLICATION. The CLASS statement is required, and it must appear before the MODEL statement. Classification variables are also called *categorical, qualitative, discrete, or nominal variables*. The values of a class variable are called *levels*. Class variables can be either numeric or character. This is in contrast to the *response (or dependent) variables*, which are continuous. Response variables must be numeric.

MODEL *dependents=effects;*

The MODEL statement names the dependent variables and independent effects. If no independent effects are specified, only an intercept term is fit. This tests the hypothesis that the mean of the dependent variable is zero. All variables in effects that you specify in the MODEL statement must appear in the CLASS statement.

BY *variables ;*

You can specify a BY statement with PROC ANOVA to obtain separate analyses on observations in groups defined by the BY variables. When a BY statement appears, the procedure expects the input data set to be sorted in order of the BY variables. The *variables* are one or more variables in the input data set. If your input data set is not sorted you should sort the data using the SORT procedure with a similar BY statement. If the BY statement is used, it must appear before the first RUN statement or it is ignored. When you use a BY statement, the interactive features of PROC ANOVA are disabled.

MEANS *effects ;*

PROC ANOVA can compute means of the dependent variables for any effect that appears on the right-hand side in the MODEL statement.