



Academic

## PROC UNIVARIATE and SAS V8

by Katrien Declercq

One of the new features of PROC UNIVARIATE in V8 is the HISTOGRAM statement. It creates high-resolution graphics of histograms and optionally superimposes parametric and nonparametric density curve estimates. A histogram breaks up the values of a variable in a number of classes and a bar represents the number of observations in that class. The histogram statement allows to specify one or more numeric variables to create a histogram for and options allow to fit densities on the histogram. The following parametric densities can be fitted: beta, exponential, gamma, lognormal, normal and Weibull. Supplementary options allow to specify shape, scale, threshold and location parameters for these distributions. Nonparametric densities can be fitted using kernel density estimation (option KERNEL). Additional options allow to specify the standardized bandwidth and the type of kernel density curve.

A set of options is available to specify the appearance of the estimated density curve (e.g. COLOR of fitted curve, ...) and the general appearance of the histogram (e.g. MIDPOINTS to list the midpoints of the histogram,...). The histograms can also be enhanced using annotation and another list of options. An output dataset may be created with information on the histogram intervals.

The following code illustrates some examples.

```
PROC UNIVARIATE DATA=sashelp.class;  
  VAR height;  
  HISTOGRAM height/NORMAL(COLOR=red FILL);  
run;
```

```
PROC UNIVARIATE DATA=sashelp.class;  
  HISTOGRAM height/GAMMA(COLOR=RED W=5) MIDPOINTS=50,55,60,65,70,75  
  HREF=60 VAXISLABEL="%";  
run;
```

And much, much more... Please check the V8 Online DOC for a full documentation of all options.