New Directions in SAS/QC® Software
Robert N. Rodriguez and Paul S. Bachtel
SAS Institute Inc.

Abstract

SAS/QC software provides a broad range of statistical tools for quality improvement of products, processes, and services. This paper describes ways in which these tools can be applied and highlights enhancements that are available in Release 6.03 of SAS/QC software for microcomputers and workstations.

Introduction

Statistical methods have played an increasingly important role in the field of industrial quality control during the 1980s. To remain competitive, many American companies have initiated programs in statistical quality control (SQC) or statistical process control (SPC), often as part of much broader organizational changes.

SAS/QC software provides a wealth of statistical and graphical tools that meet the needs for analysis, management, and presentation of data encountered by engineers, managers, and statisticians involved in large-scale applications of statistics to quality control. The software was introduced in Version 5 of the SAS System for mainframe computers and minicomputers; see Rodriguez (1986a). Release 6.03 of SAS/QC software makes these tools available on personal computers and workstations.

This paper describes new directions for SAS/QC software motivated by recent trends in statistical quality control. A series of examples illustrates these applications and highlights facilities newly available in Release 6.03.

Quality Improvement Through Designed Experimentation

The most significant trend in statistical quality control has been the increased use of designed experiments to improve the quality of products and processes from the start. The role of statistically planned experiments as an aid to quality engineering is widely recognized; see, for example, Box and Bisgaard (1987), Delmann (1989), and Taguchi and Wu (1980). While the methods of Taguchi remain controversial, there is general agreement that experimental design provides engineers with the basis for deciding which factors have an effect on the performance of a product or process, and how these factors should be adjusted.

To facilitate the design of industrial experiments, new tools have been added to Release 6.03 of SAS/QC software: the FACTEX procedure, the OPTEX procedure, the ADX macros, and the ADX menu system. Since these tools have been discussed by Tobias (1989) and are described fully in the software documentation (see the Appendix), a brief description will suffice here.

The FACTEX Procedure

The FACTEX procedure constructs orthogonally confounded fractional factorial designs, including two-level fractional factorial designs with and without blocking. These designs are commonly used for studying complicated processes that involve many variables. Because the FACTEX procedure uses a general algorithm to search for the construction rules for a design, it places no inherent limit on the number of factors or the size of the design that can be constructed. The FACTEX procedure provides extensive facilities for modifying the designs it constructs, including derived factors, replication, decoding, randomization, and inner and outer designs. These facilities can be used interactively to build designs interactively.

The OPTEX Procedure

The OPTEX procedure searches for D-optimal or A-optimal designs and is intended for situations where a standard design (such as a factorial design) is not appropriate. For example, PROC OPTEX might be used when resource limitations restrict the number of runs or when some combinations of factor levels cannot be run. The OPTEX procedure is also useful for studying designs interactively.

The ADX Macros

The ADX macros provide researchers with interactive programming facilities for standard design applications that combine the above procedures with tools in SAS/STAT® software and the SAS DATA step. The ADX macros can be used to construct fractional factorial designs; two-level screening designs (Plackett-Burman designs); orthogonal, rotatable central composite response surface designs; and mixture designs. The macros can also be used for decoding, randomization, printing data collection forms, and analysis of fractional factorial designs.

The ADX Menu System

The ADX menu system leads the user through the steps of selecting a standard design, decoding and randomizing the design, printing a data collection form, and analyzing the data with regression models, transformations, Bayes plots, cube plots, and contour plots. The designs constructed by the ADX system include two-level fractional factorial designs (with and without blocking) for as many as 11 factors and 128 runs, central composite and Box-Behnken designs for exploring response surfaces, and constrained and unconstrained mixture designs. The ADX menus were developed with SAS/AF® software and combine the FACTEX procedure with tools in SAS/STAT and SAS/GRAPH® software.

The ADX menus are intended for use by researchers and engineers with little or no SAS programming experience. Statistical training at the level of Box et al. (1978) is recommended, although complete on-line help is provided.

A Sample ADX Session

The following session illustrates how the ADX system can be used to create a fractional factorial design, customize the design, transform the response data, perform a regression analysis, and study the effects with Bayes plots. The example shown here is the half-fractional design constructed by Box et al. (1978, p. 376) to study the percentage of chemicals that react for a given chemical reaction.

The main menu for the ADX system is:
By selecting Option M we can choose to create a new design:

![Diagram of design options]

To select a two-level factorial design move the cursor to the appropriate label and press the ENTER key:

![Fractional factorial design]

A Resolution V design for five two-level factors (without blocking) is chosen by moving the cursor to D13 and pressing ENTER:

![Design selection]

This panel allows us to customize the design. To add a response variable, select Option 1 and press ENTER:

![Adding response variables]

Since this example involves a single response variable (percentage of chemicals that react) a '1' is placed in the required field. When ENTER is pressed we are prompted for the name of the response variable.
From this screen we first move to the Output Window to preview the design:

Then we return to the design summary panel (by entering the AF command) to change the default design name and add a label. We press the OK button to save the design.

Now, suppose that the experiment has been carried out. We are ready to add the values of the response variable and analyze the results. We select option W from the main menu:

Once the response data are entered we select option A from the Main Design Library panel to analyze the data. As a preliminary step we recode the design:

The next panel displays the default form of the model that we selected (main effects and two-factor interactions.)

The next panel lists the designs previously saved. The design we just constructed is available for subsequent analysis:
We choose not to modify this model and press the OK button. The next panel prompts us for the analysis:

The three default analysis components are TRANSFORM, REGRESSION, and ESTIMATES. These are executed, beginning with TRANSFORM, when we press OK:

Here a log transformation for the response is appropriate. Once this transformation is selected, the REGRESSION component is executed. The ADX menus use the REG procedure to fit the transformed data and then display a window for additional regression analysis:

For example, a stepwise regression can be of benefit when dealing with a large number of factors, only some of which may be active. (The goal is to separate the "important few" from the "trivial many"). Assuming that five factors are active, we skip this option and browse the parameter estimates displayed in the output window:

When we return to the regression analysis panel and press the OK button, the ESTIMATES component is executed:

Listed are the factors and a test of their significance (P-values). These P-values are understated because our model is saturated. The ADX menus provide graphical alternatives, such as normal plots and Bayes plots, that are useful in this situation.

To produce a Bayes plot, for example, we highlight the BAYES_PLOT button and press OK:
The Bayes plot indicates that CATALYST, TEMP, and CON are active.

Graphical Methods in Quality Improvement

Graphical methods play an increasingly important role in modern statistical quality control, particularly in the use of control charts for establishing and monitoring statistical control, and for reducing process variability.

The SHEWHART procedure creates a wide variety of Shewhart charts, including X and R charts, X and S charts, charts for individual measurements and moving ranges, p charts, np charts, u charts, and c charts. Two new statements (MCHART and MRCART) have been added to the SHEWHART procedure for creating charts for medians, as illustrated in the following display:

Sometimes process data are stratified in groups of consecutive subgroup samples. For example, suppose that samples of process output are recorded hourly, and that expert operators alternate with novice operators every ten hours. You can use the following statements to produce a p chart for the proportion of defective items produced that also displays the training level of the operator:

The MACONTROL procedure creates control charts for uniformly and exponentially weighted moving averages, and the CUSUM procedure creates cumulative sum (cusum) control charts for averages or individual measurements.

The next sections describe graphical enhancements of these procedures.

Displaying Stratified Process Data With Symbol and Block Variables

Two new features in the SHEWHART, CUSUM, and MACONTRL procedures, referred to as block variables and symbol variables, facilitate the display of stratified process data.

For example, suppose that cans are filled by two different machines, and that the weights of the cans are analyzed with X and R charts. If the samples can be classified (stratified) by machine, you can create charts that display the stratification with the following SAS statements:

The output reveals that Machine 1 is associated with smaller fill weights than Machine 2:

The next display illustrates the use of two labeled block variables (training and use of air conditioning):
Does Air Conditioning Help?

The SAS statements that produce the above display are:

```sas
proc shewhart graphics;
pchart DEFECTS * HOUR (TRAINING AIRCOND) / 
  subgroupn = 200 
  blockpos = 4 ;
label TRAINING = 'Operator skill level'
  AIRCOND = 'All-conditioning';
```

The analysis of multivariate process data is a topic of growing importance, and a variety of approaches can be used. The next display illustrates the use of the SHEWHART procedure to create a control chart for a quality variable (Paint Index) that uses stars to indicate the variation in six related environmental and process variables (see Rodriguez 1986b). The circles indicate nominal values for the six related variables, whose standardized values are given by the distance between the center and vertices of the stars. For example, the display indicates that high values of gloss (one o'clock position) and temperature (five o'clock position) are associated with high out-of-control values of average paint index. For other approaches to graphical display of multivariate process data see, for example, Blazek et al. (1987).

The analysis of multivariate process data is a topic of growing importance, and a variety of approaches can be used. The next display illustrates the use of the SHEWHART procedure to create a control chart for a quality variable (Paint Index) that uses stars to indicate the variation in six related environmental and process variables (see Rodriguez 1986b). The circles indicate nominal values for the six related variables, whose standardized values are given by the distance between the center and vertices of the stars. For example, the display indicates that high values of gloss (one o'clock position) and temperature (five o'clock position) are associated with high out-of-control values of average paint index. For other approaches to graphical display of multivariate process data see, for example, Blazek et al. (1987).

The analysis of multivariate process data is a topic of growing importance, and a variety of approaches can be used. The next display illustrates the use of the SHEWHART procedure to create a control chart for a quality variable (Paint Index) that uses stars to indicate the variation in six related environmental and process variables (see Rodriguez 1986b). The circles indicate nominal values for the six related variables, whose standardized values are given by the distance between the center and vertices of the stars. For example, the display indicates that high values of gloss (one o'clock position) and temperature (five o'clock position) are associated with high out-of-control values of average paint index. For other approaches to graphical display of multivariate process data see, for example, Blazek et al. (1987).

Analyzing Autocorrelated Process Data With TRENDVAR= Variables

Another topic of increasing interest is the analysis of autocorrelated process data, commonly encountered in chemical and food industry applications. Here, the standard Shewhart model is not valid since the process variation cannot be characterized as a series of random shocks away from a constant process mean; in other words, the mean "wanders" (see MacGregor et al. 1988.)

The next display illustrates a Shewhart chart (top half) for data with an EWMA structure after removing the autocorrelation structure (fitted with the ARIMA model plotted in the bottom half.) The model was fitted with the ARIMA procedure in SAS/ETS software, and the display was created with the SHEWHART procedure using the TRENDVAR= option (which can be used with the XCHART, BOXCHART, and MCHART statements in Release 6.02.) The dashed line indicates a forecast of the process mean and is bounded by a 95% prediction interval. For other approaches to this class of problems see MacGregor et al. (1988).
Service-Industry Application Using New BOXCHART Features

Statistical quality control methods are increasingly used in non-manufacturing applications. A potentially valuable tool in these settings is the boxplot with control limits, which you can create with the BOXCHART statement in PROC SHEWHART.

Flight Departure Analysis

This display was created with the following statements:

```plaintext
proc shewhart graphics;
   boxchart DELAY * DAY (TRAFFIC) / boxstyle = schematicid
      pctiledef = 5;
   id REASON;
```

Here DELAY is the delay in minutes between actual and posted departure times of daily subgroups of airplane flights during December (the data are hypothetical). The option BOXSTYLE=SCHENAMID specifies that values of the ID variable REASON (the recorded cause of the delay) are to be used to label unusually long delays. Here “unusually long” means a delay that exceeds an upper fence defined as the third quartile plus 1.5 times the inter-quartile distance. The upper whisker is drawn to the outermost delay inside the upper fence. (Delays less than a similarly defined lower fence do not occur in this example but would also be labeled.) A block variable (TRAFFIC) is used to differentiate between standard traffic and holiday traffic, and the method of computing the quantiles is specified with the PCTILEDEF= option.

New Graphical Features for Process Capability Analysis

The CAPABILITY procedure provides a variety of tools for process capability analysis, including histograms superimposed with fitted curves and specification limits, quantile-quantile plots, summary statistics, process capability indices, and statistical intervals. Three new statements (CDFPLOT, PROBPLOT, and PPPLOT) have been added to the CAPABILITY procedure for creating cumulative distribution function plots, probability plots, and P-P plots. Distributions available for fitting and plotting now include the beta, as well as the normal, lognormal, exponential, Weibull, and gamma.

The following examples were produced with the CDFPLOT and PROBPLOT statements, respectively:

Cumulative Distribution Plot

Normal Probability Plot
Making SAS/QC Software Broadly Accessible

In large-scale quality improvement programs it is essential to make statistical tools accessible to broad groups of users with specialized needs, limited training in statistics, and little or no SAS programming experience. For this reason, three menu-driven systems have been prepared as part of Release 6.03 of SAS/QC software. The ADX system is described above, and the other two are described below.

Menu System for Creating Ishikawa Diagram

The ISHIKAWA system provides an interactive tool for creating Ishikawa (cause-and-effect) diagrams. The screens display a skeletal representation of the trunk, branches, and stems for the diagram, and the user provides the labels for each level. The system provides on-line help and creates the diagram on a graphics device.

Menu System for Statistical Quality Control

The SQC system provides interactive tools for creating basic Shewhart charts, as well as histograms superimposed with specification limits, fitted curves, and summary statistics (process capability analysis). The system also provides screens for data entry.

Summary

The field of statistical quality control has evolved in the direction of statistical quality improvement through more frequent application of designed experimentation for building quality into products and processes from the start. Graphical methods are being used in increasingly diverse applications and make statistical methods accessible to broader groups of users. SAS/QC software has moved in these directions with the addition of new tools for design of experiments, graphical enhancements of procedures for statistical quality control, and the addition of menu systems.

References


Appendix: New Documentation for SAS/QC Software

The documentation for Version 6 SAS/QC software will consist of three books:

- SAS/QC Software: Reference, Version 6, which provides encyclopedic coverage of the software, including syntax examples and computational details but no worked-out examples.
A companion volume (not presently available) which will provide a tutorial, task-driven guide to the software and examples with graphical output.

SAS/QC Software Examples, Version 6 (Technical Report P-188) which provides comprehensive examples organized by procedure and includes graphical output.

The organization of the documentation is intended to meet the distinct needs of advanced and novice users. Note that Technical Report P-188 serves as an interim for the planned tutorial guide.

SAS, SAS/QC, SAS/STAT, SAS/GRAPH, SAS/ETS, and SAS/AF are registered trademarks of SAS Institute Inc., Cary, NC, USA.