SAS/STAT® 9.3 User's Guide
What's New in SAS/STAT 9.3
(Chapter)

Chapter 1

What's New in SAS/STAT 9.3

Contents

Overview

SAS/STAT 9.3 includes one new procedure and many enhancements.

New Experimental FMM Procedure

The experimental FMM procedure fits statistical models to data where the distribution of the response is a finite mixture of univariate distributions. These models are useful for applications such as estimating multimodal or heavy-tailed densities, fitting zero-inflated or hurdle models to count data with excess zeros, modeling overdispersed data, and fitting regression models with complex error distributions.

PROC FMM fits finite mixtures of regression models or finite mixtures of generalized linear models in which the regression structure and the covariates can be the same across components or different. Maximum likelihood and Bayesian methods are available with the FMM procedure.

Highlights of Enhancements

The following are the highlights of the enhancements in SAS/STAT 9.3:

- The EFFECT statement is now production. This statement is available in the HPMIXED, GLIMMIX, GLMSELECT, LOGISTIC, ORTHOREG, PHREG, PLS, QUANTREG, ROBUSTREG, SURVEYLOGISTIC, and SURVEYREG procedures.

- The MCMC procedure now supports the RANDOM statement.

- The METHOD=FIML option in the CALIS procedure is now production. This option specifies the full information maximum likelihood method. Instead of deleting observations with missing values, the full information maximum likelihood method uses all available information from all observations.

- The SURVEYPHREG procedure is now production.

- The HPMIXED procedure now provides a REPEATED statement and additional covariance structures.

- The MI procedure offers fully conditional specification methods for multiple imputation.

More information about the changes and enhancements follows. Details can be found in the documentation for the individual procedures in the *SAS/STAT 9.3 User's Guide*.

Highlights of Enhancements in SAS/STAT 9.22

Some users might be unfamiliar with updates made in SAS/STAT 9.22. The following are some of the major enhancements that were introduced in SAS/STAT 9.22:

- The experimental SURVEYPHREG procedure performs regression analysis based on the Cox proportional hazards model for sample survey data. The procedure provides design-based variance estimates, confidence intervals, and hypothesis tests concerning the parameters and model effects.

- The PLM procedure takes model results that are stored from SAS/STAT linear modeling procedures and performs additional postfitting inferences without your having to repeat your original analysis. The PLM procedure can perform tasks such as testing hypotheses, computing confidence intervals, producing prediction plots, and scoring new data sets by using familiar statements such as the ESTIMATE, LSMEANS, LSMESTIMATE, and SLICE statements.

- The EFFECT statement is now available in the GLIMMIX, GLMSELECT, HPMIXED, ORTHOREG, PHREG, PLS, QUANTREG, ROBUSTREG, SURVEYLOGISTIC, and SURVEYREG procedures. This statement enables you to construct a much richer family of linear models than you can traditionally define with the CLASS statement. Effect types include splines for semiparametric modeling, multimember effects for situations in which measurements can belong to more than one class, lag effects, and polynomials.

- Exact Poisson regression is now available with the GENMOD procedure.

- The MCMC procedure can create samples from the posterior predictive distribution.

- The zero-inflated negative binomial model is now available with the GENMOD procedure.

- The HPMIXED procedure is now production.

- The CALIS procedure has been completely revised and includes enhancements that were formerly available in the experimental TCALIS procedure.

ODS Graphics Changes

Producing graphs with ODS Graphics no longer requires a SAS/GRAPH® license. In addition, the family of statistical graphics procedures (SGPANEL, SGPLOT, SGRENDER, and SGSCATTER) has moved from SAS/GRAPH to Base SAS® license.

The MAXPOINTS= option has been added to the ANOVA, CLUSTER, GLM, LOGISTIC, MIXED, QUANTREG, and VARCLUS procedures. This option specifies a limit for the number of points that can be displayed on certain plots, and these plots are not created when this limit is exceeded. Note that the REG procedure already provided this option.

The frequency plots and cumulative frequency plots of PROC FREQ and the weighted frequency plot of PROC SURVEYFREQ are no longer produced automatically when ODS Graphics is enabled. You can request these graphs with the PLOTS= option.

In SAS 9.3, the default destination in the SAS windowing environment is HTML; in addition, ODS Graphics is enabled by default in the SAS windowing environment. These new defaults have several advantages. Graphs are integrated with tables, and all output is displayed in the same HTML file using a new style. This new style, HTMLBLUE, is an all-color style, which is designed to integrate tables and modern statistical graphics. You can view and modify the default settings by selecting **Tools ▶ Options ▶ Preference** from the menu at the top of the main SAS window. Then click the **Results** Tab.

Enhancements

CALIS Procedure

The following features are now production:

- METHOD=FIML option

- mean structure analysis with the COSAN model

- extended PATH modeling language that supports the specification of variances or covariances as paths

- unnamed free parameter specification in all model types

- improved RAM model specification

In addition, PROC CALIS now provides detailed analysis of the missing patterns with the FIML estimation method. With the COVPATTERN= and MEANPATTERN= options, you can specify various standard mean and covariance patterns by using keywords. PROC CALIS then generates the required covariance and mean structures automatically.

CLUSTER Procedure

The CLUSTER procedure now produces a dendrogram by default when ODS Graphics is enabled. The MAXCLUS= option enables you to right-truncate the CCC, PSF, and PST2 plots to improve readability. The MAXPOINTS= option enables you to suppress the dendrogram when there is a large number of clusters.

EFFECT Statement

The EFFECT statement is now production. This statement is available in the HPMIXED, GLIMMIX, GLMSELECT, LOGISTIC, ORTHOREG, PHREG, PLS, QUANTREG, ROBUSTREG, SURVEYLOGIS-TIC, and SURVEYREG procedures.

The NATURALCUBIC option specifies a natural cubic spline basis for the spline expansion.

EFFECTPLOT Statement

The CLUSTER option modifies the box plot display by displaying a plot for each level of the SLICEBY= classification variable.

FREQ Procedure

The FREQ procedure now produces agreement plots when the AGREE option is specified and ODS Graphics is enabled. It also offers a number of alternative confidence limits for the proportion difference, and it provides exact unconditional confidence limits for the proportion difference that are based on the Farrington-Manning score statistic.

GENMOD Procedure

The EXACTMAX option in the MODEL statement limits the number of response values for exact Poisson regression.

GLIMMIX Procedure

The EFFECT statement is now production.

GLMPOWER Procedure

The GLMPOWER procedure now produces its graphs with ODS Graphics.

GLMSELECT Procedure

The GLMSELECT procedure now provides a STORE statement which enables you to save the context and results of the statistical analysis for further processing with the PLM procedure.

The MODELAVERAGE statement, which specifies model selection on resampled subsets of the input data, is now production.

The EFFECT statement is now production.

HPMIXED Procedure

The HPMIXED procedure now provides the REPEATED statement, which defines the repeated effect and the residual covariance structure in the mixed model. The AR(1), CS, CSH, UC, UCH, and UN covariance structures are now available with the TYPE= option in the RANDOM statement.

The EFFECT statement is now production.

LIFETEST Procedure

The X axis tick marks are now aligned with the at-risk values in the survival plot.

The survival plot template is available in the SAS sample library in macro form, which makes it easier to modify.

LOGISTIC Procedure

You can now request that standardized residuals be saved in the output data set. In addition, the STDRES suboption of the INFLUENCE option in the MODEL statement includes standardized residuals and likelihood residuals in the resulting display. The FITSTAT option in the SCORE statement produces the AIC, SBC, RSq, AUC, and Brier score fit statistics. Additionally, the ODDSRATIO statement and the CLDISPLAY= suboption of the CLODDS option control the appearance of the confidence limit error bars.

The EFFECT statement is now production.

MCMC Procedure

The new RANDOM statement simplifies the construction of hierarchical random-effects models and significantly reduces simulation time while improving convergence, especially in models with a large number of subjects or clusters. This statement defines random effects that can enter the model in a linear or nonlinear fashion and supports univariate and multivariate prior distributions.

In addition to the default Metropolis-based algorithms, PROC MCMC now takes advantages of certain forms of conjugacy in the model in order to sample directly from the target conditional distributions. In

many situations, the conjugate sampler increases sampling efficiency and provides a substantial reduction in computing time.

The MCMC procedure now supports multivariate distributions including the Dirichlet, inverse Wishart, multivariate normal, and multinomial distributions.

MI Procedure

The experimental FCS statement specifies a multivariate imputation by fully conditional specification (FCS) methods. For data with an arbitrary missing data pattern, these methods enable you to impute missing values for all variables, assuming that a joint distribution for these variables exists. The FCS method requires fewer iterations than the MCMC method.

MULTTEST Procedure

The STOUFFER option in the PROC statement produces adjusted p-values by using the Stouffer-Liptak combination method.

NLIN Procedure

The NLIN procedure provides several experimental features for diagnosing your nonlinear model fit, including the PLOTS, NLINMEASURES, and BIAS options in the PROC NLIN statement, in addition to producing observation-wise statistics in the OUTPUT data set. The PLOTS option enables you to plot the fitted model, fit diagnostics, tangential and Jacobian leverage, and local influence. The NLINMEASURES displays global measures of nonlinearity, and the BIAS option computes Box's bias statistics for the parameter estimates. Finally, you can add the leverage, local influence, and residual diagnostics in the output data set that is produced with the OUTPUT statement.

ORTHOREG Procedure

The EFFECT statement is now production.

PHREG Procedure

The PHREG procedure now fits frailty models with the addition of the RANDOM statement. You often use frailty models when you analyze clustered data and want to account for the within-cluster correlation with

random effects. In addition, the NLOPTIONS statement is available with PROC PHREG, and the Zellner g-prior is now available for the piecewise exponential model.

The EFFECT statement is production.

PLS Procedure

The EFFECT statement is now production.

POWER Procedure

Graphs are now produced with ODS Graphics.

QUANTREG Procedure

The new QINTERACT option in the TEST statement enables you to test whether any difference exists among the coefficients across quantiles if several quantiles are specified in the MODEL statement.

The RANKSCORE option in the TEST statement now supports the tau score function, which is appropriate for non-iid error models.

The EFFECT statement is now production.

ROBUSTREG Procedure

The new MCDINFO suboption of the LEVERAGE option in the MODEL statement displays detailed information about the MCD covariance estimate, including the low-dimensional structure, the breakdown value, the MCD center, and the MCD covariance.

The EFFECT statement is now production.

SURVEYFREQ Procedure

You can now produce Rao-Scott chi-square tests with second-order corrections.

SURVEYLOGISTIC Procedure

Replication variance estimation is now available for domain analysis.

The EFFECT statement is now production.

SURVEYMEANS Procedure

Variance estimation based on replication methods is now available for quantiles.

SURVEYPHREG Procedure

The SURVEYPHREG procedure is now production. Also, the addition of programming statements enables you to include time-dependent covariates in the model.

SURVEYREG Procedure

The SURVEYREG procedure now provides replication variance estimation for domain analysis.

The EFFECT statement is now production.

SURVEYSELECT Procedure

Instead of specifying the total sample size to allocate among the strata, you can specify the desired margin of error for estimating the overall mean from the stratified sample.

VARCLUS Procedure

The VARCLUS procedure now produces a dendrogram by default when ODS Graphics is enabled. The MAXPOINTS= option enables you to suppress the dendrogram when there is a large number of clusters.

What's Changed

What follows are changes in software behavior from SAS/STAT 9.22 to SAS/STAT 9.3. Several of these changes are related to ODS Graphics. A few procedures have adopted the MAXPOINTS= option as a way to avoid producing plots when the number of points exceeds a specified limit. The default limit is 5,000 points.

ANOVA Procedure

Box plots, which are created with the MEANS statement or for one-way ANOVA when ODS Graphics is enabled, are not produced when the number of outlier points exceeds the limit, which is controlled by the MAXPOINTS= option.

CLUSTER Procedure

The CLUSTER procedure now produces a dendrogram by default when ODS Graphics is enabled.

FREQ Procedure

Frequency plots and cumulative frequency plots are no longer produced by default when ODS Graphics is enabled. You can request these plots with the PLOTS=FREQPLOT and PLOTS=CUMFREQPLOT options in the TABLES statement.

GLM Procedure

The fit plot, box plot, interaction plot, ANCOVA plot, and contour fit plot are not produced when the number of points exceeds the limit, which is contolled by the MAXPOINTS= option. This limit also applies to diagnostic plots and residual plots.

LOGISTIC Procedure

Plots associated with the INFLUENCE or IPLOTS= options in the MODEL statement are not produced when the number of points exceeds the limit, which is controlled by the MAXPOINTS= option.

If the ODDSRATIO statement or CLODDS= option is specified, the default "Odds Ratio" table is no longer produced, and only the requested results are displayed.

MCMC Procedure

PROC MCMC no longer produces the tuning, burn-in, and sampling history tables by default. To produce this information, specify the MCHISTORY= option in the PROC MCMC statement.

The scaled inverse chi-square distribution is parameterized in terms of *scale*^{2}, as opposed to *scale* in the previous release.

MIXED Procedure

Plots associated with the INFLUENCE, RESIDUAL, and VCIRY options are not produced when the number of points exceeds the limit, which is controlled by the MAXPOINTS= option.

QUANTREG Procedure

The fit plot is not produced when the number of points exceeds the limit, which is controlled by the MAX-POINTS= option.

The rank score test has changed.

SURVEYFREQ Procedure

The weighted frequency plot is no longer produced by default when ODS Graphics is enabled. You can request this display with the PLOTS=WTFREQPLOT option in the TABLES statement.

VARCLUS Procedure

The VARCLUS procedure now produces a dendrogram by default when ODS Graphics is enabled.

Your Turn

We welcome your feedback.

- If you have comments about this book, please send them to
  yourturn@sas.com. Include the full title and page numbers (if applicable).
- If you have comments about the software, please send them to
  suggest@sas.com.

SAS® Publishing Delivers!

Whether you are new to the work force or an experienced professional, you need to distinguish yourself in this rapidly changing and competitive job market. SAS® Publishing provides you with a wide range of resources to help you set yourself apart. Visit us online at support.sas.com/bookstore.

SAS® Press

Need to learn the basics? Struggling with a programming problem? You'll find the expert answers that you need in example-rich books from SAS Press. Written by experienced SAS professionals from around the world, SAS Press books deliver real-world insights on a broad range of topics for all skill levels.

support.sas.com/saspress

SAS® Documentation

To successfully implement applications using SAS software, companies in every industry and on every continent all turn to the one source for accurate, timely, and reliable information: SAS documentation. We currently produce the following types of reference documentation to improve your work experience:

- Online help that is built into the software.
- Tutorials that are integrated into the product.
- Reference documentation delivered in HTML and PDF – **free** on the Web.
- Hard-copy books.

support.sas.com/publishing

SAS® Publishing News

Subscribe to SAS Publishing News to receive up-to-date information about all new SAS titles, author podcasts, and new Web site features via e-mail. Complete instructions on how to subscribe, as well as access to past issues, are available at our Web site.

support.sas.com/spn

§sas | THE POWER TO KNOW®