



THE  
POWER  
TO KNOW.

# **SAS/STAT® 9.22 User's Guide**

## **What's New in SAS/STAT 9.22**

### **(Book Excerpt)**



This document is an individual chapter from *SAS/STAT® 9.22 User's Guide*.

The correct bibliographic citation for the complete manual is as follows: SAS Institute Inc. 2010. *SAS/STAT® 9.22 User's Guide*. Cary, NC: SAS Institute Inc.

Copyright © 2010, SAS Institute Inc., Cary, NC, USA

All rights reserved. Produced in the United States of America.

**For a Web download or e-book:** Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

**U.S. Government Restricted Rights Notice:** Use, duplication, or disclosure of this software and related documentation by the U.S. government is subject to the Agreement with SAS Institute and the restrictions set forth in FAR 52.227-19, Commercial Computer Software-Restricted Rights (June 1987).

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st electronic book, May 2010

SAS® Publishing provides a complete selection of books and electronic products to help customers use SAS software to its fullest potential. For more information about our e-books, e-learning products, CDs, and hard-copy books, visit the SAS Publishing Web site at [support.sas.com/publishing](http://support.sas.com/publishing) or call 1-800-727-3228.

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

# Chapter 1

## What's New in SAS/STAT 9.22

### Contents

---

Overview . . . . .	2
New Procedures . . . . .	2
Highlights of Enhancements . . . . .	2
Documentation Enhancements . . . . .	3
Highlights of Enhancements in SAS/STAT 9.2 . . . . .	3
CALIS Procedure . . . . .	4
FACTOR Procedure . . . . .	4
FREQ Procedure . . . . .	4
GAM Procedure . . . . .	4
GENMOD Procedure . . . . .	5
GLIMMIX Procedure . . . . .	5
GLM Procedure . . . . .	5
GLMSELECT Procedure . . . . .	6
HPMIXED Procedure . . . . .	6
KRIGE2D Procedure . . . . .	6
LIFEREG Procedure . . . . .	7
LIFETEST Procedure . . . . .	7
LOGISTIC Procedure . . . . .	7
MCMC Procedure . . . . .	7
MIXED Procedure . . . . .	8
ORTHOREG Procedure . . . . .	8
PHREG Procedure . . . . .	8
PLM Procedure . . . . .	9
PLS Procedure . . . . .	9
POWER Procedure . . . . .	9
QUANTREG Procedure . . . . .	10
ROBUSTREG Procedure . . . . .	10
SIM2D Procedure . . . . .	10
SURVEYFREQ Procedure . . . . .	11
SURVEYLOGISTIC Procedure . . . . .	11
SURVEYMEANS Procedure . . . . .	11
SURVEYPHREG Procedure . . . . .	11
SURVEYREG Procedure . . . . .	12
SURVEYSELECT Procedure . . . . .	12

TPSPLINE Procedure . . . . .	12
VARCOMP Procedure . . . . .	12
VARIOGRAM Procedure . . . . .	12

---



---

## Overview

SAS/STAT 9.22 includes two new procedures and many new enhancements.

---

## New Procedures

The experimental SURVEYPHREG procedure performs regression analysis based on the Cox proportional hazards model for sample survey data. Cox's semiparametric model is widely used in the analysis of survival data to estimate hazard rates when explanatory variables are available. 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. It can handle model results that are stored from the following SAS/STAT procedures: GENMOD, GLIMMIX, GLM, LOGISTIC, MIXED, ORTHOREG, PHREG, SURVEYLOGISTIC, SURVEYPHREG, and SURVEYREG.

---

## Highlights of Enhancements

Additional statements for least squares means, scoring, and other advanced postfitting inferences have been added to several modeling procedures, including the LOGISTIC, MIXED, and ORTHOREG procedures. Other highlights include:

- The EFFECT statement is now available in the HPMIXED, GLIMMIX, GLMSELECT, LOGISTIC, 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.

More information about the changes and enhancements follow. Details can be found in the documentation for the individual procedures in the *SAS/STAT 9.22 User's Guide*.

---

## Documentation Enhancements

In recent releases, this documentation has grown to contain chapters that apply to many SAS/STAT procedures, such as “Using the Output Delivery System” and “Statistical Graphics Using ODS.” “Shared Concepts and Topics” is another such chapter, and it is greatly expanded in this edition to include information about the postfitting statements that now are common to many linear modeling procedures.

---

## Highlights of Enhancements in SAS/STAT 9.2

Some users are moving directly from SAS/STAT 9.1.3 to SAS/STAT 9.22. The following are some of the major enhancements that were introduced in SAS/STAT 9.2:

- ODS Statistical Graphics became production and over 60 procedures in SAS/STAT, SAS/ETS®, SAS/QC®, and Base SAS® software have been modified to use it. Many new plots are now produced by these procedures, either by default or with the specification of procedure options.
- The GENMOD, LIFEREG, and PHREG procedures now include facilities for Bayesian analysis.
- The MCMC procedure is a flexible simulation-based procedure that is suitable for fitting a wide range of Bayesian models.
- The SEQDESIGN procedure designs interim analyses for clinical trials. The SEQTEST procedure performs the interim analyses based on design information produced by the SEQDESIGN procedure.

For more information, see [support.sas.com/whatsnewstat92](http://support.sas.com/whatsnewstat92).

---

## CALIS Procedure

The CALIS procedure now includes updates that were previously surfaced in the experimental TCALIS procedure. These capabilities include:

- new modeling languages such as LISMOD, MSTRUCT, and PATH
- multiple group analysis
- improved mean structure analysis
- general parametric function testing
- improved effect analysis

In addition, PROC CALIS introduces several experimental features, including the full information likelihood method (FIML), mean structure analysis with the COSAN model, unnamed free parameter specification, and an extended path modeling language.

---

## FACTOR Procedure

The output now includes a table with the number of observations used in the analysis.

---

## FREQ Procedure

Exact  $p$ -values are available for tests of the following measures: Kendall's tau- $b$ , Stuart's tau- $c$ , Somers' D(CIR), and Somers' D(RIC). The GAILSIMON option in the TABLE statement specifies the Gail-Simon test for qualitative interactions, and the MANTELFLEISS suboption of the CMH option requests the Mantel-Fleiss criterion for the Mantel-Haenszel statistic for stratified  $2 \times 2$  tables.

Relative risk plots and risk difference plots are now available.

---

## GAM Procedure

The LOESS smoother in the MODEL statement is now production.

---

## GENMOD Procedure

The EFFECTPLOT statement produces a display of the fitted model. The LSMESTIMATE and the SLICE statements provide additional postprocessing inferences. The STORE statement enables you to save the context and results of the statistical analysis for further processing with the PLM procedure. The LSMEANS statement has been updated to include options such as the AT, ADJUST=, STEPDOWN, and PLOTS= options.

The zero-inflated negative binomial model is now available through the ZEROMODEL statement.

New sampling methods are available with the Bayesian analysis offered in PROC GENMOD. For the normal distribution with a conjugate prior, the closed form for the posterior distribution is now used by default. The ARMS algorithm is otherwise the default, but you can now specify the Gamerman algorithm or the independent Metropolis algorithm with the SAMPLING= option in the BAYES statement.

You can now perform exact Poisson regression and exact logistic regression by using the EXACT statement in PROC GENMOD.

---

## GLIMMIX Procedure

The SLICE statement enables you to perform inferences on model effects that consist entirely of classification variables. These effects must be higher-order effects of at least two classification variables. The STORE statement enables you to save the context and results of the statistical analysis for further processing with the PLM procedure. The CPSEUDO option in the OUTPUT statement changes the way in which marginal residuals are computed when model parameters are estimated by pseudo-likelihood methods.

You can now perform a joint test under one-sided restrictions with the LSMESTIMATE statement (Silvapulle and Sen 2004); for example, you can test ordered alternatives. The GLIMMIX procedure computes a simulation-based chi-bar-square statistic and produces a  $p$ -value for the constrained joint test.

---

## GLM Procedure

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

---

## GLMSELECT Procedure

The GLMSELECT procedure now provides model averaging with the experimental MODELAVERAGE statement, which requests model selection on resampled subsets of the input data. An average model is produced by averaging the parameter estimates of the selected models that are obtained for each resampled subset of the input data.

The ADAPTIVE option of the SELECTION=LASSO method specifies adaptive lasso selection, which is a modification of lasso selection where weights are applied to each of the parameters in forming the lasso constraint.

---

## HPMIXED Procedure

The HPMIXED procedure is production with this release.

The experimental EFFECT statement enables you to construct a much richer family of linear models than you can traditionally define with the CLASS statement. The BLUP= option in the PROC HPMIXED statement creates a data set that contains the BLUE and BLUP solutions. This option is designed for users who need BLUP/BLUE solutions for random effects with many levels, up to tens of millions.

The SLICE and DIFF options are now supported in the LSMEANS statement.

---

## KRIGE2D Procedure

The RESTORE statement specifies an item store that provides spatial correlation model input for the PROC KRIGE2D prediction tasks. The KRIGE2D procedure can use only item stores that are created by PROC VARIOGRAM.

The ID statement specifies which variable to include for identification of the observations in the OUTNBHD= output data set. The ID statement variable is also used for the labels and tool tips in the observations plot and the tool tips in the prediction plot.

You can now request plots of the semivariogram model used for prediction tasks. You can also produce plots for prediction at individual points or in grids in one dimension. The LABEL= option in the GRID statement enables you to identify the prediction locations for grids in one dimension.

---

## LIFEREG Procedure

Fit criteria based on the distribution of the response on the original scale, rather than the log of the response, are reported if you specify the Weibull, exponential, lognormal, log-logistic, or gamma distribution.

---

## LIFETEST Procedure

You can now request the Breslow and Fleming-Harrington estimates of the survivor function with the `METHOD=` option in the `PROC LIFETEST` statement. The number of subjects at risk can be displayed with the product-limit estimates, the Breslow estimates, and the Fleming-Harrington estimates.

---

## LOGISTIC Procedure

The experimental `EFFECT` statement enables you to construct a much richer family of linear models than you can traditionally define with the `CLASS` statement. The `EFFECTPLOT` statement produces a graphical display of the fitted model.

The `ESTIMATE`, `LSMEANS`, `LSMESTIMATE`, and `SLICE` statements provide additional postfitting inferences. The `STORE` statement enables you to save the context and results of the statistical analysis for further processing with the `PLM` procedure.

---

## MCMC Procedure

The `PREDDIST` statement creates random samples from the posterior predictive distribution of the response variable and saves the samples to a SAS data set. The posterior predictive distribution is the distribution of unobserved observations (prediction) conditional on the observed data.

---

## MIXED Procedure

The LSMEESTIMATE and SLICE statements provide additional postfitting inferences. The STORE statement enables you to save the context and results of the statistical analysis for further processing with the PLM procedure.

---

## ORTHOREG Procedure

The ORTHOREG procedure fits general linear models by the method of least squares. Other SAS/STAT software procedures, such as GLM and REG, fit the same types of models, but PROC ORTHOREG can produce more accurate estimates than other regression procedures when your data are ill-conditioned.

PROC ORTHOREG has been greatly expanded in this release to provide postfitting inferences with the inclusion of the ESTIMATE, LSMEANS, LSMEESTIMATE, SLICE, and TEST statements. In addition, the EFFECTPLOT statement produces a graphical display of the fitted model.

PROC ORTHOREG also includes the experimental EFFECT statement, which enables you to construct a much richer family of linear models than you can traditionally define with the CLASS statement.

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

---

## PHREG Procedure

The PHREG procedure now supports the ESTIMATE, LSMEANS, LSMEESTIMATE, and SLICE statements for additional postfitting inferences.

The experimental EFFECT statement enables you to construct a much richer family of linear models than you can traditionally define with the CLASS statement. The STORE statement enables you to save the context and results of the statistical analysis for further processing with the PLM procedure.

The ATRISK option in the PROC PHREG statement displays a table that contains the number of units at risk at each event time and the corresponding number of events in the risk sets.

You can now specify the Zellner g-prior for the regression coefficients in the BAYES statement. You can also request the random walk Metropolis (RWM) algorithm to sample an entire parameter vector from the posterior distribution in a Bayesian analysis.

Likelihood ratio tests of model parameters are available with the TYPE1 and TYPE3 options in the MODEL statement except when the robust sandwich estimate for the covariance matrix is specified.

---

## PLM Procedure

The PLM procedure performs postfitting inferences for model results that are stored by one of the following SAS/STAT procedures: GENMOD, GLIMMIX, GLM, LOGISTIC, MIXED, ORTHOREG, PHREG, SURVEYLOGISTIC, SURVEYPHREG, and SURVEYREG. These procedures now include the STORE statement, which produces item stores that can then be used as input for the PLM procedure.

PROC PLM can perform tasks such as testing hypotheses, computing confidence intervals, producing prediction plots, and scoring new data sets. This enables you to separate common postfitting inferences, such as testing for treatment differences and predicting new observations under a fitted model, from the process of model building and fitting. PROC PLM offers the most advanced postfitting inference techniques available in SAS/STAT software, including new techniques such as step-down multiplicity adjustments for  $p$ -values,  $F$  tests with order restrictions, analysis of means (ANOM), and sampling-based linear inference based on Bayes posterior estimates.

The PLM procedure supports the EFFECTPLOT, ESTIMATE, FILTER, LSMEANS, LSMESTIMATE, SCORE, SHOW, SLICE, TEST, and WHERE statements.

---

## PLS Procedure

The PLS procedure now supports the experimental EFFECT statement, which enables you to construct a much richer family of linear models than you can traditionally define with the CLASS statement.

---

## POWER Procedure

You can now parameterize computations for survival analysis in terms of the expected number of events, in addition to sample size. See the EVENTSPERGROUPE=, EVENTSTOTAL=, and GROUPEVENTS= options in the TWOSAMPLESURVIVAL statement. Parameterization in terms of sample size accrued per unit time is also available in this statement with the ACCRUALRATEPERGROUPE=, ACCRUALRATETOTAL=, and GROUPACCRUALRATES= options.

---

## QUANTREG Procedure

If you specify multiple quantiles in a MODEL statement, additional analyses, such as those specified in the TEST statement, are now produced for each quantile specified.

The RANKSCORE option in the TEST statement enables you to perform rank tests. Available score functions provide normal scores, Wilcoxon scores, and sign scores, which are asymptotically optimal for the Gaussian, logistic, and Laplace location shift models, respectively.

---

## ROBUSTREG Procedure

You can specify classification effects with the LTS, S, and MM methods. PROC ROBUSTREG now computes a robust version of the Mahalanobis distance by using the generalized minimum covariance determinant (MCD) method. Leverage point analysis is updated to reflect the inclusion of classification variables.

The experimental EFFECT statement enables you to construct a much richer family of linear models than you can traditionally define with the CLASS statement.

---

## SIM2D Procedure

The RESTORE statement specifies an item store that provides spatial correlation model input for the PROC SIM2D simulation tasks. The SIM2D procedure can use only item stores that are created by PROC VARIOGRAM. You can request scatter plots, simulation plots, and plots of the semivariogram models.

The ID statement specifies which variable to include for identification of the observations in labels and tool tips for the observations plot and in tool tips for the simulation plot. The ID variable is used only when you perform conditional simulation.

You can now produce plots for simulation at individual points or in grids in one dimension. The LABEL= option in the GRID statement enables you to identify the simulation locations for grids in one dimension.



---

## SURVEYFREQ Procedure

The SURVEYFREQ procedure now provides plots created with ODS Graphics, including a weighted frequency plot, an odds ratio plot, a relative risk plot, and a risk difference plot. The CL option now offers additional confidence limit types, including the modified Clopper-Pearson (exact), modified Wilson (score), and logit confidence limits.

If you specify the DEFF option in the TABLES statement, PROC SURVEYFREQ computes design effects for the overall proportion estimates in the frequency and crosstabulation tables.

---

## SURVEYLOGISTIC Procedure

The SURVEYLOGISTIC procedure now includes the experimental EFFECT statement, which enables you to construct a much richer family of linear models than you can traditionally define with the CLASS statement.

PROC SURVEYLOGISTIC also includes the ESTIMATE, LSMEANS, LSMESTIMATE, and SLICE statements for additional postfitting inferences. The new STORE statement enables you to save the context and results of the statistical analysis for further processing with the PLM procedure.

---

## SURVEYMEANS Procedure

The SURVEYMEANS procedure now performs analysis for domain ratios. Variance estimation based on replication methods is available for domain means, totals, and ratios.

---

## SURVEYPHREG Procedure

The experimental SURVEYPHREG procedure performs regression analysis based on the Cox proportional hazards model for sample survey data. Cox's semiparametric model is widely used in the analysis of survival data to estimate hazard rates when explanatory variables are available. The procedure provides design-based variance, confidence intervals, and tests for the estimated parameters in the model.

---

## SURVEYREG Procedure

The SURVEYREG procedure now includes the LSMEANS, LSMESTIMATE, SLICE, and TEST statements for additional postfitting inferences. The new STORE statement enables you to save the context and results of the statistical analysis for further processing with the PLM procedure.

The experimental EFFECT statement enables you to construct a much richer family of linear models than you can traditionally define with the CLASS statement.

---

## SURVEYSELECT Procedure

The SAMPLINGUNIT statement names variables that identify the sampling units as groups of observations (clusters). The combinations of categories of SAMPLINGUNIT variables define the sampling units. If there is a STRATA statement, sampling units are nested within strata. The NMIN= option in the PROC SURVEYSELECT statement specifies the minimum stratum sample size for the SAMPRATE= option.

---

## TPSPLINE Procedure

The TPSPLINE procedure now provides plots created with ODS Graphics, including residual plots, diagnostic plots, and fit plots. You can now request confidence bands for the expected value of the dependent variables by using the UCLM and LCLM keywords in the SCORE statement.

---

## VARCOMP Procedure

You can now request generalized confidence limits for the parameters with the CL=GCL option in the MODEL statement.

---

## VARIOGRAM Procedure

The STORE statement requests that the procedure save the context and results of the semivariogram model fitting analysis in an item store. The contents of item stores produced by PROC VARIOGRAM

can be processed only with the KRIGE2D or the SIM2D procedure. After you save results in an item store, you can use them at a later time without having to refit the model.

The ID statement specifies which variable to include for identification of the observations in the OUTPAIR= and the OUTACWEIGHTS= output data sets. The ID statement variable is also used for labels and tool tips in the observations plot.

PROC VARIOGRAM now provides a Moran plot, which is a scatter plot of standardized observed values against weighted averages.



## 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](http://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](http://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](http://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](http://support.sas.com/spn)**



**THE  
POWER  
TO KNOW®**

