

What's New in SAS/STAT for SAS 9.2

Melodie Rush, Senior Analytical Engineer

CUSTOMER LOYALTY TEAM • Support You Can Count On



**THE
POWER
TO KNOW®**

Agenda...

What's New in SAS/STAT for SAS 9.2

- ODS Graphics
- New Procedures
- Highlights of Enhancements to Current Procedures
- Highlights of What's New in IML & QC
- Resources

Key Benefits for the Statistician

- New and improved analytical methods
 - Statistical Modeling
 - Bayesian Analysis
 - Survey Data Analysis
- Easy-to-use production quality graphics
- Many enhancements to existing procedures

What is New for SAS/Stat in SAS 9.2

- Two updates so far - 9.2 and 9.22
- When you upgrade to the current release you have both updates.
- SAS/STAT, SAS/ETS and SAS/OR can independently release updates.

ODS Statistical Graphics – Now Production!

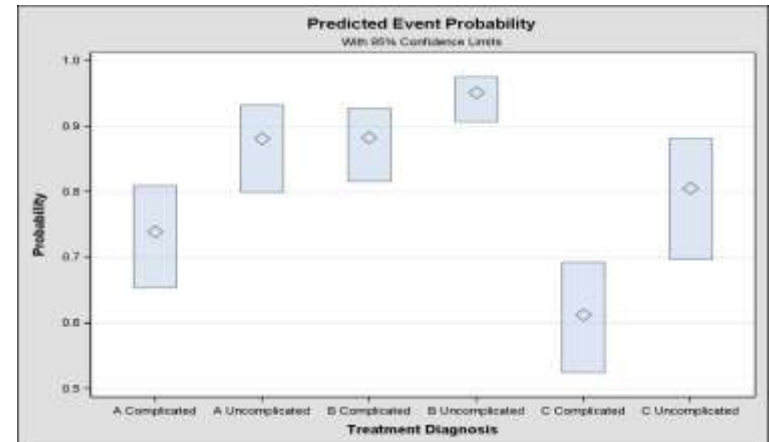
- *Automates* creation of commonly used displays in statistical procedures
- *Experimental* in SAS 9.1
- *Production* in SAS 9.2 for over 50 procedures
- *Functionality* is being used to provide modern statistical graphics in a variety of use cases
 - convenient specialized graphics in statistical procedure output
 - standalone graphics for preliminary analysis of data
 - customized graphics (using the underlying Graph Template Language)

NOTE: A SAS/GRAPH license will be required in order to use the ODS graphics functionality in SAS/STAT for 9.2.

ODS Statistical Graphics

ODS Graphics provides statistical graphics.

- Used in over 40 STAT procedures
- New SG family of procedures in SAS/GRAPH
- Customize by modifying templates
- Editor for modifying graphs on the fly



Procedures Supporting ODS Graphics

■ SAS/STAT

- ANOVA
- BOXPLOT
- CALIS
- CLUSTER
- CORRESP
- FACTOR
- GAM
- GENMOD
- GLM
- GLIMMIX
- GLMSELECT
- KDE
- LIFEREG
- LIFETEST
- LOESS
- LOGISTIC
- MDS
- MI
- MIXED
- MULTTEST
- NPAR1WAY
- PHREG
- PLS
- PRINCOMP
- PRINQUAL
- QUANTREG
- REG
- ROBUSTREG
- RSREG
- TRANSREG
- TTEST

Basics of ODS Graphics

- *Invoke* with

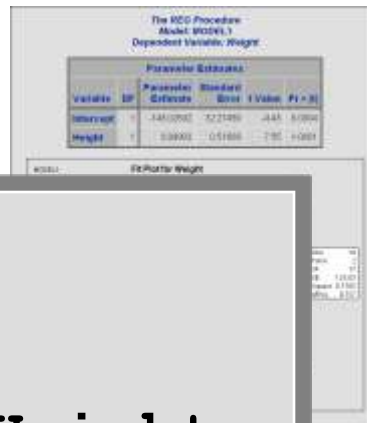
```
ODS GRAPHICS ON;
```

- *Procedure options* determine “which plot I get”
- *Destination* determines “where I see my plots”

```
ODS HTML STYLE=JOURNAL;
```

- *Style* determines “what all my plots look like”

Simple Linear Regression Is Really Simple ...




The screenshot shows the output of the PROC REG procedure in SAS. The title is "The REG Procedure" and the model is "Model: MODEL1" with a dependent variable of "Weight". The output includes a table of parameter estimates for the intercept and height, along with their standard errors, t-values, and p-values.

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	1	-141.0766	52.21491	-2.70	0.0094
Height	1	0.8903	0.51866	1.72	0.0881

```
ods graphics on;
```

```
proc reg;  
  model Weight = Height;  
quit;
```



Demonstration of ODS Graphics

New Procedures in SAS/STAT

The GLIMMIX Procedure

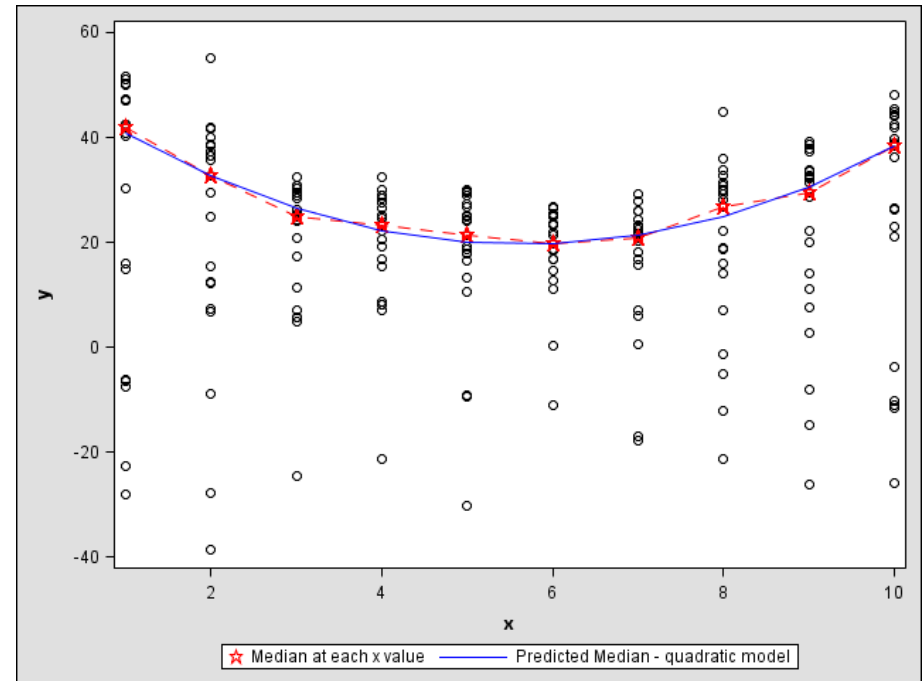
The new GLIMMIX procedure fits generalized linear mixed models, based on linearization.

- For data with correlations or nonconstant variability
- Response is not necessarily normally distributed
- Assumes normal (Gaussian) random effects
- Conditional on these random effects, data can have any distribution in the exponential family
- Default estimation is restricted pseudo-likelihood for models with random effects

The QUANTREG Procedure

Quantile regression extends the regression model to conditional quantiles of the response variable, such as the median or the 90th percentile.

- Makes no distributional assumption or link assumption
- Robust to outliers
- Robust to heterogeneity
- Used as an exploratory tool / extreme value analysis
- Equivariant to monotone transformations



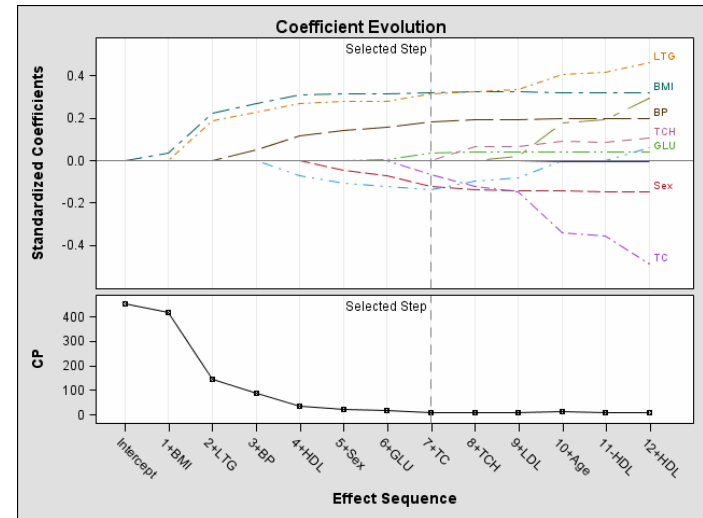


Demonstration of PROC GLIMMIX & QUANTREG

Model Selection with PROC GLMSELECT

Provides great versatility in model selection:

- Multiple selection methods
- Cross-validation
- Graphical aids
- Crossed and nested effects
- Classification effects
- Selection from thousands of effects



New MCMC Procedure

PROC MCMC

- Markov Chain Monte Carlo simulations
- Flexible simulation-based procedure that is suitable for fitting a wide range of Bayesian models
- Uses a random-walk Metropolis algorithm
- Model specification similar to PROC NLIN, and syntax similar to PROC NLMIXED.

New MCMC Procedure

You provide a likelihood function for the data and a prior distribution for the parameters

- You specify hyperparameters if you are fitting a hierarchical model.
- PROC MCMC obtains samples from the posterior distributions, produces summary and diagnostic statistics, and saves the samples to output data sets.
- PREDDIST statement enables you to create random samples from the posterior predictive distribution of the response variables.

Postfitting Bonanza

New shared architecture means more postfitting inference capabilities for many linear modeling procedures.

EFFECTPLOT	GENMOD, LOGISTIC, ORTHOREG
ESTIMATE	LOGISTIC, ORTHOREG, PHREG, SURVEYLOGISTIC
LSMEANS	LOGISTIC, ORTHOREG, PHREG, SURVEYLOGISTIC, SUREVYREG
LSMESTIMATE	GENMOD, LOGISTIC, MIXED, ORTHOREG, PHREG, SURVEYLOGISTIC, SURVEYREG
SLICE	GENMOD, GLIMMIX, LOGISTIC, MIXED, ORTHOREG, SURVEYLOGISTIC, SURVEYREG
TEST	ORTHOREG, SURVEYREG

New PLM Procedure

The new PLM procedure performs postfitting inference on model fit information stored by ten modeling procedures:

- STORE statement saves model info as a SAS item store
- PROC PLM provides the full complement of postfitting inference statements plus EFFECTPLOT.
- Saves computation time and is a tool for confidentiality governance

See **Introducing PROC PLM and Postfitting Analyses in Very General Linear Models** by Randy Tobias and Weijie Cai in the 2010 SGF proceedings.



Demonstration of PROC PLM

Other New Procedures

- Proc HPMIXED - Fits linear mixed models using sparse matrix techniques
- Proc SURVEYPHREG – fits the Cox model for proportional hazards to sample survey data.
- Proc TCALIS – Same analysis as CALIS plus multiple-group analysis, enhanced mean structure analysis, path-like model specification, support of LISREL-type models, and more.

Highlights to Enhancements in SAS/Base & SAS/STAT Procedures

SAS® 9.2 : Enhanced Base Procedures

- PROC SUMMARY/MEANS

- MODE

```
proc means data=tmp1.demo mean median mode  
  var age;  
run;
```

Demography Data

The MEANS Procedure

Analysis Variable : age Age in Years at Baseline

Mean	Median	Mode
51.0392157	49.0000000	45.0000000

Also available in PROC TABULATE, PROC REPORT

SAS® 9.2 : Enhanced Base Procedures

- **PROC FREQ**
- **TABLE** statement options
CROSSLIST

```
proc freq data=demog;  
  table gender*age/crosslist;  
run;
```

NLEVELS

```
proc freq data=demog nlevels;  
  table gender*age;
```

run;

CONTENTS=

```
ods html contents='c:\temp\whatnewc.html';  
proc freq data=demog;  
  table sex*age/contents='Demography Data for Age';  
  table sex*weight/contents='Demography Data for Weight';  
run;  
ods html close;
```



The FREQ Procedure

Table of Gender by Age

Gender	Age	Frequency	Percent	Row Percent	Column Percent
F	12	2	12.50	25.00	40.00
	13	3	12.50	25.00	66.67
	14	2	12.50	25.00	50.00
	15	2	12.50	25.00	50.00
	Total	8	50.00	100.00	
M	12	3	18.75	37.50	60.00
	13	1	6.25	12.50	11.11
	14	2	12.50	25.00	30.00
	15	3	12.50	25.00	30.00
	Total	8	50.00	100.00	
Total	12	9	31.25		100.00
	13	3	18.75		100.00
	14	4	25.00		100.00
	15	4	25.00		100.00
Total		14	100.00		



Number of Variable Levels

Variable	Levels
Gender	2
Age	6

Table of Gender by Age

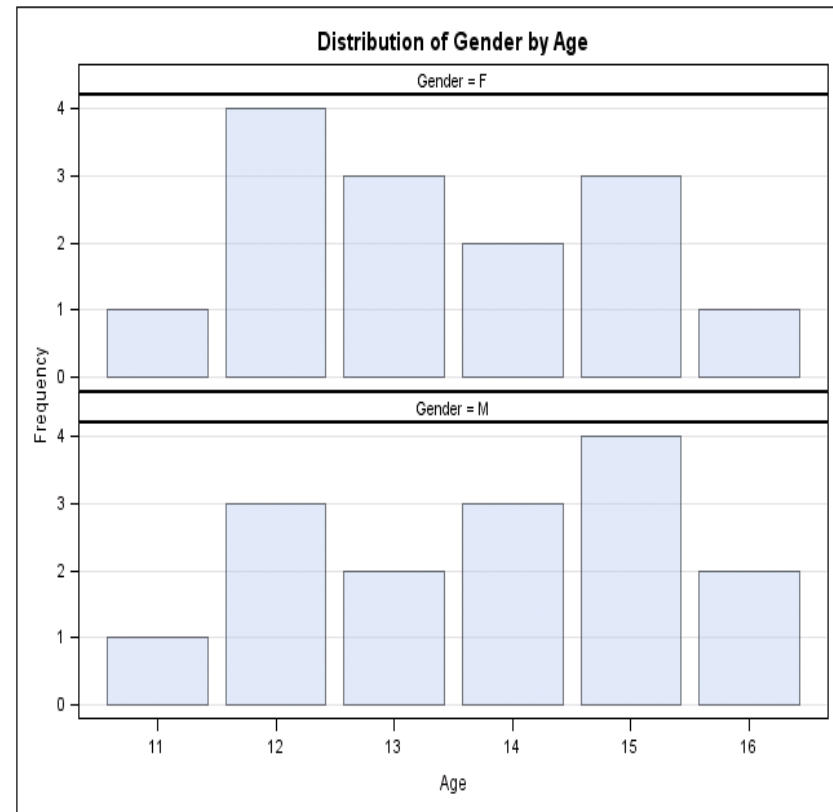
Table of Contents

1. The Freq Procedure
 - Table Sex * Age
 - [Demography Data for Age](#)
 - Table Sex * Weight
 - [Demography Data for Weight](#)

SAS® 9.2 : Enhanced Base Procedures

■ PROC FREQ

- WEIGHT statement options
ZEROS
- New plots produced:
 - Frequency
 - Deviation
 - Odds ratio
 - kappa
- Supports ODS graphics
ods graphics on;
proc freq data=demog;
table gender*age/plots=freqplot;
run;
ods graphics off;



SAS® 9.2 : Enhanced Base Procedures

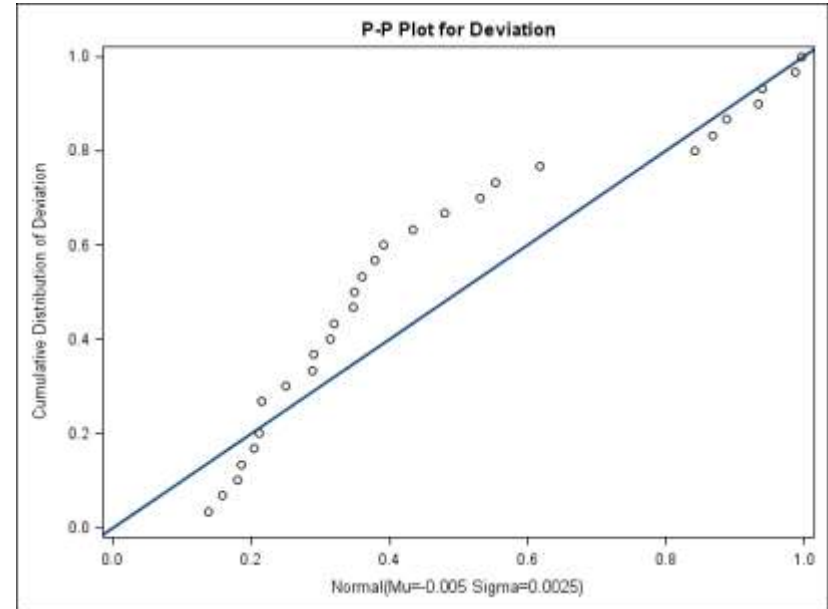
■ PROC UNIVARIATE

- CDFPLOT – Observed Cumulative Distribution Function
- PPLOT – Probability – Probability Plot

■ Both Plots Support:

- Beta
- Exponential
- Gamma
- Lognormal
- Normal
- Weibull
- Supports ODS graphics

```
ods graphics on;  
proc univariate data=aircraft;  
  var deviation;  
  ppplot /normal;  
run;  
ods graphics off;
```



Highlights of SAS/Stat Procedures

Over 200 new features have been added to SAS/STAT for the 9.2 release.

The following slides present highlights for a few procedures.

Highlights: Model Averaging

Model averaging methods provide a way to address some of the issues of automatic variable selection and make more stable inferences.

- Use resampled data as proxy for multiple data sets.
- Select a model from each resampled set of data.
- Average the predictions of these selected models.
- Resampling methods fall under the bootstrap umbrella.

The GLMSELECT procedure provides modeling averaging.

Highlights: Model Averaging - Hald Data

These data have been used to study variable selection methods for years.

- Response is heat given off per gram of cement after 180 days of hardening
- Predictors are percentage by weight of various metals

Woods, Stein, Starke 1932



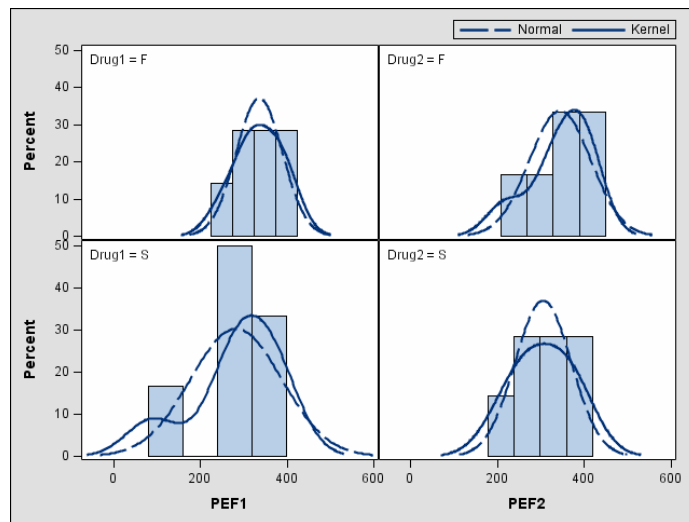
Demonstration of Model Averaging

Highlights: Survey Data Analysis

- JACKKNIFE and BRR variance estimation methods included in analysis procedures
- DOMAIN statement added to the SURVEYREG and SURVEYLOGISTIC procedures
- Allocation capabilities added to the SURVEYSELECT procedure
- percentiles in PROC SURVEYMEANS (Taylor)

Highlights: TTEST

- ODS Graphics – univariate and comparative versions of histograms, box plots, and densities
- SIDES= option for one-sided versions of tests and confidence intervals
- Equivalence tests
- Analysis of simple crossover designs

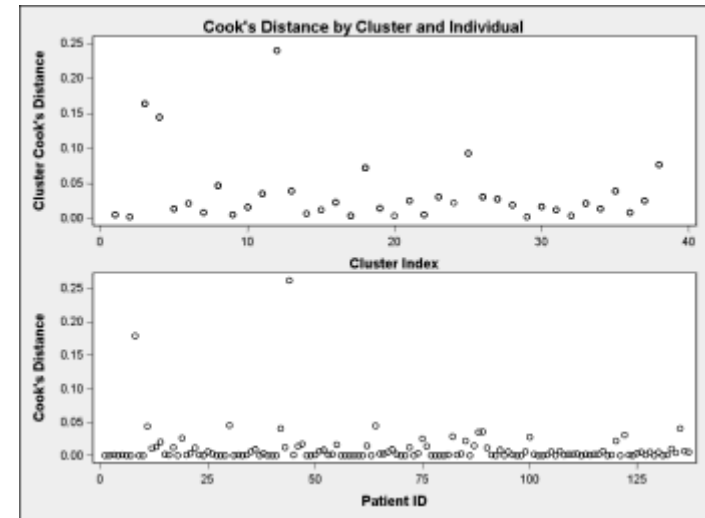


Highlights: PHREG

- Firth's method for handling monotone likelihoods
- Profile likelihood confidence limits for the hazard ratios in the Cox model
- Plot for the baseline survival functions
- CLASS statement
- HAZARDRATIO statement for computing hazards ratios especially in the presence of interactions

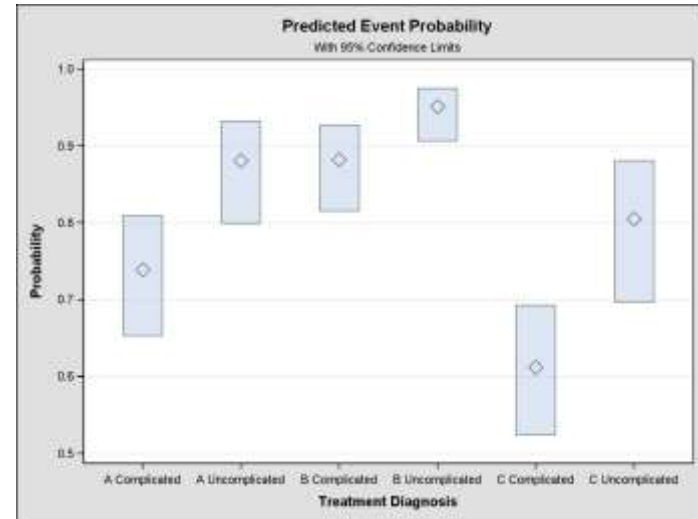
Highlights: GENMOD

- Include inverse link estimates in LSMEANS
- Deletion diagnostics for GLMs and GEEs
- Zero-Inflated Poisson models
- Computes AIC and QIC



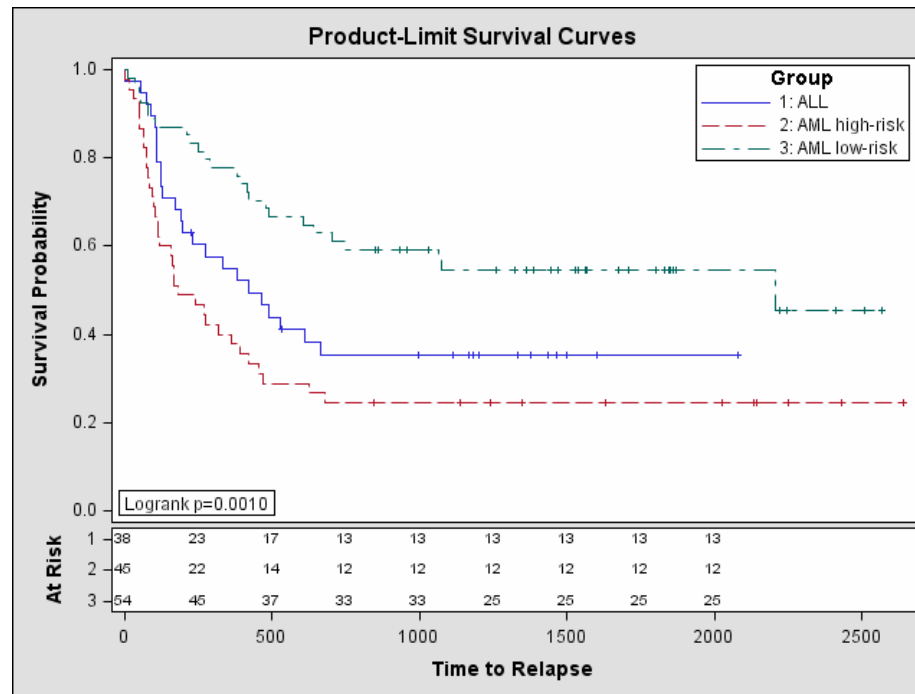
Highlights: LOGISTIC

- Add cumulative probabilities to SCORE output
- Inverse link added to CONTRAST statement
- ROC statement for ROC comparisons
- Odds ratio plots and other plot enhancements
- Fuzzy comparisons in exact computations



Highlights: LIFETEST

- Proc LIFETEST
 - SURVIVAL statement enables the creation of confidence bands for the survivor function $S(t)$
 - Number of subjects at risk can be displayed for Kaplan Meier survival curves
 - Smoother hazard function using the kernel method can be specified



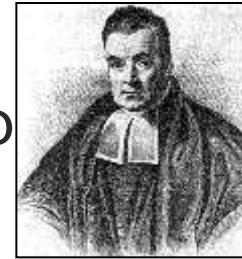


Demonstration of TTEST & LIFETEST

Highlights: Bayesian Analysis

Advantages of Modern Bayesian Methods

- Incorporation of prior information
- Flexibility for model building
- Predictive distributions easier to comp



Costs

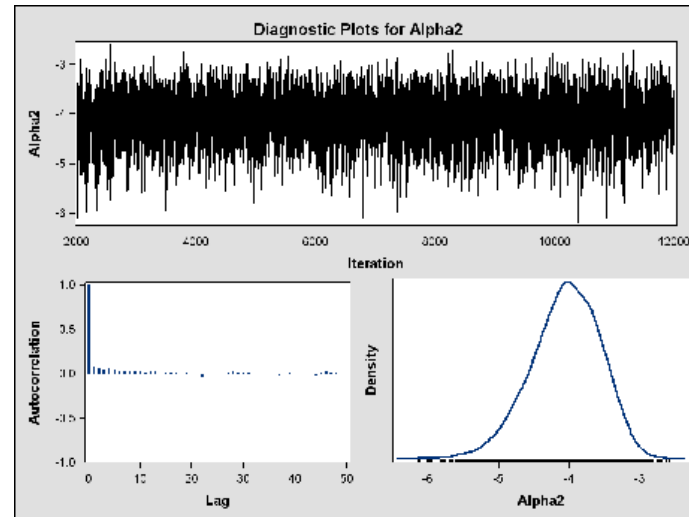
- Computational set-up
- Complex simulations

Thomas Bayes, 1701-1761

Highlights: Bayesian Analysis

In 9.2, Bayesian analyses are available in several SAS procedures.

- PHREG, GENMOD, and LIFEREG procedures
- Uses Gibbs sampler
- Diagnostic graphics



Highlights: Analyses Coverage

Thus, you can now easily obtain Bayesian results for the following analyses in SAS:

- Regression
- Poisson regression
- Logistic regression
- Loglinear models
- Accelerated failure time models
- Proportional hazards models

WHAT'S NEW IN SAS[®]/STAT 9.3



BUT FIRST

...

With the 12.1
release of
SAS/STAT, we turn
to a new release-
numbering scheme

- SAS/STAT released independently of Base SAS®
- All analytical products follow new numbering scheme
- SAS/STAT released every 12–18 months
- SAS/STAT 13.1 next full-featured release

SAS/STAT 12.1

LATEST RELEASE OF SAS/STAT SOFTWARE

SAS/STAT 12.1 became available in the 3rd quarter of 2012.

- Installed with SAS 9.3M2 maintenance release
- Begins new release-numbering protocol
- Builds on the SAS/STAT 9.3 release

- New FMM procedure (experimental) for finite mixture modeling
- Shared frailty models in PHREG procedure
- Expanded capabilities for Bayesian analysis
- Production SURVEYPHREG procedure
- New model diagnostics in NLIN procedure

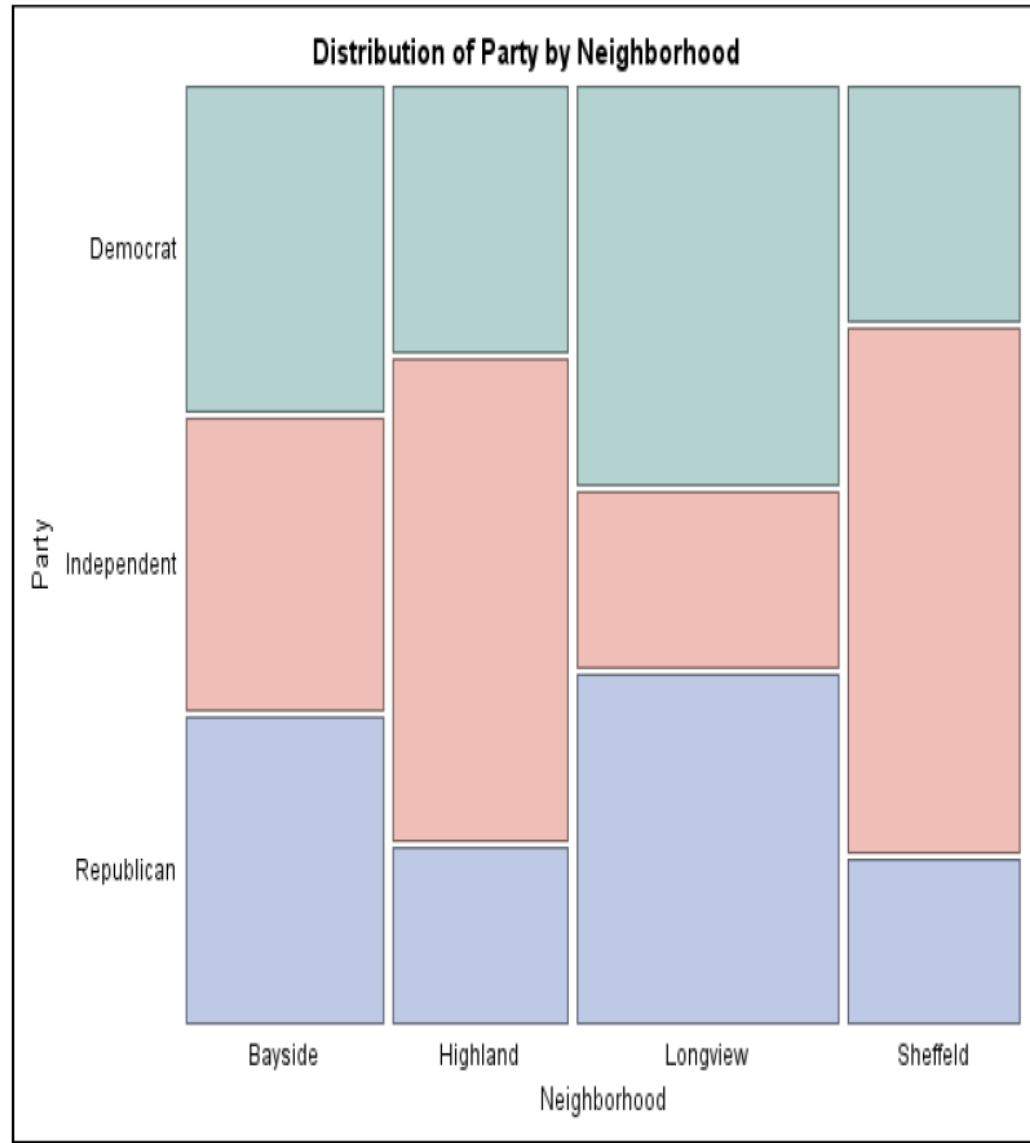
KEY FEATURES OF SAS/STAT 12.1

- Standardized rates and proportions
- Quantile regression: model selection
- Quantile selection: censored data
- Multivariate adaptive regression splines
- Updates to Bayesian software
- Additional highlights

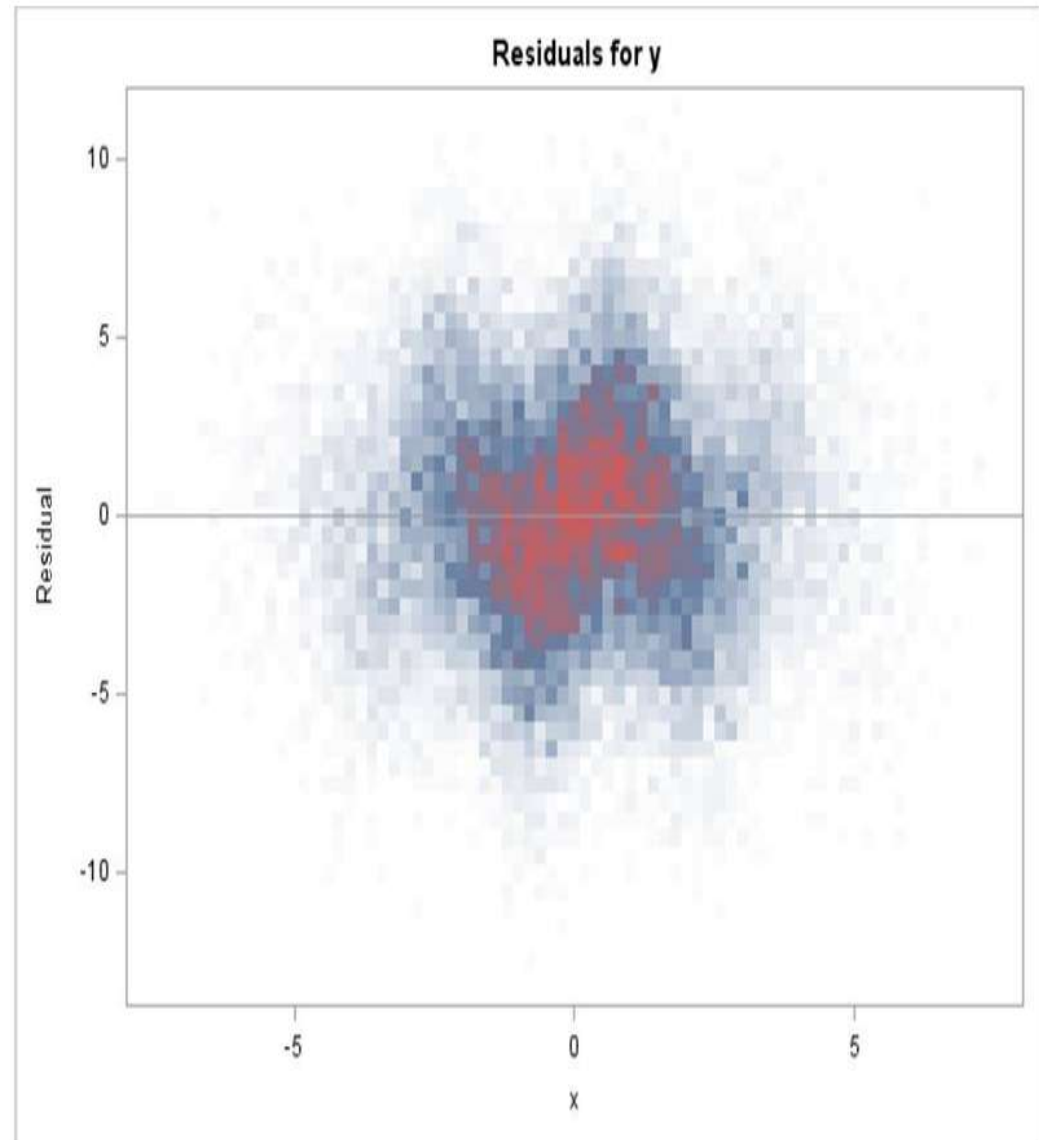
ADDITIONAL HIGHLIGHTS

- PROC FMM becomes production with SAS/STAT 12.1 and adds the truncated normal and truncated negative binomial distributions.
- Partial R-square for relative importance of parameters in PROC LOGISTIC
- Miettinen-Nurminen confidence limits for the difference of proportions in PROC FREQ
- Poisson sampling in PROC SURVEYSELECT
- Post-stratification estimation in the SURVEYMEANS procedure
- REF= option added to the CLASS statement for GLM, MIXED, GLIMMIX, and ORTHOREG procedures
- Heat map and mosaic plots

MOSAIC PLOT



HEAT MAP



RESOURCES



SAS/Stat, SAS/QC, SAS/IML videos

- Youtube.com

<http://www.youtube.com/playlist?list=PL0B05D53A5E101AA6>

- Video portal to the STAT and OR focus area.

<http://support.sas.com/rnd/app/video/index.html>

support.sas.com Resources

- What's New in SAS/Stat 9.22 User Guide

http://support.sas.com/documentation/cdl/en/statug//63347/HTML/default/viewer.htm#/documentation/cdl/en/statug/63347//HTML/default/whatsnew_toc.htm

- What's New in SAS/Stat 9.2 User Guide

<http://support.sas.com/documentation/cdl/en/whatsnew/62580/HTML/default/viewer.htm#/documentation/cdl/en/whatsnew/62580/HTML/default/statugwhatsnew.htm>

On Demand Webinars

- [Getting Started with ODS Statistical Graphics in SAS® 9.2](#)
- [Modifying ODS Statistical Graphics Templates in SAS® 9.2](#)

Current SAS Statistical Resources

- [Stat, IML, OR, ETS Papers](#)
- [Stat and OR e-newsletter](#)
- [Discussion Forum](#)

Links and Other Resources

- Ten Great Reasons Why a Statistician Should Update to SAS® 9.2
http://support.sas.com/rnd/app/da/stat_top10.html
- What's new in SAS®STAT 9.2
<http://support.sas.com/documentation/cdl/en/whatsnew/61982/HTML/default/statugwhatsnew.htm>
- An Introduction to Quantile Regression and the QUANTREG Procedure
<http://www2.sas.com/proceedings/sugi30/213-30.pdf>
- Introducing the GLMSELECT PROCEDURE for Model Selection <http://www2.sas.com/proceedings/sugi31/207-31.pdf>
- Robust Regression and Outlier Detection with the ROBUSTREG Procedure
<http://www2.sas.com/proceedings/sugi27/p265-27.pdf>
- An introduction to partial least squares regression
<http://support.sas.com/techsup/technote/ts509.pdf>
- Sample-Size Analysis in Study Planning: Concepts and Issues, with Examples Using PROC POWER and PROC GLMPOWER
<http://www2.sas.com/proceedings/sugi29/211-29.pdf>
- Updates to SAS® Power and Sample Size Software in SAS/STAT® 9.2
<http://www2.sas.com/proceedings/forum2008/368-2008.pdf>
- A Comparison of the Mixed Procedure and the Glimmix Procedure
<http://www2.sas.com/proceedings/sugi31/189-31.pdf>
- Introducing the GLIMMIX Procedure for Generalized Linear Mixed Models
<http://www2.sas.com/proceedings/sugi30/196-30.pdf>
- Growing Up Fast: SAS 9.2 Enhancements to the GLIMMIX Procedure
<http://www2.sas.com/proceedings/forum2007/177-2007.pdf>
- Old versus New: A Comparison of PROC LOGISTIC and PROC GLIMMIX
<http://www2.sas.com/proceedings/forum2008/226-2008.pdf>
- Advanced Statistical and Graphical features of SAS® PHREG
<http://www2.sas.com/proceedings/forum2008/375-2008.pdf>
- SAS Online Help for SAS®STAT, Chapter 7: Introduction to Bayesian Analysis Procedures
- Data Preparation for Analytics
http://www.sascommunity.org/wiki/Data_Preparation_for_Analytics



THE
POWER
TO KNOW.