What's new in IBM SPSS Statistics 29

Highlights

New procedures including OLS regression and AFT models

Support for Psuedo-R-squared measures

Enhancements to existing procedures

Improvements to workbook mode and usability

IBM® SPSS® Statistics is a comprehensive software platform that addresses all facets of the analytical process from data preparation and management to analysis and reporting. You can improve forecasts and plans by inputting missing values with expected values, deliver tables and visualizations to communicate results effectively, classify cases into groups and predict values of target variables based on values of predictor variables, and accurately model linear and non-linear relationships. Organizations of all types have relied on IBM SPSS Statistics to increase revenue, expand capabilities, conduct research, and drive decision-making with data.

IBM SPSS Statistics 29 includes new customer-requested linear ordinary least squares (OLS) regression and parametric accelerated failure time (AFT) model statistical procedures, improved open-source extension integration, UI enhancements, new data visualization features, and other enhancements that are designed to improve everyday usability.

The perpetual editions of SPSS Statistics 29 (Base, Standard, Professional and Premium) deliver the long-term software access required for analytics groups. These editions group essential features, functionality, and usage requirements to offer a convenient way to acquire the capabilities you need. For individuals and organizations that require pay-as-you-go software usage, SPSS Statistics Subscription and its optional add-ons provide the desired flexibility to pay on a monthly basis. The latest features and functionality are available in both the IBM SPSS Statistics Subscription and the on-premises versions of IBM SPSS Statistics 29.

New procedures

Linear OLS alternatives

1. Elastic net

The new linear elastic net extension procedure estimates regularized linear regression models for a dependent variable on one or more independent variables. The graph represents example output for the procedure.

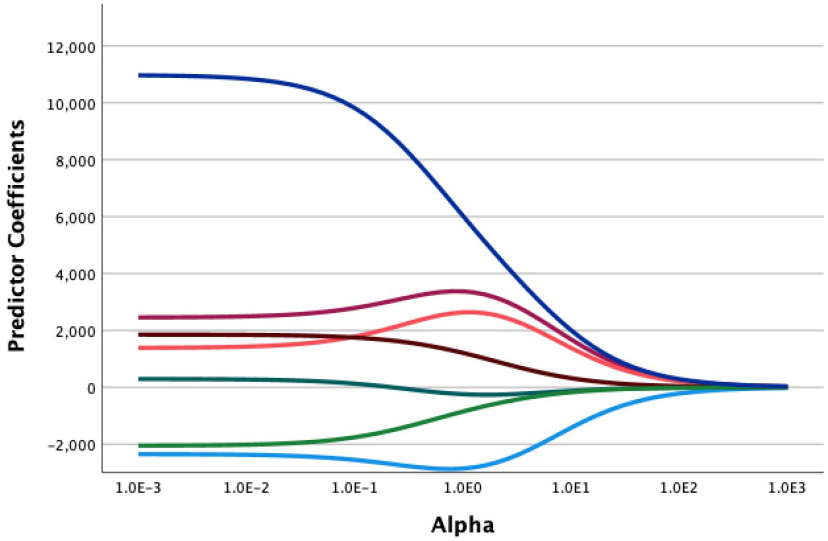

Figure 1. Elastic net example

2. Lasso

The new linear lasso extension estimates L1 loss regularized linear regression models for a dependent variable on one or more independent variables, and includes optional modes to display trace plots and to select the alpha hyperparameter value based on cross validation. The graph represents example output for the procedure.

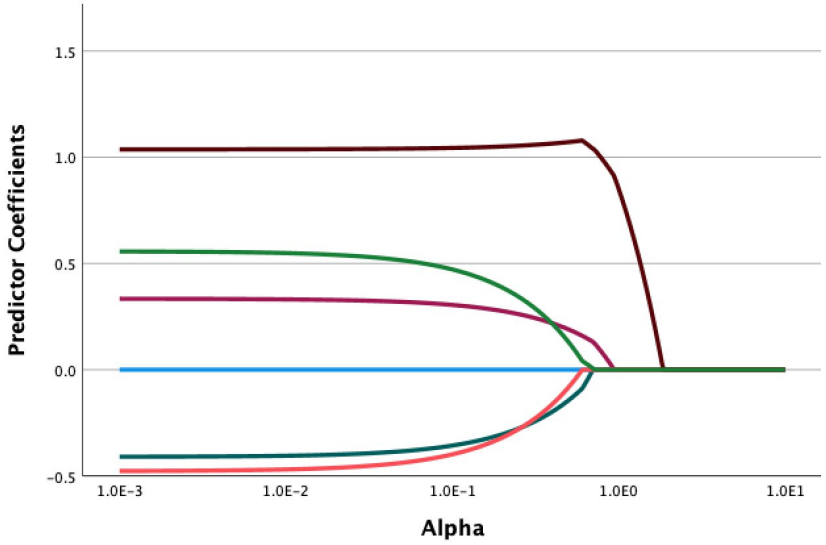

Figure 2. Lasso example

3. Ridge

The new linear ridge extension procedure estimates L2 or squared loss regularized linear regression models for a dependent variable on one or more independent variables, and includes optional modes to display trace plots and to select the alpha hyperparameter value based on cross validation. The graph represents example output for the procedure.

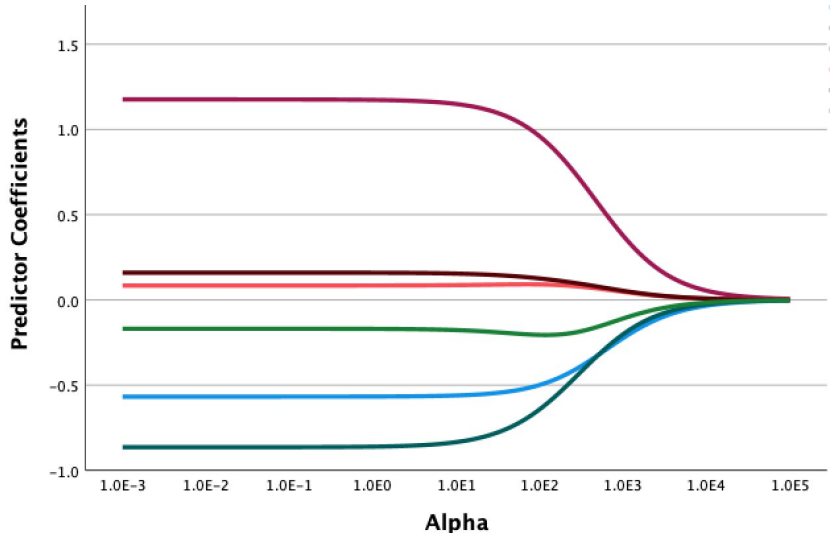

Figure 3. Ridge example

4. Parametric AFT models

The new procedure invokes the parametric survival models procedure with non-recurrent life time data. Parametric survival models assume that survival time follow a known distribution and this analysis fits accelerated failure time models with their model effects proportional with respect to survival time. The graph represents example output for the procedure.

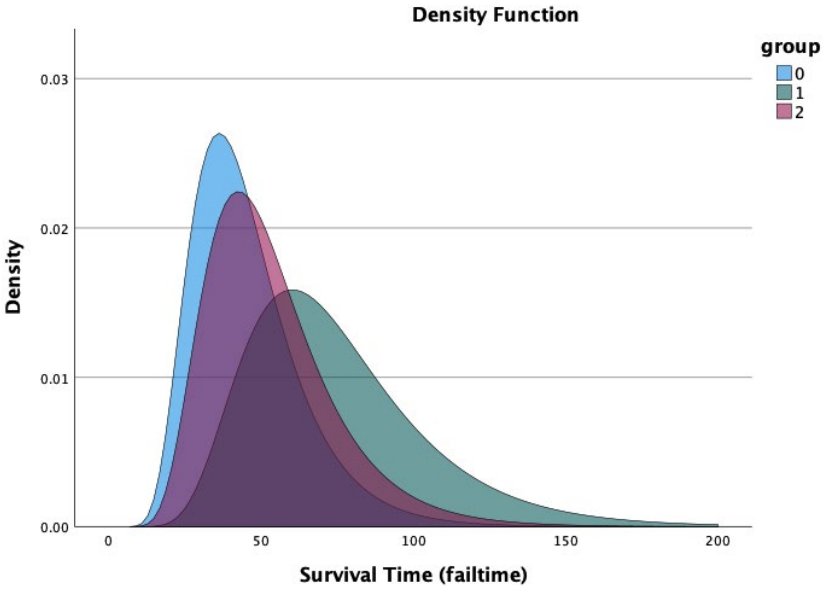

Figure 4. Parametric AFT survival model example

Pseudo-R-squared measures

Pseudo-R-squared measures and the intra-class correlation coefficient are now included in linear mixed models and generalized linear mixed models output when appropriate. The coefficient of determination R2 is a commonly reported statistic, because it represents the proportion of variance explained by a linear model. The intra-class correlation coefficient (ICC) is a related statistic that quantifies the proportion of variance explained by a random grouping factor in multilevel and hierarchical data.

Enhancements to existing procedures

Violin plots

The Graphboard Template Chooser includes a new violin plot, which is a hybrid of the box and kernel density plots. Violin plots show peaks in the data and are used to visualize the distribution of numerical data. Unlike a box plot that can only show summary statistics, violin plots depict summary statistics and the density of each variable.

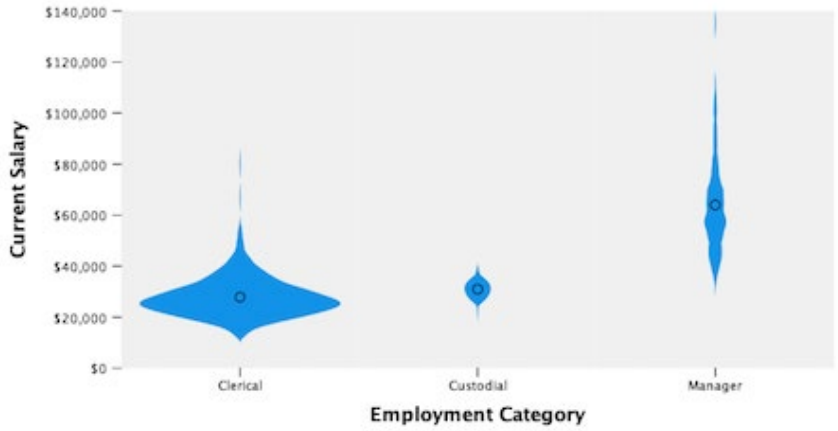

Figure 5. Violin plot example

Usability improvements

Workbook mode enhancements
Two new workbook toolbar items have been added: Show/Hide all syntax windows and Clear all output. There is also a new button on the status bar to switch between classic (output and syntax) and workbook modes.

Search enhancements
The search feature now provides options for entering terms directly in a toolbar field and for viewing results in a drop-down pane.

Removal of ability to hide unselected cases
Unselected cases are no longer hidden in the data editor when a subset of cases is selected and the unselected cases are not discarded. This represents a return to the behavior of Statistics 27.0.1 and earlier versions.

Python and R upgrades
Python 3.10.4 and R 4.2.0 are installed with IBM SPSS Statistics 29.

Conclusion

IBM SPSS Statistics 29 includes new customer-requested linear OLS and parametric AFT model procedures, improved open-source integration, usability improvements and enhancements to existing procedures. IBM SPSS Statistics enables organizations to gather rich insights from data with a powerful set of tools to validate assumptions, forecast trends and analyze performance. To ensure that the most advanced techniques are available to a broader group of analysts and business users, enhancements have been made to the features and capabilities of the IBM SPSS Statistics portfolio.

For more information

To learn more about IBM SPSS Statistics please contact your IBM representative or IBM Business Partner, or visit ibm.com/products/spss-statistics.

IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the
United States of America
September 2022