

IBM SPSS STATISTICS 29

NOVAS FUNCIONALIDADES

Na sequência do anúncio internacional pela IBM, a PSE vem apresentar as novas funcionalidades incluídas na versão 29 do software IBM SPSS Statistics.

Esta versão traz consigo novos procedimentos estatísticos; melhorias em procedimentos existentes; melhor integração com as linguagens *Open Source R* e *Python* e melhorias ao nível de usabilidade.

De seguida apresentamos as novidades introduzidas nesta versão.

Novas análises

Regularized Linear Regression models

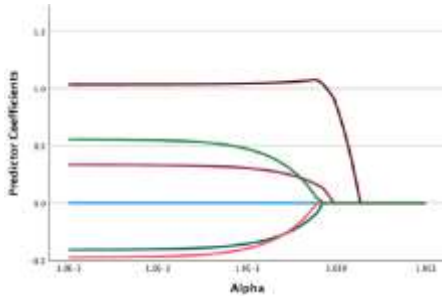
Frequentemente enfrentamos o problema de os nossos modelos apresentarem um bom desempenho nos dados de treino, mas apresentarem um desempenho fraco em dados novos. Este fenómeno é conhecido na literatura por *overfitting*.

A 'Regularização' é uma técnica para reduzir o *overfitting*. Ela atua impondo restrições aos coeficientes do modelo de modo a diminuir a influência de certas variáveis; penalizar modelos mais complexos e, desta forma, evitar o *overfitting*.

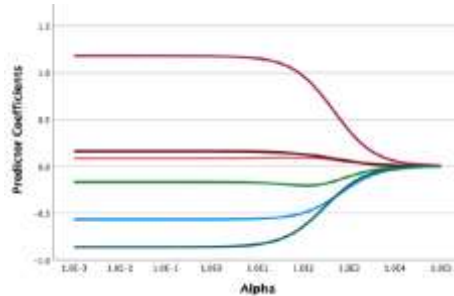
O IBM SPSS Statistics 29 introduz três técnicas de regularização para os modelos de Regressão Linear:

- **Lasso Regression (L1)**: em que é aplicada uma penalização em função da soma dos coeficientes da regressão (exceto o da interseção) e de um parâmetro *Alpha*.
- **Ridge Regression (L2)**: em que é aplicada uma penalização em função do quadrado dos coeficientes da regressão (exceto o da interseção) e de um parâmetro *Alpha*.
- **Elastic Net**: em que a penalização é obtida por uma combinação dos métodos anteriores.

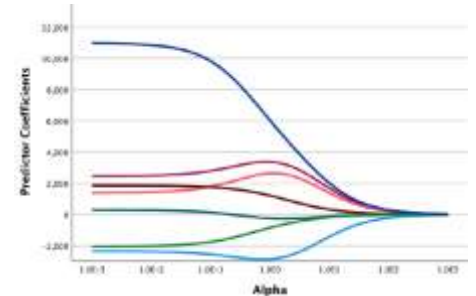
Exemplo do output deste novo procedimento como as três técnicas de Regularização



Lasso



Ridge

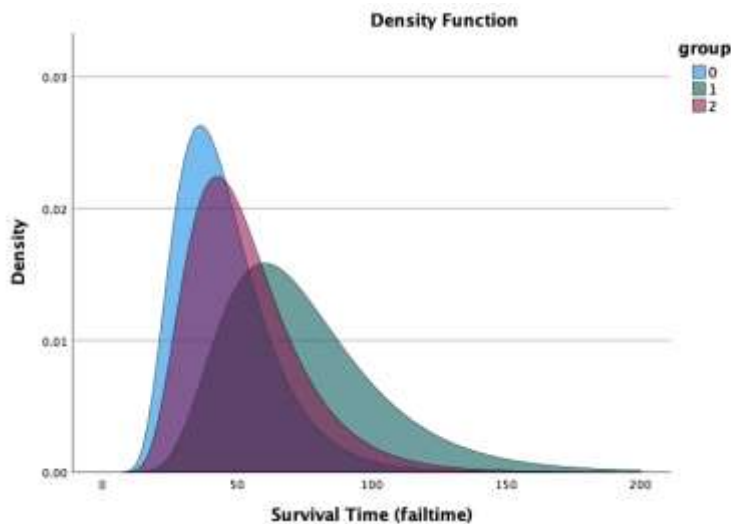


Elastic Net

Parametric Accelerated Failure Time (AFT) Models

Os modelos de sobrevivência são amplamente usados na pesquisa na área da saúde. Os modelos Accelerated Failure Time (AFT) são uma alternativa paramétrica aos clássicos modelos de Kaplan-Meier e Cox. Num modelo AFT o efeito das covariáveis atua acelerando ou desacelerando o ciclo de vida de uma doença, ou seja, aumenta ou diminui esse mesmo ciclo. Estes modelos assumem que os dados seguem uma distribuição conhecida.

Exemplo do output de um modelo AFT

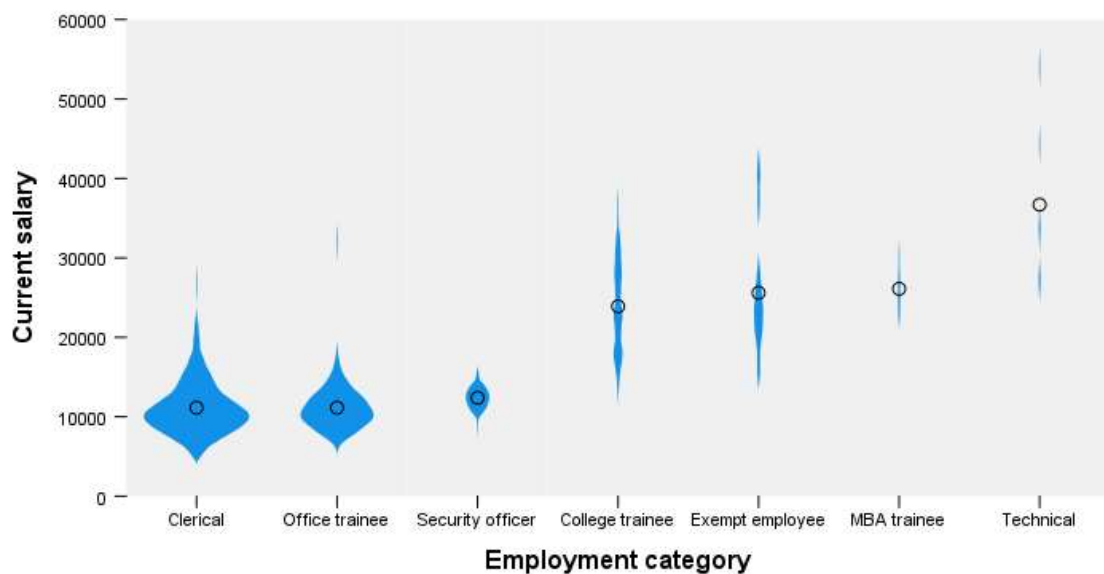


Violin Plots

Nesta versão o *Graphboard Template Chooser* foi enriquecido com um novo tipo de gráfico: o *Violin Plot*.

Um *Violin Plot* é um híbrido entre *Box Plots* e os *Kernel Density Plots*. Os *Violin Plot* são bastante usados para visualizar a distribuição de dados numéricos. Para além de apresentarem estatísticas sumárias - como os *Box Plots* - os *Violin Plot* apresentam ainda a densidade de probabilidade de uma variável de interesse como os *Kernel Density Plots*.

Exemplo de um Violin Plot



Melhorias em procedimentos

Medidas de Pseudo-R² nos Linear Mixed Models e Generalized Linear Mixed Models

O IBM SPSS Statistics 29 introduz o cálculo de medidas de *Pseudo-R²* e das Correlações Intra-Classe no output dos *Linear Mixed Models* e *Generalized Linear Mixed Models* (quando apropriado).

O Coeficiente de determinação é uma estatística reportada frequentemente pelos investigadores pois representa a proporção da variância da variável dependente que é explicada pelo modelo. O Coeficiente de Correlação Intra-classe (ICC) é uma estatística relacionada que quantifica a proporção de variância explicada pelo fator de agrupamento em dados hierárquicos.

Exemplo dos novos outputs de um *Linear Mixed Model*

Coefficients of Determination		
Pseudo-R Square Measures	Marginal	,313
	Conditional	,656

Medidas de Pseudo-R²

Intraclass Correlation Coefficients		
Overall ICCs	Adjusted	,500
	Conditional	,344

Coefficiente de Correlação Intra-classe

Melhorias na experiência de utilização

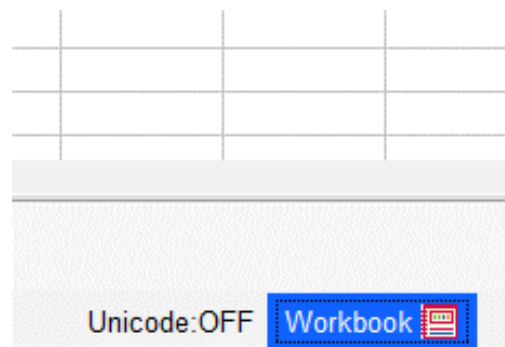
Statistics Workbook

O *Statistics Workbook* foi introduzido na versão 28. No IBM SPSS Statistics 29 esta funcionalidade é estendida. Foram introduzidos dois novos itens na barra de ferramentas do *Workbook*: *Show/Hide all syntax windows* e *Clear all output*. Também foi adicionado à barra de estado um novo botão que permite alternar entre o modo clássico (ficheiros de *Output* e de *Syntax* separados) e o novo modo de *Workbook* (um só ficheiro que junta *Outputs* e *Syntax*).

As novas funcionalidades do modo *Workbook*



Novos botões na barra de ferramentas



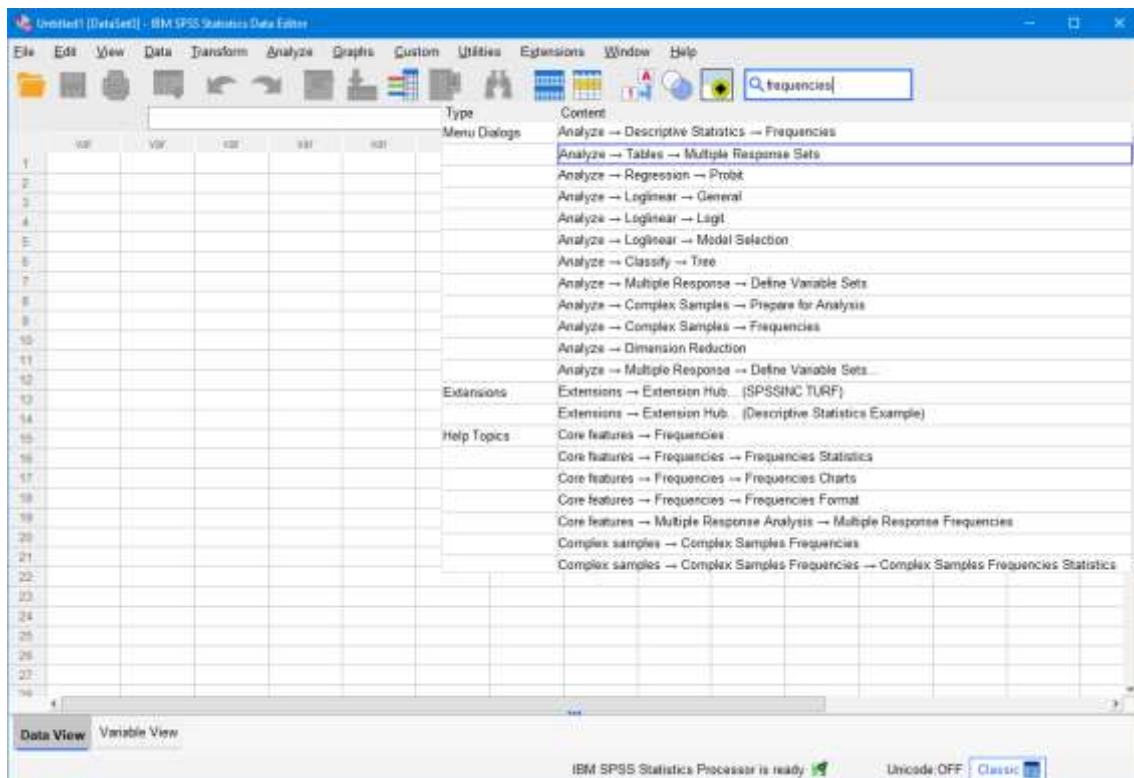
Botão para alternar entre modos

Search

A ferramenta de Pesquisa é bastante útil no trabalho do dia-a-dia pois permite encontrar, de forma rápida e fácil, vários tipos de recursos como menus, comandos de sintaxe, extensões, tópicos de ajuda ou *case studies*.

De modo a tornar a sua utilização ainda mais intuitiva é agora possível introduzir os termos que pretende pesquisar diretamente num campo na barra de ferramentas e visualizar os resultados num menu do tipo *drop-down*.

A nova caixa de pesquisa integrada com a barra de ferramentas



Outras melhorias

Integração com *R* e *Python*

Quando instala o IBM SPSS Statistics no seu computador serão automaticamente instaladas as versões 4.2.0 da linguagem *R* e 3.10.4 da linguagem *Python* o que contribui para melhorar o suporte a extensões *open source*.

Seleção de casos

Com esta nova versão, quando seleciona um conjunto de casos, os não selecionados deixam de ser escondidos no *Data View* sendo representados com um traço na diagonal no respetivo número de linha.

Outras características importantes

O SPSS Statistics 29 é apenas disponibilizado na edição de 64 bits. Necessita do sistema operativo Microsoft Windows 10 ou 11 ou macOS 10.13 a 12.0.

Lisboa, Outubro de 2022

IBM E SPSS SÃO MARCAS REGISTRADAS DA IBM CORPORATION

A PSE DISTRIBUI ESTA INFORMAÇÃO NA SUA QUALIDADE DE PARCEIRO DE NEGÓCIO IBM