

Integração do R nos menus do *PASW Statistics*¹:

Um exemplo de aplicação com o *package* ‘polycor’ do R

João Maroco, jpmaroco@ispa.pt

Unidade de Investigação em Psicologia e Saúde; Departamento de Estatística.

ISPA - Instituto Universitário

SUMÁRIO

O Coeficiente de correlação policórica tem, recentemente, vindo a ter utilização crescente na estimação da associação entre variáveis de natureza ordinal (e.g. níveis sócio-económicos; severidade de diagnóstico médico; itens de tipo Likert em escalas psicométricas, etc...). Contudo, este tipo de correlação não está ainda disponível nos principais softwares *user-friendly* de análise estatística, nomeadamente no *PASW Statistics* (até à v. 18, inclusive). Neste artigo será apresentado o coeficiente de correlação policórica, a sua implementação no *package* ‘Polycor’ do sistema R e a sua integração nos menus do *PASW Statistics* por recurso ao ‘PASW Custom Dialog Builder’. Este aplicativo permite disponibilizar as bibliotecas do R nos menus do *PASW*, libertando assim o utilizador final da necessidade de dominar a programação em R.

1. INTRODUÇÃO

As variáveis de natureza ordinal são relativamente comuns nas ciências sociais, ciências biomédicas e ciências da engenharia. Estas variáveis, de acordo com a definição original de Stevens (1946) reflectem apreciações qualitativas, com propriedades de grandeza ordenável, mas cuja, eventual, codificação numérica é desprovida de significado quantitativo. Exemplos típicos destas variáveis incluem as variáveis de tipo Likert (e.g., Concordância com políticas sociais [‘1-Discordo completamente’, ‘2-Discordo’, ‘3-Nem discordo nem concordo’, ‘4-Concordo’ e ‘5 - Concordo completamente’]; Severidade de dor [‘3-Dói muito; 2-Dói; 1-Não Dói]) e as variáveis ordinais, designadas por *rating scales*, que reflectem características de intensidade crescente (e.g. Fissuras em betão [Sem fissuras, algumas fissuras, muitas fissuras] e Dureza de minerais [1 – Talco; 2 – Gesso, 3 – Calcite; ... ; 10 – Diamante]). A estimação da associação entre variáveis ordinais é, tradicionalmente, feita por recurso ao coeficiente de correlação de Spearman. Esta medida, desenvolvida pelo psicometrista inglês Charles Spearman avalia quão bem uma função monótona arbitrária (ordenação) é capaz de descrever a associação

¹ O software R é uma marca registada de ‘The R Foundation’; O software *PASW Statistics* é uma marca registada de ‘SPSS, An IBM Company’.

entre duas variáveis cujas realizações são ordenáveis, sem contudo, apresentar algum tipo de assunção sobre a natureza dessa associação. Se X e Y representarem duas variáveis cujas ordens são R_X e R_Y respectivamente, o coeficiente de correlação de Spearman pode formular-se como:

$$r_S = \frac{\sum_{i=1}^n (R_{X_i} - \bar{R}_X)(R_{Y_i} - \bar{R}_Y)}{\sqrt{\sum_{i=1}^n (R_{X_i} - \bar{R}_X)^2 \sum_{i=1}^n (R_{Y_i} - \bar{R}_Y)^2}} \quad (1)$$

O coeficiente de correlação de Spearman é frequentemente usado em estudos de natureza descritiva e inferencial. Contudo, a sua utilização em métodos correlacionais multivariados é pouco frequente (por exemplo, na Análise Factorial de escalas psicométricas, uma técnica correlacional multivariada proposta pelo próprio C. Spearman). Mas, a expressão (1) é, simplesmente, o ubíquo coeficiente de Correlação de Pearson aplicado às ordens das variáveis. O coeficiente de correlação de Pearson, estimado por

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (Y_i - \bar{Y})^2}} = \frac{Cov(X, Y)}{S_X S_Y} \quad (2)$$

é também frequentemente usado com variáveis ordinais quando: (i) as classes destas variáveis são codificadas numericamente, (ii) se assume uma ordem implícita entre as codificações numéricas e (iii) se assume uma relação linear entre estas ordens. Contudo, a utilização do coeficiente de correlação de Pearson com variáveis ordinais é controversa. Em primeiro lugar o cálculo deste coeficiente requer o cálculo das médias das variáveis o que exige uma medida de natureza pelo menos intervalar. O cálculo de médias e desvios-padrão com variáveis ordinais é, porém, uma prática frequente em algumas áreas de estudo, e.g. na Psicometria, ainda que não esteja isenta de criticismo. Como refere Stevens: “As a matter of fact, most of the scales used widely and effectively by psychologists are ordinal scales. In the strictest propriety the ordinary statistics involving means and standard deviations ought not to be used with these scales, for these statistics imply a knowledge of something more than the relative rank order of data” (Stevens, 1946, p. 679). Vários estudos de simulação (para uma revisão ver, e.g., Bollen, 1989, p. 434-435) têm demonstrado que as estimativas da associação entre variáveis ordinais, obtidas com a correlação de Pearson, são geralmente inferiores às verdadeiras associações entre as variáveis quantitativas de cuja discretização resultaram as variáveis ordinais simuladas. Esta atenuação é tanto maior quanto menor for o número de classes (menos de 5) e quanto maior for a oposição das assimetrias das variáveis. À medida que o número de classes aumenta (7, 10, ...), a medida aproxima-se mais de uma métrica quantitativa, e as correlações de Pearson calculadas

para variáveis ordinais aproximam-se das correlações obtidas para as variáveis quantitativas correspondentes (Bollen, 1989, p. 435; Finney e Distefano, 2006, p. 276). Mais recentemente, alguns autores têm proposto que as análises correlacionais (Análise factorial, modelos de equações estruturais, etc....) com variáveis ordinais sejam baseadas não nos usuais coeficientes de correlação de Pearson e ou Spearman, mas sim no coeficiente de correlação policórica. Num estudo de simulação de larga escala, Babakus, Ferguson, & Jöreskog (1987) observaram que o estimador da correlação policórica produziu as melhores estimativas das correlações bivariadas entre variáveis ordinais, comparativamente às estimativas obtidas com os coeficientes de Pearson, Spearman ou τ de Kendall.

2. O COEFICIENTE DE CORRELAÇÃO POLICÓRICA

De uma forma geral, as escalas psicométricas produzem resultados que se podem assumir como estimativas quantitativas de habilidades cognitivas. Assim, os itens ordinais da maioria das escalas psicométricas, podem assumir-se como ‘operacionalizações’ de variáveis latentes (as habilidades) contínuas de natureza pelo menos intervalar, que só podem ser estimadas por intermédio de avaliações ordinais. De forma a aceder à verdadeira associação entre variáveis latentes, de cujos itens ordinais são manifestações, é necessário estimar a correlação não entre os itens, mas sim entre as variáveis latentes. É esta associação que o coeficiente de correlação policórica estima. O estimador do coeficiente de correlação policórica foi desenvolvido a partir dos trabalhos seminais de Karl Pearson que reconheceu a facilidade de operacionalização de variáveis latentes por intermédio de itens ordinais. Conceptualmente, uma variável ordinal X com 5 categorias pode ser interpretada como o resultado da divisão de uma variável, subjacente ou latente, ξ em 5 categorias. Por exemplo, um item de uma escala de satisfação com 5 classes de ‘1–Nada satisfeito’, a ‘5– Muito satisfeito’, pode ser conceptualizado como

$$X = \begin{cases} 1 - \text{Nada Satisfeito} & \text{se } \xi \leq \xi_1 \\ 2 - \text{Pouco Satisfeito} & \text{se } \xi_1 < \xi \leq \xi_2 \\ 3 - \text{Indiferente} & \text{se } \xi_2 < \xi \leq \xi_3 \\ 4 - \text{Satisfeito} & \text{se } \xi_3 < \xi \leq \xi_4 \\ 5 - \text{Muito Satisfeito} & \text{se } \xi > \xi_4 \end{cases} \quad (3)$$

Assim, as 5 categorias de resposta podem ser interpretadas como uma aproximação aos 5 intervalos em que a variável latente ξ , contínua, foi dividida por recurso a 4 pontos de corte ou *thresholds*: ξ_1 , ξ_2 , ξ_3 e ξ_4 sendo, por convenção, $\xi_0 = -\infty$ e $\xi_5 = +\infty$. A correlação policórica estima a associação entre duas variáveis latentes, que se assumem com distribuição normal bivariada, subjacentes a duas variáveis ordinais manifestas. A figura 1 ilustra este conceito.

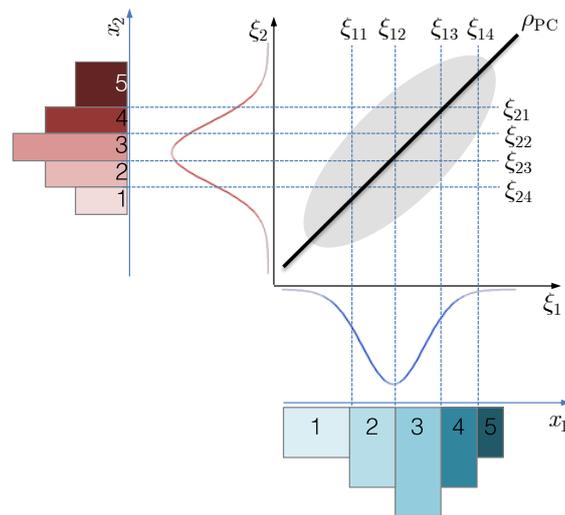


Figura 1 - Ilustração da correlação policórica entre duas variáveis contínuas latentes (ξ_1 e ξ_2) operacionalizadas por duas variáveis ordinais manifestas (X_1 e X_2) cada uma com 5 classes definidas pelos pontos-de-corte $\xi_{11}, \dots, \xi_{14}$ e $\xi_{21}, \dots, \xi_{24}$. O coeficiente de correlação policórica é ilustrado pela linha ρ_{PC} .

O coeficiente de correlação policórica pode estimar-se pelo método de máxima verosimilhança (ver, e.g., Drasgow, 2006). Neste método, a probabilidade conjunta (P_{ij}) de se observar o valor x_{1i} para a variável ξ_1 e o valor x_{2j} para a variável ξ_2 é estimada por

$$P_{ij} = \int_{\xi_{i-1}}^{\xi_i} \int_{\xi_{j-1}}^{\xi_j} \phi(\xi_1, \xi_2; \rho) d\xi_2 d\xi_1 \quad (4)$$

onde

$$\phi(\xi_1, \xi_2; \rho) = \frac{1}{2\pi\sqrt{1-\rho^2}} \times e^{\left(\frac{-1}{2(1-\rho^2)}(\xi_1^2 - 2\rho\xi_1\xi_2 + \xi_2^2)\right)} \quad (5)$$

é a função de densidade normal bivariada de ξ_1 e ξ_2 com, sem perda de generalidade, $\mu=0$, $\sigma=1$ e correlação de Pearson ρ . Se n_{ij} for o número de observações x_{1i} da variável X_1 e de x_{2j} na variável X_2 , a verosimilhança das observações amostrais é:

$$L = k \prod_{i=1}^r \prod_{j=1}^s P_{ij}^{n_{ij}} \quad (6)$$

onde k é uma constante e r e s são o número de classes de x_1 e x_2 , respectivamente. A estimativa de máxima verosimilhança do coeficiente de correlação policórica (ρ_{PC}) obtém-se derivando o $\ln(L)$ em ordem a ρ

$$\frac{\partial \ln(L)}{\partial \rho} = \sum_{i=1}^r \sum_{j=1}^s \frac{n_{ij}}{P_{ij}} \left[\phi(\xi_{1i}, \xi_{2j}; \rho) - \phi(\xi_{1i-1}, \xi_{2j}; \rho) \right] - \phi(\xi_{1i}, \xi_{2j-1}; \rho) + \phi(\xi_{1i-1}, \xi_{2j}; \rho) \quad (7)$$

Igualando a derivada parcial (7) a 0 e resolvendo a equação obtida em ordem a ρ obtém-se a estimativa de ρ_{PC} . Contudo, e de uma forma geral, a solução do estimador de máxima verosimilhança é obtida de forma iterativa já que resolução da derivada parcial em ordem a ρ exige o conhecimento dos restantes parâmetros da função de verosimilhança. A estimação destes parâmetros, por sua vez, exige a derivada parcial de $\ln(L)$ em ordem a cada um dos coeficientes de (6) (ρ , $\xi_{11}, \dots, \xi_{1r}$, $\xi_{21}, \dots, \xi_{2s}$) e a resolução do sistema com todas as derivadas parciais iguais a 0. A correlação policórica pode também estimar-se com outro algoritmo, computacionalmente mais rápido, designado por algoritmo *two-step* (Martinson & Hamdan, 1975). Neste algoritmo, começa-se por ajustar distribuições normais univariadas às distribuições marginais de x_1 e x_2 , estimando os *thresholds* ($\xi_{11}, \dots, \xi_{1r}$, $\xi_{21}, \dots, \xi_{2s}$) para cada distribuição. No passo final, a estimativa de ρ_{PC} é obtida derivando a função $\ln(L)$ em ordem a ρ , igualando a

derivada 0 e resolvendo a equação em ordem a ρ (Dragow, 2006; Martinson & Hamdan, 1975).

Se uma das variáveis for quantitativa e a outra for ordinal, a correlação entre as duas obtêm-se estimando a variável latente subjacente à variável ordinal, como anteriormente, calculando-se, em seguida, a sua associação com a variável quantitativa. Esta associação é conhecida por ‘correlação poliserial’.

3. O PACKAGE *POLYCOR* DO SISTEMA *R*

O coeficiente de correlação policórica pode obter-se quer em *software* comercial [PRELIS/LISREL, EQS, MPLUS, STATA 8 (programa de Stas Kolenikov), SAS (Polychoric Macro)] quer no sistema *R* com o *package* ‘Polycor’ da autoria de John Fox (jfox@mcmaster.ca) e disponível nos repositórios do CRAN. O *package* ‘Polycor’ permite calcular correlações policóricas, poliseriais (correlação entre uma variável ordinal e uma variável quantitativa) com estimativas opcionais dos erros-padrão e testes à distribuição normal multivariada. O *package* é composto por 3 programas. O programa ‘HetCor’ produz uma matriz de correlações constituída por correlações de Pearson (para variáveis numéricas), correlações policóricas para (variáveis ordinais – *factors*) e correlações poliseriais (para variáveis numéricas vs. variáveis ordinais). O programa ‘Polychor’ calcula as correlações policóricas e respectivos erros-padrão entre duas variáveis ordinais a partir dos valores originais ou de uma tabela de contingência. Finalmente, o programa ‘Polyserial’ calcula a correlação poliserial e o erro-padrão respectivo. Em ambos os cenários, o *package* só calcula os coeficientes policóricos e poliseriais se as distribuições das variáveis não estiverem muito afastadas da distribuição normal.

Para ilustrar a utilização do *package* ‘Polycor’ consideremos um ficheiro de dados ‘DadosAF.sav’ em formato SPSS (.sav)² constituído por 5 variáveis ordinais com 5 pontos (X_1 a X_5), uma variável intervalar (*Age*) e uma variável nominal (*Sex*). Para integrar o *package* ‘Polycor’ no *R* recorreremos ao *Mirror* do CRAN mais próximo e carregamos a biblioteca ‘Polycor’:

```
> library(polycor)
```

Para importar o ficheiro de dados ‘DadosAF.sav’:

```
> library (foreign)
> data<-read.spss("D:/DataAF.sav", to.data.frame = TRUE)
```

² Disponível se solicitado, por e-mail, ao autor.

Para que o programa HetCor calcule o coeficiente de correlação apropriado à métrica das variáveis é necessário fazer a conversão da ‘measure’ do PASW para a métrica do R:

```
> Data$X1<-ordered(Data[,1])      # Converte métrica para ordinal
> Data$X2<-ordered(Data[,2])      # Converte métrica para ordinal
  (...)
> Data$X5<-ordered (Data[,5])
> Data$Age<-as.numeric(Data[,6])  # Converte métrica para numérico
> Data$Sex<-factor(Data[,7])      # converte métrica para nominal
```

Finalmente, para obter os coeficientes de correlação (‘ML=TRUE’ para usar o método de máxima verosimilhança ou ‘ML=FALSE’ para usar o método *two-step*):

```
> R<-hetcor(Data, ML=FALSE, std.err=TRUE)
> print (R)
```

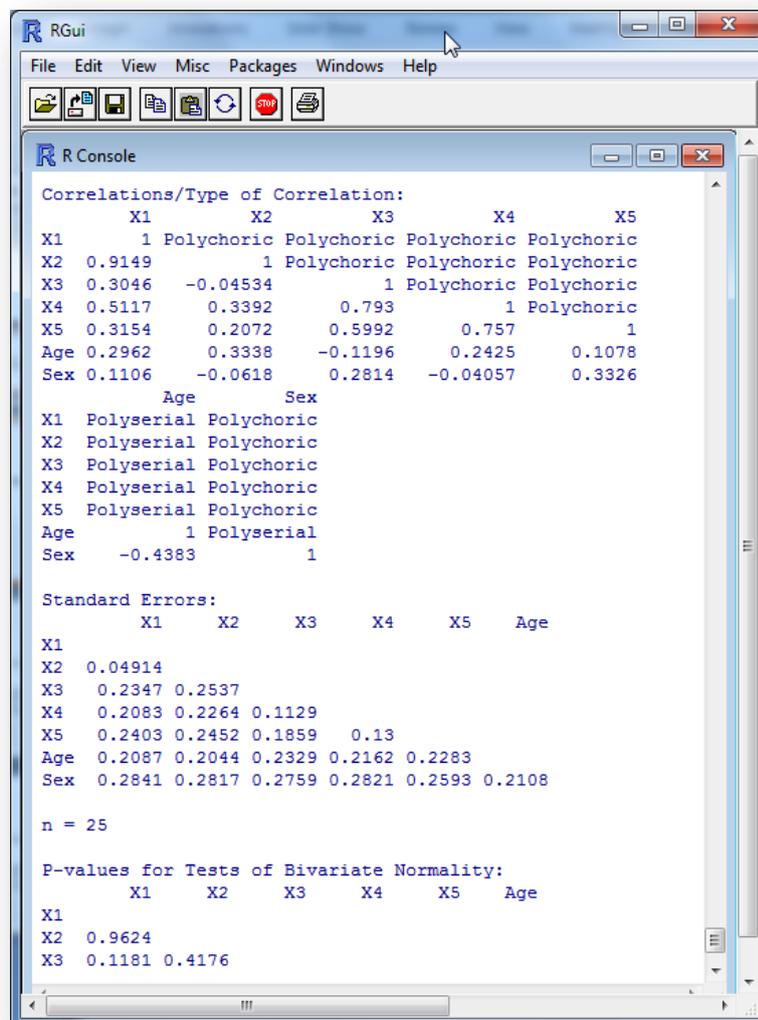


Figura 2 – Output do programa HetCor no sistema R

4. INTEGRAÇÃO DO R NOS MENUS DO *PASW Statistics* COM O ‘CUSTOM DIALOG BUILDER’

A partir da versão 16 do *PASW Statistics* é possível integrar o código do *R* (e também do *Python*) na sintaxe do *PASW*. Na versão 17 do *PASW*, foi desenvolvido um novo aplicativo – Custom Dialog Builder (CDB) – que permite incorporar as funções do *R* nos menus do *PASW* facilitando a interação do utilizador do *PASW*, não familiarizado com a linguagem do *R*, com as potencialidades do *R*. Na verdade, o utilizador final dos menus do *PASW* não necessita sequer compreender o funcionamento do *R* já que todo processo corre na memória do computador sem evocar o GUI do *R*. Para permitir a comunicação entre o *PASW* e o *R* é necessário instalar o ‘*R* integration plugin’ do *package* ‘*R* essentials’ disponível gratuitamente no endereço <http://www.spss.com/devcentral/>. O ‘*R* essentials’ disponibiliza a instalação automática do ‘*PASW: R integration Plug-in*’, do ‘*R2.8.1*’ bem como vários exemplos de bibliotecas de *R* transpostas para os menus do *PASW*. O ‘*R* integration Plug-in’ disponibiliza as funções necessárias à exportação das variáveis do *PASW* para um *dataset* de *R*; à leitura da informação das variáveis do *PASW* para o formato *R*; à transferência dos resultados de análise (via *OMS*) do *PASW* para o *R*; à escrita dos resultados das análises efectuadas no *R* para o *output* do *PASW* e à transposição dos gráficos gerados em *R* para o *output* do *PASW*.

Depois de instalado o ‘*R* integration Plug-in’ o código do *R* poderá ser incorporado na sintaxe do *PASW*, num bloco específico de sintaxe:

```
BEGIN PROGRAM R.          */ Início Bloco de Sintaxe do PASW
(...)                    # Código de R
END PROGRAM.             */ Fim bloco Sintaxe de PASW. Não esquecer os “.”
```

Dentro deste bloco, todo e qualquer código do *R* pode ser utilizado, bem como as funções do ‘*R* integration Plug-in’ para leitura de dados, de variáveis e conversão dos resultados para os *outputs* do *PASW*. Este código pode depois ser utilizado no ‘Custom Dialog Builder’ com modificação mínimas apenas ao nível da leitura das variáveis e das opções de análise.

No *PASW*, para iniciar o ‘Custom Dialog Builder’ recorreremos ao menu

- ▶ Utilities
- ▶ Custom Dialogs
- ▶ Custom Dialog Builder

A janela do ‘Custom Dialog Builder’ e a barra de ferramentas de desenho ilustra-se na figura seguinte:

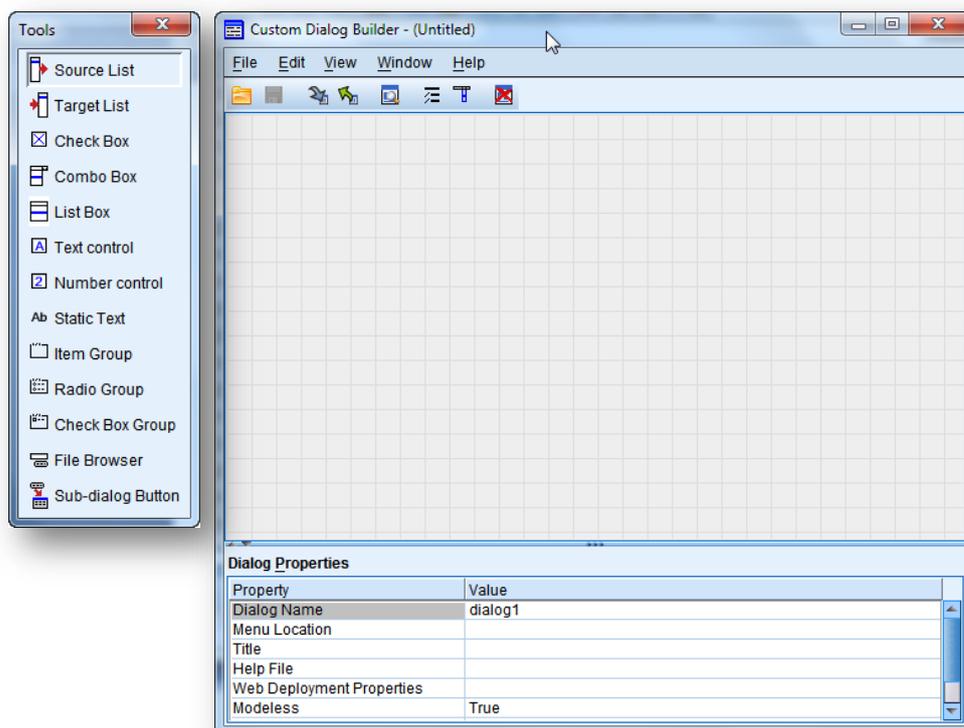


Figura 3 – ‘Custom Dialog Builder’ (CDB) com a barra de ferramentas

Para definir as propriedades do novo menu, preenche-se a caixa ‘Dialog Properties’ com o nome do menu (Dialog name); o local onde o novo menu deverá aparecer (Menu Location); o título do menu (Title) e a localização do ficheiro de ajuda (em formato .html). A opção ‘Modeless’ indica se o menu é independente de outra interação com os menus do *PASW* (True) ou se enquanto este estiver activo não é possível usar outras funcionalidades do *PASW* (False). As especificações do nosso exemplo apresentam-se na figura seguinte. Note que o novo menu designado ‘Poly’ irá aparecer no menu Analyze ► Correlate:

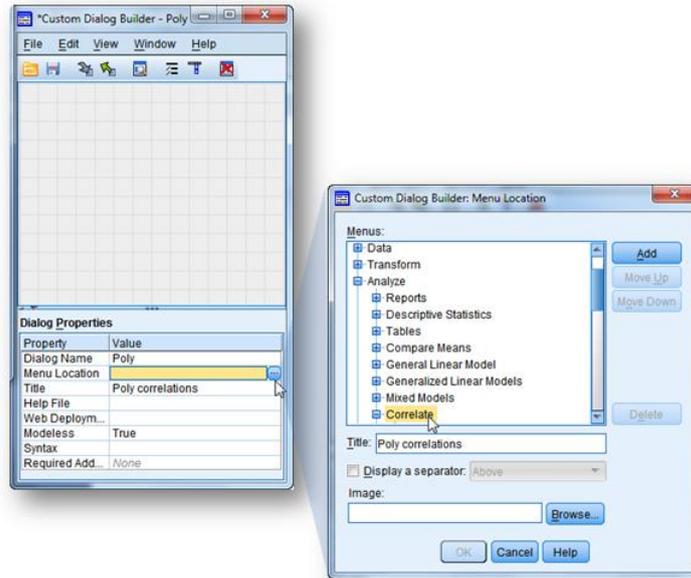


Figura 4 – Atribuição da localização do novo menu ‘Poly’. Para abrir a janela de atributos da localização clica-se no botão .

Para desenhar o menu, recorreremos às ferramentas de desenho da Fig. 2. Para definir a caixa de selecção de variáveis para análise, usa-se a ferramenta ‘Source List’ , desenha-se um rectângulo na área de desenho do CDB e definem-se as propriedades das variáveis que podem entrar na análise, nomeadamente o tipo (type) e a métrica (measurement level). Para seleccionar as variáveis a analisar usa-se a ferramenta . A Fig. 5 ilustra o menu com as caixas de selecção de variáveis e variáveis a analisar.

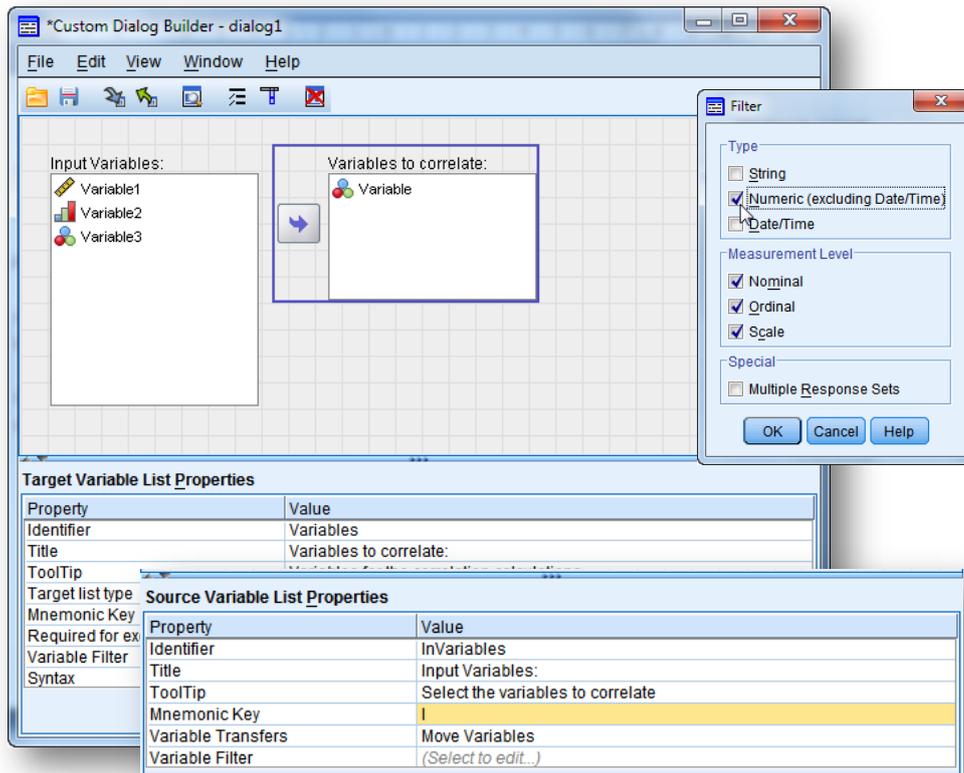


Figura 5 – Definição das variáveis a usar no menu

Para definir opções de análise, por exemplo, o método de estimação recorreremos à ferramenta [Check Box Group](#). Na janela dos 'Radio buttons' definimos as duas opções de análise como ilustra a Fig. 6.

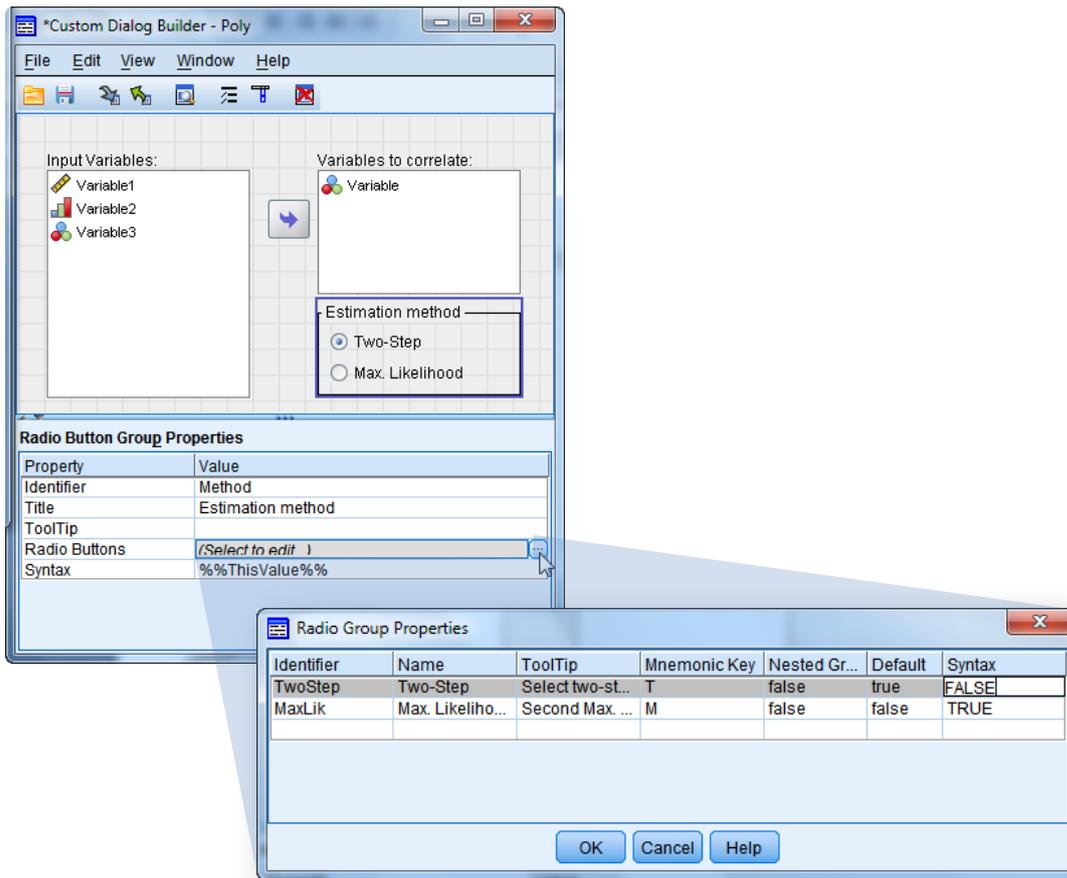


Figura 6 – Definição do método de estimação dos coeficientes. *Two-Step* (keyword FALSE na opção ‘ML=’ no programa HetCor) ou *Maximum likelihood* (keyword TRUE na opção ‘ML=’ no programa HetCor). Para abrir a janela de atributos dos botões clica-se no botão .

Finalmente, introduzimos a sintaxe do PASW com código do R para evocar o programa HetCor. No ‘CDB’ clica-se em ‘Syntax’ e digita-se a sintaxe apropriada:

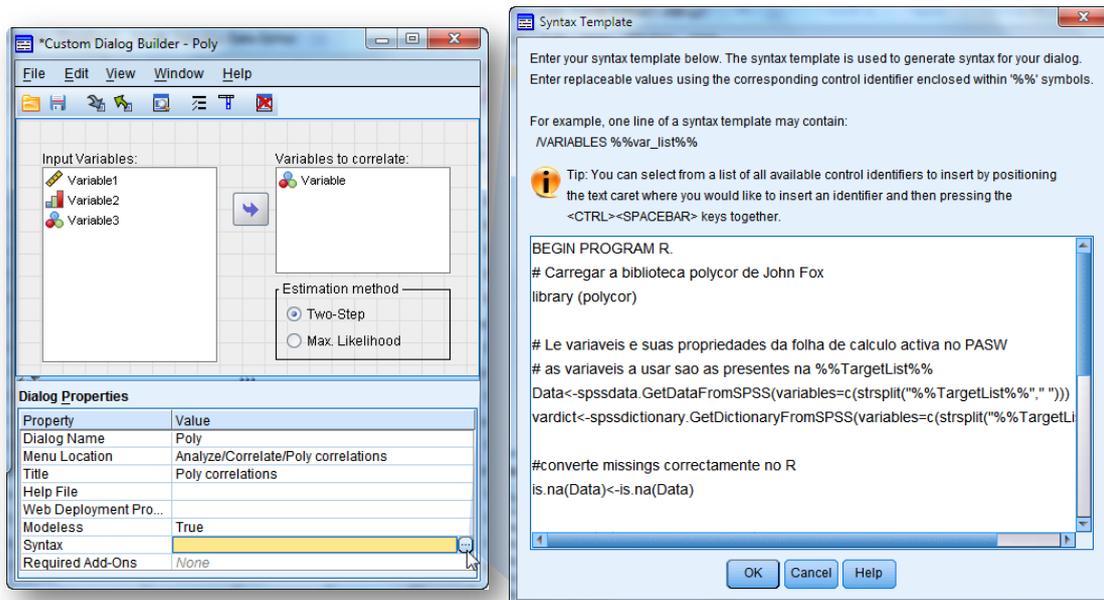


Figura 7 – Evocação da Janela para digitação da sintaxe com o código R adequado à importação das variáveis do PASW para o R, à realização dos cálculos no R e ao retorno dos resultados para o *output* do PASW.

A sintaxe, anotada, adequada ao exemplo apresenta-se a seguir:

```
BEGIN PROGRAM R.
# Carregar a biblioteca polycor de John Fox
library (polycor)

# Lê variáveis e suas propriedades da folha de calculo activa no PASW
# as variáveis a usar são as presentes na %%TargetList%%
Data<-spssdata.GetDataFromSPSS(variables=c(strsplit("%%TargetList%%"," ")))
vardict<-spssdictionary.GetDictionaryFromSPSS (variables = c(strsplit(
"%%TargetList%%"," ")))

#Converte missings correctamente no R
is.na(Data)<-is.na(Data)

# Converte tipo de variáveis nos tipos do R: numeric, non-ordered ou ordered R
# factors
Mat<-Data
for (i in 1:length(Data)) {
  if (vardict["varMeasurementLevel",i]=="scale") Mat[,i]<-Data[,i]
  else
  if (vardict["varMeasurementLevel",i]=="nominal") Mat[,i]<- factor(Data[,i])
  else
  if (vardict["varMeasurementLevel",i]=="ordinal") Mat[,i]<-ordered(Data[,i])
}

# Calcula a matriz de correlação como R List; usar str(r) para ver a
# estrutura da lista. O método de estimação e dado pela keyword %%Estimation%%
r <- hetcor(Mat, ML = %%Estimation%%)

#Cria matriz triangular com estimativas e tipos de correlação
R <- signif(r$correlations, digits=3)
R[upper.tri(R)] <- r$type[upper.tri(R)]
R <- as.data.frame(R)

# Formata as tabelas para output no PASW
spsspivottable.Display(R, title="Correlation types and Correlation
Coefficients")

# Legenda para as colunas e linhas
lab=c(strsplit("%%TargetList%%", " "))
# Dimensão da amostra
spsspivottable.Display(round(r$n), title="N", rowlabels=lab, collabels="n")
# Erros-padrão dos coeficientes
spsspivottable.Display(r[["std.errors"]], title="std. errors")
# Teste à Distribuição normal bivariada
spsspivottable.Display(r[["tests"]], title="p-values for Tests of Bivariate
Normality ")

# Limpa memória de trabalho
rm(list = ls())
END PROGRAM.
```

Finalmente, o menu ‘Poly’ é inserido no PASW, clicando no botão  como ilustra a figura seguinte:

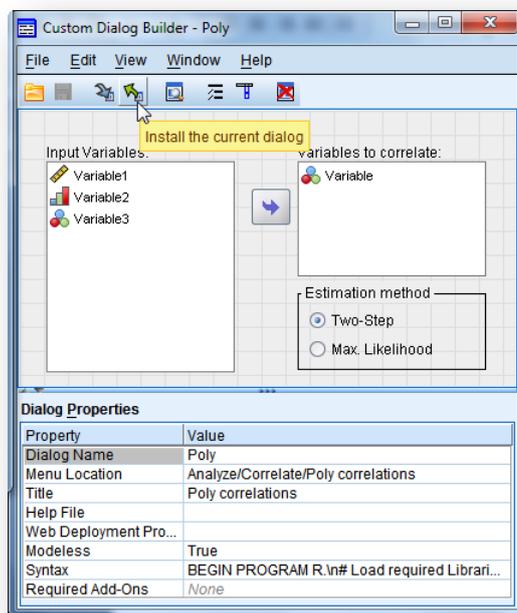


Figura 8 – Inserção do menu ‘Poly’ no PASW Statistics.

Naturalmente, pode também gravar-se o ficheiro (extensão .spd) com o código para o novo menu e cálculos associados. Este ficheiro³ poderá agora ser distribuído em conjunto com os *packages* do *R*, e instalado no *PASW* pelo utilizador final com um simples duplo-click, .

Depois de reiniciar o *PASW*, o menu ‘Poly’ acede-se por

- ▶ Analyze
- ▶ Correlate
- ▶ Poly

como ilustra a Fig. 9.

³ O ficheiro deste exemplo pode ser solicitado, via email, ao autor, ou poder-se-á descarregar do sitio da SPSS DevCentral.

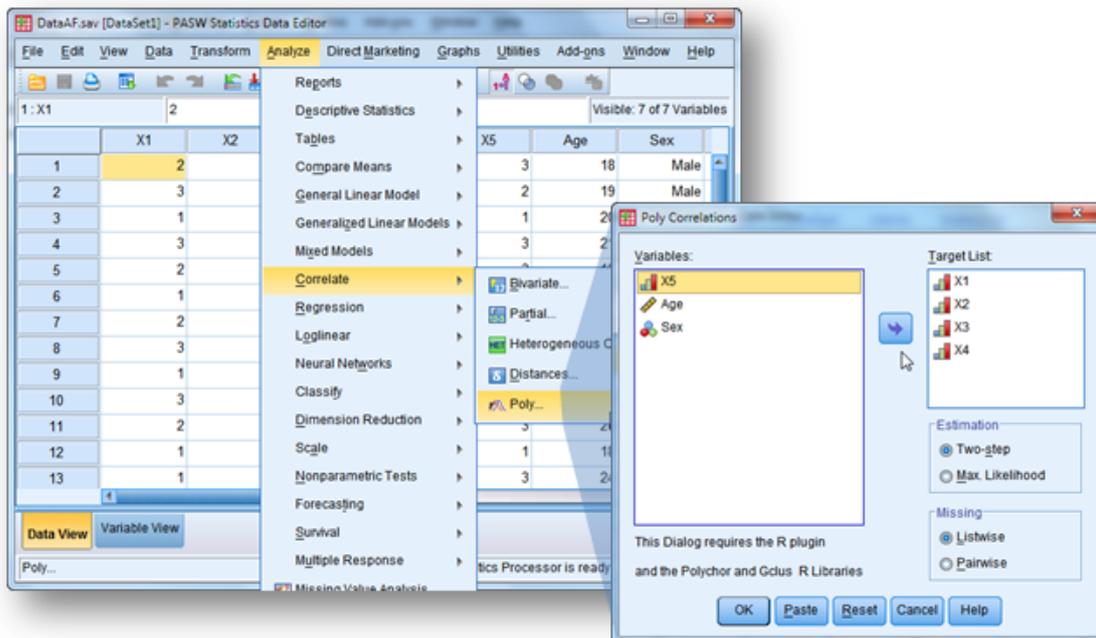
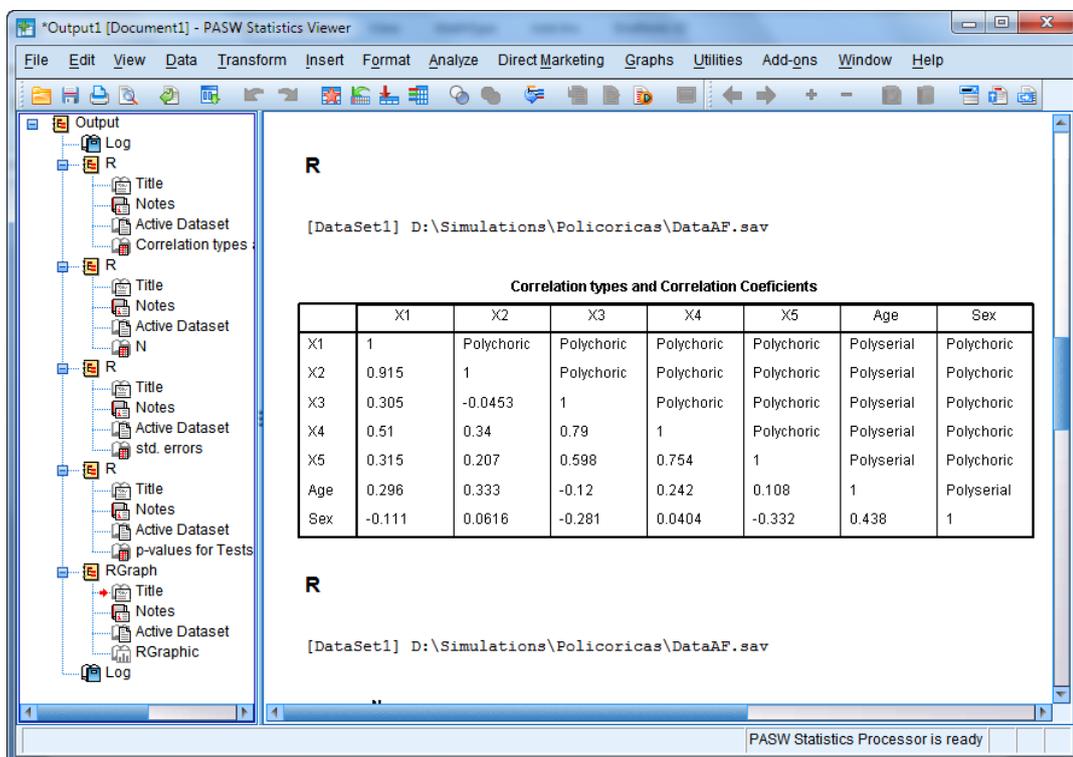


Figura 9 – PASW Statistics com o novo menu para cálculo das correlações policóricas

Depois de seleccionadas as variáveis para as quais se pretende calcular os coeficientes de correlação, o resultado final aparecerá na janela de *output* do PASW Statistics:



As tabelas (e os gráficos) do *output* do PASW podem agora ser exportadas e ou copiadas para outros aplicativos.

REFERÊNCIAS BIBLIOGRÁFICAS

- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Babakus, E., Ferguson, C. E., & Jöreskog, K. G. (1987). The Sensitivity of Confirmatory Maximum Likelihood Factor Analysis to Violations of Measurement Scale and Distributional Assumption. *Journal of Marketing Research*, 24(2), 222-228.
- Finney, S. J., & DiStefano, C. (2006). Non-normal and Categorical Data in Structural Equation Modelling. In G. R. Hancock & R. O. Mueller (Eds.), *Structural equation modeling: a second course* (pp. 269-314). Greenwich, Conn.: IAP.
- Martinson, E. O., & Hamdan, M. A. (1975). Algorithm AS 87: Calculation of the Polychoric Estimate of Correlation in Contingency Tables. *Applied Statistics*, 24(2), 272-278.
- Dragow, F. (2006). Polyserial and Polychoric correlations. In S. Kotz, N. Balakrishnan, C. Read, B. Vidakovic & N. L. Johnson (Eds.), *Encyclopedia of Statistical Sciences* (2nd ed., Vol. 9, pp. 6244-6248). New York: John Wiley & Sons.
- Stevens, S.S. (1946). On the theory of scales of measurement. *Science*, 103, 677-680.