

UAB GetDummies

Generación de variables ficticias y términos de interacción

Comando de extensión (*Extension Command*) para generar las variables ficticias (*dummies*) correspondientes a la descomposición de variables categóricas y a la generación de términos de interacción.

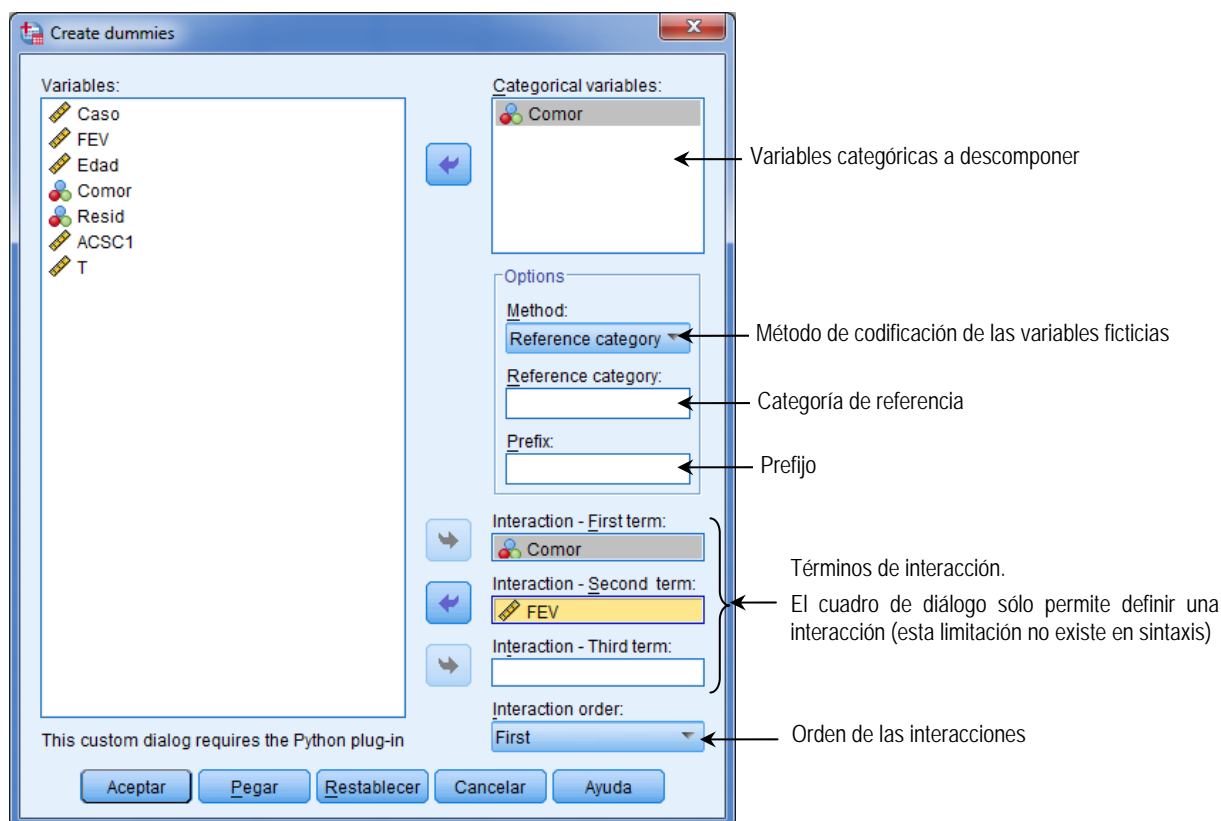
Instalación

Los comandos de extensión requieren como mínimo la versión 18 de SPSS. Antes de instalar *UAB GetDummies* es necesario instalar el complemento de integración de *Python* (*Essentials for Python*). Una vez se cumplen los requisitos previos, instalar *UAB GetDummies* es muy fácil. Ejecutando SPSS como administrador, se ejecuta la opción de menú "Utilidades | Grupos de extensión | Instalar grupo de extensión..." de la ventana de datos de SPSS. En el cuadro de diálogo que se abre debe indicar la dirección en disco del archivo .spe correspondiente. Es conveniente cerrar y reiniciar SPSS para tener la operativa completa.

Atención: En el enlace "Material" del campus virtual hay un documento (*Guia instalación Python y R.pdf*) que explica el proceso de instalación de los complementos de integración y de los comandos de extensión. Siga detalladamente las instrucciones para asegurar que realiza correctamente la instalación.

Ejecución

El comando de extensión funciona como cualquier otro procedimiento SPSS. Es accesible a través de la opción de menú "Transformar | UAB Create dummies...", que abre un cuadro de diálogo que, una vez configurado, puede ejecutar directamente o pegar la sintaxis. Con los datos del archivo de ejemplo *GetDummies_Test.sav* la siguiente imagen muestra como generar las ficticias que representan *Comor* y su interacción con *FEV*:



Como resultado se presentan en la ventana de resultados los nombres de las variables ficticias generadas y una tabla con los coeficientes de cada variable ficticia para cada valor de la variable categórica original. En caso de generarse interacciones se informa también de ello.

```
UAB GETDUMMIES /CATORICALS Comor /INTERACTIONS Comor*FEV
/OPTIONS TYPE=REFERENCE ORDER = 1.
```

UAB GETDUMMIES

Dummy variables:

Comor = Comor10 Comor20

Comor*FEV = Comor10_FEV Comor20_FEV

Type: Reference (reference category: first)

Comor coefficients

	Comor10	Comor20
0	0	0
1	1	0
2	0	1

(c) JM Domenech & JB Navarro.

Laboratori d'Estadística Aplicada - Universitat Autònoma de Barcelona.

Si para la variable *Comor* se selecciona la codificación polinómica se obtienen las dos variables ficticias que representan las tendencias lineal y cuadrática.

```
UAB GETDUMMIES /CATORICALS Comor /OPTIONS TYPE=POLINOMIAL PREFIX= @.
```

UAB GETDUMMIES

Dummy variables:

Comor = @Comor01 @Comor02

Type: Polinomial

Comor coefficients

	@Comor01	@Comor02
0	-,7581518	,3377790
1	,5222824	-1,7882420
2	1,8027166	1,0857183

(c) JM Domenech & JB Navarro.

Laboratori d'Estadística Aplicada - Universitat Autònoma de Barcelona.

A continuación se presentan algunos ejemplos de uso del comando de extensión:

- a. UAB GETDUMMIES /CATEGORICALS Comor /OPTIONS TYPE= REFERENCE REFCAT= FIRST.

Esta llamada genera 2 variables ficticias, *Comor10* y *Comor20* según la codificación respecto a una categoría de referencia, usando la primera como referencia. El algoritmo para generar estas variables es:

```
COMPUTE Comor10 = (Comor = 1).
```

```
COMPUTE Comor20 = (Comor = 2).
```

El nombre de las nuevas variables se genera automáticamente a partir del nombre de la variable original, el valor de la categoría actual y el valor de la categoría de referencia. Así, *Comor10* es la ficticia correspondiente a la segunda categoría (1) respecto a la primera (0).

ATENCIÓN: si estas variables ya existen en la matriz de datos se **sobreescriben**.

- b. UAB GETDUMMIES /CATEGORICALS Comor /OPTIONS TYPE= AVERAGE REFCAT= 2.

Esta llamada genera 2 variables ficticias, *Comor01* y *Comor21* según el método de promedio de categorías, usando la segunda categoría (valor 1 en *Comor*) como referencia. El algoritmo para generar estas variables es:

```
COMPUTE Comor01 = (Comor = 0).
```

```
IF (Comor = 1) = - n0/n1.
```

```
COMPUTE Comor21 = (Comor = 2).
```

```
IF (Comor = 1) = - n2/n1.
```

Donde n_0 , n_1 y n_2 es el número de casos de cada categoría.

- c. UAB GETDUMMIES /CATEGORICALS Comor /OPTIONS TYPE= NESTED REFCAT= FIRST.

Esta llamada genera 2 variables ficticias, *Comor10* y *Comor21* según el método de categorías anidadas, usando la anterior como referencia. El algoritmo para generar estas variables es:

```
COMPUTE Comor10 = (Comor >= 1).
```

```
COMPUTE Comor21 = (Comor >= 2).
```

Si REFCAT= LAST se utiliza la categoría posterior como referencia, según el algoritmo:

```
COMPUTE Comor01 = (Comor < 1).
```

```
COMPUTE Comor12 = (Comor < 2).
```

- d. UAB GETDUMMIES /CATEGORICALS Comor /OPTIONS TYPE= POLINOMIAL PREFIX=@.

Esta llamada genera 2 variables ficticias, *@Comor01* y *@Comor02* descomponiendo en polinomios ortogonales. En este caso los nombres de las variables se generan secuencialmente, y se les añade el prefijo @ como indica el parámetro PREFIX.

- e. UAB GETDUMMIES /CATEGORICALS Comor /INTERACTIONS Comor*FEV.

Esta llamada genera las 2 variables ficticias *Comor10* y *Comor20* del ejemplo a., y dos variables más correspondientes a la interacción: *Comor10_FEV* y *Comor20_FEV*. En el caso de las interacciones los nombres de las variables ficticias se generan uniendo los nombres de los dos componentes con el símbolo '_'. Si una de las variables es categórica debe indicarse en el comando CATEGORICALS, y de forma automática se generan los componentes de interacción para cada una de las ficticias de la variable categórica.

- f. UAB GETDUMMIES /CATEGORICALS Comor /INTERACTIONS Comor*FEV*Resid /OPTIONS ORDER=2.

Esta llamada genera las 2 variables ficticias *Comor10* y *Comor20* del ejemplo a., y dos variables más correspondientes a la interacción: *Comor10_FEV_Resid* y *Comor20_FEV_Resid*. Para respetar el principio jerárquico también se generan las variables correspondientes a todas las interacciones de primer orden: *Comor10_FEV*, *Comor20_FEV*, *Comor10_Resid*, *Comor20_Resid* y *FEV_Resid*.

Referencia de sintaxis

```
UAB GETDUMMIES /CATEGORICALS list_var
/INTERACTIONS list_varint
/OPTIONS [TYPE= {REFERENCE**}] [REFCAT= {FIRST**}] [PREFIX= literal] [ORDER={1**}]
          {AVERAGE   }                {LAST   }                {2   }
          {NESTED    }                {n     }
          {POLINOMIAL }

[HELP]

** Valor por defecto
```

- /CATEGORICALS: Variable o lista de variables **categorías** que se desean descomponer.
- /INTERACTIONS: Variable o lista de variables de **interacción** que se desean crear. El formato de la lista debe ser var1*var2.
- TYPE: Método para generar las ficticias de las categorías:
 - REFERENCE: respecto a una categoría de referencia (**defecto**).
 - AVERAGE: respecto al promedio de categorías.
 - NESTED: codificación en categorías anidadas.
 - POLINOMIAL: codificación en polinomios ortogonales.
- REFCAT: Categoría de referencia si TYPE = REFERENCE, AVERAGE o NESTED. Los valores de la variable categórica se ordenan en forma ascendente.
 - FIRST: primera categoría (**defecto**).
 - LAST: última categoría.
 - n: **posición** de una categoría intermedia (no tiene sentido si TYPE = NESTED). Este parámetro no tiene sentido si TYPE = POLINOMIAL.
- PREFIX: Prefijo en forma de texto literal, sin comillas. Se añade al nombre creado de forma automática para la variable ficticia. Por defecto no se añade nada. Este valor debe cumplir las normas para los nombres de variable en SPSS, y no puede empezar por # o \$.
- ORDER: Orden de las interacciones que se están definiendo. Puede ser 1 (primer orden, var1*var2) o 2 (segundo orden, var1*var2*var3). Por defecto es 1. No es posible definir interacciones de diferente orden en una misma llamada. Si la interacción es de segundo orden, se crean todas las interacciones de primer orden (var1*var2, var1*var3, var2*var3) necesarias para respetar el principio jerárquico.
- HELP: Imprime la ayuda y la referencia bibliográfica en la ventana de resultados.