



CORRELACIONES CON SPSS

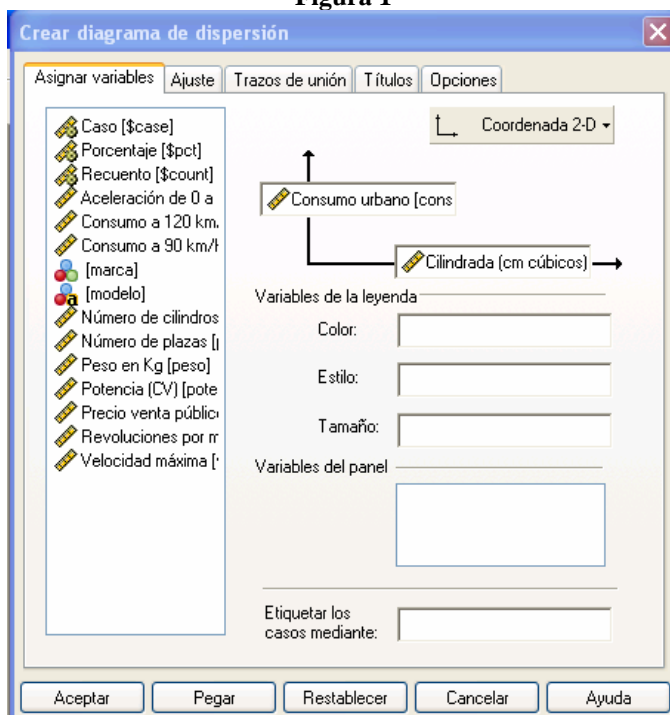
1.- INTRODUCCIÓN

El concepto de relación o *correlación* entre dos variables se refiere al grado de parecido o variación conjunta existente entre las mismas. En este apartado vamos a estudiar un tipo particular de relación llamada *lineal* y se limita a considerar únicamente el caso de dos variables cuantitativas (*correlación simple*).

Una relación lineal *positiva* entre dos variables X e Y significa que los valores de las dos variables varían de forma parecida: los sujetos que puntúan alto en X tienden a puntuar alto en Y y los que puntúan bajo en X tienden a puntuar bajo en Y. Una relación lineal *negativa* significa que los valores de ambas variables varían justamente el revés.

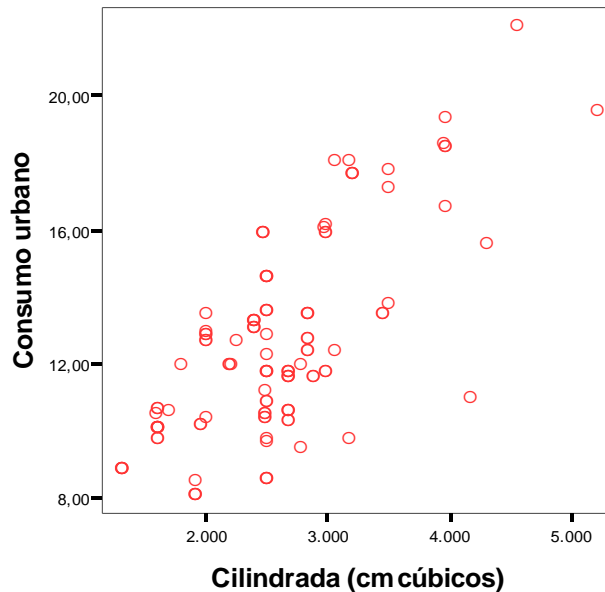
La forma más directa de formarse una primera idea sobre el tipo de relación existente entre dos variables cuantitativas es a través de un **Diagrama de dispersión**. Este tipo de diagramas puede obtenerse mediante: **Gráficos**, en las diferentes opciones que tiene: *Generador de Gráficos*, *Cuadros de diálogo antiguos* o en *Gráficos interactivos*. Ésta última opción muestra el cuadro de diálogo de la figura 1:

Figura 1



Que no da como resultado un diagrama como el que muestra la figura 2.

Figura 2



Para poder cuantificar el grado de relación lineal existente entre dos variables cuantitativas, así como medir el grado de ajuste de la nube de puntos a una recta, vamos a estudiar *coeficientes de correlación*. En el procedimiento de *Tablas de Contingencia* ya se puede obtener el coeficiente de correlación de Pearson, en este apartado estudiaremos el procedimiento **Correlaciones** que incluye tres opciones (1) **Bivariadas**, para el estudio de la relación entre dos variables cuantitativas, (2) **Parciales**, para el estudio de la relación entre dos variables cuantitativas cuando se controla o elimina el efecto de terceras variables y (3) **Distancias**, para el estudio de la relación entre dos variables cualesquiera que sea su nivel de medida.

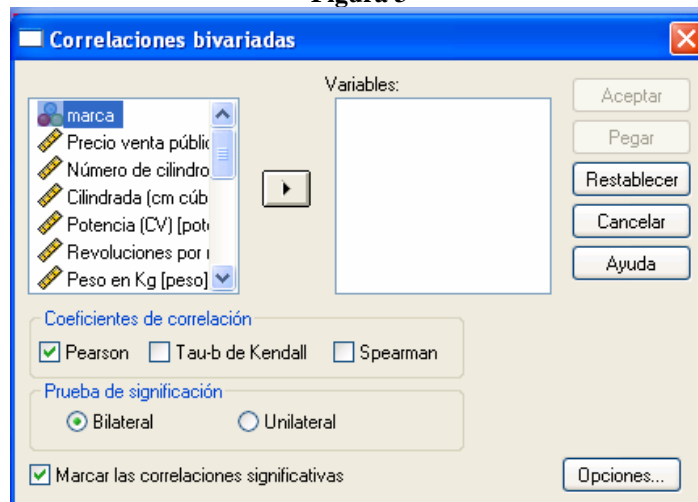
2.- CORRELACIONES BIVARIADAS

El procedimiento **Correlaciones divariadas** ofrece tres tipos de coeficientes: r_{xy} de Pearson, tau-b de Kendall y rho de Spearman. Para acceder a este procedimiento, elegir:

Analizar
Correlaciones
Divariadas...

Apareciendo el cuadro de diálogo de la Figura 3

Figura 3



La lista de variables sólo muestra las variables que poseen formato numérico. Es necesario trasladar al menos dos variables.

Coefficientes de Correlación. Pueden seleccionarse uno o más de los tres siguientes coeficientes:

- **Peason:** Es una medida de la asociación lineal entre dos variables. Los valores del coeficiente de correlación van de -1 a 1. El signo del coeficiente indica la dirección de la relación y su valor absoluto indica la fuerza. Los valores mayores indican que la relación es más estrecha.
- **Tau-b de Kendall:** Es una medida no paramétrica de asociación para variables ordinales o de rangos que tiene en consideración los empates. El signo del coeficiente indica la dirección de la relación y su valor absoluto indica la magnitud de la misma, de tal modo que los mayores valores absolutos indican relaciones más fuertes. Los valores posibles van de -1 a 1, pero un valor de -1 o +1 sólo se puede obtener a partir de tablas cuadradas.
- **Spearman:** Versión no paramétrica del coeficiente de correlación de Pearson, que se basa en los rangos de los datos en lugar de hacerlo en los valores reales. Resulta apropiada para datos ordinales, o los de intervalo que no satisfagan el supuesto de normalidad. Los valores del coeficiente van de -1 a +1. El signo del coeficiente indica la dirección de la relación y el valor absoluto del coeficiente de correlación indica la fuerza de la relación entre las variables. Los valores absolutos mayores indican que la relación es mayor.

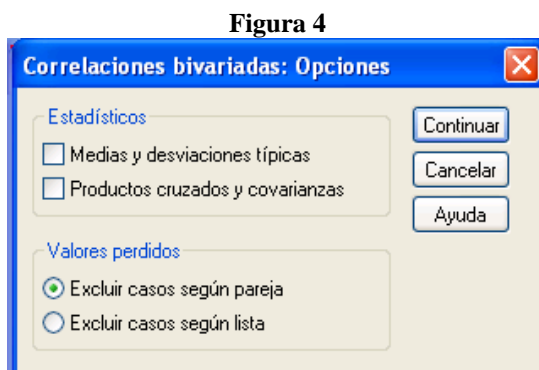
Prueba de significación. Junto con cada coeficiente de correlación, el Visor ofrece la información necesaria para contrastar la hipótesis nula de que el valor poblacional del coeficiente es cero. El SPSS permite seleccionar el nivel crítico deseado:

- **Bilateral:** Probabilidad de obtener resultados tan extremos como el obtenido, y en cualquier dirección, cuando la hipótesis nula es cierta. Un nivel de significación bilateral (de dos colas) contrasta una hipótesis nula en la que la dirección del efecto no se especifica de antemano.
- **Unilateral:** Probabilidad de obtener un resultado tan extremo como el observado, y en la misma dirección, cuando la hipótesis nula es cierta. Contrasta la hipótesis nula en la que se especifica con antelación la dirección del efecto.

Marcar las correlaciones significativas. Esta opción, que se encuentra activa por defecto, Marca con un asterisco los coeficientes de correlación significativos al nivel 0,05 y, con dos asteriscos, los significativos al nivel 0,01.

OPCIONES

El subcuadro **Opciones** permite solicitar estadísticos adicionales y determinar el tratamiento de los valores perdidos. Al pulsarlo nos aparece el cuadro de diálogo de la figura 4



Estadísticos. Si se ha elegido el coeficiente de Pearson, las opciones de este recuadro permiten seleccionar una o más de las siguientes opciones:

- **Medias y desviaciones típicas:** Muestra, para cada variable, la media, la desviación típica (insesgada) y el número de casos válidos.
- **Productos cruzados y covarianzas:** Muestra, para cada par de variables, los productos cruzados de las desviaciones de cada puntuación respecto de su media y la covarianza

Valores perdidos. Las opciones de este recuadro permiten seleccionar el tratamiento que se desea dar a los valores perdidos.

- **Excluir casos según pareja:** Excluye del análisis los casos que tengan valores perdidos en cualquiera (o en ambas) de las variables de una pareja implicada en el cálculo de un estadístico específico.
- **Excluir casos según pareja:** Excluye los casos que tienen valores perdidos en cualquiera de las variables utilizadas en cualquiera de los análisis.

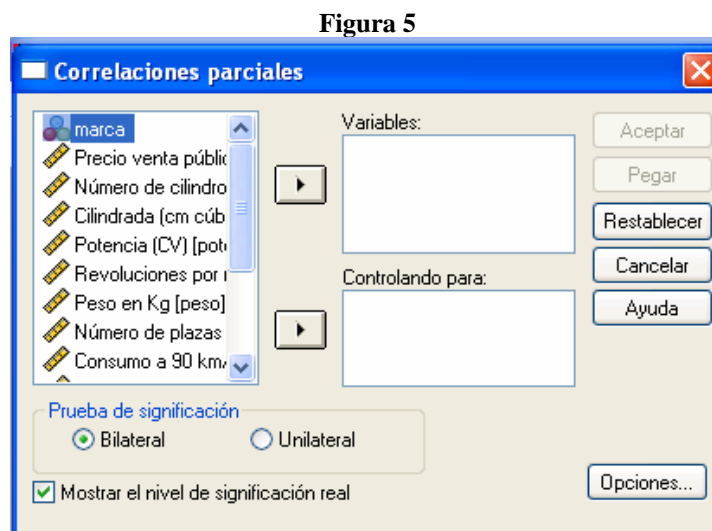
3.-CORRELACIONES PARCIALES

El procedimiento de **Correlaciones parciales** permite estudiar la relación existente entre dos variables cuantitativas controlando el posible efecto de una o más variables cuantitativas extrañas. Un coeficiente de *correlación parcial* es una técnica de control estadístico que expresa el grado de relación lineal entre dos variables tras eliminar de ambas el efecto atribuible a terceras variables.

Para acceder al procedimiento:

Analizar
Correlaciones
Parciales...

Apareciendo el cuadro de diálogo de la Figura 5



En **Variab** se trasladan las variables numéricas que se quieren correlacionar y en **Controlando para** se trasladan las variables cuyo efecto se desea controlar.

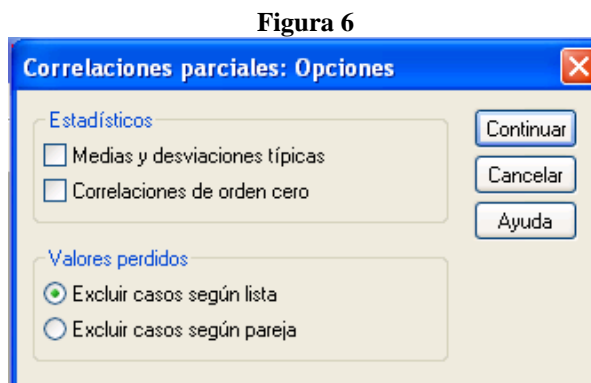
Se habla de correlación de *primer orden* ($r_{12.3}$), para indicar que se está controlando el efecto de una variable; de *segundo orden* ($r_{12.34}$), para indicar que se está controlando el efecto de dos variables; etc.

Prueba de significación. Es exactamente como en Bivariadas pero para los coeficientes parciales, dando opción a elegir el tipo de nivel crítico deseado.

Mostrar el nivel de significación real. Se muestran la probabilidad y los grados de libertad para cada coeficiente. Es una opción activa por defecto, cuando se desactiva el Visor muestra un asterisco al lado de los coeficientes de correlación significativos al nivel 0,05 y, con dos asteriscos, los significativos al nivel 0,01.

OPCIONES

El subcuadro **Opciones** permite solicitar estadísticos adicionales y determinar el tratamiento de los valores perdidos. Al pulsarlo nos aparece el cuadro de diálogo de la figura 6



Estadísticos. Las opciones de este recuadro permiten seleccionar una o más de las siguientes opciones:

- **Medias y desviaciones típicas:** Muestra, para cada variable, la media, la desviación típica y el número de casos válidos.
- **Correlaciones de orden cero:** Coeficientes de correlación ordinarios, sin variables de control. Los valores del coeficiente de correlación van de -1 a 1. El signo del coeficiente indica la dirección de la relación, y su valor absoluto indica la fuerza. Los valores mayores, en valor absoluto, indican que la fuerza es mayor.

Valores perdidos. Idéntico al recuadro para Bivariadas.

4.- DISTANCIAS

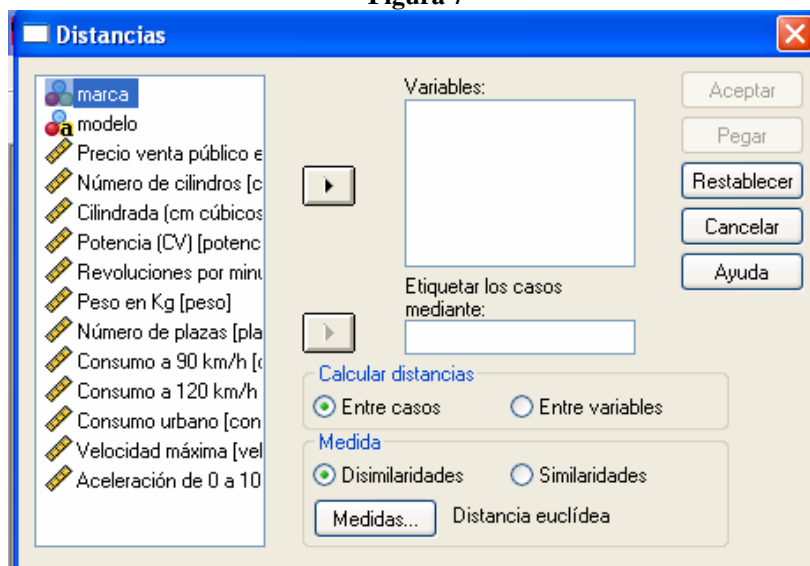
Los coeficientes de correlación vistos son formas de cuantificar la distancia entre dos variables, pero existen otras muchas formas de hacerlo. El procedimiento **Distancias** incluye un gran número de medidas que se diferencian, básicamente, por el tipo de datos para el que han sido diseñadas. Estas medidas pueden utilizarse tanto para obtener distancias entre *variables* como entre *casos*. Llamaremos *elemento* tanto a los casos como a las variables.

Para acceder al procedimiento:

Analizar
Correlaciones
Distancias...

Apareciendo el cuadro de diálogo de la Figura 7

Figura 7



En **Variables** se trasladan las variables cuya distancia se desea calcular (o las variables en las que debe basarse la distancia entre casos).

En **Etiquetar los casos mediante:** Por defecto, en los resultados los casos se identifican mediante el número de caso. De manera opcional, puede utilizar los valores de una variable de cadena para identificar los casos.

Calcular distancias. El procedimiento permite calcular distancias entre casos y entre variables, en ambos opciones las distancias se calculan a partir de las puntuaciones de los casos en el conjunto de variables seleccionadas.

Medida. Las medidas de distancias están agrupadas en dos bloques: *Disimilaridad* y *Similaridad*. El botón **Medidas...** de este recuadro conduce a un subcuadro de diálogo que permite elegir la media de distancia que se desea utilizar, tiene dos versiones (que se diferencian por el tipo de medida que ofrecen) dependiendo de la opción marcada en el recuadro:

- **Disimilaridades:** Medidas de diferencias o lejanía. Los valores más altos indican que los elementos son muy distintos o que se encuentran muy alejados.
- **Similaridades:** Medidas de parecido o cercanía. Los valores más altos indican que los elementos son muy parecidos o que se encuentran muy próximos.