

UNEDpSI: Un programa “didáctico” de análisis de datos en Psicología.

Área Temática: Universidades virtuales y centros de educación a distancia.

Ángel Villarino y Pedro Rodríguez-Miñón.

(avillarino@psi.uned.es prodriguez-minon@psi.uned.es)

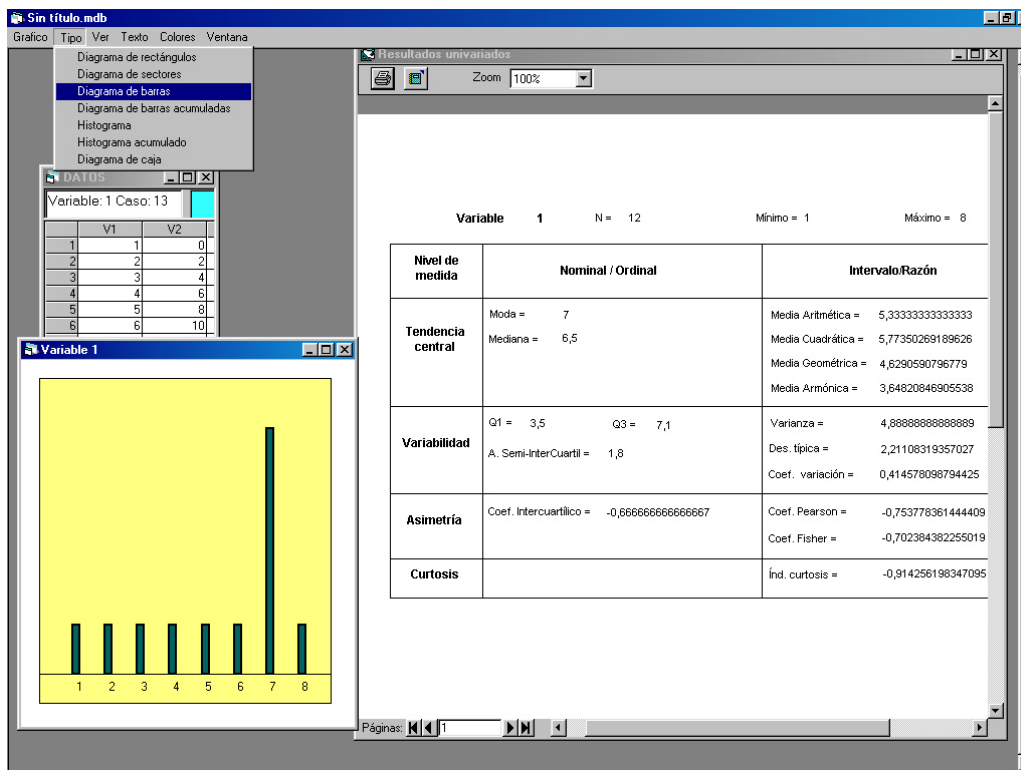
Facultad de Psicología. UNED.

RESUMEN: Actualmente es difícil concebir la docencia y la aplicación práctica del Análisis de Datos en Psicología sin el uso del ordenador, y para este fin existen en el mercado diferentes paquetes estadísticos comerciales capaces de satisfacer las necesidades más exigentes. Ahora bien, desde el punto de vista docente, y especialmente en una Universidad a distancia, el uso de dichos paquetes por parte del alumnado es más que problemático. Al margen de los evidentes problemas económicos, la versatilidad de los programas profesionales de análisis de datos puede ser contraproducente para iniciar a un alumno en su uso, puesto que el alumno debe aprender a navegar entre un conjunto de opciones que exceden con mucho su nivel de exigencia de la asignatura. De igual manera, la salida de resultados proporcionados, excede también los objetivos didácticos de la materia.

Presentamos un programa de análisis de datos, elaborado por el equipo docente de Psicología Matemática I, de la facultad de Psicología de la UNED especialmente diseñado para nuestros alumnos en un triple sentido: económico, ajustado a los objetivos didácticos y de fácil uso que le familiarizará con el “interface” de los paquetes profesionales más complejos.

El programa, ajustado a los objetivos docentes de la materia de estudio, dispone de tres bloques temáticos: El primero realiza todo el análisis de datos de una variable: tendencia central, variabilidad y forma así como las representaciones gráficas más utilizadas que se exponen en el temario de la asignatura: diagrama de barras, de sectores, histogramas, polígonos de frecuencias, utilizando tanto las frecuencias absolutas, como las frecuencias relativas o proporciones y con los datos acumulados. La figura 1 muestra una típica sesión de análisis de datos univariados donde aparece la barra de menú y tres ventanas. La barra de menú cambia según activemos la ventana de datos, de gráficos o de resultados y cada menú dispone de diferentes opciones.

Figura 1: Ventana de Análisis Univariado de Datos.

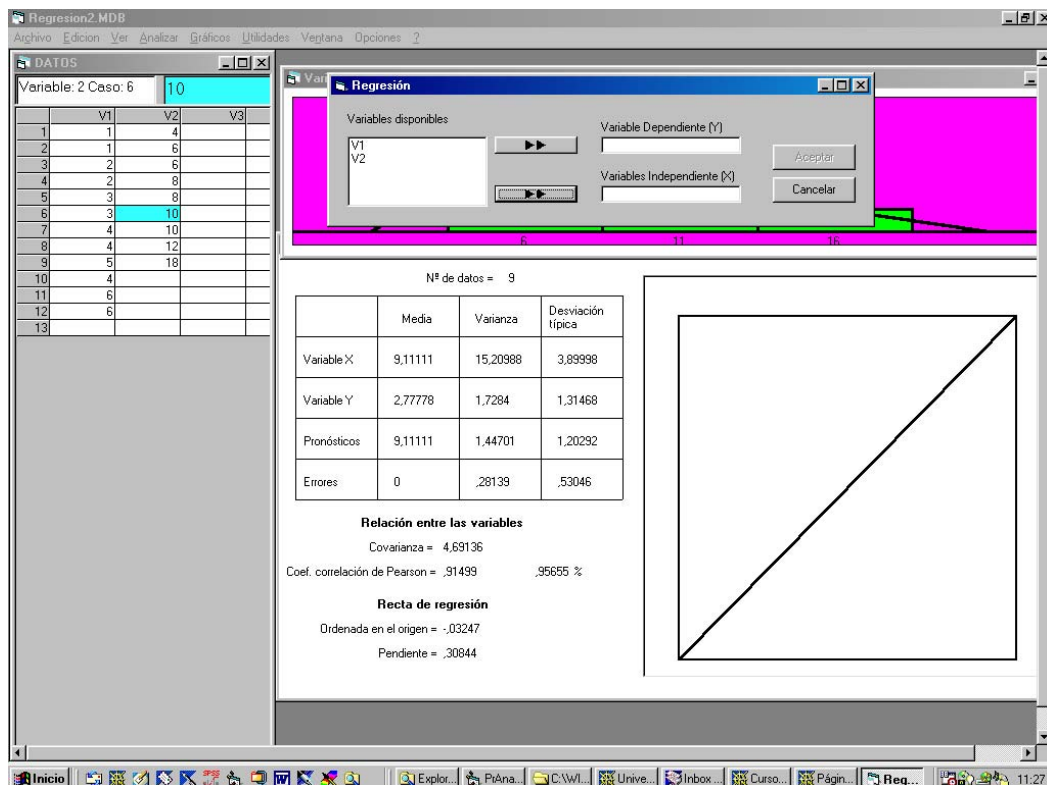


Si nos situamos en la ventana de datos y nos movemos con la flechas del cursor de una variable a

otra, la ventana de gráficos y la ventana de resultados cambia proporcionando la información correspondiente a la variable seleccionada. Igualmente si añadimos nuevos datos a una variable, tanto la ventana de gráficos como de resultados muestra las modificaciones gráficas y analíticas que se van produciendo en los resultados.

El segundo gran bloque temático es el análisis de datos bivariados, donde el programa permite obtener la correlación y regresión entre dos variables. Para ello, con la ventana de datos activa se selecciona en la barra de menú la opción *analizar* para seleccionar dentro de las opciones la relación entre variables -que permite obtener distintos índices de relación en función de la naturaleza de las variables seleccionadas- o la regresión para variables cuantitativas.

Figura 2: Ventana de Análisis Bivariado de Datos.



La figura 2 muestra el *interface* del programa para el análisis de la regresión donde previamente aparece la ventana que permite seleccionar las variables que se introducen en el análisis como variable dependiente y como variable independiente. La ventana de resultados muestra en una tabla la media, varianza y desviación típica de cada variable introducida en la regresión, así como de los pronósticos y de los errores de pronóstico. Muestra también los índices estadísticos del análisis bivariado: covarianza, correlación de Pearson, coeficiente de determinación, la ordenada y la pendiente de la recta de regresión.

Dentro de esta ventana, aparece la representación gráfica de la recta de regresión ajustada sobre el diagrama de dispersión o nube de puntos. Finalmente en esta misma ventana de resultados aparece la puntuación pronosticada en puntuaciones directas, diferenciales y típicas, para cualquier valor de la variable independiente que seleccione el usuario.

El tercer bloque temático de este programa se centra en la probabilidad y fundamentalmente en los modelos de probabilidad de variables aleatorias discretas y continuas. El usuario puede trabajar con cuatro modelos de probabilidad de variables continuas: Normal, t de Student, chi-cuadrado y F de Fisher y con dos modelos de probabilidad de variables discretas: Binomial y Poisson. Cada una de estas distribuciones permite introducir valores de la variable para determinar probabilidades asociadas a estos valores o, viceversa, entrar unos valores de probabilidad para determinar el valor de la variable correspondiente.

Para cada distribución se muestra en la parte inferior tres cajas de texto (Fig.3) que se refieren a la distribución que se está ejecutando, las puntuaciones seleccionadas y las probabilidades, cada una de las cuales con una serie de controles que puede manipular el usuario en función de la distribución de probabilidad que estemos componiendo.

Figura 3: Caja de Texto donde las ventanas abiertas son los valores que puede introducir el usuario

Distribución Binomial	Puntuación seleccionada	Probabilidades
Media <input type="text" value="4,2"/>	Nº de éxitos <input type="text" value="0"/>	Función de probabilidad <input type="text" value="0,00569"/>
Varianza <input type="text" value="2,73"/>	Probabilidad de éxito <input type="text" value="0,35"/>	Probabilidad por debajo <input type="text" value="0,00569"/>
Nº de ensayos <input type="text" value="12"/>	Probabilidad de fracaso <input type="text" value="0,65"/>	Probabilidad por encima <input type="text" value="0,99431"/>

Así, para la distribución normal podemos definirla por su media y desviación típica en la primera caja de texto (Fig. 4), y en la segunda caja seleccionar una determinada puntuación (directa o típica) dando como resultado, en la tercera caja de texto, las probabilidades de obtener puntuaciones menores o iguales a la puntuación seleccionada y la probabilidad complementaria.

Figura 4: Caja de Texto de la distribución Normal.

Distribución normal	Puntuación seleccionada	Probabilidades
Media <input type="text" value="25"/>	Puntuación directa <input type="text" value="20"/>	Función de densidad <input type="text" value=",047"/>
Desviación típica <input type="text" value="6"/>	Puntuación típica <input type="text" value="-0,8333"/>	Probabilidad por debajo <input type="text" value="0,2023"/>
		Probabilidad por encima <input type="text" value="0,7977"/>

Se puede trabajar de forma inversa, una vez definida los parámetros de la distribución seleccionar un determinado valor de probabilidad y la salida es la puntuación correspondiente y su representación gráfica.

Con la distribución t de Student de media cero y varianza en función de sus grados de libertad, se puede manipular la puntuación directa y los grados de libertad obteniendo en la tercera caja de texto (Fig. 5) de probabilidades la probabilidad de obtener puntuaciones menores o iguales que la seleccionada y la probabilidad de obtener probabilidades mayores o iguales que la seleccionada. También se puede operar de forma inversa: seleccionando una probabilidad y unos determinados grados de libertad el programa proporciona como salida la puntuación correspondiente y su representación gráfica.

Figura 5: Caja de Texto de la distribución T de Student.

Distribución T de Student	Puntuación seleccionada	Probabilidades
Media <input type="text" value="0"/>	Puntuación <input type="text" value=",4732"/>	Función de densidad <input type="text" value=",3519"/>
Varianza <input type="text" value="1,0833333"/>	Grados de libertad <input type="text" value="26"/>	Probabilidad por debajo <input type="text" value="0,68"/>
		Probabilidad por encima <input type="text" value="0,32"/>

Con la distribución chi-cuadrado cuyos parámetros vienen determinados por sus grados de libertad que introducimos en el cuadro de control de la caja de texto (Fig. 6) referida a las puntuaciones seleccionadas, podemos seleccionar una determinada puntuación para obtener sus probabilidades correspondientes o a la inversa, es decir, fijar previamente un valor de probabilidad para obtener la puntuación que divide el espacio probabilístico. En cualquier caso se obtiene también su correspondiente representación gráfica.

Figura 6: Caja de Texto de la distribución Chi-cuadrado.

Distribución Ji cuadrado	Puntuación seleccionada	Probabilidades
Media <input type="text" value="9"/>	Puntuación <input type="text" value="12,3"/>	Función de densidad <input type="text" value=",0529"/>
Varianza <input type="text" value="18"/>	Grados de libertad <input type="text" value="9"/>	Probabilidad por debajo <input type="text" value="0,8031"/>
		Probabilidad por encima <input type="text" value="0,1969"/>

Si ejecutamos la distribución F de Fisher, podemos controlar los grados de libertad del numerador y del denominador (Fig. 7) lo cual determina sus correspondientes parámetros media y varianza y posteriormente seleccionar una determinada puntuación para obtener la probabilidad o una determinada probabilidad para obtener la puntuación correspondiente.

Figura 7: Caja de Texto de la distribución F de Fisher.

Distribución F de Fisher	Puntuación seleccionada	Probabilidades
Media <input type="text" value="1,3333333"/>	Puntuación <input type="text" value="1,9133"/>	Función de densidad <input type="text" value=",1873"/>
Varianza <input type="text" value="0,81"/>	G.L. Numerador <input type="text" value="8"/>	Probabilidad por debajo <input type="text" value="0,85"/>
	G.L. Denominador <input type="text" value="12"/>	Probabilidad por encima <input type="text" value="0,15"/>

Para la distribución binomial (Fig. 8) los controles que quedan abiertos al usuario son el número de ensayos –en la primera caja de texto- el número de éxitos cuya probabilidad queremos buscar y su probabilidad–en la segunda caja de texto- y como resultado nos proporciona su representación gráfica y las probabilidades correspondientes en la tercera caja de texto.

Figura 8: Caja de Texto de la distribución binomial.

Distribución Binomial	Puntuación seleccionada	Probabilidades
Media <input type="text" value="14"/>	Nº de éxitos <input type="text" value="12"/>	Función de probabilidad <input type="text" value="0,1144"/>
Varianza <input type="text" value="4,2"/>	Probabilidad de éxito <input type="text" value="0,7"/>	Probabilidad por debajo <input type="text" value="0,22773"/>
Nº de ensayos <input type="text" value="20"/>	Probabilidad de fracaso <input type="text" value="0,3"/>	Probabilidad por encima <input type="text" value="0,77227"/>

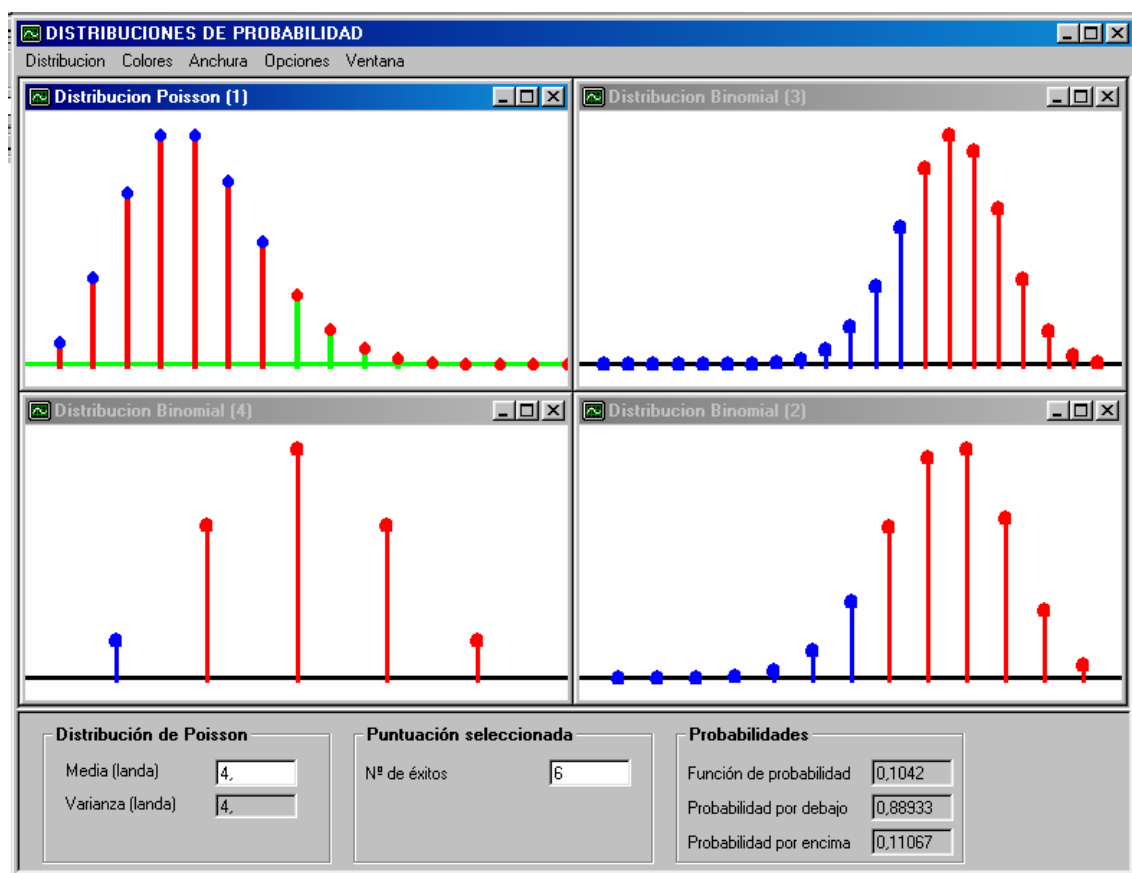
La última distribución de probabilidad que se presenta es la distribución de Poisson (Fig. 9) en la que podemos controlar su media y el número de éxitos para obtener su probabilidad y representación gráfica.

Figura 9: Caja de Texto de la distribución Binomial.

Distribución Binomial	Puntuación seleccionada	Probabilidades
Media <input type="text" value="4,2"/>	Nº de éxitos <input type="text" value="0"/>	Función de probabilidad <input type="text" value="0,00569"/>
Varianza <input type="text" value="2,73"/>	Probabilidad de éxito <input type="text" value="0,35"/>	Probabilidad por debajo <input type="text" value="0,00569"/>
Nº de ensayos <input type="text" value="12"/>	Probabilidad de fracaso <input type="text" value="0,65"/>	Probabilidad por encima <input type="text" value="0,99431"/>

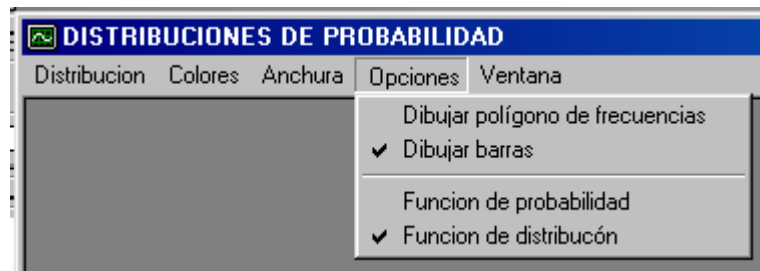
En cada caso se presenta una ventana de gráficos con cuatro opciones en la barra de menú: colores, anchura, opciones y ventana que permite al usuario cambiar, tanto las características de la presentación gráfica, como las opciones gráficas de la distribución. En la barra de menú, la opción ventana nos permite trabajar con varias distribuciones a la vez, como se muestra en la figura 10. Las distribuciones se pueden presentar en mosaico o en cascada, seleccionando con el ratón una de las distribuciones aparece en la parte inferior su correspondientes controles en las cajas de texto.

Figura 10: Cuatro representaciones gráficas presentadas en mosaico.



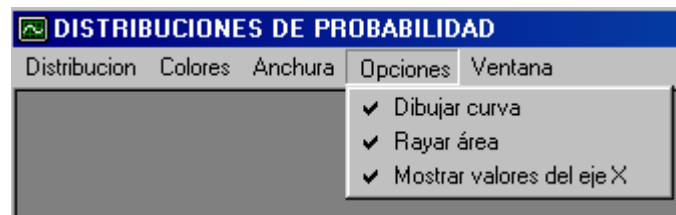
El menú opciones, permite calcular y dibujar la función de probabilidad y la función de distribución para las distribuciones discretas (Fig. 11) y presentarlo como un diagrama de barras o como polígono de frecuencias.

Figura 11: Barra de menú de las distribuciones discretas con el menú de opciones desplegado.



o bien mostrar los diferentes valores de la variable en el eje X para las distribuciones continuas.

Figura 12: Barra de menú de las distribuciones continuas con el menú de opciones desplegado.



El menú de distribución permite, no solo seleccionar la distribución que se va a ejecutar sino también guardar el gráfico que se ha generado en formato BMP y poder utilizarlo en otras aplicaciones.

De esta manera, el software se ajusta estrictamente al temario de nuestra asignatura (Análisis de Datos en Psicología I), realizando todos los análisis estadísticos y gráficas del temario, además de contar con interesantes “apoyos” didácticos no disponibles en los programas profesionales (tratando temas como, por ejemplo, combinatoria y probabilidades). También los menús de ayuda han sido realizados considerando el nivel de conocimientos de nuestros alumnos y aprovechando los años de experiencia de nuestro equipo docente. Finalmente, consideramos nuestro programa como algo “vivo” que irá modificándose progresivamente e incluyendo todo aquello que facilite el aprendizaje de nuestra materia.

BIBLIOGRAFÍA:

Hays; W.L. (1988): *Statistics*. New York, Holt, Rinehart and Winston.

Merino, J.M; Moreno, E; Padilla, M; Rodríguez-Miñón, P; Villarino, A. (2001): *Análisis de Datos en Psicología I*. UNED. Madrid.