

## **SISTEMAS ESTADÍSTICOS DE PROPÓSITOS MÚLTIPLES: UNA REVISIÓN DE PROGRAMAS GRATUITOS.**

Rubén Ledesma

*Universidad Nacional Mar de Plata*

### RESUMEN

En este trabajo se ofrece una revisión de algunos sistemas estadísticos de propósitos múltiples (SEPM) de carácter gratuito. Estos sistemas pueden ofrecer una alternativa viable cuando no se disponen de recursos suficientes para la adquisición de software comercial. Se incluyen en la revisión los programas EpiInfo, ViSta, WinIDAMS, OpenSat y DataPlot, seleccionados en base a los siguientes criterios de inclusión: a) gratuidad del software y del soporte documental, b) accesibilidad desde Internet, y c) disponibilidad de técnicas estadísticas básicas, intermedias y avanzadas. La revisión incluye una breve descripción de cada sistema y un análisis de sus ventajas y desventajas comparativas. Asimismo, se analizan ciertas limitaciones de los SEPM en general y se mencionan posibles soluciones a través de otros recursos informáticos de libre acceso. Se espera que el trabajo contribuya a una mayor difusión y uso del software gratuito, especialmente entre quienes trabajan en centros que no disponen de recursos suficientes para adquirir productos comerciales.

Palabras clave: *software gratuito, sistemas estadísticos de propósitos múltiples, revisión comparativa.*

## Introducción

La utilización de métodos de encuesta, así como la gestión y análisis racional de la información obtenida por este medio, no es algo estrictamente reservado a los centros de investigación académica o profesional. Los métodos de encuesta son una estrategia habitual para obtener información con fines prácticos o evaluativos en contextos muy diversos, como pueden ser centros de salud, pequeñas empresas, organizaciones no gubernamentales, etc. No obstante, existen importantes diferencias entre la investigación académica y la investigación que puede realizarse en contextos institucionales donde el propósito básico no es la producción de conocimiento. Una diferencia importante está dada por la disponibilidad de recursos, infraestructura y apoyo financiero destinado a tareas de investigación.

Por otro lado, también es cierto que la investigación académica o científica propiamente no siempre se desarrolla en las mismas condiciones desde el punto de vista de los recursos e infraestructura disponible. Estas diferencias son evidentes, por ejemplo, si se comparan las condiciones y el apoyo financiero de la investigación en países desarrollados y en países en vía de desarrollo. Uno de los aspectos claramente afectado por las diferencias en los recursos disponibles es la adquisición de tecnología informática, dentro de lo cual incluimos el acceso a programas de análisis estadístico, herramienta fundamental cuando se trata de gestionar y analizar datos de investigación. En particular, la necesidad de tales herramientas es clara en los estudios por encuestas, donde se requieren herramientas informáticas flexibles e integradas para gestionar, manipular, transformar y analizar grandes volúmenes de datos.

Como hemos mencionado, para quienes desempeñan sus actividades de investigación en centros o instituciones que disponen de escasos recursos económicos, la adquisición de software estadístico comercial puede verse seriamente limitada. Esta limitación es patente en países en desarrollo —como son los países latinoamericanos—, sobre todo en ámbitos de trabajo oficial como universidades, ministerios, dependencias de ayuntamientos, centros de salud, etc. En estos ámbitos, puede resultar difícil cubrir los costos asociados a los productos de origen comercial, tal el caso de sistemas como *SPSS* (Spss Inc., 2003) o *Statistica* (StatSoft Inc., 2004), por mencionar algunos de los más populares. Frente a esta limitación, el software estadístico gratuito se presenta como opción cuyo principal atractivo radica, justamente, en ofrecer herramientas alternativas cuando las restricciones económicas impiden adquirir programas comerciales. Cabe mencionar que los recursos disponibles de forma gratuita han evolucionado considerablemente en el último tiempo, existiendo una gran variedad de herramientas que incluso pueden llegar a superar, en ciertos aspectos, a los productos de origen comercial.

Muchas de estas herramientas gratuitas son bien conocidas en ámbitos específicos. Por ejemplo, *EpiInfo* (CDC, 2000) es un programa gratuito muy extendido y utilizado en instituciones relacionadas con la salud, tanto para la gestión y análisis de datos provenientes de cuestionarios o encuestas, como para el mantenimiento de sistemas de información de vigilancia epidemiológica. *EpiInfo* incluye herramientas de gestión, análisis y presentación de resultados, proporcionando al investigador un ambiente simple a la vez que poderoso para realizar procesos de análisis de datos. No obstante, se trata de un

software poco difundido entre profesionales de otras disciplinas, por ejemplo, entre los investigadores procedentes de las Ciencias del Comportamiento.

Por otro lado, existen además de EpiInfo otros programas gratuitos de diferente alcance y naturaleza, entre los que se cuentan programas estadísticos especializados, paquetes estadísticos generales e incluso entornos de programación estadística completamente gratuitos, como *R* (R Development Core Team, 2002) o *XlipStat* (Tierney, 1990). Dentro de este abanico de recursos, la alternativa más atractiva para la investigación mediante encuestas la constituyen algunos *sistemas estadísticos de propósitos múltiples (SEPM)*. La principal razón de esta afirmación es que los SEPM ofrecen entornos para realizar procesos completos de *gestión, análisis y presentación de datos estadísticos*, representando el equivalente no comercial más próximo a programas profesionales como SPSS o Statistica.

Ejemplos de SEPM gratuitos son: el propio *EpiInfo* (CDC, 2000), *WinIDAMS* (UNESCO, 2003), *ViSta* (Young, 1996), *DataPlot* (NIST, 2002) y *OS3* (Miller, 2003), entre otros. Estos sistemas se caracterizan por proporcionar un paquete integrado de aplicaciones para cubrir diversas necesidades de análisis estadístico en un mismo entorno informático. De ahí la expresión *propósitos múltiples*, que diferencia a estos sistemas de otras categorías de software, como los *programas especializados*, orientados a la aplicación de un conjunto limitado de técnicas estadísticas —por ejemplo, el programa *Arc* para modelos de regresión (Cook & Weisberg, 1999)—. A diferencia de los programas especializados, los SEPM incluyen un abanico de técnicas que pueden cubrir las necesidades básicas de un usuario estándar en lo referente a gestión y análisis de datos.

Otro aspecto a destacar de los SEPM gratuitos es que pueden incluso superar en determinados aspectos a los productos comerciales. Por ejemplo, *ViSta* (Young, 1996) incluye poderosas e innovadoras herramientas gráficas para visualizar datos y resultados de modelos estadísticos multivariados (Young, Valero Mora, Faldowsky & Bann, 2000). Estas técnicas gráficas, idiosincrásicas de *ViSta*, superan en dinamismo y sofisticación a las que pueden encontrarse en sistemas comerciales como *SPSS*. No obstante, el sistema *ViSta* —así como otros programas estadísticos gratuitos—, se encuentra escasamente difundido, aún entre quienes pueden estar potencialmente interesados en acceder a herramientas de carácter gratuito.

En síntesis, hoy en día el investigador dispone de todo un arsenal de recursos gratuitos, con menor difusión que el software comercial, pero que pueden ser utilizados en reemplazo de este último cuando no se dispone de recursos financieros suficientes para su adquisición. En este contexto, nos hemos interesado en ofrecer una revisión comprensiva de los principales SEPM gratuitos actualmente disponibles, incluyendo una descripción de sus características técnicas generales y un análisis de sus posibles ventajas y desventajas comparativas desde el punto de vista de los usuarios que utilizan métodos de investigación mediante encuestas. De forma complementaria, también mencionamos algunas opciones informáticas de utilidad para ampliar las capacidades de los SEPM, como son los programas estadísticos especializados y los lenguajes de programación estadística.

De este modo, se espera contribuir a una mayor difusión del software gratuito y propiciar su uso en la investigación, especialmente entre aquellos investigadores o profesionales que pertenecen a instituciones con recursos económicos limitados. En el caso de

aquellos usuarios que ya utilizan sistemas de libre acceso, como *EpiInfo*, entendemos que el conocimiento de otras opciones de similar naturaleza puede ser de utilidad para ampliar el repertorio de herramientas de este sistema, sin necesidad de recurrir a productos comerciales.

Por último, es importante aclarar que la expresión software gratuito se utiliza para calificar a aquellos programas que pueden adquirirse en su versión completa sin ningún tipo de costo económico. Esta característica, no necesariamente implica “software de código abierto”, ni software de “dominio público”. Estas expresiones se refieren a distintos aspectos del software, aunque en la práctica puede darse el caso de programas que son a la vez gratuitos, de código abierto y de dominio público. *Código abierto* es un término que ha sido adoptado dentro de la comunidad del software libre para evitar la confusión derivada de la acepción inglesa *free software*. El significado del término código abierto es, básicamente, que puede accederse al código fuente del programa, lo cual tiene un significado más preciso que el de software libre. En el software de código abierto el código fuente está públicamente disponible, aunque el alcance de su licenciamiento puede variar considerablemente. Por otro lado, el término “software de dominio público” se refiere exclusivamente a programas que no están protegidos por derechos de autor, es decir que no poseen copyright. Para explorar en mayor detalle estas distinciones puede consultarse Free Software Foundation (2001).

## Revisión de sistemas estadísticos de propósitos múltiples (SEPM)

Como hemos mencionado, los SEPM representan la mejor opción cuando el investigador requiere un entorno flexible e integrado que le permita realizar un proceso completo de gestión, manipulación y análisis de datos. No obstante, puesto que dentro de esta categoría existen muchos y variados programas, hemos establecido algunos criterios básicos de inclusión a los efectos de acotar el número de programas revisados. En particular, hemos tomado en consideración los SEPM que cumplen con las siguientes condiciones:

1. *Estar disponible en forma completamente gratuita desde Internet*, es decir, los programas que poseen sitio web y pueden ser obtenidos en forma completa, sin restricciones y costos por este medio. No se incluyen dentro de esta categoría las versiones de demostración (*Demo*) y las versiones estudiantiles (*Student version*) gratuitas de los SEPM comerciales.
2. *Incluir técnicas estadísticas básicas, intermedias y avanzadas*. No han sido considerados aquellos programas que solo proporcionan al usuario técnicas estadísticas básicas (frecuencias, descriptivos, etc.) e intermedias (ANOVA y regresión). Este tipo de programas, como *SSP* (Smith, 2004) y *StatCrunch* (West, Wu & Heyd, 2004), representan una buena alternativa para ser utilizados con fines didácticos, pero pueden no ser suficientes para cubrir las necesidades de los usuarios e investigadores profesionales.
3. *Tener documentación de usuario gratuita*, es decir, que el usuario pueda acceder sin costo a las guías, manuales de referencia u otra documentación sobre el software. De este modo, han quedado fuera de la revisión algunos programas gratui-

tos que vienen integrados a textos de distribución comercial, como el programa *Arc* (Cook & Weisberg, 1999).

Finalmente, los SEPM identificados y que cumplen con estos criterios son: *EpiInfo*, (CDC, 2000) *WinIDAMS* (UNESCO, 2002), *OpenStat* (Miller, 2003), *ViSta* (Young, 1996) y *DataPlot* (NIST, 2002). A continuación, se ofrece una breve descripción de cada sistema y un análisis de sus posibles ventajas y desventajas comparativas. Los interesados en obtener o documentarse sobre estos SEPM pueden consultar las direcciones URL listadas en la tabla 1.

Tabla 1: direcciones URL de los SEPM gratuitos incluidos en la revisión

<i>Denominación (Autor)</i>	<i>Dirección URL</i>
DataPlot (NIST, 2002)	<a href="http://www.itl.nist.gov/div898/software/dataplot/">http://www.itl.nist.gov/div898/software/dataplot/</a>
EpiInfo (CDC, 2000)	<a href="http://www.cdc.gov/epiinfo/">http://www.cdc.gov/epiinfo/</a>
OS3 (Miller, 2003)	<a href="http://www.statpages.org/miller/openstat/OS3.html">http://www.statpages.org/miller/openstat/OS3.html</a>
ViSta (Young, 1996)	<a href="http://forrest.psych.unc.edu/research/index.html">http://forrest.psych.unc.edu/research/index.html</a>
WinIDAMS (UNESCO, 2002)	<a href="http://www.unesco.org/idams">http://www.unesco.org/idams</a>

### Breve descripción de EpiInfo

*EpiInfo* es quizá el sistema estadístico más popular entre los investigadores de Ciencias de la Salud. Creado por el CDC de Atlanta (USA), es un programa disponible en varios idiomas y su última versión (*EpiInfo* 2000) funciona en Windows (95, 98, NT, o 2000), con algunos requerimientos básicos (32 Mb de Memoria RAM, 50 Mb de espacio libre en el disco duro y Procesador recomendado de 200 megahercios).

*EpiInfo* incluye técnicas de gestión, análisis y presentación de resultados y aplicaciones específicas para epidemiología y otras áreas de la salud. Posee una interfaz de usuario estructurada gráficamente —o interfaz gráfica de usuario (*IGU*)— resultando sencillo de utilizar aún para los menos expertos. Entre sus componentes básicos más destacados encontramos un asistente para desarrollar cuestionarios (*MakeView*), un programa de ingreso y validación de datos (*Enter*), un módulo para depurar, transformar y analizar datos (*Analysis*) y un programa de mapeo compatible con los algunos sistemas GIS. Sus módulos se pueden instalar individual o colectivamente, incluyendo en la instalación un manual completo y un tutorial de ayuda.

### *Ventajas y limitaciones de EpiInfo*

Para usuarios en Ciencias de la Salud, *EpiInfo* posee ventajas comparativas con respecto a otros SEPM, pues fue diseñado y desarrollado propiamente dentro del ámbito de la salud pública, encontrándose ampliamente difundido en este medio. Además, aunque su última versión solo funciona en Windows, está disponible en varios idiomas, posee

documentación de usuario completa y es un sistema abierto. Este último punto resulta de gran interés, ya que implica que puede ser extendido por los usuarios expertos interesados en ampliar o adaptar las capacidades del sistema a sus propias necesidades.

Como desventaja de *EpiInfo*, debemos mencionar su limitación en cuanto a disponibilidad de métodos multivariados y modelos estadísticos avanzados. *EpiInfo* no posee métodos multivariados de uso habitual en la investigación mediante encuestas, como Análisis Factorial o de Clasificación. Esto representa una limitación importante para los usuarios o analistas avanzados, quienes pueden requerir este tipo de técnicas. Por otro lado, en comparación con otros SEPM gratuitos, *EpiInfo* también presenta ciertas limitaciones en lo referente a técnicas de visualización de datos, especialmente en la categoría de métodos gráficos multivariados.

### **Breve descripción de WinIDAMS**

WinIDAMS es un programa creado por la UNESCO en cooperación con varios expertos de diferentes países. WinIDAMS es la versión de IDAMS “Internationally Developed Data Analysis and Management Software Package” para Windows. El programa está disponible en Inglés, Francés y Español, y corre en Windows 95, 98, NT, 2000 y XP. Para su correcto funcionamiento, se recomiendan algunos requisitos básicos (procesador equivalente a Pentium II o superior, 64 Mb de memoria RAM y 11 Mb de espacio libre en disco).

Este sistema incluye una serie de métodos de gestión y técnicas de análisis que se ejecutan de forma interactiva, mediante IGU. Además, incluye una gran variedad de técnicas estadísticas avanzadas, entre las que se cuentan Análisis de Clasificación, Análisis Factorial, Análisis Discriminante y Escalamiento Multidimensional. Dichas técnicas se ejecutan por medio de subprogramas sin interfaz gráfica. Estos programas permiten definir y controlar las opciones de los diferentes análisis utilizando una sintaxis específica que, si bien no es compleja y está bien documentada, puede representar una limitación para los usuarios acostumbrados a utilizar programas con IGU.

### *Ventajas y limitaciones de WinIDAMS*

WinIDAMS, en comparación con otros SEPM como *EpiInfo*, incluye un gran variedad de técnicas multivariadas y modelos estadísticos avanzados, aspecto ventajoso considerando las necesidades de los analistas o usuarios expertos. Además, el proyecto es de reciente desarrollo y se muestra muy activo, con lo cual es esperable que siga avanzando y ampliando las capacidades de gestión y análisis de datos. Por otro lado, WinIDAMS está muy bien documentado y se encuentra disponible en castellano.

Con respecto a sus posibles desventajas, debemos mencionar que la mayoría de sus técnicas de análisis no disponen de IGU para su aplicación, razón por la cual el usuario debe ocupar algo de tiempo en aprender a familiarizarse con la sintaxis de cada método para definirla según sus necesidades. Otra desventaja comparativa de WinIDAMS es que solo está disponible para ser utilizado con sistema operativo Windows.

## Breve descripción del programa OpenStat (OS3)

OS3 es un sistema desarrollado totalmente por Bill Miller (Miller, 2003) en la Universidad Estatal de Iowa (USA). El sistema funciona en Windows (95/98/NT) y Linux, y solo está disponible en idioma inglés. Fue creado originalmente para Ciencias del Comportamiento y con fines educativos, motivo por el cual resulta altamente interactivo y simple de utilizar. El programa cubre una extensa familia de técnicas estadísticas, diferenciándose de otros programas de esta categoría por sus recursos de análisis para cuestionarios, escalas, test y otras medidas basadas tanto en Teoría Clásica de los Tests como en Teoría de Respuesta al Item. Entre estas opciones de análisis incluye métodos multivariados avanzados, como Análisis Factorial, de Clasificación y Modelos Causales. Todas sus opciones de análisis estadístico, incluyendo los métodos multivariados, se ejecutan mediante una IGU simple e intuitiva. El programa se encuentra bien documentado, posee un manual de usuario en línea completo desarrollado por su propio creador.

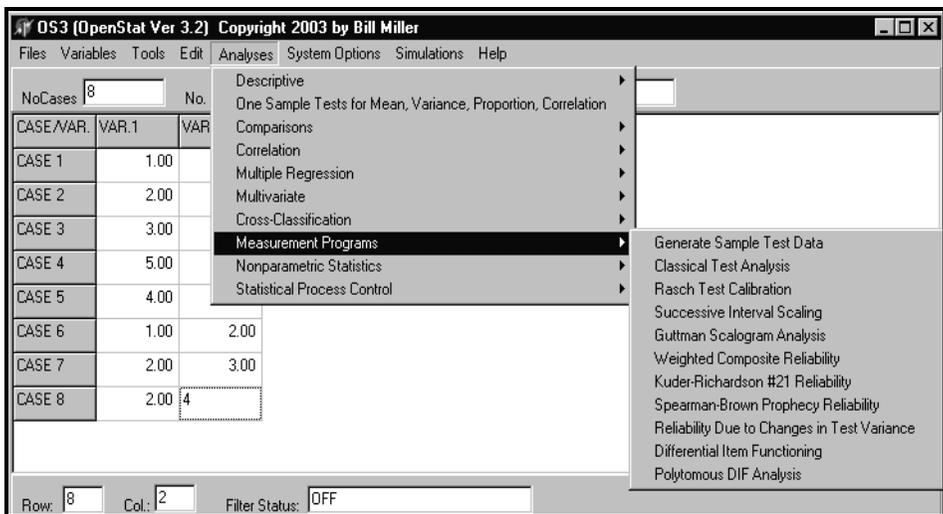


Figura 1: imagen parcial del programa Open Stat (OS3).

### Ventajas y limitaciones de OpenStat (OS3)

Este sistema es muy recomendable para investigadores provenientes de las Ciencias del Comportamiento. La interfaz de usuario es muy similar a los sistemas comerciales más populares en estas disciplinas (como SPSS o Statistica), con lo cual puede resultar muy accesible para los profesionales o estudiantes familiarizados con estos sistemas. En tal sentido, también hemos desatacado que posee un amplio repertorio de métodos de análisis de uso habitual en Ciencias del Comportamiento, entre los que se cuentan aplicaciones para propósitos psicométricos. Por su lado, los usuarios del ámbito de la salud familiarizados con programas como EpiInfo, pueden no encontrar en OS3 un entorno apropiado para realizar ciertos análisis específicos, como mapeos o cálculos estadísticos propios de epidemiología. Además, el programa solo está disponible en idioma inglés.

Por último, en comparación con otros SEPM, *OS3* también es bastante limitado en lo que respecta a sus métodos de visualización estadística de datos.

### Breve descripción del programa ViSta “The Visual Statistics System”

*ViSta* es un programa creado por el profesor Forrest W. Young en la Universidad de Carolina del Norte en Chapel Hill (Young, 1996). Posee versiones en diferentes idiomas (Inglés, castellano y Francés) y para diferentes plataformas (Mac, Windows y Unix). Se trata de un programa diseñado originalmente como herramienta para la investigación y docencia en métodos de visualización estadística multivariada, aunque actualmente posee características que lo definen como un SEPM.

*ViSta* posee una interfaz completamente gráfica y muy interactiva, que puede ser utilizada sin problemas por analistas noveles. El programa destaca por sus métodos gráficos dinámicos, que se utilizan para visualizar simulaciones de procesos aleatorios, diferentes estructuras de datos, transformaciones y resultados de modelos estadísticos. *ViSta* es abierto y extensible, esto significa que proporciona acceso al código fuente y herramientas de programación para que los usuarios avanzados puedan expandir las capacidades del programa. Por último, el sistema dispone de documentación básica de usuario que puede obtenerse en línea.

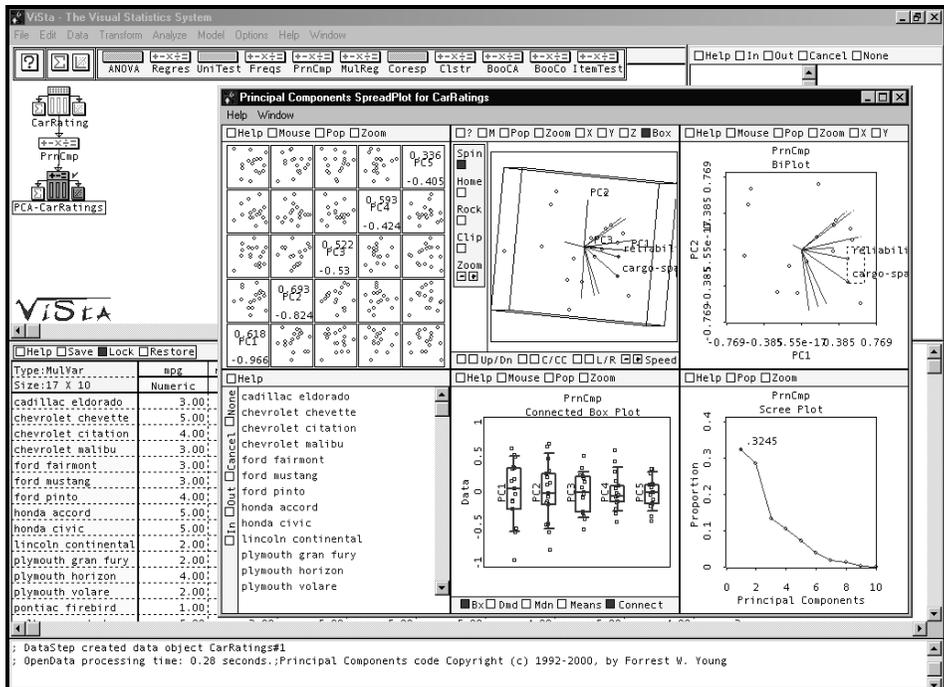


Figura 2: imagen parcial del sistema ViSta “The Visual Statistics System”.

### *Ventajas y limitaciones de ViSta*

*ViSta* es un programa multiusuario, puede ser utilizado por principiantes, docentes o investigadores avanzados. En este sentido, ofrece una interfaz completamente gráfica adecuada para analistas noveles, pero también un poderoso entorno de programación para expertos. La principal ventaja comparativa de este sistema está dada por sus métodos de visualización multivariada, altamente interactivos e innovadores. Estos métodos incluso superan a los disponibles en la mayoría de los SEPM comerciales. Como en el caso de *OS3*, *ViSta* puede ser un sistema adecuado para investigadores en Ciencias del Comportamiento. Además, los usuarios de otros programas especializados como *EpiInfo*, pueden encontrar en *ViSta* una poderosa herramienta complementaria, sobre todo si se requieren técnicas más sofisticadas de visualización y análisis exploratorio de datos.

En relación a las posibles desventajas comparativas de este sistema, debemos mencionar ciertas limitaciones en sus capacidades de gestión de datos, aspecto fundamental para el proceso de datos en la investigación por encuestas. Si bien *ViSta* ofrece herramientas básicas para crear, editar y realizar operaciones con archivos de datos, no es equiparable en este aspecto a otros SEPM. Otro punto débil de este programa es su documentación de usuario, ya que solo está disponible en idioma inglés y se encuentra algo desactualizada e incompleta.

### **Breve descripción del programa DataPlot**

*DataPlot* es un programa desarrollado por el National Bureau of Standards de los EEUU (actualmente el National Institute of Standards and Technology), está disponible en idioma inglés y funciona en varias plataformas (Unix, Linux, Windows, etc.). Los desarrolladores de *DataPlot* recomiendan equipos de nivel Pentium para utilizar el programa con IGU, se exige además aproximadamente 100 MB de espacio libre en disco.

El programa puede ser utilizado mediante un sistema de comandos en línea, pero también mediante IGU en ciertas plataformas. Se trata además de un sistema extensible, que incluye opciones para crear macros y extender la interfaz gráfica del programa, entre otras posibilidades. En lo referente a sus capacidades de análisis, dispone de gran cantidad y variedad de métodos estadísticos básicos y avanzados, y proporciona una amplia gama de opciones de cálculo y análisis matemático para investigadores en ciencias e ingeniería. En este sentido, puede decirse que *DataPlot* es un programa que en muchos aspectos excede las capacidades de análisis de los SEPM convencionales. Por último, el programa se encuentra muy bien documentado, posee un manual de usuario completo disponible en línea.

### *Ventajas y limitaciones de DataPlot*

En comparación con otros SEPM, se trata de un sistema realmente amplio y completo en sus capacidades de análisis, que excede las aplicaciones de un sistema estadístico convencional. *DataPlot* es un sistema multiplataforma y puede ser utilizado mediante comandos o mediante una IGU. En ambos casos, el usuario dispone de documentación técnica completa y detallada, aunque solo disponible en idioma inglés. No obstante ser un sistema poderoso, *DataPlot* puede presentar ciertas desventajas desde

un sistema poderoso, *DataPlot* puede presentar ciertas desventajas desde el punto de vista de los usuarios acostumbrados a programas estadísticos estándar, por ejemplo, el programa abunda en métodos que son de escaso interés en la investigación por encuestas. En este sentido, muchos investigadores pueden no encontrar en este software un entorno amigable para realizar un proceso de análisis de datos.

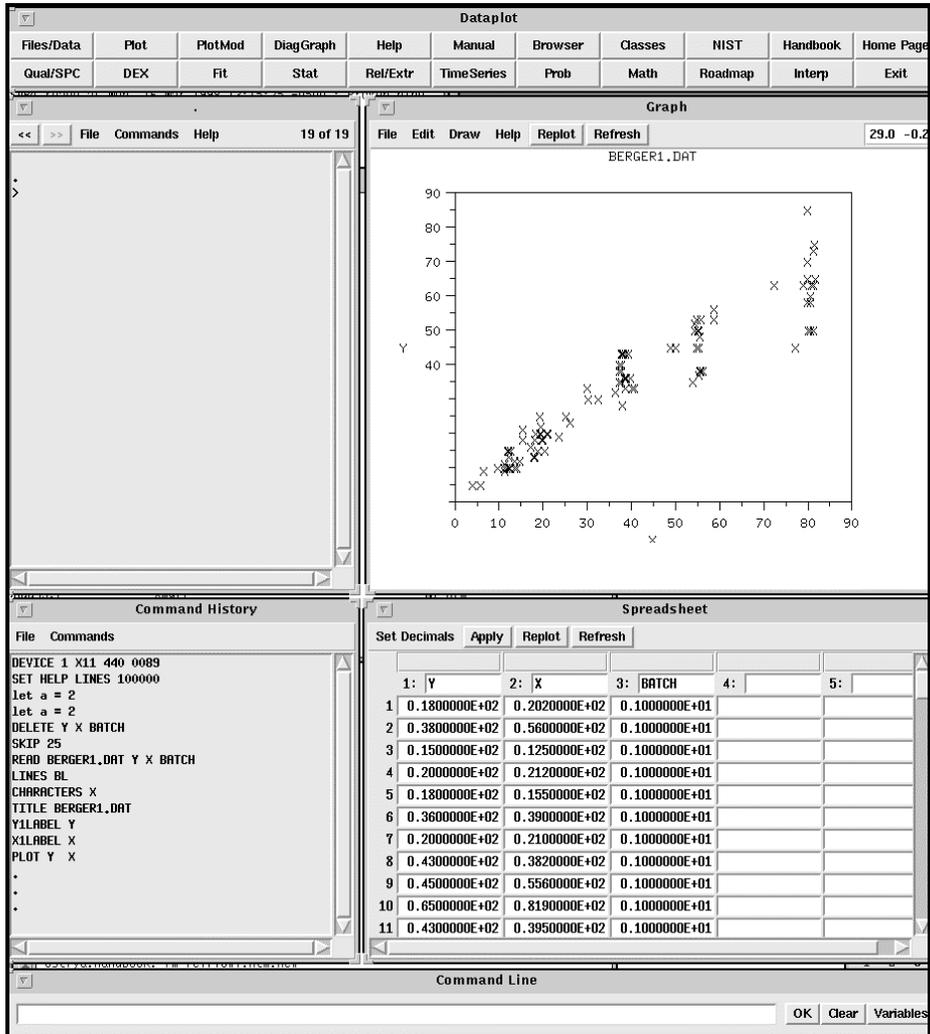


Figura 3: imagen parcial del programa DataPlot.

## Limitaciones de los SEPM gratuitos y posibles soluciones

Si bien los SEPM poseen la ventaja de brindar un entorno básico y general de trabajo para realizar procesos completos de análisis de datos, pueden no cubrir en algunos casos las necesidades de los usuarios que buscan ciertas aplicaciones o métodos específicos. Es decir, puede suceder que el investigador no encuentre en los SEPM mencionados las técnicas estadísticas específicas para satisfacer sus necesidades (de hecho, resulta prácticamente imposible que un sistema cubra todos los métodos de análisis estadísticos disponibles en la actualidad).

Frente a esta situación, es importante mencionar la existencia de otros recursos informáticos gratuitos que pueden ser de utilidad para complementar o ampliar las capacidades de los SEPM, o incluso para ser utilizadas por sí mismas. Estos recursos pueden agruparse en dos categorías: a) los *Programas estadísticos especializados*, que incluyen grupos de técnicas específicas para responder a ciertos objetivos de análisis, y b) los *Entornos de programación estadística*, que pueden utilizarse para extender algunos de los sistemas existentes o para desarrollar nuevas aplicaciones.

Dentro de la categoría de *Programas estadísticos especializados* encontramos una gran variedad de recursos gratuitos, entre los que podemos mencionar programas como *ADE-4* (Thioulouse, Chessel, Dolédec & Olivier, 1997) con métodos de análisis multivariado, *Demetra* (Eurostat, 2002) para ajustar modelos estacionales, *PS* (Dupont & Plummer, 1997) para el cálculo de la potencia y el tamaño muestral, etc. Estos programas llegan a superar a los sistemas comerciales en lo referente a sus técnicas específicas, pudiendo ser utilizados para complementar los métodos de algunos de los SEPM descriptos anteriormente. Por ejemplo, *ADE-4* constituye una alternativa interesante para complementar las opciones multivariadas de EpiInfo, ya que proporciona al usuario técnicas tales como Análisis Factorial, Métodos para datos espaciales, Análisis Discriminante y Métodos de visualización multivariada. No obstante, también tenemos que mencionar que este tipo de programas especializados solo representan una alternativa viable para el investigador cuando este requiere un conjunto limitado de métodos estadísticos para resolver sus problemas de análisis. En caso contrario, el usuario debe identificar, aprender y utilizar varios programas diferentes, según sus necesidades.

Cuando no se encuentra lo que se necesita en los SEPM o en los programas especializados, existe la opción avanzada de recurrir a ciertas herramientas de programación estadística, cuya limitación práctica es la necesidad de cierto dominio en la materia. Con este fin, están disponibles ciertos entornos de programación estadística que pueden obtenerse directamente desde Internet, como *Xlisp-Stat* (Tierney, 1990) o *R* (R Development Core Team, 2002). *Xlisp-Stat* implementa un lenguaje estadístico basado en *Xlisp*, una versión libre del lenguaje *Lisp*. El sistema *ViSta* y otros de acceso gratuito como *Arc* (Cook & Weisberg, 1999) están escritos en este lenguaje. Otra alternativa similar es *R*, un lenguaje y entorno de programación que actualmente se muestra muy activo y en pleno desarrollo. Ambos casos, *Xlisp-Stat* y *R*, resultan de gran interés para los programadores y usuarios expertos, quienes pueden desarrollar y distribuir libremente sus propias herramientas de análisis estadístico.

## Conclusiones

En la actualidad, el software estadístico gratuito constituye una alternativa viable frente al software de origen comercial. En el último tiempo, las herramientas informáticas gratuitas han evolucionado considerablemente, pudiendo incluso llegar a superar en ciertos aspectos a los productos comerciales. Este desarrollo es de especial interés para aquellas entidades o centros de investigación que disponen de recursos económicos limitados o escasos, lo cual restringe el acceso a los productos informáticos profesionales de distribución comercial.

Dentro de los recursos informáticos gratuitos, los denominados *Sistemas Estadísticos de Propósitos Múltiples (SEPM)* constituyen la alternativa más atractiva para el investigador que utiliza métodos de encuesta, ya que brindan ambientes amigables y poderosos para realizar procesos de carga, edición, transformación, análisis y presentación de datos estadísticos. Así, estos sistemas ofrecen al investigador un paquete integrado de aplicaciones que le permiten cubrir muchas de sus necesidades de análisis estadístico en un mismo entorno informático. En este trabajo nos hemos centrado en la descripción y análisis comparativo de algunos de estos programas, seleccionados por ser completamente gratuitos, estar disponibles en Internet, poseer técnicas estadísticas básicas, intermedias y avanzadas, y disponer de documentación gratuita para usuarios.

Los SEPM revisados han sido *EpiInfo*, *WinIdams*, *OpenStat*, *Vista* y *DataPlot*. Además de las ventajas y desventajas comparativas de estos sistemas, se han mencionado ciertas limitaciones generales de los SEPM, como la dificultad evidente de cubrir todas y cada una de las necesidades que pueden presentarse en la práctica. Frente a tal situación, hemos querido hacer referencia a otras herramientas gratuitas complementarias, que pueden servir de apoyo a los SEPM. En particular, se ha comentado brevemente la existencia de ciertos programas especializados de carácter gratuito, a partir de los cuales puede accederse a técnicas especializadas no incluidas en los SEPM. Asimismo, se ha hecho una breve referencia a determinados entornos de programación estadística gratuitos que el usuario experto puede utilizar para desarrollar sus propios métodos estadísticos, cuando estos no se encuentran disponibles en los SEPM o en los programas especializados.

En cualquier caso, podemos presuponer que las limitaciones actuales de los SEPM gratuitos serán progresivamente superadas, considerando la vitalidad y el crecimiento continuo que muestran algunos de estos proyectos. En tal sentido, esperamos que el presente trabajo contribuya a una mayor difusión, conocimiento y uso de estas herramientas, en especial, entre quienes desarrollan sus actividades de investigación en centros o entidades que poseen limitaciones para adquirir software estadístico de distribución comercial.

## Referencias

- Centers for Disease Control and Prevention (2000). *EpiInfo 2000, Manual de usuario*. Atlanta: CDC.
- Cook, R. D. & Weisberg, S. (1999). *Applied regression including computing and graphics*. New York: John Wiley & Sons.

- Dupont, W. D & Plummer, W. D. (1997). PS power and sample size program available for free on the Internet. *Controlled Clinical Trials*, 18, 274.
- Eurostat (2002). *Demetra 2.03*. [programa informático] [en línea]. Disponible: [URL: http://forum.europa.eu.int/irc/dsis/eurosam/info/data/demetra.htm](http://forum.europa.eu.int/irc/dsis/eurosam/info/data/demetra.htm) [consulta: 02/9/04]
- Free Software Foundation (2001). Categories of Free and Non-Free Software [en línea]. Disponible: URL: <http://www.gnu.org/philosophy/categories.es.html> [consulta: 12/11/04]
- Miller, W. G. (2003). *OpenStat 3* [programa informático] [en línea]. Disponible: URL: <http://www.statpages.org/miller/openstat/OS3.html> [consulta: 09/09/04]
- NIST-National Institute of Standards and Technology (2002). *DataPlot software*. [programa informático] [en línea]. Disponible: URL: <http://www.itl.nist.gov/div898/software/dataplot/>
- R Development Core Team (2002). *R1.7.1* [programa informático] [en línea]. Disponible: URL: <http://www.r-project.org/> [consulta: 12/10/04]
- Smith, G. (2004). *Smith's Statistical Package*. [programa informático] [en línea]. Disponible: URL: <http://www.economics.pomona.edu/StatSite/SSP.html> [consulta: 22/04/04]
- Spss Inc. (2003) *SPSS 12.0 Base user's guide*. NJ: Prentice Hall.
- StatSoft, Inc. (2004). *Electronic Statistics Textbook*. Tulsa, OK: StatSoft [en línea] Disponible: URL: <http://www.statsoft.com/textbook/stathome.html> [consulta: 18/04/04]
- Thioulouse J., Chessel D., Dolédec S., & Olivier, J. M. (1997). ADE-4: a multivariate analysis and graphical display software. *Statistics and Computing*, 7(1), 75-83.
- Tierney, L. (1990). *Lisp-Stat An object-oriented environment for statistical computing and dynamic graphics*. NY: John Wiley & Sons.
- UNESCO (2002). *WinIDAMS Reference Manual*. 6ta. Edición. París: The United Nations Educational, Scientific and Cultural Organization.
- West, R. W., Wu, Y. & Heyd, D. (2004). An Introduction to StatCrunch 3.0. *Journal of Statistical Software*, 9(5), sn, [en línea] Disponible: URL: <http://www.jstatsoft.org/> [consulta: 12/05/04]
- Young, F. W. (1996). *ViSta: The Visual Statistics System*. UNC L.L. Thurstone Psychometric Laboratory, Research Memorandum 94-1, 1996.
- Young, F., Valero Mora, P., Faldowsky, R. A. & Bann, C. (2000). *SpreadPlots*. UNC L.L. Thurstone Psychometric Laboratory, Report Number 4, May 2000.