

## 4. INTRODUCCIÓN A LOS PROGRAMAS DE CÓMPUTO

“Había una vez...” Así empiezan los cuentos. Pero ésta es una historia que se vivió hasta mediados del siglo XX. Durante muchos años, el ejercicio de la estadística incluyó el trabajo penoso de tabular datos durante horas antes de realizar la primera operación, lo cual por lo general se hacía con la ayuda de una calculadora (pero también con los dedos). Este trabajo era tardado, cansado y, en ocasiones, aburrido. A lo anterior se agregaban los errores frecuentes que obligaban a repetir los conteos, o... ¡A redondear los datos! Las ecuaciones estadísticas estaban limitadas al cálculo de proporciones y promedios, y las regresiones no pasaban de dos variables. Los cálculos de probabilidad ( $p$ ) en que se basaban las inferencias se apoyaban en las tablas de los anexos al final del libro, y sólo algunos pocos matemáticos expertos se atrevían a realizar pruebas más avanzadas. La entrega de resultados siempre se retrasaba y con frecuencia quedaba poco tiempo para su discusión. Y así reinó el caos durante años, hasta que un día llegaron las computadoras; al principio las cosas no fueron más fáciles. Las primeras computadoras, que eran equipos muy grandes y más costosos, estaban al cuidado de un grupo selecto de iniciados: los programadores. Los humildes mortales sólo las podían ver a través de los grandes cristales que las protegían del polvo y del calor. Fue en esa época, cuando las computadoras podían “equivocarse”, que aparecieron los primeros programas de cómputo especialmente diseñados para el análisis estadístico, pero ese recurso estaba dedicado a cumplir funciones administrativas y todavía se encontraba alejado del investigador. Por tanto, la entrega de resultados aún se retrasaba y con frecuencia tampoco había tiempo para su discusión. Pero luego sucedió lo que nadie esperaba: las computadoras llegaron a los escritorios de las oficinas y a las mesas de las casas y no sólo ellas, también llegaron los programas de cómputo. Los programadores salieron de la vista y los propios investigadores podían hacer los análisis sin necesidad de intermediarios, pero la entrega de resultados sigue retrasándose y con frecuencia sigue sin haber tiempo para la discusión. Sólo que ahora ya se han encontrado a las culpables: las computadoras. Gracias al contacto con los humanos, estos equipos han adquirido características propias de ellos: además de “equivocarse” también se entretienen “echando a perder” los archivos de datos.

Hemos querido empezar este capítulo ironizando porque es importante dejar bien claro que la computadora es una herramienta muy útil en el análisis estadístico de los datos, pero que no piensa ni actúa por sí sola. La computadora sigue instrucciones en lenguaje binario, y las cumple a una velocidad que no deja de ser impresionante. Si se le alimenta correctamente y las instrucciones que se le dan son las adecuadas, los resultados serán sorprendentes. Pero si esto se hace incorrectamente los resultados también serán sorprendentes por la magnitud de los errores. Vale la pena insistir: la computadora es una herramienta y en ningún momento sustituye la labor del investigador en el análisis estadístico. Con la ayuda de la computadora se agiliza la tabulación, y también las operaciones, pero es el investigador quien tiene que decidir qué análisis es el más adecuado para sus datos, y es él quien tiene que interpretar los resultados. Si lo anterior ha quedado claro, entonces ya se puede continuar con el tema de los programas de cómputo que pueden ser de gran ayuda en el análisis estadístico.

Durante el análisis estadístico el investigador tiene que realizar varias tareas:

- Presentación de la propuesta o proyecto.
- Búsqueda y registro de datos.
- Captura y transformación de datos.
- Revisión de la captura.
- Tabulación de datos.
- Cálculo de estadísticos.
- Interpretación de resultados.
- Reporte final del trabajo.

Para realizar estas tareas el investigador llega a utilizar varios programas de cómputo. Algunos son exclusivos del trabajo estadístico, otros son menos específicos. En términos generales, entre los programas de cómputo que se llegan a utilizar se incluyen los procesadores de texto, los administradores de bases de datos, las hojas de cálculo, los programas para presentaciones y los programas estadísticos. A esta lista también podríamos agregar algunas utilerías y páginas de la WEB que funcionan como calculadoras estadísticas/epidemiológicas. El uso que se hace de ellos difiere de muchas maneras, pero con frecuencia realizan tareas comunes. A continuación se harán comentarios sobre algunos programas de cómputo. Es necesario aclarar que no se trata de una revisión sistemática de los programas existentes. Más bien, se escribe sobre algunos de los que los autores han tenido oportunidad de utilizar.

## **PROCESADORES DE TEXTO**

Estos programas son los más genéricos de todos los que se utilizan. Básicamente sirven para escribir las propuestas iniciales, reportes finales del trabajo y formas para captar datos (cuestionarios, cédulas de captura). Algunos incluyen utilerías capaces de elaborar cuadros y gráficos. Otros hasta permiten escribir fórmulas como las que se presentaron en este libro. Entre estos programas destaca Word para Windows.

## **ADMINISTRADORES DE BASES DE DATOS**

Cumplen una función muy importante durante el trabajo estadístico: ayudan a capturar datos en los archivos de cómputo en los que se almacenan. Además de generar la estructura de la base y de permitir capturar los datos, estos programas también permiten editar y transformar datos, así como generar nuevas variables y asignarles valores a partir de los existentes. Pueden ayudar a generar pantallas de captura que faciliten el trabajo de introducir datos limitando el número de errores que se pueden cometer. La verificación de la captura se puede realizar comparando los datos en pantalla contra la forma en la que se registraron los datos. Si existiera algún error el programa permite corregirlo. Si alguien cuenta con un poco de experiencia en programación también puede utilizar estos programas para tabular datos y realizar operaciones estadísticas, como el cálculo de proporciones, promedios y desviaciones estándar, o pruebas de chi-cuadrada y *t* de Student, por ejemplo. El potencial de estos programas en el análisis estadístico es enorme, pero la limitación más importante reside en que es necesario conocer los programas con mucha profundidad y saber programar. Entre los programas de esta categoría están dBase, Fox y Access.

## HOJAS DE CÁLCULO

Su estructura de celdas, construidas a partir de columnas y renglones, además de las funciones que incluyen, las hace muy útiles para el trabajo estadístico. Entre estos programas los más conocidos son Excel y Lotus. Su primera aplicación suele ser la captura de datos, y es mucho más sencilla de la que se puede hacer en los administradores de bases de datos, aunque menos versátil. Para capturar datos en una hoja de cálculo se procede a identificar las columnas con las variables y a los renglones con los registros. En las celdas del primer renglón se anotan los nombres de las variables y, a partir del segundo renglón, se capturan los datos que corresponden a cada elemento del grupo. Por brevedad en la captura suelen utilizarse códigos y no etiquetas. Por ejemplo, en vez de escribir “masculino” en la columna de sexo se puede anotar “1”, y en lugar de “femenino” el código puede ser “2”. De esta manera se ahorra mucho tiempo y se reduce el número de errores.

### Ejemplo explicativo 4-1. -----

Recuérdese el Cuadro 1-1 en el que se presentaron los datos de un grupo de trabajadores de un taller de automóviles. Los mismos datos en una hoja de cálculo tendrían la presentación del cuadro 4-1, en el cual los renglones (identificados con un número en la primera columna, en gris) y las columnas (identificadas con letras mayúsculas en el primer renglón, también en gris) definen las referencias de cada celda. En el renglón 1 se identifican las variables: “registro” en la columna “A”, “sexo” en la columna “B”, “edad” en la columna “C” y hemoglobina con la etiqueta “hg” en la columna “D”. Bajo la etiqueta “registro” se anota el número *i*-ésimo que corresponde a la captura, generalmente un número progresivo, en sustitución del nombre del sujeto o elemento del conjunto que se estudia. Bajo la etiqueta “sexo” el código “1” significa “varón” mientras que el “2” se refiere a “mujer”. Para las celdas debajo de “edad” y “hg” se anotan los valores que corresponden a cada lectura.

Cuadro 4-1. Representación de la captura de datos del Cuadro 1-1 en una hoja de cálculo.

	A	B	C	D	E
1	registro	sexo	edad	hg	
2	1	1	45	12	
3	2	2	23	13	
4	3	1	32	11	
5	4	1	18	15	
6	5	1	21	14	

-----

La verificación de datos puede realizarse de varias maneras. En primer lugar, los datos en la pantalla se pueden comparar contra la forma de la cual se registraron los datos. Otro procedimiento más ventajoso por su facilidad es la utilización de filtros, a partir de los cuales se pueden ordenar los datos según diferentes criterios. Por ejemplo, si los datos del cuadro 4-1 fueran muy extensos, entonces se podría pedirle al programa que los ordenara según los valores anotados en la columna del sexo. Si por algún motivo alguien hubiera anotado un tres en esa columna se podría identificar al principio o al final de ella después de ordenar los registros según un criterio ascendente o

descendente, respectivamente. Los filtros reconocen todos los criterios o secuencia de datos capturados en cada variable, permitiendo identificar cualquier error de registro, como el número tres capturado en la variable sexo del ejemplo anterior. Una tercera opción para verificar la captura incluye registrar dos veces los datos en la misma secuencia: la primera en una hoja y la segunda en otra hoja del mismo archivo. Para comparar las capturas se utilizaría una tercera hoja que contrastaría las dos primeras y destacara las capturas que no fueran iguales.

**Ejemplo explicativo 4-2.** -----

Supóngase que se han capturado los datos del ejemplo 4-1 por duplicado, tal como se presentan en el cuadro 4-2. Si se observa con detenimiento se puede encontrar, en primer plano, la misma captura que en el cuadro 4-1. En segundo plano se aprecian los primeros dos renglones con los mismos datos excepto dos cambios: en las celdas B2 y C3 los datos no son los mismos que en la primera hoja. En el tercer plano se muestra una hoja que identifica la diferencia de valores y los señala con la palabra "ERROR". Para que el programa Excel muestre las diferencias de la captura en la tercera hoja se puede utilizar la función =SI() en cada una de las celdas de la tercera hoja (que es la que compara): en la celda A1 se escribe la fórmula =SI(Hoja1!A1<>Hoja2!A1;"ERROR";" "), y luego se copia a todas las demás celdas en la hoja de cálculo. Una vez que las capturas diferentes han sido identificadas se busca en las formas de captura cuál de las dos es la correcta para cambiarla en la hoja en la cual se tiene el error.

Cuadro 4-2. Representación de la captura de datos del Cuadro 1-1, por duplicado, en hojas de cálculo.

	A	B	C	
1	registro	sexo	edad	
2		ERROR		
3			ERROR	

	A	B	C	D
1	registro	sexo	edad	hg
2	1	2	45	12
3	2	2	26	13

	A	B	C	D	E
1	registro	sexo	edad	hg	
2	1	1	45	12	
3	2	2	23	13	
4	3	1	32	11	
5	4	1	18	15	
6	5	1	21	14	

Las hojas de cálculo incluyen funciones que, a partir de un bloque de datos, permiten realizar varias operaciones, entre las que se encuentran los cálculos de la media, mediana, moda, varianza, desviación estándar, coeficiente de correlación, intercepción y pendiente de la regresión lineal. También puede mostrar los valores de varias

distribuciones de probabilidad, como la normal, binomial, Poisson,  $t$ ,  $F$ ,  $\chi^2$ , así como realizar las pruebas de chi-cuadrada,  $t$  de Student,  $F$  y  $z$ . Varias funciones matemáticas y trigonométricas también están incluidas en las hojas de cálculo, y éstas son muy útiles cuando se desea redondear los resultados o transformar los valores de las variables mediante logaritmos o antilogaritmos, por ejemplo. Una función que puede ser de mucha utilidad está relacionada con la generación de números aleatorios. Las hojas también pueden servir para simplificar el trabajo, de tal manera que una vez que se ha introducido una fórmula ya no sea necesario capturarla nuevamente, sólo es necesario cambiar los valores que resulten de procesos intermedios para actualizar el resultado, derivado de la fórmula, automáticamente.

**Ejemplo explicativo 4-3.** -----

El investigador con frecuencia tiene que hacer muchas veces las mismas operaciones, pero con valores diferentes. Por ejemplo, cuando se prepara un reporte final podría existir interés en presentar los resultados mediante diferencias de proporciones con intervalos de confianza de 95%. En la hoja de cálculo se pueden poner los valores y las fórmulas necesarias una vez, y después sólo modificar los valores para que cambien los intervalos de confianza de acuerdo con cada cuadro. En el cuadro 4-3 se representa una hoja de cálculo mediante la cual es posible calcular los intervalos de confianza. En las celdas de la columna B se anotan las proporciones y el número de observaciones en las muestras A y B, mientras que en la columna E se colocan las ecuaciones para los resultados: la diferencia de proporciones se obtendría mediante la fórmula (en Excel) =B1-B3 en la celda E1; el límite inferior del intervalo de confianza se obtendría en la celda E3 mediante =E1-1.96\*(B1\*(1-B1)/B2+B3\*(1-B3)/B4)^0,5; mientras que el límite superior se tendría en la celda E4 con la fórmula =E1+1.96\*(B1\*(1-B1)/B2+B3\*(1-B3)/B4)^0,5. Estas fórmulas se presentan con mayor detalle en el Capítulo 16.

Cuadro 4-3. Presentación de proporciones e intervalos de confianza del 95%.

	A	B	C	D	E
1	Proporción A	0.5		Pa - Pb	0.2
2	n de A	150		IC 95%	
3	Proporción B	0.3		límite inferior	0.095
4	n de B	175		límite superior	0.305
5					
6					

Una vez que ya se tiene la hoja con los valores y las ecuaciones en posición, para el resto de los cálculos sólo es necesario cambiar los valores de B1, B2, B3 y B4 para obtener las diferencias de proporciones y los intervalos de confianza de 95% que correspondan a los datos.

Las hojas de cálculo también son de gran ayuda para realizar los gráficos necesarios en el análisis estadístico. La facilidad con la cual se puede realizar un gráfico en estos programas permite ensayar con varias formas diferentes hasta que se encuentre la que mejor presenta los resultados.

Además de todas las facilidades que brindan las hojas de cálculo, también se debe mencionar la gran capacidad que tienen para importar y exportar archivos generados en otros formatos, lo cual los hace muy útiles cuando se trabaja en varias plataformas de datos.

Parecería que un programa de hoja de cálculo cubriría todas las necesidades para el trabajo. Sin embargo, la hoja de cálculo también tiene algunas limitaciones inherentes a su estructura en forma de hoja: es muy difícil tabular datos cuando existe interés en más de dos variables, y esto limita el uso de las hojas de cálculo sólo a la realización de análisis muy elementales.

## **PROGRAMAS PARA PRESENTACIONES**

Estos programas ayudan a elaborar la presentación de los datos, principalmente a través de proyecciones o carteles. Para hacerlo, generalmente se le concede preferencia al uso de gráficos o cuadros. Entre estos programas se encuentran Power Point y Harvard Graphics.

## **PROGRAMAS ESTADÍSTICOS**

De todos los programas que se comentan en este capítulo, éstos son los que más han revolucionado el trabajo en la estadística. Entre ellos existe una gran variedad de funciones y costos. Los hay genéricos o especializados y también gratuitos o muy caros. Es difícil decir cuál es el mejor, pero no cabe duda de que se debe tener el que mejor se conozca y que permita realizar los análisis estadísticos que se requieren para el trabajo. La presentación que aquí se hace se concentra en tres programas: Epi Info, SPSS y R.

### **Epi Info**

Este programa ha sido desarrollado y distribuido por el CDC de Atlanta. En un principio fue concebido como una herramienta auxiliar de los epidemiólogos de campo para ser utilizado en equipos portátiles de cómputo, pero con el tiempo ha encontrado su lugar en la mayoría de los grupos de trabajo que laboran en el campo de la salud, entre los que sin duda es uno de los programas más populares. Varias de sus características han contribuido a brindarle ese lugar privilegiado, entre las que destaca su facilidad de manejo; pero ninguna le ha dado tanto impulso como el hecho de que el programa se distribuye libremente y sin costo a través de Internet desde el CDC de Atlanta ([www.cdc.gov](http://www.cdc.gov)) y otros sitios de la red. Esta gran difusión ha facilitado la traducción del programa y sus manuales al español, entre otros idiomas. Epi Info ha sido desarrollado para ejecutarse en dos plataformas diferentes: MSDOS y Windows. Originalmente fueron escritas en inglés, pero existen opciones en español. La versión para MSDOS, que se ejecutaba bastante bien en Windows XP o anteriores, es Epi Info 6 y aún se puede descargar en <http://huespedes.cica.es/huespedes/epiinfo/>. La primera versión para Windows se conoció como Epi Info 2000 y se ejecuta bastante bien en Windows XP o posteriores. La versión más reciente del programa es Epi Info 7 que se ejecuta, de preferencia, en Windows 7. Las tres versiones en inglés se pueden descargar de <http://www.cdc.gov/epiinfo/>.

**Epi Info 7.** En el Anexo K se presentan las instrucciones mínimas necesarias para su utilización. El programa corre en ambiente Windows, y ésta es una gran ventaja porque se maneja de manera semejante a otros programas de este ambiente gráfico;

así, si ya se conoce uno, todos los demás resultan familiares y más fáciles de aprender. De manera general, la interface del Menú nos muestra las utilerías que ofrece el programa: Crear formas de captura (Create Forms), capturar datos (Enter Data), analizar datos (Analyze Data) y crear mapas (Create Maps). A estas mismas opciones se puede llegar a través de "Tools" en el menú que se encuentra en el borde superior de la ventana. En el mismo menú superior se encuentra "StatCalc", esta opción incluye una serie de calculadoras epidemiológicas y el enlace a la página WEB de OpenEpi.com que se describe líneas abajo.

El análisis de los datos (en Analyze Data) se puede realizar tanto en archivos propios de Epi Info 7, como archivos con formatos de Access, Excel, SQL y ASCII.

### **SPSS (Statistical Package for Social Science)**

Este programa tiene una larga historia en el análisis estadístico. Las primeras versiones se hicieron para correr en equipos muy grandes, y fue uno de los primeros programas de estadística disponibles en las computadoras personales. En la actualidad, el programa corre en varias plataformas, entre las que se encuentra Windows. En las primeras versiones de SPSS para PC el usuario tenía que saber mucho de programación porque cada comando se escribía en una pantalla negra en la que no se veían ayudas. Ahora el ambiente gráfico facilita el manejo a tal punto que el usuario puede aprender a utilizarlo en horas (o en minutos si se tiene alguna experiencia en otros programas de cómputo). Al entrar al programa se ve una pantalla cuadriculada muy semejante a una hoja de cálculo. En esta pantalla se puede empezar a capturar datos de la misma manera como se señaló para las hojas de cálculo, con las mismas dificultades, pero sin la facilidad de poder realizar operaciones en las celdas. Los datos capturados de esta forma pueden guardarse en un archivo de SPSS y después pueden llamarse para continuar la captura o iniciar el análisis. El programa también permite leer bases de datos generadas por otros programas, como Excel o Fox, por ejemplo.

### **R**

El programa R es un ambiente de programación para realizar gráficos y cálculos estadísticos. Su gran ventaja es que es un programa de acceso abierto y gratuito en constante actualización, el cual puede descargarse y encontrar temas de ayuda en la siguiente dirección, <http://www.r-project.org/index.html>. Es un proyecto de colaboración en el cual los colaboradores donan códigos de acceso libre, actualizan el programa y sus paquetes, corrigen errores de programación y documentan las distintas funciones de R. Entre la variedad de cálculos estadísticos que pueden realizarse en R se encuentran los análisis estadísticos clásicos, modelación lineal y no lineal, análisis de series de tiempo, análisis de clasificación y estadística multivariada, por mencionar algunos. Las personas que conocen el lenguaje de programación pueden generar sus propios códigos para realizar cálculos específicos (e.g. simulaciones Monte Carlo). Otra de sus ventajas es la calidad del diseño de sus gráficos para la publicación de los mismos. En R pueden importarse bases de datos guardadas en formato de texto (\*.txt) ó formato CVS (\*.csv) con columnas delimitadas por comas o tabulaciones; también es posible importar y utilizar los archivos creados en Excel (\*.xls).

La gran desventaja de este programa radica en que la consola de R trabaja con un lenguaje de programación y no cuenta con los menús de selección a los que estamos habituados los usuarios de la plataforma de Windows; es necesario conocer su

lenguaje, basado en códigos de programación específicos, para poder comenzar a utilizar la consola blanca. Otro de los inconvenientes es que los códigos son extremadamente sensibles a los errores tipográficos, por lo que una simple coma (,) fuera de lugar interrumpe el proceso del análisis generando mensajes de error.

Actualmente existen programas gratuitos que funcionan como editores de código R y que trabajan bajo la plataforma de Windows. Estos programas facilitan la escritura de los comandos en un documento llamado “script”, el cual puede ser archivado como texto (\*.txt) o código R (\*.r) para un subsecuente uso y edición. Estos programas se vinculan con la consola de R con lo cual al tiempo que se escribe el código se pueden enviar las instrucciones a R y observar los resultados del mismo. Los autores de este libro recomiendan como editores de código los programas RStudio (<http://www.rstudio.com/>) y Tinn-R (<http://sourceforge.net/projects/tinn-r/>), los cuales tienen ventajas adicionales como la ventana para observar y guardar los gráficos y las ventanas de fácil acceso a la instalación de paquetes y a la búsqueda de temas de ayuda en R.

**R Commander.** La consola de R puede requerir paquetes de comandos para realizar análisis específicos, los cuales deben ser instalados en R previo a su uso. El paquete necesario para realizar análisis estadísticos clásicos se instala automáticamente al instalar el programa R (“stats”). El paquete ODBC Database Access (“RODBC”) debe ser instalado por el usuario si éste desea importar a R las bases de datos creadas en formato Excel (\*.xls).

R Commander (“Rcmdr”) es un paquete que funciona como un programa de análisis estadístico dentro del programa R. La ventaja de trabajar con R Commander radica en que éste aporta todas las ventajas de R (*e.g.* gráficos) por medio de una interface mucho más amigable para el usuario, basada en menús y ventanas de selección bajo el perfil de Windows. En este libro, se guiará al lector a través de los pasos básicos necesarios para el análisis y creación de los modelos lineales generalizados en R Commander (capítulo 27). Los autores recomiendan el uso de R Commander para aquellos lectores que comienzan a estudiar e implementar este tipo de análisis estadístico con sus datos, sin embargo, los modelos generalizados pueden generarse directamente en la consola de R sin requerir ningún paquete adicional, tal como se mencionará en el capítulo 27. Es importante señalar que la aplicabilidad de R Commander no se limita a los modelos generalizados. Los lectores pueden iniciarse en el lenguaje de R utilizando dicho paquete como otra herramienta estadística (*e.g.* análisis de varianza, análisis de regresión) precisa, accesible y de bajo costo.

#### **Ejemplo explicativo 4–4.** -----

Una vez descargado de la red e instalado el programa R, es necesario instalar el paquete R Commander. Por simplicidad, lo anterior lo haremos por medio del programa RStudio siguiendo los siguientes pasos:

- 1.- Instalar el programa R (en su versión más reciente).
- 2.- Instalar el programa RStudio.
- 3.- Al abrir RStudio, éste solicitará que se seleccione el “CRAN mirror”, servidor a partir del cual se instalarán los componentes necesarios para la sesión de trabajo. Seleccionar el que corresponda al país en el que se encuentre. Es probable que la primera vez que se utilice el programa, éste solicite instalar ciertos paquetes automáticamente.



- 4.- Una vez abierta la consola buscar en la ventana derecha la pestaña “Packages” la cual contiene la lista de los paquetes que utiliza R (figura 4-1). Una vez abierta esa ventana buscar en la lista a R Commander. Al palomear la casilla “Rcmdr” el programa comenzara a instalarlo automáticamente, esto ocurrirá solamente la primera vez que se utiliza el programa.

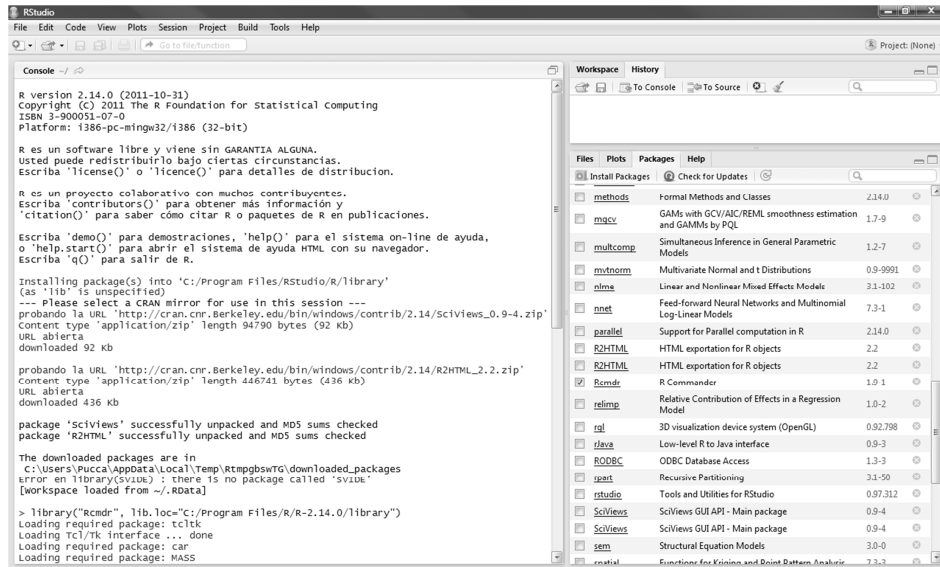


Figura 4-1. Programa RStudio. A la izquierda se muestra la consola de R versión 2.14.0 (2011). A la derecha se muestra la ventana “Packages” con la lista de los paquetes y la opción “Rcmdr” seleccionada.

- 5.- En las siguientes ocasiones en que se utilice RStudio, solo será necesario seleccionar el paquete “Rcmdr” de la lista de paquetes y el programa se encargara de cargarlo en la consola de R y abrirlo (figura 4-1).

## OpenEpi

OpenEpi es una página en la WEB que se puede consultar a través de un browser (en [http://openepi.com/v37/Menu/OE\\_Menu.htm](http://openepi.com/v37/Menu/OE_Menu.htm)), pero que también se puede utilizar a nivel local si previamente se descargó el programa en la computadora. Está conformado por una serie de calculadoras epidemiológicas y con enlaces a muchas páginas especializadas en análisis estadístico y epidemiológico.

## ¿EL MEJOR PROGRAMA?

No existe. Cada programa puede tener sus ventajas y sus aplicaciones. Por ejemplo, SPSS tiene una gran cantidad de opciones para el análisis estadístico, pero pocos lo han utilizado completamente. Epi Info brinda menos posibilidades, pero incluye opciones que no están presentes en SPSS, al igual que el programa R el cual, a través de su paquete R Commander, complementa los estadísticos de SPSS y facilita el análisis de los modelos generalizados. En opinión de los autores, los administradores de bases de datos, como dBase o Fox, son mejores para transformar variables, pero no mejores que Epi Info para capturar o verificar datos. Las hojas de cálculo, como Excel, son insustituibles cuando se trata de hacer operaciones; los programas de presentaciones

no tienen comparación cuando se desea mostrar los datos en un evento científico; y los procesadores de texto son una verdadera bendición cuando se describe el proyecto, se realiza el informe final o se escribe un artículo en el que se dan a conocer los resultados de la investigación. Así que no existe un “mejor programa”. En general, la mayoría son buenos, en particular los más reconocidos. Pero es preciso tener cuidado, porque sí podría existir un mal programa y su uso tendría consecuencias muy desagradables para el trabajo y la interpretación de los resultados.

## **REFERENCIAS**

**R Development Core Team** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013. ISBN 3-900051-07-0, URL: <http://www.R-project.org/>.