

# Apuntes de Estudio

## *Procesamiento Estadístico de Datos con Minitab y Harvard Graphics*

Julio Angeles Olivera  
Jorge Rubio Donet  
Yván Soto Rodríguez  
Jorge Toma Inafuko



UNIVERSIDAD DEL PACIFICO

***Procesamiento  
Estadístico  
de Datos con  
Minitab y  
Harvard Graphics***

**Julio Angeles Olivera  
Jorge Rubio Donet  
Yván Soto Rodríguez  
Jorge Toma Inafuko**



**UNIVERSIDAD DEL PACIFICO**

**LIMA-PERÚ  
1995**

© Universidad del Pacífico  
Centro de Investigación  
Avenida Salaverry 2020  
Lima 11, Perú

**UNIVERSIDAD DEL PACIFICO**  
**BUP - CENDI**

**PROCESAMIENTO ESTADÍSTICO DE DATOS  
CON MINITAB Y HARVARD GRAPHICS**

**40672**

**Julio Angeles Olivera**

**Jorge Rubio Donet**

**Yván Soto Rodríguez**

**Jorge Toma Inafuko**

1a. Edición: marzo 1995

Diseño de la carátula: M & B CREATIVOS.

**BUP - CENDI**

**Procesamiento estadístico de datos con Minitab y Harvard  
Graphics / Julio Angeles Olivera [et al.] -- Lima : Universidad  
del Pacífico, 1995. -- (Apuntes de Estudio) ; 19.**

**/ESTADÍSTICA/PROCESO ELECTRÓNICO DE DATOS/ PRO-  
GRAMAS DE COMPUTADORAS/**

**681.3:31 (CDU)**

Miembro de la Asociación Peruana de Editoriales Universitarias y de Escuelas Superiores (APESU) y miembro de la Asociación de Editoriales Universitarias de América Latina y el Caribe (EULAC).

El Centro de Investigación de la Universidad del Pacífico no se solidariza necesariamente con el contenido de los trabajos que publica.

Derechos reservados conforme a Ley.

# Índice

<b>Prólogo</b> .....	<b>11</b>
<b>I. Ingreso a Minitab y normas básicas</b> .....	<b>13</b>
Introducción .....	13
1. La hoja de cálculo .....	13
2. Constantes .....	15
3. Ingreso a Minitab y manejo básico .....	15
4. Comandos .....	16
4.1 Subcomandos .....	17
4.2 Argumentos .....	17
4.3 Prompts .....	17
4.4 Formas abreviadas .....	18
4.5 Archivos .....	19
4.6 Símbolos de Minitab .....	20
4.7 Help o ayuda .....	21
5. Limitaciones de Minitab .....	21
<b>II. Comandos para el manejo de datos</b> .....	<b>22</b>
1. Ingreso de datos .....	22
1.1 Ingreso de datos desde un archivo ASCII .....	23
1.2 Asignación de nombre a una columna .....	24
1.3 Obtención de información de la hoja de cálculo. ....	25

1.4	Ingreso de datos a la hoja de cálculo desde un archivo Minitab .....	25
1.5	Ingreso de datos desde el teclado .....	26
2.	Guardando la información .....	29
	Ejercicios .....	31
<b>III.</b>	<b>Manejo, edición y transformación de datos .....</b>	<b>38</b>
1.	Corrección de datos .....	38
2.	Creación de nuevas variables o datos .....	39
3.	Uso de funciones para crear nuevas variables .....	40
4.	Eliminación de datos .....	42
5.	Copia de datos seleccionados .....	44
6.	Asignación de códigos .....	46
7.	Concatenación de columnas .....	46
8.	División de columnas .....	47
	Ejercicios .....	48
<b>IV.</b>	<b>Aplicaciones: resúmenes de datos .....</b>	<b>50</b>
1.	Resumen de datos de columnas .....	50
2.	Resumen de datos de filas .....	52
3.	Otros comandos para resumir datos numéricos .....	54
	Ejercicios .....	58
<b>V.</b>	<b>Aplicaciones: representaciones gráficas de datos .....</b>	<b>60</b>
1.	Histogramas horizontales .....	61
2.	Histogramas verticales .....	64
3.	Gráficos de cajas .....	66
4.	Gráficos en dos dimensiones .....	66
	Ejercicios .....	69
<b>VI.</b>	<b>Aplicaciones: distribuciones de probabilidades y números aleatorios .....</b>	<b>71</b>
1.	Cálculo de una probabilidad acumulada .....	71
2.	Cálculo de percentiles .....	73
3.	Generación de números aleatorios .....	74
4.	Uso de una porción de un conjunto de datos .....	78
	Ejercicios .....	80

<b>VII. Aplicaciones: inferencia estadística</b> .....	<b>81</b>
1. Inferencia estadística para la media poblacional .....	81
1.1 Inferencia cuando se conoce la variancia poblacio- nal .....	81
1.2 Inferencia cuando no se conoce la variancia poblacio- nal .....	83
2. Inferencia estadística para la diferencia de medias pobla- cionales .....	85
2.1 Inferencia cuando las variancias poblacionales son desconocidas y las muestras son independientes ....	85
2.2 Inferencia con dos muestras dependientes .....	92
3. Análisis de datos categóricos .....	94
3.1 Pruebas sobre frecuencias de k categorías .....	94
3.2 Pruebas de bondad de ajuste .....	96
3.3 Pruebas de independencia y de homogeneidad ...	102
Ejercicios .....	106
 <b>VIII. Aplicaciones: análisis de correlación y regresión</b> .....	 <b>108</b>
1. Correlación .....	108
2. Regresión lineal simple y múltiple .....	109
2.1 Ejemplo sobre un análisis de regresión lineal simple .....	110
2.2 Ejemplo sobre un análisis de regresión a través del origen .....	112
2.3 Ejemplo sobre un análisis de regresión lineal múltiple .....	113
Ejercicios .....	123
 <b>IX. Breve descripción del menú de Minitab</b> .....	 <b>124</b>
1. El menú principal .....	124
2. La opción File .....	125
3. La opción Edit .....	128
4. La opción Calc .....	130
5. La opción Stat .....	133
6. La opción Graph .....	135
7. Ejemplos, utilizando el menú de Minitab .....	137
7.1 Ingreso a la hoja de cálculo .....	138
7.2 Asignar nombre a las columnas .....	138
7.3 Obtener resumen de la información que contiene la hoja de cálculo .....	138

7.4	Construcción de histograma	140
7.5	Cálculo de medidas estadísticas	141
7.6	Intervalo de confianza	142
7.7	Prueba de hipótesis	143
7.8	Regresión lineal	144
<b>X.</b>	<b>Uso del Harvard Graphics</b>	<b>147</b>
	Introducción	147
1.	Ingreso al Harvard Graphics	148
2.	Creación de un gráfico de sectores (Pie Chart)	149
3.	Creación de un gráfico de barras (Bar Chart)	153
4.	Creación de histogramas y polígonos de frecuencias	157
5.	Creación de un gráfico de líneas	159
6.	Recuperación de un gráfico	161
7.	Impresión de gráficos	161
<b>Anexo:</b>	<b>Equivalencia entre los comandos y las opciones del menú de Minitab</b>	<b>165</b>

# Prólogo

Cuando se realiza un trabajo estadístico, especialmente el que se realiza a través de encuestas, se recoge usualmente una gran cantidad de datos. Estos últimos deben ser procesados; esto quiere decir, que a los datos deben aplicarse procedimientos estadísticos para ser organizados, graficados, utilizados en cálculos de medidas estadísticas, etc. Con la finalidad de hacer más sencilla y rápida la aplicación de tales procedimientos, se han desarrollado paquetes o programas estadísticos de cómputo.

En esta publicación se explicará el manejo de dos paquetes de cómputo: el paquete estadístico Minitab versión 8 y el paquete Harvard Graphics.

El paquete Minitab versión 8 tiene aplicaciones para estadística básica, gráficos, distribuciones probabilísticas, inferencia estadística paramétrica, análisis de variables categóricas, regresión y correlación, entre otros. Además, con este paquete se puede trabajar de dos formas: la primera, utilizando directamente los llamados comandos o instrucciones para realizar determinado procesamiento; y la segunda, mediante un menú de opciones similar al que tienen otros paquetes como Lotus 123, Excel, etc. Aquí se dará mayor énfasis al uso de comandos, ya que permite un trabajo más interactivo, además de ser más directo y rápido.

Con Harvard Graphics se explicará la construcción de las diferentes representaciones gráficas, que son necesarias para una adecuada descripción de los datos estadísticos. Harvard Graphics permite construir gráficos de mejor calidad que los construidos con Minitab, por esta razón se explica su manejo en este manual.



Desde el capítulo I hasta el IX se explicará el manejo del paquete Minitab. Entre el I y el VIII, la explicación estará referida al uso de comandos; mientras que el capítulo IX estará dedicado al uso de las diferentes opciones del menú del Minitab. En el capítulo X se explicará el manejo del paquete Harvard Graphics. Además, esta publicación incluye un conjunto de ejercicios, que permitirá al alumno afianzarse en los conocimientos aprendidos.

# I

## Ingreso a Minitab y normas básicas

### Introducción

Minitab es un paquete estadístico que tiene alrededor de 200 comandos que permiten examinar y analizar datos. Mediante estos comandos o instrucciones sencillas, es posible realizar análisis estadísticos, generar histogramas, construir gráficos y cuadros sencillos, y hacer transformaciones matemáticas con los datos a procesar.

Los comandos Minitab son accesibles a través de menús y también pueden trabajarse en forma interactiva, mediante el uso de comandos. Se puede trabajar con cualquiera de las dos formas, de manera independiente o intercalando las dos.

Para el procesamiento se han establecido algunas normas que, a continuación, se describen. Estas normas, que se explicarán y utilizarán desde este capítulo hasta el capítulo VIII, son las que corresponden al trabajo con el lenguaje interactivo. Los procedimientos correspondientes a la forma de trabajo a través de menús, se presentarán en el capítulo IX.

### 1. La hoja de cálculo

Minitab trabaja con una hoja de cálculo. Ésta es un arreglo de filas (observaciones) y columnas (variables), tal como se muestra en la figura 1.1. Las columnas son denotadas como C1, C2, C3, etc., o por un nombre; por ejemplo, 'salario'. Los nombres deben estar encerrados entre apóstrofes.

La hoja de cálculo puede contener hasta tres mil quinientos datos en un máximo de 100 columnas; es decir, el producto del número de filas por el número de columnas (menor o igual a 100) no debe ser mayor de tres mil quinientos.

La hoja de cálculo es una área temporal de trabajo, la cual se borra automáticamente cuando se finaliza la sesión de Minitab. Para no perder la información, en dicha hoja, se puede crear un archivo en el disco, mediante dos formas: como un archivo de hoja de cálculo (para uso único de Minitab) o como un archivo de datos en código ASCII (para ser usado por otros programas).

Para ingresar a la hoja de cálculo se debe presionar, simultáneamente, las teclas [Alt] y [E] y luego, seleccionar en el menú la opción DATA SCREEN o presionar, simultáneamente, las teclas [Alt] y [D].

La hoja de cálculo acepta caracteres alfabéticos o numéricos. En el capítulo siguiente se explica el ingreso de datos a la hoja de cálculo de Minitab.

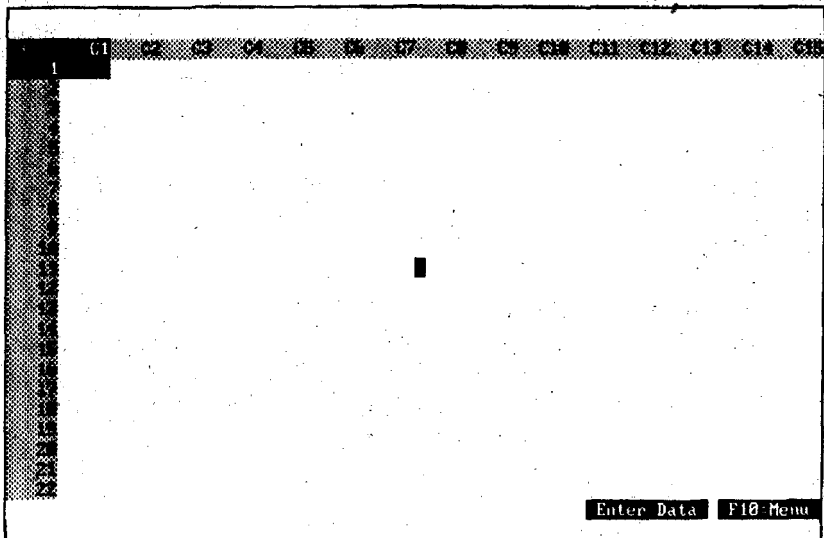


Figura 1.1 Hoja de Cálculo de Minitab

## 2. Constantes

En el paquete Minitab, las constantes se denotan como K1, K2, K3, etc. Una constante es un área de memoria, en el cual se guarda el valor de un dato importante o el resultado de un cálculo. Este almacenamiento se realiza mediante el comando LET. Por ejemplo, si se realiza el cálculo de la media de los datos de la columna 10 (C10), el resultado se puede almacenar (mediante el comando LET) en la constante K1, para su uso posterior. Minitab tiene 100 constantes, pero asigna automáticamente valores a las tres últimas constantes.

K98	=	*	(el símbolo de valor o dato perdido),
K99	=	2.718...	(el valor del número neperiano e)
K100	=	3.141...	(el valor de $\pi$ ).

## 3. Ingreso a Minitab y manejo básico

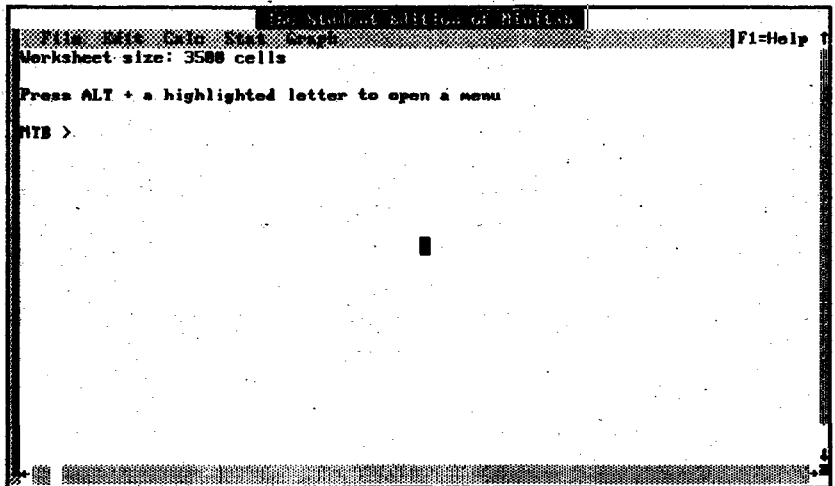
Para utilizar Minitab, en un disco duro, es necesario crear un directorio que contenga los archivos de programas. Este directorio, usualmente, tiene como nombre MINITAB. Si desea iniciar una sesión desde cualquier otro directorio, es necesario adicionar en el comando PATH del archivo AUTOEXEC.BAT la ruta C:\MINITAB, con lo cual la línea de comando de rutas podría verse como sigue:

```
PATH C:\C:\DOS;C:\WP51;C:\MINITAB
```

Este comando DOS, del AUTOEXEC.BAT, supone que los programas se han cargado en el directorio MINITAB del disco duro C. Luego de efectuar esta modificación, arranque nuevamente su computador a fin de hacer activos los cambios.

Para iniciar la sesión de Minitab ingrese al directorio MINITAB y digite el comando MINITAB, luego presione [Enter]. En este apunte, los comandos se han expresado con letras mayúsculas, a fin de poder diferenciarlos fácilmente. Los resultados son los mismos si se usan letras minúsculas.

Al ejecutar el comando anterior, aparece en pantalla el Menú Principal del Minitab con las opciones File (F), Edit (E), Calc (C), Stat (S) y Graph (G). También aparece en pantalla el logotipo de Minitab (ver figura 1.2). Se puede trabajar con el menú o, directamente, en forma interactiva con el lenguaje de comandos.



**Figura 1.2 Pantalla inicial de Minitab**

El logotipo Minitab que aparece en pantalla (MTB >), al inicio de cada línea, es llamado PROMPT y sirve para visualizar el estado de ejecución de las instrucciones o comandos.

#### **Final de la sesión Minitab**

Para finalizar la sesión, a la derecha del prompt del Minitab (MTB >) se debe ejecutar el comando STOP (digite STOP y presione [Enter]).

#### **4. Comandos**

Una sesión de Minitab se basa en la ejecución de instrucciones o comandos. Éstos indican, a Minitab, qué acción ejecutar. Cada comando debe empezar en una nueva línea y puede ser digitada en mayúsculas o minúsculas. Por ejemplo, el comando INFO:

```
MTB > INFO
```

lista información de las columnas y constantes existentes en la hoja de cálculo en uso.

#### 4.1 Subcomandos

La mayoría de los comandos tiene subcomandos opcionales, que proporcionan a Minitab instrucciones adicionales. Los subcomandos deben ser ingresados uno por línea y en cualquier orden. Para ingresar un subcomando, el comando o subcomando anterior debe haber terminado en punto y coma. El último subcomando debe terminar en punto. Por ejemplo:

```
MTB > copy c1 c3;  
SUBC > use 1 4.
```

copia el contenido de las filas 1 y 4 de la columna C1, en las filas 1 y 4 de la columna C3.

#### 4.2 Argumentos

Los argumentos indican a Minitab qué información o datos debe procesar. Pueden ser columnas, constantes y/o archivos. Una descripción de los argumentos apropiados, de cada comando o subcomando, se puede obtener mediante el comando `HELP`; por ejemplo.

```
MTB > HELP READ
```

muestra la sintaxis y los argumentos del comando `READ`.

#### 4.3 Prompts

El prompt es la secuencia de caracteres al inicio de cada línea y sirve para indicar qué información debe ingresar. Minitab utiliza los siguientes prompts:

**MTB >** Es el prompt de comandos.

**SUBC >** Es el prompt de los subcomandos. Aparece cuando el comando de la línea anterior termina en punto y coma.

**DATA >** Es el prompt de datos. Aparece después de los comandos `READ`, `SET` e `INSERT`.

**CONTINUE ?** Es el prompt de página. Aparece cuando un reporte es extenso. Presione [Enter] para ver la siguiente página.

Presione la tecla [N] (No) y luego [Enter] para terminar el listado del reporte.

**CONT >** Es el prompt de continuación. Cuando un comando, subcomando o línea de datos tiene una extensión mayor a una línea, cada línea debe terminar en &, excepto la última; a fin de continuar el tipeo de un comando, subcomando o línea de datos.

**STOR >** Aparece después que el comando STORE ha sido usado, e indica que Minitab está a la espera de un archivo de comandos o de un macro.

#### 4.4 Formas Abreviadas

1. Para ejecutar un comando o subcomando, sólo es necesario tipear sus cuatro primeras letras.
2. Puede abreviar una serie de columnas mediante el símbolo -. Por ejemplo:

```
MTB > PRINT C1-C5
```

en lugar de:

```
MTB > PRINT C1 C2 C3 C4 C5
```

3. Puede abreviar una serie de números con el símbolo :. Por ejemplo:

```
MTB > delete rows 1:5 c1
```

en lugar de:

```
MTB > delete rows 1 2 3 4 5 C1
```

4. Puede abreviar el ingreso de números con el comando SET. Por ejemplo:

```
MTB > set c1
DATA > 3(1:4)
```

en lugar de:

```
MTB > set c1
DATA > 1 2 3 4 1 2 3 4 1 2 3 4
```

#### 4.5 Archivos

Los nombres de los archivos tienen dos partes: el "Nombre de archivo", que puede tener un máximo de 8 caracteres, y la "extensión", que puede tener un máximo de tres caracteres. Estas dos partes deben estar separadas por un "punto". Para hacer referencia a un archivo, mediante un comando Minitab, debe ponerse entre apóstrofes simples su "nombre", incluyendo si es necesario la ruta para definir la unidad y el directorio donde está el archivo. No es necesario tipear la "extensión", a menos que ésta sea diferente a la definida por defecto. Por ejemplo, se tiene:

```
MTB > retrieve 'c:\datos\caso1'
```

Mediante el comando RETRIEVE se recupera (carga a la memoria RAM) la hoja de cálculo grabada con el nombre CASO1.MTW, la cual se encuentra en el directorio DATOS de la unidad C. No es necesario especificar la ruta (c:\datos) si el archivo se encuentra en el directorio activo.

Minitab diferencia los archivos por su extensión, la cual es asignada automáticamente. Los diferentes tipos de archivos que reconoce Minitab son:

Nombre de archivo	Usado por el comando	Descripción
Nombre.LIS	OUTFILE NOOUTFILE	Contiene reportes de Minitab. Pueden ser editados o enviados a una impresora.
Nombre.DAT	WRITE READ SET INSERT	Identifica archivos de datos en código ASCII. Pueden ser editados, impresos y leídos por otros programas.



Nombre.MTW	SAVE RETRIEVE	Identifica una hoja de cálculo archivada por Minitab. No puede ser editada ni enviada a una impresora.
Nombre.MTP	SAVE; PORTABLE. RETRIEVE; PORTABLE.	Identifica hojas de cálculo portables de Minitab. Sólo pueden ser usadas en otras computadoras como Main-frame, etc. No puede ser editada ni impresa.
Nombre.MTB	STORE EXECUTE	Define un archivo de comandos o un macro que contiene comandos Minitab. Puede ser editado y enviado a una impresora.

Si desea almacenar toda la información de una hoja de cálculo de Minitab (columnas, nombres de columnas y constantes), debe usar el comando **SAVE**. El archivo tendrá una extensión **MTW** y sólo podrá ser usado por Minitab. El comando **RETRIEVE** permite recuperar archivos creados con el comando **SAVE**.

Si desea almacenar la información de una hoja de cálculo en código **ASCII**, debe usar el comando **WRITE**. Los archivos generados por este comando sólo contienen datos (no se almacenan los nombres de columnas y constantes) y pueden ser editados y usados por otros programas. Para recuperar un archivo de este tipo (en código **ASCII**), debe usar los comandos **READ**, **SET** e **INSERT**.

#### 4.6 Símbolos de Minitab

- \* Es el símbolo del valor perdido. Puede ser usado en comandos y subcomandos, y como un valor de datos.
- & Es el símbolo de continuación. Es útil cuando el comando que se está tipeando debe continuar en la siguiente línea.
- # Es el símbolo de comentario. Minitab ignora todo el texto entre dicho símbolo y el final de la línea.
- ' ' Los apóstrofes simples sirven para encerrar los nombres de un archivo y los nombres de las columnas.

- ( ) Los paréntesis son usados para separar grupos de columnas.
- Se usa para abreviar una serie de columnas.
- : Se usa para abreviar una serie de números.
- ; El punto y coma debe figurar al final de una línea y sirve para indicar que luego viene un subcomando.
- . Un punto finaliza el último subcomando usado con un comando.

#### 4.7 Help o ayuda

El comando HELP explica la sintaxis y los argumentos necesarios de un comando o subcomando. Para usarlo debe tipear HELP y luego, el nombre del comando para el cual desea alguna ayuda.

### 5. Limitaciones de Minitab

1. Minitab permite trabajar con cualquier número en el rango de  $-1e^{+18}$  a  $+1e^{+18}$ . Cualquier valor fuera de este rango es convertido al símbolo de valor perdido (\*).
2. Ningún comando o subcomando puede contener más de 50 argumentos (número total de nombres de archivos, columnas y constantes). Por ejemplo, el comando:

```
MTB > READ 'C:\DATOS\ENCUESTA' C1-C7
```

tiene 8 argumentos (el nombre del archivo, más 7 columnas).

# II

## Comandos para el manejo de datos

Minitab puede trabajar con datos que están en la memoria interna (RAM) del computador. En esta sección se describe cómo cargar datos a la memoria (Ingreso de datos) y como se puede almacenar datos de la memoria en un disco (Salida de datos).

### 1. Ingreso de datos

Existen dos formas de ingresar datos a la memoria: mediante datos existentes en un archivo y por digitación directa de la información. Para apreciar ambas formas de ingreso de datos, suponga que al aplicar una encuesta a 5 clientes se capta la información siguiente:

1	28	26	3	3	1	1	2	0	350	2500	1	50
2	32	33	3	3	2	4	3	1	480	3600	1	30
3	44	40	3	3	1	8	4	2	550	5000	2	65
4	36	31	3	3	3	4	3	1	490	3000	1	40
5	42	40	3	3	4	1	4	2	600	4600	1	45

Estos datos están almacenados en el archivo texto (en código ASCII) con nombre ENCU.DAT y también en la hoja de cálculo de Minitab con nombre ENCU.MTW.

## 1.1 Ingreso de Datos desde un archivo ASCII

### Comando READ

Una vez que inicia una sesión de Minitab y tiene activado el prompt "MTB>", se puede ingresar un comando. Para ingresar la información contenida en el archivo ENCU.DAT, en código ASCII, se puede ejecutar el comando siguiente:

```
MTB > READ 'ENCU' C1-C13
```

con lo cual se leerá las 13 columnas de datos del archivo ENCU.DAT, obteniéndose el mensaje siguiente:

```
5 ROWS READ
```

ROW	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
1	1	28	26	3	3	1	1	2	0	350	2500	1	50
2	2	32	33	3	3	2	4	3	1	480	3600	1	30
3	3	44	40	3	3	1	8	4	2	550	5000	2	65
4	4	36	31	3	3	3	4	3	1	490	3000	1	40

### Comando FORMAT

Cuando se desea leer datos de acuerdo con un formato o estructura particular, se debe usar el subcomando FORMAT de la siguiente manera:

```
MTB > READ 'ENCU' C1-C13;
SUBC > FORMAT (A2,2X,F2.0,2X,F2.0,6(2X,F1.0),2X,F3.0,2X,
F4.0,2X,F1.0,2X,F2.0)
```

El subcomando FORMAT indica que los dos primeros caracteres a leer son alfanuméricos (letras y/o números); luego, se tiene 2 espacios en blanco; en seguida, un número con un máximo de 2 dígitos y sin decimales; a continuación, 2 espacios en blanco; luego, un número entero de 2 dígitos; en seguida, 6 números de un dígito precedidos por 2 espacios en blanco; y así, sucesivamente. El resultado obtenido, con este subcomando, es el mismo al descrito anteriormente.

Dentro del subcomando **FORMAT** se pueden usar las siguientes opciones para definir el proceso de lectura de datos:

- Fw.d** Para leer un número que ocupa "W" caracteres, de los cuales "d" son decimales.
- Aw** Para leer "w" caracteres alfanuméricos.
- nX** Para saltar "n" espacios entre datos.
- ,** Para separar campos de lectura.

## 1.2 Asignación de nombre a una columna

### Comando NAME

Hasta ahora, para hacer referencia a una columna o variable, la forma de hacerlo es como C1, C2, C3, etc. Sin embargo, es posible asociar un **NOMBRE** a cada columna, a fin de identificar mejor la variable asociada a ella. El nombre de la columna o variable puede constar hasta de 8 caracteres como máximo. Para asignar nombres a las columnas, se usa el comando **NAME**; por ejemplo:

```
MTB > name c1 'codigo' c2 'edadjf' c3 'edadco' c4 'instjf' &
CONT > c5 'instco' c6 'ocupjf' c7 'ocupco' c8 'numper' &
CONT > c9 'nhijoesc' c10 'gastosem' c11 'ingreso'
MTB > name c12 'numvehi' c13 'gastocom'
```

Como se deben asignar los nombres a las 13 columnas, es necesario tipear más de una línea, por lo cual se puede usar el caracter "&", como se puede apreciar al final de las dos primeras líneas. También es posible hacer la asignación de nombres mediante varios comandos **NAME**, como se aprecia en la última línea donde se asigna nombres a las columnas 12 y 13.

Minitab no distingue letras mayúsculas de minúsculas, por ejemplo: 'CODIGO', 'codigo' y 'Codigo' son nombres equivalentes. Por lo tanto, se debe asegurar la asignación de nombres diferentes a las columnas y que ninguno de ellos sea igual al nombre del archivo.

Una vez asignados los nombres, se puede hacer referencia a las columnas por su número (C1, C2, etc.) o por su nombre.

### 1.3 Obtención de información de la hoja de cálculo

#### Comando INFO

Para obtener un listado descriptivo de la información existente en una hoja de cálculo de Minitab, se puede usar el comando INFO, por ejemplo:

```
MTB > info
```

El computador proporciona lo siguiente:

COLUMN	NAME	COUNT
		15
C1	codigo	5
C2	edadjf	5
C3	edadco	5
C4	instjf	5
C5	instco	5
C6	ocupjf	5
C7	ocupco	5
C8	numper	5
C9	nhjoesc	5
C10	gastosem	5
C11	ingreso	5
C12	numvehi	5
C13	gastocom	5

CONSTANTS USED: NONE

### 1.4 Ingreso de datos a la hoja de cálculo desde un archivo Minitab

#### Comando RETRIEVE

Para ingresar la información contenida en el archivo ENCU.MTW, desde una hoja de cálculo de Minitab, se puede usar el comando RETRIEVE de la siguiente manera:

```
MTB > retrieve 'encu'
```

El computador responde:

```
WORKSHEET SAVED 3/11/1994
```

```
Worksheet retrieved from file: encu.MTW
```

Como puede apreciarse, mediante este comando no es necesario indicar las columnas a leer ni el formato de lectura. Basta indicar tan sólo el nombre del archivo Minitab.

Si ejecuta el comando INFO, obtendrá un listado descriptivo de la información existente en esta hoja de cálculo, tal como se obtuvo anteriormente.

### **Nota importante**

Es importante verificar que no se tiene información en memoria, antes de usar el comando RETRIEVE o READ (para lectura de un archivo en código ASCII). Puede hacer esto mediante el comando INFO. Si tuviese información debe eliminarla, con el comando ERASE, antes de leer un archivo Minitab. Por ejemplo, si aún existen en memoria las columnas C1 a C5 y las constantes K1 a K3, se puede eliminar esta información mediante:

```
MTB > erase c1-c5
```

```
MTB > erase k1-k3
```

```
MTB > info
```

```
* ALL COLUMNS EMPTY AND UNNAMED
```

```
CONSTANTS USED: NONE
```

Se debe tener presente, sobre todo cuando se lee un archivo texto, que el comando READ solamente carga la información del archivo a las columnas especificadas en el comando, sin eliminar la información existente en otras columnas y constantes.

## **1.5 Ingreso de datos desde el teclado**

### **Comando READ**

Una forma de ingresar datos, desde el teclado, es utilizando el comando READ, indicando solamente las columnas que se desean crear. Por ejemplo,

la información anteriormente presentada puede ingresarse a la hoja de cálculo de la siguiente manera:

```
MTB > READ c1-c13
DATA > 1 28 26 3 3 1 1 2 0 350 2500 1 50
DATA > 2,32,33,3,3,2,4,3,1,480,3600,1,30
DATA > 3 44 40 3 3 1 8 4 &
CONT > 2 550 5000 2 65
DATA > 4 36 31 3 3 3 4 3 1 490 3000 1 40
DATA > 5 42 40 3 3 4 1 4 2 600 4600 1 45
DATA > END
```

Como puede apreciarse, los datos pueden separarse mediante coma o con espacios.

El carácter "&" se utiliza para indicar que no se ha concluido el ingreso de datos en una línea y que ésta continúa, en la línea siguiente (como se aprecia en la cuarta línea). Para indicar a Minitab que se ha concluido el ingreso de datos, se debe finalizar con END. Al ejecutar la última línea (presionando la tecla [Enter]), Minitab retorna el prompt "MTB >" para continuar con otros comandos.

### Comando SET

Cuando se desea ingresar datos con el teclado, columna por columna, se debe utilizar el comando SET. Por ejemplo, para ingresar los datos de la primera y segunda columna o los datos de las dos primeras variables, se puede usar:

```
MTB > SET into c1
DATA > 1,2,3,4,5
DATA > END
MTB > SET c2
DATA > 28 32 44 36 42
DATA > END
```

Como se aprecia, el comando se puede digitar como "SET into" o como "SET". La palabra "into" es descriptiva y puede omitirse.

Existen algunos casos donde los números a ingresar muestran cierta secuencia. En éstos, se puede abreviar el ingreso usando paréntesis o el carácter ":". Por



ejemplo, para la primera columna, los valores son números consecutivos del 1 al 5 y pueden ser ingresados como:

```
MTB > SET into c1
DATA > 1:5
DATA > END
```

Por otro lado,

```
MTB > SET into c2
DATA > 3(2 3 4)
DATA > END
```

pone en la columna C2, la secuencia de números: 2 3 4 2 3 4 2 3 4; es decir, 3 veces la secuencia 2 3 4.

Del mismo modo, el comando:

```
MTB > SET into c3
DATA > 3(2 3 4)2
DATA > END
```

pone en la columna C3, la secuencia de números: 2 2 3 3 4 4 2 2 3 3 4 4 2 2 3 3 4 4 ; es decir, cada número dentro del paréntesis se repite 2 veces y la secuencia resultante ( 2 2 3 3 4 4 ) se repite, a su vez, 3 veces.

Si, por alguna razón, en una posición no se tiene dato (valor perdido), basta poner en esa posición un asterisco. Por ejemplo,

```
MTB > SET c4
DATA > 12 14 11 * 15
DATA > END
```

indica a Minitab que el cuarto valor, de la columna C4, es un valor perdido.

### Comando INSERT

Este comando permite insertar "filas" de datos para una o más columnas, sin tener que volver a digitar toda la información. Por ejemplo,

```
MTB > INSERT 3 4 C5-C7
```

```
DATA > 12 24 31
DATA > 13 22 36
DATA > 11 20 32
DATA > END
```

permite insertar, entre las filas 3 y 4, tres filas de datos en las columnas C5, C6 y C7. Si en el comando anterior no se indica las filas, entre las cuales se hace la inserción (3 y 4), los datos son añadidos al final de las columnas. Si desea añadir datos al inicio, debe indicar 0 y 1 como las filas entre las cuales debe añadirse la información.

## 2. Guardando la información

### Comando SAVE

Una vez que termina una sesión de Minitab, toda la información que existe en memoria se pierde. Si desea preservar su hoja de cálculo, como un archivo Minitab, debe hacer uso del comando SAVE. Para esto, debe indicar entre comillas simples el nombre del archivo (la extensión es asignada, por defecto, como MTW) y si es necesario debe, además, indicar la unidad y el directorio donde se van a almacenar los datos, nombres de columnas y constantes. Por ejemplo,

```
MTB > SAVE 'CURSO'
MTB > SAVE 'B:CURSO'
MTB > SAVE 'C:DATA\CURSO'
```

En el primer caso, la hoja de cálculo se almacena en el directorio y unidad de disco activas, con el nombre CURSO.MTW; en el segundo caso, se almacena en el diskette de la unidad B, con el nombre CURSO.MTW; y en el tercer caso, se almacena en el directorio DATA de la unidad de disco duro C, con el nombre CURSO.MTW.

El archivo almacenado no puede ser editado ni impreso directamente, siendo reconocido solamente por Minitab.

### Comando WRITE

Si desea almacenar sus datos en un archivo texto (en código ASCII), debe usar el comando WRITE. En este caso, solamente, se guardan los datos de las columnas (no se almacenan los nombres de las columnas ni las constantes).

En forma similar al comando anterior, sólo es necesario indicar entre comillas simples, el nombre del archivo (la extensión es asignada por defecto como DAT). Por ejemplo,

```
MTB > WRITE 'CURSO' C1-C4, C6
MTB > WRITE 'B:CURSO' C1-C4, C6
MTB > WRITE 'C:\DATA\CURSO' C1-C4, C6
```

Con estos comandos se almacena la información de las columnas C1, C2, C3, C4 y C6, en un archivo con nombre CURSO.DAT que contiene información en código ASCII, la cual puede ser editada e impresa, siendo además posible usar esta información con otros programas.

De manera similar al comando SAVE, el destino del archivo CURSO.DAT depende de la unidad y directorio indicados.

### Comando PRINT

Este comando se utiliza para listar, en pantalla, los datos que existen en una o más columnas. Por ejemplo,

```
MTB > retrieve 'encu'
WORKSHEET SAVED 3/11/1994
```

Worksheet retrieved from file: encu.MTW

```
MTB > print c2-c4, c6
```

ROW	edadjf	edadco	instjf	ocupjf
1	28	26	3	1
2	32	33	3	2
3	44	40	3	1
4	36	31	3	3
5	42	40	3	4

Como puede apreciarse, con el comando PRINT se ha listado en pantalla los datos existente en las columnas C2, C3, C4 y C6.

### Comando OUTFILE

Si se desea guardar en un archivo todos los comandos ejecutados y los resultados obtenidos con ellos, tal como aparecen en pantalla durante una sesión de Minitab, puede usarse el comando OUTFILE. Por ejemplo,

```
MTB > OUTFILE 'DIA15'
```

Todos los siguientes comandos que se ejecuten y sus resultados son almacenados en código ASCII en el archivo llamado DIA15.LIS, el cual puede ser editado e impreso. Para finalizar el envío de resultados a este archivo, basta ejecutar el comando NOOUTFILE, sin otro argumento adicional.

### Comando PAPER

*Este comando efectúa lo mismo que el comando anterior, con la diferencia que todo lo que aparece en la pantalla es, a la vez, impreso directamente. Para iniciar la impresión, de lo que obtiene en una sesión de Minitab, debe ejecutar:*

```
MTB > PAPER
```

y para terminar la impresión de los resultados de una sesión, basta ejecutar el comando NOPAPER.

### Ejercicios

1. Suponga que la Cooperativa de Crédito CREDITAR desea proporcionar tarjetas de crédito y otros beneficios a sus socios, para lo cual aplica a 50 clientes, seleccionados al azar, la siguiente encuesta:

#### ENCUESTA FAMILIAR A SOCIOS DE LA COOPERATIVA CREDITAR

1. Código de la encuesta: \_\_\_\_\_
2. Edad del jefe de familia: \_\_\_\_\_ años
- Edad de la cónyuge: \_\_\_\_\_ años

3. Indique el grado de instrucción de: (marcar con X)

3.1 Del jefe de familia

Primaria	1
Secundaria	2
Superior	3
Maestría	4
Doctorado	5

3.2 De la cónyuge

Primaria	1
Secundaria	2
Superior	3
Maestría	4
Doctorado	5

4. Indique la ocupación principal de: (marcar con X)

4.1 Del jefe de familia

Industrial	1
Comerciante	2
Funcionario de empresa	3
Médico	4
Abogado	5
Asesor	6
Profesor universitario	7
Otras	8

4.2 De la cónyuge

Su casa	1
Industrial	2
Comerciante	3
Funcionario de empresa	4
Médico	5
Abogado	6
Asesor	7
Profesor universitario	8
Otras	9

5. Indique el número de personas que conforman la familia: \_\_\_\_\_

6. Indique el número de hijos en edad escolar . . . . . \_\_\_\_\_

7. Indique su gasto familiar semanal en comida . . . : \_\_\_\_\_  
nuevos soles

8. Indique su ingreso familiar mensual promedio . . : \_\_\_\_\_  
nuevos soles

9. Indique el número de vehículos de la familia . . . : \_\_\_\_\_

10. Indique su gasto familiar semanal en combustible: \_\_\_\_\_  
nuevos soles

Supóngase que la información obtenida es la siguiente:

01	28	26	3	3	1	1	2	0	350	2500	1	50
02	32	33	3	3	2	4	3	1	480	3600	1	30
03	44	40	3	3	1	8	4	2	550	5000	2	65
04	36	31	3	3	3	4	3	1	490	3000	1	40
05	42	40	3	3	4	1	4	2	600	4600	1	45
06	40	40	3	2	2	4	3	1	420	3200	1	55
07	40	36	3	2	4	7	4	2	480	3800	1	35
08	48	49	4	2	7	4	5	2	630	4700	1	30
09	45	35	3	2	1	1	3	1	500	5200	1	65
10	48	42	4	2	3	3	3	1	430	3800	1	40
11	35	33	3	2	3	1	2	0	250	2800	1	35
12	48	46	2	2	2	3	5	2	600	3900	1	25
13	36	36	2	2	2	1	4	2	550	4200	1	60
14	44	39	4	3	3	3	3	1	470	4100	1	30
15	50	43	3	2	5	7	4	2	525	4600	1	45
16	46	45	4	3	7	3	4	2	485	4800	1	50
17	42	42	3	2	1	7	3	1	410	4300	1	50
18	48	44	3	2	3	1	5	2	470	4200	1	35
19	52	49	3	2	5	3	7	4	560	5100	1	70
20	32	24	2	2	2	6	2	0	280	3000	1	30
21	55	50	4	3	3	1	4	1	490	3900	1	35
22	39	39	4	4	6	3	3	1	485	4500	1	40
23	53	46	2	2	2	6	4	2	495	4800	1	45
24	50	48	5	3	6	3	4	2	510	5200	1	65
25	36	36	3	2	1	6	3	1	380	4600	1	38
26	49	49	3	3	3	1	4	2	460	4300	1	28
27	30	25	4	3	6	6	2	0	310	2800	1	25
28	47	45	3	3	2	9	4	2	445	5100	1	68
29	34	30	3	3	4	3	3	1	380	3200	1	30
30	46	46	3	3	3	1	4	2	465	4500	1	43
31	29	28	3	3	2	5	2	0	220	2900	1	27
32	39	38	3	3	1	1	4	2	720	5500	1	75
33	38	39	3	3	2	4	4	2	640	5200	1	68
34	42	38	3	3	3	3	5	3	800	6800	2	70
35	31	29	3	3	8	1	3	1	430	4500	1	55
36	44	43	3	3	2	4	4	2	680	5900	1	58
37	45	42	4	3	3	1	5	2	850	6500	2	90
38	28	29	3	2	1	5	3	1	460	5400	1	62
39	51	50	5	5	3	4	4	2	620	6500	2	95

40	31	29	4	4	2	1	3	1	410	4200	1	45
41	53	50	4	3	3	2	5	2	750	6600	1	95
42	48	47	3	2	1	3	4	2	620	4800	1	42
43	55	52	3	2	3	4	6	2	750	5500	1	80
44	46	44	3	3	2	1	4	2	610	4900	1	40
45	33	30	4	4	5	5	3	1	420	3600	1	32
46	42	38	3	2	4	1	4	2	540	3800	1	30
47	44	39	3	2	3	8	5	3	580	4900	1	40
48	36	32	3	2	4	2	3	1	390	3950	1	35
49	44	42	3	3	5	5	4	2	435	4200	1	38
50	46	45	3	3	4	5	5	3	625	4700	1	32

- a) Crear, con estos datos, un archivo ASCII con nombre PCA1.DAT.
  - b) Ingresar a Minitab y ejecutar el comando OUTFILE, para almacenar los resultados de su sesión en el archivo PRAC1.LIS.
  - c) Leer el archivo PCA1.DAT.
  - d) Asignar los nombres de cada columna, indicados en el ejemplo del comando NAME.
  - e) Verificar su hoja de cálculo con el comando INFO.
  - f) Guardar la información en el archivo EN9401.MTW del Minitab.
2. Ingresar, en la hoja de cálculo de Minitab, los datos que se encuentran en la sección 1 del presente capítulo. Luego, guardar la información con el nombre ENCU en un archivo Minitab y en un archivo texto (código ASCII).
  3. Ingresar, en la hoja de cálculo de Minitab, los datos del archivo ENCU.MTW. Luego, utilizar el comando PRINT.
  4. Asignar nombre a cada una de las columnas, del archivo ENCU, del ejercicio anterior. Utilizar los nombres dados, en el ejemplo desarrollado, con el comando NAME. Con el comando PRINT, pedir los datos de todas las columnas. ¿Qué diferencia encuentra con lo obtenido utilizando el comando PRINT del ejercicio 2?

5. Utilizar el comando INFO para el archivo ENCU. Luego, aplicar el comando STOP.
6. Recuperar, en la hoja de cálculo del Minitab, el archivo ENCU que fue almacenado en código ASCII en el ejercicio 2.
7. Activar el comando OUTFILE con el nombre PRAC11. Luego, realizar las mismas operaciones indicadas en los ejercicios 5 y 6, y ejecutar los comandos NOOUTFILE y STOP.
8. Ingresar al paquete Word Perfect versión 5.1. Recuperar el archivo PRAC11.LIS.
9. Supóngase que la encuesta mencionada, en el ejercicio 1, también se aplicó a 40 clientes de la ciudad de Piura y que los datos obtenidos fueron:

1	24	26	2	2	3	4	3	0	450	2400	1	35
2	37	32	5	5	4	7	2	0	360	2300	2	70
3	43	47	3	2	7	1	2	0	420	3700	1	50
4	39	33	4	5	3	6	3	1	580	3500	1	35
5	45	48	2	3	8	3	2	0	430	2700	1	45
6	48	49	3	3	2	2	4	2	590	3300	1	40
7	43	36	4	2	3	1	3	1	530	3600	1	50
8	48	43	2	2	7	1	5	2	650	5800	0	0
9	40	36	5	5	3	3	2	0	300	1800	1	25
10	42	47	4	2	7	5	3	1	430	2500	1	45
11	34	38	2	3	4	3	4	1	450	2200	1	35
12	56	49	3	2	8	7	3	1	580	3700	1	50
13	38	33	4	3	5	1	6	3	730	5200	1	80
14	42	37	2	3	7	1	4	2	480	3800	1	45
15	58	42	3	3	1	7	3	1	430	3500	2	50
16	43	47	4	2	2	3	6	3	580	3800	1	65
17	45	44	2	2	1	3	2	0	340	2600	1	45
18	48	40	5	2	4	1	3	1	490	3500	1	40
19	53	43	4	3	8	2	5	3	640	4700	1	70
20	36	26	2	3	3	1	4	2	430	2400	1	35
21	58	58	3	3	6	1	2	0	250	1300	0	0
22	33	33	2	2	2	1	3	1	360	2700	1	35
23	56	49	3	2	5	1	5	2	680	4500	1	80
24	59	43	3	3	8	1	2	0	430	2600	2	30



25	33	36	3	2	3	3	3	1	420	2700	1	35
26	46	47	4	3	5	7	4	2	560	3500	1	40
27	39	23	4	5	7	5	3	0	330	2400	1	25
28	40	44	3	2	5	4	5	2	670	4700	1	55
29	32	34	3	3	8	8	3	1	250	1800	1	20
30	45	44	2	2	2	1	4	2	440	3400	1	40
31	23	25	2	3	3	6	2	0	360	2500	1	35
32	38	33	2	2	6	1	4	2	580	3800	2	80
33	34	36	2	3	4	8	5	3	640	4500	1	40
34	46	33	3	2	7	3	3	1	360	2600	1	30
35	37	25	3	2	3	3	4	1	540	3400	1	35
36	43	42	3	3	4	1	2	0	270	1700	1	25
37	48	43	3	3	2	1	3	1	450	2500	1	40
38	24	26	4	2	4	1	4	2	260	2700	2	30
39	59	57	4	4	6	4	5	3	780	6800	1	85
40	34	25	3	3	3	5	3	1	440	2800	1	25

- a) Ingresar esta información a la hoja de cálculo de Minitab.
  - b) Asignar nombre a cada una de las columnas.
  - c) Ejecutar el comando INFO.
  - d) Grabar la información como archivo Minitab, con el nombre ENC9402.
10. Formar un archivo Minitab y otro, en código ASCII con el nombre ENC9403, con los datos de los archivos ENC9401.MTW y ENC9402.-MTW . Los datos provenientes del archivo ENC9401.MTW deben ocupar las primeras trece columnas del nuevo archivo y los datos provenientes de ENC9402.MTW deben ocupar las trece últimas columnas del nuevo archivo. Los nombres de las columnas del nuevo archivo ENC9403.MTW deben ser, respectivamente: código1 (columna C1), edadjf1 (C2), edadco1 (C3), instjf1 (C4), instco1 (C5), ocupjf (C6), ocupco1 (C7), numper1 (C8), nhijos (C9), gastose1 (C10), ingreso1 (C11), numveh1 (C12), gastoco1 (C13), código2 (C14), edadjf2 (C15), edadco2 (C16), instjf2 (C17), instco2 (C18), ocupjf2 (C19), ocupco2 (C20), numper2 (C21), nhijos2 (C22), gastose2 (C23), ingreso2 (C24), numveh2 (C25) y gastoco2 (C26).

- 
11. Recuperar el archivo ENC9401.MTW y realizar lo siguiente:
- a) Adicionar, a las trece columnas, los 40 datos del archivo ENC940-2.MTW para formar un archivo con trece columnas y 90 filas. No modifique los nombres de las columnas.
  - b) Grabar el nuevo conjunto de datos con el nombre ENC9404, en las dos formas; es decir, como archivo Minitab y en código ASCII.
  - c) Recuperar el archivo ENC9404.MTW y verificar mediante el comando INFO.
  - d) Recuperar el archivo ENC9404.DAT y verificar su contenido.

# III

## Manejo, edición y transformación de datos

Una de las bondades notables de Minitab es el manejo sencillo de los datos. Durante el proceso de información, usualmente, se realizan tareas de corrección de errores, eliminación o adición de datos y creación o transformación de datos. En esta sección se muestra cómo realizar dichos procesos.

### 1. Corrección de datos

#### Comando LET

Una vez que se ingresa datos a la memoria, es usual detectar valores errados. Cuando esto ocurre, dichos valores se pueden corregir mediante el comando LET.

Para hacer esto, se debe identificar en qué columna y fila se encuentra el dato errado. Por ejemplo, supóngase que primero se efectúa el siguiente ingreso de datos:

```
MTB > READ c1-c5
DATA > 1 3 12 25 8 900
DATA > 2 8 15 32 4 450
DATA > 3 2 19 42 6 750
DATA > 4 7 17 34 9 890
DATA > END
```

En este conjunto de datos se detecta que el tercer valor (750) de la columna C5 está errado y el valor correcto es 650. Para corregirlo, debe realizar:

```
MTB > LET c5(3)=650
```

Para verificar que esta corrección ha sido efectuada, puede usar el comando PRINT. Por ejemplo,

```
MTB > print c1-c5
```

ROW	C1	C2	C3	C4	C5
1	3	12	25	8	900
2	8	15	32	4	450
3	2	19	42	6	650
4	7	17	34	9	890

Como puede apreciarse, el comando LET permite el ingreso de un valor numérico en cualquier celda de la hoja de cálculo.

## 2. Creación de nuevas variables o datos

### Comando LET

Además de lo mostrado anteriormente, el uso más frecuente del comando LET es para la creación de nuevas variables, a partir de los datos que ya tiene en la hoja de cálculo. Asimismo, Minitab ofrece numerosas funciones matemáticas y lógicas, las cuales pueden ser utilizadas para transformar variables y crear así otras nuevas. Por ejemplo, si para el caso anterior, se desea expresar los valores de la columna C3 en porcentajes, conociendo que la suma de los valores en dicha columna es 133; entonces, se debe ejecutar:

```
MTB > LET C6 = (C3/133)*100
MTB > NAME C6 'PORCENT'
```

Con esto se crea una nueva variable en la columna C6, con nombre 'PORCENT', a partir de los datos de la columna C3.

Es importante tener en cuenta que la columna referenciada, a la izquierda del símbolo "=" en el comando LET, debe ser diferente a las columnas ya existentes en memoria; de lo contrario, se pierde la información. Por ejemplo,

$$\text{MTB} > \text{LET C3} = (\text{C3}/133)*100$$

efectúa lo mismo que la sentencia anterior, con la diferencia que la información existente en la columna C3 se ha perdido y ahora, esta columna contiene los valores porcentuales deseados. Si no se desea que esto ocurra, debe indicar una columna diferente.

### 3. Uso de funciones para crear nuevas variables

Las funciones matemáticas más importantes, usadas por Minitab, para transformar datos o crear nuevas variables son:

<b>ABSOLUTE</b>	Genera el valor absoluto del argumento.
<b>SQRT</b>	Genera la raíz cuadrada del argumento.
<b>LOGTEN</b>	Genera el logaritmo decimal del argumento.
<b>LOGE</b>	Genera el logaritmo neperiano del argumento.
<b>EXPO</b>	Calcula el valor del número "e", elevado a la potencia indicada en el argumento.
<b>ANTILOG</b>	Calcula el valor de 10, elevado a la potencia indicada en el argumento.
<b>ROUND</b>	Redondea el valor del argumento al entero más cercano.
<b>SIN</b>	Calcula el seno del argumento, el cual debe estar expresado en radianes.
<b>COS</b>	Calcula el coseno del argumento, el cual debe estar expresado en radianes.
<b>TAN</b>	Calcula la tangente del argumento que debe de estar expresado en radianes.
<b>ASIN</b>	Calcula el arco seno del argumento y expresa el resultado en radianes.

<b>ACOS</b>	Calcula el arco coseno del argumento y expresa el resultado en radianes.
<b>ATAN</b>	Calcula el arco tangente del argumento y expresa el resultado en radianes.
<b>SIGNS</b>	Genera +1, 0 y -1 si el argumento es positivo, cero o negativo, respectivamente.
<b>NSCORES</b>	Calcula los valores estandarizados de los datos. Los valores estandarizados tienen media cero y variancia uno. Un valor estándar se obtiene restando al dato el valor de la respectiva media y luego, se le divide entre la desviación estándar correspondiente.
<b>PARSUMS</b>	Calcula sumas parciales o acumulativas.
<b>PARPRODUCTS</b>	Calcula productos parciales o acumulativos.

El argumento de una función es lo que sigue a continuación del nombre de la función y debe estar encerrado entre paréntesis. Se puede usar solamente funciones u operadores, o una combinación de ellos. El resultado producido por una función puede servir, a su vez, como argumento de otra función. Enseguida se presentan algunos ejemplos del uso de funciones, para transformar datos o crear nuevas variables.

```
MTB > LET C6 = COS( C3*K100/(90*2))
MTB > LET C4 = SQRT( C3 + 4 )
MTB > LET C8 = SQRT(SIN(C3*K100/(90*2)) + 0.5 )
```

En el primer caso, se obtiene el coseno de los valores de la columna C3 que son multiplicados por la constante K100 (3.1415..) y divididos entre el producto de 90 por 2.

Para usar estas funciones hay que tener en cuenta el orden en el cual se ejecutan las operaciones matemáticas: primero, son evaluadas las funciones; a continuación, las expresiones entre paréntesis; luego, la potenciación (\*\*), que es ejecutada antes que la multiplicación y división (\* y /), las cuales se ejecutan antes que la adición y resta (+ y -); y, finalmente, los operadores de la misma prioridad (\* y /, + y -) que se ejecutan de izquierda a derecha.

#### 4. Eliminación de datos

Minitab permite eliminar datos de una hoja de cálculo. Es posible eliminar datos de toda una columna o fila, de una parte de una columna o fila, o de un segmento rectangular de la hoja de cálculo. Los comandos para eliminar datos son ERASE y DELETE.

##### Comando ERASE

Este comando permite eliminar columnas de datos. Por ejemplo,

```
MTB > READ 'ENCU' C1-C13
      5 ROWS READ
```

ROW	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
1	1	28	26	3	3	1	1	2	0	350	2500	1	50
2	2	32	33	3	3	2	4	3	1	480	3600	1	30
3	3	44	40	3	3	1	8	4	2	550	5000	2	65
4	4	36	31	3	3	3	4	3	1	490	3000	1	40

```
MTB > ERASE C1 C3
MTB > PRINT C1-C13
```

ROW	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
1		28		3	3	1	1	2	0	350	2500	1	50
2		32		3	3	2	4	3	1	480	3600	1	30
3		44		3	3	1	8	4	2	550	5000	2	65
4		36		3	3	3	4	3	1	490	3000	1	40
5		42		3	3	4	1	4	2	600	4600	1	45

Como se puede apreciar, con el comando ERASE se han eliminado las columnas C1 y C3. Para eliminar varias columnas consecutivas se puede usar, por ejemplo,

```
MTB > ERASE C5-C10
MTB > PRINT C1-C13
```

ROW	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
1		28		3						2500	1	50	
2		32		3						3600	1	30	
3		44		3						5000	2	65	
4		36		3						3000	1	40	
5		42		3						4600	1	45	

Las instrucciones realizadas, mediante el comando ERASE, han eliminado los datos de las columnas C5, C6, C7, C8, C9 y C10.

### Comando DELETE

Este comando permite eliminar filas específicas o valores particulares, de una o más filas, para una o más columnas. Por ejemplo,

```
MTB > PRINT C1-C13
```

ROW	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
1	1	28	26	3	3	1	1	2	0	350	2500	1	50
2	2	32	33	3	3	2	4	3	1	480	3600	1	30
3	3	44	40	3	3	1	8	4	2	550	5000	2	65
4	4	36	31	3	3	3	4	3	1	490	3000	1	40
5	5	42	40	3	3	4	1	4	2	600	4600	1	45

```
MTB > DELETE ROWS 4 OF C1-C13
```

```
MTB > PRINT C1-C13
```

ROW	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
1	1	28	26	3	3	1	1	2	0	350	2500	1	50
2	2	32	33	3	3	2	4	3	1	480	3600	1	30
3	3	44	40	3	3	1	8	4	2	550	5000	2	65
4	5	42	40	3	3	4	1	4	2	600	4600	1	45

Como puede apreciarse, todos los valores de la fila 4 han sido eliminados. Ahora, suponga que se desea eliminar los valores de las filas 1 y desde la 3 a la 5, correspondientes a las columnas de C1 a C5 y de C8 a C13; luego, se tendrá.



MTB > PRINT C1-C13

ROW	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
1	1	28	26	3	3	1	1	2	0	350	2500	1	50
2	2	32	33	3	3	2	4	3	1	480	3600	1	30
3	3	44	40	3	3	1	8	4	2	550	5000	2	65
4	4	36	31	3	3	3	4	3	1	490	3000	1	40
5	5	42	40	3	3	4	1	4	2	600	4600	1	45

MTB > DELETE ROWS 1, 3:5 OF C1-C5, C8-C13

MTB > PRINT C1-C13

ROW	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13
1	2	32	33	3	3	1	1	3	1	480	3600	1	30
2						2	4						
3						1	8						
4						3	4						
5						4	1						

Debe notar que para indicar una serie de filas se usó dos puntos y para indicar una serie de columnas, se usa un guión.

Es conveniente señalar que al usar los comandos ERASE y DELETE, la información se pierde y no hay forma de recuperarla. Si desea conservar sus datos, antes de usar estos comandos, debe almacenar la información en un archivo, tal como se vio anteriormente.

## 5. Copia de datos seleccionados

### Comando COPY

Algunas veces, es necesario copiar algunos datos de una o más columnas a nuevas columnas. Mediante el comando COPY, resulta sencillo hacer este proceso. Por ejemplo,

MTB > PRINT C1-C5

ROW	C1	C2	C3	C4	C5
1	15	8			6
2	16	4			9
3	14	3			8
4	15	5			7

```
MTB > COPY C2 C5 to C3 C6;
SUBC > USE ONLY 1:3.
MTB > PRINT C1-C6
```

ROW	C1	C2	C3	C4	C5	C6
1	15	8	8		6	6
2	16	4	4		9	9
3	14	3	3		8	8
4	15	5			7	

En este caso, solo para las filas 1 a 3, se ha copiado los valores de la columna C2 a la columna C3 y los de la columna C5 a la columna C6.

El subcomando USE ONLY tiene una opción de lógica. Por ejemplo,

```
MTB > COPY C2 C5 to C3 C6;
SUBC > USE ONLY C1=14:15.
MTB > PRINT C1-C6
```

ROW	C1	C2	C3	C4	C5	C6
1	15	8	8		6	6
2	16	4	3		9	8
3	14	3	5		8	7
4	15	5			7	

Con las instrucciones dadas, se ha copiado los valores de la columna C2 a la columna C3 y los de la columna C5 a la columna C6; pero la copia se ha efectuado sólo para los valores de las filas donde el valor de la columna C1 está comprendido entre 14 y 15.

## 6. Asignación de códigos

### Comando CODE

Este comando permite crear una columna de códigos, en base a los valores de otra columna. Por ejemplo,

```
MTB > CODE (10:13) 1, (6:9) 2, (0:5) 3 IN C1 store in C4
MTB > PRINT C1-C4
```

ROW	C1	C2	C3	C4
1	12			1
2	6			2
3	10			1
5	3			3
6	8			2

Como puede verse, si el valor de C1 es menor que 6, se asigna el código 3, si está entre 6 y 9, se asigna 2; y si está entre 10 y 13, se asigna el código 1.

## 7. Concatenación de columnas

### Comando STACK

Este comando permite reunir los datos, de una o más columnas, en una nueva columna, colocándolas una a continuación de la otra. Por ejemplo,

```
MTB > STACK 'DIA1' 'DIA2' store in C3
MTB > PRINT C1-C3
```

ROW	DIA1	DIA2	C3
1	15	8	15
2	16	4	16
3	14	3	14
4	15	5	15
5			8
6			4
7			3
8			5

El comando **STACK** ha permitido colocar los datos de las columnas 'DIA1' y 'DIA2' en la columna C3.

## 8. División de columnas

### Comando **UNSTACK**

Este comando hace lo contrario que el comando **STACK**. Toma los valores de una columna y lo divide en columnas más cortas. Se debe indicar dónde poner las nuevas columnas, especificándolas en el mismo comando, y señalar qué valores pertenecen a cada una, mediante una columna de subíndices (entradas con el mismo subíndice se asignan a una misma columna).

Si no especifica una columna de subíndices (**SUBSCRIPT**), Minitab creará una nueva columna (con un solo valor), con cada dato de la columna a dividir.

Para mostrar cómo funciona este comando, retomemos el ejemplo del comando **CODE**, con el cual se creó una columna de códigos o subíndices. Sobre la base de dichos resultados se tiene,

```
MTB > UNSTACK (C1) into (C5) (C6) (C8);  
SUBC > SUBSCRIPT C2.  
MTB > PRINT C1-C2 C5-C8
```

ROW	C1	C2	C5	C6	C7	C8
1	12	1	12	6		3
2	6	2	10	8		
3	10	1	11			
4	3	3				
5	8	2				
6	11	1				

La columna C5 contiene los valores de la columna C1, para los cuales el código asignado en la columna C2 es igual a 1; la columna C6 contiene los valores de la columna C1, para los cuales el código asignado en la columna C2 es igual a 2 y la columna C8 contiene los valores de la columna C1, para los cuales el código asignado en la columna C2 es igual a 3.

## Ejercicios

1. Ingresar a Minitab y ejecutar el comando **OUTFILE**, para almacenar los resultados de su sesión en el archivo **PRACT2.LIS**
2. Leer el archivo **ENCU.DAT** y ejecutar lo siguiente:

- a) Insertar al final del archivo las dos líneas de datos siguientes:

6	33	29	4	4	2	3	4	3	500	1600	2	90
7	41	37	3	4	3	9	2	0	400	2400	1	50

- b) Eliminar las columnas de datos **C6**, **C7** y **C9**.
- c) Eliminar la segunda fila de datos.
- d) Eliminar los datos de las filas 3 y 4, para las columnas **C10** a **C11**.
- e) Copiar los datos de las columnas **C3** y **C5** a las columnas **C14** y **C15**, respectivamente.
- f) Crear una columna de códigos, de acuerdo con los datos de la columna **C13** (gasto semanal en combustible) y de acuerdo con la siguiente definición:

Intervalo	Código
Menos de 30	1
De 30 a menos de 40	2
De 40 a menos de 50	3
50 a más	4

- g) Colocar en una sola columna (**C17**), los valores de las columnas **C2** y **C3**.
- f) Almacenar su hoja de cálculo Minitab, en el archivo, con nombre **PRACT21**.

3. Supóngase que el siguiente cuadro muestra las cantidades vendidas de tres artículos, en cinco días de la semana pasada, y los precios de venta.

Día	Artículos			
	A	B	C	D
Lunes	32	32	20	25
Martes	24	94	35	30
Miércoles	31	77	44	24
Jueves	28	62	75	38
Viernes	32	37	30	21
Sábado	20	20	56	45
Precio S/.	21	15	35	50

- Ingresar estos datos mediante el comando READ, a las columnas C1 a C4.
- Asignar los precios a las constantes K1 a K4.
- Obtener los ingresos, por día y por artículo, y asignarlos a las columnas C6 a C9.
- Obtener los ingresos diarios y asignarlos a la columna C11.
- Almacenar su hoja de cálculo Minitab, en el archivo, con nombre PRACT22.

# IV

## Aplicaciones: resúmenes de datos

Una de las primeras tareas, en el procesamiento de datos, es resumir la información en valores estadísticos simples. Minitab permite efectuar este proceso, para la información de cada columna o de cada fila.

### 1. Resumen de datos de columnas

Cada comando genera un solo valor, el cual es mostrado en pantalla o puede ser asignado a una constante para cualquier cálculo posterior. Los comandos para resumir datos de columnas son:

<b>COUNT</b>	Cuenta el número de valores en una columna, incluyendo valores perdidos.
<b>N</b>	Cuenta el número de valores (no perdidos) de una columna.
<b>NMISS</b>	Cuenta el número de valores perdidos de una columna.
<b>SUM</b>	Obtiene la suma de los valores no perdidos de una columna.
<b>MEAN</b>	Calcula el promedio de los valores no perdidos de una columna.
<b>STDEV</b>	Calcula la desviación estándar de los valores no perdidos de una columna.

**MEDIAN** Calcula la mediana de los valores no perdidos de una columna.

**MINIMUN** Halla el mínimo valor de una columna.

**MAXIMUN** Halla el máximo valor de una columna.

**SSQ** Calcula la suma de cuadrados sin corregir de los valores de una columna.

Para ver como trabaja Minitab, supóngase que los ingresos provenientes de la venta de 4 artículos, durante 6 días de una semana, son los siguientes:

Día	Artículos			
	A	B	C	D
Lunes	145	235	180	75
Martes	122	290	165	60
Miércoles	133	275	148	64
Jueves	118	265	170	58
Viernes	142	287	190	80
Sábado	160	320	95	

Los datos se pueden ingresar a la hoja de cálculo de la siguiente manera:

```

MTB > read c1-c4
DATA > 145 235 180 75
DATA > 122 290 165 60
DATA > 133 275 148 64
DATA > 118 265 170 58
DATA > 142 287 190 80
DATA > 160 320 * 95
DATA > END
6 ROWS READ

```

Puede apreciarse que el valor de las ventas del artículo C, correspondiente al día sábado, no existe; es decir, es un valor perdido. Durante el proceso de ingreso, Minitab espera 4 datos para cada día. La forma de indicar que un



dato específico no existe es usando un asterisco (\*), como se muestra en el ejemplo.

Sobre la base de los datos ingresados, se pueden ejecutar los siguientes comandos:

```
MTB > MEAN C1
      MEAN = 136.67
MTB > MEAN C1 IN K1
      MEAN = 136.67
MTB > PRINT K1
K1      136.667
MTB > LET K1=MEAN(C1)
MTB > LET K2=SUM(C2)
MTB > LET K3=SSQ(C3)
MTB > PRINT K1 K2 K3
K1      136.667
K2      1672.00
K3      146529
```

En el primer comando, se pide el promedio de los valores de la columna C1 y su resultado es mostrado en pantalla. En el segundo comando, se solicita lo mismo y, además, se indica a Minitab que el resultado se asigne a la constante K1.

Los comandos también pueden usarse como funciones en combinación con el comando LET. En estos casos el resultado no se muestra en pantalla y es necesario usar el comando PRINT como se muestra en las 7 últimas líneas del ejemplo.

## 2. Resumen de datos de filas

Los comandos, para resumir los datos de filas, trabajan de manera similar a los comandos para el manejo de columnas. Los comandos generan valores que pueden ser asignados a una nueva columna de datos. Los comandos para resumir datos de filas son:

**RCOUNT** Cuenta el número de valores en una fila, incluyendo valores perdidos, para las columnas que han sido especificadas.

<b>RN</b>	Cuenta el número de valores (no perdidos) de una fila, para las columnas especificadas.
<b>RNMISS</b>	Cuenta el número de valores perdidos de una fila, para las columnas especificadas.
<b>RSUM</b>	Obtiene la suma de los valores no perdidos de una fila, para las columnas especificadas.
<b>RMEAN</b>	Calcula el promedio de los valores no perdidos de una fila, para las columnas especificadas.
<b>RSTDEV</b>	Calcula la desviación estándar de los valores no perdidos de una fila, para las columnas especificadas.
<b>RMEDIAN</b>	Calcula la mediana de los valores no perdidos de una fila, para las columnas especificadas.
<b>RMINIMUN</b>	Halla el mínimo valor de una fila.
<b>RMAXIMUN</b>	Halla el máximo valor de una fila.
<b>RSSQ</b>	Calcula la suma de cuadrados sin corregir de los valores de una fila.

Para observar cómo funcionan estos comandos, se utilizarán los datos del ejemplo anterior. Puede ser de interés, también, obtener valores estadísticos asociados a cada día de la semana; además de los obtenidos para cada artículo (columnas) y que se mostró anteriormente.

Sobre la base de los datos ingresados en la hoja de cálculo, se pueden ejecutar los siguientes comandos:

```

MTB > RMEAN C1-C4 IN C5
MTB > RSUM C1-C4 IN C6
MTB > RN C1-C4 IN C7
MTB > NAME C5 'PROM.' C6 'TOTAL' C7 'NOBS.'
MTB > PRINT C1-C7

```

El paquete Minitab proporciona los siguientes resultados:

ROW	C1	C2	C3	C4	PROM.	TOTAL	NOBS.
1	145	235	180	75	158.750	635	4
2	122	290	165	60	159.250	637	4
3	133	275	148	64	155.000	620	4
4	118	265	170	58	152.750	611	4
5	142	287	190	80	174.750	699	4
6	160	320	*	95	191.667	575	3

En el primer comando se pide el promedio de los valores de cada fila, asociada a las columnas C1 a C4, y los resultados se almacenan en la columna C5. En el segundo comando se pide la suma de los valores de cada fila, asociada a las columnas C1 a C4, y los resultados se almacenan en la columna C6; y en el tercer comando se pide el número de valores de cada fila, asociada a las columnas C1 a C4, y los resultados se almacenan en la columna C7.

Para ver los resultados, es necesario listar en pantalla, con el comando PRINT, los valores de las nuevas columnas.

### 3. Otros comandos para resumir datos numéricos

Además de los comandos para resumir información por columnas y por filas, descritos anteriormente, Minitab ofrece comandos para generar estadísticos de los datos existentes en una o más columnas. Estos comandos son: DESCRIBE y TALLY.

#### Comando DESCRIBE

Este comando permite generar estadísticos descriptivos de la información de una o más columnas.

Para ver el uso de este comando se utilizará el archivo Minitab 'ENCU.MTW', el cual puede ser leído de la siguiente manera:

```
MTB > RETRIEVE 'ENCU'
WORKSHEET SAVED 3/11/1994
```

Worksheet retrieved from file: ENCU.MTW

Enseguida se aplica el comando DESCRIBE,

```
MTB > DESCRIBE C2 C3
```

El computador proporciona los siguientes resultados:

	N	MEAN	MEDIAN	TRMEAN	STDEV	SEMEAN
edadjf	5	36.40	36.00	36.40	6.69	2.99
edadco	5	34.00	33.00	34.00	6.04	2.70
	MIN	MAX	Q1	Q3		
edadjf	28.00	44.00	30.00	43.00		
edadco	26.00	40.00	28.50	40.00		

Como puede apreciarse, el comando DESCRIBE permite obtener los valores del número total de observaciones (N), el promedio (MEAN), la mediana (MEDIAN), el promedio central del 90% o TRMEAN (obtenido eliminando el 5% de los valores más pequeños y el 5% de valores más altos y promediando, el 90% de los valores restantes), la desviación estándar (STDEV), la desviación estándar de promedios (SEMEAN), el valor mínimo (MIN), el valor máximo (MAX), el primer cuartil (Q1) y el tercer cuartil (Q3). Cuando existen valores perdidos, aparece una columna adicional titulada N\*.

También es posible obtener estadísticos para una o más columnas, agrupando la información de acuerdo con una columna que contiene números enteros y que sirve como criterio de clasificación. Por ejemplo, para obtener estadísticos para las variables 'EDADJF' y 'EDADCO', según la variable "número de personas que conforman la familia" cuya información está en la columna 'NUMPER', se debe realizar lo siguiente:

```
MTB > DESCRIBE 'EDADJF' 'EDADCO';
SUBC > BY 'NUMPER'.
```

	numper	N	MEAN	MEDIAN	TRMEAN	STDEV	SEMEAN
edadjf	2	1	28.000	28.000	28.000	*	*
	3	2	34.00	34.00	34.00	2.83	2.00
	4	2	43.00	43.00	43.00	1.41	1.00
edadco	2	1	26.000	26.000	26.000	*	*
	3	2	32.00	32.00	32.00	1.41	1.00
	4	2	40.000	40.000	40.000	0.000	0.000

	number	MIN	MAX	Q1	Q3
edadjf	2	28.000	28.000	*	*
	3	32.00	36.00	*	*
	4	42.00	44.00	*	*
edadco	2	26.000	26.000	*	*
	3	31.00	33.00	*	*
	4	40.000	40.000	*	*

En ocasiones, es necesario calcular estadísticos que no son generados por el comando DESCRIBE. Minitab brinda la flexibilidad suficiente para obtenerlos, mediante el uso de sus funciones. Por ejemplo, para obtener el coeficiente de asimetría de Pearson, se puede operar de la siguiente manera:

```
MTB > LET K1= 3*(MEAN(C2)-MEDIAN(C2))/STDEV(C2)
MTB > PRINT K1
K1      0.179285
```

Si prefiere calcular el coeficiente de asimetría sobre la base de la definición de momentos, se puede operar como sigue.

```
MTB > LET K2=MEAN((C2-MEAN(C2))**2) # Segundo
momento muestral
MTB > LET K3=MEAN((C2-MEAN(C2))**3) # Tercer momento
muestral
MTB > LET K4=K3/(K2**1.5) # Coeficiente de asimetría
MTB > PRINT K4
K4      -0.0590607
```

Obviamente, éste es un resultado diferente al anterior y su interpretación depende de la definición del indicador.

Además, en las instrucciones anteriores se ha utilizado el caracter "#". Este símbolo sirve para indicar a Minitab que, a continuación de las instrucciones correspondientes al comando en la respectiva línea, se tiene un comentario.

### Comando TALLY

Con este comando se puede formar tablas de frecuencias para variables cualitativas y cuantitativas discretas. Por ejemplo, para resumir la información de las variables "ocupación del jefe de familia" y "ocupación de la cónyuge" se puede ejecutar:

```
MTB > TALLY 'OCUPJF' 'OCUPCO'
```

El computador proporciona:

ocupjf	COUNT	ocupco	COUNT
1	2	1	2
2	1	4	2
3	1	8	1
4	1	N=	5
N=	5		

Obyiamente, para analizar estos resultados es necesario conocer que el código 1 de la variable 'OCUPJF' se refiere a la ocupación "Industrial", que el código 2, se refiere a la ocupación "Comerciante" y así, sucesivamente.

El comando TALLY tiene, además, los siguientes subcomandos:

<b>COUNTS</b>	Permite obtener las frecuencias absolutas.
<b>PERCENTS</b>	Permite obtener los porcentajes no acumulados.
<b>CUMCOUNTS</b>	Permite obtener las frecuencias acumuladas.
<b>CUMPERCENTS</b>	Permite obtener los porcentajes acumulados.
<b>ALL</b>	Genera las frecuencias absolutas, porcentajes no acumulados, frecuencias acumuladas y porcentajes acumulados.

Si en el ejemplo anterior, adicionamos el subcomando ALL:

```
MTB > TALLY 'OCUPJF' 'OCUPCO';
SUBC > ALL.
```

se obtiene los siguientes resultados:

ocupjf	COUNT	CUMCNT	PERCENT	CUMPCT
1	2	2	40.00	40.00
2	1	3	20.00	60.00
3	1	4	20.00	80.00
4	1	5	20.00	100.00
N=	5			

ocupco	COUNT	CUMCNT	PERCENT	CUMPCT
1	2	2	40.00	40.00
4	2	4	40.00	80.00
8	1	5	20.00	100.00
N=	5			

### Ejercicios

1. Ingresar a Minitab, activar la opción de generación de un archivo de reportes, mediante el comando `OUTFILE` usando el nombre 'PRACT4', y ejecutar lo siguiente:
  - a) Leer el archivo `ENC9401.MTW`.
  - b) Obtener valores estadísticos para las variables: "ingreso familiar mensual", "gasto familiar semanal en comida" y "gasto familiar semanal en combustible".
  - c) Obtener lo pedido en el punto anterior, para los datos agrupados por "ocupación del jefe de familia".
  - d) Mediante el comando `TALLY` resumir la información de las variables: "grado de instrucción del jefe de familia" y "grado de instrucción de la cónyuge".
  - e) Repetir, lo pedido en el punto anterior, utilizando cada uno de los subcomandos de la instrucción `TALLY`.
  
2. Recuperar el archivo `ENC9403.MTW` y realizar lo siguiente:
  - a) Activar el comando `OUTFILE` y generar un archivo con el nombre `PRAC41`.
  - b) Obtener un resumen de resultados (medidas estadísticas) para las columnas 'ingreso1', 'ingreso2', 'instjf1' e 'instjf2'.
  - c) Desactivar el comando `OUTFILE` y salir del paquete Minitab.
  - d) Ingresar al paquete Word Perfect 5.1 y recuperar el archivo `PRAC41.LIS`.

- e) Analizar el resumen de resultados (medidas estadísticas) de las variables 'ingreso1' e 'ingreso2'. Hacerlo en forma individual y de manera comparativa.
- f) Analizar el resumen de resultados (medidas estadísticas) de las variables 'instjf1' e 'instjf2'.
- g) Redactar las conclusiones, en el archivo PRAC41.LIS, para los análisis realizados anteriormente. Luego, hacer la impresión del informe respectivo.



# V

## Aplicaciones: representaciones gráficas de datos

Minitab proporciona varios procedimientos para obtener resúmenes gráficos de un conjunto de datos. Para ilustrar el uso de los comandos, que generan estos resúmenes, se utilizarán los datos de la encuesta archivados en ENC9401.MTW. Para ingresar los datos del archivo, antes mencionado, se ejecuta:

```
MTB > retrieve 'enc9401'  
WORKSHEET SAVED 3/16/1994
```

Worksheet retrieved from file: enc9401.MTW

Con el comando INFO, se obtiene la información básica del conjunto de datos ingresados a la hoja de cálculo de Minitab.

```
MTB > info`
```

COLUMN	NAME	COUNT
C1	codigo	50
C2	edadjf	50
C3	edadco	50
C4	instjf	50
C5	instco	50
C6	ocupjf	50
C7	ocupco	50

C8	numper	50
C9	nhijoesc	50
C10	gastosem	50
C11	ingreso	50
C12	numvehi	50
C13	gastocom	50

CONSTANTS USED: NONE

## 1. Histogramas horizontales

### Comando HISTOGRAM

Minitab genera histogramas sencillos sobre la base de intervalos de igual amplitud, mostrando los puntos medios de cada intervalo y el número de observaciones que cae dentro de cada intervalo. Por ejemplo,

```
MTB > histogram 'edadjf'
```

proporciona en la figura 5.1, el histograma para la variable edad del jefe de familia.

Histogram of edadjf N = 50

Midpoint	Count	
28	3	***
32	6	*****
36	6	*****
40	5	*****
44	11	*****
48	11	*****
52	6	*****
56	2	**

Figura 5.1

Para la variable ingreso familiar mensual se tiene,

```
MTB > histogram 'ingreso'
```

Histogram of ingreso N = 50

Midpoint	Count	
2500	1	*
3000	7	*****
3500	2	**
4000	11	*****
4500	10	*****
5000	11	*****
5500	3	***
6000	1	*
6500	3	***
7000	1	*

Figura 5.2

Como no se ha especificado el número de intervalos, Minitab lo establece automáticamente. La primera columna muestra los puntos medios de cada intervalo y, la segunda columna, el número de observaciones o frecuencia absoluta que corresponde a cada intervalo. Mediante subcomandos, es posible controlar la amplitud de cada intervalo y el punto medio del primer intervalo. Por ejemplo, para obtener la figura 5.3, se debe ejecutar:

```
MTB > histogram 'ingreso';
SUBC > increment=1000;
SUBC > start 2500.
```

Histogram of ingreso N = 50

Midpoint	Count	
2500	4	****
3500	12	*****
4500	20	*****
5500	10	*****
6500	4	****

Figura 5.3

El subcomando INCREMENT=1000, indica a Minitab que la amplitud de cada intervalo es 1000 y el subcomando START 2500, indica que el primer intervalo debe tener como punto medio el valor de 2500.

### Comando GHISTOGRAM

Este comando permite obtener histogramas de alta resolución; es decir, construye los histogramas como conjuntos de rectángulos, sin espacios vacíos entre los rectángulos que representan a los intervalos de clase. Este comando trabaja de la misma forma que el comando HISTOGRAM, anteriormente presentado. Por ejemplo, al ejecutar:

```
MTB > ghistogram 'edadjf' 'ingreso'
```

se obtienen los gráficos proporcionados en las figuras 5.4 y 5.5, respectivamente.

El comando GHISTOGRAM también trabaja con los subcomandos INCREMENT y START.

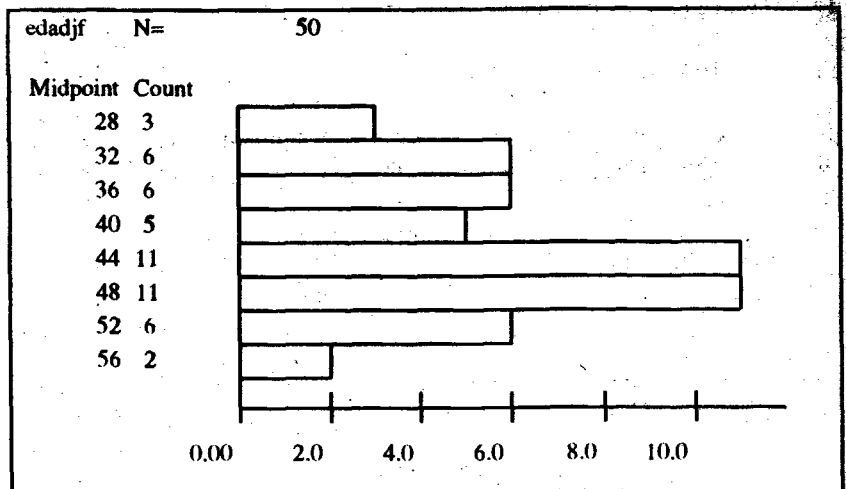


Figura 5.4

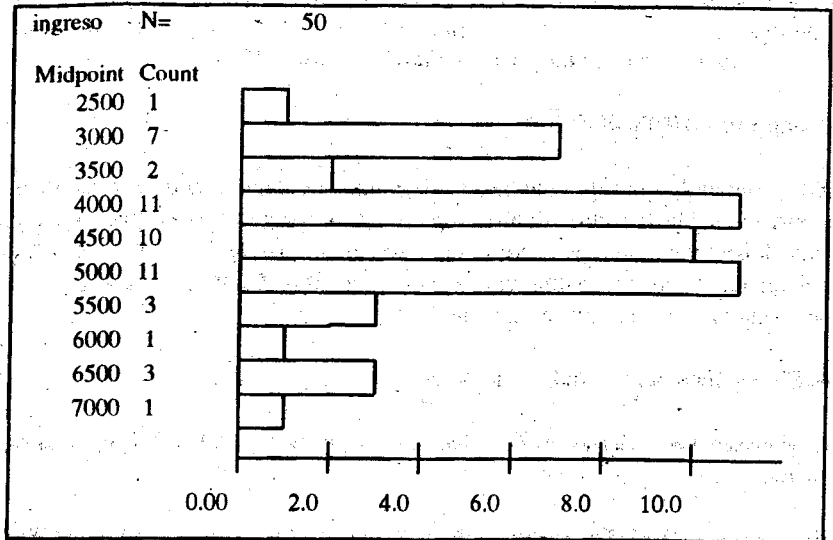


Figura 5.5

## 2. Histogramas verticales

### Comando DOTPLOT

Este comando genera un histograma similar al proporcionado por el comando HISTOGRAM, pero las barras son verticales y están representadas por puntos. Por ejemplo,

```
MTB > dotplot 'edadjf'
```

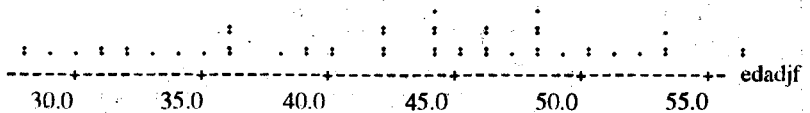


Figura 5.6

Los puntos alineados, verticalmente, muestran las frecuencias de ocurrencia. Debe tenerse presente que cada punto no está asociado a una observación.

```
MTB > dotplot 'edadjf' 'edadco';
SUBC > same.
```

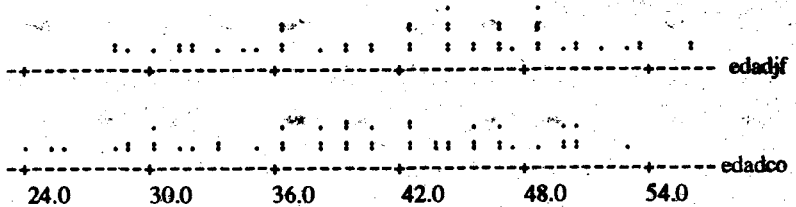


Figura 5.7

En este ejemplo, se ha usado el subcomando SAME, para indicar a Minitab que los intervalos deben tener la misma amplitud. Esto es de utilidad, cuando se desea obtener dos o más gráficos para fines de comparación.

```
MTB > dotplot 'ingreso'
```

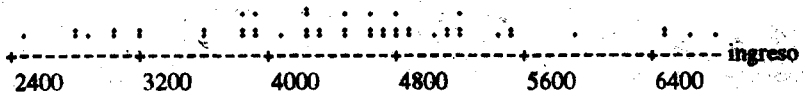


Figura 5.8

En este ejemplo también se muestra la aplicación del comando DOTPLOT. La presentación se mejora con los subcomandos INCREMENT y START, que se utilizan de manera similar al comando HISTOGRAM. Por ejemplo, para tener intervalos de una amplitud de 1000 y que el punto medio del primer intervalo sea igual a 2500, se debe ejecutar:

```
MTB > dotplot 'ingreso';
SUBC > increment=1000;
SUBC > start 2500.
```

Se obtiene el gráfico de la figura 5.9.

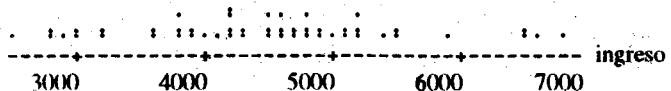


Figura 5.9

### 3. Gráficos de cajas

#### Comando BOXPLOT

Este comando permite obtener un gráfico de cajas. También tiene los subcomandos INCREMENT y START. Por ejemplo, si se ejecuta:

```
MTB > boxplot 'ingreso';
SUBC > start 2400;
SUBC > increment=800.
```

El computador proporciona el gráfico de la figura 5.10.

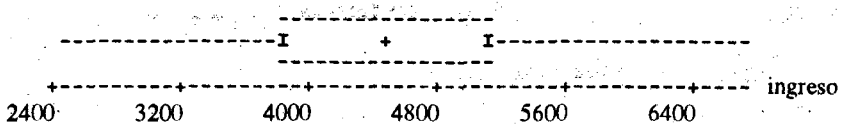


Figura 5.10

El gráfico de cajas para la variable INGRESO, muestra la posición de la mediana mediante el símbolo "+". La caja cubre el rango de la mitad central de los datos, terminando a la izquierda y derecha con cuantiles.

### 4. Gráficos en dos dimensiones

#### Comando PLOT

Este comando permite generar gráficos bidimensionales, para mostrar las relaciones entre dos variables. Por ejemplo,

```
MTB > plot 'edadco' 'edadjf'
```

proporciona la figura 5.11.

La primera variable o columna, descrita en el comando (EDADCO), se presenta en el eje vertical y la segunda variable o columna (EDADJF), en el eje horizontal. Cada asterisco (\*) representa una observación. Los números que aparecen en el gráfico (2,3, etc.) indican el número de observaciones que coincide en dicha posición.

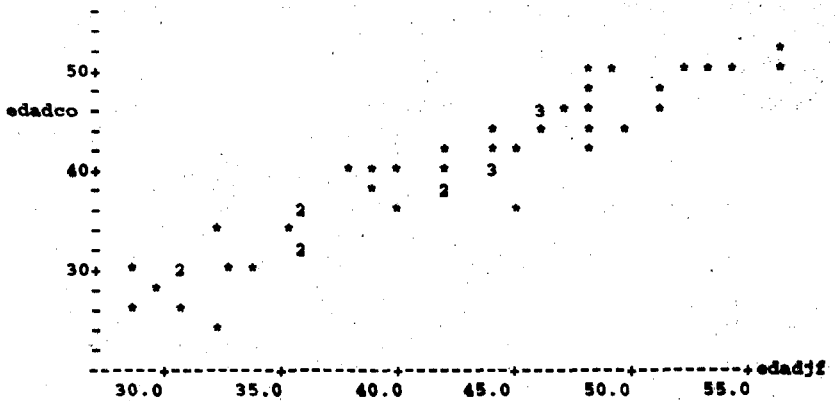


Figura 5.11

### Comando M PLOT

Este comando permite generar gráficos bidimensionales, en ellos se representan dos o más gráficos individuales como los descritos con el comando PLOT. Es de utilidad para comparar las relaciones entre dos o más subconjuntos de datos. Por ejemplo, el siguiente comando proporciona el gráfico de la figura 5.12.

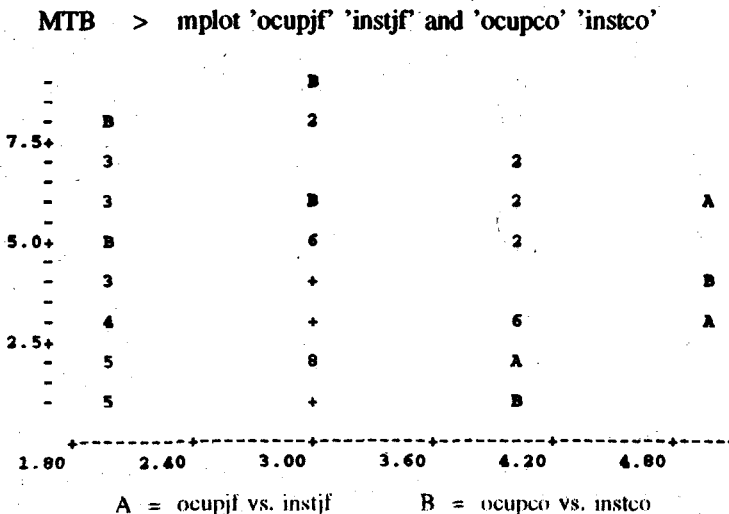


Figura 5.12



- a) Obtener el histograma para la variable "gasto familiar semanal en comida". Utilizar los comandos HISTOGRAM y GHISTOGRAM. ¿Cuáles son las diferencias?
- b) Obtener, con el comando DOTPLOT, el histograma para la variable "gasto familiar semanal en comida".
- c) Obtener el gráfico de cajas con los datos de "gasto familiar mensual en combustible".
- d) Obtener el gráfico de dispersión de datos de "gasto familiar mensual en comida" y "gasto familiar semanal en combustible". ¿Para qué sirve este gráfico?
- e) Obtener el gráfico de dispersión de datos de "gasto familiar mensual en comida" contra "gasto familiar semanal en combustible", agrupado por el "número de hijos en edad escolar".
- f) Desactivar el comando OUTFILE; luego, mediante el paquete WP5.1, recuperar el archivo PRAC3.LIS y realizar el análisis de los gráficos construidos.

# VI

## Aplicaciones: distribuciones de probabilidades y números aleatorios

En este capítulo se describe el uso de comandos que permiten obtener valores de probabilidad o percentiles, asociados a las distribuciones de probabilidades más usadas.

### 1. Cálculo de una probabilidad acumulada

#### Comando CDF

Este comando permite obtener la probabilidad acumulada para las distribuciones: Binomial, Poisson, Uniforme (discreta y continua), Normal, T, Chi-cuadrado y F. El tipo de distribución se indica en el subcomando, con los parámetros siguientes:

Distribución	Parámetro	Comentario
BINOMIAL	$n=k$ $\pi=m$	$\pi=P(\text{Éxito})$
POISSON	mean= $k$	
INTEGER	$k$ to $k$	Dist. Discreta desde $k$ hasta $k$ .
UNIFORM	$k$ to $k$	Dist. Continua desde $k$ hasta $k$ .
NORMAL	$\mu=k$ $\sigma=m$	
T	df= $k$	Grados de libertad.
CHISQUARE	df= $k$	Grados de libertad.
F	df= $k$ df= $m$	Grados de libertad del numerador y del denominador.

Por ejemplo,

```
MTB > CDF 2.3;
SUBC > NORMAL MU=0 SIGMA=1.
      2.3000  0.9893
```

El valor 0.9893 indica la probabilidad de elegir al azar un valor menor o igual a 2.3, en una distribución normal estándar.

Si no se especifica el argumento (parámetros), el computador asume una distribución normal con media 0 y variancia 1.

Con las otras distribuciones se opera de manera similar. Por ejemplo,

```
MTB > CDF 8;
SUBC > BINOMIAL N=12 P=0.4.
      K P( X LESS OR = K)
      8.00  0.9847
```

```
MTB > CDF 8;
SUBC > POISSON MEAN=8.2.
      K P( X LESS OR = K)
      8.00  0.5647
```

```
MTB > CDF 18;
SUBC > INTEGER 11 20.
      18.00  0.8000
```

```
MTB > CDF 15;
SUBC > UNIFORM 13 17.
      15.0000  0.5000
```

```
MTB > CDF 2.16;
SUBC > T 25.
      2.1600  0.9797
```

```
MTB > CDF 33.4;
SUBC > CHIS 15.
      33.4000  0.9959
```

```

MTB > CDF 2.15;
SUBC > F 21 35.
          2.1500          0.9781

```

## 2. Cálculo de percentiles

### Comando INVCDF

Este comando permite obtener el percentil de una distribución; es decir, el valor de la variable de cierta función de probabilidad, asociada a una probabilidad acumulada.

De manera similar al comando CDF, se puede usar el comando INVCDF con las distribuciones Binomial, Poisson, Uniforme (discreta y continua), Normal, T, Chi-cuadrado y F. El tipo de distribución se indica en el subcomando, donde los parámetros son los mismos que los indicados para el comando CDF.

Por ejemplo,

```

MTB > INVCDF 0.75;
SUBC > NORMAL 200 15.
          0.7500 210.1174

```

El valor 210.1174 viene a ser el percentil  $P_{0.75}$  de la distribución normal, con media 200 y desviación estándar 15; es decir, que el 75% de los valores de esta distribución son menores a 210.1174.

De manera similar, se opera con las otras distribuciones. Por ejemplo,

```

MTB > INVCDF 0.5;
SUBC > BINOMIAL N=15 P=0.3.
          K P(X LESS OR = K)    K P(X LESS OR = K)
          3 0.2969              4 0.5155

```

```

MTB > INVCDF 0.6;
SUBC > POISSON MEAN=12.5.
          K P(X LESS OR = K)    K P(X LESS OR = K)
          12 0.5190             13 0.6278

```

```

MTB > INVCDF 0.4;
SUBC > INTEGER 12 26.
           K P(X LESS OR = K)   K P(X LESS OR = K)
           17           0.4000   18           0.4667

```

```

MTB > INVCDF 0.25;
SUBC > UNIFORM 21 35.
           0.2500  24.5000

```

```

MTB > INVCDF 0.4;
SUBC > T 21.
           0.4000  -0.2566

```

```

MTB > INVCDF 0.7;
SUBC > CHISQ 40.
           0.7000  44.1649

```

```

MTB > INVCDF 0.8;
SUBC > F 13 27.
           0.8000  1.4517

```

### 3. Generación de números aleatorios

#### Comando RANDOM

Mediante este comando se pueden generar números aleatorios, de cualquiera de las distribuciones mencionadas anteriormente. Esto equivale a tomar una muestra aleatoria de una población que tiene una distribución particular. Se puede usar datos aleatorios para verificar y explorar los resultados de los teoremas y conceptos estadísticos.

Para usar el comando, debe indicarse el número de valores aleatorios a generar, la columna donde se guardan estos valores y la distribución (con los parámetros descritos anteriormente). Por ejemplo, para generar números aleatorios de una distribución binomial con parámetros  $n=12$  y  $\pi=0.7$ , se debe ejecutar:

```

MTB > RANDOM 60 INTO C1;
SUBC > BINOMIAL N=12 P=0.7.

```

Los valores aleatorios se colocan en la columna C1. El gráfico, para el conjunto de valores colocados en C1, se obtiene con el comando DOTPLOT, tal como se muestra a continuación.

```
MTB > DOTPLOT C1
```

El gráfico se proporciona en la figura 6.1.

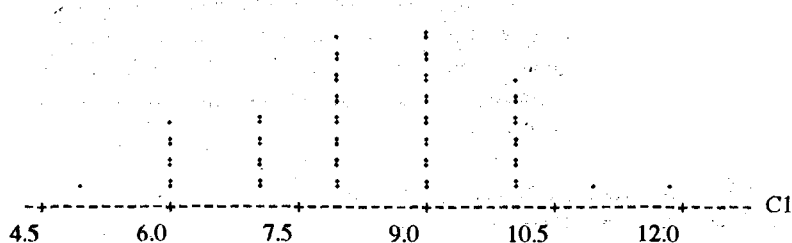


Figura 6.1

Para generar números aleatorios de una distribución F, con 5 grados de libertad para el numerador y 11 grados de libertad para el denominador, se debe ejecutar:

```
MTB > RANDOM 90 INTO C2;  
SUBC > F 5 11.
```

El gráfico respectivo, que se proporciona en la figura 6.2, se obtiene con:

```
MTB > DOTPLOT C2
```

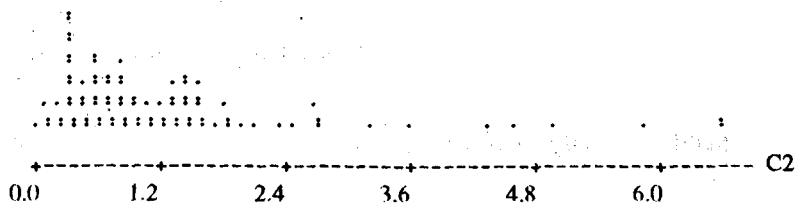


Figura 6.2

Como se muestra en los ejemplos, se ha utilizado el comando DOTPLOT para graficar las frecuencias de los números aleatorios generados. En el ejemplo, correspondiente a la distribución F; puede apreciarse un sesgo o asimetría a la derecha, tal como era de esperarse para esta distribución.

Aun cuando no es posible, por razones de espacio, presentar en este documento un caso práctico de la verificación del cumplimiento del Teorema del Límite Central, a continuación se muestra un ejemplo de los pasos que se pueden seguir para hacerlo. El primer paso consiste en obtener cierto número de muestras de una distribución. En este ejemplo se obtienen 10 muestras de tamaño 40, de una distribución uniforme continua con parámetros 8 y 21 (valor inicial y final del intervalo), donde los valores de cada muestra se guardan en las columnas C1 a C10.

```
MTB > RANDOM 40 IN C1-C10;
SUBC > UNIFORM 12 28.
```

Luego, se obtienen los promedios de cada muestra. Estos valores se guardan en la columna C12.

```
MTB > LET C12(1)=MEAN(C1)
MTB > LET C12(2)=MEAN(C2)
MTB > LET C12(3)=MEAN(C3)
MTB > LET C12(4)=MEAN(C4)
MTB > LET C12(5)=MEAN(C5)
MTB > LET C12(6)=MEAN(C6)
MTB > LET C12(7)=MEAN(C7)
MTB > LET C12(8)=MEAN(C8)
MTB > LET C12(9)=MEAN(C9)
MTB > LET C12(10)=MEAN(C10)
```

Como último paso, se pueden obtener estadísticos descriptivos de los datos generados.

```
MTB > DESCRIBE C1-C10 C12
```

	N	MEAN	MEDIAN	TRMEAN	STDEV	SEMEAN
C1	40	20.414	19.921	20.473	4.943	0.782
C2	40	20.502	20.823	20.576	4.431	0.701
C3	40	20.032	20.468	20.023	4.376	0.692

C4	40	19.915	19.326	19.928	4.923	0.778
C5	40	21.126	21.362	21.186	4.331	0.685
C6	40	18.969	18.925	18.856	4.610	0.729
C7	40	19.809	20.128	19.764	4.511	0.713
C8	40	18.906	18.131	18.860	4.362	0.690
C9	40	20.205	20.236	20.290	3.887	0.615
C10	40	20.502	20.394	20.584	3.825	0.605
C12	10	20.038	20.118	20.043	0.689	0.218

	MIN	MAX	Q1	Q3
C1	12.043	27.816	16.808	25.000
C2	12.094	27.060	17.264	23.730
C3	12.092	27.700	16.204	23.527
C4	12.177	27.455	16.020	24.037
C5	13.494	27.920	18.043	25.094
C6	12.072	27.863	14.069	22.120
C7	12.838	27.766	15.668	22.840
C8	12.054	26.692	15.086	23.398
C9	12.102	26.636	17.515	23.034
C10	12.967	26.312	17.122	23.727
C12	18.906	21.126	19.599	20.502

Se puede verificar que los resultados, para la columna C12, están próximos de los valores que deben obtenerse según el Teorema del Límite Central.

Para una verificación más real, debe obtenerse un mayor número de muestras. En este caso, los cálculos de promedios con el comando LET, como se mostró en el ejemplo anterior, pueden resultar tediosos; sin embargo, utilizando el comando OUTFILE, se puede generar un archivo que contenga los promedios de cada muestra. Luego, con la ayuda de un editor, se puede eliminar los títulos y, así, crear un archivo de datos para ser leído por Minitab, con el propósito de ejecutar los cálculos necesarios para completar la verificación del teorema.

### Comando BASE

Este comando se usa con el comando RANDOM y sirve para indicar, a Minitab, el valor inicial (semilla) que debe utilizarse para generar los números aleatorios. Si no se usa este comando, antes de ejecutar el comando RANDOM, Minitab elige automáticamente el valor inicial. Por ejemplo,



```
MTB > BASE 15
MTB > RANDOM 10 INTO C1;
SUBC > UNIFORM 8 21.
```

```
MTB > RANDOM 10 INTO C2;
SUBC > UNIFORM 8 21.
```

```
MTB > BASE 15
MTB > RANDOM 10 INTO C3;
SUBC > UNIFORM 8 21.
MTB > PRINT C1 C2 C3
```

ROW	C1	C2	C3
1	8.0180	8.1963	8.0180
2	10.2519	19.5336	10.2519
3	16.4922	19.7033	16.4922
4	16.5268	14.9139	16.5268
5	20.8535	14.2437	20.8535
6	15.6817	8.4647	15.6817
7	19.2152	14.0826	19.2152
8	18.9001	14.3198	18.9001
9	18.5118	17.9803	18.5118
10	8.9769	20.5415	8.9769

Como puede verse, las columnas C1 y C3 contienen los mismos números aleatorios. Esto se debe a que se usó el comando BASE, para especificar el valor inicial de 15, antes de pedir los números aleatorios con el comando RANDOM. Los valores de la columna C2 contienen números aleatorios para la misma distribución, generados con un valor inicial fijado por Minitab. Si se desea generar los mismos números aleatorios, se debe fijar el valor inicial o semilla del generador de números aleatorios con el comando BASE.

#### 4. Uso de una porción de un conjunto de datos

##### Comando SAMPLE

Este comando permite trabajar con una porción de los datos de una o más columnas. Es de utilidad cuando, al disponer de un número elevado de datos, se desea trabajar sólo con una muestra de ellos.

Para usar el comando se debe indicar el número de filas que se deben elegir al azar, en qué columnas se efectúa la elección y en qué columnas se guardan los valores elegidos. Por ejemplo, suponga que se tiene en memoria datos para las columnas C1 y C2, tal como se muestra a continuación con el comando INFO.

```
MTB > INFO
```

COLUMN	NAME	COUNT
	C1	60
	C2	90

CONSTANTS USED: NONE

Para tomar una muestra de 15 valores de la columna C2 y guardar estos datos en la columna C3, debe procederse de la siguiente manera:

```
MTB > SAMPLE 15 ROWS C2 STORE IN C3
```

```
MTB > INFO
```

COLUMN	NAME	COUNT
	C1	60
	C2	90
	C3	15

CONSTANTS USED: NONE

Como puede observarse, la columna C3 contiene sólo 15 datos. A esta muestra, se le puede aplicar otros comandos. Por ejemplo, se tiene:

```
MTB > DESCRIBE C2 C3
```

	N	MEAN	MEDIAN	TRMEAN	STDEV	SEMEAN
C2	90	1.294	0.904	1.122	1.233	0.130
C3	15	0.939	0.748	0.868	0.676	0.175
	MIN	MAX	Q1	Q3		
C2	0.013	6.030	0.508	1.525		
C3	0.161	2.641	0.357	1.482		

## Ejercicios

1. Ingresar a Minitab. Activar el comando **OUTFILE**, utilizando el nombre 'PRAC6', y ejecutar lo siguiente:

- a) Leer el archivo 'ENC9401.MTW'.
- b) Hallar los siguientes valores de probabilidad acumulada

Distribución	Parámetros	Probabilidad
Normal	$\mu=30, \sigma=3$	$P[X \leq 31]$
T	$g.l.=24$	$P[T \leq 0.5]$
Chi-cuadrado	$g.l.=21$	$P[X^2 \leq 38.5]$
Poisson	$\mu=3.4$	$P[X \leq 8]$

c) Hallar los siguientes percentiles:

Distribución	Parámetros	Percentil= $w$
Normal	$\mu=500, \sigma=25$	$P[X \leq w]=0.85$
T	$g.l.=14$	$P[T \leq w]=0.25$
Chi-cuadrado	$g.l.=28$	$P[X^2 \leq w]=0.45$

- d) Obtener 10 números aleatorios de una distribución normal estándar y el gráfico respectivo.
  - e) Verificar el Teorema del Límite Central con 40 muestras de tamaño 15 de una distribución Chi-cuadrado, de 20 grados de libertad.
2. Utilizando los comandos respectivos, determinar el valor del coeficiente de asimetría para las distribuciones Chi-cuadrado con 15 y 35 grados de libertad. ¿Cuál tiene mayor asimetría?
  3. Determinar el valor del coeficiente de curtosis para: una distribución normal con media 40 y variancia 36, una T de student con 56 grados de libertad y una T de student con 20 grados de libertad. Compare los valores hallados. ¿Qué conclusiones pueden obtenerse?

# VII

## Aplicaciones: inferencia estadística

### 1. Inferencia estadística para la media poblacional

En esta sección se describe el uso de comandos para efectuar procesos de inferencia, sobre una media poblacional; primero, considerando una muestra aleatoria proveniente de una población con variancia conocida; y segundo, cuando se tiene una muestra extraída de una población normal con variancia desconocida.

#### 1.1 Inferencia cuando se conoce la variancia poblacional

##### Comando ZINTERVAL

Mediante este comando es posible realizar la estimación por intervalo de confianza, para una media poblacional. Es necesario indicar el valor de confianza (probabilidad), el valor de la desviación estándar y las columnas que contienen los datos de la muestra. El reporte genera un intervalo para cada columna o variable. Si no se indica el valor de probabilidad, Minitab asume el valor 0.95 para obtener intervalos del 95% de confianza.

Por ejemplo, para obtener un intervalo del 90% de confianza para la media del "gasto familiar semanal en comida" (columna C10), asumiendo que la desviación estándar es 125, se debe ejecutar:

```
MTB > RETRIEVE 'ENC9401'
```

WORKSHEET SAVED 3/16/1994

MTB > ZINT 0.9 125 C10

THE ASSUMED SIGMA =125

	N	'MEAN	STDEV	SE MEAN	90.0 PERCENT C.I.
gastosem	50	510.0	134.5	17.7	(480.9, 539.1)

Este resultado indica que se tiene un 90% de confianza que la verdadera media, de los "gastos familiares semanales en comida", esté contenida en el intervalo comprendido entre 480.9 y 539.1 soles. Si en el comando se indica más de una columna, Minitab genera un intervalo para cada columna; sin embargo, debe tenerse presente que Minitab asumirá (con este comando) un valor igual de " $\sigma$ " para cada columna, lo cual no ocurre con frecuencia.

### Comando ZTEST

Este comando permite probar una hipótesis sobre una media poblacional. Es necesario indicar el valor, hipotético, de la media y el valor de la desviación estándar poblacional. Adicionalmente, mediante el subcomando ALTERNATIVE, se puede indicar el tipo de hipótesis a probar, tal como se indica a continuación.

**ALTERNATIVE=1** Para una prueba unilateral a la derecha; es decir, para probar  $H_p: \mu=k$  contra  $H_a: \mu>k$ .

**ALTERNATIVE=0** Para una prueba bilateral; es decir, para probar  $H_p: \mu=k$  contra  $H_a: \mu\neq k$ .

**ALTERNATIVE=-1** Para una prueba unilateral a la izquierda; es decir, para probar  $H_p: \mu=k$  contra  $H_a: \mu<k$ .

Por ejemplo, suponga que la empresa CREDITAR conoce, por experiencias anteriores, que los "gastos familiares semanales en comida" de sus clientes, tenían un comportamiento promedio de 450 nuevos soles y una desviación estándar de 125 nuevos soles. Asimismo, sospecha que el gasto promedio se ha incrementado, pero se sabe que la desviación estándar sigue siendo 125 soles. Para probar la hipótesis sobre la media, se debe ejecutar:

```
MTB > ZTEST MU=450 SIGMA=125 C10;
SUBC > ALTER =1.
```

```
TEST OF MU = 450.000 VS MU G.T. 450.000
THE ASSUMED SIGMA = 125
```

	N	MEAN	STDEV	SE MEAN	Z	P VALUE
gastosem	50	510.000	134.483	17.678	3.39	0.0004

Este resultado indica que en la columna C10 ('GASTOSEM') se tiene una muestra de tamaño 50, que el promedio muestral es 510 nuevos soles, que la desviación estándar muestral es 134.483 nuevos soles, que la desviación estándar poblacional de promedios es 17.678 nuevos soles, que el valor calculado de la prueba estadística es 3.39 y que el valor de probabilidad (P VALUE) es 0.0004. En general, se rechaza la hipótesis planteada, siempre que el valor P VALUE sea menor al nivel de significación " $\alpha$ " elegido. En este ejemplo, si supuestamente se hubiera establecido  $\alpha=0.01$ , la hipótesis  $H_0$  habría sido rechazada.

Por otro lado, puede apreciarse que Minitab presenta, como primer reporte, la hipótesis planteada  $MU = 450$  ( $\mu = 450$ ) y la hipótesis alternante,  $MU G.T. 450$  ( $\mu > 450$ ); es decir, se tiene una prueba "unilateral a la derecha", tal como se solicitó con el subcomando ALTERNATIVE.

También puede verse que el subcomando ALTERNATIVE se digitó como ALTER, pues debe recordarse que es válido utilizar los cuatro primeros caracteres de un comando o subcomando.

## 1.2 Inferencia cuando no se conoce la variancia poblacional

### Comando TINTERVAL

Este comando trabaja de manera similar al comando ZINTERVAL. Se utiliza cuando se desea estimar por intervalo la media poblacional, siendo desconocida la variancia poblacional. Si no se indica el valor de probabilidad, Minitab asume el valor 0.95 para obtener intervalos del 95% de confianza.

Por ejemplo, para obtener un intervalo del 90% de confianza, para la media del "gasto familiar semanal en comida" y la media del "gasto semanal en combustible" (columnas C10 y C13), se debe ejecutar:

MTB > TINT 0.9 C10 C13

El computador proporciona el reporte:

	N	MEAN	STDEV	SEMEAN	90.0 PERCENT C.I.
gastosem	50	510.0	134.5	19.0	(478.1, 541.9)
gastocom	50	48.22	18.59	2.63	(43.81, 52.63)

Este reporte, generado por Minitab, es similar al que se genera con el comando ZINTERVAL, con la diferencia que el valor SEMEAN se refiere a la desviación estándar muestral de promedios.

Como puede verse, en el comando TINTERVAL (también en ZINTERVAL) se puede especificar más de una columna, generándose un intervalo para cada columna indicada.

### Comando TTEST

Este comando trabaja de manera similar al comando ZTEST y se usa para realizar prueba de hipótesis sobre la media poblacional, cuando no se conoce la variancia poblacional. Con este comando también se puede utilizar el subcomando ALTERNATIVE, tal como se describió anteriormente para el comando ZTEST.

Por ejemplo, para probar la misma hipótesis presentada para la aplicación del comando ZTEST, se debe ejecutar:

```
MTB > TTEST MU=450 C10;
SUBC > ALTER=1.
```

El computador proporciona el siguiente reporte:

```
TEST OF MU = 450.000 VS MU G.T. 450.000

      N   MEAN   STDEV   SEMEAN   T   P VALUE
gastosem  50  510.000  134.483   19.019   3.15  0.0014
```

Como puede verse, el reporte generado es similar al obtenido con el comando ZTEST y en él, se indica que: el tamaño de muestra es 50, el promedio muestral es 510 nuevos soles, la desviación estándar muestral es 134.483 nuevos soles, la desviación estándar muestral de promedios es de 19.019

nuevos soles, el valor calculado de la prueba estadística es 3.15 y el valor de probabilidad (P VALUE) es 0.0014. Del mismo modo, se rechaza la hipótesis planteada, siempre que el valor P VALUE sea menor al nivel de significación " $\alpha$ " elegido.

También, al igual que con el comando ZTEST, Minitab presenta, como primer reporte, la hipótesis planteada  $MU = 450$  ( $\mu = 450$ ) y la hipótesis alternante  $MU > 450$  ( $\mu > 450$ ); es decir, se tiene una prueba "unilateral a la derecha", tal como se solicitó con el subcomando ALTERNATIVE.

## 2. Inferencia estadística para la diferencia de medias poblacionales

En esta sección se describe el uso de comandos para comparar las medias de dos poblaciones y para obtener intervalos de confianza, para la diferencia de medias de dos poblaciones.

Para determinar el procedimiento estadístico a usar y, por tanto, el comando Minitab necesario, se requiere poner atención a las siguientes interrogantes:

- ¿Son las muestras dependientes o independientes?
- ¿Las variancias poblacionales son iguales o diferentes?
- ¿Cómo están organizados los datos en la hoja de cálculo Minitab?

### 2.1 Inferencia cuando las variancias poblacionales son desconocidas y las muestras son independientes

#### Comando TWOSAMPLE, Comando TWOT

Estos comandos permiten efectuar pruebas de hipótesis para comparar las medias de dos poblaciones y obtener un intervalo de confianza para la diferencia de medias, sobre la base de dos muestras que son independientes. La diferencia entre el uso de TWOSAMPLE y TWOT, radica en la organización de los datos. Esto último se explicará más adelante.

Para apreciar el uso de estos comandos, supóngase que la empresa CREDITAR ha aplicado la misma encuesta a sus clientes de la ciudad de PIURA y



que, con la información de las dos encuestas, se ha creado el archivo ENC9403.MTW. En él, se ha colocado los datos de la encuesta de Lima, en las columnas C1 a C13, y los de la encuesta de Piura, en las columnas C21 a C33.

Para determinar qué procedimiento se debe usar, debe responderse a las tres preguntas indicadas anteriormente.

Respecto de la independencia, puesto que las encuestas han sido aplicadas en zonas geográficas diferentes, no hay razones para pensar que las respuestas a las encuestas en Lima, se vean afectadas por las respuestas obtenidas en las encuestas de Piura. Es decir, podemos considerar, de manera razonable, que las muestras son independientes. Los comandos Minitab TWOSAMPLE y TWOT son los indicados para muestras independientes.

En relación con la variabilidad, considerando que las variancias poblacionales son desconocidas, mediante la prueba F, para la comparación de dos variancias, se puede decidir si las variancias son iguales o diferentes. Minitab en su versión estudiante no presenta un procedimiento para prueba sobre dos variancias. Sin embargo, es posible hacerlo ejecutando las operaciones necesarias, utilizando los comandos y funciones adecuados.

Cuando las dos muestras tienen igual variabilidad, según la prueba F, se debe indicar a Minitab que calcule la variancia muestral ponderada de las dos muestras, para estimar la variancia común poblacional. En este caso, se debe usar el subcomando POOLED, en asociación a los comandos TWOSAMPLE y TWOT.

En cuanto a la organización de los datos, existen dos formas en las que Minitab maneja los datos:

- Los datos de las dos muestras están ubicados en columnas separadas. En estos casos, debe usarse el comando TWOSAMPLE.
- Los datos de las dos muestras están colocados en una misma columna (los datos de una muestra a continuación de los datos de la otra muestra), con otra columna conteniendo un índice que indica a qué muestra pertenece cada dato. En estos casos, debe usarse el comando TWOT.

**Ejemplo sobre el uso del comando TWOSAMPLE**

En el archivo ENC9403.MTW se tienen los datos, de dos muestras independientes, organizados en columnas separadas. Para tener este archivo en la hoja de cálculo se debe ejecutar:

```
MTB > retri 'enc9403'  
WORKSHEET SAVED 3/29/1994
```

Worksheet retrieved from file: enc9403.MTW

```
MTB > info
```

COLUMN	NAME	COUNT
C1	codigo1	50
C2	edadjf1	50
.	.	.
C11	ingreso1	50
C12	numvehi1	50
C13	gastoco1	50
.	.	.
C21	codigo2	40
C22	edadjf2	40
.	.	.
C31	ingreso2	40
C32	numvehi2	40
C33	gastoco2	40
.	.	.

```
CONSTANTS USED: NONE
```

Por razones de espacio, sólo se muestra las columnas de interés para este ejemplo. Puede verse que las columnas C1 a C13 contienen los datos de la muestra de Lima (tamaño de muestra 50), y las columnas C21 a C33 contienen los datos de la muestra de Piura (tamaño de muestra 40).

Para comparar el ingreso promedio de los clientes de Lima (columna C11) con el ingreso promedio de los clientes de la ciudad de Piura (columna 31), mediante el comando TWOSAMPLE, se puede hacer lo siguiente:

```
MTB > twosample t 90% confidence using 'ingreso1' 'ingreso2';
SUBC > alternative=0.
```

TWOSAMPLE T FOR ingreso1 VS ingreso2

	N	MEAN	STDEV	SE MEAN
ingreso1	50	4473	1015	144
ingreso2	40	3248	1143	181

90 PCT CI FOR MU ingreso1 - MU ingreso2: (841, 1610)

TTEST MU ingreso1 = MU ingreso2 (VS NE): T= 5.31 P=0.0000 DF=78

Como puede verse, luego del nombre del comando se indica: "t", para pedir una prueba de hipótesis; "90%", para pedir una estimación por intervalo para la diferencia de medias con 90% de confianza; y los nombres de las columnas que contienen los datos de las muestras. En el subcomando se indica "alternative=0", para señalar que es una prueba bilateral. En los resultados se muestran, primero, las estadísticas básicas para el ingreso, en ambas muestras; el intervalo del 90% de confianza, para la diferencia de las medias de los ingresos mensuales de los clientes de Lima y Piura; y los resultados de la prueba de hipótesis, donde se presenta una probabilidad  $P = 0.000$ , la cual será menor que cualquier " $\alpha$ " elegido. Por lo tanto, se debe rechazar la hipótesis planteada (ingresos promedios semejantes en ambas ciudades).

Por otro lado, como no se ha utilizado el subcomando POOLED, se está indicando a Minitab que las variancias son diferentes. En este caso, Minitab usa el estadístico:

$$T' = \frac{X_1 - X_2}{S_{\Delta X}}, \text{ donde}$$

$$S_{\Delta X} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

$S_1$  y  $n_1$  son, respectivamente, la desviación estándar y el tamaño de la muestra de la primera columna indicada; y  $S_2$  y  $n_2$ , son la desviación estándar y el tamaño de la muestra, respectivamente, de la segunda columna indicada. El estadístico  $T'$  tiene, aproximadamente, una distribución  $T$  con grados de libertad establecidos por la parte entera de la expresión:

$$gl = \frac{(S_1^2 + S_2^2)^2}{\frac{S_1^4}{n_1 - 1} + \frac{S_2^4}{n_2 - 1}}$$

Cuando se asume que existe homogeneidad de variancias (variancias iguales), debe utilizarse el subcomando POOLED. Por ejemplo, para hacer lo mismo que el ejemplo anterior, pero indicando que las variancias son iguales, debe ejecutarse:

```
MTB > twosample t 90% 'ingreso1' 'ingreso2';
SUBC > alter=0;
SUBC > pooled.
```

El computador proporciona el siguiente reporte:

TWOSAMPLE T FOR ingreso1 VS ingreso2

	N	MEAN	STDEV	SE MEAN
ingreso1	50	4473	1015	144
ingreso2	40	3248	1143	181

90 PCT CI FOR MU ingreso1 - MU ingreso2: (847, 1604)

TTEST MU ingreso1 = MU ingreso2 (VS NE):T= 5.38 P=0.0000 DF= 88

POOLED STDEV = 1074

Los resultados obtenidos son similares al obtenido sin el subcomando POOLED, con la diferencia que los cálculos efectuados se hacen considerando la variancia ponderada estimada. Como último valor reportado (1074), se presenta la desviación estándar ponderada estimada.

### Ejemplo sobre el uso el comando TWOT

El comando TWOT, a diferencia del comando TWOSAMPLE, se utiliza cuando los datos de las dos muestras están apilados en una sola columna. Es decir, los datos de las dos muestras, correspondientes a una misma variable, son ingresados en una misma columna.

En el archivo "ENC9404.MTW" se tienen almacenados los datos, de las muestras de Lima y Piura, en forma apilada. Al leer este archivo se aprecia:

```
MTB > retri 'enc9404'
      WORKSHEET SAVED 3/29/1994
```

```
Worksheet retrieved from file: enc9404.MTW
```

```
MTB > info
```

COLUMN	NAME	COUNT
C1	muestra	90
C2	edadjf	90
C3	edadco	90
C4	instjf	90
C5	instco	90
C6	ocupjf	90
C7	ocupco	90
C8	numper	90
C9	nhijoesc	90
C10	gastosem	90
C11	ingreso	90
C12	numvehi	90
C13	gastocom	90

```
CONSTANTS USED: NONE
```

Como puede verse, cada columna contiene 90 datos (50 de la encuesta de Lima y 40 de la encuesta de Piura). Además, la columna C1 ("muestra") contiene un índice para especificar a qué muestra pertenece cada fila de datos (1 para los datos de Lima y 2 para los datos de Piura).

Para comparar el ingreso promedio de los clientes de Lima con el ingreso promedio de los clientes de la ciudad de Piura, asumiendo que las variancias son diferentes, se utiliza el comando TWOT de la siguiente manera:

```
MTB > twot t 90% 'ingreso' groups in 'muestra';
SUBC > alter=0.
```

El computador produce el siguiente reporte:

TWOSAMPLE T FOR ingreso

muestra	N	MEAN	STDEV	SE MEAN
1	50	4473	1015	144
2	40	3248	1143	181

90 PCT CI FOR MU 1 - MU 2: (841, 1610)

TTEST MU 1 = MU 2 (VS NE): T= 5.31 P=0.0000 DF= 78

A la derecha del nombre del comando se indica: "t", para pedir una prueba de hipótesis; "90%", para pedir una estimación por intervalo con 90% de confianza; 'ingreso', para indicar la columna de datos con los cuales se hace la inferencia y 'muestra', para indicar la columna que contiene los índices que sirven para especificar a qué muestra pertenece cada fila de datos. En el subcomando se indica "alternative=0", para señalar que es una prueba bilateral. Como puede verse, los resultados son similares a los obtenidos con el comando TWOSAMPLE.

Para hacer la misma comparación, pero asumiendo que las variancias son iguales, se debe utilizar el comando POOLED. Por ejemplo,

```
MTB > twot t 90% 'ingreso' 'muestra';
SUBC > alter=0;
SUBC > pooled.
```

TWOSAMPLE T FOR ingreso

muestra	N	MEAN	STDEV	SE MEAN
1	50	4473	1015	144
2	40	3248	1143	181

90 PCT CI FOR MU 1 - MU 2: (847, 1604)

TTEST MU 1 = MU 2 (VS NE): T= 5.38 P=0.0000 DF= 88

POOLED STDEV = 1074

Estos resultados son similares a los obtenidos con el comando TWOSAMPLE y el subcomando POOLED, mostrados anteriormente.

## 2.2 Inferencia con dos muestras dependientes

### Comando TTEST, Comando TINTERVAL

Los comandos anteriores, TWOSAMPLE y TWOT, asumen que las muestras son independientes. Cuando los datos son apareados o son dependientes, se debe utilizar una prueba de T apareada, para probar la respectiva hipótesis sobre la media.

Por ejemplo, en los datos de la encuesta de Lima se captó información sobre el "ingreso familiar mensual" y el "gasto semanal en combustible". En consecuencia, los datos para estas variables, para cada unidad de muestreo, están "apareados"; por lo tanto, son dependientes porque el gasto en combustible depende del ingreso, entre otras variables.

Para este caso, suponga que resulta de interés probar si los "gastos familiares mensuales en combustible" son inferiores a la décima parte del "ingreso familiar mensual". Para probar esta hipótesis, primero, debemos tener datos mensuales comparables, lo cual se logra ejecutando lo siguiente:

```
MTB > retri 'enc9401'
      WORKSHEET SAVED 3/16/1994
```

Worksheet retrieved from file: enc9401.MTW

```
MTB > let c14=c13*4
MTB > name c14 'gastcome'
```

Mediante una transformación de datos se ha creado la columna C14, la cual contiene los "gastos mensuales en combustible". Ahora, si "los gastos mensuales en combustible" equivalen al 10% de los "ingresos mensuales", se

pueden obtener datos comparables, multiplicando por 10 a los datos de la columna C14 ("gasto mensual en combustible"); es decir, se debe ejecutar:

```
MTB > let c15=c14*10
```

Con esto se crea una nueva columna de datos, los cuales se han de comparar con los de la columna C11 (ingreso familiar mensual). Puesto que se trata de datos apareados, para compararlos, se genera una nueva columna de datos tomando la diferencia entre las columnas a comparar; es decir, se debe realizar lo siguiente:

```
MTB > let c16=c11-c15
MTB > name c16 'diferenc'
```

### Ejemplo sobre el uso del Comando TTEST

Con los datos de la columna C16 ("diferenc"), que contiene las diferencias entre los "ingresos familiares mensuales" y los "gastos familiares mensuales en combustible", se puede hacer la prueba de hipótesis correspondiente. Por último, si se supone que los "gastos en combustible" son menores al 10% de los "ingresos mensuales", es de esperar que la media de los ingresos sea superior a la media de los gastos mensuales en combustible multiplicado por 10 (columna C15); es decir, se tiene una prueba unilateral a la derecha. Por lo tanto, se debe ejecutar:

```
MTB > ttest mu=0 'diferenc';
SUBC > alter=+1.
```

El computador produce el siguiente reporte:

```
TEST OF MU = 0.000 VS MU G.T. 0.000
```

N	MEAN	STDEV	SE MEAN	T	P VALUE	diferenc
50	2544.200	647.091	91.513	27.80		0.0000

Como puede verse, el valor de probabilidad (P VALUE) es 0, menor a un " $\alpha$ " (nivel de significación elegido), lo cual indica que las evidencias muestrales confirman la sospecha de que los "gastos familiares mensuales en combustible" son inferiores a la décima parte de los "ingresos familiares mensuales" (se rechaza la hipótesis planteada).



### Ejemplo sobre el uso del comando TINTERVAL

Los datos de la columna C16, utilizados en el ejemplo para el comando TTEST, se usarán con el comando TINTERVAL, tal como se muestra a continuación.

```
MTB > tinterval 90% for 'diferenc'
```

N	MEAN	STDEV	SE MEAN	90.0 PERCENT C.I. diferenc
50	2544.2	647.1	91.5	(2390.7, 2697.7)

### 3. Análisis de datos categóricos

En esta sección, se describe el uso de comandos para el análisis de datos categóricos. Es decir, para datos, con un número relativamente pequeño de valores enteros, obtenidos mediante procesos de conteo en variables cualitativas o por agrupamiento de valores discretos o continuos en clases o categorías.

#### 3.1 Pruebas sobre frecuencias de k categorías

Minitab no brinda un comando para hacer esta prueba de una manera directa; por lo tanto, será necesario ejecutar los cálculos requeridos de una manera interactiva.

Por ejemplo, supóngase que se desea probar si el "número de hijos en edad escolar" (0, 1, 2, 3, 4), de los clientes de la ciudad de Lima, están en la relación 3:5:4:3:1. Para hacer esto, primero, se debe obtener las frecuencias observadas de la siguiente forma:

```
MTB > retri 'enc9401'
WORKSHEET SAVED 3/16/1994
```

```
Worksheet retrieved from file: enc9401.MTW
```

```
MTB > histo 'nhijoesc'
```

```
Histogram of nhijoesc N = 50
```

Midpoint	Count	
0	5	*****
1	16	*****
2	25	*****
3	3	***
4	1	*

Conociendo estas frecuencias observadas, se puede realizar la prueba indicada, ejecutando lo siguiente:

```

MTB > set in c14          # Ingreso de datos de la relación
DATA > 3 5 4 3 1
MTB > end
MTB > set in c15         # Ingreso de frecuencias observadas
DATA > 5 16 25 3 1
MTB > end

MTB > let c16=50*c14/16 # Cálculo de frecuencias esperadas

```

Este último comando es necesario para tener frecuencias comparables; es decir, para que la suma de las frecuencias esperadas y observadas sean iguales. A continuación, se debe hallar el valor calculado de Chi-cuadrado, con cada par de frecuencias (observada y esperada), tal como se muestra a continuación.

```

MTB > let c17=(c15-c16)**2/c16 # Cálculo de valores chi-cuadrado
MTB > let k1=sum(c17)          # Valor chi-cuadrado calculado
MTB > print k1
K1
20.3307
MTB > cdf k1 k2;              # Cálculo de  $P\{X^2 \leq 20.3307\}$ 
SUBC > chisquare 4.           #  $gl=5-1$ 
MTB > print k2
K2
0.999570
MTB > let k3=1-k2             # Cálculo de  $P\{X^2 > 20.3307\}$ 
MTB > print k3
K3
0.000429928

```

Si el nivel de significación elegido es 0.05, entonces se debe rechazar la hipótesis planteada, la cual expresa que las frecuencias observadas del

"número de hijos en edad escolar" no difieren, significativamente, de las frecuencias esperadas establecidas.

### 3.2 Pruebas de bondad de ajuste

Minitab tampoco brinda un comando para realizar la prueba de bondad de ajuste de una manera directa. Por lo tanto, se deben utilizar varios comandos y funciones para realizar los cálculos requeridos de una manera interactiva, tal como se hizo en el caso de la prueba de frecuencias para k categorías.

#### Aplicación a una distribución discreta

Por ejemplo, suponga que se desea probar si el "número de hijos en edad escolar", de los clientes de la ciudad de Lima, tienen una distribución que se ajusta (aproxima) a una distribución Binomial. Para hacer esto, primero, se debe obtener las frecuencias observadas, de la manera siguiente:

```
MTB > reti 'enc9401'
WORKSHEET SAVED 3/16/1994
```

```
Worksheet retrieved from file: enc9401.MTW
```

```
MTB > histo 'nhijoesc'
```

```
Histogram of nhijoesc N = 50
```

Midpoint	Count	
0	5	*****
1	16	*****
2	25	*****
3	3	***
4	1	*

Conociendo estas frecuencias observadas y que el tamaño de muestra es 50, se debe ejecutar:

```
MTB > set in c14 # ingreso de frecuencias observadas
DATA > 5 16 25 3 1
MTB > end
MTB > set in c15 # ingreso de valores de X
```

```

DATA > 0 1 2 3 4
MTB > end
MTB > let k1=mean('nhijoesc')
MTB > print k1
K1      1.58000
MTB > let k2=k1/4 # Cálculo de valor estimado de pi=P[éxito]
MTB > print k2
K2      0.395000

```

Para estimar la probabilidad de éxito, se ha dividido el promedio muestral entre el mayor valor observado de la variable (4), el cual será considerado igual al valor del parámetro "n" en la distribución Binomial.

A continuación, se realizará el cálculo de las probabilidades acumuladas.

```

MTB > cdf c15 c16; # Cálculo de probabilidades acumuladas de X
SUBC > binomial n=4 p=0.395.

```

Puesto que las probabilidades, generadas en la columna C16, son acumulativas, es necesario obtener, por diferencia, la probabilidad de ocurrencia de cada valor posible de la variable. Para ello, se genera una nueva columna (C17), tal como se muestra a continuación.

```

MTB > let c17(1)=c16(1) # Cálculo de P[X=0]
MTB > let c17(2)=c16(2)-c16(1) # Cálculo de P[X=1]
MTB > let c17(3)=c16(3)-c16(2) # Cálculo de P[X=2]
MTB > let c17(4)=c16(4)-c16(3) # Cálculo de P[X=3]
MTB > let c17(5)= 1 - c16(4) # Cálculo de P[X=4]

```

Con estas probabilidades se pueden obtener las frecuencias esperadas, multiplicando los valores de la columna C17 por 50 (tamaño de muestra); es decir, se debe ejecutar:

```

MTB > let c18=50*c17 # Cálculo de frecuencias esperadas
MTB > print c14-c18

```

ROW	C14	C15	C16	C17	C18
1	5	0	0.13397	0.133974	6.6987
2	16	1	0.48386	0.349883	17.4942

```

MTB > set in c14 # Ingreso de frecuencias observadas
DATA > 3 2 11 15 8 6 3 2
MTB > end

MTB > set in c15 # Ingreso de límites superiores de intervalos
DATA > 300 380 460 540 620 700 780 860
MTB > end

```

Puesto que no se conocen los valores de  $\mu$  y  $\sigma^2$  (parámetros de la función de densidad de la normal), es necesario estimarlos. Para ello, se ejecuta:

```

MTB > LET k1=mean('gastosem')
MTB > let k2=stdev('gastosem')
MTB > print k1 k2 # Valores estimados de  $\mu$  y  $\sigma$ 
K1      510.000
K2      134.483

```

A continuación, se realiza el cálculo de probabilidades acumuladas.

```

MTB > cdf c15 c16; # Cálculo de probabilidades acumuladas
SUBC> normal mean=510 deviation=134.483.

```

Las probabilidades acumuladas, para cada uno de los intervalos, fueron almacenadas en la columna C16. Con los valores de esta columna, se deben obtener las probabilidades no acumuladas de cada intervalo, mediante la respectiva diferencia de los valores de las probabilidades acumuladas de los intervalos. El último intervalo debe ser considerado como "desde a más de".

```

MTB > let c17(1)=c16(1) # Cálculo de  $P[X < 300]$ 
MTB > let c17(2)=c16(2)-c16(1) # Cálculo de  $P[300 < X < 380]$ 
MTB > let c17(3)=c16(3)-c16(2) # Cálculo de  $P[380 < X < 460]$ 
MTB > let c17(4)=c16(4)-c16(3) # Cálculo de  $P[460 < X < 540]$ 
MTB > let c17(5)=c16(5)-c16(4) # Cálculo de  $P[540 < X < 620]$ 
MTB > let c17(6)=c16(6)-c16(5) # Cálculo de  $P[620 < X < 700]$ 
MTB > let c17(7)=c16(7)-c16(6) # Cálculo de  $P[700 < X < 780]$ 
MTB > let c17(8)=1-c16(7) # Cálculo de  $P[X > 780]$ 

```

Para obtener las frecuencias esperadas, las probabilidades de la columna C17 se multiplican por 50 (tamaño de muestra).

```

MTB > let c18=50*c17 # Cálculo de frecuencias esperadas
MTB > print c14-c18

```

ROW	C14	C15	C16	C17	C18
1	3	300	0.059199	0.059199	2.9599
2	2	380	0.166856	0.107657	5.3829
3	11	460	0.355023	0.188167	9.4084
4	15	540	0.588262	0.233239	11.6619
5	8	620	0.793306	0.205044	10.2522
6	6	700	0.921145	0.127839	6.3919
7	3	780	0.977662	0.056517	2.8258
8	2	860	0.995373	0.022338	1.1169

Como las frecuencias esperadas de los intervalos 1, 7 y 8 son menores de 5, es necesario unir intervalos. El intervalo 1 se une con el intervalo 2 y los intervalos 7 y 8, se unen con el intervalo 6.

```

MTB > let c18(2)=c18(2)+c18(1) # suma de frec. esp. de 1 y 2
MTB > let c14(2)=c14(2)+c14(1) # suma de frec. obs. de 1 y 2
MTB > let c18(6)=c18(6)+c18(7)+c18(8) # suma frec esp 6,7 y 8
MTB > let c14(6)=c14(6)+c14(7)+c14(8) # suma frec obs 6,7 y 8
MTB > delete rows 7:8 c14 c18 # Eliminación de int. 7 y 8
MTB > delete rows 1 c14 c18 # Eliminación de intervalo 1
MTB > print c14-c18

```

ROW	C14	C15	C16	C17	C18
1	5	300	0.059199	0.059199	8.3428
2	11	380	0.166856	0.107657	9.4084
3	15	460	0.355023	0.188167	11.6619
4	8	540	0.588262	0.233239	10.2522
5	11	620	0.793306	0.205044	10.3346
6		700	0.921145	0.127839	
7		780	0.977662	0.056517	
8		860	0.995373	0.022338	

Ahora, se tienen 5 intervalos. Con los valores de la columna C14 (frecuencias observadas) y de la columna C18 (frecuencias esperadas), se puede obtener el valor calculado del estadístico Chi-cuadrado.

esperadas. Otras opciones para este subcomando son: "code 1", para solicitar sólo las frecuencias observadas y "code 3", para solicitar las frecuencias observadas, esperadas y la desviación relativa de las frecuencias observadas respecto de las esperadas.

A continuación, con los últimos valores del reporte, debe ejecutarse:

```

MTB > cdf 40.819 k1; # Cálculo de  $P[ X^2 \leq 40.819 ]$ 
SUBC > chisquare 21.
MTB > let k1=1-k1 # Cálculo de  $P[ X^2 > 40.819 ]$ 
MTB > print k1
K1 0.00590146

```

Para un valor  $\alpha = 0.05$ , por ejemplo, puesto que el valor de probabilidad encontrado (0.0059016) es menor a 0.05, debe rechazarse la hipótesis planteada que existe independencia entre las variables "ocupación del jefe de familia" y "grado de instrucción del jefe de familia".

### Comando CHISQUARE

Este comando permite efectuar pruebas de independencia y de homogeneidad, cuando se dispone de datos de las frecuencias observadas en un cuadro de doble entrada, el cual es suministrado por el usuario.

Utilizando las frecuencias observadas del ejemplo anterior, se procede a ejecutar,

```

MTB > read c1-c4
DATA > 0 8 0 0
DATA > 4 7 1 0
DATA > 0 8 5 1
DATA > 0 6 0 0
DATA > 0 3 1 0
DATA > 0 0 2 1
DATA > 0 0 2 0
DATA > 0 1 0 0
DATA > end

```

8 ROWS READ

```

MTB > chisquare c1-c4

```

El computador proporciona el siguiente reporte:

Expected counts are printed below observed counts

	C1	C2	C3	C4	Total
1	0 0.64	8 5.28	0 1.76	0 0.32	8
2	4 0.96	7 7.92	1 2.64	0 0.48	12
3	0 1.12	8 9.24	5 3.08	1 0.56	14
4	0 0.48	6 3.96	0 1.32	0 0.24	6
5	0 0.32	3 2.64	1 0.88	0 0.16	4
6	0 0.24	0 1.98	2 0.66	1 0.12	3
7	0 0.16	0 1.32	2 0.44	0 0.08	2
8	0 0.08	1 0.66	0 0.22	0 0.04	1
Total	4	33	11	2	50

$$\begin{aligned} \text{ChiSq} = & 0.640 + 1.401 + 1.760 + 0.320 + \\ & 9.627 + 0.107 + 1.019 + 0.480 + \\ & 1.120 + 0.166 + 1.197 + 0.346 + \\ & 0.480 + 1.051 + 1.320 + 0.240 + \\ & 0.320 + 0.049 + 0.016 + 0.160 + \\ & 0.240 + 1.980 + 2.721 + 6.453 + \\ & 0.160 + 1.320 + 5.531 + 0.080 + \\ & 0.080 + 0.175 + 0.220 + 0.040 = 40.819 \end{aligned}$$

df = 21

\* WARNING \* 20 cells with expected counts less than 1.0

\* Chisquare approximation probably invalid



29 cells with expected counts less than 5.0.

Como puede verse, para usar el comando CHISQUARE, solamente es necesario indicar las columnas que contienen a los datos.

A continuación, con los valores ChiSq y df del reporte, debe ejecutarse,

```
MTB > cdf 40.819 k1; # Cálculo de probabilidad acumulada
SUBC > chisquare 21.
MTB > let k1=1-k1 # Cálculo de  $P[ X^2 > 40.819 ]$ 
MTB > print k1
K1      0.00590146
```

Los resultados obtenidos, con los comandos TABLE y CHISQUARE, son iguales. La diferencia es que, éste último, genera avisos de advertencia cuando existen celdas con frecuencia observada 0 o con frecuencias esperadas menores de 5. En estos casos, una forma de lograr frecuencias más altas, en las celdas, puede ser eliminando clases o categorías para evitar una excesiva desagregación de los datos.

### Ejercicios

1. Ingresar a Minitab. Activar la opción de generación de un archivo de reportes, mediante el comando OUTFILE, utilizando el nombre 'PRACT6'. Luego, ejecutar lo siguiente:

a) Leer el archivo 'ENC9401.MTW'.

b) Si la desviación estándar de los ingresos es 1020 soles:

- Hallar un intervalo del 95% de confianza para el "ingreso mensual" promedio.
- Probar con  $\alpha = 0.05$  si el "ingreso mensual" promedio es mayor de 4000 soles.

c) Si la variancia de los ingresos es desconocida:

Hallar un intervalo del 90% de confianza para el "ingreso mensual" promedio.

- Probar con  $\alpha = 0.1$  si el "ingreso mensual" promedio es mayor de 4100 soles.
2. Ingresar a Minitab. Activar la opción de generación de un archivo de reportes, mediante el comando `OUTFILE`, utilizando el nombre 'PRAC7'. Luego, ejecutar lo siguiente:
- a) Con la información del archivo 'ENC9403.MTW', utilizando  $\alpha = 0.05$ , probar si el promedio del "gasto familiar semanal en combustible" de los clientes de la ciudad de Lima es diferente del promedio correspondiente a los clientes de la ciudad de Piura.
  - b) Con la información del archivo 'ENC9404.MTW', utilizando  $\alpha = 0.10$ , probar si el promedio del "gasto familiar semanal en combustible" de los clientes de la ciudad de Lima es mayor que el promedio correspondiente a los clientes de la ciudad de Piura.
  - c) Con la información del archivo 'ENC9401.MTW', considerando  $\alpha = 0.05$ , probar si el promedio del "gasto familiar semanal en comida" de los clientes de la ciudad de Lima es inferior al 60% de sus "ingresos familiares mensuales".
3. Ingresar a Minitab. Activar la opción de generación de un archivo de reportes, mediante el comando `OUTFILE`, utilizando el nombre 'PRAC71'. Luego, ejecutar lo siguiente:
- a) Leer el archivo 'ENC9403.MTW'.
  - b) Probar, con  $\alpha = 0.05$ , si la distribución del "número de personas por familia" de los clientes de la ciudad de Lima se ajusta a una distribución Binomial.
  - c) Probar, con  $\alpha = 0.05$ , si la distribución del "ingreso familiar mensual" de los clientes de la ciudad de Lima se ajusta a una distribución Normal.
  - d) Probar, con  $\alpha = 0.05$ , si la "ocupación principal del jefe de familia" es independiente de la "ocupación principal de la cónyuge".

# VIII

## Aplicaciones: análisis de correlación y regresión

### 1. Correlación

#### Comando CORRELATE

Este comando permite obtener los coeficientes de correlación de Pearson, entre dos o más variables. Para usar el comando, sólo se requiere indicar las variables (columnas) que intervienen en el análisis.

Por ejemplo, para obtener el coeficiente de correlación estimado entre el "ingreso familiar mensual" y el "gasto familiar semanal en comida" de los clientes de la ciudad de Lima debe ejecutarse:

```
MTB > retri 'enc9401'  
WORKSHEET SAVED 3/16/1994
```

```
Worksheet retrieved from file: enc9401.MTW
```

```
MTB > correlate 'ingreso' 'gastosem'
```

```
Correlation of ingreso and gastosem = 0.816
```

Este reporte indica una correlación positiva relativamente alta, entre el "ingreso familiar mensual" y el "gasto familiar semanal en comida", tal como era de esperarse.

Para pedir a Minitab las correlaciones, entre más de dos variables, sólo es necesario digitar o listar, luego del nombre del comando (CORRELATE), las columnas (o nombres de columnas) que contienen los datos. Aun, cuando algunas variables del archivo ENC9401.MTW no son continuas, a modo de ilustración del uso de este comando, se pedirá determinar los valores de los coeficientes de correlación para las variables: "número de personas que conforman la familia", "número de hijos en edad escolar", "gasto semanal en comida", "ingreso familiar mensual" y "número de vehículos".

```
MTB > correlate c8-c13
```

El computador proporciona el siguiente reporte:

	numper	nhijoesc	gastosem	ingreso	numvehi
nhijoesc	0.914				
gastosem	0.762	0.721			
ingreso	0.640	0.661	0.816		
numvehi	0.212	0.239	0.432	0.507	
gastocom	0.406	0.366	0.610	0.772	0.509

Como puede verse, el reporte es generado en forma matricial y muestra los coeficientes de correlación estimados, entre las variables (columnas) indicadas en el comando.

## 2. Regresión lineal simple y múltiple

### Comando REGRESS

Este comando permite ajustar una ecuación lineal, a los datos, para una variable dependiente y un conjunto de variables explicativas o independientes, mediante el método de mínimos cuadrados. Los argumentos que se requieren para usar el comando son: la variable dependiente, el número de variables independientes, las columnas que contienen los datos de las variables independientes y, opcionalmente, una columna para almacenar los **residuales** estandarizados y otra, para almacenar los valores ajustados para la **variable dependiente**.

A su vez, el comando REGRESS tiene subcomandos. Éstos son:

<b>NOCONSTANT</b>	Para indicar que en el modelo no se debe considerar el coeficiente de intercepción (regresión a través del origen).
<b>COEFFICIENTS in C</b>	Para indicar la columna donde se guardan los coeficientes de regresión estimados.
<b>PREDICT for E...E</b>	Para pedir valores estimados, sobre la base de la ecuación de regresión estimada. Para usar este subcomando, se debe indicar un argumento E (columna o constante), para cada variable independiente en el modelo. Se pueden usar varios subcomandos PREDICT, con un mismo comando REGRESS. El reporte que se genera, con este subcomando, incluye el valor ajustado o estimado y los intervalos del 95% de confianza, para el valor medio e individual.
<b>RESIDUALS into C</b>	Para pedir que los residuales se almacenen en la columna C.
<b>DW</b>	Para pedir el estadístico de Durbin-Watson de la prueba de autocorrelación.

## 2.1 Ejemplo sobre un análisis de regresión lineal simple

Para efectuar el análisis de regresión simple del "gasto familiar semanal en comida" (C10) sobre el "ingreso familiar mensual" (C11), y pedir que los residuales estandarizados se almacenen en la columna C15 y los valores ajustados en la columna C16, se debe ejecutar:

```
MTB > régress c10 1 c11 c15 c16
```

El computador proporciona el siguiente reporte:

```
The regression equation is
gastosem = 26.8 + 0.108 ingreso
```

Predictor	Coef	Stdev	t-ratio	p
Constant	26.80	50.73	0.53	0.600
ingreso	0.10803	0.01106	9.76	0.000

$s = 78.64$        $R\text{-sq} = 66.5\%$        $R\text{-sq(ajd)} = 65.8\%$

#### Analysis of Variance

SOURCE	DF	SS	MS	F	p
Regression	1	589393	589393	95.32	0.000
Error	48	296807	6183		
Total	49	886200			

#### Unusual Observations

Obs.	ingreso	gastosem	Fit	Stdev.Fit	Residual	St.Resid
34	6800	800.0	761.4	28.0	38.6	0.53 X

X denotes an obs. whose X value gives it large influence.

Como puede verse, el reporte generado con el comando REGRESS muestra:

- La línea de regresión estimada.
- Para cada coeficiente de regresión: su valor estimado, la desviación estándar estimada, el valor t calculado para probar  $H_0: \beta = 0$  contra  $H_a: \beta \neq 0$  y el complemento de la probabilidad acumulada para el valor t.
- La desviación estándar residual estimada ( $s =$  raíz cuadrada del cuadrado medio residual), el coeficiente de determinación ( $R\text{-sq}$ ) y el coeficiente de determinación ajustado ( $R\text{-sq(ajd)}$ ).
- El cuadro de análisis de variancia.
- Una tabla de observaciones inusuales; es decir, un dato fuera de lo normal o un punto de influencia.

En los resultados también puede verse que el valor de probabilidad, asociado al t para probar el intercepto (Constant), es igual a 0.6. Si el nivel de significación es 0.05, entonces, la conclusión para la prueba es aceptar la hipótesis de que el valor del coeficiente de intercepción es igual a cero.

## 2.2 Ejemplo sobre un análisis de regresión a través del origen

Para realizar esta regresión, y pedir que los residuales se almacenen en la columna C18 y que se genere el estadístico de Durbin-Watson, se debe ejecutar:

```
MTB > regress c10 1 c11 c15 c16;
SUBC > noconstant;
SUBC > residuals c18;
SUBC > dw.
```

El computador presenta el siguiente reporte:

The regression equation is  
 gastosem = 0.114 ingreso

Predictor	Coef	Stdev	t-ratio	p
Noconstant				
ingreso	0.113730	0.002408	47.23	0.000

s = 78.05

### Analysis of Variance

SOURCE	DF	SS	MS	F	p
Regression	1	13592668	13592668	2231.05	0.000
Error	49	298532	6092		
Total	50	13891200			

### Unusual Observations

Obs.	ingreso	gastosem	Fit	Stdev.Fit	Residual	St.Resid
12	3900	600.0	443.5	9.4	156.5	2.02R
38	5400	460.0	614.1	13.0	-154.1	-2.00R

R denotes an obs. with a large st. resid.

Durbin-Watson statistic = 1.53

```
MTB > print c15-c18
```

ROW	C15	C16	C17	C18
1	0.84393	284.324		65.676
2	0.90979	409.426		70.574
3	-0.24180	568.648		-18.648
4	1.91473	341.189		148.811

Por razones de espacio, no se muestra lo que sigue del reporte.

Al comparar estos resultados, con los anteriores, se puede ver, por ejemplo, que la desviación estándar estimada del coeficiente de regresión (b1) cambia de 0.01106 -en el modelo con intercepto- a 0.002408 -en el modelo sin intercepto-; es decir, la estimación del coeficiente de regresión se hace más confiable.

### 2.3 Ejemplo sobre un análisis de regresión lineal múltiple

Para hacer un estudio de regresión lineal múltiple, se procede de manera similar. Solamente se debe adicionar, en el comando inicial, todas las variables independientes que intervienen en el análisis.

Para estudiar la regresión lineal del "gasto familiar semanal en comida" sobre el "ingreso familiar mensual" y el "número de personas que conforman la familia" (2 variables independientes), para los clientes de la ciudad de Lima, se debe ejecutar: -

```
MTB > regress 'gastosem' 2 'ingreso' 'numper'
```

El computador presenta el siguiente reporte:

The regression equation is  
 $gastosem = -16.7 + 0.0735 \text{ ingreso} + 52.6 \text{ numper}$

Predictor	Coef	Stdev	t-ratio	p
Constant	-16.67	44.22	-0.38	0.708
ingreso	0.07352	0.01223	6.01	0.000
numper	52.61	11.93	4.41	0.000

s = 66.83

R-sq = 76.3%

R-sq(adj) = 75.3%



## Analysis of Variance

SOURCE	DF	SS	MS	F	p
Regression	2	676275	338137	75.71	0.000
Error	47	209925	4466		
Total	49	886200			

SOURCE	DF	SEQSS
ingreso	1	589393
numper	1	86881

## Unusual Observations

Obs.	ingreso	gastosem	Fit	Stdev.Fit	Residual	St.Resid
19	5100	560.00	26.55	35.53	-166.55	-2.94RX

R denotes an obs. with a large st. resid.

X denotes an obs. whose X value gives it large influence.

En este reporte puede apreciarse, después del cuadro de análisis de variancia, una tabla de suma de cuadrados secuenciales. Cada SEQ SS representa la reducción en la suma de cuadrados residual, que resulta de la adición de la respectiva variable independiente en el modelo de regresión, estando ya presentes las otras variables explicativas. Por ejemplo, el valor 86881 para "numper" viene a ser la reducción de la suma de cuadrados residual, debido a la adición de la variable "numper", estando ya presente en el modelo la variable "ingreso". Minitab establece este cuadro, de acuerdo con el orden establecido de las variables en el comando REGRESS. Si tiene interés en evaluar las contribuciones de cada variable explicativa, debe considerar el orden deseado en el comando REGRESS.

En cuanto a los resultados, comparándolos con los obtenidos en el primer ejemplo de regresión lineal simple, puede verse entre otras cosas, que el coeficiente de determinación se incrementa de 66.5% a 76.3%.

**Comando BRIEF**

Este comando puede usarse conjuntamente con el comando REGRESS, a fin de controlar la cantidad de información generada en el reporte. Para usarlo, debe indicarse uno de los tres niveles siguientes:

- 1 Para generar un reporte con información básica.
- 2 Para generar un reporte con información, usualmente, requerida. Este nivel es fijado por Minitab al iniciar una sesión.
- 3 Para obtener, además de la información reportada con el nivel 2, una tabla completa de valores ajustados y de residuales.

Al usar este comando (BRIEF), debe tenerse presente que permanece activado hasta que uno fija un nuevo nivel o usa los comandos STOP o RESTART (este comando permite iniciar una nueva sesión, sin necesidad de salir del programa Minitab).

El siguiente ejemplo muestra los reportes que se obtienen utilizando el nivel 3, para el mismo caso del ejemplo anterior.

```
MTB > brief 3
MTB > regress 'gastosem' 2 'ingreso' 'numper'
```

The regression equation is  
 gastosem = - 16.7 + 0.0735 ingreso + 52.6 numper

Predictor	Coef	Stdev	t-ratio	p
Constant	-16.67	44.22	-0.38	0.708
ingreso	0.07352	0.01223	6.01	0.000
numper	52.61	11.93	4.41	0.000

s = 66.83                  R-sq = 76.3%                  R-sq(adj) = 75.3%

#### Analysis of Variance

SOURCE	DF	SS	MS	F
Regression	2	676275	338137	75.71
Error	47	209925	4466	
Total	49	886200		

SOURCE	DF	SEQSS
ingreso	1	589393
numper	1	86881

Obs.	ingreso	gastosem.	Fit	Stdev.Fit	ResidualSt.	Resid
1	2500	350.00	272.35	21.55	77.65	1.23
2	3600	480.00	405.83	12.72	74.17	1.13
3	5000	550.00	561.37	10.75	-11.37	-0.17
4	3000	490.00	361.72	16.95	128.28	1.98
48	3950	390.00	431.57	11.76	-41.57-0.63	
49	4200	435.00	502.55	11.00	-67.55-1.02	
50	4700	625.00	591.92	16.23	33.08	0.51

R denotes an obs. with a large st. resid.

X denotes an obs. whose X value gives it large influence.

Por razón de espacio, no se presentan las líneas correspondientes a las observaciones desde la 5 hasta la 47.

Comparando este reporte con el anterior, se puede notar que ahora se muestra para cada observación: su valor estimado, su valor residual y sus respectivas desviaciones estándar estimadas.

### Comando STEPWISE

Con este comando, Minitab ofrece varios procedimientos de selección de variables. Todos los procedimientos tienen el mismo propósito: seleccionar de un grupo de variables independientes, aquellas que en conjunto constituyen las mejores variables explicativas. Para esto se ejecuta un procedimiento por pasos, en cada uno de los cuales se ingresa o remueve (elimina) una variable del modelo de regresión.

Para usar este comando deben indicarse como argumentos: la columna (o su nombre) que contiene los datos de la variable dependiente y las columnas (o sus nombres) que contienen los datos de las variables independientes.

Los subcomandos que se pueden usar son:

- FENTER = K** Para indicar el mínimo valor F, requerido para ingresar una variable independiente al modelo. El valor por defecto es 4. El valor de FENTER debe ser mayor o igual que el valor de REMOVE.
- FREMOVE = K** Para indicar el valor F, debajo del cual una variable independiente será removida o eliminada del modelo. El valor por defecto es 4.
- FORCE C...C** Para indicar las variables independientes que deben estar presentes en los subsiguientes modelos. Posteriormente, es posible modificar este estado con los subcomandos REMOVE o ENTER.
- ENTER C...C** Para indicar las variables independientes que deben ingresar en el modelo actual. Cualquiera de estas variables puede ser removida en los siguientes pasos, si su valor F es menor al valor de REMOVE.
- REMOVE C...C** Para indicar las variables independientes que deben ser removidas o eliminadas del modelo actual. Cualquiera de estas variables puede reingresar al modelo en los siguientes pasos, si su valor F es mayor al valor de FENTER.
- BEST K** Para indicar que, en cada paso en el cual una variable es adicionada al modelo, las siguientes k mejores alternativas de variables explicativas (en términos de su valor F) serán impresas. Cada valor calculado de la prueba estadística t, que es mostrado, se obtiene reemplazando la variable elegida por la variable alternativa. El valor de K, por defecto, es 0.
- STEPS = K** Para indicar el número de pasos reportados antes que aparezca en pantalla el mensaje MORE? El valor de K puede ser cualquier número, desde 1 hasta el máximo permitido por la pantalla.

Adicionalmente, con los comandos REGRESS y STEPWISE, se pueden usar los siguientes subcomandos:

- CONSTANT** Para indicar a Minitab que se deben estimar los parámetros del modelo de regresión con intercepto.
- NOCONSTANT** Para indicar a Minitab que se deben estimar los parámetros del modelo de regresión sin intercepto (regresiones a través del origen).

Con el comando **STEPWISE**, Minitab genera reportes por pantallas. Cada vez que la pantalla se llena de reportes aparece el mensaje **MORE?** El usuario puede responder **NO (N)**, para concluir o terminar el proceso, o puede responder **YES (Y)**, para ejecutar uno o más subcomandos indicados anteriormente.

Además, el comando **STEPWISE** permite ejecutar uno de los procedimientos siguientes:

#### **Regresión por pasos (Stepwise Regression)**

Este es el método por defecto. El método alterna pasos para adicionar y remover variables (independientes), hasta que ya no pueden entrar o salir variables del modelo. En un primer paso, Minitab calcula un valor **F** para cada variable, que ya figura en el modelo, y si existen valores menores al valor de **FREMOVE**, se remueve del modelo la variable con el menor valor **F**. En un segundo paso, Minitab calcula un valor **F** para cada variable, que no está presente en el modelo, y si existen valores mayores al valor de **FENTER**, se adiciona al modelo la variable con el mayor valor **F**.

#### **Selección hacia adelante (Forward Selection)**

Este método inicia su desarrollo desde un modelo sin variables y adiciona, en pasos sucesivos, nuevas variables, de manera similar al segundo paso descrito en el método anterior. Mediante este método, nunca se remueven o eliminan variables del modelo. Para utilizar este procedimiento debe indicarse **FREMOVE=0**.

#### **Selección hacia atrás (Backward Elimination)**

Este método inicia su desarrollo con un modelo que contiene todas las variables independientes posibles. Luego, remueve en pasos sucesivos una variable cada vez, de manera similar al primer paso descrito en el método de

selección por pasos. Mediante este método, una variable nunca reingresa al modelo. Para utilizar este procedimiento hacer **FENTER = 100000** y listar todas las variables independientes en el subcomando **ENTER**.

Para mostrar cómo funciona el comando **STEPWISE**, en sus diferentes procedimientos, se utilizarán los datos del archivo **ENC9401.MTW**. Se considera como variable dependiente al "gasto familiar semanal en comida" y como variables independientes al "número de personas que conforman la familia", "número de hijos en edad escolar", "ingreso familiar mensual", "número de vehículos de la familia" y "gasto familiar semanal en combustible".

### Ejemplos sobre regresión por pasos (Stepwise Regression)

```
MTB > stepwise c10, c8 c9 c11 c12 c13
```

```
STEPWISE REGRESSION OF gastosem ON 5 PREDICTORS, WITH
N = 50
```

STEP	1	2
CONSTANT	26.80	-16.67
ingreso	0.108	0.074
T-RATIO	9.76	6.01
numper		53
T-RATIO		4.41
S	78.6	66.8
R-SQ	66.51	76.31

```
MORE? (YES, NO, SUBCOMMAND, OR HELP)
SUBC> NO
```

Como puede verse, Minitab reporta el intercepto estimado, el coeficiente de regresión estimado y su valor estadístico T para cada variable independiente; y para cada paso (modelo), el valor de la desviación estándar residual (s) y su coeficiente de determinación.

El método elegido fue el de regresión por pasos y los resultados muestran que: en el primer paso, se ingresa al modelo la variable 'ingreso' obteniéndose un  $r^2$  de 66.51% y, en el segundo paso, se ingresa al modelo la variable 'numper' obteniéndose un  $r^2$  de 76.31%. Es decir, según estos resultados, se puede establecer el "gasto familiar semanal en comida" como una función del "ingreso familiar mensual" y del "número de personas que conforman la familia".

El siguiente ejemplo muestra el uso de los subcomandos FENTER, FREMOVE y ENTER.

```
MTB > stepwise c10, c8 c9 c11 c12 c13;
SUBC > fenter 1;
SUBC > fremove 1;
SUBC > enter c8 c9 c11 c12 c13.
```

STEPWISE REGRESSION OF gastosem ON 5 PREDICTORS, WITH  
N = 50

STEP	1	2	3
CONSTANT	-65.64	-65.25	-38.99
numper	67	67	55
T-RATIO	2.85	2.93	4.53
nhijoesc	-19	-19	
T-RATIO	-0.61	-0.64	
ingreso	0.068	0.068	0.066
T-RATIO	3.41	4.70	4.70
numvehi	44	44	44
T-RATIO	1.03	1.06	1.07
gastocom	-0.01		
T-RATIO	-0.01		
S	67.9	67.2	66.7
R-SQ	77.10	77.10	76.89

MORE? (YES, NO, SUBCOMMAND, OR HELP)  
SUBC> NO

En el primer paso están presentes todas las variables independientes (tal como se indicó con el subcomando ENTER); en el segundo paso, se remueve del modelo a la variable 'gastocom' y, en el tercer paso, se remueve a la variable 'nhijoesc'. Con relación al ejemplo anterior, se aprecia que el haber fijado FENTER=1 y FREMOVE=1, ha permitido que una variable más ('numvehi') permanezca en el modelo. La ecuación de regresión estimada, de acuerdo con el reporte, queda como:

$$\text{gastosem} = -38.99 + (55)(\text{numper}) + (0.066)(\text{ingreso}) + (44)(\text{numvehi}),$$

y el coeficiente de determinación, correspondiente, es 76.89%.

#### Ejemplo sobre el método de selección hacia adelante (Forward Selection)

MTB > stepwise c10, c8 c9 c11 c12 c13;  
SUBC > fremove 0.

STEPWISE REGRESSION OF gastosem ON 5 PREDICTORS, WITH  
N = 50

STEP	1	2
CONSTANT	26.80	-16.67
ingreso	0.108	0.074
T-RATIO	9.76	6.01
numper		53
T-RATIO		4.41
S	78.6	66.8
R-SQ	66.51	76.31

MORE? (YES, NO, SUBCOMMAND, OR HELP)  
SUBC> NO

Como puede observarse, los resultados obtenidos son similares a los del método de regresión por pasos.



**Ejemplo para el método de selección hacia atrás (Backward Elimination)**

```

MTB > stepwise c10, c8 c9 c11 c12 c13;
SUBC > fenter 100000;
SUBC > enter c8 c9 c11 c12 c13.

```

STEPWISE REGRESSION OF gastosem ON 5 PREDICTORS, WITH  
N = 50

STEP	1	2	3	4
CONSTANT	-65.64	-65.25	-38.99	-16.67
numper	67	67	55	53
T-RATIO	2.85	2.93	4.53	4.41
nhijoesc	-19	-19		
T-RATIO	-0.61	-0.64		
ingreso	0.068	0.068	0.066	0.074
T-RATIO	3.41	4.70	4.70	6.01
numvehi	44	44	44	
T-RATIO	1.03	1.06	1.07	
gastocom	-0.01			
T-RATIO	-0.01			
S	67.9	67.2	66.7	66.8
R-SQ	77.10	77.10	76.89	76.31

MORE? (YES, NO, SUBCOMMAND, OR HELP)

SUBC> NO

En el primer paso están presentes todas las variables independientes; luego, en los pasos 2, 3 y 4 se remueven del modelo las variables 'gastocom', 'nhijoesc' y 'numvehi', respectivamente. El resultado final es similar a los obtenidos con los métodos de regresión por pasos y de regresión hacia adelante.

## Ejercicios

1. Ingresar a Minitab. Activar la opción de generación de un archivo de reportes, mediante el comando `OUTFILE`, utilizando el nombre 'PRAC8'. Luego, ejecutar lo siguiente:
  - a) Leer el archivo 'ENC9402.MTW', el cual contiene la información muestral de los clientes de la ciudad de Piura.
  - b) Obtener los coeficientes de correlación estimados, entre las variables de las columnas C8, C10, C11 y C13.
  - c) Efectuar el análisis de regresión del "gasto familiar semanal en comida" sobre "ingreso familiar mensual", con  $\alpha = 0.05$ .
  - d) Efectuar el análisis de regresión del "gasto familiar semanal en comida" (C10) sobre las variables de las columnas C8, C9, C11 y C13, con  $\alpha = 0.05$ . Utilice el método de selección por pasos.
  - e) Efectuar el análisis de regresión del "gasto familiar semanal en comida" (C10) sobre las variables de las columnas C8, C9, C11 y C13, con  $\alpha = 0.05$ . Utilice el método de selección hacia adelante.
  - f) Realizar el análisis de regresión del "gasto familiar semanal en comida" (C10) sobre las variables de las columnas C8, C9, C11 y C13, con  $\alpha = 0.05$ . Utilice el método de selección hacia atrás.
2. Recuperar el archivo PRAC8.LIS, con el paquete WP 5.1. Luego, realizar un discusión sobre los resultados obtenidos en las subpreguntas d, e y f, de la pregunta anterior, y hacer la impresión del informe respectivo.

# IX

## Breve descripción del menú de Minitab

### 1. El menú principal

Al ingresar al paquete Minitab, en la pantalla del monitor aparece lo siguiente:

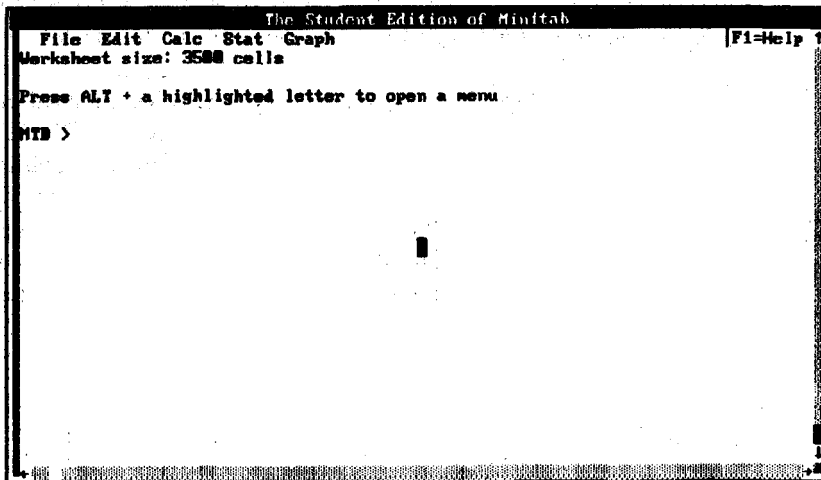


Figura 9.1

En la parte superior izquierda, presenta las opciones: File, Edit, Calc, Stat y Graph. Cualquiera de estas opciones puede ser elegida presionando,

simultáneamente, la tecla [Alt] y la tecla correspondiente a la primera letra del nombre de la opción. Así, por ejemplo, para elegir la opción Edit debe presionarse las teclas [Alt] y [E].

## 2. La opción File

Si se presionan, simultáneamente, las teclas [Alt] y [F] aparecerá en pantalla, en lado superior izquierdo, el menú de la opción FILE, tal como se muestra a continuación.

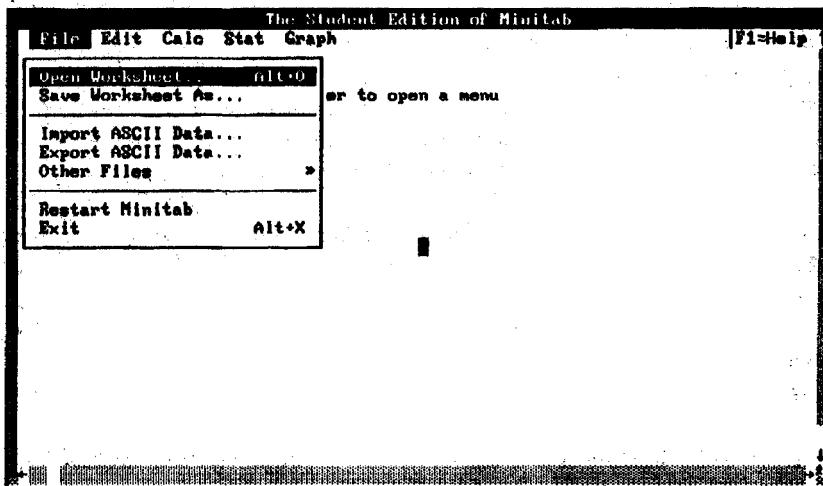


Figura 9.2

La opción FILE contiene los comandos necesarios para almacenar y recuperar archivos Minitab. Además, permite exportar e importar archivos en código ASCII. Su menú contiene las opciones Open Worksheet... , Save Worksheet As..., Import ASCII Data... , Export ASCII Data ... , Others Files , Restart Minitab y Exit.

### Open Worksheet ... (Alt+O)

Este comando permite recuperar archivos guardados previamente. Al ejecutarse esta opción, se producen tres alternativas: Minitab worksheet, Minitab portable worksheet y Lotus 1-2-3 worksheet.

La alternativa Minitab worksheet permite recuperar archivos Minitab; es decir, aquellos archivos almacenados con extensión MTW.

Minitab portable worksheet recupera un archivo Minitab, creado por Minitab, en un sistema de cómputo diferente y guardado usando un subcomando portable. Por ejemplo, permite transferir hojas de cálculo entre sistemas de cómputo de IBM PC a sistemas VAX.

La opción Lotus 1-2-3 worksheet, recupera hojas de cálculo en el formato de Lotus 1-2-3

### **Save Worksheet As...**

Esta opción permite guardar hojas de trabajo, incluyendo columnas, nombres de columnas, constantes, entre otros, dentro de un archivo. Tiene las alternativas Minitab worksheet, Minitab portable worksheet y Lotus 1-2-3.

La opción Minitab worksheet guarda las hojas de trabajo como archivo Minitab, es decir, con extensión MTW.

La alternativa Minitab portable worksheet guarda la hoja de trabajo en un formato llamado portable, que puede ser recuperado con otro sistema de cómputo.

El Lotus 1-2-3 worksheet es una alternativa que permite guardar archivos que pueden ser recuperados con Lotus 1-2-3

### **Import ASCII Data ...**

Esta opción del menú de FILE permite importar archivos que contienen la información en código ASCII (ver figura 9.2). Por defecto, al activarse "Import ASCII Data...", el computador selecciona "Replace existing data, if any". Este último, reemplaza los datos existentes en la hoja de cálculo del Minitab por los datos del archivo importado.

La alternativa "Insert before existing data", permite colocar los datos del archivo importado antes del conjunto de datos que se tiene en la hoja de cálculo del Minitab.

"Insert between rows [ ] and [ ]", es la alternativa que se usa para insertar datos importados, entre las filas que se indique, en la hoja de cálculo existente en la memoria.

"Insert after existing data", es la opción que permite ubicar los datos importados después de los datos almacenados en la hoja de cálculo del Minitab. Esta opción está asociada con el subcomando FORMAT. Este último sirve para especificar dónde y cómo aparecerán los datos en las líneas de datos de la hoja de cálculo de Minitab y se utiliza de manera similar a lo señalado en la explicación del uso del comando READ, para el ingreso de datos desde un archivo que tiene información en código ASCII.

### **Export ASCII Data...**

Es la opción que permite exportar datos en código ASCII. El archivo exportado puede ser editado, impreso o usado en otros paquetes de cómputo diferentes a Minitab. Por defecto, asigna la extensión DAT al archivo con datos en código ASCII. Tiene cuatro alternativas: Columns to export, Blank delimited y User-supplied format y Export to session window.

"Columns to export" es la alternativa que indica las columnas de datos que serán guardados en el archivo con extensión DAT. Si las columnas de datos son de desigual longitud, Minitab asigna el símbolo de valor perdido (\*) a los casilleros de las columnas que no tienen datos.

Por defecto, en la alternativa "Blank delimited", Minitab asigna espacios en blanco para separar columnas.

"User supplied format" es la alternativa asociada al subcomando FORMAT. Su uso es similar al señalado en la alternativa "Format", de la opción "Import ASCII Data".

"Export to session window" es la alternativa que se utiliza para tener los datos en la pantalla del monitor.

### **Others Files**

Esta opción tiene un menú con las alternativas: Start Recording Session, Start Recording History, Star Storing Macro y Execute macro.

La alternativa "Start Recording Session" tiene cuatro opciones. La primera, "Put output in file and session window", muestra por defecto la sesión Minitab, con los comandos utilizados y los resultados obtenidos y los graba en un archivo. La segunda, "Put output in file only", envía los comandos y sus resultados a un archivo, pero no a la sesión Minitab. La tercera, "Set output width to:", tiene por defecto 79 caracteres para la amplitud o ancho de una línea de pantalla, pero puede ser cambiado a cualquier valor en el rango de 30 a 132 caracteres. Sin embargo, algunos comandos no producen resultados con una longitud mayor a 70 caracteres. La cuarta, "Set output height to:", permite cambiar el número de líneas por página. Además, tiene la alternativa "Select File" que permite observar los archivos que se tiene y, adicionalmente, grabar lo establecido con alguna de las alternativas, antes mencionadas, en algún archivo con extensión LIS o grabarlo en un nuevo archivo, al cual el computador le asignará automáticamente la extensión LIS.

"Start Recording History" es la alternativa que guarda (en código ASCII) los comandos ejecutados, pero no los resultados. Minitab asigna al archivo la extensión MT.

"Start Storig Macro" es la alternativa que permite guardar los comandos más utilizados. Minitab asigna, al respectivo archivo, la extensión MTB.

"Execute Macro" es la alternativa que sirve para ejecutar un macro que fue previamente archivado. Al activarse esta alternativa, por defecto, el paquete establece el valor uno para el número de veces que se ejecutará el macro; sin embargo, esta cantidad puede cambiarse.

### **Restart Minitab**

Esta opción del menú de FILE, sirve para borrar los resultados de la sesión Minitab y dejar limpia la hoja de cálculo.

### **Exit**

Sirve para salir del paquete Minitab.

## **3. La opción Edit**

Esta opción, del menú principal, proporciona las alternativas necesarias para procesar los datos y ejecutar los comandos en una sesión de trabajo con

Minitab. Con esta opción se puede copiar, insertar y eliminar datos, hacer aparecer en pantalla los datos de la hoja de cálculo, generar muestra de datos, retornar a comandos ejecutados y guardar los colores preferidos para las pantallas de Minitab. Las alternativas que tiene son: Data screen, Cut, Copy Commands Only, Copy Entire Selection, Paste, Clear, Get Worksheet Info..., Display Data..., Set Patterned Data..., Edit Last Command Dialog y Set Colors. En la figura 9.3 se presenta el menú de opciones para Edit.

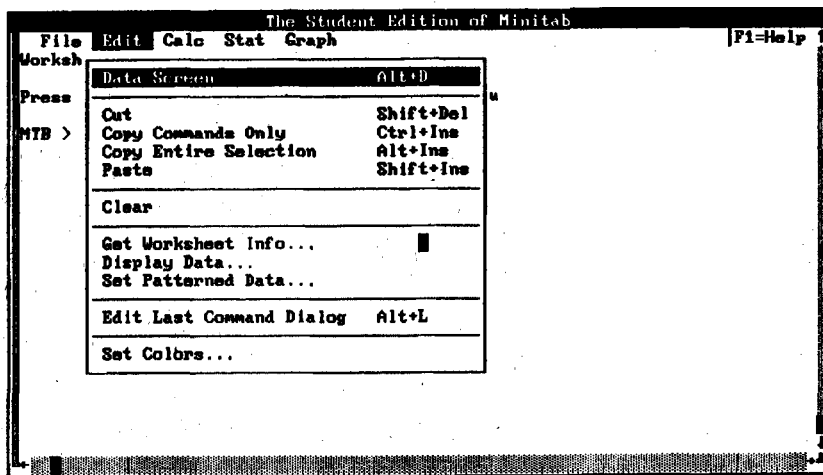


Figura 9.3

"Data screen" ([Alt]+[D]) permite exhibir en pantalla los datos que se encuentran en la sesión de trabajo. Presionando [Alt]+[M] desaparecen los datos de la pantalla.

"Cut"([Shift]+[Del]) se usa para eliminar el comando indicado después del prompt MTB>, que todavía no ha sido ejecutado.

"Copy Commands Only" ([Ctrl]+[Ins]) se usa para copiar únicamente comandos que se encuentran en la sesión de trabajo. Dichos comandos son guardados hasta que se seleccione Edit-Paste.

"Copy Entire Selection" ([Alt]+[Ins]) realiza el mismo trabajo que "Copy Commands Only" con la diferencia que, también, incluye resultados.



"Paste" sirve para insertar el contenido de lo guardado con "Copy Commands Only" o "Copy Entire Selection". Para usarlo, debe activarse el prompt MTB>.

"Clear" tiene un uso similar a "Cut".

"Get Worksheet Info..." permite observar en pantalla la información sobre las columnas de datos y las constantes.

"The Get Worksheet Info Dialog Box..." proporciona información sobre todas las columnas seleccionadas, nombres de las columnas, número de filas, número de valores perdidos y constantes guardadas.

"Display Data" proporciona en pantalla los datos de las columnas que se seleccionen y que se encuentran en la hoja de cálculo del Minitab.

"Set Patterned Data" permite ingresar datos repetidos en una columna.

"Edit Last Command Dialog" ([Alt]+[L]) permite que aparezca en pantalla la caja o relación de opciones de la última alternativa usada.

"Set colors" proporciona un conjunto de alternativas de colores para la pantalla del monitor.

#### 4. La opción Calc

Esta opción, del menú principal, contiene tres grupos de comandos: para distribuciones y datos aleatorios, transformaciones matemáticas y operaciones de trabajo en la hoja de cálculo. En la figura 9.4 se presenta el submenú de opciones para Calc, tal como aparece en la pantalla del monitor.

"Set Base" es un comando que fija el punto inicial para un generador de números aleatorios del Minitab.

"Random Data" permite seleccionar números aleatorios de distribuciones como Chi-cuadrado, Normal, F, T, Uniforme Continua, Bernoulli, Binomial, entre otros. También permite seleccionar números aleatorios de cualquier columna de la hoja de cálculo.

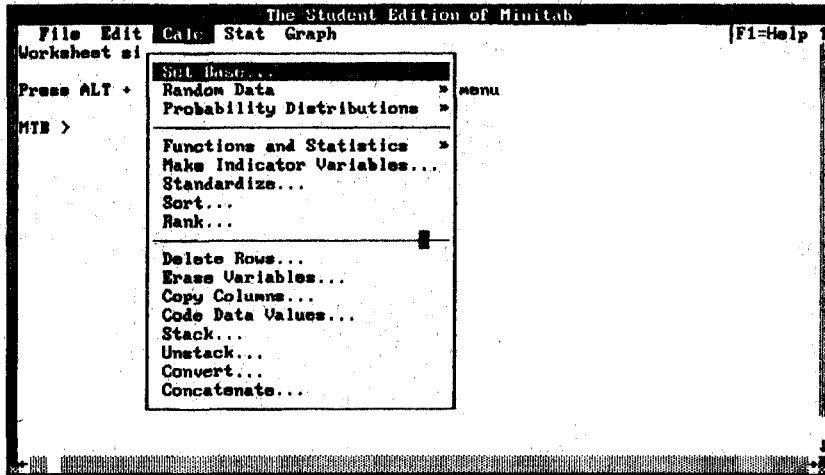


Figura 9.4

"Probability Distributions" permite calcular probabilidades, valores de las densidades de probabilidades, probabilidades acumuladas y percentiles de cualquiera de las distribuciones probabilísticas que se encuentran en el submenú.

"Functions and Statistics" se usa para realizar transformaciones matemáticas en columnas, filas, constantes guardadas y números. Los comandos respectivos se agrupan en: Expressions General (operadores aritméticos, operadores de comparaciones y operadores lógicos), Functions (Sqrt, Logt, Antilog, Sin, Cos, etc.), Column Statistics (Count, N, Mean, Stdev, etc.) y Row Statistics (Rcount, RN, Rmean, Rstdev, etc.).

"Standardize" es el comando que permite estandarizar datos de una variable, según las siguientes opciones: transforma datos de una columna restándoles su media, transforma datos de una columna dividiéndolos entre su desviación estándar, transforma datos de una columna restándoles cierto valor específico y dividiéndolos entre otro valor específico, realiza transformaciones de los datos de una columna en el rango que se señale y guarda los resultados de las transformaciones en la columna que se especifique.

"Sort" permite clasificar datos (numéricos o alfabéticos) por filas, de acuerdo con los valores de la columna que se indique. La clasificación se almacena en la columna que se especifique.

"Rank" es un comando que rankea los valores que se encuentran en la columna de interés. Asigna "1" al valor más pequeño, "2" al valor siguiente en orden de magnitud y así, sucesivamente. Además, permite almacenar los valores asignados en la columna que se indique.

"Delete Rows" permite eliminar filas para las columnas que se indique y reacomoda los datos de las columnas que han sido afectadas por la eliminación de filas.

"Erase variables..." es un comando que elimina cualquier conjunto de columnas que se encuentren en la hoja de cálculo.

"Copy columns..." copia todos los datos de una columna o sólo una porción de los datos de una columna a otra columna.

"Code Data Values" es un comando que permite codificar los datos de una columna, almacenando los códigos en otra columna. Así, por ejemplo, para los datos de la variable "edad del jefe de familia" almacenados en la columna 2 del archivo ENC9401.MTW, se puede indicar al computador asignar el código 1 al rango de edades 28-30 años, el código 2 al rango 31-35 años, el código 3 al rango 36-40 y el código 4 al rango 41-44 años. Debe tenerse presente que, la variable "edad del jefe de familia" presenta un valor mínimo igual a 28 y un valor máximo igual a 44.

"Stack" es un comando que permite formar bloques de datos con dos o más columnas. Estos bloques son colocados en una columna, uno a continuación de otro. A los datos de cada uno de los bloques, le asigna el mismo código; es decir, a los datos del primer bloque le asigna el código 1, a los datos del segundo bloque le asigna el código 2 y así, sucesivamente.

"Unstack" es un comando que realiza la operación contraria a la del comando "Stack". Forma bloques con los valores de una columna y coloca cada uno de los bloques en otras columnas.

"Convert" es un comando que convierte datos numéricos en alfabéticos y viceversa. Para realizar esta operación usa los códigos que se le especifique.

"Concatenate", es un comando que combina dos o más columnas alfabéticas, fila por fila. Estas combinaciones son ubicadas en las respectivas filas, de una nueva columna.

## 5. La opción Stat

Esta opción muestra un menú con los comandos necesarios para realizar análisis estadísticos. Contiene las alternativas: Basic Statistics, Regression, Analysis of Variance, Statistical Process Control, Time Series, Contingency Table Analysis y Nonparametric Statistics. El menú de la opción Stat se presenta en la figura 9.5, tal como aparece en la pantalla del monitor.

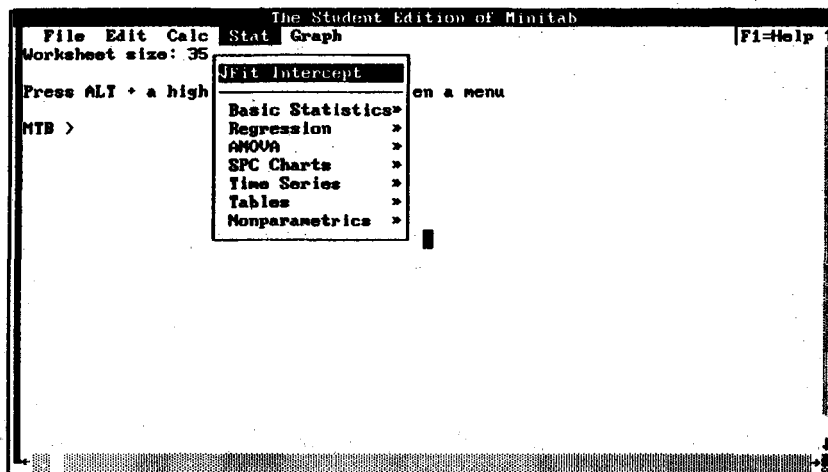


Figura 9.5

"Basic Statistics" contiene comandos que ejecutan cálculo de estadísticas, realizan inferencias utilizando Z o T y realizan cálculo de correlaciones y covariancias.

"Regression" es una alternativa que contiene comandos para dos opciones. La primera, es el ajuste a un modelo de regresión (simple, polinomial o múltiple). La segunda, corresponde al procedimiento de selección de variables, llamada Stepwise, con sus diversas alternativas (ver figura 9.6).

"ANOVA contiene comandos que sirven para ejecutar un análisis de variancia. Además, tiene opciones para efectuar pruebas de comparaciones múltiples.

"SPC charts" es una opción que tiene la capacidad de producir once clases diferentes de cartas de proceso de control estadístico de calidad.

"Time Series" contiene comandos que permiten investigar el comportamiento de los datos coleccionados a intervalos regulares de tiempo.

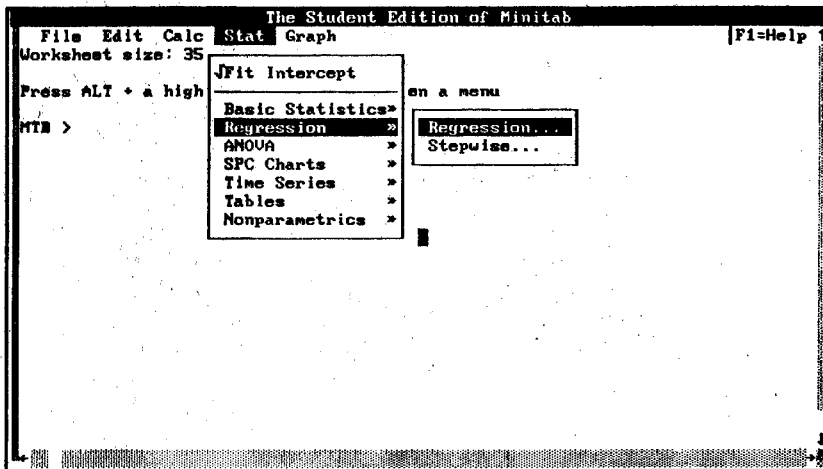


Figura 9.6

"Tables" es una opción que contiene un conjunto de comandos que permiten analizar datos categóricos. Tiene tres alternativas: "Cross Tabulation...", "Tally..." y "Chi-square Test..." (ver figura 9.7). La primera alternativa sirve para formar tablas de clasificación simple o múltiple con los datos de las variables de interés, para obtener valores de estadísticos básicos (media, mediana, etc.), y para calcular el valor de Chi-cuadrado para pruebas de independencia o de homogeneidad. La segunda alternativa proporciona información sobre frecuencia absoluta, frecuencia absoluta acumulada, porcentaje no acumulado y porcentaje acumulado para cada variable de interés. La tercera alternativa contiene comandos que permiten calcular el valor Chi-cuadrado, para prueba de independencia o de homogeneidad de una tabla ya formada.

"Nonparametrics" es una opción que permite analizar datos que no están distribuidos normalmente. Contiene comandos que permiten realizar la prueba del signo, Wilcoxon, Mann-Whitney, entre otros.

## 6. La opción Graph

Esta opción contiene cinco grupos de comandos. Éstos son: para establecer o fijar las opciones de gráficos (Set Options), gráficos para una columna, gráficos para varias columnas, gráficos de series de tiempo y gráficos seudo tridimensionales. Algunos de estos grupos, contienen subcomandos que permiten seleccionar colores, tipo de líneas, notas, leyendas y tamaño. La figura 9.8 presenta el menú de Graph, tal como aparece en la pantalla del monitor.

"Set Options..." permite especificar la altura y ancho de los gráficos correspondientes a series de tiempo, diagramas de dispersión, seudo tridimensionales y de control de calidad.

El grupo de comandos para gráficos de columnas simples corresponden a las alternativas Histogram, Ghistogram, Boxplot y Dotplot. Éstos fueron explicados en el capítulo V.

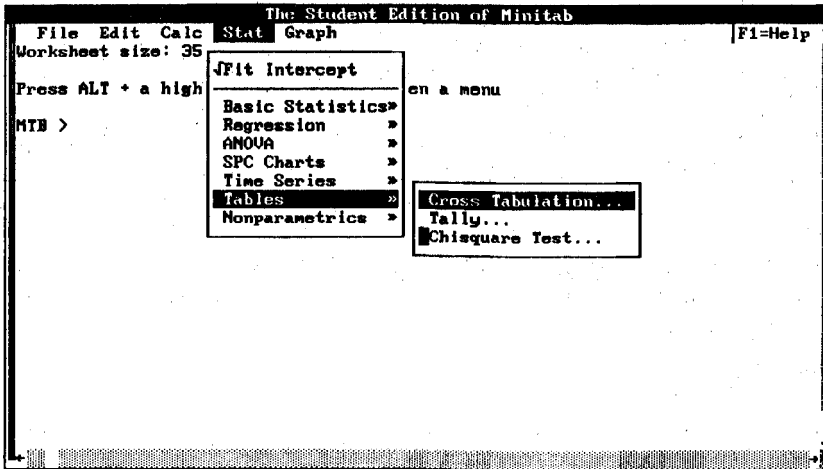


Figura 9.7

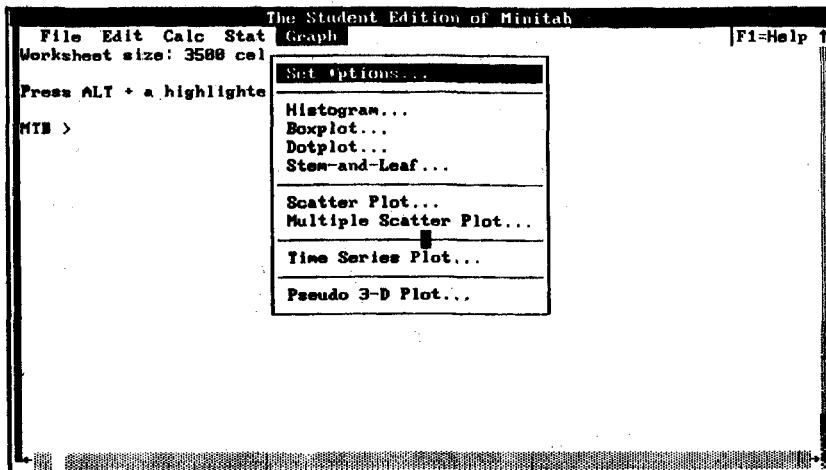


Figura 9.8

El grupo de comandos para gráficos, para varias columnas, corresponden a Stem-and-Leaf, Scatter Plot y Multiple Scatter Plot. Algunos de ellos, se explicaron en el capítulo V.

Con el conjunto de comandos, para Series de Tiempo, se pueden construir gráficos para una o más columnas (variables), cuyos datos corresponden a períodos regulares de tiempo. En el eje horizontal, el computador establecerá los valores enteros 1, 2, etc., los cuales corresponden a los períodos de tiempo; y en el eje vertical, fijará una escala adecuada (que puede ser modificada) para los valores de las columnas involucradas.

Para gráficos seudo tridimensionales, Minitab establece que a la primera columna le corresponde el eje Y, a la segunda columna le corresponde el eje X y a la tercera columna (que correspondería al eje Z) la representa por un conjunto especial de símbolos. Estos símbolos son:

- "X", si el valor de la variable correspondiente a la tercera columna es mayor que la media más una desviación estándar.
- "/", si el valor de la variable está en el intervalo [media, media + desviación estándar].

si el valor de la variable está en el intervalo [media-desviación estándar, media].

"0", si el valor de la variable es menor que la media menos una desviación estándar.

## 7. Ejemplos, utilizando el menú de Minitab

Se realiza una encuesta a 20 familias. Se obtuvo datos para las variables: "edad del jefe de familia" (C2), "número de miembros que tiene la familia" (C3), "número de hijos en edad escolar" (C4), "gasto semanal en comida" (C5), "ingreso mensual" (C6) y "gasto semanal en combustible" (C7). Los datos obtenidos se presentan a continuación.

Familia	C2	C3	C4	C5	C6	C7
1	28	2	0	350	1500	50
2	32	3	1	480	2600	30
3	44	4	2	550	2000	65
4	36	3	1	490	1600	40
5	42	4	2	600	1800	45
6	40	3	1	420	2200	55
7	40	4	2	480	2400	35
8	48	5	2	630	1800	30
9	45	3	1	500	2200	65
10	48	3	1	430	2400	40
11	35	2	0	250	1300	35
12	48	5	2	600	1600	25
13	36	4	2	550	1850	60
14	44	3	1	470	2100	30
15	50	4	2	525	2300	45
16	46	4	2	485	1900	50
17	42	3	1	410	2250	50
18	48	5	2	470	1350	35
19	52	7	4	560	2100	70
20	32	2	0	280	1450	30



## 7.1 Ingreso a la hoja de cálculo

Para tener la hoja de cálculo en la pantalla del monitor, presionar [Alt] y [M]; luego, seleccionar la opción [Data Screen] o presionar [Alt] y [D].

El ingreso de datos debe hacerse casillero por casillero. Después de ingresar un dato, debe presionarse [ENTER]. Al finalizar el ingreso de datos de una fila, ubicar el cursor en el inicio de la siguiente fila, mediante las teclas de flechas.

Si se desea modificar un dato, cuando aún no se ha presionado [Enter], utilizar la tecla [Backspace] para borrar el dato equivocado; luego, ingrese el dato correcto y presione [Enter].

Si se desea modificar un dato, después de presionar [Enter], colocar el cursor en el casillero correspondiente, usando las teclas de flechas. Presionar la barra espaciadora para borrar el dato equivocado y, luego, digitar el dato correcto y presionar [Enter].

Finalizado el ingreso de datos a la hoja de cálculo, presionar [F10], para ingresar a un menú de opciones y seleccionar la que le interese.

## 7.2 Asignar nombre a las columnas

Escoger la opción [Name Columns] o presionar [Alt] y [N]. Aparecerá en pantalla la hoja de cálculo con un indicador encima de C1. En el espacio, en el que se ubica el indicador, se iniciará el proceso de asignación de nombres a las columnas.

Dar el nombre a cada columna. No se debe utilizar más de 8 caracteres. Después de dar el nombre a una columna, presionar [Enter]. Para el ejemplo, dar los nombres: 'familia' a C1, 'edadif' a C2, 'numper' a C3, 'nhijoesc' a C4, 'gastosem' a C5, 'ingreso' a C6 y 'gastocom' a C7.

## 7.3 Obtener resumen de la información que contiene la hoja de cálculo

Para obtener un resumen de la información, de todas las columnas, efectuar el siguiente procedimiento:

- Presionar [F10] para tener el menú de opciones. Luego, presionar [Alt] y [M] para regresar a la sesión Minitab.
- Presionar [Alt] y [E]. Aparecerá el menú de opciones. Luego, seleccionar la opción [Get Worksheet Info...]. Aparecerá un panel con opciones.
- Por defecto (default), el computador marca la opción [Include all columns and constants]. Presionar [Tab]. El computador marcará la opción [≤ OK ≥]. Presionar la tecla [Enter]. En la pantalla aparecerá el resumen correspondiente y el prompt MTB>, tal como se muestra a continuación.

MTB > Info.

COLUMN	NAME	COUNT
C1	familia	20
C2	edadf	20
C3	numper	20
C4	nhijoesc	20
C5	gastosem	20
C6	ingreso	20
C7	gastocom	20

CONSTANTS USED: NONE

Si se desea obtener un resumen de la información de algunas columnas, digamos C2 y C6, se debe proceder de la siguiente manera:

En el panel de opciones de [Get Worksheet Info...] marcar la opción [Include Only]. Para lograrlo, considerando que por defecto (default) el computador tiene marcada la opción [Include all columns and constants], se debe presionar [Alt] e [I]. El computador marcará la opción [Include Only].

Presionar la tecla [Tab]. El cursor se ubicará en el recuadro en el que se deben digitar los nombres de las columnas de interés. Digitar C2 y C6, dejando un espacio en blanco entre ellos.

- Presionar la tecla [Tab]. El computador marcará la opción [≤ OK ≤].
- Presionar [Enter]. Aparecerá en pantalla el resumen solicitado y el prompt.MTB>, tal como se muestra a continuación.

MTB > Info c2 c6.

COLUMN	NAME	COUNT
C2	edadjf	20
C6	ingreso	20

#### 7.4 Construcción de histograma

Se construirá el histograma para la variable "ingreso mensual", cuyos datos se encuentran en la columna 6.

- Presionar las teclas [Alt] y [G]. En la pantalla aparecerá un menú de opciones de gráficos.
- Seleccionar la opción [Histogram]. Aparecerá un panel con opciones. En el panel, el cursor estará ubicado en la opción Variables.
- Digitar C6. Presionar dos veces la tecla [Tab]. El cursor se ubicará en la opción [First midpoint o marca de clase del primer intervalo], en él se puede asignar un valor. Para este ejemplo, no se asignará valor alguno; por lo tanto, presionar [Enter]. El cursor se ubicará en la opción [Last midponit o marca de clase del último intervalo]. Como no se desea asignar un valor, presionar [Enter]. Luego, el cursor se ubicará en la opción [Interval width o tamaño de intervalo de clase]. Tampoco se desea asignar un valor; entonces, presionar [Enter]. El cursor se ubicará en la opción [High Resolution]
- Presionar dos veces la tecla [Tab]. El cursor se ubicará en la opción [≤ OK ≥]. Presionar [Enter]. En la pantalla aparecerá el gráfico del histograma (ver figura 9.9).
- Para borrar de la pantalla el histograma, presionar la tecla [Q]. Aparecerá en pantalla el prompt MTB>.

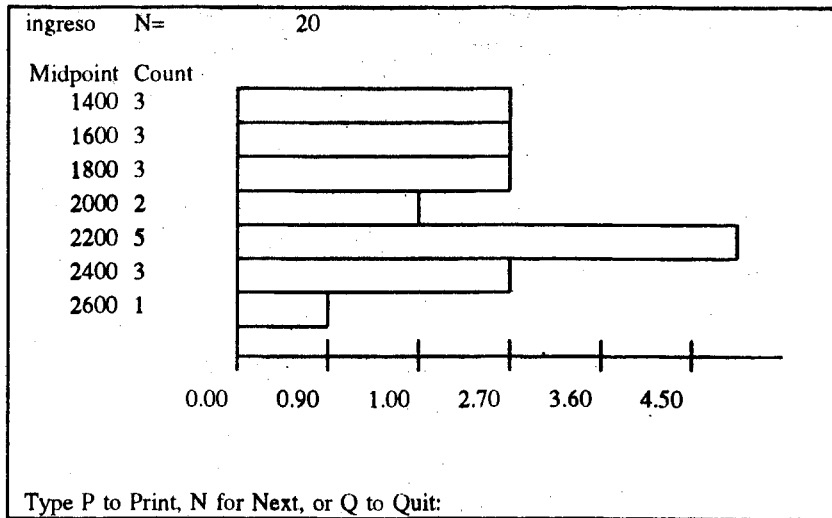


Figura 9.9

## 7.5 Cálculo de medidas estadísticas

Para obtener los valores de las medidas estadísticas, para las variables "edad del jefe de familia" (C2) e "ingreso" (C6), se debe ejecutar el siguiente procedimiento.

- Presionar las teclas [Alt] y [S]. Aparecerá en pantalla un menú de opciones.
- Seleccionar la opción [Basic Statistics]. Aparecerá otro menú de opciones.
- Seleccionar la opción [Descriptive Statistics]. Aparecerá en la pantalla un panel con dos recuadros pequeños en su interior. El cursor estará ubicado en el recuadro de variables. En este recuadro digitar C2 (o 'edadjf') y C6 (o 'ingreso'), separados por un espacio en blanco.
- Presionar la tecla [Tab] dos veces. El cursor se ubicará en la opción [≤ OK ≥]. Presionar [Enter]. En pantalla aparecerá, para las columnas referidas, los valores de N (número de observaciones), MEAN (media), MEDIAN (mediana), TRMEAN (media del 90% central de observacio-

nes), STDEV (desviación estándar), SEMEAN (desviación estándar de promedios), MIN (valor mínimo), MAX (valor máximo), Q1 (primer cuartil) y Q3 (tercer cuartil). En pantalla aparecerá lo siguiente:

MTB > Describe c2 c6.

	N	MEAN	MEDIAN	TRMEAN	STDEV	SEMEAN
edadjf	20	41.80	43.00	42.00	6.75	1.51
ingreso	20	1935.0	1950.0	1933.3	380.1	85.0

	MIN	MAX	Q1	Q3
edadjf	28.00	52.00	36.00	48.00
ingreso	1300.0	2600.0	1600.0	2237.5

## 7.6 Intervalo de confianza

Se desarrollará el procedimiento para estimar, con 98% de confianza, la media del "gasto mensual en combustible" (C7), considerando que la desviación estándar poblacional es igual a 20. Los pasos que se deben seguir son:

Presionar [Alt] y [S]. En pantalla aparecerá un menú de opciones.

Seleccionar la opción [Basic Statistics]. En pantalla aparecerá otro menú de opciones.

Seleccionar la opción [1. Sample Z]. Aparecerá un panel con opciones. El cursor está ubicado en el recuadro para Variables. En este recuadro, digitar C7 o 'gastocom'.

Presionar [Tab]. Por defecto, el computador marca la opción [Confidence interval].

Presionar [Tab]. Por defecto, el computador asignará un grado de confianza del 95%. Se desea una estimación con 98% de confianza; por lo tanto, digitar 98.

Presionar [Tab]. El computador resaltará la opción [desv. estándar]. Digitar el valor 20.

- Presionar [Tab]. El computador marcará la opción [ $\leq$  OK  $\geq$ ].
- Presionar [Enter]. Aparecerá en pantalla el siguiente reporte.

MTB > ZInterval 98 20 c7.

THE ASSUMED SIGMA =20.0

	N	MEAN	STDEV	SE MEAN	98.0 PERCENT C.I.
gastocom	20	44.25	13.50	4.47	(33.83, 54.67)

### 7.7 Prueba de hipótesis

Para la variable ingreso mensual (C6), se desarrollará el procedimiento para probar la hipótesis  $H_p: \mu = 2000$  contra  $H_a: \mu \neq 2000$ , suponiendo que la desviación estándar poblacional es igual a 100, y el nivel de significación 0.05. Los pasos, a seguir, son:

- Presionar [Alt] y [S]. En pantalla aparecerá un menú de opciones.
- Seleccionar la opción [Basic Statistics]. En pantalla aparecerá otro menú de opciones.
- Seleccionar la opción [1. Sample Z]. Aparecerá un panel con opciones. El cursor está ubicado en el recuadro para variables. En este recuadro digitar C6 o 'ingreso'.
- Presionar [Tab]. Por defecto, el computador marca la opción [Confidence interval]. Para pasar a la opción [Test], presionar [Alt] y [T]. El computador marcará la opción [Test].
- Presionar [Tab]. El cursor se ubicará en la zona para el valor hipotético de la media. Digitar 2000.
- Presionar [Tab]. El cursor se ubicará en la zona de la hipótesis alternante [Alternative]. Por defecto, el computador especifica la hipótesis alternativa correspondiente a una prueba bilateral [not equal]. Con la tecla de flecha hacia arriba [ $\uparrow$ ] puede seleccionarse una prueba unilateral-izquierda [less than]. Con la tecla [ $\downarrow$ ] puede seleccionarse una

prueba unilateral derecha [greater than]. Para el ejemplo, la prueba es bilateral; por lo tanto, no es necesario realizar cambio alguno.

- Presionar [Tab]. El cursor se ubicará en la zona de la desviación estándar poblacional. Digitar el valor 100.
- Presionar [Tab]. El cursor se ubicará en la opción [ $\leq$  OK  $\geq$ ].
- Presionar [Enter]. El computador proporcionará el siguiente reporte:

```
MTB > ZTest 2000 100 c6;
SUBC > Alternative 0.
```

```
TEST OF MU = 2000.000 VS MU N.E. 2000.000
THE ASSUMED SIGMA = 100
```

	N	MEAN	STDEV	SE MEAN	Z	P VALUE
ingreso	20	1935.000	380.132	22.361	-2.91	0.0037

## 7.8 Regresión lineal

En esta aplicación se considerará a la variable "gasto semanal en comida" (C5) como la dependiente. Las variables explicativas serán: "número de personas que conforman la familia" (C3), "número de hijos en edad escolar" (C4), "ingreso mensual" (C6) y "gasto mensual en combustible" (C7).

El procedimiento a seguir es el siguiente:

- Presionar [Alt] y [S]. Aparecerá en pantalla un menú de opciones.
- Seleccionar la opción Regression. Aparecerá en pantalla un menú con las opciones Regression y Stepwise. Seleccionar la opción Regression. Aparecerá en pantalla un panel, con el cursor ubicado en el recuadro [Response o variable dependiente].
- Digitar C3 o gastosem
- Presionar [Tab] para pasar al recuadro de [Predictors o variables explicativas]. Digitar, separados por un espacio en blanco, los nombres de las columnas o C3, C4, C6 y C7.

Presionar [Tab] para pasar a la opción [Residuals o residuales]. Como en este ejemplo no se desea tener los valores de los residuales, ni los valores correspondientes a las opciones siguientes, presionar tres veces [Tab]. El computador resaltará la alternativa [ $\leq$  OPTIONS $\geq$ ]. Esta alternativa corresponde, principalmente, al estadístico Durbin-Watson. Como en este ejemplo no interesa el estadístico antes mencionado, presionar [Tab]. El computador resaltará la opción [ $\leq$  OK  $\geq$ ]. Presionar [Enter]. En la pantalla del monitor aparecerá el siguiente reporte.

MTB > Regress c5 4 c3 c4 c6 c7.

The regression equation is

gastosem = 336 + 2.6 numper + 81.9 nhijoesc + 0.0302 ingreso - 1.04 gastocom

Predictor	Coef	Stdev	t-ratio	p
Constant	336.1	211.4	1.59	0.133
numper	2.60	63.23	0.04	0.968
nhijoesc	81.88	87.00	0.94	0.362
ingreso	0.03019	0.05137	0.59	0.565
gastocom	-1.045	1.431	-0.73	0.477

s=68.36

R-sq=63.3%

R-sq(adj)=53.5%

Analysis of Variance

SOURCE	DF	SS	MS	F	p
Regression	4	120712	30178	6.46	0.003
Error	15	70093	4673		
Total	19	190805			

SOURCE	DF	SEQSS
numper	1	107943
nhijoesc	1	7960
ingreso	1	2318
gastocom	1	2491

Unusual Observations

Obs.	numper	gastosem	Fit	Stdev.Fit	Residual	St.Resid
19	7.00	560.0	672.1	51.0	-112.1	-2.46R



R denotes an obs. with a large st. resid.  
 MTB > Retrieve 'EJEMEN1.MTW'.  
 WORKSHEET SAVED 2/17/1995

Worksheet retrieved from file: EJEMEN1.MTW  
 MTB > Regress c5 4 c3 c4 c6 c7 .

The regression equation is  
 $\text{gastosem} = 336 + 2.6 \cdot \text{numper} + 81.9 \cdot \text{nhijoesc} + 0.0302 \cdot \text{ingreso} - 1.04 \cdot \text{gastocom}$

Predictor	Coef	Stdev	t-ratio	p
Constant	336.1	211.4	1.59	0.133
numper	2.60	63.23	0.04	0.968
nhijoesc	81.88	87.00	0.94	0.362
ingreso	0.03019	0.05137	0.59	0.565
gastocom	-1.045	1.431	-0.73	0.477

s=68.36      R-sq=63.3%      R-sq(adj)=53.5%

#### Analysis of Variance

SOURCE	DF	SS	MS	F	p
Regression	4	120712	30178	6.46	0.003
Error	15	70093	4673		
Total	19	190805			

SOURCE	DF	SEQSS
numper	1	107943
nhijoesc	1	7960
ingreso	1	2318
gastocom	1	2491

#### Unusual Observations

Obs.	numper	gastosem	Fit	Stdev.Fit	Residual	St.Resid
19	7.00	560.0	672.1	51.0	-112.1	-2.46R

R denotes an obs. with a large st. resid.

# X

## Uso del Harvard Graphics

### Introducción

Las computadoras han revolucionado el mundo de las presentaciones gráficas. Con la computadora y el **Harvard Graphics** se pueden crear gráficos de alta resolución, en corto tiempo.

Los gráficos permiten transmitir y describir, rápidamente, no sólo las ideas sino también la información.

Los gráficos tienen las siguientes ventajas:

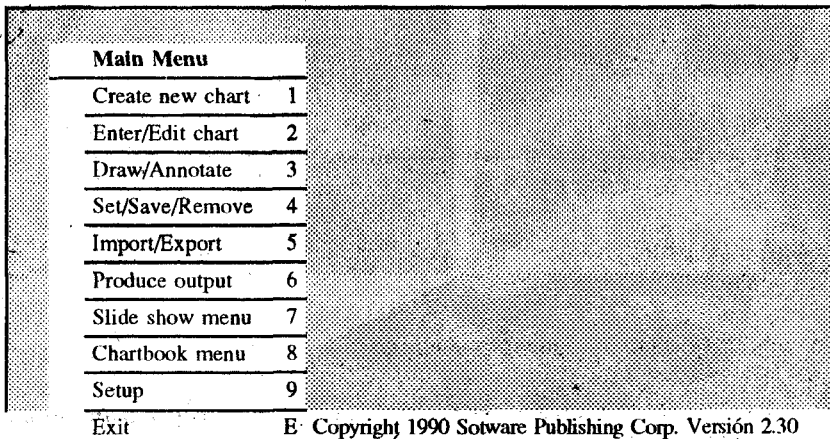
- Facilitan la captación y retención de la información por parte de un auditorio.
- Dinamizan las presentaciones
- Establecen la relación entre los datos e ideas, que de otra forma sería difícil explicar.
- Enfatizan ideas importantes.
- Proyectan una imagen profesional de la persona que la usa y de su organización.

La ventaja de utilizar **Harvard Graphics**, es que es un programa que posee características especiales, brinda excelentes opciones para crear diversidad de gráficos y, además, es fácil de usar.

## 1. Ingreso al Harvard Graphics

Si el directorio Harvard Graphics (HG) se encuentra en el disco duro, ingresar a dicho directorio. Digitar **HG**, luego pulsar **ENTER**. Aparecerá el **menú principal**. Éste contiene diez opciones: Create new chart (1), Enter/Edit Chart (2), Draw/Annotate (3), Get/Save/Remove (4), Import/export (5), Produce output (6), Slide show menu (7), Chart book menu (8), Setup (9) y Exit.

En la figura 10.1 se observa el **menú principal** del Harvard Graphics, tal como aparece en pantalla del monitor.



F1-Help                      F3-Applications  
 F2-Draw chart              F4-Spell check              F8-Options              ↵ Continue

**Figura 10.1** Menú principal del Harvard Graphics

### Ejemplo

En el capítulo I, se presentó la información proporcionada por la encuesta realizada por la cooperativa de crédito CREDITAR. Dicha información fue utilizada y procesada en los capítulos siguientes, mediante el uso de los comandos de MINITAB. Algunos de estos comandos son HISTOGRAM, GHISTOGRAM y TALLY. Con ellos, se obtienen resultados que permitirán

la construcción de gráficos mediante el paquete Harvard Graphics. Estos gráficos son de mayor calidad que los obtenidos con MINITAB.

## 2. Creación de un gráfico de sectores (Pie Chart)

Suponga que se tienen los siguientes resultados, para la variable "ocupación principal del jefe de familia":

OCUPACIÓN	No. DE JEFES DE FAMILIA
Industrial	8
Comerciante	12
Médico	6
Abogado	4
Asesor	3
Prof.Univ.	2
Otras	1

Con estos resultados se construirá un gráfico de sectores (Pie Chart), éste es uno de los gráficos más fáciles de crear y entender. Para ello, se debe efectuar los siguientes pasos:

1. En el menú principal, seleccionar la opción **Create new chart** (Crear un gráfico nuevo). Aparecerá en pantalla el menú de la opción **Create new chart**, tal como se observa en la figura 10.2.
2. Elegir la opción 2 (Pie). En la pantalla aparecerá una hoja para la introducción de datos del primer gráfico de sectores -[Pie Chart 1, Data Page 1 of 2]-, con el cursor posicionado en el campo Título -[Title]-, tal como se observa en la figura 10.3.
3. Escribir el título del gráfico (**OCUPACIÓN DEL JEFE DE FAMILIA**) y pulsar [Enter]. El cursor se ubicará en el campo de subtítulo. Digitar el nombre del subtítulo y presionar [Enter]. Si no se desea colocar nombre al subtítulo, presione [Enter]. El cursor se ubicará en el campo correspondiente a Pie de página [Footnote].

Main Menu		Create New Chart	
Create new chart		Text	1
Enter/Edit chart		Pie	2
Draw/Annotate		Bar/Line	3
Set/Save/Remove		Area	4
Import/Export		High/Low/Close	5
Produce output		Organization	6
Slide show menu		Multiple chart	7
Chartbook menu		From gallery	8
Setup		Clear value	9
Exit			

Publishing Corp. Versión 2.30

F1-Help

F3-Applications

F2-Draw chart

F4-Spell check

F8-Options

↵ Continue

Figura 10.2

4. Escribir **Cooperativa CREDITAR** en el campo para el Pie de página [Footnote] y pulsar [Enter], con lo cual el cursor se ubicará en la línea para la primera categoría ocupacional, bajo la columna Label Name -[Nombre de etiqueta]-.
5. Introducir los nombres de las categorías de la variable "ocupación principal de los jefes de familia". Éstas son: Industrial, Comerciante, Func. de Empresa, Médico, Abogado, Asesor, Prof. Univ. y otras. Debe pulsar [ENTER] después del ingreso del nombre de cada categoría.

## Pie Chart 1 Data Page 1 of 2

Title:

Subtitle:

Footnote:

Slice	Label Name	Value Series 1	Cut Slice		Color	Pattern
			Yes	No		
1			No		2	1
2			No		3	2
3			No		4	3
4			No		5	4
5			No		6	5
6			No		7	6
7			No		8	7
8			No		9	8
9			No		10	9
10			No		11	10
11			No		12	11
12			No		13	12

F1-Help

F3-Save

F9-More series

F2-Draw chart

F4-Draw/Annot

F6-Colors

F8-Options

F10-Continue

Figura 10.3

6. Trasladar el cursor a la segunda columna (Value), pulsando la tecla [Tab]. En esta columna, digitar los valores de la frecuencia correspondiente a cada categoría de la variable "ocupación". Presionar [Enter] después de digitar el valor de la frecuencia correspondiente.
7. Si le interesa construir otro gráfico de sectores, debe pasar a la segunda pantalla de introducción de datos -[Pie Chart 2 Data Page 2 of 2]-, esto se logra pulsando la tecla [Pg Dn o Av Pag]. En este ejemplo, no interesa un segundo gráfico de sectores.
8. Presionar [F8] para dar opciones al gráfico. Existen dos páginas para el conjunto de opciones. Con el cursor en la página 1, pulsar F7. Aparecerá, en la parte superior izquierda de la pantalla, las opciones para tamaño y lugar de título, footnote, etc., tal como se observa en la figura 10.4. Estas últimas opciones se desactivan pulsando la tecla [Esc].

9. Dar opciones al gráfico, según interese. Por ejemplo, si se quiere tener un gráfico tridimensional, ubique el cursor en 3D Effect y presione la barra espaciadora en Yes.
10. Pulsar [F2] para observar el gráfico en la pantalla del monitor. En esta publicación, el gráfico correspondiente se encuentra en la figura 10.5.

Pie Chart Titles & Options Page 1 of 2						
Size	Place					
8	L	R	Title:	OCUPACIÓN DEL JEFE DE FAMILIA		
6	L	C	R	Subtitle:		
6	L	C	R			
2.5	L	C	R	Footnote:	COOPERATIVA CREDITAR	
2.5	L	C	R			
2.5	L	C	R			
5				Pie 1 title:		
5						
5				Pie 2 title:		
5						
				3D effect	Yes	No
				Link pies	Yes	No
				Proportional pies	Yes	No
				Fill style	Color	Pattern Both
F1-Help			F5-Attributes		F7-Size/Place	
F2-Draw chart					F8-Data F10-Continue	

Figura 10.4

11. Para grabar el gráfico, pulsar [Esc] y luego [F10] para regresar al menú principal [Main Menu]-. Luego, elegir la opción 4 [Get/Save/Remove - Recuperar/Almacenar/Remover]. En seguida, seleccionar la opción Save chart-[Almacenar gráfico]-; entonces, aparece la caja de diálogo para almacenar gráficos -[Save Chart]-. El cursor comenzará a parpadear, indicándole que le debe dar un nombre de archivo al gráfico. Escribir el nombre conveniente (por ejemplo: **Graff1**) para el archivo y pulse [Enter]. El cursor avanzará hasta la línea donde Harvard Graphics ya ha colocado una descripción del gráfico para el usuario. Por definición (default), Harvard Graphics utiliza el título del gráfico como la descripción. Se puede cambiar la descripción del gráfico, escribiendo

sobre lo que allí aparece. Esta operación no modificará el título de su gráfico. Pulsar [Enter] para finalizar el proceso de grabación.

12. Pulsar [Esc] para regresar al menú principal, desde el cual se puede iniciar el proceso de recuperación y modificación del archivo.

### 3. Creación de un gráfico de barras (Bar Chart)

Se tienen los siguientes resultados para la variable "grado de instrucción"

GRADO DE INSTRUCCIÓN	No. DE JEFES DE FAMILIA	No. DE CÓNYUGES
Primaria	0	0
Secundaria	4	21
Superior	33	25
Maestría	11	3
Doctorado	2	1

Estos resultados se utilizarán para construir un gráfico de barras. La construcción de este gráfico es muy flexible. Se pueden crear gráficos de barras horizontales, verticales, apiladas y superpuestas.

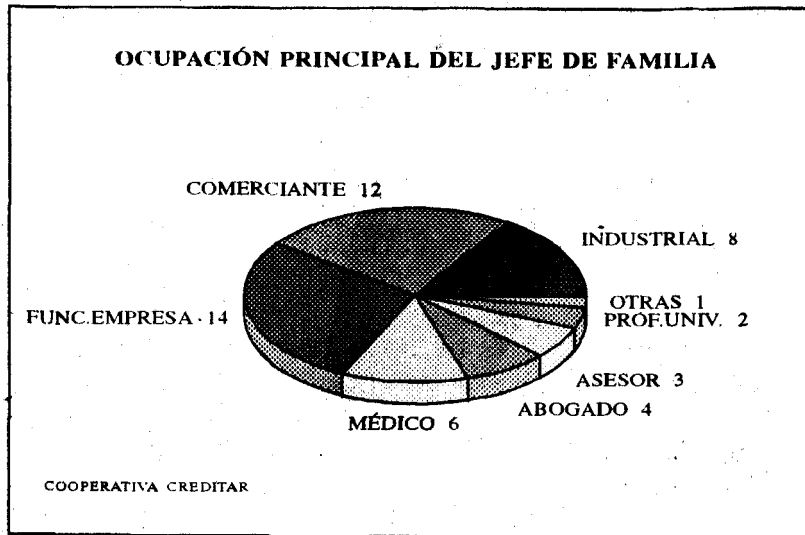
Los pasos para la construcción de los gráficos de barras, para los resultados dados, son:

1. Seleccionar la opción Create new chart [Crear un gráfico nuevo]; luego, elegir la opción 3 [Bar/Line-Barra/Línea] en el submenú Create new chart. Como ya se trabajó en la construcción de un gráfico (en este caso el gráfico Pie), aparecerá el siguiente mensaje:

Keep current data:                    Yes No  
 [Conservar los datos actuales]    Si No

Como se desea utilizar otro conjunto de datos, seleccionar No.





**Figura 10.5**

2. A continuación pulsar [Enter] 4 veces. El cursor debe ubicarse en la opción Title [Título].
3. En el campo de título escribir **Grado de instrucción** y presionar dos veces [Enter]. El cursor pasará al campo de footnote.
4. En el campo de footnote escribir **Cooperativa Creditar** y pulsar [Enter]. El cursor se ubicará en la primera fila de la columna [X Axis Name].
5. En la columna X Axis Name, escriba las categorías de la variable grado de instrucción: es decir: **Primaria, Secundaria, Superior, Maestría y Doctorado**. El nombre de cada categoría debe colocarse en una línea y después de terminar de escribir el nombre de la categoría se debe pulsar [Enter].
6. Con la tecla [Tab], desplazar el cursor a la columna Series 1. En ella, digitar las frecuencias ("No. de jefes de familia") para cada categoría (ver figura 10.6). Pulsar [Enter] después de digitar cada frecuencia.

Realizar la misma operación en la columna Series 2, para los valores de "No. de cónyuges".

### Bar/Line Chart Data

Title: Grado de instrucción

Subtitle:

Footnote: Cooperativa Creditar

Pt	X Axis Name	Series 1	Series 2	Series 3	Series 4
1	Primaria	0	0		
2	Secundaria	4	21		
3	Superior	33	25		
4	Maestría	11	3		
5	Doctorado	2	1		
6					
7					
8					
9					
10					
11					
12					

F1-Help

F3-Save

F5-Set X type

F9-More series

F2-Draw chart

F4-Draw/Annot

F6-Calculate

F8-Options

F10-Continue

Figura 10.6

7. Presionar la tecla [F8], para ingresar a la hoja de opciones para el gráfico de barras; luego, pulsar [F7]. En la parte superior izquierda de la pantalla aparecerá el cuadro de opciones de tamaño y lugar para el título, footnote, etc. Realizar los cambios necesarios y presionar [Esc], para salir de la opción de tamaño de letra y lugar.
8. Ubicar el cursor en el campo X Axis Title y escribir **Grado de Instrucción**. Luego, presionar [Enter].
9. Colocar el cursor en el campo Y1 Axis Title y escribir **# de socios**. Luego, presionar [Enter].
10. Ubicar el cursor en el campo Series 1 y escribir **Jefe de familia**; luego, colocar el cursor en el campo Series 2 y escribir **Cónyuge**.

11. Presionar [PgDn] y ubicarse en la opción **Bar enhancement**; a continuación, seleccionar la opción 3D presionando la barra espaciadora. Esto permitirá tener las barras en la forma tridimensional. Con la barra espaciadora puede seleccionar otra opción diferente a 3D.
12. Seleccionar las siguientes opciones presionando la barra espaciadora.
  - En Legend location, seleccionar Right.
  - En Legend justify, seleccionar Center.
  - En Legend placement, seleccionar In.
  - En Legend Frame, seleccionar Shadow.
13. Presionar [F2] para ver el gráfico en la pantalla del monitor. Debe aparecer el gráfico correspondiente, tal como el que se muestra en la figura 10.7.
14. Almacenar el gráfico con la opción Get/Save/Remove [Recuperar/Almacenar/ Remover] bajo el nombre **Graff2**.

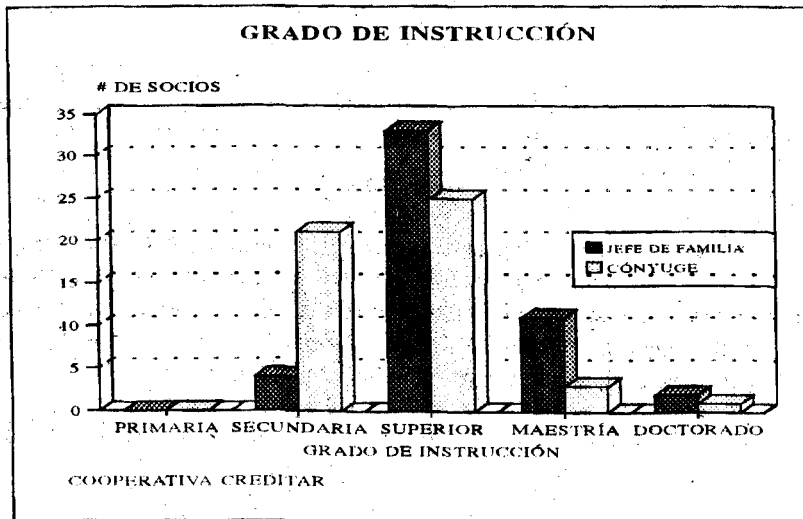


Figura 10.7

#### 4. Creación de histogramas y polígonos de frecuencias

Para explicar el procedimiento de construcción, se utilizará los resultados brindados por el Minitab para la variable "ingreso familiar mensual"

M.clase	# DE FAMILIAS
2500	1
3000	7
3500	2
4000	11
4500	10
5000	11
5500	3
6000	1
6500	3
7000	1

El procedimiento para la construcción del histograma y el polígono de frecuencias es el siguiente:

1. En el menú principal, seleccionar la opción **Create New Chart**; luego, en el submenú de **Create New Chart**, elegir la opción **Bar/Line**.
2. Presionar 6 veces [Enter]. El cursor se ubicará en el campo de **Footnote** (Pie de página).
3. En el campo de **Footnote**, escribir **Distribución de los Ingresos**.
4. Ubicar el cursor en la columna **X axis Name** y digitar los siguientes valores de marcas de clase: 2000, 2500, 3000, 3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000 y 7500. Como se observa, se han considerado los valores 2000 y 7500 que no figuran en la información dada por Minitab. El valor 2000 se considera como marca de clase de un intervalo anterior al primer intervalo, y el valor 7500, se considera como marca de clase del intervalo siguiente al intervalo nueve. Los referidos valores tienen frecuencia CERO (tal como se señala en el siguiente paso). Estos agregados se realizan con la finalidad de "cerrar" el polígono de frecuencias.

5. Colocar el cursor en la columna Series 1 y digitar: 0, 1/50, 7/50, 2/50, 11/50, 10/50, 11/50, 3/50, 1/50, 3/50, 1/50 y 0.
6. Ubicar el cursor en la columna Series 2 y digitar los mismos valores que se digitaron en la columna Series 1.
7. Presionar [F8] para ingresar a la hoja de opciones para el gráfico; luego, ubique el cursor en X axis Title y digite **Marca de clase**. A continuación, ubique el cursor en Y1 axis Title y escriba **Frecuencia relativa**.
8. Presionar [Enter] 3 veces, luego, reemplazar Series 1 por **Histograma** y Series 2 por **Polígono de frec.**.
9. Presionar [Tab] y seleccionar **Line** con la barra espaciadora.
10. Presionar [PgDn] para observar más opciones. Con la barra espaciadora, en **Bar Style**, seleccionar **Step**; en la opción **Bar Fill Style**, seleccionar **Pattern**.
11. Pulsar dos veces [PgDn], para ubicar el cursor debajo de **Marker/Pattern**; luego, digitar la opción 5. Esto permitirá que el histograma tenga cierto tramado.

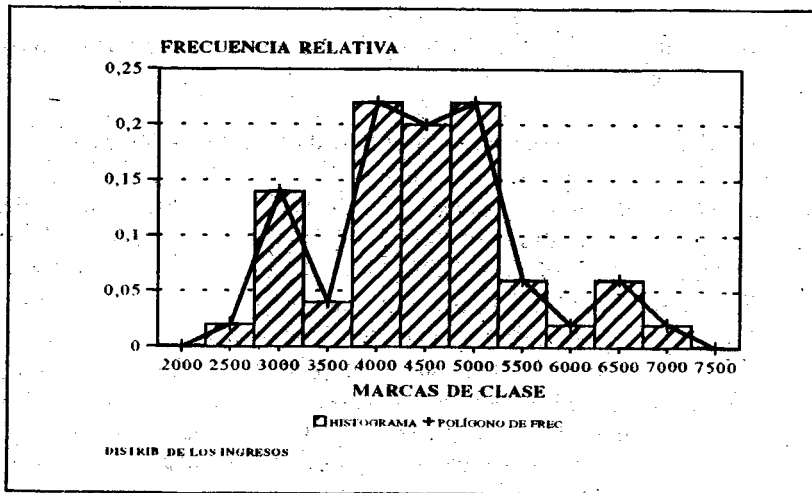


Figura 10.8

12. Presionar [F2]. Esto permitirá observar el gráfico en la pantalla del monitor, tal como el que se observa en la figura 10.8.
  13. Almacenar el gráfico con el nombre **Graff3**.
- ### 5. Creación de un gráfico de líneas

La explicación de la construcción de este tipo de gráfico se realizará utilizando los resultados proporcionados por Minitab para la variable "número de personas por familia". Estos resultados se dan a continuación:

No. DE PERSONAS	No. DE FAMILIAS
2	5
3	15
4	20
5	8
6	1
7	1

El procedimiento es el siguiente:

1. Elegir la opción **Create new chart** -[Crear un gráfico nuevo]-; luego, seleccionar la opción **Bar/Line** -[Barra/Línea]-; a continuación, pulse cuatro veces la tecla [Enter] para ubicar el cursor en la opción **Title** -[Título]-.
2. Escribir **Número de personas por familia** en el campo de título. Enseguida, presionar dos veces [Enter]. El cursor se ubicará en la opción **Footnote**.
3. Escribir **Cooperativa Creditar** en el campo de la opción **Footnote**. A continuación, presionar [Enter].
4. Ubicar el cursor en la columna **X Axis Name** y digitar los valores siguientes: 2, 3, 4, 5, 6 y 7. Debe presionar la tecla [Enter], después de digitar cada valor.

5. Con la tecla [Tab], colocar el cursor en la columna Series 1. Luego, digitar los siguientes valores: 5, 15, 20, 8, 1 y 1 (ver figura 10.6). Presionar [Enter], después de digitar cada valor.
6. Presionar la tecla [F8] para ingresar a la hoja de opciones del gráfico. Luego, pulsar [F7] para realizar los cambios necesarios en tamaño y lugar para las diferentes opciones. Presionar [Esc] al finalizar los cambios efectuados.
7. Ubicar el cursor en X Axis Title y escribir **Número de personas**, luego presione [Enter].

#### Bar/Line Chart Data

Title:       Número de personas por familia

Subtitle:

Footnote:   Cooperativa Creditar

Pt	X Axis Name	Series 1	Series 2	Series 3	Series 4
1	2	5			
2	3	15			
3	4	20			
4	5	8			
5	6	1			
6	7	1			
7					
8					
9					
10					
11					
12					

F1-Help

F3-Save

F5-Set X type

F9-More series

F2-Draw chart

F4-Draw/Annot

F6-Calculate

F8-Options

F10-Continue

**Figura 10.9**

8. Colocar el cursor en Y1 Axis Title y escribir **# de familias**, luego, presione [Enter].
9. Ubicar el cursor en Series 1 y escribir **Número de personas**.

10. Presionar [PgDn] y ubicar el cursor en Bar Width; a continuación, escriba el valor 1.
11. Con la barra espaciadora seleccionar:
  - En Legend location, la opción Right.
  - En Legend justify, la opción Center.
  - En Legend placement, la opción In.
  - En Legend frame, la opción Shadow.
12. Presionar [F2]. En la pantalla del monitor aparecerá el gráfico correspondiente, tal como el que se observa en la figura 10.10.
13. Almacenar el gráfico con la opción [Get/Save/Remove, Recuperar/Almacenar/ Remover] con el nombre **Graff4**.

## 6. Recuperación de un gráfico

Para recuperar un gráfico, previamente almacenado, se deben ejecutar las siguientes instrucciones:

1. En el menú principal (Main Menu), seleccionar Get/Save/Remove. Este es el mismo menú que utilizó para almacenar el archivo anterior.
2. En el submenú de Get/Save/Remove, elegir la opción **Get chart**. Aparecerá un directorio mostrando los nombres de los archivos almacenados en el disco. Los nombres de los archivos de Harvard Graphics tienen la extensión **CHT**, el cual se agrega en forma automática al realizar el proceso de grabación.
3. Con la tecla [Flecha hacia abajo] o [↓] seleccionar el archivo de interés y presionar [Enter] o [F10]. El computador mostrará el gráfico en el monitor.
4. Para regresar al menú principal, pulsar dos veces la tecla [Esc].

## 7. Impresión de gráficos

Para realizar la impresión de un gráfico, que se encuentra en la memoria RAM del computador, se debe ejecutar el siguiente procedimiento:



1. En el menú principal, seleccionar la opción **Produce Output** (Producir salida). Aparecerá en pantalla el submenú de **Produce Output**. Luego, elegir la opción **Printer** (Impresora), que permitirá que en pantalla aparezca un recuadro con el menú para imprimir gráficos (ver figura 10.11).
2. En dicho recuadro ubicar (con la barra espaciadora) la alternativa correspondiente para cada una de las opciones. Después de seleccionar la alternativa, de una opción, debe presionar [Enter].

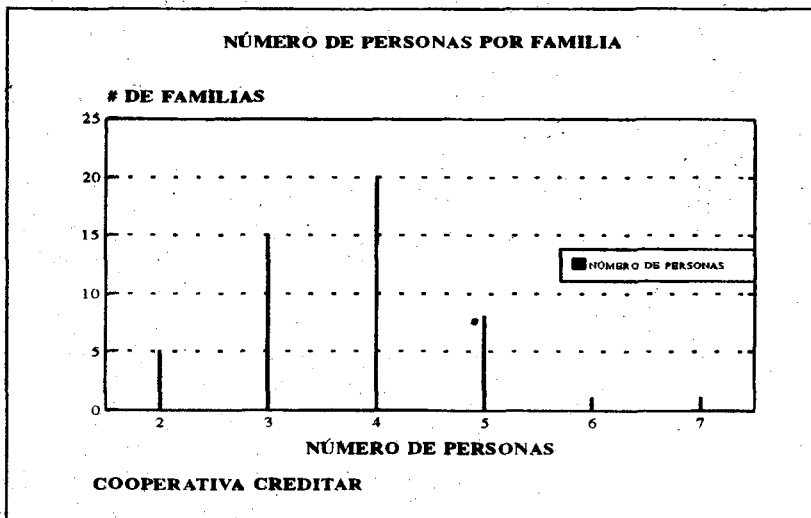


Figura 10.10

3. El computador iniciará la impresión del gráfico, después de presionar [Enter], para la alternativa seleccionada en la opción Número de copias.

### Breve descripción de las opciones de impresión de gráficos

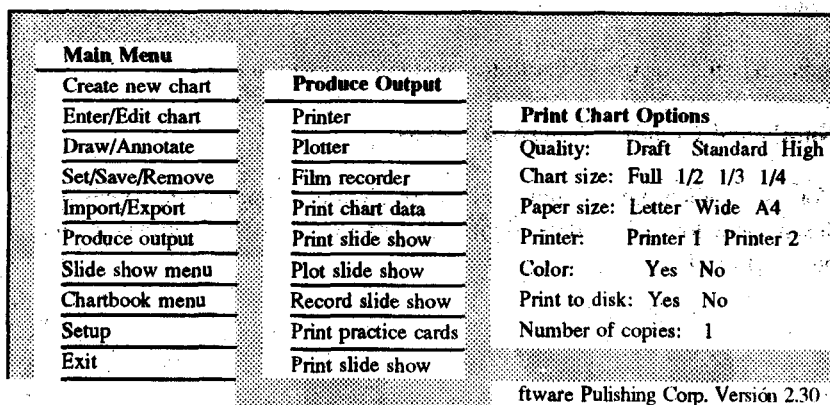
Las opciones de impresión de gráficos, por definición (Default), están identificadas por una pequeña marca en su lado izquierdo. A continuación, se revisará, brevemente, cada una de ellas.

### Quality (Calidad de impresión)

Si elige **Draft**, los gráficos son impresos rápidamente, pero tienen una apariencia rústica.

Con la alternativa **Estándar**, los gráficos se ven mejor que en el caso Draft, pero tarda más en imprimirse.

La alternativa **High** tiene un proceso de impresión más lento, con relación a las dos primeras alternativas, pero los gráficos tendrán una apariencia profesional.



F1-Help

F3-Applications

F2-Draw chart

F4-Spell check

F8-Options

Continue

Figura 10.11

### Chart Size (Tamaño del gráfico)

Esta opción determina el tamaño del gráfico y la posición que éste ocupará en el papel.

La alternativa **Full (Tamaño completo)**, permite tener el gráfico en una página completa.

Con la alternativa **1/2**, el gráfico ocupará la mitad superior de la página.

<b>AOVONEWAY</b>	Stat ▶ ANOVA ▶ Onrway (Unstacked)..
<b>ARIMA</b>	Stat ▶ Time Series ▶ ARIMA...
<b>ASIN</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Functions...
<b>ATAN</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Functions...

**B**

<b>BASE</b>	Calc ▶ Set Base...
<b>BOXPLOT</b>	Graph ▶ Boxplot...
<b>BRIEF</b>	No tiene equivalencia.

**C**

<b>CCF</b>	Stat ▶ Time Series ▶ Cross Correlation...
<b>CCHART</b>	Stat ▶ SPC Charts ▶ C...
<b>CDF</b>	Calc ▶ Probability Distributions ▶ distribution
<b>CENTER</b>	Calc ▶ Standarize...
<b>CHISQUARE</b>	Stat ▶ Tables ▶ Chisquare Test...
<b>CODE</b>	Calc ▶ Code Data Values...
<b>CONCATENATE</b>	Calc ▶ Concatenate...
<b>CONSTANT</b>	Stat ▶ Fit Intercept
<b>CONVERT</b>	Calc ▶ Convert...
<b>COPY</b>	Calc ▶ Copy Columns...
<b>CORRELATION</b>	Stat ▶ Basic Statistics ▶ Correlation...

<b>COS</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Functions...
<b>COUNT</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Column Statistics...
<b>COVARIANCE</b>	Stat ▶ Basic Statistics ▶ Covariance...

**D**

<b>DELETE</b>	Calc ▶ Delete Rows...
<b>DESCRIBE</b>	Stat ▶ Basic Statistics ▶ Descriptive Statistics...
<b>DIFFERENCES</b>	Stat ▶ Time Series ▶ Differences...
<b>DIVIDE</b>	Calc ▶ Functions and Statistics ▶ General Expressions...
<b>DOTPLOT</b>	Graph ▶ Dotplot...

**E**

<b>END</b>	File ▶ Other Files ▶ Stop Storing Macro
<b>ERASE</b>	Calc ▶ Erase Variables...
<b>EWMA</b>	Stat ▶ SPC Charts ▶ EWMA...
<b>EXECUTE</b>	File ▶ Other Files ▶ Execute Macro...
<b>EXPONENTIATE</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Functions...

**G**

<b>GBOXPLOT</b>	Graph ▶ Boxplot...
<b>GCCHART</b>	Stat ▶ SPC Charts ▶ C...
<b>GEWMA</b>	Stat ▶ SPC Charts ▶ EWMA...

<b>GHISTOGRAM</b>	Graph ▶ Histogram...
<b>GICHART</b>	Stat ▶ SPC Charts ▶ Individuals...
<b>GLPLOT</b>	Graph ▶ Scatter Plot...
<b>GMACHART</b>	Stat ▶ SPC Charts ▶ Moving Average...
<b>GMPLOT</b>	Graph ▶ Multiple Scatter Plot...
<b>GMRCHART</b>	Stat ▶ SPC Charts ▶ Moving Range...
<b>GNPCHART</b>	Stat ▶ SPC Charts ▶ NP...
<b>GPCHART</b>	Stat ▶ SPC Charts ▶ P...
<b>GPlot</b>	Graph ▶ Scatter Plot...
<b>GRCHART</b>	Stat ▶ SPC Charts ▶ R...
<b>GSCHART</b>	Stat ▶ SPC Charts ▶ S...
<b>GTPLOT</b>	Graph ▶ Pseudo 3-D Plot...
<b>GUChart</b>	Stat ▶ SPC Charts ▶ U...
<b>GXBARCHART</b>	Stat ▶ SPC Charts ▶ Xbar...

## H

<b>HEIGHT</b>	Graph ▶ Set Options...
<b>HELP</b>	Data Screen Menu ▶ Help...
<b>HISTOGRAM</b>	Graph ▶ Histogram...

## I

<b>ICHART</b>	Stat ▶ SPC Charts ▶ Individuals...
<b>INDICATOR</b>	Calc ▶ Make Indicator Variables...

<b>INFO</b>	Edit ▶ Get Worksheet Info...
<b>INSERT</b>	File ▶ Import ASCII data...
<b>INVCDF</b>	Calc ▶ Probability Distributions ▶ distribution

**J**

<b>JOURNAL</b>	File ▶ Other Files ▶ Start Recording History..
----------------	--

**K**

<b>KRUSKAL-WALLIS</b>	Stat ▶ Nonparametrics ▶ Kruskal-Wallis...
-----------------------	---

**L**

<b>LAG</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Stat ▶ Time Series ▶ Lag...
<b>LET</b>	Calc ▶ Functions and Statistics ▶ General Expressions...
<b>LOGE</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Functions...
<b>LOGTEN</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Functions...
<b>LPLOT</b>	Graph ▶ Scatter Plot...

**M**

<b>MACHART</b>	Stat ▶ SPC Charts ▶ Moving Average...
<b>MANN-WHITNEY</b>	Stat ▶ Nonparametrics ▶ Mann-Whitney...
<b>MAXIMUM</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Column Statistics...

<b>MEAN</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Column Statistics...
<b>MEDIAN</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Column Statistics...
<b>MINIMUM</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Column Statistics...
<b>MPILOT</b>	Graph ▶ Multiple Scatter Plot...
<b>MRCHART</b>	Stat ▶ SPC Charts ▶ Moving Range...
<b>MTSPLOT</b>	Stat ▶ Time Series ▶ Time Series Plot... Graph ▶ Time Series Plot...
<b>MULTIPLY</b>	Calc ▶ Functions and Statistics ▶ General Expressions...
<b>N</b>	
<b>N</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Column Statistics...
<b>NAME</b>	Con los datos en hoja de cálculo ▶ F10 ▶ Alt N
<b>NMISS</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Column Statistics...
<b>NOCONSTANT</b>	Stat ▶ Fit Intercept...
<b>NOJOURNAL</b>	File ▶ Other Files ▶ Stop Recording History
<b>NOOUTFILE</b>	File ▶ Other Files ▶ Stop Recording History
<b>NPCHART</b>	Stat ▶ SPC Charts ▶ NP...
<b>NSCORES</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Functions...

**O**

<b>ONEWAY</b>	Stat ▶ ANOVA ▶ Oneway...
<b>OUTFILE</b>	File ▶ Other Files ▶ Start Recording Session...

**P**

<b>PACF</b>	Stat ▶ Time Series ▶ Partial Autocorrelation...
<b>PARPRODUCTS</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Functions...
<b>PARSUMS</b>	Calc ▶ Functions and Statistics ▶ General Expressions.. Calc ▶ Functions and Statistics ▶ Functions...
<b>PCHART</b>	Stat ▶ SPC Chart ▶ P...
<b>PDF</b>	Calc ▶ Probability Distributions ▶ distribution
<b>PLOT</b>	Graph ▶ Scatter Plot...
<b>PRINT</b>	Edit ▶ Display Data...

**R**

<b>RAISE</b>	Calc ▶ Functions and Statistics ▶ General Expressions...
<b>RANDOM</b>	Calc ▶ Random Data ▶ distribution
<b>RANK</b>	Calc ▶ Rank... Calc ▶ Functions and Statistics ▶ General Expressions...
<b>RCHART</b>	Stat ▶ SPC Charts ▶ R...
<b>RCOUNT</b>	Calc ▶ Functions and Statistics ▶ Row Statistics...
<b>READ</b>	File ▶ Import ASCII data...
<b>REGRESS</b>	Stat ▶ Regression ▶ Regression...



<b>RESTART</b>	File ▶ Restart Minitab
<b>RETRIEVE</b>	File ▶ Open Worksheet...
<b>RMAXIMUM</b>	Calc ▶ Functions and Statistics ▶ Row Statistics...
<b>RMEAN</b>	Calc ▶ Functions and Statistics ▶ Row Statistics...
<b>RMEDIAN</b>	Calc ▶ Functions and Statistics ▶ Row Statistics...
<b>RMINIMUM</b>	Calc ▶ Functions and Statistics ▶ Row Statistics...
<b>RN</b>	Calc ▶ Functions and Statistics ▶ Row Statistics...
<b>RNMISS</b>	Calc ▶ Functions and Statistics ▶ Row Statistics...
<b>ROUND</b>	Calc ▶ Functions and Statistics ▶ General Expressions...
<b>RSSQ</b>	Calc ▶ Functions and Statistics ▶ Row Statistics...
<b>RSTDEV</b>	Calc ▶ Functions and Statistics ▶ Row Statistics...
<b>RSUM</b>	Calc ▶ Functions and Statistics ▶ Row Statistics...
<b>RUNS</b>	Stat ▶ Nonparametrics ▶ Runs Test...

### S

<b>SAMPLE</b>	Calc ▶ Random Data ▶ Sample from Columns...
<b>SAVE</b>	File ▶ Worksheet As...
<b>SCHART</b>	Stat ▶ SPC Charts ▶ S...
<b>SET</b>	Edit ▶ Set Patterned Data...
<b>SIGN</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Functions...
<b>SIN</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Functions...

<b>SINTERVAL</b>	Stat ▶ Nonparametrics ▶ 1-Sample Sign...
<b>SORT</b>	Calc ▶ Sort...
<b>SQRT</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Functions...
<b>SSQ</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Column Statistics...
<b>STACK</b>	Calc ▶ Stack...
<b>STDEV</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Column Statistics...
<b>STEM-AND-LEAF</b>	Graph ▶ Stem-and-Leaf...
<b>STEPWISE</b>	Stat ▶ Regression ▶ Stepwise...
<b>STEST</b>	Stat ▶ Nonparametrics ▶ 1-Sample Sign...
<b>STOP</b>	File ▶ Exit
<b>STORE</b>	File ▶ Other Files ▶ Start Storing Macro...
<b>SUBTRACT</b>	Calc ▶ Functions and Statistics ▶ General Expressions...
<b>SUM</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Column Statistics...
<b>T</b>	
<b>TABLE</b>	Stat ▶ Tables ▶ Cross Tabulation...
<b>TALLY</b>	Stat ▶ Tables ▶ Tally...
<b>TAN</b>	Calc ▶ Functions and Statistics ▶ General Expressions... Calc ▶ Functions and Statistics ▶ Functions...
<b>TINTERVAL</b>	Stat ▶ Basic Statistics ▶ 1-Sample t...

<b>TPLOT</b>	Graph ▶ Pseudo 3-D Plot...
<b>TSPLIT</b>	Stat ▶ Time Series ▶ Time Series Plot... Graph ▶ Time Series Plot...
<b>TTEST</b>	Stat ▶ Basic Statistics ▶ 1-Sample t...
<b>TWOSAMPLE</b>	Stat ▶ Basic Statistics ▶ 2-Sample t...
<b>TWOT</b>	Stat ▶ Basic Statistics ▶ 2-Sample t...
<b>TWOWAY</b>	Stat ▶ ANOVA ▶ Twoway...

## U

<b>UCHART</b>	Stat ▶ SPC Charts ▶ U...
<b>UNSTACK</b>	Calc ▶ Unstack...

## W

<b>WALSH</b>	Stat ▶ Nonparametrics ▶ Pairwise Averages...
<b>WDIFF</b>	Stat ▶ Nonparametrics ▶ Pairwise Differences...
<b>WIDTH</b>	Graph ▶ Set Options...
<b>WINTERVAL</b>	Stat ▶ Nonparametrics ▶ 1-Sample Wilcoxon...
<b>WRITE</b>	File ▶ Export ASCII data...
<b>WSLOPE</b>	Stat ▶ Nonparametrics ▶ Pairwise Slopes...
<b>WTEST</b>	Stat ▶ Nonparametrics ▶ 1-Sample Wilcoxon...

## X

<b>XBARCHART</b>	Stat ▶ SPC Charts ▶ Xbar...
------------------	-----------------------------

**Z****ZINTERVAL**

Stat ► Basic Statistics ► 1-Sample Z...

**ZTEST**

Stat ► Basic Statistics ► 1-Sample Z...

cuadernos  
de  
investigación

**CICLOS Y TENDENCIAS EN LA  
ECONOMÍA PERUANA:  
1950-1989**

Bruno Seminario De Mera  
César Bouillon Buenda



UNIVERSIDAD DEL PACÍFICO  
CENTRO DE INVESTIGACIONES



Instituto de Investigación Económica

cuadernos  
de  
investigación

*Nuevas publicaciones:*

- SEMINARIO, Bruno y César Bouillon:  
Ciclos y tendencias en la economía peruana: 1950 - 1989
- PORTOCARRERO, Felipe y Luis Torrejón:  
Modernización y atraso en las haciendas de la élite  
económica. Perú: 1916 - 1932
- AMAT Y LEÓN, Carlos y otros:  
Análisis y recomendaciones de política para el sistema lácteo
- BOTTERI, Giancarlo y otros:  
Ensayos sobre la realidad económica peruana I
- FERNÁNDEZ-BACA, Jorge y Lorena Canalle:  
Rigideces nominales y dinámica de la inflación en  
el Perú: 1950-1990. Una aproximación empírica
- URRUNAGA, Roberto, Guillermo Berastain y Luis Bravo:  
Desempeño empresarial durante el proceso de  
estabilización
- YAMADA, Gustavo:  
Autoempleo e informalidad urbana: teoría y evidencia  
empírica de Lima Metropolitana, 1985-86 y 1990

En venta en las principales librerías del país.  
Informes en la Librería de la Universidad del Pacífico.  
Teléfonos: 71-2277, 72-9635. Fax: 706121



UNIVERSIDAD DEL PACÍFICO

# documento de trabajo



- José Luis Sardón:  
Estado, política y gobierno
- Rosario Gómez:  
La comercialización de mango fresco en el mercado norteamericano
- Héctor Malarín y Paul Remy:  
La contaminación de aguas superficiales en el Perú: una aproximación económico-jurídica
- Elsa Galarza y Héctor Malarín:  
Lineamientos para el manejo eficiente de los recursos en el sector pesquero industrial peruano
- Gustavo Yamada:  
Estrategias de desarrollo, asistencia financiera oficial e inversión privada directa: la experiencia japonesa.
- Julio Velarde y Martha Rodríguez:  
El programa de estabilización peruano: evaluación del periodo 1991-1993
- Felipe Portocarrero y María Elena Romero:  
Política social en el Perú 1990-1994: una agenda para la investigación.
- Jürgen Schuldt:  
La enfermedad holandesa y otros virus de la economía peruana.

En venta en las principales librerías del país.  
Informes en la Librería de la Universidad del Pacífico.  
Teléfonos: 71-2277. 72-9635. Fax: 706121



UNIVERSIDAD DEL PACIFICO