



UNIVERSIDAD NACIONAL  
AUTONOMA DE MEXICO

FACULTAD DE ESTUDIOS SUPERIORES  
CUAUTITLAN



35  
MEXICO DE  
ESTADISTICA  
ANTOLOGIA  
LA UN  
ALIAS

**"MANUAL DE USO DEL PAQUETE ESTADISTICO MICROSTAT  
CON APLICACIONES A LA CARRERA DE MEDICO  
VETERINARIO ZOOTECNISTA"**

T E S I S  
Que para obtener el Título de  
**MEDICA VETERINARIA ZOOTECNISTA**  
P r e s e n t a

**SILVIA ESCOBAR RAMIREZ**

Asesor: **M. en C. Benito López Baños**

**Cuautitlán Izcalli, Edo. de Méx.**

**1995**



## **UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso**

### **DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL  
AUTÓNOMA DE MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES  
CUAUILTÁN



35  
29

TRABAJO DE USO DEL PAQUETE ESTADÍSTICO  
CON APLICACIONES A LA CARRERA DE  
VETERINARIO ZOOTECNISTA

1995  
1995

T E S I S  
Que para obtener el Título de  
MEDICA VETERINARIA ZOOTECNISTA  
P R E S E N T A

SILVIA ESCOBAR RAMÍREZ

Asesor: M. en C. Benito López Baños

Cuautitlán Izcalli, Edo. de Méx.

1995



UNIVERSIDAD NACIONAL  
AVENIDA DE  
MÉXICO

FACULTAD DE ESTUDIOS SUPERIORES CUAUTITLAN  
UNIDAD DE LA ADMINISTRACION ESCOLAR  
DEPARTAMENTO DE EXAMENES PROFESIONALES

U. N. A. M.  
FACULTAD DE ESTUDIOS  
SUPERIORES-CUAUTITLAN



DEPARTAMENTO DE  
EXAMENES PROFESIONALES

ASUNTO: VOTOS APROBATORIOS

DR. JAIME KELLER TORRES  
DIRECTOR DE LA FES-CUAUTITLAN  
P R E S E N T E .

AT'N: Ing. Rafael Rodríguez Ceballos  
Jefe del Departamento de Exámenes  
Profesionales de la F.E.S. - C.

Con base en el art. 23 del Reglamento General de Exámenes, nos permitimos comunicar a usted que revisamos la TESIS TITULADA:

"Anual de uso del paquete estadístico "Microstat", con aplicaciones  
a la carrera de médico veterinario zootecnista".

que presenta la pasante: Silvia Escobar Ramírez  
con número de cuenta: 3356127-0 para obtener el TITULO de:  
Médico Veterinaria Zootecnista .

Considerando que dicha tesis reúne los requisitos necesarios para ser discutida en el EXAMEN PROFESIONAL correspondiente, otorgamos nuestro VOTO APROBATORIO.

A T E N T A M E N T E .  
"POR MI RAZA HABLARA EL ESPIRITU"  
Cuautitlán Izcalli, Edo. de Méx., a 5 de Julio de 1994

PRESIDENTE	<u>MVZ. Benito López Baños</u>	
VOCAL	<u>MVZ. Fernando Altamirano Abarca</u>	
SECRETARIO	<u>MVZ. Miguel Angel Carrona Hadero</u>	
PRIMER SUPLENTE	<u>MVZ. Geriberto Contreras Angles</u>	
SEGUNDO SUPLENTE	<u>M.C. Jorge Bermúdez Estívez</u>	

EL SUEÑO SE HACE A MANO Y  
SIN PERMISO, ARANDO EL PORVENIR  
CON VIEJOS BUEYES.

CAMINANTE NO HAY CAMINO... SE HACE CAMINO AL ANDAR.

**D E D I C A T O R I A S**

DOY GRACIAS A DIOS  
POR AYUDARME A LLEGAR  
A OTRO PUNTO EN MI CAMINO Y  
TENER UN MOTIVO MAS PARA ALABARLE Y GLORIFICARLE.

CON PROFUNDO AGRADECIMIENTO A MI PADRE:

RAMIRO ESCOBAR ZAMORA

GRACIAS...

POR BRINDARME SIEMPRE LO MEJOR DE TÍ.  
POR TU AMOR, APOYO Y COMPRESIÓN.  
POR ENSEÑARME A APRECIAR LA VIDA EN TODO MOMENTO Y BAJO  
CUALQUIER CIRCUNSTANCIA.  
POR DARME UN HOGAR DONDE LOS LAZOS DE UNIÓN SON TAN FUERTES Y  
LOS SENTIMIENTOS DE AMOR SON TAN PROFUNDOS Y SINCEROS, QUE  
HE APRENDIDO A AMAR, LUCHAR, PERDONAR Y TENER FÉ EN DIOS Y EN  
MÍ MISMA.  
POR NEGARTE A TÍ MISMO PARA QUE YO LOGRE SER QUIEN SOY.  
POR SER MI MÁXIMO EJEMPLO, ESPERANDO VEAS CUMPLIDAS LAS  
ESPERANZAS QUE SEMBRASTE EN MÍ.

¡ERES LO MAXIMO Y MI MAS GRANDE Y PURO AMOR!

TE AMO

**MADRE:**

NO OBSTANTE TU SUFRIMIENTO, ME HAS DADO UN GRAN EJEMPLO, CON TU TRABAJO, TENACIDAD, HONESTIDAD, PERSEVERANCIA Y RESPONSABILIDAD CONSTANTES, HE LOGRADO LLEGAR A ESTE PUNTO Y DESEO CONSIDERES COMO TUYA LA SATISFACCIÓN DE HABER LOGRADO UNO DE MIS SUEÑOS Y DEBERES, QUE SIN TU CONFIANZA, AMOR Y CARINO NO HUBIESE REALIZADO.

**GRACIAS.**

ROSSY:

POCOS SERES POSEEN LA DICHA DE  
TENER A ALGUIEN QUIEN REPRESENTA  
RESPECTO, AMOR Y ADMIRACIÓN, LO CUAL  
ES UNA BENDICIÓN QUE DIOS ME HA REGALADO  
EN TÍ. ME ES DIFÍCIL EXPRESAR LO BENÉFICO  
QUE HA SIDO PARA MÍ TU LLEGADA A MI VIDA,  
POR ELLO SIMPLEMENTE ¡GRACIAS! (TU ENTIENDES LA  
MAGNITUD) Y SÓLO ME RESTA DECIRTE QUE SI LA VIDA

ME DIERA A ESCOGER A MIS SERES AMADOS,

INDUDABLEMENTE TÚ SERÍAS DE LOS PRIMEROS.

A MIS HERMANOS

RAMIRO:

POR SER COMO ERES TE AMO  
INFINITAMENTE Y POR QUE ESTOY  
SEGURA QUE NOS SENTIMOS MÁS  
UNIDOS QUE NUNCA. TAL VEZ TE  
RESULTE DIFÍCIL COMPRENDER LO  
IMPORTANTE QUE HA SIDO PARA  
MÍ TU EJEMPLO, EL CUAL SE  
REFLEJA EN PARTE EN ESTAS  
HOJAS. ESPERO QUE MUY PRONTO  
SIENTAS ESTA SATISFACCIÓN QUE  
AHORA SIENTO, PORQUE ¿SABES?  
LO VAS A LOGRAR TODO.  
¡SIGUE ADELANTE!

BELINDA:

POR APOYARME SIEMPRE, SIN  
IMPORTAR LA DISTANCIA QUE NOS  
SEPARARA NI AÚN NUESTRAS GRANDES  
DIFERENCIAS. TENGO TANTO QUE  
AGRADECERTE, QUE NO BASTARÍA  
NINGUNA FRASE PARA EXPRESARLO.  
GRACIAS POR ENSEÑARME A SER COMO  
SOY, POR SER MI CRÍTICA MÁS  
SEVERA Y AL MISMO TIEMPO LA PORRA  
MÁS ENTUSIASTA QUE PUDIERA TENER.

IVAN JOSUE:

QUE ESTE TRABAJO SEA PARA TÍ UN EJEMPLO A SEGUIR O UN RETO A SUPERAR. NO SABES LO FELICES QUE HARÁS A NUESTROS PADRES. SIGUE ADELANTE, SUPÉRATE, CÚMPLELES EL DESEO DE VERTE FORMADO COMO UN EXCELENTE HOMBRE Y PROFESIONISTA.

ALDO:

GRACIAS POR ENSEÑARME LA ESPONTANEIDAD DEL AMOR INFANTIL Y A COMPRENDER QUE NADIE SOMOS PERFECTOS.

INGRID:

A QUIEN NO NECESITO DECIRLE NADA MÁS QUE ES UN ÁNGEL Y QUE AGRADEZCO A LA VIDA POR PERMITIRME TENER LA OPORTUNIDAD DE AYUDARLA A CRECER, SIENDO LA MÁS PEQUEÑA DE NUESTRA FAMILIA. GRACIAS POR ENTREGARME TU CARIÑO CON TERNURA E INOCENCIA.

A MIS ABUELITOS

FELIPE SANTOS Y ESTELA ORDAZ:  
POR SU INCONDICIONAL Y GRAN AMOR.

DR. BENITO LOPEZ BAÑOS:

GRACIAS POR SU COMPRESIÓN, PACIENCIA Y APOYO, QUE A LO LARGO DE ESTE TIEMPO FUERON UN ALIENTO PARA CULMINAR ESTE TRABAJO QUE POR FIN SE VE REALIZADO. LAS PALABRAS FALTAN PERO EL AGRADECIMIENTO ES INFINITO.

OFICIAL F.A.A.M.A. J.A.D. PINEDA:

UN AGRADECIMIENTO ESPECIAL A TÍ, PORQUE ME SIENTO AFORTUNADA DE HABER TENIDO UNA LUZ QUE ME GUIÓ EN LA OBSCURIDAD, MANTENIÉNDOME A FLOTE CON TU AYUDA Y APOYO, SIN EL CUAL, AÚN NO LLEGARÍA A ESTE PUNTO. LO DEMÁS QUE PUDIERA DECIR, YA LO SABES.

## INDICE

	Pág.
RESUMEN .....	1
INTRODUCCION .....	2
OBJETIVO .....	6
PROCEDIMIENTO .....	7
CAPITULO 1. Presentación del programa .....	8
CAPITULO 2. Manejo de Archivos .....	13
CAPITULO 3. Estadística Descriptiva .....	44
CAPITULO 4. Tablas de Frecuencia .....	52
CAPITULO 5. Permutación y combinación .....	60
CAPITULO 6. Distribución de Probabilidad .....	66
CAPITULO 7. Pruebas de Hipótesis .....	77
CAPITULO 8. Análisis de Varianza .....	88
CAPITULO 9. Gráficos .....	101
CAPITULO 10. Matriz de Correlación .....	105
CAPITULO 11. Análisis de Regresión .....	111
CAPITULO 12. Pruebas de CHI cuadrada .....	124
CAPITULO 13. Estadísticas no Paramétricas .....	134
DISCUSION .....	147
CONCLUSION .....	150
BIBLIOGRAFIA .....	151
APENDICE .....	153

## RESUMEN

El siguiente manual se ha hecho en base al contenido del propio paquete "Microstat" y su asociación con el programa de Bioestadística que cursan los estudiantes de la carrera de Medicina Veterinaria y Zootecnia, que se imparte en la FES-Cuautitlán, seleccionando ejemplos prácticos para cada uno de los 13 capítulos de que consta este trabajo.

Capítulo 1. Presentación del programa. Descripción general de la utilización del paquete "Microstat", el método de acceso al mismo y la elaboración de un respaldo de dicho paquete.

Capítulo 2. Manejo de Archivos. Bases necesarias para la creación de archivos, así como para copiar, renombrar, modificar o borrar los ya existentes.

Capítulo 3. Estadística Descriptiva. Descripción del método para obtener cálculos de: media aritmética, varianza, desviación estándar, coeficiente de variación, error estándar y suma de cuadrados.

Capítulo 4. Tablas de Frecuencia. Pasos a seguir para la elaboración de tablas de frecuencia para datos agrupados o con valores individuales.

Capítulo 5. Permutación y Combinación. Manejo del programa para la obtención del factorial de un  $N$ , las permutaciones de  $N$  objetos proporcionando el valor de  $R$  y el  $\#$  de combinaciones de  $N$  objetos dando también un valor  $R$ .

Capítulo 6. Distribución de Probabilidad. Descripción del método para obtener la distribución binomial de  $x$ , comportamiento de la distribución de Poisson, así como la distribución normal para cálculos de probabilidad proporcionando el valor de  $Z$  y cálculos de  $Z$  proporcionando la probabilidad.

Capítulo 7. Pruebas de Hipótesis. Método para la realización de pruebas de hipótesis con las siguientes opciones: comparando medias de una muestra con valor hipotético, comparando medias con observaciones pareadas, comparando 2 medias con varianza conocida o comparando 2 medias con varianza común.

Capítulo 8. Análisis de Varianza. Descripción de la secuencia que se lleva a cabo para hacer análisis de Varianza con un solo camino de clasificación (modelo totalmente aleatorizado) o con 2 criterios de clasificación (modelo de bloques al azar).

Capítulo 9. Gráficos. Pasos para la obtención de gráficos de punto con trazo de la línea de regresión.

Capítulo 10. Matriz de Correlación. Método para la obtención del análisis de correlación simple, con 2 ó más variables.

Capítulo 11. Análisis de Regresión. Manejo del paquete para realizar análisis de regresión simple o múltiple, esta última además, con un modelo completo de regresión o con un modelo de regresión por pasos.

Capítulo 12. Pruebas de CHI cuadrada. Metodología para realizar pruebas de CHI cuadrada bajo los modelos de: tablas de contingencia a partir de un archivo ya existente, tablas de contingencia introduciendo la información directamente desde el teclado de la computadora y pruebas de bondad de ajuste.

Capítulo 13. Estadísticas no Paramétricas. Desarrollo del método que permite hacer cálculos no paramétricos por medio de pruebas como: prueba de rango con signo, suma de rangos para 2, 3 ó más grupos y prueba de Kolmogorov-Smirnov para 2 grupos.

## INTRODUCCION

El uso de las computadoras ha tenido un crecimiento explosivo en los últimos 50 años. Conforme ha evolucionado la computadora, se ha difundido su uso y pasa a ser un eficiente procesador de datos utilizado en todos los campos relacionados con el almacenamiento, procesamiento y recuperación de datos, siendo éste el uso más difundido de la computación en ésta época, además de que las tendencias actuales van encaminadas a que la computadora se convierta en un procesador del conocimiento, es decir, una extensión del cerebro humano (3).

Las computadoras tienen un sin número de aplicaciones, como pueden ser militares, administrativas, industriales científicas, de investigación, educativas, médicas y hasta en la vida diaria. Ha sido tal la proliferación de las computadoras, que resulta poco probable encontrar una determinada actividad humana en que no esté involucrada de una u otra forma la computadora (3,5,9). En el área de la investigación, el proceso estadístico genera información exacta y definida respecto de las características de los datos, de forma tal que pueda ser comunicada con facilidad ya que reduce masas de datos a sencillos términos descriptivos y extrae inferencias de ellos (15,17).

Así pues, el área de medicina veterinaria y zootecnia no puede ser excluida del uso de la computadora, siendo ésta muy

útil en aplicaciones administrativas (para automatización de los diferentes procesos rutinarios que se presentan en cualquier empresa, institución o granja, como procesos contables y financieros de la organización , proveedores y/o distribuidores, control de adquisiciones, inventarios, almacenes, pie de cría, semen, etc.), industriales (para realizar mediciones, analizarlas y tomar acciones específicas, lo que se traduce en el seguimiento y control automático de procesos industriales, verbigracia, alimenticios), en la investigación y medicina (para llevar a cabo monitoreo en la línea de pacientes, nutrición, inventario de fármacos, pruebas de laboratorio, análisis e investigación, manejo genético y administración de clínicas), etcétera (2,7,8,18). Es por ello que tanto al alumnado, así como los profesionistas, profesores e investigadores de la medicina veterinaria y zootecnia, deben estar habilitados para el aprovechamiento de los resultados obtenidos por otros investigadores y por ellos mismos (4).

En la actualidad el médico veterinario zootecnista se ha visto en la necesidad de involucrarse en el área de la informática por el uso de la computadora como un instrumento de trabajo. Los avances logrados en dicha área han puesto a su disposición una poderosa herramienta que poco a poco está ocasionando importantes cambios en todas las áreas de la práctica veterinaria, por ejemplo, dentro de las aplicaciones más relevantes de la computación en la medicina veterinaria y zootecnia es considerado el análisis estadístico de datos (13), ya que se cuenta con los sistemas de registros y expedientes

tanto para consultorios como para granjas, los sistemas de planeación, control y análisis de granjas, los sistemas de nutrición y balanceo de raciones de costo mínimo, diagnóstico asistido por computadoras, sistemas de ayuda en el establecimiento de estrategias terapéuticas para diversos estados patológicos y uso en el manejo e interpretación de estudios estadísticos para evaluación de parámetros productivos, reproductivos y de salud, lo que orienta, facilita y hace posible una actitud acertada en la toma de decisiones y en la solución de problemas (1,11,14). La estadística esta asociada con el método científico bajo la forma, organización, recopilación, presentación y análisis de datos, tanto para la deducción de razonamientos así como para la toma de decisiones lógicas con dicho análisis (10,16).

Es por tanto, que para este efecto se utilizan los paquetes estadísticos y de ahí que la finalidad de este trabajo, es brindar al médico veterinario zootecnista, un manual de uso de un paquete estadístico que le ayude a integrarse al aprovechamiento de las computadoras en su práctica diaria, mejorando a través de ello la eficiencia productiva, financiera y económica de las empresas pecuarias y áreas involucradas a su cargo, como puede verse en evaluaciones de mortalidad, productividad, producción, peso de canales, etc., en que ya ha sido utilizado el paquete estadístico "Microstat", lo cual redundará en importantes beneficios para el país (2,9,12,14).

Así pues, el trabajo que a continuación se describe es un manual para el programa de computación estadística enfocado a las aplicaciones generales que dicho programa puede tener en el área de medicina veterinaria y zootecnia. El desarrollo se lleva a cabo a través del planteamiento de ejemplos cotidianos asociados a la práctica veterinaria, proporcionando la metodología descriptiva para su resolución por medio del uso del paquete estadístico "Microstat", el cual puede ser adquirido en la siguiente dirección:

Product: Microstat-II (Rel.2.5)

\*Data Sources Report COPYRIGHT Ziff-Davis Publishing Co.199

-----

Company: Ecosoft, Inc.  
Address: 8295 Indy Court  
Indianapolis, IN 46214  
800-952-0472; 317-271-5551  
FAX: 317-271-5561

-----

Category: Software, Applications  
Statistics and Mathematics

Specs: Release date: 1978  
Application: Statistics & Mathematics  
Compatible with: PC-MS/DOS  
Minimum RAM required: 512 KB  
Source language: C

-----

## OBJETIVO

Elaborar un manual de uso del paquete estadístico "Microstat", por medio de una computadora personal, que permita el análisis estadístico de la información obtenida en la práctica de las diversas áreas de la medicina veterinaria y zootecnia.

## PROCEDIMIENTO

Se realizó una revisión bibliográfica de libros, revistas técnicas, tesis y textos del área de medicina veterinaria y zootecnia con la finalidad de integrar un paquete de ejercicios o problemas relacionados con algunas asignaturas de la carrera de medicina veterinaria y zootecnia. Los problemas seleccionados se resolvieron utilizando la computadora P.C. y el programa estadístico del "Microstat".

El manual se integró con las órdenes de manejo del programa "Microstat" para resolver cada uno de los ejercicios, así como sus resultados.

## C A P I T U L O 1

### PRESENTACION DEL PROGRAMA

**MICROSTAT** es un paquete estadístico elaborado para ser usado en computadoras de tipo personal como son PC XT y PC AT y aún en los modelos 386 ó superiores en los que ha sido probado con muy buena eficiencia. La cantidad de memoria que necesita este paquete es un mínimo de 512 kbytes de memoria RAM y puede ser usado en PC con uno o dos drives e incluso en disco duro. La capacidad máxima de memoria que utiliza MICROSTAT es del orden de unos 200 kbytes de memoria RAM, lo que permite utilizar archivos con el resto de memoria.

**MICROSTAT** como su nombre lo indica, es un micropaquete, sin embargo, por su capacidad de trabajo, es un potente paquete estadístico capaz de hacer cualquier análisis estadístico de uso común. Puede ser de utilidad para cualquier profesionista, independientemente del área en que labore, desde investigadores en el área biológica hasta el área economico-administrativa, pasando por las áreas de sociología y psicología, entre otras. La facilidad de manejo del paquete, así como el poco requerimiento de equipo, le hacen un paquete versátil que seguramente tendrá una buena aceptación entre los estudiantes del nivel de licenciatura y posgrado.

## **1.1 ¿QUE HACE MICROSTAT?**

**MICROSTAT** puede hacer desde sus propios archivos para posteriormente ser utilizados, hasta el análisis estadístico descriptivo, distribuciones de probabilidad, pruebas de hipótesis, estudios de correlación y regresión lineal simple y múltiple, análisis de chi cuadrada, pruebas de ANDEVA e inclusive análisis estadísticos no paramétricos.

**Nota:** El presente manual no pretende substituir al manual original que acompaña al paquete ni ser un curso de computación, ya que se da por hecho que el usuario conoce los conceptos básicos de computación que le permitirán moverse con soltura frente a la PC. Tampoco se pretende hacer un manual estadístico, aunque para explicar el uso de algunos conceptos estadísticos del paquete, se utilizarán ejercicios que remitirán constantemente a algún texto estadístico.

## **1.2 INICIANDO CON MICROSTAT.**

El programa **MICROSTAT** normalmente viene en un formato de diskette flexible de 5  $\frac{1}{4}$  . El primer paso a seguir con el paquete original de **MICROSTAT** es hacer un respaldo. A continuación se ofrecen dos opciones para trabajar con este paquete y desarrollar el respaldo.

La primera opción considera un equipo con dos "drive", el A y el B, con formatos de 5 ¼ :

1. Encender la máquina.
2. Cargar el Sistema Operativo (MS-DOS).
3. Estando el cursor en posición de espera en el drive A, introducir en el drive B un diskette virgen.
4. Teclear la siguiente orden: **FORMAT B:/S**. El diskette del Sistema Operativo debe continuar en el drive A, al dar "Enter" la máquina iniciará el formateo en el drive B. Con ésta orden se cargará el Sistema Operativo y algunos archivos del Comand. Com. en el diskette B, que es donde se recibirá el respaldo del paquete MICROSTAT.
5. Una vez terminado el formateo del diskette instalado en el drive B, debe retirarse el diskette que contiene el Sistema Operativo (drive A) y se introduce en ese drive el diskette que contiene el paquete original del programa MICROSTAT y se teclaea la siguiente orden: **COPY \*.\* B:** y dar "Enter".

A partir de aquí, la máquina iniciará el copiado del paquete original. Una vez terminado éste proceso, se debe guardar el paquete original e iniciar el trabajo con la copia del mismo, instalándola en el drive A.

Ya para trabajar con el paquete estadístico, se teclaea la palabra completa **MICROSTAT** (estando el diskette con el programa en el drive A), se da "Enter" e inmediatamente aparece en pantalla el MENU PRINCIPAL, que está compuesto de 16 opciones enumeradas con letras de la A a la P, como se observa en la fig.

1.

**FIG. 1 MENU PRINCIPAL**

----- MICROSTAT -----

Copyright (c) 1984 by Ecosoft, Inc.

**OPTIONS:**

- |                              |                                    |
|------------------------------|------------------------------------|
| A. DATA MANAGEMENT SUBSYSTEM | I. TIME SERIES ANALYSIS            |
| B. DESCRIPTIVE STATISTICS    | J. NONPARAMETRIC STATISTICS        |
| C. FREQUENCY DISTRIBUTIONS   | K. CROSSTAB / CHI-SQUARE TESTS     |
| D. HYPOTHESES TESTS: MEAN    | L. PERMUTATIONS / COMBINATIONS     |
| E. ANALYSIS OF VARIANCE      | M. PROBABILITY DISTRIBUTIONS       |
| F. SCATTERPLOT               | N. HYPOTHESIS TESTS: PROPORTIONS   |
| G. CORRELATION MATRIX        | O. <Identification / Installation> |
| H. REGRESSION ANALYSIS       | P. <Terminate>                     |

-----  
ENTER: OPTION: \_

**Donde:**

- A. SUBSISTEMA PARA EL MANEJO DE DATOS
- B. ESTADISTICA DESCRIPTIVA
- C. TABLAS O DISTRIBUCIONES DE FRECUENCIA
- D. PRUEBAS DE HIPOTESIS PARA MEDIAS
- E. ANALISIS DE VARIANZA
- F. GRAFICAS SENCILLAS
- G. MATRIZ DE CORRELACION
- H. ANALISIS DE REGRESION
- I. ANALISIS DE SERIES DE TIEMPO
- J. ESTADISTICAS NO PARAMETRICAS
- K. PRUEBAS DE CHI CUADRADA

- L. PERMUTACIONES Y COMBINACIONES
- M. DISTRIBUCIONES DE PROBABILIDAD
- N. PRUEBAS DE HIPOTESIS DE PROPORCIONES
- O. INSTALAR E IDENTIFICAR EL PAQUETE
- P. SALIR DEL PROGRAMA

-----  
INTRODUZCA LA OPCION: \_

La segunda opción para trabajar con MICROSTAT es para un equipo que posee disco duro y un drive A, con formato de 5 ¼:

1. Encender la máquina (el sistema operativo se cargará directamente del disco duro).
2. Estando el cursor en posición de espera en el drive C, teclear la siguiente orden: MD STAT, que es el nombre del archivo con el que se identificará el paquete estadístico en el disco duro (no debe rebasar ocho caracteres). Al dar "Enter", la máquina creará un directorio especial para copiar el paquete estadístico.
3. Introducir en el drive A el diskette original del paquete. Teclear desde el drive C:, CD\STAT y dar "Enter", anotar entonces COPY A:\*. \* y dar "Enter". Con ésta orden, la máquina copiará todos los archivos que estén en el diskette del drive A al drive C, concretamente en el subdirectorío STAT.
4. Una vez terminado este proceso, se retira el diskette original y desde esa misma posición teclear la palabra MICROSTAT, dar "Enter" y aparecerá en pantalla el menú principal mostrado anteriormente en la fig. 1.

## C A P I T U L O 2

### MANEJO DE ARCHIVOS

Como es posible observar, la opción A del menú principal, es un SUBSISTEMA DE MANEJO DE DATOS. Al seleccionar este inciso, aparece en pantalla un submenú con 15 alternativas, como lo muestra la fig. 2.

FIG. 2

-----  
ENTER: OPTION: A

----- DATA MANAGEMENT SUBSYSTEM -----

DATA FILE OPTIONS:

A. ENTER DATA	H. DELETE CASES
B. LIST DATA	I. VERTICAL AUGMENT
C. EDIT DATA	J. SORT
D. RENAME FILE / EDIT HEADER	K. RANK-ORDER
E. FILE DIRECTORY	L. LAG TRANSFORMATIONS
F. DESTROY FILES	M. READ/WRITE EXTERNAL FILES
G. RECODE/TRANSFORM/SELECT	N. TRANSPOSE FILE
	O. <Terminate>

-----  
ENTER: OPTION: \_

Donde cada opción permite realizar las siguiente funciones:

- A. Ingresar datos para hacer archivos nuevos.
- B. Listar datos de los archivos.

- C. Corregir datos de archivos ya elaborados.
- D. Renombrar archivos o corregir encabezados de archivos creados.
- E. Mostrar el directorio de archivos contenidos en un diskette determinado.
- F. Eliminar archivos caducos.
- G. Recodificar, transformar o seleccionar variables de un archivo determinado.
- H. Eliminar casos en particular de algún archivo.
- I. Aumentar los archivos en forma vertical (aumentar variables).
- J. Clasificar u ordenar datos de un archivo.
- K. Ordenar los datos de un archivo por medio de rangos.
- L. Transformar variables de archivos.
- M. Leer o escribir archivos externos.
- N. Transponer archivos.
- O. Salir del subsistema y pasar al menú principal.

Como puede observarse, la opción **A** del menú principal ofrece una variada gama de opciones para manejar archivos de datos. A continuación se muestra, por medio de un ejemplo, cómo crear un archivo de datos.

## **2.1 CREACION DE ARCHIVOS.**

Asumiendo que se tiene el paquete MICROSTAT instalado en un diskette con formato de 5  $\frac{1}{4}$  introducido en el drive **A**, se coloca en el drive **B** un diskette formateado y como anteriormente se

mencionó, se teclea completa la palabra **MICROSTAT**, se da "Enter" y aparece en pantalla el menú principal (fig. 1), del cual se elige la opción **A**, que despliega en pantalla el submenú de el **SUBSISTEMA DE MANEJO DE DATOS** (fig. 2). De este submenú se selecciona la opción **A** (**ENTER DATA**) y aparece en pantalla otro submenú con 4 opciones, como se muestra en la fig. 3:

FIG. 3

```
-----  
ENTER: OPTION: A  Module being loaded...  
----- DATA MANGEMENT SUBSYSTEM -----  
  
A. ENTER DATA.  
  
OPTIONS: A. START NEW FILE  
          B. ADD DATA TO EXISTING FILE  
          C. INSERT CASE(S) INTO EXISTING FILE  
          D. <Terminate>  
-----  
ENTER: OPTION:  _
```

cuyas aplicaciones son las siguientes: con **A** se inicia un nuevo archivo, **B** permite adicionar datos a un archivo existente, **C** permite insertar nuevos casos a un archivo existente y **D** sirve para salir de ese submenú.

Se elige la opción **A** (**START NEW FILE**) y el programa pide se anote el nombre del archivo que se va a crear, teniendo siempre en cuenta que **es necesario indicar el drive en que se desea guardar cada archivo que sea creado**, en este caso el archivo se guardará en el diskette del drive **B** y el nombre con el que se va a identificar es **EJMO** (cabe recordar que el nombre no debe rebasar 8 dígitos, por lo que se debe escribir **B:EJMO**), se da "Enter" e inmediatamente la máquina pide en pantalla una etiqueta, que es un breve párrafo que permite identificar el

tipo de archivo que se va a construir, en este caso se anota PESO Y ESTATURA DE ALUMNOS (si se desea, la etiqueta puede ser omitida con sólo dar "Enter"). Enseguida pregunta el número de variables que se van a introducir, para este ejemplo serán 2, por lo tanto, se anota el # 2, se da "Enter" y pide el nombre de cada variable, asignándole a la variable 1 el nombre "PESO" y a la 2 "ESTATURA". La máquina ofrece la opción de corregir los nombres, preguntando si son correctos los nombres anotados, si se confirma con la letra Y (yes=si), inmediatamente la máquina indica que ese archivo ha sido creado (fig. 4).

FIG. 4

-----  
ENTER: OPTION: A (opción)

ENTER: FILE NAME: B:EJMO (nombre del archivo)

ENTER: FILE LABEL: PESO Y ESTATURA DE ALUMNOS

ENTER: NUMBER OF VARIABLES: 2 (no. de variables)

ENTER: NAME FOR VARIABLE 1: PESO (nombre de la variable 1)

ENTER: NAME FOR VARIABLE 2: ESTATURA (nombre de la variable 2)

NAMES OK (Y,N)? YES (nombres correctos)

FILE: B:EJMO HAS BEEN CREATED (ha sido creado el archivo EJMO y se halla guardado en el drive B)

PRESS ANY KEY TO CONTINUE. (presione cualquier tecla para continuar)  
-----

Al presionar cualquier tecla, aparece en pantalla un sumario de caracteres de entrada, "Input Character Summary", como el que se muestra en la fig. 5 (si se requiere de mayor información, se

debe consultar el manual original).

FIG. 5

---

Input Character Summary (see Manual for details)

Character(s)	Result
B	Back up one entry
R	Restart at beginning of case
=	Enters case number
, or space	Enters value for previous case
.	Enters missing data code
RETURN	Enters 0 or previously entered data
END	Terminates data input
other	Enters a number if valid, else error message

PRESS ANY KEY TO CONTINUE (presionar cualquier tecla para continuar)

---

Al presionar de nuevo una tecla, aparece en pantalla otro mensaje, que contiene tres caracteres de uso importante en la creación de un archivo (éstos tres caracteres pertenecen al sumario de caracteres de entrada de la fig. 5), los cuales se muestran en la fig. 6, donde E indica al programa que se ha terminado de introducir datos, R permite regresar un bloque (caso) o renglón equivocado y corregirlo aún después de haber dado "Enter" y B se usa para corregir solamente el último dato que acaba de ingresarse. En este punto la máquina pide el dato 1 de la variable 1, el cual debe ser anotado y dar "Enter", a

continuación pide el dato 1 de la variable 2, se anota y se da "Enter", posteriormente pide el dato 2 de la variable 1, el dato 2 de la variable 2 y así sucesivamente. En este ejemplo se anotan los datos como aparecen en la fig. 6.

FIG. 6

-----  
ENTER DATA AS PROMPTED or

('E' TO STOP, 'R' TO RE-ENTER A CASE, 'B' TO RE-ENTER A NUMBER)

1, 1, PESO : 50  
1, 2, ESTATURA: 155

2, 1, PESO : 65  
2, 2, ESTATURA: 169

3, 1, PESO : 84  
3, 2, ESTATURA: 185

4, 1, PESO : 72  
4, 2, ESTATURA: 178

5, 1, PESO : 55  
5, 2, ESTATURA: 155

6, 1, PESO : E

END OF A NEW FILE

PRESS ANY KEY TO CONTINUE  
-----

Cuando la máquina pide el dato 6 de la variable 1, se indica que se ha terminado de introducir datos por medio de la letra E, se da "Enter" y la máquina manda un mensaje que reconoce el fin del archivo, en seguida pide que sea presionada una tecla regresando así a la opción A del menú principal (SUBSISTEMA DE

MANEJO DE DATOS) y poder realizar cualquier otra elección.

**Nota:** Al final de este manual, se encuentra un apéndice de ejemplos, cuyos datos están destinados para crear los archivos que se utilizan en los ejercicios de cada capítulo.

**2.1.1 Adicionar datos a un archivo.** Si se desea insertar o adicionar nuevos datos a un archivo existente, estando en el SUBSISTEMA DE MANEJO DE DATOS (opción A del menú principal), se elige la opción A (ENTER DATA) que mostrará el submenú de 4 opciones descrito en la fig. 3. En este punto se opta por la letra B (ADD DATA TO EXISTING FILE), la máquina pregunta el nombre del archivo al que se desea adicionar datos y se anota, indicando el drive en que se encuentra ese archivo, se da "Enter" y aparece en pantalla la descripción general de dicho archivo (nombre, etiqueta, número de casos y número de variables).

Al pisar cualquier tecla, la máquina muestra en pantalla el sumario de caracteres de entrada de la fig. 5, de nuevo se pisa cualquier tecla y puede observarse en la pantalla que aparece la primera variable después del último dato que se tiene en el archivo y allí es donde se continúa introduciendo los datos requeridos. Nuevamente, para dar fin a la adición de datos, se oprime la letra E y se da "Enter" para regresar al SUBSISTEMA DE MANEJO DE DATOS.

**2.1.2 Insertar casos en un archivo.** Si se requiere

insertar casos en un archivo existente, se utiliza la opción C (INSERT CASE(S) INTO EXISTING FILE) del submenú mostrado en la fig. 3. Enseguida, la máquina pide el nombre del archivo que se desea trabajar, previa indicación del drive en donde se encuentra el diskette que almacena tal archivo, para ejemplificar esto, se anota B:EJMO, se da "Enter" y la máquina muestra los datos generales del archivo y luego pide se indique el punto (caso) donde se va a iniciar la introducción de los datos, anotando el número 3 (caso 3) y se da "Enter" para que aparezca en pantalla la primera variable de ese caso, que es donde se iniciará la introducción de tantos datos como se desee. Para este caso, se introduce en "PESO" el número 80 y en "ESTATURA" el número 184.

Al iniciar otra serie de datos o caso, se teclea la letra E y se da "Enter" para salir del archivo y regresar al SUBSISTEMA DE MANEJO DATOS. De esta manera, se ha agregado en la posición 3 (fila 3) un juego de variables (80 para PESO y 184 para ESTATURA).

Si se desea ver como quedó el archivo, estando en el SUBSISTEMA DE MANEJO DE DATOS (fig. 2), se elige la opción B (LIST DATA), al hacerlo, la máquina pregunta el nombre del archivo que se desea listar, se anota B:EJMO, se da "Enter" y la máquina despliega en pantalla los datos generales del archivo y un submenú con 6 opciones como se muestra en la fig. 7.

**Nota:** Esta figura es referida en todos los capítulos siguientes de este manual.

FIG. 7

-----  
ENTER: FILE NAME: B:EJMO

HEADER DATA FOR: B:EJMO LABEL: PESO Y ESTATURA DE ALUMNOS  
NUMBER OF CASES: 6 NUMBER OF VARIABLES: 2

VARIABLE NUMBERS AND NAMES FOR: B:EJMO  
1. PESO 2. ESTATURA

OPTIONS: A: SCREEN OUTPUT  
B: PRINTER OUTPUT WITH FORMFEEDS  
C: PRINTER OUTPUT WITHOUT FORMFEEDS  
D: TEXT FILE OUTPUT  
  
E: OUTPUT PRINTER SET-UP CODES  
F: CHANGE PRINTER WIDTH. CURRENT VALUE: 80

ENTER: OPTION: \_  
-----

De lo cual se entiende que **A** sirve para ver los datos en pantalla, **B** ó **C** permiten imprimirlos (si se cuenta con impresora) y **D** se usa para hacer un archivo o mandar datos a un archivo de texto, lo cual permite, por ejemplo, imprimirlo en una tesis. Para el uso de las opciones **E** y **F** se recomienda consultar el manual original.

Para efectos de este manual, siempre que se requiera la orden de salida de datos, se hará a la pantalla, es decir, se opta por la letra **A** (SCREEN OUT PUT), en la orden de listar todos los casos (opción **A**) o sólo algunos en particular (opción **B**), se elige opción **A** (OUTPUT ALL CASES) y en la orden de listar todas las variables (**A**) o sólo algunas en particular (**B**) se elige también la opción **A**, para así poder ver o utilizar los archivos completos.

Inmediatamente después, la máquina pregunta con cuántos decimales se desea sean presentados los datos en pantalla. Si se da "Enter" sin indicar un número, la máquina por de faul presupone que se desean 2 decimales (pero puede anotarse el número que se desee, por ejemplo 0 y se da enter) y enseguida aparecen desplegadas en pantalla 2 columnas (PESO Y ESTATURA), con los datos que contiene cada una de esas variables, como lo muestra la fig. 8.

FIG.8

	PESO	ESTATURA
1.	50	155
2.	65	169
3.	80	184
4.	84	185
5.	72	178
6.	55	155

## 2.2 CORRECCION DE DATOS DE UN ARCHIVO.

Para modificar o corregir un archivo de datos, estando en el MENU PRINCIPAL (fig. 1), se selecciona la letra **A** (DATA MANAGEMENT SUBSYSTEM) y de este se elige la opción **C** (EDIT DATA). La máquina pregunta el nombre del archivo que se desea modificar, ofreciendo como archivo abierto (OPEN FILE) el último archivo con el que se halla trabajado y para utilizarlo, sólo es necesario dar "Enter", pero si no se desea trabajar con el archivo abierto, se anota el nombre del archivo que se requiera, sin olvidar indicar el drive en que se encuentra ese archivo, se

da "Enter" e inmediatamente la máquina pide se le indique el número de caso que se desea corregir y al hacerlo la máquina muestra el nombre de la variable y el dato que se tiene en ese caso, si no se desea modificarlo solamente se da "Enter" para conservarlo, pero si se desea cambiar, simplemente se anota el número que va a sustituir al anterior y se da "Enter", la máquina pregunta entonces si las correcciones hechas estan bien y se tecllea Y (yes) o N(no), según sea el caso. Si se selecciona la letra Y, la máquina pregunta si se desea corregir algún otro caso, pero si se selecciona la letra N, ofrece el número de caso que se tenía anteriormente para poder ser corregido (fig. 9).

FIG. 9

-----  
ENTER: OPTION: C

-----DATA MANAGEMENT SUBSYSTEM-----

C. EDIT.

OPEN FILE: B:EJMO (PRESS 'RETURN' TO USE OPEN FILE)  
(presione "Enter" para usar archivo abierto)

ENTER: FILE NAME: B:EJMO

HEADER DATA FOR: B:EJMO LABEL: PESO Y ESTATURA DE ALUMNOS  
NUMBER OF CASES: 6 NUMBER OF VARIABLES: 2

ENTER: CASE NO. TO BE EDITED: \_

EDIT ANOTHER CASE (Y,N)? \_  
-----

Cuando se indica que no se desea corregir algún otro dato, la máquina regresa al submenú del SUBSISTEMA DE MANEJO DE DATOS que es la opción A del menú principal.

### 2.3 RENOMBRANDO ARCHIVOS.

Si se selecciona la opción **D** (RENAME FILE / EDIT HEADER) del submenú del **SUBSISTEMA DE MANEJO DE DATOS**, es posible renombrar cualquier archivo que se halla elaborado con **MICROSTAT**. Inicialmente la máquina pide el drive y nombre del archivo que se va a renombrar, al proporcionarlo, la máquina lo muestra en pantalla, así como su etiqueta (si contiene), el número de casos que tiene, el número de variables, el nombre de las variables y la opción de confirmar si se va a renombrar el archivo o no. Si se selecciona la letra **Y**, la máquina pide el nuevo nombre con el que se va a llamar el archivo, se indica, se da "Enter" y la máquina muestra que ha quedado aceptado el nuevo nombre, así como los otros archivos con extensión **MSD** que se tengan, que son la extensión que por de faul el paquete **MICROSTAT** pone a todos sus archivos; también indica el espacio libre que se tiene en el diskette y permite renombrar algún otro archivo, preguntando si se desea hacerlo o no, si se escoge la letra **Y**, pide el nuevo nombre del archivo por corregir, pero si se elige **N**, la máquina sale de ese menú y regresa al **SUBSISTEMA DE MANEJO DE DATOS**.

**2.3.1** Si se desea **corregir el encabezado**, cambiar la etiqueta de un archivo, adicionársela si no la tiene e incluso cambiar el nombre de las variables, después de anotar el nombre del archivo con que se va a trabajar y aparece en pantalla su etiqueta, número de casos y variables, se opta por la letra **N** (no renombrar el archivo), entonces la máquina pregunta si se desea editar el encabezado y si se escoge la letra **Y** se podrán realizar los cambios mencionados.

## 2.4 DIRECTORIO DE ARCHIVOS.

La opción **E** (FILE DIRECTORY) del SUBSISTEMA DE MANEJO DE DATOS, permite desplegar en pantalla el directorio de archivos contenidos en un diskette. Al elegir esta opción, aparece en pantalla un submenú compuesto de 6 opciones como el de la fig. 7. Al seleccionar la opción **A**, de salida en pantalla, la máquina pregunta si sólo se desean ver los archivos de MICROSTAT o cualquier tipo de archivo que se encuentre. Si se opta únicamente por los hechos con MICROSTAT, se debe indicar a la máquina el drive en que se encuentran, para que despliegue en pantalla todos los nombres de esos archivos e indique el espacio que tiene libre, además de ofrecer tres nuevas opciones: **A**) si se desea lista de detalles de ese directorio, **B**) preguntas a la máquina para eliminar archivos y **C**) para terminar (fig. 10).

FIG. 10

-----  
ENTER: OPTION: A

OPTIONS: A. MICROSTAT FILES ONLY  
          B. OTHER TYPES OF FILES

ENTER: OPTION: A

.MSD FILES ON DRIVE B:

EJMO

SPACE REMAINING: # BYTES

OPTIONS: A. DETAILED DIRECTORY LISTING  
          B. QUERY FOR DESTROY  
          C. [Terminate]

ENTER: OPTION: \_  
-----

## 2.5 BORRAR ARCHIVOS.

La opción **F** (DESTROY FILES) del SUBSISTEMA DE MANEJO DE DATOS, permite borrar archivos que ya no se desea conservar. Al elegirla, la máquina pregunta si se desea o no ver el directorio, por medio de las letras **Y** o **N**, lo cual es recomendable hacer para evitar equivocaciones y borrar algún archivo de importancia. Al dar la opción **Y**, la máquina pide se le indique el drive donde se halla el diskette que se desea consultar, al hacerlo, aparecen en pantalla todos los archivos que están escritos con MICROSTAT, factibles de ser borrados, indicando el espacio que se tiene disponible en ese diskette y preguntando de nueva cuenta si se desea ver el directorio, a lo que se responde negativamente, el programa pide entonces que se le indique el nombre del archivo que se desea borrar.

Si no desea borrar archivo alguno, se anota "End" (fin), pero si se decide destruir o borrar algún archivo, se debe anotar el drive y el nombre y dar "Enter", inmediatamente la máquina vuelve a pedir el nombre de algún otro archivo por borrar, de esta manera es factible destruir al mismo tiempo varios archivos. Cuando ya no se desea borrar algún otro archivo, se tecléa "End" (fin) y la máquina pregunta si se está realmente seguro de los archivos por borrar, confirmando con la letra **Y**.

La máquina procede a borrar los archivos indicados, manda un mensaje cuando ha terminado y vuelve a preguntar si se desea ver el directorio de archivos, lo cual es recomendable hacer para ver el efecto de las órdenes dadas (fig. 11). Obviamente ya no aparecen en pantalla los nombres de los archivos borrados, sólo

aparecerán aquellos que se ha decidido dejar y la indicación del espacio disponible en el diskette, preguntando si se desea ver de nuevo el directorio, se indica que no y entonces la máquina vuelve a preguntar si se desea borrar más archivos, de nuevo se indica que no e inmediatamente la máquina regresa al submenú del SUBSISTEMA DE MANEJO DE DATOS.

FIG. 11

```
-----  
ENTER: OPTION: F  Module being loaded...  
-----DATA MANAGEMENT SUBSYSTEM -----  
  
F. DESTROY FILE(S):  
  
VIEW DIRECTORY (Y,N)? YES  
  
-E  
  
ENTER DRIVE A, B,...: B  
  
.MSD FILES ON DRIVE B:  
  
EJMO  
  
SPACE REMAINING: # BYTES  
  
VIEW DIRECTORY (Y,N)? NO  
  
ENTER: NAMES OF FILES TO BE DESTROYED ('END' TO STOP ENTRY):  
  
ENTER: FILENAME: EJMO  
ENTER: FILENAME: END  
  
ARE YOU SURE YOU WANT THESE FILES DESTROYED (N,Y)? YES  
  
THE ABOVE FILES HAVE BEEN DESTROYED.  
  
VIEW DIRECTORY (Y,N)? YES  
  
.MSD FILES ON DRIVE B:  
  
SPACE REMAINING: # BYTES  
  
VIEW DIRECTORY (Y,N)? N  
  
DESTROY MORE FILES (Y,N)? N  
-----
```

## 2.6 SELECCIONAR, TRANSFORMAR O RECODIFICAR UN ARCHIVO.

La opción G (RECODE/TRANSFORM/SELECT) del SUBSISTEMA DE MANEJO DE DATOS permite seleccionar, transformar o recodificar un archivo. Si se elige esta opción, la máquina le ofrece dos alternativas: A. trabajar con un sólo archivo ó B. trabajar con dos archivos. Si se opta por la letra A, inmediatamente la máquina pregunta el nombre del archivo con el que se va a trabajar. En el momento que se anota el nombre de dicho archivo y se da "Enter", la máquina despliega una vez más el número de casos que contiene, el nombre del archivo, el número de variables, si tiene una etiqueta también lo muestra, así como los nombres de las variables que el archivo contiene y pregunta si se desea crear alguna nueva variable con este comando. Si se teclea la letra Y, para confirmar, inmediatamente la máquina despliega un código de transformación, que es una serie de opciones que van de la letra A a la Z, permitiendo seleccionar la que se requiera (fig. 12).

FIG. 12

-----  
ENTER: OPTION: G Module being loaded...

-----DATA MANAGEMENT SUBSYSTEM-----

G: RECODE/TRANSFORM/SELECT

OPTION: A. ONE INPUT FILE  
          B: TWO INPUT FILES

ENTER: OPTION: A

OPEN FILE: B:EJMO (PRESS 'RETURN' TO USE OPEN FILE)

ENTER: FILE NAME: B:EJM1

E

HEADER DATA FOR: B:EJM1 LABEL: PRIMER EJEMPLO PARA MICROSTAT  
NUMBER OF CASES: 6 NUMBER OF VARIABLES: 1

VARIABLE NUMBERS AND NAMES FOR: B:EJM1

1. PESO

WILL NEW VARIABLES BE CREATED WITH RECODE/TRANSFORMATIONS (Y,N)? YES

TRANSFORMATION CODES:

A. 1 / X	L. X1 + X2	T. CASE NO.
B. LOG (X)	M. X1 - X2	U. COPY
C. LN (X)	N. X1 * X2	V. SCALING
D. EXP (X)	O. X1 / X2	
E. X ^ a	P. SUMX1:X2	W. DUMMY
F. a + b*X		X. RECODE
G. Z-TRANS	Q. RND NO.	
H. ABS (X)	R. RND INT	Y. [REVIEW]
I. ROUND (X)	S. RND NORM	Z. [EXIT]
J. TRUNC (X)		
K. FRAC (X)		

ENTER MENU SELECTION FOR RECODE/TRANSFORMATION NO. 1. (MAX= 250).  
-----

La máquina asume que se realizarán varias transformaciones, por lo que creará varias nuevas variables. La primera será la opción uno y cada vez que se termine de seleccionar un código, cambia al número siguiente. Es importante notar que en este menú de códigos de transformaciones, se tienen varias opciones, que

van desde transformar una variable a su recíproco que consiste en dividir uno entre esa variable (opción A), transformar una variable a un logaritmo en base diez (opción B), obtener el logaritmo natural de esa variable (opción C), usar esa variable como un exponente (opción D), usar a la variable elevada a una potencia  $n$  (opción E), linearizar una variable por medio de una regresión lineal simple (opción F), sumar dos variables, restarlas, multiplicarlas y hasta dividir las, etc. Para una mayor información del uso de este comando, es recomendable consultar el manual original, señalando únicamente que la letra Z permite salir de este comando.

A continuación se ofrecen algunos ejercicios para mostrar como usar este comando.

**2.6.1** Tras haber seleccionado la opción A, para trabajar con un solo archivo dentro del comando G del SUBSISTEMA DE MANEJO DE DATOS, cuando la máquina pregunta el nombre del archivo con el que se va a trabajar, se anota el drive en que se encuentra y **EJM1**, se da "Enter" y la máquina muestra como encabezado que el archivo **EJM1** tiene una etiqueta llamada "PRIMER EJEMPLO PARA MICROSTAT", con 6 casos y una sola variable, llamada **PESO**, numerada con el número 1, fig. 13.

-----  
G. RECODE/TRANSFORM/SELECT

OPTIONS: A. ONE INPUT FILE  
B. TWO INPUT FILES

ENTER: OPTION: A

OPEN FILE: B:EJMO (PRESS 'RETURN' TO USE OPEN FILE)

ENTER: FILE NAME: B:EJM1

HEADER DATA FOR: B:EJM1 LABEL: PRIMER EJEMPLO PARA MICROSTAT  
NUMBER OF CASES: 6 NUMBER OF VARIABLES: 1

VARIABLE NUMBERS AND NAMES FOR: B:EJM1  
1. PESO

WILL NEW VARIABLES BE CREATED WITH RECODE/TRANSFORMATION (Y,N)? \_  
-----

A continuación pregunta si se desea crear una nueva variable, se elige la letra Y e inmediatamente muestra el código de transformación señalado en la fig. 12. Se elige el código B y la máquina pregunta si se desea hacer una transformación simple o utilizando un rango, se escoge la opción B (simple) y enseguida pregunta si se va a crear una nueva variable o se va a reemplazar la existente, eligiendo la opción N, que indica "nueva variable".

La máquina muestra que va a usar la función de **logaritmo de X** y la variable a utilizar es la opción número 1 (única que existe) y pide el nombre de la variable número 2 que se va a crear, que va a ser denominada "LOGPESO", con el objeto de identificarla mejor. Al dar "Enter", la máquina pregunta si los nombres son correctos y se debe escribir Y o N, según sea el caso, como se muestra en la fig. 14.

FIG. 14

-----  
ENTER: CODE: B

VARIABLE NUMBERS AND NAMES FOR: B:EJMI  
1. PESO

-----  
LOG (X) , VAR. NO.: 1

ENTER NAMES FOR VARIABLES OR 'RETURN' TO KEEP ORIGINAL NAMES.

[Transformation = LOG (X) ]

ENTER: NAME FOR VARIABLE 2, PESO : LOGPESO

NAMES OK (Y,N)? \_  
-----

Al anotar que los nombres son correctos, la máquina vuelve a mostrar el código de transformación de la fig. 12 y pregunta el código de una segunda transformación, ahora se elige la letra C y una vez más la máquina pregunta si se hará una transformación simple o de un rango de variables. Nótese que también ofrece la variable 2, denominada LOGPESO que es la que se acaba de crear. Se elige la opción S y se indica que se desea crear una nueva variable (N). Inmediatamente después, pide el número de variable que se va a transformar a logaritmos naturales, se elige la opción 1, se da "Enter" y la máquina pregunta el nombre de la variable 3, ofreciendo como opción el nombre "PESO, pero este nombre lo tiene la variable 1, por lo tanto se utiliza el nombre LNPEO y al dar "Enter", la máquina pregunta si el nombre es correcto (fig. 15).

FIG. 15

-----  
ENTER MENU SELECTION FOR RECODE/TRANSFORMATION NO. 3. (MAX= 250).  
ENTER: CODE: C

VARIABLE NUMBERS AND NAMES FOR: B:EJMI  
1. PESO

VARIABLE NUMBERS AND NAMES FOR: NEW VAR  
2. LOGPESO

-----  
LN (X) , VAR. NO.: 1

ENTER NAMES FOR VARIABLES OR 'RETURN' TO KEEP ORIGINAL NAMES.

{Transformation = LN (X)

ENTER: NAME FOR VARIABLE 3, PESO :LNPESO

NAMES OK (Y,N)? Y  
-----

Al teclear la letra Y, una vez más la máquina ofrece el código de transformaciones y ahora se utiliza la opción 2, para indicar al programa que se desea terminar con las transformaciones. La máquina pregunta si las transformaciones elegidas son correctas y se teclea la letra Y, entonces pregunta el programa el número de variables con el que se formará el nuevo archivo, que en este caso es el 3. Ahora, la máquina ofrece un menú de opciones, que va de la letra A a la E, como lo muestra la fig. 16.

FIG. 16

-----  
OPTIONS: A. OUTPUT ALL CASES (6)  
          B. OUTPUT SUBSET OF CASES  
          C. SELECT INDIVIDUAL CASES  
          D. SELECT BY VALUE OF KEY VARIABLE  
          E. EXCLUDE BY VALUE OF KEY VARIABLE

ENTER: OTION: A

E OPEN FILE: B:EJM1 (PRESS RETURN TO USE OPEN FILE)

ENTER: NAME OF OUTPUT FILE: B:EJM1

ENTER: FILE LABEL: PESO DE ALUMNOS

PRESS ANY KEY TO CONTINUE.  
-----

Por ahora, se selecciona la letra **A**, que permite trabajar con todos los casos y el programa demanda enseguida el nombre del nuevo archivo por crear, que en este caso será **EJM2**, se da "Enter" y pregunta si se desea anotar alguna etiqueta, que para este caso será "**PESO LOGPESO LNPESO**", con lo cual se ejemplifica cómo la variable sencilla **PESO**, se puede transformar al logaritmo en base 10 y también al logaritmo natural (logaritmo neperiano). Al dar "Enter", la máquina pide teclear cualquier letra e inmediatamente pide seleccionar el orden o la secuencia en el que deben aparecer estas tres variables. Se seleccionan los números 1, 2 y 3, en ese orden e inmediatamente la máquina muestra que la variable 1 es **PESO**, la 2 es **LOGPESO** y la variable 3 es **LNPESO**, además pregunta si la secuencia es correcta y si lo es, se teclea la letra **Y**. Enseguida el programa manda un mensaje indicando que el archivo **EJM2** ha sido creado y pregunta si se desea destruir o borrar el archivo original con el que se iniciaron las transformaciones, es decir, el **EJM1**, anotando en

este caso que no (N). Inmediatamente después, la máquina procede a guardar el nuevo archivo y regresa al SUBSISTEMA DE MANEJO DE DATOS.

2.6.2 En el siguiente ejemplo, se utilizan 2 archivos para crear uno solo, seleccionando variables de ambos. Estando en la orden G de la fig. 2, cuando la máquina ofrece las opciones A y B, se elige la opción B para manejo de 2 archivos y enseguida pregunta el nombre del primer archivo que se va a utilizar, que en este caso es **EJMO**, al dar "Enter" la máquina muestra dicho archivo, su etiqueta e indica que contiene 2 variables con 6 casos cada una, la variable 1 se llama PESO y la 2 se llama ESTATURA (fig. 17).

FIG. 17

```
-----  
OPTIONS:  A. ONE INPUT FILE  
          B. TWO INPUT FILES  
  
ENTER: OPTION: B  
  
ENTER: FILE NAME: B:EJMO  
  
HEADER DATA FOR: B:EJMO LABEL: PESO Y ESTATURA DE ALUMNOS  
NUMBER OF CASES: 6 NUMBER OF VARIABLES: 2  
  
VARIABLE NUMBERS AND NAMES FOR: B:EJMO  
1. PESO 2. ESTATURA  
  
SELECT SECOND INPUT FILE.  
  
ENTER: FILE NAME: _____  
-----
```

De la misma forma, pide el nombre del segundo archivo que se

va a utilizar y en este caso es **EJM2** y la máquina indica que ese archivo tiene como etiqueta "PESO LOGPESO LNPESO", con 3 variables y 6 casos en cada una, siendo la variable 1 PESO, la variable 2 LOGPESO y la variable 3 es LNPESO. También pide se le indique si se desea crear una nueva variable, indicando que no, como se muestra en la fig. 18.

FIG. 18

-----  
SELECT SECOND INPUT FILE.

ENTER: FILE NAME: B:EJM2

HEADER DATA FOR: B:EJM2 LABEL: PESO LOGPESO LNPESO  
NUMBER OF CASES: 6 NUMBER OF VARIABLES: 3

VARIABLE NUMBERS AND NAMES FOR: B:EJM2  
1. PESO 2. LOGPESO 3. LNPESO

WILL NEW VARIABLES BE CREATED WITH RECODE/TRANSFORMATIONS  
(Y,N)? NO  
-----

Después aparece en pantalla el número y nombres de las variables del archivo **EJM0**, así como los del archivo **EJM2** y el programa pide se anote el número de variables que se van a utilizar, en este caso son 3, por lo tanto se anota el número 3. Al dar "Enter", aparece un submenú del cual se utiliza la opción **A**, para indicar a la máquina que se desea utilizar todos los casos. Seguido a esto, la máquina pide el nombre del nuevo archivo que se va a formar, es decir, el que va a contener las 3 variables que se acaban de seleccionar. Aquí se tecléa **EJM3** como nuevo nombre del archivo (recordando anteponer el drive al que se manda el archivo), al dar "Enter" la máquina pregunta que

etiqueta se desea poner a ese nuevo archivo, entonces, se teclea "COMBINACION EJMO-EJM2" y al dar "Enter" la máquina pide se pise cualquier tecla, fig. 19.

FIG. 19

-----  
VARIABLE NUMBERS AND NAMES FOR: B:EJMO

1. PESO                    2. ESTATURA

VARIABLE NUMBERS AND NAMES FOR: B: EJM2

3. PESO                    4. LOGPESO                5. LNPESO

ENTER: NUMBER OF VARIABLES TO BE OUTPUT (MAX = 255): 3

OPTIONS:    A. OUTPUT ALL CASES(6)  
              B. OUTPUT SUBSET OF CASES  
              C. SELECT INDIVIDUAL CASES  
              D. SELECT BY VALUE OF KEY VARIABLE  
              E. EXCLUDE BY VALUE OF KEY VARIABLE

ENTER: OPTION: A

E OPEN FILE: B:EJMO (PRESS RETURN TO USE OPEN FILE)

ENTER NAME OF OUTPUT FILE: B:EJM3

ENTER FILE LABEL: COMBINACION EJMO-EJM2

PRESS ANY KEY TO CONTINUE  
-----

al teclear cualquier letra, inmediatamente se observa en pantalla que se tienen 5 variables, 2 en el archivo EJMO y 3 en el archivo EJM2, procediendo entonces a seleccionar las variables que van a ser utilizadas para conformar el nuevo archivo: la variable 2 en primer lugar, en segundo la variable 4 y la número 5 en tercer lugar, dando "Enter" al ir anotando cada una. Con esto se indica a la máquina que se desea crear un nuevo archivo con las variables 2, 4 y 5 (obsérvese que la variable 2

pertenece al archivo **EJMO** y que las variables 4 y 5 son del archivo **EJM2**). El nuevo archivo que se ha creado, llevará en el orden señalado 1, 2 y 3, las variables **ESTATURA**, **LOGPESO** y **LNPEO**; la máquina pregunta si la secuencia es correcta y si lo es, se indica con la letra **Y** y después envía un mensaje, indicando que el nuevo archivo **EJM3** ha sido creado, pregunta también si se desea borrar el archivo **EJMO**, indicando que no, por medio de **N** y lo mismo para no borrar el archivo **EJM2** (fig. 20).

FIG. 20

```
-----  
VARIABLE NUMBERS AND NAMES FOR: B:EJMO  
1. PESO          2. ESTATURA  
  
VARIABLE NUMBERS AND NAMES FOR: B:EJM2  
3. PESO          4. LOGPESO    5. LNPEO  
  
SELECT 3 VARIABLE NUMBERS IN ANY SEQUENCE.  
  
1:  2. ESTATURA  
2:  4. LOGPESO  
3:  5. LNPEO  
  
OUTPUT SEQUENCE OK (Y,N)? YES  
FILE: B:EJM3 IS NOW BEING OUTPUT. . .  
  
DESTROY INPUT FILE B:EJMO (N,Y)? NO  
DESTROY INPUT FILE B:EJM2 (N,Y)? NO  
-----
```

Después de un breve momento, la máquina regresa al **SUBSISTEMA DE MANEJO DE DATOS** y puede ser comprobada la existencia del archivo que se acaba de crear, utilizando la opción **E(FILE DIRECTORY)** de ese submenú e incluso, pueden listarse los datos de

ese archivo usando la opción **B** de éste mismo submenú. Si se opta por listar los datos del archivo **EJM3**, estos deberán aparecer como se muestra en la fig. 21.

FIG. 21

```
-----  
      ESTATURA           LOGPESO           LNPESO  
1. 155.00              1.70              3.91  
2. 169.00              1.81              4.17  
3. 184.00              1.90              4.38  
4. 185.00              1.92              4.43  
5. 178.00              1.86              4.28  
6. 155.00              1.74              4.01
```

PRESS ANY KEY TO CONTINUE  
-----

**2.6.3** Por último, el siguiente ejemplo consiste en seleccionar **2** archivos para transformar o crear una nueva variable a partir de las variables existentes en ellos. Nuevamente, cuando el programa ofrece la alternativa de trabajar con uno o dos archivos, dentro del comando **G** de selección y transformación de un archivo (fig. 2), se utiliza la opción **B** para manejo de dos archivos. Cuando la máquina pide el nombre del primer archivo, se tecldea **EJMO** y **EJM3** para el segundo; cuando pregunta si se desea crear una nueva variable, se indica que sí con la letra **Y**. Cuando aparece el código de transformación para la primer variable que se va a crear, se utiliza la letra **L**. Cuando el programa pregunta cuál será la variable **x1** se elige el número **2** y el número **3** para variable **x2**. Nótese que lo que se desea hacer es sumar las variables **ESTATURA** del archivo **EJMO** y del archivo **EJM3**; también debe observarse que juntando todas las

variables de los dos archivos, el total es de 5, por lo que la máquina pide, en un siguiente paso, el nombre de la variable 6, ofreciéndole el nombre **ESTATURA** como opción, sin embargo, el nombre que va a utilizarse es **ESTATUR2**, para indicar que es una variable compuesta (fig. 22).

FIG. 22

```
-----  
ENTER MENU SELECTION FOR RECODE/TRANSFORMATION NO.1 (MAX= 250).  
ENTER: CODE: L  
  
VARIABLE NUMBERS AND NAMES FOR: B:EJMO  
1. PESO 2. ESTATURA  
  
VARIABLE NUMBERS AND NAMES FOR :B:EJM3  
3. ESTATURA 4. LOGPESO 5. LNPESO  
  
-----  
X1 + X2 , VARIABLE X1: 2, VARIABLE X2: 3  
ENTER NAMES FOR VARIABLE OR 'RETURN' TO KEEP ORIGINAL NAMES.  
  
[Transformation = X1 + X2]  
  
ENTER: NAME FOR VARIABLE 6, ESTATURA:ESTATUR2  
  
NAMES OK (Y,N,)? Y  
-----
```

Después pregunta si el nombre es correcto y si lo es, se tecléa la letra **Y**. Enseguida el programa regresa al código de transformación y ahora se elige la opción **N** y se pide el nombre de la primera variable **x1**, se selecciona el número 4 y se da "Enter" y para la segunda variable se selecciona el 5 y se da "Enter", a continuación, la máquina pide el nombre de la variable 7 y se tecléan las letras **LOGLN** (variable compuesta de logaritmo en base 10 multiplicada por el logaritmo natural), se da "Enter" y se tecléa la letra **Y** cuando la máquina pregunta si el nombre es

correcto, como se ve en la fig. 23.

FIG. 23.

-----  
ENTER MENU SELECTION FOR RECODE/TRANSFORMATION NO. 2 (MAX=250).  
ENTER: CODE: N

VARIABLE NUMBERS AND NAMES FOR: B:EJMO  
1. PESO            2. ESTATURA

VARIABLE NUMBERS AND NAMES FOR: B:EJM3  
3. ESTATURA    4. LOGPESO    5. LNPESO

VARIABLE NUMBERS AND NAMES FOR: NEW VAR  
6. ESTATUR2

-----  
X1 x X2 ,    VARIABLE X1: 4,    VARIABLE X2: 5  
ENTER NAMES FOR VARIABLES OR 'RETURN' TO KEEP ORIGINAL NAMES.

[Transformation = X1 x X2]

ENTER: NAME FOR VARIABLE 7: LOGLN

NAMES OK (Y,N)? Y  
-----

Enseguida el programa regresa de nueva cuenta al código de transformaciones, del que se elige la letra Z, para indicar que se desea salir o que ya se ha terminado de hacer la selección. La máquina pregunta si la información es correcta y se indica que sí, con la letra Y. Ahora la máquina ofrece en pantalla las 7 variables de las que se dispone y se pueden seleccionar de todas ellas las que se desea que conformen el nuevo archivo. Llegado este punto, se debe indicar cuantas variables van a ser utilizadas, anotando el número 2 para indicar que solamente se desean 2 y por medio de un submenú con 5 opciones, pregunta si se desea utilizar todos los casos o algunos en particular. Se elige la opción A y enseguida pide el nombre del nuevo archivo, que

será **EJM4**, sin olvidar teclear antes el drive de archivos. Al dar "Enter", la máquina pide la etiqueta de ese nuevo archivo y se anota "COMBINACION DE EJMO-EJM3", se da "Enter" y pide pisar cualquier tecla, al hacerlo, la máquina pide se le indique la secuencia u orden en que se desea que aparezcan las 2 variables que se van a seleccionar, eligiendo las variables 6 y 7, que son las dos variables nuevas que se crearon. Enseguida, la máquina pregunta si la secuencia de salida es correcta, indicando que sí con la letra Y y aparece entonces el mensaje de que el archivo **EJM4** ha sido creado. Después el programa pregunta si se desea borrar los archivos originales (EJMO y EJM3), indicando que no, con la letra N.

Después de un breve momento la máquina regresa al submenú del SUBSISTEMA DE MANEJO DE DATOS. Una vez más, se puede corroborar o ver el nuevo archivo que acaba de ser creado, usando la opción B, para listar los datos que contiene ese archivo. Si se utiliza esta opción, los datos deben aparecer en pantalla como lo muestra la fig. 24.

FIG. 24.

---

	ESTATUR2	LOGLN
1.	310	6.65
2.	338	7.57
3.	368	8.34
4.	370	8.53
5.	356	7.94
6.	310	6.97

PRESS ANY KEY TO CONTINUE

---

Como puede observarse en los 3 ejemplos anteriores, la

opción A (DATA MANAGEMENT SUBSYSTEM) del MENU PRINCIPAL en su opción G (TRANSFORM /RECODE /SELECT), ofrece un potente comando de transformación.

Se recuerda que para salir de este submenú, se utiliza la opción O e inmediatamente la máquina regresa al MENU PRINCIPAL (fig. 1).

## C A P I T U L O 3

### ESTADISTICA DESCRIPTIVA

Por medio de la opción B del MENU PRINCIPAL del paquete estadístico MICROSTAT, es posible dar rápida solución a numerosos problemas estimando los estadísticos descriptivos de archivos previamente contruidos, ya que por medio de esta opción puede obtenerse: media aritmética, desviación estándar, varianza y coeficiente de variación de una muestra y de la población, entre otros cálculos. Después de elegida la opción B, la máquina pide el nombre del archivo con el que se desea trabajar. A continuación se describe el archivo llamado DESCRIP1, que consta de una sola variable con 70 casos y tiene como etiqueta "# DE GAZAPOS POR CAMADA"

#### EJEMPLO 3

---

En un módulo de conejos de cierta granja de la Cd. de Irapuato, Gto., se obtuvo el número de gazapos vivos por camada de 70 hembras, arrojando los datos que contiene el archivo DESCRIP1.

(López, B. B., 1994)

---

De las siguientes opciones que aparecen en pantalla, se

selecciona la letra **A** (INPUT ALL CASES) para indicar a la máquina que se desea trabajar con todos los casos de esa variable y así, la máquina ofrece 3 opciones: **A** para trabajar con un método corto en donde sólo se estime la media, la desviación estándar, el dato mínimo y el dato máximo, **B** para trabajar con una opción extendida y **C** para terminar (fig. 25).

FIG. 25

```
-----  
ENTER: FILE NAME: B:DESCRIP1  
  
HEADER DATA FOR: B:DESCRIP1 LABEL: # DE GAZAPOS POR CAMADA  
NUMBER OF CASES: 70 NUMBER OF VARIABLES:1  
VARIABLE NUMBERS AND NAMES FOR: B:DESCRIP1  
1. #GAZAPOS  
  
OPTIONS: A. INPUT ALL CASES  
B. INPUT SUBSET OF CASES  
  
ENTER: OPTION: A  
  
OPTIONS: A. SHORT FORM OUTPUT (MEAN, STD. DEV., MIN, MAX)  
B. EXTENDED OUTPUT OF SELECTED VARIABLES  
C. [Terminate]  
  
ENTER: OPTION: _  
-----
```

### 3.1 METODO CORTO.

Primeramente se utilizaremos la opción A (SHORT FORM OUTPUT) y del menú siguiente se selecciona también la opción A para que la máquina muestre los resultados en pantalla, enseguida pide un título para esos resultados, que en este caso será "ESTADISTICA DESCRIPTIVA", se da "Enter" y después de un momento, la máquina muestra en pantalla los resultados de las operaciones que hizo con los datos de ese archivo, indicando que ha usado una variable que se llama "#GAZAPOS", cuyo tamaño es N=70, la media aritmética es 7.7286, la desviación estándar es igual a 1.9849, el dato mínimo que existe en ese archivo es 2 y el máximo 13, como se muestra en la fig. 26.

FIG. 26

```
-----  
-----DESCRIPTIVE STATISTICS-----  
HEADER DATA FOR: B:DESCRIPT1 LABEL:# DE GAZAPOS POR CAMADA  
NUMBER OF CASES: 70 NUMBER OF VARIABLES: 1
```

#### ESTADISTICAS DESCRIPTIVAS

NO.	NAME	N	MEAN	STD.DEV.	MINIMUM	MAXIMUM
1	#GAZAPOS	70	7.7286	1.9849	2.0000	13.0000

PRESS ANY KEY TO CONTINUE  
-----

Al oprimir cualquier tecla, la máquina ofrece nuevamente tres opciones, donde A es para repetir la misma operación, B

permite realizar más operaciones con este u otro archivo, es decir, otros cálculos y **C** es para salir de este submenú.

Como siguiente ejercicio, se utiliza la opción **B**, para hacer más análisis y cuando el programa pide el nombre del archivo con que se va a trabajar, solamente se da "Enter" para utilizar el archivo abierto (**DESCRIP1**). Enseguida pregunta si se desea trabajar con todos los casos o sólo con algunos casos subsecuentes, eligiendo la letra **A** (**INPUT ALL CASES**) e inmediatamente después requiere se anote la opción del método con el que se va a trabajar (**A** corto, **B** largo).

### **3.2 METODO EXTENDIDO.**

Al seleccionar la opción **B** (**EXTENDED OUTPUT OF SELECTED VARIABLES**), la máquina muestra en pantalla un menú con seis opciones, de las que se elige la **A** para que el programa muestre en pantalla los resultados que se obtengan. También pide el título con el cual mostrar los resultados, dando "Enter" únicamente sin hacer anotación alguna. Enseguida se muestra una serie de opciones, de las cuales se seleccionan las deseadas por medio de la letra **Y** que son las siguientes:

- media aritmética
- desviación estándar
- varianza para la muestra y para la población
- coeficiente de variación para la muestra y para la población
- error estándar
- valor máximo

- valor mínimo
- suma de cuadrados
- desviación de la suma de cuadrados
- momentos de la media
- curtosis

indicando a todos que sí, **excepto a curtosis**, señalando que no se desea obtener ese valor.

Después el programa pregunta si se desea hacer una prueba de bondad de ajuste, es decir, hacer la prueba para conocer si la distribución de los datos del archivo que se está usando, se ajustan o no a una distribución normal, a lo cual se contesta afirmativamente; enseguida la máquina pregunta si la selección hecha es correcta y si lo es, se indica Y y en un momento la máquina indica en pantalla el drive en donde se encuentra el archivo con el que se está trabajando, su nombre y etiqueta, que en este caso es, como ya se dijo, es el archivo **DESCRIP1**, con la etiqueta "**# DE GAZAPOS POR CAMADA**", con una sola variable que contiene 70 casos. Enseguida requiere se indique cuál es la variable que se desea utilizar, que en este caso, únicamente puede ser la número 1. Cabe señalar que si el archivo contiene dos o más variables, puede utilizarse cualquiera de ellas individualmente, pero nunca dos variables a la vez, ya que la máquina lo rechazaría. En el momento que se selecciona la variable 1, aparece en pantalla el valor de la media aritmética, que en este caso es 7.728571. Después, pide se presione cualquier tecla y ofrece la desviación estándar de la muestra, que es igual a 1.9848, la varianza de la muestra que es 3.93975, así como el

coeficiente de variación igual a 25.68%, después la máquina pide se presione una tecla para inmediatamente ofrecer los valores de desviación estándar, varianza y coeficiente de variación de la población, que son respectivamente 1.9706, 3.8834 y 25.4982. Al pisar nuevamente una tecla, aparece el valor del error estándar de la media, que es igual a 0.2372, al oprimir otra tecla aparece el valor mínimo que es igual a 2 y el valor máximo que es igual a 13. Al pisar otra tecla la máquina ofrece la suma de todos los datos, la suma de todos los datos al cuadrado, así como la desviación de la suma de cuadrados, que son respectivamente 541, 4452.999 y 271.8428. Finalmete, al pisar una vez más una tecla, la máquina hace la prueba de bondad de ajuste para una distribución normal y manda un mensaje donde indica que "la hipótesis de que la población es normal, con una media de 7.728571, una desviación estándar de 1.98488 puede ser rechazada a un 95% de nivel de confianza". Como el valor de chi cuadrada es igual a 29.429 con 5 grados de libertad y con una probabilidad de que este valor se dé por azar, igual a  $1.866 \times 10^{-5}$ , se concluye que hay evidencia de que la población a la que pertenece la muestra estudiada, no se distribuye normalmente, como se muestra en la fig. 27.

FIG. 27

-----  
ENTER: VARIABLE TO BE OUTPUT (E=End, L=List. VAR. NAMES):  
1. #GAZAPOS

VARIABLE NAME: #GAZAPOS            N = 70

ARITHMETIC MEAN = 7.728571

PRESS ANY KEY TO CONTINUE

SAMPLE STD.DEV. = 1.9848

SAMPLE VARIANCE = 3.93975

COEFFICIENT OF VARIATION = 25.68%

POPULATION STD.DEV. = 1.9706

POPULATION VARIANCE = 3.8834

COEFFICIENT OF VARIATION = 25.49%

STANDAR EEROR OF THA MEAN = 0.2372

MINIMUM = 2

MAXIMUM = 13

SUM = 541

SUM OF SQUARES = 4452.999

DEVIATION S S = 271.8428

THE HIPOTHESIS THAT THE POPULATION IS NORMAL OF MEAN 7.728571 AND  
STD. DEV. 1.98488            CAN BE REJECTED AT THE 95% CONFIDENCE LEVEL

CHI SQUARE= 29.429,            D.F.= 5,            P= 1.866E-05

ENTER: VARIABLE TO BE OUTPUT (E=End, L=List    VAR NAMES):\_

Finalmente, seleccionando la letra E (end) para indicar que se ha finalizado, la máquina da la opción de repetir las operaciones (A); de hacer más análisis (B) ó confirmar que se desea salir, usando para esto último la letra C y así, regresar al MENU PRINCIPAL (fig. 1).

## C A P I T U L O 4

### TABLAS DE FRECUENCIA.

Entre las múltiples aplicaciones del programa MICROSTAT, se halla la posibilidad de realizar fácilmente tablas de frecuencia, para lo cual se utiliza la opción C del menú principal.

Una vez hecha la selección, cuando el programa pide el nombre del archivo de trabajo, como ejercicio se anota DESCRIPi (sin olvidar anteponer el drive en que se encuentra el archivo), al dar "Enter" la máquina ofrece en pantalla el título del archivo que se acaba de seleccionar, la etiqueta, el número de variables, el número de casos que contiene cada una de ellas y su nombre. Así mismo ofrece 3 opciones para continuar trabajando, donde A sirve para construir tablas de frecuencia para datos agrupados, B permite hacer tablas de frecuencia con valores individuales y la opción C permite salir de ese submenú (fig. 28).

FIG. 28

```
-----  
ENTER: OPTION:  C Module being loaded. . .  
-----FREQUENCY DISTRIBUTIONS-----  
ENTER: FILE NAME:  DESCRI1  
HEADER DATA FOR: B:DESCRI1 LABEL: # DE GAZAPOS VIVOS POR CAMADA  
NUMBER OF CASES:  70   NUMBER OF VARIABLES:  1  
VARIABLE NUMBERS AND NAMES FOR: B:DESCRI1  
  1. #GAZAPOS  
OPTIONS: A: GROUPED FREQUENCY DISTRIBUTION  
        B: COUNT INDIVIDUAL VALUES  
        C: [Terminate]  
ENTER: OPTION:  _  
-----
```

#### 4.1 TABLA DE FRECUENCIA PARA DATOS AGRUPADOS.

Al seleccionar la opción **A** del submenú anterior (GROUPED FREQUENCY DISTRIBUTION), el programa pregunta a dónde se desea que salga la información solicitada, se selecciona de nuevo la letra **A**, que indica que la muestre en pantalla. Después la máquina pregunta el número de la variable con el que se hará la distribución de frecuencia, como en este ejemplo en particular sólo se tiene una, debe teclearse el número 1 y dar "Enter", nuevamente la máquina por medio de 2 opciones pregunta si se desea trabajar con todos los casos o solo algunos en especial, seleccionando la letra **A** (todos los casos) para este ejercicio. Después pide el título con el que se desea sean mostrados los resultados, tecleando "GAZAPOS VIVOS POR CAMADA" al dar "Enter" la máquina pide se introduzcan los parámetros para agrupar los datos en una distribución de frecuencia.

Primeramente pide el intervalo con el cuál se harán los grupos, cabe mencionar que este archivo fué trabajado en el tema anterior, por lo que se conoce que el dato menor que contiene es 2 y el dato mayor es 13, por lo tanto, el rango que hay entre el dato menor y el mayor es de 11 unidades. Si se desea construir una distribución de frecuencia apropiada, se deben tener entre 5 y 20 grupos, por consiguiente el rango de 11 debe estar dividido entre ese número. Por ejemplo, un mínimo de 6 grupos es un número apropiado para agrupar esos datos. Al dividir 11 entre 6, es posible estimar el intervalo aproximado que debe tener cada grupo, este intervalo es constante, por lo tanto un valor apropiado puede ser 1.9. Enseguida la máquina pregunta el límite inferior para el primer intervalo y se anota el valor mínimo que es el 2. Después de un breve tiempo la máquina utiliza los 2 parámetros anotados y ofrece en pantalla el cuadro de frecuencia que lleva como título "GAZAPOS VIVOS POR CAMADA", que muestra el límite de clase inferior y superior para cada una de las clases obtenidas, una columna que lleva por título **FRECUENCIA**, donde es posible notar cómo se distribuyen los datos en cada una de las clases o grupos que se acaban de formar, así como el porcentaje que representa cada una de las frecuencias en las clases, la frecuencia acumulada y la frecuencia acumulada relativa expresada en porcentajes.

Al presionar cualquier tecla, la máquina ofrece en pantalla una tabla conteniendo los límites de clase y las frecuencias de cada clase, junto con una gráfica de barras en donde la barra con mayor frecuencia, que es de 33, es la más alta y pertenece a la cuarta clase, teniendo como límite inferior 7.7 y como límite

superior 9.6 (fig. 29).

FIG. 29

-----

GAZAPOS VIVOS POR CAMADA				....CUMULATIVE...	
====CLASS LIMITS====	FREQUENCY	PERCENT	FREQUENCY	PERCENT	
2.00 < 3.90	2	2.86	2	2.86	
3.90 < 5.80	8	11.43	10	14.29	
5.80 < 7.70	17	24.29	27	38.57	
7.70 < 9.60	33	47.14	60	85.71	
9.60 < 11.50	9	12.86	69	98.57	
11.50 < 13.40	1	1.43	70	100.00	
	TOTAL	70			

VARIABLE: 1. #GAZAPOS

====CLASS LIMITS====	FREQUENCY .....
2.00 < 3.90	2: CC
3.90 < 5.80	8: CCCCCCCC
5.80 < 7.70	17: CCCCCCCCCCCCCCCC
7.70 < 9.60	33: CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
9.60 < 11.50	9: CCCCCCCC
11.50 < 13.40	1:

PRESS ANY KEY TO CONTINUE

-----

Al oprimir una tecla, la máquina ofrece de nuevo 3 opciones: con **A** se repiten los cálculos que se acaban de ver (con la posibilidad de mandar los datos a un archivo), **B** permite hacer más cálculos, en donde pueden ser variadas las constantes introducidas anteriormente y la opción **C** es para terminar y salir de ese submenú.

#### 4.2 TABLAS DE FRECUENCIA CON VALORES INDIVIDUALES.

Si se desea hacer una tabla de distribución de frecuencia con datos en forma individual, se utiliza la opción B (COUNT INDIVIDUAL VALUES) del submenú que aparece como en la fig. 28 y enseguida se elige la letra A del submenú de salida de datos (fig. 7), para enseguida tomar la opción uno para indicar que es la variable uno la que se desea trabajar. Después se toma como opción la letra A, que indica se utilicen todos los casos, se omite el título con sólo dar "Enter" y por último la máquina pregunta cuántos valores individuales se desea que el cálculo considere, recordando que el dato menor es 2 y el mayor es 13 para ese archivo y que del 2 al 13, incluyendo el 2, se tienen 12 datos, por lo tanto, se selecciona el número 12. Inmediatamente después, la máquina pide se introduzcan los valores de los datos a considerar, el valor uno debe ser el 2, se da "Enter" e inmediatamente después el valor 2 debe ser el número 3 y así sucesivamente hasta llegar al valor 12 que en este caso corresponde al número 13 (fig. 30).

FIG. 30

-----  
OPTIONS: A. INPUT ALL CASES  
          B: INPUT SUBSET OF CASES

ENTER: OPTION: A

ENTER: JOB TITLE:

HOW MANY INDIVIDUAL VALUES ARE TO BE COUNTED? 12

ENTER: VALUES TO BE COUNTED:

VALUE 1: 2  
VALUE 2: 3  
VALUE 3: 4  
VALUE 4: 5  
VALUE 5: 6  
VALUE 6: 7  
VALUE 7: 8  
VALUE 8: 9  
VALUE 9: 10  
VALUE 10: 11  
VALUE 11: 12  
VALUE 12: 13  
-----

Al dar "Enter", después de una pausa la máquina muestra en pantalla la fig. 31-A que contiene los valores individuales, la frecuencia y porcentaje individuales y las frecuencias y porcentajes acumulados de los datos del archivo DESCRIP1.

FIG. 31-A

-----  
 ENTER: OPTION: A

VARIABLE: 1. #GAZAPOS

===== VALUE =====	FREQUENCY	PERCENT	....CUMULATIVE....	
			FREQUENCY	PERCENT
2.00	1	1.43	1	1.43
3.00	1	1.43	2	2.86
4.00	3	4.29	5	7.14
5.00	5	7.14	10	14.29
6.00	5	7.14	15	21.43
7.00	12	17.14	27	38.57
8.00	18	25.71	45	64.29
9.00	15	21.43	60	85.71
10.00	6	8.57	66	94.29
11.00	3	4.29	69	98.57
12.00	0	.00	69	98.57
13.00	1	1.43	70	100.00
	TOTAL 70	100.00		

PRESS ANY KEY TO CONTINUE.  
 -----

Si se presiona cualquier otra tecla, aparece en pantalla una tabla como la que muestra la fig. 31-B, donde se observan los límites de clase y una gráfica simple de barras.

FIG. 31-B

VARIABLE: 1. #GAZAPOS

```
=====CLASS LIMITS===== FREQUENCY.....
      2.00                      1: C
      3.00                      1: C
      4.00                      3: CCC
      5.00                      5: CCCCC
      6.00                      5: CCCCC
      7.00                     12: CCCCCCCCCCCC
      8.00                     18: CCCCCCCCCCCCCCCCCC
      9.00                     15: CCCCCCCCCCCCCCCC
     10.00                      6: CCCCC
     11.00                      3: CCC
     12.00                      0:
     13.00                      1: C
```

PRESS ANY KEY TO CONTINUE.

Esta selección permite que el paquete MICROSTAT utilice todos los datos que hay en el archivo, que van del 2 al 13.

Por último, al oprimir de nuevo cualquier tecla, la máquina ofrece 3 opciones, donde **A** permite repetir los cálculos, **B** permite realizar más cálculos y **C** es para terminar. En este caso, se elige la opción **C** y la máquina regresa al MENU PRINCIPAL de MICROSTAT. Es importante recordar que con la opción **P** del MENU PRINCIPAL (fig. 1), se puede terminar una sesión de trabajo.

## C A P I T U L O 5

### PERMUTACION Y COMBINACION.

Para trabajar con distribuciones de probabilidad, es necesario tener conocimiento de las técnicas básicas de conteo, como son la PERMUTACION Y COMBINACION. El paquete estadístico MICROSTAT, posee la capacidad de llevar a cabo, de manera cómoda y sencilla, el conteo del número de eventos que satisfacen algún conjunto de condiciones, que es de gran utilidad al realizar el cálculo de la probabilidad de un evento. Para poder realizar estos cálculos, estando en el MENU PRINCIPAL (fig. 1), se utiliza la opciones L o M (en forma indistinta) y se despliega en pantalla un submenú de 3 opciones, donde A permite estimar factoriales, permutaciones o combinaciones, B permite desarrollar algunas distribuciones de probabilidad y C permite salir del submenú (fig. 32).

FIG. 32

-----  
ENTER: OPTION: L o M Module being loaded  
-----

OPTIONS    A. FACTORIAL, PERMUTATIONS, COMBINATIONS  
           B. PROBABILITY DISTRIBUTIONS  
           C. [Terminate]

ENTER: OPTION: \_  
-----

Para explicar el funcionamiento de este submenú, a continuación se presentan algunos ejercicios.

### 5.1 FACTORIAL.

Considerando que se desea calcular el factorial de un número  $N$ , primeramente se elige la opción A (FACTORIAL, PERMUTATIONS, COMBINATIONS) del submenú indicado en la fig. 32. Enseguida aparece un submenú como el que se muestra en la fig. 33, del que se elige la opción A (FACTORIAL).

FIG. 33

```
-----  
ENTER: OPTION: A  
-----FACTORIAL, PERMUTATIONS, COMBINATIONS-----  
  
OPTIONS: A. FACTORIAL  
          B. PERMUTATIONS  
          C. COMBINATIONS  
          D. [Terminate]  
  
ENTER: OPTION: A  
  
      FACTORIAL PROGRAM  
  
      ENTER: N: _  
-----
```

Así entonces, la pantalla muestra el PROGRAMA FACTORIAL y pide se introduzca el valor de  $N$ , es decir, el valor del # del cual se desea estimar el factorial, que será tomado del ejemplo 5.1.

### EJEMPLO 5.1

---

Supóngase que se tienen diez recipientes de medios de cultivo, cada uno de los cuales está inoculado con un organismo diferente. En cuántas formas diferentes pueden colocarse en línea sobre un estante?

(Daniel, W. W., 1983 modificado).

---

Así entonces, será el factorial de 10 el valor que se desea estimar. Al anotararlo se da "Enter" y el paquete pregunta a manera de opciones a dónde se desea que envíe el resultado, utilizando la opción A, para que lo muestre en pantalla y automáticamente se observa que el factorial de 10 es igual a 3'628,800 lo que indica el número de formas diferentes en línea en que pueden ser colocados los 10 recipientes.

Inmediatamente después la máquina pide se pise una tecla y al hacerlo ofrece 3 opciones: A para repetir la misma operación (incluso mandar el resultado a un archivo), B para hacer algún otro cálculo, ya sea con factorial, con combinación o con permutación y C para terminar (fig. 33-A).

FIG. 33-A

---

OPTIONS: A. REPEAT OUTPUT  
B. ANOTHER COMPUTATION  
C. [Terminate]

ENTER: OPTION: \_

---

Si se selecciona la opción B (ANOTHER COMPUTATION), para hacer algún otro cálculo, la máquina envía automáticamente al submenú de la fig. 33.

## 5.2 PERMUTACION.

Para realizar cálculos de PERMUTACION, se elige la opción B del menú que aparece en pantalla (fig 33), enseguida, la máquina pide los valores de N y R, que se obtendrán del ejemplo 5.2

### EJEMPLO 5.2

---

Un Centro Antirrábico ha proporcionado a un M.V.Z. diez carteles diferentes para exhibir en la fachada de su consultorio, pero unicamente tiene espacio para acomodar cinco de ellos. ¿En cuántas formas diferentes los puede disponer?

(Daniel, W. W., 1983 modificado)

---

Tomando en cuenta los datos del ejemplo anterior, como valor N se anota el número 10, se da "Enter" y como valor R el 5 y se da "Enter". El programa muestra entonces el menú de opciones de la fig. 7, del cual se selecciona la opción A para que los resultados sean enviados a pantalla e inmediatamente después se observa que el número de permutaciones (arreglos posibles) de 10 objetos (carteles), tomando (exhibiendo) 5 a un tiempo, es igual a 30,240.

Al presionar cualquier tecla, la máquina vuelve a ofrecer el submenú de 3 opciones señalado en la fig. 33-A, del cual se elige, para este caso, la letra C (terminar) con la cual el programa regresa al submenú de 3 opciones, mostrado en la fig.32.

### 5.3 COMBINACION.

Quando se desean realizar cálculos de COMBINACION, una vez seleccionada la opción L o M del MENU PRINCIPAL (fig.1) y seleccionar la opción A (FACT. PERM. COMB.) del submenú que aparece en pantalla (fig. 32), se toma la opción C (COMBINATIONS) del siguiente submenú (fig.33). Inmediatamente después la máquina pide el valor de N y R, que serán obtenidos del ejemplo 5.3.

#### EJEMPLO 5.3

---

Un investigador tiene quince medicamentos que desea poner a prueba, pero sólo cuenta con los animales experimentales para probar cinco de los medicamentos. ¿Cuántas combinaciones de medicamentos puede poner a prueba?

(Daniel, W. W., 1983 modificado).

---

Una vez anotados los valores  $N=15$  y  $R=5$  al dar "Enter" la máquina pregunta de nuevo a dónde se desea que muestre el resultado y se elige la opción A, para indicar que lo muestre en

pantalla, observando entonces que el número de combinaciones que se puede hacer con 15 objetos (medicamentos) y 5 plazas o lugares (animales) es igual a 3,003.

Al presionar cualquier tecla, el programa regresa al submenú de 3 opciones mostrado en la fig 33-A, se elige la opción C y el programa se traslada al submenú de la fig. 32, del cual se toma la opción C, si lo que se desea es salir de este submenú, para regresar al MENU PRINCIPAL (fig. 1).

## C A P I T U L O 6

### DISTRIBUCION DE PROBABILIDADES

El paquete estadístico MICROSTAT, permite realizar cálculos de diversos tipos de DISTRIBUCION DE PROBABILIDAD, por medio de la opciones L o M del MENU PRINCIPAL (fig. 1). Inmediatamente después, aparece en pantalla un submenú como el mostrado en la fig. 32, del que se elige la opción B y de esta forma se indica a la máquina que se desea desarrollar distribuciones de probabilidad. Inmediatamente la máquina ofrece un submenú que va de las letras A a la I como el que se muestra en la fig. 34.

FIG. 34

-----  
ENTER: OPTION: B

----- PROBABILITY DISTRIBUTIONS -----

OPTIONS: A. BINOMIAL  
          B. HYPERGEOMETRIC  
          C. POISSON  
          D. EXPONENTIAL  
          E. NORMAL  
          F. F. DISTRIBUTION  
          G. STUDENT'S t  
          H. CHI SQUARE  
          I. [Terminate]

-----

**DONDE:**

**A=** Distribución binomial  
**B=** " hipergeométrica  
**C=** " de Poisson  
**D=** " exponencial  
**E=** " normal  
**F=** " F o de Fisher  
**G=** " t de student  
**H=** " de chi cuadrada  
**I=** terminar.

A continuación se realizan algunos ejercicios para mejor comprensión del uso de este comando.

### **6.1 DISTRIBUCION BINOMIAL.**

Para realizar cálculos de **DISTRIBUCION BINOMIAL**, primeramente se selecciona la letra **A** del submenú de la fig. 34. A continuación se da un ejemplo como base de datos para ilustrar la operación:

### EJEMPLO 6.1

---

Supóngase que la probabilidad de recuperación de una enfermedad es de 40%; si 20 animales enferman, ¿cuál es la probabilidad de que:

- a) Se recuperen cuando mucho 6
- b) Se recuperen exactamente 12
- c) Se recuperen cuando menos 7
- d) Se recuperen entre 5 y 10

(López, B.B., 1994).

---

Por consiguiente, cuando el programa pide que se introduzca el valor de  $N$  (tamaño de la muestra) se anota el # 20, el valor  $p$  (probabilidad) es 0.40, el valor mínimo o  $X_1$  es igual a 0 y el valor máximo o  $X_2$  es 12, dando "Enter" después de anotar cada valor. Enseguida aparece en pantalla el submenú que pide se indique a donde debe enviarse la información y se elige la letra A, apareciendo así en pantalla toda la distribución que corresponde para una distribución binomial con  $N=20$  y  $P=0.4$ . (fig. 35).

FIG. 35

ENTER: OPTION: A

BINOMIAL DISTRIBUTION  
 N = 20            P = .4

X	P (X)	CUMULATIVE PROBABILITY
0	.00004	.00004
1	.00049	.00052
2	.00309	.00361
3	.01235	.01596
4	.03499	.05095
5	.07465	.12560
6	.12441	.25001
7	.16588	.41589
8	.17971	.59560
9	.15974	.75534
10	.11714	.87248
11	.07099	.94347
12	.03550	.97897

$E(X) = 8.00000$ ,     $STD. DEV. = 2.19089$ ,     $VARIANCE = 4.80000$

PRESS ANY KEY TO CONTINUE

Como puede observarse en la fig. anterior, se muestran tres columnas, la primera columna son los valores de  $X$  que van de 0 a 12, la segunda columna es la probabilidad de cada valor  $X$  (probabilidades individuales) y la tercera columna es la probabilidad acumulada de 0 hasta 12, además se muestra un renglón en donde se ofrece la esperanza de  $X$  correspondiente a la media aritmética que es igual a 8, de la desviación estándar que es igual a 2.19089 y una varianza de 4.8.

Así entonces, tenemos las siguientes respuesta a los incisos del ejemplo 6.1:

- a) Se recuperen cuando mucho 6. Tomar el valor acumulado de 6 en la tercera columna: .25001
- b) Se recuperen exactamente 12. Tomar el valor exacto de la distribución binomial de 12 (segunda columna): .03550
- c) Sobrevivan cuando menos 7. Restar a la unidad el valor acumulado de 6 (tercera columna):  $1 - .25001 = .74999$
- d) Sobrevivan entre 5 y 10. Restar el valor acumulado de 4 (tercera columna) al valor acumulado de 10 (tercera columna):  
 $.87248 - .05095 = .82153$

Al pulsar cualquier tecla, el programa regresa al submenú de 3 opciones como el ilustrado en la fig. 33-A, del cual se selecciona la letra C para regresar al bloque de las distribuciones de probabilidad.

## 6.2 DISTRIBUCION DE POISSON

Del submenú de la fig. 34, ahora se selecciona la opción **C**, para describir cómo se comporta la distribución de Poisson, para lo cual se proporciona el siguiente ejemplo:

### EJEMPLO 6.2

---

Supóngase que se sabe que en cierta área de una gran ciudad, el número promedio de ratas por manzana es de 5. Suponiendo que el número de ratas se distribuye según Poisson; encontrar la probabilidad de que en una manzana elegida aleatoriamente:

- a) se encuentren 5 ratas
- b) se encuentren más de 5 ratas
- c) se encuentren menos de 5 ratas
- d) se tengan entre 5 y 7

(López, B.B., 1994)

---

De los datos anteriores podemos obtener los valores que el programa requiere para este caso, así entonces:

- valor de la media aritmética de ocurrencia de los eventos = 5
- valor inferior o  $X_1 = 0$
- valor máximo o  $X_2 = 7$

al dar "Enter" después de anotar cada uno de ellos, aparecen en pantalla las opciones de salida, se elige la opción **A** y así se

observa esta distribución desplegada en tres columnas, como en la fig. 36 se observa.

FIG. 36

-----  
ENTER: OPTION: C

POISSON DISTRIBUTION

MEAN RATE OF OCCURENCE = 5

X	P(X)	CUMULATIVE PROBABILITY
0	.00674	.00674
1	.03669	.04043
2	.08422	.12465
3	.14037	.26503
4	.17547	.44049
5	.17547	.61596
6	.14622	.76218
7	.10444	.86663

E(X) = 5.00000, STD. DEV. = 2.23607, VARIANCE = 5.00000

PRESS ANY KEY TO CONTINUE.  
-----

Cabe observar que esta distribución de Poisson es muy semejante a la distribución binomial, ya que se muestra esta distribución en 3 columnas, la primera columna muestra los valores de  $X$  que van del 0 a 7, la segunda la probabilidad de  $X$  para cada uno de los valores de  $X$ , y la tercera la probabilidad acumulada de todos los valores, también muestra un renglón en donde se da la esperanza de  $X$ , que corresponde a la media que es igual a 5, la desviación estándar que es igual a 2.23607 y la varianza que es igual a 5. Ya con estos valores, es posible resolver correctamente los incisos del ejemplo 6.2:

a) Se tengan 5 ratas. Tomar el valor exacto de 5 (segunda columna): .17547.

b) Se encuentren más de 5 ratas. Restar a la unidad el valor acumulado de 5:  $1 - .61596 =$  .38404.

c) Se encuentren menos de 5 ratas. Tomar el valor acumulado de 4 (tercera columna): .44049.

d) Se tengan entre 5 y 7 ratas. Restar el valor acumulado de 4 al valor acumulado de 7:  $.86663 - .44049 =$  .42614.

Al presionar cualquier tecla, el programa envía a un submenú de 3 opciones, como el mostrado en la fig. 33-A, del cual se selecciona la letra C para salir al submenú donde está la distribución de probabilidades.

### 6.3 DISTRIBUCION NORMAL.

Ahora se selecciona la letra E, cuyo uso para distribución normal se demuestra por medio de ejercicios. Al seleccionar la letra que corresponde a la distribución normal, en pantalla aparecen 2 opciones, A para CALCULOS DE PROBABILIDAD si se proporcionan valores de "Z" y B para CALCULOS DE "Z" proporcionando la probabilidad.

6.3.1. Primeramente se opta por la letra A (CALCULOS DE PROBABILIDAD) mencionada en el punto 6.3 y el programa pide se le indique el valor correspondiente a "Z". Para realizar esta prueba tómense como base los resultados que se muestran en la fig. 42, en donde se observa un valor de "Z" igual a 2.6484, al anotarlo y dar "Enter" la máquina pregunta a dónde se desea que muestre el resultado, indicando que en pantalla por medio de A, observando así que para un valor de  $Z = 2.6484$ , la probabilidad de ocurrencia de los eventos es de 0.9960 y la probabilidad de 1-P, que corresponde al área fuera de "P", es de 0.0040 (fig. 37), lo que al sumar da un valor igual a la unidad.

FIG. 37

-----  
ENTER: OPTION: A

NORMAL DISTRIBUTION

ENTER: OBSERVED Z-VALUE: 2.6484

- OPTIONS: A. SCREEN OUTPUT  
B. PRINTER OUTPUT WITH FORMFEEDS  
C. PRINTER OUTPUT WITHOUT FORMFEEDS  
D. TEXT FILE OUTPUT  
  
E. OUTPUT PRINTER SET-UP CODES  
F. CHANGE PRINTER WIDTH. CURRENT VALUE:

ENTER: OPTION: A

NORMAL DISTRIBUTION

Z = 2.6484

P = .9960

1-P = .0040

PRESS ANY KEY TO CONTINUE  
-----

Al presionar cualquier tecla, la máquina pregunta si se desea repetir la misma operación o hacer algún otro cálculo.

6.3.2. Para este punto, se elige la letra B (ANOTHER COMPUTATION) del menú que se tiene en pantalla, con la cual el programa muestra un submenú del que se utiliza la opción B para realizar CALCULOS DE "Z" (indicada en el punto 6.3). La máquina pide se le proporcione la probabilidad acumulada de "p". Para este ejemplo se toma el valor mostrado en la fig. 73, que es igual a 0.0001182, enseguida pregunta a dónde se desea que envíe el resultado, seleccionando la letra A, para así mostrar en pantalla, si  $P = .0001182$ , que es la opción de área que se dió, "z" es igual a  $-3.67656$  (fig. 38).

FIG. 38

-----  
ENTER: OPTION: B

NORMAL DISTRIBUTION

ENTER CUMULATIVE PROBABILITY, P: .0001182

OPTIONS: A. SCREEN OUTPUT  
B. PRINTER OUTPUT WITH FORMFEEDS  
C. PRINTER OUTPUT WITHOUT FORMFEEDS  
D. TEXT FILE OUTPUT  
  
E. OUTPUT PRINTER SET-UP CODES  
F. CHANGE PRINTER WIDTH. CURRENT VALUE: 80

ENTER: OPTION: A

NORMAL DISTRIBUTION

Z = -3.6765587922085  
P = .0001, 1-P = .9999

PRESS ANY KEY CONTINUE.  
-----

Al presionar una tecla la máquina regresa nuevamente a un submenú con 3 opciones, del que se utiliza C para salir, la cual lleva a otro submenú en donde estan las distribuciones de probabilidad, tecleando la letra I para terminar. Por último, del siguiente submenú se selecciona la opción C para terminar, con lo cual la máquina regresa al menú principal de MICROSTAT. Se recuerda que para una mejor comprensión de el uso de las opciones L y M, debe consultarse el manual principal del paquete MICROSTAT, tomando en consideración que para terminar una sesión de trabajo se usa la opción P del menú principal.

## C A P I T U L O 7

### PRUEBAS DE HIPOTESIS

El programa MICROSTAT ofrece la posibilidad de realizar 4 diferentes tipos de PRUEBAS DE HIPOTESIS PARA MEDIAS, como lo muestra la fig. 39, que aparece al seleccionar la opción D del MENU PRINCIPAL (fig. 1):

FIG. 39

```
-----  
ENTER: OPTION: D  Module being loaded...  
-----HYPOTHESIS TESTS FOR MEANS-----  
  
SELECT TEST TO BE PERFORMED:  
  
OPTIONS:  A. MEAN VS. HYPOTHESIZED VALUE  
          B. DIFFERENCE BETWEEN MEANS: PAIRED OBSERVATIONS  
          C. DIFFERENCE BETWEEN TWO GROUP MEANS: KNOWN VARIANCE  
          D. DIFFERENCE BETWEEN TWO GROUP MEANS: POOLED ESTIMATE  
            OF VARIANCE  
          E. [Terminate]  
  
ENTER: OPTION:  _  
-----
```

en el entendido que: **A** sirve para comparar medias de una muestra con un valor hipotético, **B** sirve para comparar 2 medias o diferencias entre 2 medias con observaciones pareadas, **C** se utiliza para comparar 2 medias o 2 grupos de medias con varianza conocida, con **D** se obtienen diferencias entre 2 grupos de medias con una varianza común y **E** permite salir de este submenú.

## 7.1 COMPARACION DE MEDIAS CON UN VALOR HIPOTETICO.

Si se selecciona la opción A del menú anterior, inmediatamente el programa pide en pantalla el nombre del archivo con que se va a trabajar, el cual debe ser proporcionado, indicando el drive donde se encuentra dicho archivo, en este caso se usa el archivo PRUEHIP. Al dar "Enter", la máquina muestra el nombre del archivo, su etiqueta (M.ALEATORIA GRASA EN LECHE) y el número de casos (10) y variables (1) que se tiene en el archivo cuya descripción es la siguiente:

### EJEMPLO 7.1

-----

De una población de 100 vacas Holstein en producción, se toma una muestra aleatoria. Demostrar que el promedio de grasa de la leche (en kilos) producida durante un mes, no es diferente de 18.1 Kg. Tomar un nivel de significancia igual a .05.

(López, B.B., 1994)

-----

El programa también ofrece las opciones de trabajar con todos los casos o sólo algunos en particular, se ordena A (ALL CASES) y el programa muestra el nombre con que se tiene identificada la (las) variables y los casos numerados de el 1 hasta el 10. Enseguida pregunta por el número de variable que se va a utilizar, que para este ejemplo sólo puede ser la 1 y al indicarlo, la máquina pregunta cuál es el valor hipotético con

el que se va a comparar esa media (generalmente es la media de la población), que en este caso es 18.1 y así entonces, se está comparando la muestra que se tiene o la media de la muestra, contra la media de la población. Al anotar dicho valor hipotético, la máquina pide un título con el cual mostrar los resultados, que puede o no ser anotado, se da "Enter" e inmediatamente la máquina muestra los resultados como en la fig. 40.

FIG. 40

-----  
DATA INPUT IN PROCESS...

----- HYPOTHESIS TESTS FOR MEANS -----

MEAN VS. HYPOTHESIZED VALUE

HEADER DATA FOR: B:PRUEHIP LABEL: M. ALEATORIA GRASA EN LECHE

NUMBER OF CASES: 10

NUMBER OF VARIABLES: 1

HYPOTHESIZED VALUE = 18.1000

MEAN = 16.5000

STD. DEV. = 7.3829

STD. ERROR = 2.3347

N = 10 (CASES = 1 TO 10)

T = -.6853

(D.F. = 9)

VARIABLE TESTED: GRASA

PROB. = .2552

PRESS ANY KE TO CONTINUE.  
-----

Es importante notar que en esta figura se ofrece el nombre del archivo, drive, etiqueta, número de casos que hay en esa variable y número de variables en total que tiene el archivo. También ofrece el valor hipotético empleado, media aritmética de la muestra, desviación estándar y número de casos que tiene la muestra, además del valor "t", los grados de libertad, la probabilidad de que los resultados se den por azar y el nombre de la variable que se seleccionó, con lo cual puede concluirse que:

No existe diferencia significativa entre el promedio de grasa en leche de la muestra y la media de la población, ya que la probabilidad obtenida es alta ( $P > .5$ ).

Al pisar cualquier tecla, se ofrece un submenú de 3 opciones como el de la fig. 33-A, del cual se elige ahora la opción C para terminar.

## **7.2 DIFERENCIA ENTRE MEDIAS CON OBSERVACIONES PAREADAS.**

Si se elige la opción B del submenú de PRUEBAS DE HIPOTESIS PARA MEDIAS (fig. 39), es posible comparar las diferencias entre 2 medias con observaciones pareadas y al hacerlo, el programa pide se indique el nombre del archivo con el que se desea trabajar y en este caso se selecciona el archivo PRUEHIP1 (previa indicación del drive en que se halla el diskette que lo guarda), que a continuación se describe:

## EjemPlo 7.2

---

De una explotación al pastoreo, se seleccionaron al azar 12 terneras de 7 meses de edad en promedio, a las cuales se les pesó, desparasitó y administró 5 cm. de un complejo de vitamina "B" y fueron pesadas nuevamente a los 35 días. En base a los datos obtenidos, ¿podría deducirse que la administración desparasitante-vitamínico, incrementó el peso de las terneras en el periodo estudiado? Tómese un nivel de significancia de .05.

(López, B. B., 1994)

---

Una vez anotado el nombre del archivo, se da "Enter" y en pantalla se muestran los datos generales de ese archivo, cuya etiqueta es "COMPARACION DE 2 MEDIAS. PRUEBA PAREADA", conteniendo 12 casos y 2 variables. Al ofrecer el programa las opciones de trabajar con todos los casos o sólo algunos en eligiendo la alternativa A (ALL CASES) para este caso. Enseguida el programa pide se le indique cuales son las variables a utilizar y se introducen los números 1 y 2 respectivamente pulsando "ENTER" después de anotar cada número. Después el programa pregunta cuál es la diferencia hipotética que se va a utilizar para comparar las 2 medias. Normalmente, se introduce el valor "0" para indicar que las medias son idénticas (lo cual se hará en este caso) y obviamente esta será una hipótesis nula, pero si se desea, puede seleccionarse algún otro valor. El programa pide un título con el cuál mostrar los resultados (que

puede ser omitido al dar "Enter") y los resultados aparecen en pantalla como lo muestra la fig. 41.

FIG. 41

DATA INPUT PROCESS...

12

HYPOTHESIS TESTS FOR MEANS

DIFFERENCE BETWEEN MEANS: PAIRED OBSERVATIONS

HEADER DATA FOR: B:PRUEHIPI LABEL: COMPARACION DE 2 MEDIAS  
NUMBER OF CASES: 12 NUMBER OF VARIABLES: 2

HYPOTHESIZED DIFF. = .0000  
MEAN = -1.8333  
STD. DEV. = 5.8284  
STD. ERROR = 1.6825  
N = 12 (CASES = 1 TO 12)

T = -1.0896 (D.F. = 11) GROUP 1: ANTES  
GROUP 2: DESPUES

PROB. = .1496

PRESS ANY KEY TO CONTINUE.

Como puede observarse, en pantalla aparece el valor hipotético de diferencia que es 0, la media que es -1.8333, la desviación estándar que es 5.8284 y el error estándar de 1.6825. El valor "t" es de -1.0896, con 11 grados de libertad, la probabilidad de que este valor "t" se encuentre por azar es de 0.1496, lo cual indica que a un nivel de significancia de .05 se debe aceptar que el grupo 2 es igual al grupo 1, es decir, que no existe diferencia significativa entre medias, debido probablemente a que las terneras se hallan en periodo de adaptación después del destete.

A continuación, se pulsa cualquier tecla y aparece en pantalla un submenú (fig. 33-A) del que se elige la opción que se desee.

### 7.3 DIFERENCIA ENTRE MEDIAS CON VARIANZA CONOCIDA.

Si se selecciona la opción C del submenú de pruebas de hipótesis, es posible COMPARAR 2 MEDIAS O DOS GRUPOS DE MEDIAS CON VARIANZA CONOCIDA, así el programa pregunta inmediatamente el nombre del archivo que contiene los datos que se requieren para correr esta prueba. Sirva para este ejemplo el archivo PRUEHIP2, que a continuación se describe:

#### EjemPlo 7.3

---

2 diferentes tipos de alimento fueron probados en borregos de engorda, para lo cual se formaron al azar 2 grupos de 12 animales cada uno, administrándole a un grupo el alimento "A" y a otro el alimento "B". Con los datos obtenidos, ¿se podría concluir que el alimento "A" es mejor que el alimento "B" a un nivel de significancia de .05?

(López B. B. 1994)

---

Una vez que se anota el archivo y se da "Enter" el programa muestra en pantalla los datos generales de ese archivo, cuya etiqueta es "PRUEBA DE 2 MUEST. CON IGUAL MEDIA Y VAR" además de las opciones de trabajar con grupos representados por variables (A) o con grupos representados por algún caso en particular (B). Se elige la primera alternativa y se introducen las variables requeridas, que en este caso son 1 y 2, dando "Enter después de anotar cada una, se pide un título para mostrar los datos, que se omite al dar "Enter" y aparecen los resultados como lo muestra la fig. 42.

FIG. 42

-----  
 ENTER JOB TITLE:

0

DATA INPUT IN PROCESS...

32

----- HYPOTHESIS TESTS FOR MEANS -----

DIFFERENCE BETWEEN TWO GROUP MEANS: KNOWN VARIANCE

	GROUP 1	GROUP 2
MEAN =	31.7500	28.6667
STD. DEV. =	3.1945	2.4618
N =	32	32
	DIFFERENCE =	3.0833
STD. ERROR OF	DIFFERENCE =	1.1642

Z = 2.6484

-----  
 PRESS ANY KEY TO CONTINUE.

En este cuadro es posible observar la media, desviación estándar y número de casos por grupo, así como la diferencia entre grupos, el error estándar de las diferencias y el valor "Z"

de la prueba, con el cual es posible obtener la probabilidad por medio de la opción M del programa MICROSTAT (específicamente punto 6.3.1). Así entonces, al realizar el cálculo se observa que la probabilidad correspondiente a "Z" igual a 2.6484 es .9960 y  $1-p = .004$ , siendo este último valor la probabilidad de encontrar estos valores por azar, por lo que es posible concluir que: el alimento A es mejor que el alimento B ( $P < .05$ ). Al presionar cualquier tecla, el programa ofrece las alternativas de continuar (A Y B) o salir de esa prueba (C).

#### 7.4 DIFERENCIA ENTRE MEDIAS CON VARIANZA COMUN.

Si se desea conocer la DIFERENCIA ENTRE MEDIAS CON VARIANZA COMUN, se selecciona la opción D del submenú de pruebas de hipótesis para medias, siguiendo las instrucciones del punto 7.4 y utilizando el archivo PRUEHIP3 que a continuación se describe:

##### EJEMPLO 7.4

-----  
En un estudio realizado en 32 mujeres con ciclos menstruales normales y 32 que presentaron alteraciones en dichos ciclos, fueron medidos los niveles de hormona LH, cuyos valores se observan en el archivo PRUEHIP3. Analizando esta información, ¿es posible concluir que los niveles de hormona LH de las pacientes con alteraciones menstruales son superiores a los niveles de las pacientes normales? Tomar un nivel de significancia ( $\alpha$ ) de .05.

-----  
(López, B. B., 1994)

Una vez corrida la prueba, los datos se muestran en pantalla como en la fig. 42-A

FIG. 42-A

-----  
ENTER JOB TITLE:

0

DATA INPUT IN PROCESS...

32

----- HYPOTHESIS TESTS FOR MEANS -----

DIFFERENCE BETWEEN TWO GROUP MEANS: POOLED ESTIMATE OF VARIANCE

	GROUP 1	GROUP 2	
MEAN =	5.9678	8.4612	
STD. DEV. =	3.2686	7.5841	
N =	32	32	
	DIFFERENCE =	-2.4934	
STD. ERROR OF DIFFERENCE =		1.4599	
T=	-1.7080	(D.F = 62)	GROUP 1: LHN GROUP 2: LHP

PROB. = .0463

PRESS ANY KEY TO CONTINUE.  
-----

Como se puede observar, con esta prueba se realizan los cálculos mediante una prueba "t", indicando así mismo la probabilidad de que este valor sea encontrado por azar, siendo posible concluir que: efectivamente los niveles de LH en las pacientes con alteraciones menstruales son superiores a los niveles de las pacientes con ciclos normales.

Si se desea imprimir los resultados de cualquiera de estas pruebas, al pulsar cualquier tecla, el programa ofrece 3 opciones

ya descritas anteriormente: A) repetir la salida, B) hacer más cálculos o C) terminar y salir de esa prueba. Se elige la opción A (repetir la salida) y en ese momento la máquina ofrece un submenú, del que se puede elegir volver a mostrar los datos en pantalla, imprimirlos o enviarlos a un archivo de textos.

## C A P I T U L O 8

### ANALISIS DE VARIANZA

El programa MICROSTAT permite realizar varios tipos de ANALISIS DE VARIANZA por medio de la opción E del menú principal (fig. 1), ofreciendo las opciones que se muestran en la fig. 43:

FIG. 43

-----  
ENTER: OPTION: E    Module being loaded...  
-----

-----ANALYSIS OF VARIANCE-----  
-----

OPTIONS:    A. ONE-WAY ANOVA  
            B. RANDOMIZED BLOCK ANOVA  
            C. TWO-WAY ANOVA  
            D. [Terminate]

ENTER: OPTION: \_  
-----

donde: **A** sirve para hacer pruebas de "ANDEVA" con un solo camino de clasificación, **B** para hacer pruebas de "ANDEVA" para bloques aleatorios, **C** para hacer pruebas de "ANDEVA" con 2 caminos de clasificación y **D** salir de ese submenú.

#### 8.1 PRUEBAS CON UN CAMINO DE CLASIFICACION.

Si se selecciona la opción **A** del submenú anterior, nos

permite hacer pruebas de ANDEVA para modelos totalmente Aleatorizados, ya sean balanceados o desbalanceados. A manera de ejercicio, se anota el drive donde está dicho archivo ANDEVA1 y la máquina muestra que ese archivo tiene 4 variables con 4 casos cada variable, cuya descripción se hace a continuación:

**EJEMPLO 8.1**

Cuatro intervalos de riego fueron probados en 4 parcelas de caña, una por intervalo (sin riego, riego cada 21, 28 y 35 días) cuyo rendimiento se muestra en la fig. 44. ¿Se podría concluir que al menos una de las medias del intervalo de riego es diferente, a un nivel de significancia de .05?

(López, B. B., 1994)

FIG. 44.

ANDEVA1

	TEST.	RIE21D	RIE28D	RIE35D
1	1850.0	2874.0	3049.0	2822.0
2	1960.0	2620.0	2980.0	3050.0
3	2000.0	2500.0	2491.0	2790.0
4	1990.0	2450.0	3000.0	2790.0

PRESS ANY KEY TO CONTINUE.

Cuando el programa pregunta el número de grupos que se van a utilizar para correr la prueba, se asigna, en este caso, el número 4. Luego ofrece 2 opciones, donde A permite trabajar con grupos definidos por variables o B con grupos definidos por casos subsecuentes.

8.1.1 GRUPOS DEFINIDOS POR VARIABLES. Si se opta por la alternativa A, el programa ofrece los 4 grupos en pantalla y se pueden seleccionar en el orden que se desee, cualquiera de los grupos con sólo anotar el número que identifica a cada nombre de variable. Una vez seleccionados los 4 grupos que ofrece la pantalla, independientemente del orden en que se hallan seleccionado, la máquina pide algún título para ofrecer los resultados (puede no anotarse si así se desea), se da "Enter" y acto seguido la máquina ofrece 2 opciones: A para mostrar los resultados con los promedios de los tratamientos y B suprimiendo estas medias. Si se selecciona la opción A, inmediatamente se muestra el menú de salida de datos, como el de la fig. 7, del cual se elige la primera opción para que la máquina muestre los resultados en pantalla, como se observa en la fig. 45.

FIG. 45

-----  
 ENTER: OPTION: A

DATA INPUT IN PROCESS. . .

4

-----ANALYSIS OF VARIANCE-----

ONE-WAY ANOVA

	GROUP	MEAN			
	1	1950.000	4		
	2	2611.000	4		
	3	2880.000	4		
	4	2928.000	4		
	GRAN MEAN	2592.250	16		
SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F RATIO	PROB.
BETWEEN	2433459.000	3	811153.000	25.216	1.815E-05
WITHIN	386022.000	12	32168.500		
TOTAL	2819481.000	15			

-----  
 PRESS ANY KEY TO CONTINUE.

En base a los resultados anteriores, con los valores obtenidos de  $F=25.216$  y  $P=.00001815$ , es posible concluir que efectivamente, al menos una de las "medias" es diferente.

Al presionar cualquier tecla, el programa ofrece un submenú de 4 opciones como se muestra en la fig. 46:

FIG. 46

```
-----  
OPTIONS:  A. REPEAT OUTPUT  
          B. MORE COMPUTATION  
          C. CHANGE TEST OR FILE  
          D. [Terminate]
```

```
ENTER:  OPTION:  _  
-----
```

donde A permite repetir la salida de los resultados anteriores, B es para hacer más análisis, C permite cambiar de prueba o archivo y D es para terminar y salir de ese submenú. Por ahora se selecciona la opción C y el programa regresa al menú de la fig. 43.

**8.1.2 GRUPOS DEFINIDOS POR CASOS SUBSECUENTES.** Nuevamente se toma la primera alternativa (ONE-WAY ANOVA) del submenú de la fig. 43. Al preguntar el programa el nombre de archivo se tecldea el drive y ANDEVA1A, se da "Enter" y se observan en pantalla los datos generales del archivo, cuyos datos se aprecian en la fig. 47 y corresponden a la descripción del ejemplo 8.1a

FIG. 47

	RANCHO.	PROD.
1	1	1850
2	1	1960
3	1	2000
4	2	2874
5	2	2620
6	2	2500
7	3	3029
8	3	2980
9	3	2491
10	4	2822
11	4	2680
12	4	3050

PRESS ANY KEY TO CONTINUE.

#### EJEMPLO 8.1a

Cuatro regeguerías en zona tropical fueron comparadas por su producción de leche, para lo cual se seleccionaron al azar 3 vacas de cada regeguería, considerando la producción total de leche, ajustada a 305 días. Podría concluirse, que al menos el promedio de producción de uno de los ranchos es diferente a los otros, tomando en cuenta un nivel de significancia igual a .05?

(López, B. B., 1994)

Cuando el programa pide el número de grupos, se anota el 4, que es el número de tratamientos que hay en este archivo (fig. 47). -Nótese que los valores por analizar están en una sola columna llamada producción. Enseguida el programa ofrece las siguientes opciones, A para grupos definidos por variables y B para grupos definidos por casos subsecuentes con una variable, optando ahora por la segunda opción.

Cuando el programa pregunta el número de caso o fila donde inicia el grupo 1, se tecléa el 1 y el 3 como fin de caso en ese grupo. A continuación para el grupo 2 se indica que inicia en el caso 4 y finaliza en el caso 6, el grupo 3 inicia en el caso 7 y termina en el caso 9 y el grupo 4 inicia en el caso 10 y termina en el 12. Al dar "Enter" la máquina demanda oprimir cualquier tecla y pregunta el número de variable que se desea analizar, que en este caso es la variable 2 (producción). En seguida el programa pide un título que puede omitirse al dar "Enter", después se selecciona la opción A del menú que aparece en pantalla para que ofrezca las medias de los tratamientos y del menú siguiente se elige la letra A para que los resultados aparezcan en pantalla como se muestran en la fig. 48.

FIG.48

-----  
 DATA INPUT IN PROCESS. . .

4

-----ANALYSIS OF VARIANCE-----

ONE-WAY ANOVA

GROUP	MEAN	N
1	1936.667	3
2	2664.667	3
3	2840.000	3
4	2850.667	3
GRAND MEAN	2573.000	12

VARIABLE 2: PROD.

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F RATIO	PROB.
BETWEEN	1685132.000	3	561710.667	13.226	1.815E-03
WITHIN	339762.000	8	42470.250		
TOTAL	2024894.000	11			

PRESS ANY KEY TO CONTINUE.  
 -----

En base a los datos obtenidos, de  $F = 13.226$  y  $P = .001815$ , es posible concluir que por lo menos una de los promedios es diferente a los otros, ya que la probabilidad obtenida es  $< .05$ .

Finalmente al oprimir cualquier tecla, el programa ofrece las opciones descritas en la fig. 46, de las que se opta por la D para regresar al submenú de la fig. 43.

## 8.2 PRUEBAS PARA BLOQUES ALEATORIOS.

Si lo que se desea es hacer un análisis de varianza con un modelo de bloques al azar, entonces se debe seleccionar la opción B de la fig. 43 y al dar "Enter" el programa pide el nombre del archivo con que se desea trabajar, que en este caso es el archivo ANDEVA3, anteponiendo el drive en que se encuentra. Al dar "Enter" la máquina muestra la etiqueta del archivo así como el número de casos y variables que contiene y que se describe en el ejemplo 8.2.

### EjemPlo 8.2

---

3 razas de bovinos para carne fueron evaluadas en 4 ranchos, para lo cual se tomaron las ganancias diarias de peso en Kgs. Si se remueve el efecto de Rancho, ¿se podría concluir a un nivel de significancia de .05, que al menos la "media" de uno de los genotipos es diferente que los otros?. Los datos obtenidos se muestran en la fig. 49.

(López B. B. 1994)

---

FIG. 49

	geno1	geno2	geno3
1	.9580	.9860	.9250
2	.9710	1.0510	.9520
3	.9270	.8910	.8290
4	.9710	1.0100	.9550

PRESS ANY KEY TO CONTINUE.

En la línea donde se pide el número de bloques debe anotarse el número de renglones (casos) que el archivo contiene, que en este caso son 4 y en la línea de número de tratamientos se anota el número de variables que se tienen y que en este caso es 3. Enseguida la máquina pide el orden en que se procesarán los 3 grupos, si no se desea cambiar la orden ya dada, solamente se oprime "Enter" tantas veces como sea necesario, luego pide un título para mostrar los resultados en pantalla, que puede ser suprimido si se da "Enter", inmediatamente después ofrece la alternativa de sacar las medias de los tratamientos (A) o suprimirlas (B), en este caso se opta por la opción A y la máquina pregunta a dónde se desea que mande los datos, oprimiendo la letra A para que los muestre en pantalla y pueden observarse como en la fig. 50.

FIG. 50

## RANDOMIZED BLOCKS ANOVA

TREATMENT		MEAN	N		
1		.957	4		
2		.985	4		
3		.915	4		
BLOCK		MEAN	N		
1		.956	3		
2		.991	3		
3		.882	3		
4		.979	3		
GRAND MEAN		.952	12		
SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F RATIO	PROB.
TREATMENT	9.7172E-03	2	4.8586E-03	6.968	.0273
BLOCK	.021	3	7.1303E-03	10.226	8.967E-03
ERROR	4.1835E-03	6	6.9725E-04		
TOTAL	.035	11			

PRESS ANY KEY TO CONTINUE.

Como puede verse, a nivel de tratamiento se obtiene una probabilidad  $< .05$ , lo que nos indica que efectivamente por lo menos una de las "medias" es diferente.

Al presionar cualquier tecla, el programa muestra el submenú de la fig. 46 y se elige la opción D para terminar y regresar al menú de ANALISIS DE VARIANZA (fig. 43).

### 8.3 PRUEBAS CON 2 CAMINOS DE CLASIFICACION.

Si se desea hacer un análisis de varianza con 2 criterios de clasificación (2 vías), se debe utilizar la opción C de la fig. 43. Al dar "Enter" la máquina pide el nombre del archivo con el que se desea trabajar, que para este caso es **ANDEVA5**, recordando indicar el drive en que se halla el archivo, al dar "Enter" la máquina ofrece información general sobre el archivo, que en este ejemplo es un modelo factorial 2x2x4, cuya descripción se hace en el ejemplo 8.3.

#### EjemPlo 8.3

---

2 tipos diferentes de alimentos, combinados con 2 concentraciones de desparasitante fueron estudiados en cuatro razas de borregos, para lo cual se utilizaron grupos de 5 animales por dieta-tratamiento y 20 animales por raza, quedando distribuidos como se muestra en la fig. 51. Se podría concluir que existe efecto de raza, de tratamiento y de interacción en el experimento? Tomar un nivel de significancia de .01.

(López, B. B., 1994)

---

FIG. 51

PESO DE BORREGOS EN KGS.

		RAZAS					
		X1	X2	X3	X4		
I	A	1	20.0	25.0	24.0	28.0	D
		2	25.0	30.0	28.0	31.0	
		3	22.0	29.0	24.0	26.0	
		4	27.0	28.0	25.0	29.0	
		5	21.0	30.0	30.0	32.0	
	B	6	30.0	30.0	39.0	40.0	I
		7	45.0	29.0	42.0	45.0	
		8	30.0	31.0	36.0	50.0	
		9	35.0	30.0	42.0	45.0	E
		10	36.0	30.0	40.0	60.0	
II	A	11	31.0	32.0	41.0	42.0	T
		12	30.0	35.0	45.0	50.0	
		13	40.0	30.0	40.0	40.0	
		14	35.0	40.0	40.0	55.0	A
		15	30.0	30.0	35.0	45.0	
	B	16	20.0	23.0	24.0	29.0	S
		17	21.0	25.0	25.0	30.0	
		18	20.0	28.0	30.0	28.0	
		19	20.0	30.0	26.0	27.0	
		20	19.0	31.0	23.0	30.0	

I y II = TRATAMIENTOS

A y B = CELDAS

En el punto donde el programa pide se indique el número de replicaciones por celda, se anota el 5, enseguida demanda el número de renglones que es 4 y en número de columnas se anota el 4. Después de dar "Enter" la máquina pide el orden en que las variables se van a analizar, si no se desea modificar dicho orden, sólo se da "Enter" las veces que sea necesario. Luego pide un título, que puede suprimirse con sólo dar "Enter" y entonces se ofrece la opción siguientes: A) calcular las medias de los

tratamientos o B) suprimir dichas medias; se selecciona A y del menú siguiente también A para mandar los resultados a pantalla (fig. 52).

FIG. 52

-----  
 ENTER: OPTION: A  
 -----

-----ANALYSIS OF VARIANCE-----

TWO-WAY ANOVA

COL	MEAN	N
1	27.850	20
2	29.800	20
3	32.950	20
4	38.100	20
ROW	MEAN	N
1	26.700	20
2	38.250	20
3	38.300	20
4	25.450	20

PRESS ANY KEY TO CONTINUE.

COL	MEANS	MEAN	N
ROW	COL		
1	1	23.000	5
2	1	35.200	5
3	1	33.200	5
4	1	20.000	5
1	2	28.400	5
2	2	30.000	5
3	2	33.400	5
4	2	27.400	5
1	3	26.200	5
2	3	39.800	5
3	3	40.200	5
4	3	25.600	5
1	4	29.200	5
2	4	48.000	5
3	4	46.400	5
4	4	28.800	5
GRAND MEAN		32.175	80

PRESS ANY KEY TO CONTINUE

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F RATIO	PROB.
COLS	1201.050	3	400.350	27.269	1.765E-11
ROWS	2997.450	3	997.483	67.943	1.000E-13
INTERACTION	608.450	9	67.606	4.605	1.047E-04
ERROR	939.600	64			
TOTAL	5741.550	79			

-----  
PRESS ANY KEY TO CONTINUE

Así entonces, de la tabla de ANALISIS DE VARIANZA, se puede inferir que existe un efecto en los pesos tanto de raza como de tratamiento e interacción.

Al pulsar cualquier tecla, se observa en pantalla el menú de la fig. 46, del cual se opta por la opción D para terminar y regresar al de la fig. 43, donde se toma también la opción D y así poder regresar al MENU PRINCIPAL (fig. 1)

## CAPITULO 9

### GRAFICOS

El paquete MICROSTAT es limitado en la elaboración de gráficos, sin embargo permite hacer gráficos sencillos de punto, donde si se desea, se puede trazar la línea de regresión. Si se quiere utilizar esta opción, se selecciona la letra F del menú principal. Al dar "Enter" el programa pide el nombre del archivo, el cual, para explicar su uso, será el REGSIM, descrito en el ejemplo 9.1 y cuyos datos aparecen en la fig. 53.

#### EJEMPLO 9.1

---

Con el objeto de construir una curva patrón para determinar la concentración de proteína en solución, se hicieron por triplicado las diluciones siguientes, utilizando albúmina de huevo disuelta en agua destilada.

Diluciones:	11 = 2.5	mg/ml
	10 = 2.25	"
	9 = 2.0	"
	8 = 1.75	"
	7 = 1.5	"
	6 = 1.25	"
	5 = 1.0	"
	4 = .75	"
	3 = .5	"
	2 = .25	"
	1 = .0	"

(López, B. B., 1994)

FIG: 53

CURVA PATRON PARA DETERMINAR LA CONCENTRACION DE UNA SUSTANCIA.

HEADER DATA FOR: B:REGSIM LABEL: curva patron  
NUMBER OF CASES: 33 NUMBER OF VARIABLES: 2

	con	abs
1	.000	.000
2	.250	.008
3	.500	.020
4	.750	.032
5	1.000	.049
6	1.250	.060
7	1.500	.072
8	1.750	.080
9	2.000	.094
10	2.250	.100
11	2.500	.110
12	.000	.000
13	.250	.008
14	.500	.020
15	.750	.032
16	1.000	.049
17	1.250	.060
18	1.500	.072
19	1.750	.081
20	2.000	.094
21	2.250	.100
22	2.500	.109
23	.000	.000
24	.250	.005
25	.500	.020
26	.750	.032
27	1.000	.050
28	1.250	.061
29	1.500	.078
30	1.750	.086
31	2.000	.095
32	2.250	.102
33	2.500	.108

Después de anotarlo y dar "Enter", la máquina proporciona datos generales del archivo y las siguientes 2 opciones:

- A: INPUT ALL CASES.
- B: INPUT SUBSET OF CASES.

si se elige la opción A (introducir todos los casos) aparece un submenú, del cual se elige la letra A para mandar los datos a pantalla. Luego el programa pide el número de variable del eje de las "x" (horizontal) que en este caso es el número 1 (concentración) y luego pide el número de variable para el eje de las "y" (abscisa vertical) que será el 2. Después el programa pregunta si se desea una escala automática, aceptando por medio de la letra Y, y pregunta también si se desea la línea de regresión indicando que sí por medio de la letra Y nuevamnete. Así, al dar "Enter" la máquina ofrece los valores de la línea de regresión, de "r" y "r2" como puede observarse en las figuras. 54 y 54-A. En la gráfica la línea de regresión puede trazarse con una recta, uniendo los dos signos "+" que en ella se encuentran.

FIG. 54

-----  
DO YOU WANT AUTOMATIC SCALING (Y, N) ? YES

DO YOU WANT LINEAR REGRESSION LINE PLOTTED (Y, N) ? YES

INPUT:

PRE-PLOT COMPUTATIONS:

HORIZONTAL AXIS: con

LEFT ENDPOINT: 0 RIGHT ENDPOINT: 2.5

VERTICAL AXIS: abs

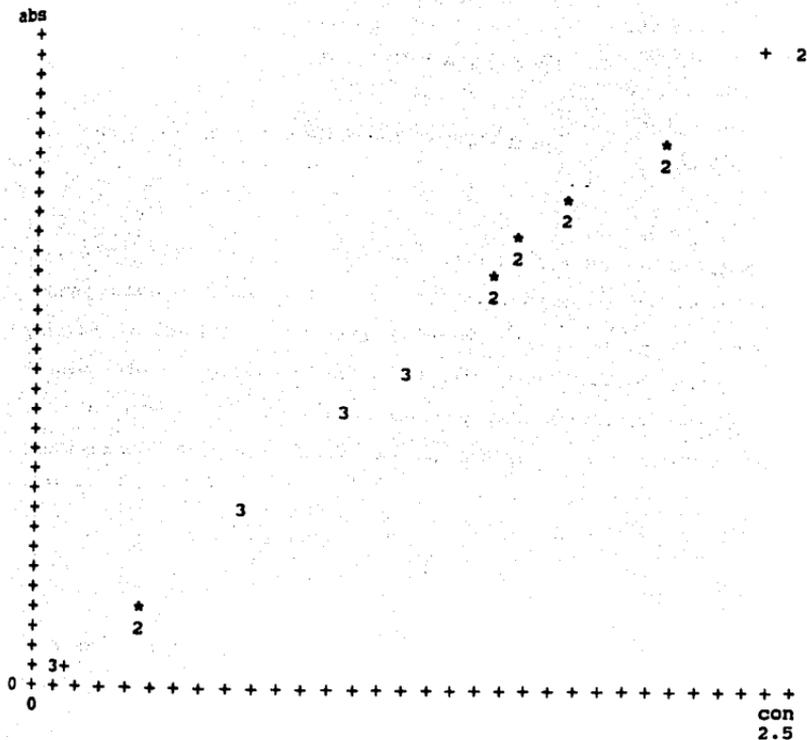
LOWER ENDPOINT: 0 UPPER ENDPOINT: .11

REGRESSION EQUATION (shown by +'s on scatterplot):

INTERCEP= -4.5454545454546E-04 SLOPE= 4.6109090909091E-02  
r = .9949 r squared = .9898

-----  
PRESS ANY KEY TO CONTINUE.

FIG. 54-A



Para continuar debe darse "Enter" y la máquina muestra un menú con las siguientes 3 opciones: A, realizar más gráficos con el archivo abierto; B, repetir la salida y C permite salir de este comando.

## C A P I T U L O 10

### MATRIZ DE CORRELACION

Con el programa MICROSTAT es posible hacer análisis de correlación simple con 2 o más variables, para lo cual se utiliza la opción G del menú principal (fig. 1). A continuación se describe el procedimiento a seguir para utilizar esta orden, con los datos del archivo REGLM1 que se muestran en la fig. 55 y corresponden a 4 variables que se midieron en la combustión de hoja de tabaco:

1. tiempo de combustión
2. concentración de Cloro
3. " de Nitrógeno
4. " de Potasio

FIG. 55

---

 COMBUSTION DE HOJA DE TABACO

HEADER DATA FOR: B:REGLM1 LABEL: combustion de hoja de tabaco

NUMBER OF CASES: 30 NUMBER OF VARIABLES: 4

	t.comb.	nitro	cloro	potasio
1	.3400	3.0500	1.4500	5.6700
2	.1100	4.2200	1.3500	4.8600
3	.3800	3.3400	.2600	4.1900
4	.6800	3.7700	.2300	4.4200
5	.1800	3.5200	1.1000	3.1700
6	.0000	3.5400	.7600	2.7600
7	.0800	3.7400	1.5900	3.8100
8	.1100	3.7800	.3900	3.2300
9	1.5300	2.9200	.3900	5.4400
10	.7700	3.1000	.6400	6.1600
11	1.1700	2.8600	.8200	5.4800
12	1.0100	2.7800	.6400	4.6200
13	.8900	2.2200	.8500	4.4900
14	1.4000	2.6700	.9000	5.5900
15	1.0500	3.1200	.9200	5.8600
16	1.1500	3.0300	.9700	6.6000
17	1.4900	2.4500	.1800	4.5100
18	.5100	4.1200	.6200	5.3100
19	.1800	4.6100	.5100	5.1600
20	.3400	3.9400	.4500	4.4500
21	.3600	4.1200	1.7900	6.1700
22	.8900	2.9300	.2500	3.3800
23	.9100	2.6600	.3100	3.5100
24	.9200	3.1700	.2000	3.0800
25	1.3500	2.7900	.2400	3.9800
26	1.3300	2.6100	.2000	3.6400
27	.2300	3.7400	2.2700	6.5000
28	.2600	3.1300	1.4800	4.2800
29	.7300	3.4900	.2500	4.7100
30	.2300	2.9400	2.2200	4.5800

---

Una vez seleccionada la opción G del menú principal y haber anotado el drive y nombre del archivo con que se va a trabajar, se da "Enter" y el programa muestra las 4 variables que tiene el archivo y los 30 casos que cada una contiene, así como la

etiqueta que se refiere al tipo de datos que se manejan en el archivo. Inmediatamente después se selecciona la opción A (input all cases = incluir todos los casos) de las que se ofrecen en pantalla y nuevamente la opción A (correlate all variables) del siguiente desplegado, que indica que se desea correlacionar todas las variables). Enseguida el programa pide un título, que puede ser suprimido dando "Enter" e inmediatamente después muestra las opciones de la fig. 7, de la cual se selecciona la letra A y al pulsar "Enter" de nuevo se pide hacer una selección de las 3 opciones que a continuación se indican:

- A. OUTPUT CORRELATION MATRIX  
(mostrar la matriz de correlación).
- B. OUTPUT SSCP AND VAR-COVAR  
(mostrar SSCP, covarianza y correlación).
- C. ALL OF THE ABOVE  
(mostrar todo lo anterior).

#### 10.1 MATRIZ DE CORRELACION

Si se elige la opción A, la máquina muestra los resultados de las operaciones realizadas como lo muestra la fig. 56.

FIG. 56

-----  
 COMPUTATIONS IN PROCESS. . .

	t.comb.	nitro	cloro	potasio
t.comb.	1.00000			
nitro	-.71773	1.00000		
cloro	-.49964	.20940	1.00000	
potasio	.17937	.09256	.40740	1.00000

CRITICAL VALUE (1-TAIL, .05) = + Or - .30645

CRITICAL VALUE (2-TAIL, .05) = +/- .36034

N = 30

-----  
 PRESS ANY KEY TO CONTINUE.

## 10.2 COVARIANZA Y CORRELACION.

Si se opta por la opción B (Out put SSCP & VAR-COVAR), la máquina mostrará los resultados de los cálculos realizados como se muestra en la fig. 57.

FIG. 57

-----  
 DATA INPUT AND COMPUTATIONS IN PROCESS. . .

ROW	COL.	RAW SSCP	ADJUSTED SSCP	VAR-COVAR	CORR
t.comb.	t.comb.	2.080740E+01	6.689520E+00	2.306731E-01	1.00000
nitro	t.comb.	6.165020E+01	-5.824760E+00	-2.008538E-01	-.71773
cloro	t.comb.	1.241030E+01	-4.211480E+00	-1.452234E-01	-.49964
potasio	t.comb.	9.844080E+01	2.668340E+00	9.201172E-02	.17937

ROW	COL.	RAW SSCP	ADJUSTED SSCP	VAR-COVAR	CORR
nitro	nitro	3.323352E+02	9.845547E+00	3.395016E-01	1.00000
cloro	nitro	8.158340E+01	2.141307E+00	7.383816E-01	.20940
potasio	nitro	4.594052E+02	1.670547E+00	5.760506E-02	.09256

ROW	COL.	RAW SSCP	ADJUSTED SSCP	VAR-COVAR	CORR
cloro	cloro	3.019070E+01	1.062094E+01	3.662392E-01	1.00000
potasio	cloro	1.203950E+02	7.636657E+00	2.633330E-01	.40740

ROW	COL.	RAW SSCP	ADJUSTED SSCP	VAR-COVAR	CORR
potasio	potasio	6.827813E+02	3.308290E+01	1.140790E+00	1.00000

PRESS ANY KEY TO CONTINUE.

CRITICAL VALUE (1-TAIL, .05) = + or - .30645

CRITICAL VALUE (2-TAIL, .05) = +/- .36034

N= 30  
 -----

### 10.3 MATRIZ DE CORRELACION Y COVARIANZA.

Si se elige la opción C (ALL OF THE ABOVE), el programa muestra juntos todos los datos que aparecen en las figs. 56 y 57.

Como puede observarse en la fig. 56, usando el criterio de dos colas, el valor que se obtiene a un nivel de significancia de 0.05 es igual a  $\bullet$  .36034, valor que al ser comparado con los

valores de correlación de la matriz de correlación, determina que la variable "t. comb." está correlacionada con el Nitrógeno en  $-.71773$ , así como con Cloro en  $-.49964$ , pero no así con Potasio, ya que la correlación de  $.17937$  es inferior a  $.36034$ . Cabe hacer notar también, que las variables Cloro y Potasio se asocian con una valor igual a  $.40740$ .

Al presionar cualquier tecla después de que aparecen los datos en pantalla, independientemente de la selección que se haya realizado, el programa muestra en pantalla un menú de opciones, del que por medio de la letra C, se obtiene la salida de esta función para regresar al menú principal del programa.

## C A P I T U L O 11

### ANALISIS DE REGRESION

Por medio de la opción **H** del menú principal de este programa (fig. 1), es posible hacer Análisis de Regresión Simple o Múltiple, lo que a continuación se describe por medio de 2 ejercicios. Al hacer la selección, el programa pide el nombre del archivo con que se va a trabajar, que es en este ejercicio el **B:REGSIM**, cuyos datos se muestran en la fig. 53 del capítulo correspondiente a "Gráficos". Al anotarlo y dar "Enter" aparecen en pantalla los datos generales del archivo, cuya etiqueta es "Curva patrón" con 2 variables y 33 casos, donde la variable 1 es "concentración" y la 2 es "absorvancia". Inmediatamente después se selecciona la opción **A** de cada uno de los 2 siguientes submenús que aparecen en pantalla, que indican que se va a trabajar con todos los casos y todas las variables.

#### 11.1 REGRESION SIMPLE.

Llegado este punto, el programa requiere se le indique cual es la **variable dependiente** con que se va trabajar, que para este caso es la 2, pide un título para mostrar los datos en pantalla (se omite dando "Enter") y aparecen en pantalla el índice,

nombre, promedio y desviación estándar de dicha variable dependiente. Posteriormente requiere se anote el número de las variables de predicción o independientes, que para este caso sólo es la 1 (concentración) y pregunta el índice de dichas variables, que en este caso de nueva cuenta es sólo 1. Pregunta el programa después si la selección hecha es correcta y si lo es, se confirma con la letra Y. En el momento en que el programa pregunta si se desea que calcule los residuales, se accede por medio de la letra Y, también pregunta si se desea realizar la prueba de DURBIN-WATON, indicando que no por medio de N y finalmente pregunta el número de decimales con que se desea que aparezcan los resultados, que en este caso por de "faul" es 4, se da "Enter" y enseguida pregunta a dónde se desea destinar la información, tecleando la opción A (fig. 7), para que finalmente, después de un breve momento, muestre índice, promedio, nombre y desviación estándar de las variables que contiene el archivo y al presionar cualquier tecla, se muestran los resultados como aparecen en la fig. 58.

FIG. 58

-----REGRESSION ANALYSIS-----

HEADER DATA FOR: B:REGSIM LABEL: curva patron  
 NUMBER OF CASES: 33 NUMBER OF VARIABLES: 2

INDEX	NAME	MEAN	STD.DEV
1	con	1.2500	.8028
DEP. VAR.:	abs	.0572	.0372

PRESS ANY KEY TO CONTINUE.

DEPENDENT VARIABLE: abs

VAR.	REGRESSION COEFFICIENT	STD. ERROR	T(DF= 31)	PROB.
con	.0461	8.40019E-04	54.891	.00000
CONSTANT	-4.5455E-04			

STD. ERROR OF EST. = .0038

r SQUARED = .9898  
 r = .9949

-----  
 PRESS ANY KEY TO CONTINUE.

Puede verse en esta figura, que el coeficiente de regresión para la variable "concentración" es de 0.0461 que sería en la ecuación de la recta ( $y=a+bx$ ) el valor de "b" y la constante el valor de "a" igual a -0.000454 Así mismo, el programa muestra un valor "T" y la probabilidad de encontrar al azar dicho resultado

menor a .0001, con lo que es posible concluir que la variable "absorvancia" puede ser explicada en función de la variable "concentración".

Al presionar cualquier otra tecla, también el programa proporciona el **Coefficiente de correlación (r)**, el **Coefficiente de determinación (r<sup>2</sup>)** y una tabla de ANDEVA para el modelo de la recta (fig. 59), además de los valores observados, esperados y el residual que resulta de la diferencia de las dos medidas anteriores, así como un pequeño gráfico de estandarización de los residuales (fig. 59-A).

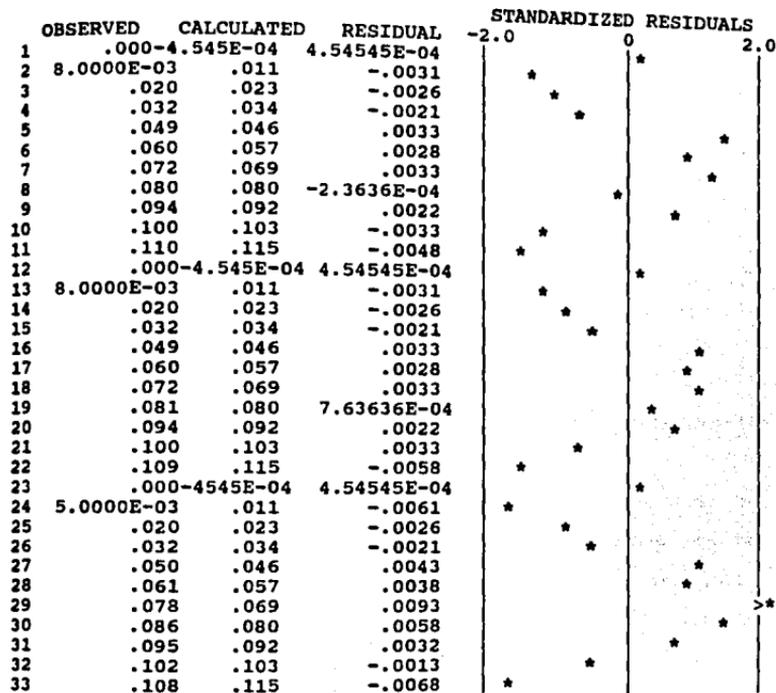
FIG. 59

-----  
ANALYSIS OF VARIANCE TABLE

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F RATIO	PROB.
REGRESSION	.0438	1	.0438	3012.969	.000E+00
RESIDUAL	4.51164E-04	31	1.45537E-05		
TOTAL	.0443	32			

-----

FIG. 59-A



Al pulsar cualquier tecla, aparece en pantalla un menú como el que se muestra en la fig. 59-B:

FIG. 59-B

OPTIONS: A. ANOTHER SET OF PREDICTORS FOR DEP. VAR  
B. CHANGE DEPENDENT VARIABLE  
C. REPEAT OUTPUT FROM PREVIOUS ANALYSIS  
D. CALCULATE PREDICTED VALUES  
E. OUTPUT RESIDUALS TO DATA FILE  
F. [Terminate]

Seleccionando la opción F de este submenú, es posible salir de este comando y regresar al menú principal del programa.

## 11.2 REGRESION MULTIPLE.

En el siguiente ejemplo se hace un Análisis de Regresión Múltiple, para el cuál se usa el archivo **REGLM1** mostrado en la fig. 55, donde se utiliza el ejemplo descrito en el capítulo 10, unicamente que en este caso se desea explicar, a través de una ecuación de Regresión Lineal Simple, el tiempo de combustión (var. dependiente) por medio de las variables independientes (nitrógeno, cloro y potasio). La ecuación que se propone para explicar la variable dependiente es:

$$Y_{ijkl} = a + b_1x_i + b_2x_j + b_3x_k + E_l$$

donde:

$Y_{ijkl}$  indica el tiempo de combustión de una hoja de tabaco en la  $i$ -ésima concentración de nitrógeno, con la  $j$ -ésima concentración de cloro y la  $k$ -ésima concentración de potasio, más un  $E_l$ -ésimo error (aleatorio).

a,  $b_1$ ,  $b_2$  y  $b_3$  son los valores estimados del coeficiente de regresión del modelo.

Una vez seleccionada la letra **H** del menú principal, se anota el nombre **B:REGLM1**, se selecciona la opción **A** para incluir todos los casos y de nuevo la **A** del siguiente submenú para incluir todas las variables. Cuando pregunta el nombre de la **variable dependiente** se anota el **# 1** (tiempo de combustión), el título se omite dando "Enter" y en pantalla se muestran los datos de las variables de ese archivo (índice, nombre, media y desviación estándar), luego pide el número de las **variables de predicción o independientes**, que en este caso para introducirlas todas se anota el número 3, al dar "Enter" el programa pregunta el orden de las variables, que en esta ocasión llevarán el mismo orden que se observa en pantalla, el programa pregunta si la selección es correcta, lo cual se confirma con la letra **Y** y al hacerlo ofrece 2 opciones para realizar los cálculos, indicadas en la fig. 59-C

FIG. 59-C

-----  
A: FULL MODEL REGRESSION (modelo completo de regresión) y  
B: STEPWISE REGRESSION (modelo de regresión por pasos).  
-----

### 11.2.1 Modelo completo de regresión.

Si se opta por la letra **A**, el programa pregunta si se desea que calcule residuales, se anota la letra **N** para indicar que **no**. Enseguida pide el número de decimales para los resultados, que en este caso será 6 y ofrece el menú de salida de datos de la fig.

7, se elige A y enseguida el programa ofrece los datos generales del archivo, con la media y desviación estándar de cada variable y al presionar cualquier tecla se muestran los resultados de los cálculos realizados (fig.60).

FIG. 60

----- REGRESSION ANALYSIS -----

HEADER DATA FOR: B:REGLM1 LABEL:COMBUSTION DE HOJA DE TABACO  
 NUMBER OF CASES: 30 NUMBER OF VARIABLES:4

INDEX	NAME	MEAN	STD.DEV
1	NITRO	3.278667	.582668
2	CLORO	.807667	.605177
3	POTASIO	4.653667	1.068077
DEP. VAR.:	T. COMB.	.686000	.480284

PRESS ANY TO CONTINUE.

DEPENDENT VARIABLE: t. comb.

VAR.	REGRESSION COEF.	STD.ERROR	T(DF= 26)	PROB	PARTIAL r^2
nitro	-.531455	.069577	-7.638	.00000	.6917
cloro	-.439636	.073037	-6.019	.00000	.5822
potasio	.208975	.040640	5.142	.00002	.5042
CONSTANT	1.811043				

STD. ERROR OF EST. = .213468

ADJUSTED R SQUARED = .802453

R SQUARED = .822889

MULTIPLE R = .907132

-----

Al presionar nuevamente una tecla, en pantalla aparece una tabla de análisis de varianza como la que se muestra en la fig. 61.

FIG. 61

## ANALYSIS OF VARIANCE TABLE

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F RATIO	PROB.
REGRESSION	5.504734	3	1.834911	40.267	6.453E-10
RESIDUAL	1.184786	26	.045569		
TOTAL	6.689520	29			

PRESS ANY KEY TO CONTINUE.

En la figura 60, pueden observarse los aportes del coeficiente de regresión por cada una de las variables independientes. Para  $b_1$  es Nitrógeno, Cloro es para  $b_2$ , para  $b_3$  es Potasio, además del aporte de la constante  $a$  del modelo descrito. Si se observa la probabilidad de  $T$  para cada coeficiente, puede notarse que las 3 variables explican el tiempo de combustión, ya que  $p < .0001$ . Así mismo, se obtienen los valores de  $R$  múltiple (.907132),  $R_2$  (.822889) y  $R_2$  ajustada (.802453). Por último, este modelo es evaluado por medio de un Análisis de Varianza (fig. 61), donde por medio del valor  $F$  (40.267) y  $p < .0001$ , se confirma que la ecuación matemática propuesta es eficiente.

Al presionar cualquier tecla, nuevamente aparece el submenú de la fig. 59-B, ya explicado en ese punto.

### 11.2.2 Modelo de regresión por pasos.

Si se opta por la letra  $B$  de la fig. 59-C, el modelo a usar será el de "STEPWISE REGRESSION". Al hacerlo, la máquina pide un valor  $F$  para correr el modelo, que por de faul, si se da "Enter",

la máquina anota el número 3; luego pide otro valor  $F$  para correr la prueba, removiendo la variable independiente que menos aporte tenga en el modelo y que por de fault es de nuevo 3, también pide el máximo valor de tolerancia que en este caso es .01 y los pasos máximos que por de fault son 5 y cuando pregunta si se desea el cálculo de residuales se anota N (no). Pregunta el número de decimales y se anota el 6, enseguida ofrece el menú de salida de pantalla (fig. 7) y se elige A, al hacerlo y dar "Enter" la máquina inicia el proceso de cálculo, ofreciendo primeramente el índice, nombre, media y desviación estandar de las variables que contiene el archivo, así como los criterios para correr el modelo, expresados en valores  $F$  para incluir o remover una variable, y la probabilidad o tolerancia. Al presionar cualquier tecla la máquina corre el primer paso y da los valores del coeficiente de regresión y la constante con que participa la primera variable incluida en el modelo, que en este caso es Nitrógeno, así mismo, muestra la tabla de análisis de varianza del modelo y los valores parciales de  $R^2$  y la  $F$  de las variables no incluidas en el modelo (fig. 62).

FIG. 62

```

-----
INDEX          NAME          MEAN          STD.DEV.
  1          nitro          3.278667      .582668
  2          cloro          .807667      .605177
  3          potasio        4.653667      1.068077
DEP. VAR.:    t.comb.          .686000      .480284
  
```

F TO ENTER = 3, F TO REMOVE = 3, TOLERANCE = .01

STEP 1. VARIABLE: nitro ENTERED.

DEPENDENT VARIABLE: t.comb.

```

VAR.    REGRESSION COEFFICIENT    STD. ERROR    F(1, 28)    PROB.
nitro   -.591614                    .108470       29.748      .00001
CONSTANT 2.625704
  
```

STD. ERROR OF EST. = .340352

```

r SQUARED = .515135
r = -.717729
  
```

PRESS ANY KEY TO CONTINUE.

ANALYSIS OF VARIANCE TABLE

```

SOURCE      SUM OF SQUARES    D.F.    MEAN SQUARE    F. RATIO    PROB.
REGRESSION  3.446008           1        3.446008       29.748      8.025E-06
RESIDUAL    3.243512           28        .115840
TOTAL       6.689520           29
  
```

VARIABLES NOT IN EQUATION:

```

NAME      PARTIAL r^2      TOLERANCE      F TO ENTER      PROB.
cloro     .2632            .9562          9.647           4.424E-03
potasio   .1257            .9914          3.881           .0592
  
```

PRESS ANY KEY TO CONTINUE.

Al presionar cualquier tecla, el programa realiza una segunda corrida, adicionando la siguiente variable con mayor valor F de las no incluidas en la primera corrida, que en este caso es Cloro, mostrando los resultados como en la fig. 63.

FIG. 63

-----  
 STEP 2. VARIABLE: cloro ENTERED.

DEPENDENT VARIABLE: t. comb.

VAR.	REGR. COEF.	STD. ERROR	F(1, 27)	PROB.	PARTIAL r <sup>2</sup>
nitro	-.528549	.096962	29.714	.00001	.5239
cloro	-.289964	.093356	9.647	.00442	.2632
CONSTANT	2.653132				

STD. ERROR OF EST. = .297500

ADJUSTED R SQUARED = .616313

R SQUARED = .642774

MULTIPLE R = .081732

PRESS ANY KEY TO CONTINUE.

ANALYSIS OF VARIANCE TABLE

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F. RATIO	PROB.
REGRESSION	4.299852	2	2.149926	24.291	9.220E-07
RESIDUAL	2.389668	27	.088506		
TOTAL	6.689520	29			

VARIABLES NOT IN EQUATION

NAME	PARTIAL r <sup>2</sup>	TOLERANCE	F TO ENTER	PROB.
potasio	.5042	.8340	26.441	2.311E-05

-----  
 PRESS ANY KEY TO CONTINUE.

En el siguiente, paso el programa realiza la tercera corrida, donde se incluyen las 3 variables del archivo, mostrando los resultados que aparecen en la fig. 64. Cabe notar, que estos valores son iguales a los de las figuras 59 y 60, cuyos resultados ya fueron discutidos.

FIG. 64

-----  
 STEP 3. VARIABLE: potasio ENTERED

DEPENDENT VARIABLE: t. comb.

VAR	REGR. COEF.	STD. ERROR	F(1, 26)	PROB.	PARTIAL r <sup>2</sup>
nitro	-.531455	.069577	58.345	.00000	.6917
cloro	-.439636	.073037	36.232	.00000	.5822
potasio	.208975	.040640	26.441	.00002	.5042
CONSTANT	1.811043				

STD. ERROR OF EST. = .213468

ADJUSTED R SQUARED = .802453

R SQUARED = .822889

MULTIPLE R = .907132

ANALYSIS OF VARIANCE TABLE

SOURCE	SUM OF SQUARES	D.F.	MEAN SQUARE	F RATIO	PROB
REGRESSION	5.504734	3	1.834911	40.267	6.453E-10
RESIDUAL	1.184786	26	.045569		
TOTAL	6.689520	29			

-----

Para terminar se selecciona la letra F del submenú de la fig. 59-B, que aparece en pantalla al pulsar cualquier tecla, y así salir de este comando y regresar al menú principal del programa (fig. 1).

## C A P I T U L O 12

### PRUEBAS DE JI CUADRADA

El programa MICROSTAT, ofrece la posibilidad de realizar pruebas de JI CUADRADA por medio de la letra K del menú principal (FIG. 1). Una vez seleccionada esta opción, la máquina ofrece un submenú de 5 opciones como el que aparece en la fig. 65.

FIG. 65

```
-----  
ENTER: OPTION: k  Module being loaded. . .  
----- CROSSTAB / CHI-SQUARE TEST -----
```

```
OPTIONS:  A. CROSSTAB  
          B. CONTINGENCY TABLE (DATA FILE INPUT)  
          C. CONTINGENCY TABLE (KEYBOARD INPUT)  
          D. GOODNESS OF FIT TEST  
          E. [Terminate]
```

```
ENTER: OPTION: -  
-----
```

donde cada opción permite realizar lo siguiente:

```
A: TABLAS CRUZADAS  
B: TABLAS DE CONTINGENCIA a partir de archivos  
C: TABLAS DE CONTINGENCIA desde el teclado  
D: PRUEBAS DE BONDAD DE AJUSTE  
E: Terminar
```

## 12.1 TABLAS DE CONTINGENCIA A PARTIR DE ARCHIVOS.

A continuación se muestra un ejemplo para uso de la prueba de JI CUADRADA en tablas de contingencia a partir de archivos, seleccionando la opción B de las mostradas en la fig. 65 y para lo cual se utiliza el archivo JI2A, cuyos datos aparecen en la fig. 66 cuya descripción es la siguiente:

### EJEMPLO 12.1

---

" Un grupo de investigación, estudiando la relación entre el tipo de sangre y el grado de cierta afección en una población, reunió datos sobre mil quinientos sujetos y deseaban saber si estos datos eran compatibles con la hipótesis de que el grado de la afección y el tipo de sangre son independientes."

(Daniel, W. W., 1983.)

---

FIG. 66

---

#### GRUPOS SANGUINEOS DE 3 COMUNIDADES

HEADER DATA FOR: B:JI2A LABEL: ejm para ji cuadrada prueb. indep.  
NUMBER OF CASES: 3 NUMBER OF VARIABLES: 4

	a	b	ab	o
1	543	211	90	476
2	44	22	8	31
3	28	9	7	31

---

\* Los grados de afección son indicados por números: 1. ninguno, 2. leve y 3. severo.

Después de anotar el nombre del archivo y dar "Enter" en pantalla aparecen los datos generales del mismo. Enseguida el programa pide un título para el cuadro que se mostrará en pantalla (y que puede omitirse dando "Enter"). El paquete ofrece entonces 4 opciones de salida, como lo muestra la fig. 67.

FIG. 67

```
-----  
OPTIONS:      A. OUTPUT OBSERVED FREQUENCIES  
              B. OUTPUT EXPECTED FREQUENCIES  
              C. OUTPUT OBSERVED PERCENTAGES  
              D. OUTPUT EXPECTED PERCENTAGES  
-----
```

donde:

```
A: SALIDA CON FRECUENCIA OBSERVADA  
B: SALIDA CON FRECUENCIA ESPERADA  
C: SALIDA CON PORCENTAJES OBSERVADOS  
D: SALIDA CON PORCENTAJES ESPERADOS
```

De las opciones anteriores, puede seleccionarse cualquiera, para este caso se selecciona la letra A, e inmediatamente el programa muestra los resultados en pantalla como pueden verse en la fig. 68.

-----  
ENTER: OPTION: A-----  
CROSSTAB / CHI-SQUARE TESTS -----

## OBSERVED FREQUENCIES

	1	2	3	4	TOTAL
1	543	211	90	476	1320
2	44	22	8	31	105
3	28	9	7	31	75
TOTAL	615	242	105	538	1500

CHI-SQUARE = 5.116,

D.F.= 6,

PROB. = .5290

-----  
PRESS ANY KEY TO CONTINUE.

Así entonces, puede observarse que es muy alta la probabilidad de que los datos se hallan por azar (prob.= .5290), por lo tanto se acepta la hipótesis de que el grado de afección y el tipo de sangre son independientes.

Al oprimir cualquier tecla, el programa ofrece el siguiente menú de 3 opciones, del cual opta por la alternativa B.

A: repetir salida  
B: regresar al menú de pruebas de JI CUADRADA  
C: terminar

**12.2 TABLAS DE CONTINGENCIA DESDE EL TECLADO.**

Si se selecciona la opción C del menú de pruebas de ji cuadrada, el programa permite hacer dichas pruebas bajo el modelo de tablas de contingencia introduciendo la información directamente desde el teclado, para lo cual se utiliza el ejemplo siguiente:

**EJEMPLO 12.2**

" Una muestra de 500 niños de cierta escuela primaria, fueron evaluados respecto a su estado de nutrición y su desempeño académico. En base a los resultados que se observan en la fig. 69, ¿es posible concluir si existe una relación entre el estado nutricional y el desempeño académico?

(Daniel, W. W., 1983)

**FIG. 69****ESTADO NUTRICIONAL Y DESEMPEÑO ACADEMICO DE 500 NIÑOS**

ESTADO DE NUTRICION			
Desempeño Académico	Pobre	Bueno	Total
Malo	105	15	120
Satisfactorio	80	300	380
<b>TOTAL</b>	<b>185</b>	<b>315</b>	<b>500</b>

Cuando el paquete demanda el número de renglones, se anota el 2 y en número de columnas también 2, como título se anota "Estado nutricional y desempeño académico de 500 niños". Después pide se anote, para la celda 1-1 el valor que le corresponde, que en este caso es 105, al dar "Enter" pide el valor de la celda 1-2 que es 15, como valor de la celda 2-1 se da 80 y para la celda 2-2 se da 300. Al dar "Enter" la máquina demanda la salida de datos por medio de un submenú de 4 opciones (fig.67) del cual se elige la opción B para este caso y al dar "Enter" el resultado se muestra en pantalla como se observa en la fig. 70. FIG. 70

-----  
 ----- CROSSTAB / CHI-SQUARE TESTS -----  
 -----

Estado Nutricional y Desempeño Académico de 500 niños

EXPECTED FREQUENCIES

	1	2	TOTAL
1	44.40	75.60	120.00
2	140.60	239.40	380.00
TOTAL	185.00	315.00	500.00

CHI-SQUARE WITH CONTINUITY CORRECTION FACTOR =169.907,  
 PROB.= 7.508E-11

CHI-SQUARE WITHOUT CONTINUITY CORRECTION FACTOR =172.746,  
 PROB.= 7.500E-11

D.F. = 1

CALCULATE FISHER EXACT PROBABILITY (Y,N)?  
 -----

Al observar los datos de esta tabla, puede notarse que la probabilidad de que los datos se hayan dado por azar es sumamente pequeña, por lo que puede concluirse que efectivamente existe relación entre el estado nutricional y el desempeño académico.

En esa misma figura, la máquina pregunta si se desea que estime la probabilidad exacta de Fisher, en este caso se opta por la opción **N** (no) e inmediatamente después el programa muestra un menú de 3 opciones descrito en el punto anterior, del cual se elige la letra **B** para regresar al menú de pruebas de ji cuadrada.

**12.3 PRUEBAS DE BONDAD DE AJUSTE.**

Si se selecciona la letra D del menú de pruebas de  $\chi^2$  cuadrada (goodness of fit test) el programa ofrece 2 opciones:

- A: INPUT OBSERVED FREQUENCIES  
(introducir frecuencias observadas)
- B: INPUT OBSERVED PROPORTIONS  
(introducir proporciones observadas)

eligiendo, en este caso, la opción A y del siguiente menú también se selecciona la opción A para introducir las frecuencias esperadas. Así puede estimarse el valor de  $\chi^2$  cuadrada si se tienen frecuencias observadas y esperadas. Para ejemplificar esta prueba, se utiliza el ejemplo siguiente:

#### **EJEMPLO 12.3**

---

" Un investigador reunió datos sobre una muestra de 250 hospitales, lo cual permite calcular, para cada uno, la tasa de ocupación de pacientes para un periodo de 12 meses. Se observan 8 intervalos de clase, así como las frecuencias observadas y esperadas en la fig. 71. ¿Es posible determinar si la muestra estudiada, provino de una población normalmente distribuida?

(Daniel, W. W., 1983).

---

FIG. 71

INTERVALO DE CLASE	FRECUENCIA OBSERVADA (O <sub>i</sub> )	FRECUENCIA ESPERADA (E <sub>i</sub> )
< 40.0	16	14.55
40.0 a 49.9	18	22.18
50.0 a 59.9	22	38.65
60.0 a 69.9	51	49.62
70.0 a 79.9	62	50.48
80.0 a 89.9	55	38.38
90.0 a 99.9	22	21.88
100.0 y mayor	4	14.27
TOTAL	250	250.0

**Nota:** los valores de la frecuencia esperada, se obtienen estandarizando las frecuencias observadas a valores  $Z$ , para lo cual se utiliza la media y la desviación estándar de los datos de la tabla.

Como el número de parámetros utilizados en una tabla de frecuencia es 2 (la media y la desviación estándar), éste es el número que se indica cuando el programa lo demanda. El tamaño de la muestra es igual a 250 y como título se anota "Ajuste de una distribución real a una normal". Al dar "Enter" la máquina pide se anoten las frecuencias observadas y esperadas (fig. 71), lo cual se hace en el orden siguiente:

Fig. 72

CLASE	FRECUENCIA OBSERVADA	FRECUENCIA ESPERADA
1	16	14.55
2	18	22.18
.	.	.
.	.	.
8	4	14.27

Al finalizar, la máquina pregunta si los datos anotados son correctos, lo que se afirma por medio de la letra Y y al dar "Enter" la máquina muestra el resultado como en la fig. 73.

FIG. 73

-----  
 ----- CROSSTAB / CHI-SQUARE TESTS -----  
 -----

GOODNESS OF FIT TEST

CLASS	FREQUENCIES		PROPORTIONS	
	OBSERVED	EXPECTED	OBSERVED	EXPECTED
1	16.00	14.55	.0640	.0582
2	18.00	22.18	.0720	.0887
3	22.00	38.65	.0880	.1546
4	51.00	49.62	.2040	.1985
5	62.00	50.48	.2480	.2019
6	55.00	38.38	.2200	.1535
7	22.00	21.88	.0880	.0875
8	4.00	14.27	.0160	.0571
TOTALS	250.00	250.01	1.0000	1.0000

CHI-SQUARE= 25.361,      D.F. = 5,      PROB. = 1.182E-04

-----

PRESS ANY KEY TO CONTINUE.

Con los datos que se observan en la figura anterior, considerando el valor de  $x_2 = 25.362$  y la probabilidad de 0.000118, puede concluirse que la muestra estudiada no proviene de una población normalmente distribuida.

Al oprimir cualquier tecla, el programa pregunta si se desea calcule la prueba de bondad de ajuste de KOLMOGOROV-SMIRNOV, a lo cual se responde que no, por medio de N y al dar "Enter" la máquina ofrece un submenú de 3 opciones, como el que se muestra en la fig. 74:

FIG. 74

-----  
OPTIONS: A. REPEAT OUTPUT  
          B. MORE COMPUTATIONS  
          C. [Terminate]

ENTER: OPTION:  
-----

donde:

A = repetir la salida  
B = hacer mas pruebas de ji cuadrada  
C = terminar

## C A P I T U L O 13

### ESTADÍSTICAS NO PARAMÉTRICAS

Si se desea hacer cálculos de estadística no paramétrica, el paquete MICROSTAT ofrece esta posibilidad por medio de la letra J del menú principal (fig. 1). A continuación se desarrollan ejemplos sobre la utilización de algunas de estas pruebas.

Al hacer la selección de esta alternativa, el paquete despliega en pantalla un submenú de opciones como el que aparece en la fig. 75.

FIG. 75

```
-----  
ENTER: OPTION: J Module being loaded. . .  
----- NONPARAMETRIC TESTS -----  
  
OPTIONS: A. WALD-WOLFOWITZ RUNS TEST  
          B. WILCOXON RANK-SUM TEST FOR TWO GROUPS  
          C. KRUSKAL-WALLIS TEST  
          D. KOLMOGOROV-SMIRNOV GOODNESS OF FIT TEST  
          E. KOLMOGOROV-SMIRNOV TWO GROUP TEST  
          F. WILCOXON SIGNED RANK TEST  
          G. ABSOLUTE NORMAL SCORES TEST  
          H. FRIEDMAN TEST  
          I. KENDALL COEFFICIENT OF CONCORDANCE  
          J. SIGN-TEST  
          K. FISHER EXACT TEST  
          L. SPEARMAN RANK-ORDER CORRELATION  
          M. [Terminate]  
  
ENTER: OPTION: -  
-----
```

### 13.1 PRUEBA DE RANGO CON SIGNO (WILCOXON SIGNED RANK-TETS)

Esta prueba se ejecuta utilizando la opción F del submenú que aparece en la fig. 75. Acto seguido, la máquina pregunta el nombre del archivo con el que se desea trabajar, en este caso es el B:NPST, que a continuación se describe:

#### EJEMPLO 13.1

---

" 2 dietas fueron probadas en dos grupos de 10 cerdos cada una, durante la etapa de crecimiento. Los datos de la fig. 76 muestran el aumento de peso en kgs. en los 2 grupos. Por medio de la prueba de rango con signo, se desea probar si existen diferencias significativas entre las dietas, a un nivel de significancia ( $\alpha$ ) igual a .05"

(López, B. B., 1994).

---

Este archivo contiene 2 variables, una de ellas representa las diferencias entre las variables "X" y "Y" respetando el signo y la otra es la clasificación de rangos de estas diferencias (fig. 77).

-----  
AUMENTO DE PESO CON 2 DIETAS EN CERDOS.

X	Y
DIETA 1	DIETA 2
17	14
17	21
21	36
18	20
17	24
14	12
19	28
17	16
16	21
12	20

-----

FIG. 77

DIFERENCIAS	RANGO
3	4
- 4	5
- 15	10
- 2	2.5
- 7	7
2	2.5
- 9	9
1	1
- 5	6
- 8	8

-----

Al dar "Enter" el programa muestra la etiqueta del archivo, las 2 variables y los 10 casos que contiene. Inmediatamente después pregunta cuál de las 2 variables contiene las diferencias, en este caso es la 1 y al dar "Enter" pide el rango, que es en este caso el número 2, luego pide un título, que puede omitirse dando "Enter" y al hacerlo el programa muestra el resultado en pantalla como se muestra en la fig. 78.

FIG. 78

-----  
NUMBER OF CASES: 10                    NUMBER OF VARIABLES: 2  
VARIABLE NUMBERS AND NAMES FOR: B:NPSRT  
1. DIFEREN.                            2. RANGOS  
ENTER: VARIABLE CONTAINING DIFFERENCES: 1 DIFEREN.  
ENTER: VARIABLE CONTAINING                RANKS: 2 RANGOS  
ENTER: JOB TITLE

----- NONPARAMETRIC TESTS -----

WILCOXON SIGNED RANK TEST

$$V = 7.5$$

$$Z = -2.039, \quad \text{PROB.} = .0207$$

PRESS ANY KEY TO CONTINUE.  
-----

Así pues, al obtener una probabilidad  $< .05$ , es posible concluir que sí hay diferencias significativas entre estas dos dietas.

Al presionar cualquier tecla aparece un submenú, con el cual, por medio de la opción D, se puede salir de esta prueba.

### 13.2 PRUEBA DE SUMA DE RANGOS.

Seleccionando la opción B del submenú de la fig. 75 puede realizarse la prueba de suma de rangos (WILCOXON RANK-SUM TEST FOR TWO GROUPS). Al dar "Enter" la máquina pregunta el nombre del archivo con que se desea trabajar, que en este caso es el B:MPRST2G. Este archivo fue ordenado por rangos utilizando la orden K de la fig. 2, usando como archivo base los datos de la fig. 79.

FIG. 79

CALIFICAC. DE 2 METODOS DE ENSEÑANZA

HEADER DATA FOR: B:NPRST2G

LABEL: PRUEBA DE SUM. DE RANG.  
2 GRUP.WILCOXON

NUMBER OF CASES: 40

NUMBER OF VARIABLES: 1

	CALIF	
1	80	
2	70	
3	81	
4	80	
5	85	
6	88	
7	82	
8	80	
9	78	
10	75	>
11	83	METODO 1
12	85	
13	78	
14	65	
15	90	
16	80	
17	91	
18	75	
19	85	
20	83	
21	81	
22	85	
23	84	
24	87	
25	82	
26	85	
27	86	
28	85	
29	75	
30	71	>
31	82	METODO 2
32	88	
33	80	
34	71	
35	80	
36	85	
37	93	
38	70	
39	87	
40	79	

En este archivo se muestran las calificaciones otorgadas a 2

grupos de 20 alumnos cada uno con 2 métodos distintos de enseñanza y se desea probar si existen diferencias significativas entre estos dos métodos de enseñanza a un nivel de significancia igual a .05.

Al anotar el nombre del archivo y dar "Enter", la máquina muestra los datos generales del archivo; cabe notar que sólo se muestra una variable, de tal manera que cuando el programa pregunta el número de variable por analizar, se indica el 1, al dar "Enter" la máquina pregunta donde inicia el grupo 1 y se anota el número 1 y el 20 donde termina ese grupo. Después pregunta donde inicia el grupo 2 anotando el número 21 y el 40 como fin de ese grupo. Al dar "Enter" la máquina pide un título que puede omitirse dando "Enter" y al hacerlo, el programa ofrece los resultados como se muestran en la fig. 80.

FIG. 80

-----  
ENTER: JOB TITLE:

DATA INPUT IN PROCESS. . .

40

----- NONPARAMETRIC TESTS -----

WILCOXON RANK-SUM TEST FOR TWO GROUPS

VARIABLE TESTED: CALIF

SUM OF RANKS, GROUP 1 = 381.5 N1 = 20

SUM OF RANKS, GROUP 2 = 438.5 N2 = 20

Z = -.771, PROB. = .2204

-----  
PRESS ANY KEY TO CONTINUE.

Así entonces, con los datos obtenidos, es posible concluir que no existen diferencias significativas entre los dos métodos de enseñanza.

Al oprimir cualquier tecla, aparece un submenú de terminación del cual se elige la letra D para salir de ese comando y regresar al menú principal.

### 13.3 PRUEBA DE KRUSKAL-WALLIS (ANALISIS DE VARIANZA POR UNA VIA DE CLASIFICACION).

Si se opta por la letra C del menú de la fig. 75 es posible hacer pruebas de sumas de rangos (KRUSKAL-WALLIS TEST) para 3 o más grupos. Al dar "Enter" la máquina demanda el nombre del archivo con el que se va a trabajar, siendo para este ejercicio el B:NPSRE, el cual fue ordenado por rangos a través de la opción K de la fig. 2, para lo cual se usó como archivo base, los datos de la fig. 81, provenientes de 3 grupos de ovejas que fueron alimentadas con 3 diferentes dietas. El primer grupo con 5 animales y los 2 restantes con 4 ovejas cada uno, midiendo los aumentos de peso durante el experimento (López, B. B., 1994). En base a los datos obtenidos y por medio de la prueba de KRUSKAL-WALLIS ¿se podría probar que por lo menos una de las dietas es mejor que las otras, a un nivel de significancia igual a .05?

FIG. 81

AUMENTO DE PESO DE TRES GRUPOS DE OVEJAS ALIMENTADAS CON TRES DIFERENTES DIETAS.

PESO		
GRUPO 1	GRUPO 2	GRUPO 3
23	12	20
24	13	19
25	17	14
26	18	15
27	-	-

Al anotar el nombre del archivo y dar "Enter", la máquina ofrece los datos generales del archivo, nótese que sólo existe una variable, ya que los 3 grupos fueron incluidos en ella. Cuando el programa pregunta el número de variables por analizar se anota el 1, al dar "Enter" el programa pregunta el número de grupos con que se va a trabajar, que en este caso es 3, al dar "Enter" de nuevo, el programa pide en donde inicia el grupo 1 y se anota el número 1 y el 5 como final de dicho grupo; como inicio del grupo 2 se anota el número 6 y el 9 como final y los números 10 y 13 como principio y fin del grupo 3. Al dar "Enter", el programa pide un título y se anota "Aumento de peso para 3 grupos de ovejas con 3 dietas diferentes". A continuación se da "Enter" y en pantalla aparece el valor  $\#$  de la prueba, con los grados de libertad empleados y su probabilidad (fig.82).

FIG. 82

-----  
ENTER: BEGINNING CASE NO.: 10                    ENDING CASE NO.: 13  
ENTER: JOB TITLE:  
AUMENTO DE PESO PARA 3 GRUPOS DE BORREGOS CON 3 DIETAS DIFERENTE  
DATA INPUT IN PROCESS. . .  
13

-----  
NONPARAMETRIC TEST -----

KRUSKAL-WALLIS TEST

AUMENTO DE PESO PARA 3 GRUPOS DE BORREGOS CON 3 DIETAS DIFERENTE

VARIABLE TESTED: PESO

H = 9.099                    D.F. = 2                    PROB. = .0106

-----  
PRESS ANY KEY TO CONTINUE.

Así entonces, al analizar los resultados de la tabla anterior, es posible concluir que por lo menos una de las dietas probadas es mejor que las otras.

Al presionar cualquier tecla el programa brinda 4 opciones, de las cuales D sirve para terminar esta prueba.

#### 13.4 PRUEBA DE KOLMOGOROV-SMIRNOV PARA 2 GRUPOS.

Esta prueba puede sustituir a la de Ji-cuadrada como prueba de bondad de ajuste de una distribución real a una teórica. Utilizando la opción E de la fig. 75, puede realizarse la prueba de KOLMOGOROV-SMIRNOV para 2 grupos. Al dar "Enter" la máquina pide se anote el número de clases que se van a crear. Como puede notarse, esta prueba requiere que se introduzcan los datos desde el teclado. Para ejemplificar el uso de esta prueba se utilizan los datos de la fig. 83, que pertenecen a 80 datos de una muestra que fueron agrupados en 6 clases, donde se proporciona la Frecuencia absoluta (observada) y la Frecuencia teórica esperada. Esta última se obtiene del estadístico E para áreas bajo la curva. Se desea, por medio de la prueba de KOLMOGOROV-SMIRNOV para 2 grupos, saber si la Frecuencia absoluta se ajusta a una distribución normal, a un nivel de significancia igual a .05. (López, B. B., 1994).

FIG. 83

---

#### FRECUENCIAS DE DOS MUESTRAS ESTUDIADAS

HEADER DATA FOR: B:NPKS2GT LABEL: PRUEBA DE KOLMOGOROV-SMIRNOV  
2 GRUPOS  
NUMBER OF CASES: 6 NUMBER OF VARIABLES: 2

	FREC. O	FREC. E
1	13	6
2	18	21
3	20	28
4	18	18
5	6	6
6	5	1

---

En este caso, en el número de clases se asigna el dato 6, al dar "Enter" la máquina demanda el valor de los grupos 1 y 2, anotando 13 y 6 respectivamente; en la clase 2 se anotan los valores 18 y 21 y así sucesivamente hasta la última clase que es 5 para el grupo 1 y 1 para el grupo 2. Al introducir el último dato y dar "Enter" la máquina pide un título que se omite al dar "Enter" y en seguida la pantalla muestra los resultados como se observa en la fig. 84.

FIG. 84

----- NONPARAMETRIC TESTS -----

KOLMOGOROV-SMIRNOV TWO GROUP TEST

CLASS	OBSERVED FRECUENCIES		CUMULATIVE RELATIVE FREQUENCIES		*
	GROUP 1	GROUP 2	GROUP 1	GROUP 2	
1	13	6	.1625	.0875	
2	18	21	.3875	.3375	
3	20	28	.6375	.6875	
4	18	18	.8625	.9125	
5	6	6	.9375	.9875	
6	5	1	1.0000	1.0000	
TOTALS	80	80			

\* D MAX = .0875

CRITICAL VALUE AT .05 LEVEL = .2150

CRITICAL VALUE AT .01 LEVEL = .2656

CHI-SQUARE = 1.225, D.F. = 2, PROB. = .5420

-----  
PRESS ANY KEY TO CONTINUE.

De acuerdo a los requerimientos de la prueba, es necesario hacer un planteamiento de hipótesis, que sería el siguiente:

$H_0$  :  $F_0$  se distribuye normalmente

$H_a$  :  $F_0$  no se distribuye normalmente

por lo tanto, analizando los resultados obtenidos al correr la prueba, se observa que la probabilidad obtenida es alta, por lo tanto se acepta  $H_0$ , es decir, que las frecuencias observadas se ajustan a una distribución normal.

Al oprimir cualquier tecla el programa ofrece un submenú de 4 opciones, del cuál la letra **D** permite terminar esta prueba y regresar al menú principal (fig. 1) del cual, por medio de la opción **F**, es posible salir del programa **M I C R O S T A T**.

## DISCUSION

El paquete estadístico "Microstat" ofrece una gran variedad de opciones de trabajo para hacer cualquier análisis estadístico de uso común, como puede observarse a través del desarrollo de cada uno de los capítulos de este manual, proporcionando la ventaja de ser un paquete de fácil manejo y de poco requerimiento de equipo, lo cual le hace ser un paquete versátil. Ofrece la posibilidad de cambiar sencillamente las opciones de trabajo, de elegir el destino de salida de la información y de reutilizar el último archivo con el que se trabajó, con sólo oprimir una tecla.

En cuanto al manejo de archivos, ofrece un sin número de opciones, permitiendo ingresar datos para hacer archivos nuevos, renombrarlos, insertarle o adicionarle casos, listar y editar datos de archivos ya elaborados, mostrar el directorio de los archivos contenidos en un diskette, eliminar los ya caducos, recodificar, transformar o seleccionar variables de un archivo, eliminar casos en particular, clasificar u ordenar datos por medio de rangos, etc.

El programa realiza veloz y eficientemente el cálculo de los elementos que conforman la estadística descriptiva, elaborando tablas de frecuencia en las que es posible observar los límites de clase y con sólo oprimir una tecla, muestra sencillas gráficas de barras que permiten una mejor observación de los resultados.

En lo concerniente al manejo de permutación y combinación, el programa permite realizar estos cálculos probabilísticos de manera por demás sencilla con sólo proporcionar los datos requeridos. El cálculo de la distribución de probabilidad, permite hacer ciertas afirmaciones de probabilidad acerca de variables aleatorias elegidas y en ese mismo rubro, se ofrecen los valores de esperanza de una variable correspondiente a la media aritmética, la desviación estándar y la varianza.

En el caso de pruebas de hipótesis, se obtiene además de un valor hipotético empleado, la media aritmética de la muestra, la desviación estándar y la probabilidad de que los datos se den por azar.

En el análisis de varianza, al seleccionar un sólo camino de clasificación, el programa permite además trabajar con grupos definidos por variables o con grupos definidos por casos subsecuentes.

En cuanto a la elaboración de gráficos, el paquete se halla muy limitado, permitiendo solamente realizar gráficos sencillos de punto.

El punto referido a matriz de correlación permite además calcular el SSCP y varianza y covarianza, mientras que el análisis de regresión, una vez realizados los cálculos, con sólo presionar una tecla, el programa ofrece una tabla de análisis de varianza del modelo usado, además de hacerlo para modelos de regresión lineal múltiple, punto que no contempla el programa de

Biostatística para la carrera de médico veterinario zootecnista.

En las pruebas de CHI cuadrada, el programa proporciona fácilmente la medida de la discrepancia entre las frecuencias observadas y esperadas de ocurrencia de los valores de ciertas categorías, con la ventaja de que proporciona la probabilidad. Un punto muy favorable de esta opción, es que se puede trabajar tanto con archivos previamente creados como con información manejada directamente desde el teclado de la P.C.

Uno de los puntos mas relevantes de este manual, es la inclusión de un capítulo de estadística no paramétrica (que el programa de Biostatística de la carrera de M.V.Z. no incluye), ya que estas pruebas pueden llevarse a cabo cuando no se realizan las suposiciones que fundamentan las pruebas paramétricas o bien, cuando los datos por analizar se miden sobre una escala demasiado débil para los procedimientos aritméticos necesarios de la pruebas paramétricas, es decir, permite estimar categorías, cualidades, clasificaciones o proporciones y pueden ser usadas cuando se desconoce la forma de distribución de la población muestreada, lo cual es de suma importancia, ya que no siempre se cuenta con datos basados en una escala de medición lo suficientemente fuerte como para permitir la realización de las operaciones aritméticas necesarias para realizar los procedimientos paramétricos.

## CONCLUSION

El paquete estadístico "Microstat" realmente es una poderosa herramienta que permite la realización de toda una serie de análisis estadísticos con la finalidad de poder tomar decisiones relacionadas al ejercicio de la medicina veterinaria y zootecnia, por estudiantes , profesionistas, profesores e investigadores de esta área.

## BIBLIOGRAFIA.

1. Acosta, J.J.L. : Uso del método "Fecha Centro" para estimar la producción de leche de vacas a diferentes intervalos de medición en el altiplano mexicano. Tesis de Licenciatura. F.E.S. Cuautitlán. U.N.A.M. México, 1992.
2. Anónimo: Importancia de un área de informática en la escuela de Medicina Veterinaria y Zootecnia de la U.M.S.N.H. LXIX reunión de la A.M.E.F.M.V.Z. Colima, Méx. 1990.
3. Ayala, S. M. G. : Computación I. Introducción a la computación. U.N.A.M. Ed. Porrúa. México, 1987.
4. Bancroft, H. : Introducción a la Bioestadística. 8a. ed. Editorial Universitaria de Buenos Aires. Argentina, 1974.
5. Castillo, J.H. : El uso de la computación electrónica en Medicina Veterinaria y Zootecnia. Algunas consideraciones importantes. U.A.M. Xochimilco D.P.A.A. México, 1981.
6. Castillo, O.A.A.: Alteraciones en la función reproductiva de la mujer y su relación con el perfil hormonal ginecológico. Tesis de Licenciatura. F.E.S. Cuautitlán. U.N.A.M. México, 1994.
7. Daniel, W. W. : Biostatística. Base para el análisis de las ciencias de la salud. 3a. ed. LIMUSA. México, 1983.
8. González, F.V.E. : Utilización de prostaglandina f2 alfa (lutalyse) para sincronizar el parto en cerdas, evaluando algunos parámetros productivos, en el altiplano mexicano. Tesis de Licenciatura. F.E.S. Cuautitlán. U.N.A.M. México, 1992.
9. Izquierdo, E. C. La informática, una formación veterinaria emergente. F.M.V.Z. U. de Colima. México, 1991.

10. Lecroft, E. A. : Biostatística. Editorial C.E.C.S.A. México, 1969.
11. Little, M. T. y Jackson, H. F. : Métodos estadísticos para la investigación en la Agricultura. Editorial Trillas. México, 1983.
12. López, B. B. : Apuntes del curso de Biostatística para la carrera de Médico Veterinario Zootecnista, semestre 94-I. F.E.S. Cuautitlán. U.N.A.M. México, 1994.
13. López, B. B. : Consideraciones generales del Programa de Cómputo de la F. E. S. Cuautitlán. LXVIII Reunión de la A.N.F.M.V.Z. México, 1991.
14. López, B. G. : Informática Veterinaria. F.M.V.Z. U.N.A.M. México, 1991.
15. Martínez, R.A.R.: Evaluación en canal de cuatro grupos genéticos porcinos. Tesis de Licenciatura. F.E.S. Cuautitlán. U.N.A.M. México, 1992.
16. Seybold, P., O'Keffe, L., Kagle, J. : Paquetes estadísticos para la familia IBM PC y compatibles. McGraw-Hill Byte Books. España, 1987.
17. Shefler, W.C. : Biostatística. Editorial Fondo Educativo Interamericano. México, 1981.
18. Spiegel, M. R. : Estadística. Teoría y problemas. Serie Shaum. McGraw-Hill. México, 1970

## APENDICE

### ARCHIVO EJMO.MSD

Los siguientes datos, corresponden al peso (kilogramos) y estatura (centímetros) de alumnos de la clase de Bioestadística del semestre 94-I. (López, B.B., 1994).

PESO	ESTATURA
50	155
65	169
80	184
84	185
72	178
55	155

**ARCHIVO EJMI.MSD**

Los datos de este archivo muestran el peso de alumnos de la clase de Bioestadística del Semestre 94-I, expresado en kgs. (López, B.B., 1994).

**PESO**

50  
65  
80  
84  
72  
55

**ARCHIVO DESCRIP1.MSD.**

Los datos que a continuación se muestran, pertenecen a un estudio realizado en una granja de conejos de la Cd. de Irapuato, Gto., del que se obtuvieron datos sobre el número de gazapos vivos por camada de 70 hembras.

(López, B.B., 1994).

**No. DE GAZAPOS POR CAMADA**

7	11
9	5
3	7
10	9
9	7
6	4
10	10
8	9
8	8
5	8
8	7
6	9
9	7
2	6
6	9
11	8
8	9
9	8
7	5
8	7
7	7
4	9
5	7
7	8
8	13
9	6
8	8
10	7
9	10
8	8
10	11
9	8
9	5
4	9
8	8

**ARCHIVO PRUEHIP.MSD**

De una población de 100 vacas Holstein en producción, fue tomada una muestra aleatoria de 10 animales, para medir el promedio de grasa de la leche (en kilos), producida durante un mes y compararlo con el promedio de grasa en leche de la población total que es igual a 18.1 kg. Los resultados obtenidos son los siguientes (López, B.B., 1994):

**GRASA EN KILOS**

29.5  
16.8  
7.7  
15.4  
23.1  
8.1  
20.9  
16.3  
6.8  
20.4

**ARCHIVO PRUEHIP1.MSD**

De una explotación al pastoreo, se seleccionaron al azar 12 terneras de 7 meses de edad en promedio, a las cuales se les pesó, desparasitó y administró 5 cm. de un complejo de vitamina "B" y fueron pesadas nuevamente a los 35 días, obteniendo los siguientes datos.

(López, B.B., 1994)

**PESO (KG.)**

<b>ANTES</b>	<b>DESPUES</b>
120	128
124	131
130	131
118	127
140	132
128	125
140	141
135	137
126	118
130	132
126	129
127	135

**ARCHIVO PRUEHIP2.HSD**

Dos diferentes tipos de alimento fueron probados en borraeos de engorda, para lo cual se formaron al azar 2 grupos de 12 animales cada uno, administrándole a un grupo el alimento "A" y a otro el alimento "B", obteniendo las siguientes ganancias de peso en hilogramos, durante el periodo de prueba.

(López, B.B., 1994)

**ALIMENTO A**

31  
34  
29  
26  
32  
35  
38  
34  
30  
29  
32  
31

**ALIMENTO B**

26  
24  
28  
29  
30  
29  
32  
26  
31  
29  
32  
28

**ARCHIVO PRUEHIP3.MSD**

En un estudio realizado en 32 mujeres con ciclos menstruales normales y 32 que presentaron alteraciones en dichos ciclos, fueron medidos los niveles de hormona LH, cuyos valores se muestran a continuación.

(Castillo, O.A.A., 1994)

LHN	LHP
.86	6.62
2.47	7.74
5.65	3.58
4.84	10.21
4.05	6.31
11.82	1.75
6.40	2.99
8.27	4.48
5.61	10.77
1.11	21.08
4.54	28.44
3.31	.84
3.44	6.15
3.94	3.26
7.05	6.12
2.91	3.73
12.01	6.79
7.87	4.08
4.20	5.80
4.27	10.13
2.04	11.91
13.03	1.85
10.19	35.73
2.50	2.54
6.48	9.62
9.92	2.35
7.59	11.12
4.71	5.25
4.10	10.35
9.76	9.04
9.83	5.40
6.20	14.73

**ARCHIVO ANDEVAL.MSD**

Cuatro intervalos de riego fueron probados en cuatro parcelas de caña, una intervalo (sin riego y riego a los 21, 28 y 35 días), cuyo rendimiento reportado es el siguiente:

(López, B.B., 1994)

TETS	RIE21D	RIE28D	RIE35D
1850	2874	3049	2822
1960	2620	2980	3050
2000	2500	2491	3050
1990	2450	3000	2790

**ARCHIVO ANDEVALA.MSD**

Cuatro regeguerías en zona tropical fueron comparadas por su producción de leche, para lo cual se seleccionaron al azar 3 vacas de cada regeguería, considerando la producción total de leche en litros, ajustada a 305 días.

(López, B.B., 1994)

RANCHO	PROD.
1	1850
1	1960
1	2000
2	2874
2	2620
2	2500
3	3049
3	2980
3	2491
4	2822
4	2680
4	3050

**ARCHIVO ANDEVA3.MSD**

3 razas de bovinos para carne fueron evaluadas en 4 ranchos, a travez de las ganancias diarias de peso en kilogramos, obteniendo los siguientes resultados.

(López, B.B., 1994)

GENO1	GENO2	GENO3
.96	.99	.93
.97	1.05	.95
.93	.89	.83
.97	1.01	.96

**ARCHIVO ANDEVAS.MSD**

2 tipos diferentes de alimentos, combinados con dos concentraciones de desparasitante, fueron estudiados en cuatro razas de borregos, para lo cual se utilizaron grupos de 5 animales por tratamiento y 20 animales por raza, arrojando los siguientes resultados.

(López, B.B., 1994)

RAZAS				
I	II	III	IV	
20	25	24	28	
25	30	28	31	
22	29	24	26	
27	28	25	29	
21	30	30	32	
30	30	39	40	
45	29	42	45	
30	31	36	50	
35	30	42	45	
36	30	40	60	
31	32	41	42	
30	35	45	50	
40	30	40	40	
35	40	40	55	
30	30	35	45	
20	23	24	29	
21	25	25	30	
20	28	36	28	
20	30	26	27	
19	31	23	30	

ARCHIVO REGSIK.SMD

Con el objeto de construir una curva patrón para determinar la concentración de proteína en solución, se hicieron diluciones por triplicado, utilizando albúmina de huevo disuelta en agua destilada.

DIL	CONC
1	.00
2	.25
3	.50
4	.75
5	1.0
6	1.25
7	1.5
8	1.75
9	2.0
10	2.25
11	2.5
1	.00
2	.25
3	.50
4	.75
5	1.0
6	1.25
7	1.5
8	1.75
9	2.0
10	2.25
11	2.5
1	.00
2	.25
3	.50
4	.75
5	1.0
6	1.25
7	1.5
8	1.75
9	2.0
10	2.25
11	2.5

**ARCHIVO JI2A.SMD**

Un grupo de investigación, estudiando la relación entre el tipo de sangre y el grado de cierta afección en una población, reunió datos sobre mil quinientos sujetos y deseaban saber si estos datos eran compatibles con la hipótesis de que el grado de la afección y el tipo de sangre son independientes.

(Daniel, W. W., 1983)

	A	B	AB	O
543	543	211	90	476
44	44	22	8	31
28	28	9	7	31

**ARCHIVO NPRST.MSD**

Dos dietas fueron probadas en 2 grupos de 10 cerdos cada uno, durante la etapa de crecimiento, obteniendo los siguientes aumentos de peso en kilogramos.

(López, B. B., 1994)

X DIETA 1	Y DIETA 2
17	14
17	21
21	36
18	20
17	24
14	12
19	28
17	16
16	21
12	20

## ARCHIVO NPRST2G

En este archivo se muestran las calificaciones otorgadas a 2 grupos de 20 alumnos cada uno con dos métodos distintos de enseñanza.

### CALIFICACIONES

14.5  
2.5  
18.5  
14.5  
29  
36.5  
21  
14.5  
9.5  
7  
23.5  
29  
9.5  
1  
38  
14.5  
39  
7  
29  
23.5  
18.5  
29  
25  
34.5  
21  
29  
33  
29  
7  
4.5  
21  
36.5  
14.5  
4.5  
14.5  
29  
40  
2.5  
34.5  
11

**ARCHIVO NPSRH.MSD**

Los datos de este archivo provienen de 3 grupos de ovejas que fueron alimentadas con 3 diferentes dietas. El primer grupo con 5 animales y los dos restantes con 4 ovejas cada uno, midiendo los aumentos de peso durante el experimento.

GRUPO 1	GRUPO 2	GRUPO 3
23	12	20
24	13	19
25	17	14
26	18	15
27	-	-

Una vez ordenados por rangos los datos quedan así:

RANGO	
1	9
2	10
3	11
4	12
5	13
6	1
7	2
8	5
9	6
10	8
11	7
12	3
13	4