

REVISTA PROYECCIONES N° 9: 133-177
Julio 1985 - I.S.S.N. 0716-0917.

EL PROGRAMA STAT 11

PATRICIO ROJAS MORERA *

En la actualidad existe una gran cantidad de programas interactivos; todos tienen virtudes y defectos, no por su construcción, sino desde el punto de vista del usuario, de lo que el usuario espera del programa, en este caso se analiza el programa Stat 11, disponible en el Centro de Computación de la Universidad del Norte; que reside en el disco de usuarios del equipo VAX II/750 y que es accesible por medio de terminales del tipo VT 101 - VT 102.

Este programa es útil para investigadores que usan la estadística como una herramienta básica para sus investigaciones, en el sentido que permite hacer una gran cantidad de análisis, ya sea del tipo paramétrico, o no paramétrico, como también análisis descriptivos. Una de sus grandes virtu-

* DEPARTAMENTO DE MATEMATICAS, FACULTAD DE CIENCIAS, UNIVERSIDAD DEL NORTE.

des es su potencialidad en gráficos de distintos tipos. La utilidad de este trabajo radica en que se ha ordenado en forma muy simple las distintas secuencias de comandos para ingresar datos como también para realizar los diferentes análisis.

Este programa puede trabajar con una matriz de datos definida al iniciar cada análisis, como también con submatrices, que se obtienen al eliminar columnas; también permite cambiar la variable de respuesta en forma muy sencilla, lo que hace que el programa sea muy dinámico, que es lo que generalmente interesa al investigador, probar, o como se dice a veces "jugar" un poco con los datos, buscando las relaciones entre las variables en estudio.

Todo lo anterior unido a una gran rapidez de respuesta que puede ser por terminal o por impresora hace que el STAT 11 sea recomendable para quienes deseen hacer análisis estadísticos en sus investigaciones.

La idea original al comenzar este trabajo era explicar cada análisis desde su concepción teórica hasta el análisis de resultados, en los formatos que entrega el STAT 11, algunos de los cuales no son muy explicativos, y que pueden ser cambiados; para concretar esta idea original significaría triplicar el volumen del anexo en que se encuentra explicado el programa STAT 11.

El programa STAT 11, además del ingreso de datos, puede realizar las siguientes funciones, entre otras:

1. Listado de datos.
2. Cambio de matriz de datos.
3. Cálculo de medias y desviaciones típicas
4. Prueba t.
5. Recta de regresión.
6. Recta de regresión de Bartlettts.
7. Matriz de correlación.
8. Anova de 1 criterio.

9. Anova a 2 criterios.
10. Medias y desviaciones de poblaciones logarítmicas.
11. Prueba chi-cuadrado.
12. Prueba de la mediana.
13. Prueba de Wilcoxon
14. Prueba de Mann-Whitney.
15. Prueba de Kruskall-Wallis.
16. Prueba de Spearman
17. Anova de Friedman
18. Integración numérica.
19. Tabulación cruzada.
20. Gráficos.
21. Conversión de columnas.
22. Regresión múltiple.
23. Anova multicriterio.
24. Regresión polinómica.
25. Regresión múltiple (Step wise).
26. Prueba de Newman-Keul.

En el Anexo titulado "Manual de Usuario . STAT 11", se puede encontrar la forma de trabajar con el STAT 11.

MANUAL DE USUARIO STAT 11.

En este trabajo se describe las potencialidades del paquete STAT 11 disponible en el disco de usuarios del CECUN (Centro de Computación de la Universidad del Norte), del equipo VAX II/750 y accesible por terminal del tipo VT101 - VT102.

Se ha tratado de ser lo más explicativo posible en cuanto a la forma de usar este paquete, el deseo inicial del autor era explicar cada análisis en forma detallada en cuanto a su concepción teórica, incluyendo las explicaciones de los resultados, pero este trabajo demandaría una publicación que triplicaría a ésta, por lo que se ha dejado esta idea para un próximo trabajo.

- - - - -

Para conectarse al paquete STAT 11, digite STAT 11, al hacerlo aparece la siguiente pantalla:

```
MENU STAT-11
(1) Documentación del paquete STAT
(2) Ejemplo entrada de datos
(3) Opciones para la entrada de datos
*(4) ENTRADA DE DATOS (OUTLAY)
(5) Ejemplo análisis de datos
(6) Opciones para el análisis de datos
(7) Análisis estadísticos posibles
*(8) ANALISIS (STAT)
(9) Datos desde un editor (OUTCHK)
(10) Análisis especiales (TABELL)

(99) Fin
OPCION:
```

Este es el menú principal de STAT II, en ella aparecen 11 opciones, que revisaremos secuencialmente.

1. (1) DOCUMENTACION DEL PAQUETE STAT.

Al digitar 1 aparece la siguiente pantalla

```

OPCION: DOCUMENTACION DEL PAQUETE STAT
        Proporcionada por:
                - Manual del usuario
                - Help de sistema
                - Help interactivo

        Indique opción:
                ( 1) Manual de Usuario
                ( 2) Help de sistema
                ( 3) Help interactivo
                (99) Fin documentación

OPCION:
  
```

PANTALLA 2

1.1. (1) Manual de Usuario.

Por el momento esta opción no está implementada, sin embargo puede ser solicitada a Universidad del Norte, CECUN. Casilla 1280 - Antofagasta.

Al digitar 1, STAT vuelve a la pantalla 2.

1.2. (2) Help del sistema.

Por el momento esta opción tampoco está implementada.

Al digitar 2 aparece en pantalla el listado del Help del sistema VAX, a la pregunta Topic? que aparece, digite <RETURN>, aparece nuevamente la pantalla 2.

1.3. (3) Help Interactivo.

Al digitar 3, STAT despliega la siguiente pantalla:

```
DOCUMENTACION STAT
HELP INTERACTIVO
Nota: Solo puede accederse en ambiente de
      ANALISIS (STAT) (opción 8 del menu principal)
Ejemplo:
  El programa STAT consultará por el análisis a realizar
  de la manera siguiente:
      ANALYSIS NO ? (USE 45 FOR HELP)
      Digite 45 para entrar al help interactivo.

Ingrese el dígito 1 para retornar menu Documentación.
Dígito:
```

PANTALLA 3

en el cual se indica como usar el Help Interactivo. Al digitar 1 después de esta pantalla, STAT vuelve a la pantalla 2.

1.4. (99) Fin Documentación.

Al digitar 99, STAT vuelve a la pantalla 1.

2. (2) EJEMPLO ENTRADA DE DATOS.

Al digitar 2 en la pantalla 1, STAT despliega la siguiente pantalla:

OPCION: EJEMPLO ENTRADA DE DATOS

Nota: Se debe estar en ambiente ENTRADA DE DATOS
(OUTLAY) (opción 4 del menu principal).

Dependiendo de las opciones que Ud. indicó para
la entrada de datos, el sistema consulta por:

- + código para Missing values
- + identificación (heading) del trabajo
- + tipo de formato a usar para el archivo de salida
en el caso de indicar formato no outlay.
- + archivo de entrada de datos a anexar al que se crea
- + heading para los grupos

Además pregunta por:

- + número de grupos en que se ha dividido el trabajo
(división de la muestra en submuestras)
- + número de columnas para el trabajo
(las columnas representan las variables en estudio
para la muestra)

Presione <CR> para continuar:

SIGUE:

Al presionar <RETURN> después de esta pantalla, STAT despliega la siguiente:

Si se tiene la siguiente información de peso, edad, estatura (variables= peso, edad, estatura), para hombres y mujeres (grupos= hombres, mujeres), y se desea hacer análisis estadísticos tanto de cada uno de los grupos por separado como de todos los grupos en forma conjunta, la manera de entrar los datos al archivo es:

a la consulta GROUPS? ingresar 2
COLUMNS? ingresar 3

ingresar los datos del grupo uno (hombres)

Variable1,variable2,variable3

" " " " " "

" " " " " "

" " " " " "

CTRL/¿ (fin de datos)

ingresar los datos del grupo dos (mujeres)

Variable1,variable2,variable3

" " " " " "

" " " " " "

" " " " " "

CTRL/¿ (fin de datos)

Presione <CR> para continuar!

SIGUE:

PANTALLA 5

Al presionar <RETURN> después de esta pantalla, STAT despliega la siguiente:

IMPORTANTE: Si desea heading para las columnas (variables) debe crear mediante un editor un archivo con nombres de columnas, o mediante el análisis 44 cuando está en el ambiente ANALISIS (STAT). En el primer caso los nombres de las columnas son permanentes para toda la corrida del programa y en el segundo son nombres temporales (designación dura hasta un análisis 50 ó 0 de STAT). Si se tiene el archivo de headings de columnas y se está en ambiente ANALISIS (STAT) se debe colocar la opción /c y además dar el nombre del archivo de headings.

En el ejemplo anterior se puede tener un archivo de nombres de columnas llamado col.dat que contiene lo siguiente:

```
top-file--> peso
             edad
             estatura <-- bottom-file
```

Digite un 1 para volver al menu principal

DIGITO?:

PANTALLA 6

Las pantallas 4, 5 y 6 explican el uso de la opción entrada de datos, este procedimiento se explica más adelante. Al digitar 1 después de esta pantalla, STAT vuelve a la pantalla 1.

3. (3) OPCIONES PARA LA ENTRADA DE DATOS.

Al digitar 3 en la pantalla 1, STAT despliega la siguiente pantalla:

```

OPCION: OPCIONES PARA LA ENTRADA DE DATOS

Nota: Las opciones son ingresadas de la siguiente
manera:
      [/s1][/s2].... [   ] = opcional
                        /sn   = calificador n

Si no desea ingresar opciones presiones <CR>
Con esto, asume:
      - código para missing es 0
      - entrada solo por teclado
      - se crea archivo salida con
        formato outlay.

Calificadores:
      /c :Consulta código para missing values
      /h :Consulta headings para cada grupo
      /a :   " "   archivo datos de entrada
      /q :El archivo de salida no tiene formato outlay
      /f :Consulta por un formato de datos no outlay

Digite 1 para retornar al menu principal

DIGITO?:

```

PANTALLA 7

Esta pantalla explica el uso de calificadores y valores que asume por defecto. Al digitar 1 después de esta pantalla STAT vuelve a la pantalla 1.

4. *(4) ENTRADA DE DATOS(OUTLAY).

Al digitar 4, STAT despliega la siguiente pantalla:

OPCION: ENTRADA DE DATOS (OUTLAY)

Inmediatamente después que en su pantalla tenga un párrafo de 4 líneas con información no relevante, Ud. opcionalmente puede indicar lo siguiente:

+ lista de opciones para la entrada de datos
o <CR> (Ver opción 3 del menú principal)

+ nombre del archivo de salida (debe ser con extensión .dat) o <CR> con lo que asume que el archivo de salida será el for007.dat

Digite 1 para retornar al menu principal o <CR> para continuar

PANTALLA 8

Esta pantalla indica los pasos a seguir después de digitar 4 en la pantalla 1. Al digitar 1 después de esta pantalla, STAT vuelve a la pantalla 1. Al digitar <RETURN> STAT despliega las siguientes 4 líneas que indica la pantalla 8

OUTLAY: VAX VERSION G.51- 10000. ON 1985- 3-14

NOTE !!

YES(JA) =1 (>1)

NO(NEJ) =0. (BLANK LINE)

Después de estas 4 líneas STAT queda esperando un string de calificadores que son los indicados en la pantalla 7 (ver apartado 4.1) en este momento, se puede digitar el string de caracteres, o poner <RETURN> con lo que STAT asume los valores por defecto indicados también en la pantalla 7, después de cualquiera de las opciones anteriores (string de calificadores o <RETURN>), STAT queda esperando el nombre del archivo de salida que debe ser del XXXXXXXXXXXX.DAT, donde las x's indican el máximo de caracteres alfanuméricos.

Después de poner el nombre del archivo de salida o <RETURN> (con lo que asume que el archivo de salida tendrá el nombre de FOR007.DAT), STAT consulta por el número de grupos (NO OF GROUPS), según se explica en la pantalla 5, después de digitar el número de grupos, STAT consulta por el número de columnas (NO OF COLUMNS). Después de digitar el número de columnas STAT consulta por el título que identifica los datos (MAIN HEADING:) que puede ser de cualquier extensión, después de digitar el MAIN HEADING, aparecen las siguientes dos líneas

WAIT

INPUT X-VALUES FOR GROUP 1 AND FINISH WITH EOF <^Z>

Y STAT consulta por los datos del primer grupo que deben ser ingresados por filas y separados por comas (cada quien debe contener tantas columnas como se haya indicado en la consulta NO OF COLUMNS). Después de terminado el ingreso de datos del grupo 1, se digitan las teclas CTRL y Z simultáneamente, con lo cual se indica a STAT que ha terminado de ingresar el primer grupo. Después de esta operación (<^Z>), si a la consulta NO OF GROUPS, se ha contestado con un dígito mayor que 1, STAT despliega las siguientes líneas:

INPUT DATA END

N POWS TOTAL N ROWS IN BUFFER

INPUT X-VALUES FOR GROUP 2 AND FINISH WITH EOF <^Z>

donde N indica el número de filas que Ud. ingresó en el grupo anterior. Ahora STAT espera el ingreso de datos para el grupo siguiente.

Después de ingresar el último grupo de datos STAT despliega las siguientes líneas

INPUT DATA END

N ROWS TOTAL M ROWS IN BUFFER

LIST? <0/1>

donde N indica el total de filas ingresadas y M el total de filas ingresadas en el último grupo. A la consulta LIST? <0/1>, poner 0 indica que no desea listado de los datos por pantalla, poner 1 indica que se desea un listado por pantalla que es presentado de la siguiente forma:

RON GROUP N₁ N₂ N₃

donde ROW indica la fila, GROUP el grupo y N_1, N_2, N_3, \dots los datos de la fila. En cualquiera de los dos casos (0 ó 1) STAT consulta si se ha cometido errores en la entrada de datos de la siguiente forma:

ERRORS?

ROW,COL,VALUES... <END= ROW 0>

Aquí se pueden arreglar los errores cometidos al digitar los datos, para lo cual debe colocarse un string de números en el orden ROW,COL,VALUES es decir la fila y la columna en que se ha cometido el error y luego el valor correcto, todo separado por comas. Después de corregir todos los posibles errores, se pone un cero al iniciar la líneas y nuevamente STAT despliega la líneas LIST? <0/1> si se responde con cero, consulta por ERRORS? y si se vuelve a digitar cero, STAT vuelve a la pantalla 1 y crea el archivo de datos con el nombre que Ud. ha ingresado previamente o le asigna el nombre FOR007.DAT si Ud. no ha ingresado el nombre del archivo de datos.

4.1. Uso de los calificadores en la opción 4.

Los calificadores que se pueden usar son los indicados en la pantalla 7, es decir

/C: código para datos missing
 /H: títulos para cada grupo
 /A: archivo de datos de entrada
 /Q: formato del archivo de salida
 /F: entrada de datos con formato especial

4.1.1. Uso de /C (código para datos missing)

Al usar /C, STAT consulta por

NO OF GROUPS
 NO OF COLUMNS
 MAIN HEADING

que se responden según lo que se ha explicado anteriormente.

título de los datos, para lo cual STAT hace las consultas

NO OF GROUPS

NO OF COLUMNS

MAIN HEADING

Después de responder a estas consultas STAT, despliega las dos líneas siguientes

WAIT

INPUT FILE UNIT NO; MUST BE 8 TO 9; <0 IF NON>? (2)

con lo cual consulta si existe o no un archivo de datos.

4.1.3.1.1. Uso de 0 en la línea (2)

Al digitar 0 en la línea (2) se está indicando que no existe un archivo de datos. Después de digitar 0 STAT despliega las siguientes dos líneas

NEW INPUT?

[TTY=4; OTHERS STATE UNIT NR <8 TO 9] (3)

con lo cual consulta si desea agregar nuevos datos ya sea por terminal o por archivos.

4.1.3.1.1.1. Uso de <RETURN> en la línea (3)

Al digitar <RETURN> en la línea (3) se está indicando que no se ingresarán nuevos datos, después de digitar <RETURN> en la línea (3), STAT despliega la siguiente línea

O.K. TO CLOSE FILE WITH NO OUTPUT? (4)

a la cual se debe responder obligadamente YES, con lo cual STAT vuelve a la pantalla 1. Cualquier otra respuesta distinta de YES, produce un error.

4.1.3.1.1.2. Uso de 4 en la línea (3)

Al digitar 4 en la línea (3) se está indicando que se ingresarán nuevos datos por terminal. Al digitar 4 en la línea (3) STAT despliega la siguiente línea

ENTER GROUP NR AND X VALUES END WITH BLANK LINES (5)

con la cual pide que se ingresen los datos, los que se deben ingresar por filas separados por comas y el primer dígito debe corresponder al grupo al cual pertenece la fila de datos. Se termina el ingreso de datos con una línea en blanco, después de lo cual STAT despliega las siguientes líneas

INPUT DAT END

N ROWS TOTAL N ROWS IN BUFFER

LIST? <0/1> (6)

y se sigue el procedimiento indicado en el apartado 4, con la diferencia que al digitar 0 en la línea

ROW,COL,VALUES... <END=ROW 0>

STAT despliega las siguientes líneas

WAIT

NEW INPUT?

ONLY TTY (UNIT 4) ALLOWED (7)

con lo cual consulta si desea ingresar nuevos datos.

4.1.3.1.1.2.1. Uso de <RETURN> en la línea (7)

Al digitar <RETURN> en la línea (7) se está indicando que no hay nuevos datos, con lo cual STAT despliega la pantalla 1 y cierra el archivo de datos.

4.1.3.1.1.2.2. Uso de 4 en la línea (7)

Al digitar 4 en la línea (7) se está indicando que se digitarán nuevos datos por terminal. Al digitar 4 en la línea (7) STAT despliega la línea (5) y se sigue el procedimiento indicado en el apartado 4.1.3.1.1.2.

4.1.3.1.1.3. Uso de los números 7, 8 ó 9 en la línea (3)

Al digitar los números 7, 8 ó 9 (todos producen el mismo efecto) en la línea (3) se está indicando que se ingresarán nuevos datos por archivos. Después de digitar algunos de estos números, STAT espera por el

nombre del archivo, una vez ingresado este y si el formato de los datos es correcto, STAT los agrega a la cola del archivo de datos que se está generando y después despliega la línea (6) y se sigue el procedimiento indicado en el apartado 4, con la diferencia que al responder con 0 a la línea ROW,COL,VALUES... <END=ROW 0>

STAT despliega la línea 3 y se sigue el procedimiento indicado en este apartado y en los apartados 4.1.3.1.1.1. y 4.1.3.1.1.2.

4.1.3.1.2. Uso de 8 ó 9 en la línea (2)

Al digitar los números 8 ó 9 (ambos producen el mismo efecto) en la línea (2) se está indicando que los datos están en un archivo, después de digitar 8 ó 9 STAT queda esperando el nombre del archivo (el archivo debe estar en el formato OUTLAY que usa STAT 11, cualquier otro formato provocará un error en la lectura). Después de digitar el nombre del archivo y si el formato de los datos está correcto, STAT despliega la línea (6) y se sigue el procedimiento indicado en el apartado 4, con la diferencia que al digitar 0 en la línea ROW,COL,VALUES ... <END=ROW 0> STAT despliega la línea (2) y se sigue el procedimiento indicado en los apartados 4.1.3.1.1. y 4.1.3.1.2.

4.1.3.2. Uso de 4 en la línea (1)

Al digitar 4 en la línea (1) se está indicando esencialmente lo mismo que con la opción 0, pero además STAT consulta por los títulos de cada grupo con la línea

N GROUP HEADING ON SEPARATE LINES (8)

Después de ingresar los títulos de cada grupo, STAT despliega la línea (2) y se sigue el procedimiento indicado en los apartados 4.1.3.1.1. y 4.1.3.1.2.

4.1.3.3. Uso de 8 ó 9 en la línea (1)

Al digitar 8 ó 9 en la línea (1) se está indicando que los datos están en un archivo de datos (en este caso se copian los datos tal

como aparecen en el archivo). Después de digitar 8 ó 9 STAT queda esperando que se ingrese el nombre del archivo, una vez ingresado éste, STAT despliega la línea (2), a la cual se debe responder con 8 ó 9 (según lo que se haya digitado en la línea (1)), después de esto STAT despliega la línea (6) y se sigue el procedimiento indicado en el apartado 4, con la diferencia que al digitar 0 en la línea

```
ROW,COL,VALUES ... <END=ROW 0>
```

STAT despliega la línea (2) y se sigue el procedimiento indicado en los apartados 4.1.3.1.1. y 4.1.3.1.2.

4.1.4. Uso de /Q (formato de archivo de salida)

Al usar /Q STAT consulta sólo por el número de columnas (NO OF COLUMNS), es decir bajo este calificador no se puede trabajar con grupos, después de ingresar el número de columnas, STAT despliega la línea

```
INPUT X-VALUES FOR GROUP 1 AND FINISH WITH EOF <^Z >
```

con lo cual pide que se ingresen los datos por filas y separados por comas, se termina el set de datos con <CTRL>Z y STAT despliega la línea (6) y se sigue el procedimiento indicado en el apartado 4, con la diferencia que al digitar 0 en la fila ROW,COL,VALUES... <END=ROW 0> STAT despliega la pantalla 1.

4.1.5. Uso de /F (entrada de datos con formato especial)

Este calificador hace lo mismo que /A

5. (5) EJEMPLO DE ANALISIS DE DATOS.

Al digitar 5 en la pantalla 1 STAT despliega las dos pantallas siguientes:

OPCION: EJEMPLO ANALISIS DE DATOS

Nota: Se debe estar en ambiente ANALISIS (STAT)
(Opción 8 del menu principal)

Ejemplo: El programa STAT consultará por el analisis
a realizar de la manera siguiente:

ANALYSIS NO ? (USE 45 FOR HELP)

Ud. debe digitar un string de la forma:
T,A1,A2,A3,....

En que T es un dígito que indica el análisis
a realizar (Ver help interactivo o manual de usuario)

A1,A2,A3... son dígitos cuyo significado varía de
un análisis a otro (Ver help interactivo o manual de
usuario).

Presiones <CR> para continuar

SIGUE:

PANTALLA 9

Si se indica:

- 0 se desea considerar otro grupo y otras columnas
para efectuar los análisis
- 1 se pide la media y el error para los grupos y
columnas consideradas
- 5,A1 se pide análisis de correlación
- 22,A1,A2 pide transformaciones temporales de los
datos

Digite 1 para volver al menu principal

DIGITO?:

PANTALLA 10

Al digitar 1 en la pantalla 10 STAT despliega la pantalla 1.

6. (6) OPCIONES PARA EL ANALISIS DE DATOS.

Al digitar 6 en la pantalla 1 se despliega la siguiente pantalla:

OPCION: OPCIONES PARA EL ANALISIS DE DATOS

Nota: Las opciones son ingresadas como se indica:

[/S1][/S2]....[/Sn] []= opcional
/Sn = calificador n

Si desea no ingresar opciones presione <CR>, con lo que asume salida por terminal y los análisis a efectuar se darán por teclado.

Calificadores:

/L salida por impresora
/n si se tiene además /L, inhibe salida por impresora y el resultado de los análisis se guarda en archivo creado por STAT.
/p no paginación
/k los análisis a efectuar se dan por archivo
/c los headings de columna se dan por archivo
/q el archivo de datos no tiene formato Outlay

Obs: Una lista completa de los calificadores se encuentra en el Help del sistema.

Digite 1 para retornar al menu principal

DIGITO?:

PANTALLA 11

Al digitar 1 en la pantalla 11 STAT despliega la pantalla 1.

7. (7) PRINCIPALES ANALISIS DE STAT.

Al digitar 7 en la pantalla 1, STAT despliega la siguiente pantalla:

OPCION: Principales análisis de STAT			
ANALISIS	SUBJECT	ANALISIS	SUBJECT
Media y error standar	summary	Chi-cuadrado	non_para
Media geométrica	" "	Test mediana de Fisher	" "
T test	t tes	Test Wilcoxon	" "
Regresion lineal	regress	Test Mann Whitney	" "
Regresión de bartlette	" "	Análisis Kruskal Wallis	" "
Regresión múltiple	" "	Correlación Spearman	" "
Regresión polinomial	" "	Análisis Friedman	" "
Regre. mult. paso a paso	" "	Mediana y grado	" "
Correlación	correlat	Test múltiple Neuman-Keul	" "
Análisis varianza un lado	anova	Análisis regresión Probit	" "
Análisis. vnza. sheffe	" "	Integración numérica	integrat
Análsis. vnza. dos lados	" "	Histograma y crosstabs	descript
Análsis. vnza. multilado	" "	Scatter plot	" "
Análsis. vnza. anidado	" "	Crosstabs múltiple	" "
Análsis. vnza. 2 nivel	" "	Reducción y clas.matriz	data_tra

NOTA: Ver formato en help interactivo mediante campo subject!

Digite 1 para retornar al menu principal

DIGITO?:

PANTALLA 12

Al digitar 1 en la pantalla 12, STAT despliega la pantalla 1.

8. (8) ANALISIS (STAT)

Al digitar 8 en la pantalla 1, STAT despliega la siguiente pantalla:

OPCION: ANALISIS DE DATOS (STAT)

ENTER PARAMETER(S)
 {IT MAY BE AS FOLLOWS: [/L), [/N] ([] = OPTIONAL)}

PANTALLA 13

Esta pantalla está pidiendo los parámetros que determinarán el ámbito de los cálculos. Estos parámetros se colocan como un string de la siguiente forma

$$/P_1/P_2 \dots / P_n$$

donde $P_1, P_2 \dots P_n$ son parámetros (todos opcionales) y que se detallan a continuación.

8.1. Uso de Parámetros.8.1.1. Uso de /L.

Esta opción deja los resultados en un archivo generado por STAT, con el nombre STAT LIST.LST; no muestra los resultados por pantalla. Vuelve iterativamente a la pantalla 15.

8.1.2. Uso de /K.

Esta opción indica que los análisis a efectuar se dan por archivo, para lo cual previamente se debe generar un archivo con la siguiente estructura

NUMERO DE GRUPOS
NUMERO DE COLUMNAS
ANALISIS 1
ANALISIS 2
:
:
:
:
:
:

Cuando se activa este parámetro STAT consulta por el nombre de este archivo y luego por el archivo de datos. Con esta opción se muestran los resultados por pantalla, pero no genera archivo de resultados. Vuelve a la pantalla 1.

8.1.3. Uso de /N.

Esta opción permite ver los resultados por pantalla, pero no genera archivo de resultados. Vuelve iterativamente a la pantalla 15.

8.1.4. Uso de /P.

Esta opción indica a STAT que no numere las páginas del archivo de datos. Vuelve iterativamente a la pantalla 15.

8.1.5. Uso de /C.

Esta opción indica que los títulos (headings) de las columnas se entregan por archivos, para lo cual previamente se debe generar un archivo con la siguiente estructura:

TITULO DE LA COLUMNA 1
TITULO DE LA COLUMNA 2
:
:
:
:
:

Cuando se activa este parámetro, STAT consulta por el nombre de este archivo y luego por el archivo de datos. Vuelve iterativamente a la pantalla 15.

8.2. Ingreso de Nombre de Archivos.

Después de haber ingresado los parámetros STAT despliega la siguiente pantalla

```

ENTER FILE(S): THE FIVE ORDER IS:
      REQ-FILE  (OPTIONAL)
      HEAD-FILE (OPTIONAL)
      DATA-FILL (NO OPTIONAL)

```

PANTALLA 14

en la cual pide el nombre de archivos en el siguiente orden:

- (i) Archivos con títulos de columnas, si se ha activado el parámetro /C.
- (ii) Archivo de análisis, si se ha activado el parámetro /K.
- (iii) Nombre del archivo de datos (este es obligatorio)

8.3. Determinación de la matriz de datos.

Se puede trabajar con todo o parte del archivo de datos. Después de ingresar los nombres de los archivos, STAT consulta por el número de grupos (GROUPS?) y el número de columnas (COLUMNS?), determinando así la matriz de datos, para ambos casos se usan los siguientes formatos:

- (i) n: indica que se consideraron en los análisis los n primeros grupos o columnas.
- (ii) n,m: indica que se considerarán los m primeros grupos o columnas a partir de n, por ejemplo: si la consulta COLUMNS? digitamos 4, 3; STAT considerará las columnas 4, 5 y 6.

8.4. Análisis a realizar.

Después de determinar la matriz de datos STAT despliega la siguiente pantalla (siempre que no se haya activado /K)

```

MFANS:      J,NOR -2,F,M,L,U,D
REGRESSION: 3,K1,K2,01,02 or 4,K1,K2 or 30,K1,IOPT
ANALYSIS NO? (use 45 for HELP)

```

PANTALLA 15

Después de solicitar el análisis con el formato adecuado STAT vuelve iterativamente a la pantalla 15, salvo que se indique un análisis 50 ó 0 (ver más adelante). En todos los análisis que se explican a continuación los parámetros marcados con (< >) son opcionales.

8.4.1. Listado de datos.

Los análisis -1, -2, -3 generan distintos tipos de listados de los datos.

8.4.1.1. Uso de -1.

El análisis -1, genera un listado de los datos.

FORMATO: -1, <FORM>,<LOWER>,<UPPER>,<DECIMAL5>

FORM= 0 : consulta por el formato de salida de los datos (en formato por cada columna)

FORM= 1 : usa formato F6.1

FORM= 2 : usa formato F10.4

FORM= 3 : usa formato E12.4 (sólo válido con los análisis -2 y -3)

LOWER= n : considera todas las columnas a partir de la n-ésima

UPPER= m : considera desde la columna n hasta la columna m

DECIMAL5= α : lista los datos con α lugares decimales, independientemente del formato utilizado en el parámetro FORM, siempre que α sea

menor que el número de lugares decimales indicados en el formato.

Los valores por defecto son los siguientes

FORM : consulta por el formato de salida
 LOWER : considera desde la primera columna
 UPPER : considera hasta la última columna
 DECIMAL5 : considera el número de decimales indicados en formato.

8.4.1.2. Uso de -2.

Genera un listado de datos en las mismas condiciones que -1, pero agrega al final una estadística de los datos.

FORMATO: -2, <FORM>, <MISS>, <LOWER>, <UPPER>, <DECIMALS>

Los parámetros se usan según se indicó en el apartado 8.4.1.1., para MISS usamos

MISS = -1: lista las líneas que no contengan valores missing (los valores missing se indican al general el archivo de datos).
 MISS = 1 lista todas las líneas, indicando los datos missing con con la palabra MISSING.

El valor por defecto es

MISS : considera MISS = 1.-

8.4.1.3. Uso de -3.

Genera un listado de datos, sin la estadística final, y en las mismas condiciones que -2.

FORMATO: -3, <FORM>, <MISS>, <LOWER>, <UPPER>, <DECIMALS>

Los parámetros se usan según lo indicado en los apartados 8.4.1.1. y 8.4.1.2.

8.4.2. Uso de 0.

Permite determinar una nueva matriz de datos a partir del archivo de datos con que se está trabajando. Después de solicitar el análisis 0, seguir el procedimiento que se indica en el apartado 8.3.

8.4.3. Cálculo de medias y desviaciones típicas.

El análisis 1 permite realizar varios cálculos relacionados con la media y las desviaciones típicas.

8.4.3.1. Uso de 1.

El análisis 1 calcula media, desviación típica, coeficiente de variación, valor mínimo, valor máximo para cada columna y además indica el número de valores missing en el set de datos.

8.4.3.2. Uso de 1, -1.

El análisis 1, -1, calcula lo mismo que el análisis 1, pero no considera las filas que contengan valores missing.

8.4.3.2. Uso de 1, K.

El análisis 1, K donde K es un número de columna calcula Σx y Σx^2 para los datos de la columna K.

8.4.3.3. Uso de 1, K_1 , K_2 .

El análisis 1, K_1 , K_2 donde K_1 y K_2 son dos números de columnas calcula Σx , Σy , Σxy , $\Sigma(x-y)$, Σx^2 , Σy^2 , $\Sigma(x-y)^2$ donde x son los datos de la columna 1, e y los datos de la columna 2.

8.4.3.4. Uso de 1, K_1 , K_2 , -1.

El análisis 1, K_1 , K_2 , -1 hace lo mismo que el análisis 1, K_1 , K_2 , pero no considera las filas que contienen datos missing.

8.4.4. Uso de 2.

El análisis 2 realiza una prueba 5

FORMATO: 2, K_1 , K_2 , N.

Los parámetros indican lo siguiente:

$K_1 = n$: indica la columna n

$K_2 = m$: indica la columna m

N = 1 : la prueba t se realiza según la fórmula standard, en la cual se suponen varianzas agudas para cada columna.

N = 2 : la prueba 5 se realiza suponiendo que las varianzas no son iguales. Internamente realiza una prueba de igualdad de varianzas previa a la realización del análisis.

N = 3 : la prueba t se realiza considerando los datos como datos aparejados. En este caso se requiere que las columnas tengan el mismo número de datos.

8.4.5. Uso de 3.

El análisis 3 determina los parámetros de la recta de regresión sus desviaciones típicas, además de la suma de cuadrados explicada y no explicada, el valor t para la prueba de hipótesis.

FORMATO : 3, K_1 , K_2 , <IOPT1>, <IOPT2>

Los parámetros son los siguientes

$K_1 = n$: indica la columna n

$K_2 = m$: indica la columna m

- IOPT1 = 1 : agrega un listado con los valores de la columna m, calcula según la recta de regresión con los datos de la columna n.
- IOPT2 = 2 : consulta por valores específicos de la variable independiente (m) y entrega el valor de la variable dependiente (n). Este último valor es entregado como un intervalo de confianza.

Los valores por defecto son los siguientes

- IOPT1 : no agrega listado de valores esperados para la variable de respuesta.
- IOPT2 : no consulta por valores específicos de la variable independiente.

8.4.6. Uso de 4.

El análisis 4 entrega la recta de regresión de Bartlett y los intervalos de confianza para la pendiente, además del valor 5 para la prueba de hipótesis.

FORMATO: 4, KOLX, KOLY

Los parámetros son

- KOLX = n : indica que n es la columna que contiene los datos de la variable independiente.
- KOLY = m : indica que m es la columna que contiene los datos de la variable dependiente.

8.4.7. Uso de 5.

El análisis 5 entrega una matriz de correlación de todas las columnas y una matriz de datos aparejados.

FORMATO : 5,<IOPT>

IOPT = -1 : calcula la matriz de conclusiones eliminando las filas que contienen datos missing.

8.4.8. Uso de 6.

El análisis 6 calcula una tabla de análisis de varianza (ANOVA) típica de un factor.

FORMATO : 6

8.4.9. Uso de 8.

El análisis 8 calcula una tabla de análisis de varianza (ANOVA) de dos factores. No considera las filas con datos missing.

FORMATO : 8

8.4.10. Uso de 9.

El análisis 9 calcula la media y desviación típica de datos provenientes de poblaciones logarítmicas. Además entrega intervalos de confianza para la media.

FORMATO : 9

8.4.11. Uso de 11.

El análisis 11, realiza una prueba chi-cuadrado para datos conjugados por terminal, en contraste con los otros análisis, éste pide que los datos se ingresen en el momento de realizar el análisis.

FORMATO : 11, NCOL, NROW

NCOL = n : indica que la matriz de datos tiene n columnas.

NROW = m : indica que la matriz de datos tiene m filas.

8.4.12. Uso de 12.

El análisis 12, realiza el test de la mediana, usando Fisher o chi-cuadrado sobre muestras independientes. Fisher se usa si $N < 40$ ó si alguna frecuencia es menor que 5, en caso contrario usa chi-cuadrado.

FORMATO : 12, K_1 , K_2 , <IOPT1>, <IOPT2>

IOPT1 = 1 : tabula medianas y cuartiles

IOPT1 = -1 : no usa la mediana común para realizar la prueba sino que solicita algún valor que desee el usuario usar como "break-value"

IOPT1 = -2 : solicita por terminal un campo de 4 caracteres numéricos para realizar el test

IOPT2 = 1 : siempre usa chi-cuadrado

IOPT2 = -1 : siempre usa Fisher

Los valores por defecto son

IOPT1 :

IOPT2 :

8.4.13. Uso de 13.

El análisis 13 realiza el test de Welcoxon, que es el equivalente no paramétrico del test t de datos aparejados.

FORMATO : 13, K_1 , K_2 .

8.4.14. Uso de 14.

El análisis 14 realiza el test de Mann-Whitney, que es el equivalente no paramétrico del test t de datos no aparejados.

FORMATO 14, K_1 , K_2 .

8.4.15. Uso de 15.

El análisis 15, realiza el test de Kruskal-Wallis que es el equivalente no paramétrico del ANOVA de un factor.

FORMATO : 15, <IOPT>

IOPT = -1 : elimina las filas con datos missing y agrega una tabla de diferencias.

8.4.16. Uso de 16.

El análisis 16 realiza el test de correlación de rangos de Spearman.

FORMATO : 16, K_1 , K_2 .

8.4.17. Uso de 17.

El análisis 17 realiza el análisis de Friedman, que es el equivalente no paramétrico del ANOVA de dos factores.

FORMATO : 17

8.4.18. Uso de 18.

El análisis 18, realiza integración numérica.

FORMATO : 18, K_1 , K_2 , <ARG1>, <ARG2>

$K_1 = n$: indica que la columna n será usada como los valores de x .
 $K_2 = m$: indica que la columna m será usada como los valores de y .
 ARG1 = 0 : entrega sólo el área final.
 ARG2 = 1 : además entrega una tabla de áreas parciales
 ARG3 = 1 : calcula el área por el método de trapecios modificados, des_
 pues de un "break-valur" solicitado por terminal.

Los valores por defecto son

ARG1 : entrega sólo el área final
 ARG2 : no pide "break-value"

8.4.19. Uso de 19.

El análisis 19, calcula mediana para cada columna y ordena las columnas (rank).

FORMATO : 19, <IOPT>

IOPT = 0 : calcula las medianas para todas las columnas, sin considerar los datos missing.
 IOPT = n : ordena la columna (rank) n .

El valor por defecto es

IOPT : IOPT = 0

8.4.20. Uso de 20.

El análisis 20, realiza una tabulación cruzada o un histograma.

FORMATO : 20, K_1 , K_2 , <NINT>

$K_1 \neq K_2$: hace la tabulación cruzada entre las columnas K_1 y K_2

$K_1 = K_2$: hace el histograma para la columna K_1

NINT = k : usa k intervalos tanto para la tabulación cruzada como para el histograma, ($3 \leq k \leq 10$)

El valor por defecto es

NINT : NINT = 5

8.4.21. Uso de 21.

El análisis 21, realiza gráficos de una variable contra otra.

FORMATO: 21, K_1 , K_2 , <IOPT1>, <IOPT2>, <IOPT3>

$K_1 = n$: los datos de la columna n se usan como eje horizontal. Este eje se divide en 50 partes.

$K_2 = m$: los datos de la columna m, se usan como eje vertical. Este eje puede dividirse desde 20 hasta 50 partes.

IOPT1 = k: divide en k partes los datos de la columna K_2 (eje vertical) $10 \leq k \leq 50$.

IOPT2 = 1: permite cambiar de ubicación el punto de mayor coordenadas, produciendo un reordenamiento del gráfico.

IOPT3 = -1: permite trabajar con una escala logarítmica en el eje vertical.

Los valores por defecto son:

IOPT1 : divide los datos de la columna K_2 (eje vertical) en 25 partes.

IOPT2 : ubica la coordenada mayor en la parte superior del gráfico.

IOPT3 : no usa escala logarítmica.

8.4.22. Uso de 22.

El análisis 22 permite cambiar los datos en el archivo de trabajo temporalmente, los datos originales permanecen invariables.

FORMATOS : (1) 22, K_1

$K_1 = n$: la columna n es ordenada en forma ascendente. Los valores miss
ing se ponen al final de la lista

(2) 22, $-K_1$

$K_1 = n$: la columna n es ordenada en forma descendente. Los valores miss
ing se ponen al final de la lista

(3) 22, NK, K_1 , -1

NK = n : indica el número n total de columna del archivo original.

$K_1 = m$: reduce el número n total de columnas a m columnas

(4) 22, K_1 , K_2 , CODE

KOL1 = n : la columna n es reemplazada por la columna m (ingresada por
KOL2 = m) según lo indicado por CODE.

KOL2 = m : los datos de la columna m, se ponen en la columna n (ingresa-
da por KOL1 = n) según lo indicado por CODE.

KOL2 = -1: en este caso se ignora la columna n (ingresada por KOL1 = n)
y el cambio se hace en todas las columnas.

CODE = k : produce el cambio indicado por k según la lista siguiente:

1 raíz cuadrada

2 exponencial

3 logaritmo natural

4 logaritmo decimal

5 cuadrado

6 igual (copia la columna n en la columna m)

- 7 diferencia entre las columnas m y n
- 8 suma de las columnas m y n
- 9 cociente de las columnas
- 10 producto de las columnas
- 11 numera las columnas
- 12 pone la columna n constante
- 13 indica el número de observaciones no missing por fila
- 14 suma las observaciones no missing por filas.

8.4.23. Uso de 23.

El análisis 23 produce un intercambio de filas y columnas.

FORMATO : 23

8.4.24. Uso de 30.

El análisis 30 realiza una regresión múltiple, usando todos los datos en la matriz de trabajo.

FORMATO : 30, K <IOPT>

K = n : indica que la columna n es la variable de respuesta

IOPT=k : donde $k > 0$, imprime la tabla de residuos.

Valor por defecto

IOPT : no imprime la tabla de residuos.

8.4.25. Uso de 31.

El análisis 31, realiza un análisis de varianza de dos factores y para varios niveles.

FORMATO ; 31, n

n : indica el número de niveles.

8.4.26. Uso de 32.

El análisis 32 realiza una regresión polinomial.

FORMATO : 32, NX, NY, MAX, <IOPT>

NX = n : indica que la columna n es considerada como la variable dependiente.

NY = m : indica que la variable m es considerada como la variable dependiente.

MAX = k : $1 \leq n \leq 4$, indica el grado del polinomio de regresión.

IOPT = ℓ : ℓ ≠ 0, imprime al final una table de residuos.

El valor por defecto es:

IOPT : no imprime la tabla de residuos.

8.4.27. Uso de 34.

El análisis 34, realiza una regresión múltiple paso a paso (STEPWISE).

FORMATO : 34, <IOPT1>, <IOPT2>

IOPT1 = n : n ≠ 0, genera tabla de residuos.

IOPT2 = m : m ≠ 0, consulta por un valor de la varianza explicada bajo la cual no continúa entrando variables forzadas.

Después de introducir el formato, consulta por un string de nk caracteres, (nk = número de columnas), con los siguientes códigos:

- Variable independiente forzada a entrar en el análisis.
- 2 Variable que no debe considerarse en el análisis.
- 3 Variable dependiente.

Los valores por defecto son:

IOPT1 : no genera tabla de residuos.

IOPT2 : el valor que asume es 00 (es decir entran todas las variables).

8.4.27. Uso de 46.

El análisis 46, permite cambiar las probabilidades con que trabaja STAT para los análisis.

8.4.28. Uso de 47.

El análisis 47 permite cambiar el valor para datos missing.

8.4.29. Uso de 48.

El análisis 48 permite hacer comentarios (escribir textos) al final del listado de resultados.

8.4.30. Uso de 50.

El análisis 50 produce la salida a la pantalla 1.

8.4.31. Uso de 70.

El análisis 70 permite leer un nuevo archivo de trabajo, tal como se pide en la pantalla 14. (ver página 159).

9. (9) DATOS DESDE UN EMISOR (OUTCHK)

Al digitar 9, aparece la siguiente pantalla

OPCION: DATOS DESDE UN EDITOR (OUTCHK)

La entrada de datos (OUTLAY) puede ser proporcionada por un archivo creado por un editor. El programa OUTCHK verifica la corrección de los datos que han sido creados bajo esta modalidad.

La salida proporcionada por este programa es un listado de todas las líneas que presentan errores.

Los datos que deben ser dados al programa son los siguientes (y en este orden):

Indicar opcion 1
SWITCHES
NOMBRE ARCHIVO A ANALIZAR
NUMERO DE COLUMNAS
FORMATO (si coloca /f)

SWITCHES:

/A = Liste todas las líneas (asume listar solo líneas incorrectas)
/L = Salida por impresora (implica /A)
/f = Formato para lectura del archivo (asume 15F11.0)

INDIQUE OPCION:

(1) Análisis archivo
(99) Retornar menu principal

OPCION :

PANTALLA 16

9.1. Uso de 1.

Al digitar 1 en la pantalla 16, STAT consulta por SWITCHES, NOMBRE DE ARCHIVO A ANALIZAR, NUMERO DE COLUMNAS y FORMATO (sólo si se ha usado el switch /f) un parámetro en cada línea. Los switches son

/A : lista todas las líneas
 /L : salida por impresora (cuando se ha usado /A)
 /f : formato de lectura por archivo

Los valores por defecto son:

/A : solo listas líneas incorrectas
 /L : muestra los resultados por pantalla
 /f : usa el formato 15F11.0

9.2. Uso de 99.

Al digitar 99 en la pantalla 16, vuelve a la pantalla 1.

10. (10) ANALISIS ESPECIALES (TABELL)

Al digitar 10 en la pantalla 1, STAT despliega la siguiente pantalla

OPCION: ANALISIS ESPECIALES (TABELL)

Este programa es usada en lugar de una tabla estadística. Permite conocer la probabilidad de un valor obtenido por un test estadístico.

La entrada al programa es la siguiente:

!P,N1,N2

En que ! = indica el tipo de tabla a usar puede ser:

T para T-valor
 F para F-valor
 X para CHI-2 (probabilidad)
 Z para valor normal
 C cálculo de chi-2 (Ccolumna, fila)
 E FIN DEL PROGRAMA

P = es la probabilidad

N1 y N2 = son los grados de libertad (N2 sólo para F, N1 no usado para Z)

INDIQUE OPCION:

(1) análisis (TABELL)

(99) retorno menu principal

OPCION :

10.1. Uso de 1.

Al digitar 1 en la pantalla 17, STAT consulta la probabilidad de seada, según se explica en la pantalla 17.

10.2. Uso de 99.

Al digitar 99 en la pantalla 17, STAT vuelve a la pantalla 1.

11. (99) FIN.

Al digitar 99 en la pantalla 1, STAT libera el terminal y lo deja a disposición del usuario.

12. RECOMENDACIONES PARA EL USUARIO.

La forma más cómoda para trabajar con STAT es mandar a digitar los datos a CECUN, indicando el nombre del archivo. Este archivo puede ser digitado en formato libre, pues usando la opción 4 de STAT, se deja en formato OUTLAY, con la cual trabaja STAT. Para realizar análisis sobre los datos, sólo para ver en pantalla use la siguiente secuencia de instrucciones:

```

$ STAT 11
8
/N
<NOMBRE DEL ARCHIVO DE DATOS>
ANALISIS 1
ANALISIS 2
      :
      :
ANALISIS N
00 (fin del análisis)
99 (salida de STAT)

```

Si desea generar archivo de resultados, use la siguiente secuencia de instrucciones:

```
$ STAT 11
8
/L
<NOMBRE DEL ARCHIVO DE DATOS>
ANALISIS 1
ANALISIS 1
      .
      .
      .
ANALISIS N
50 (fin de análisis)
99 (salida de STAT)
```

Ahora los resultados están disponibles en un archivo
STAT LIST. LST.