

24
9



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

FACULTAD DE CIENCIAS

**UNA APLICACION DE LA COMPUTACION EN LA
ENSEÑANZA DEL MUESTREO**

T E S I S

QUE PARA OBTENER EL TITULO DE

A C T U A R I O

P R E S E N T A

JOSE PABLO CASTAÑEDA MARTINEZ

MEXICO, D. F.

1986



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.

UNA APLICACION DE LA COMPUTACION EN
LA ENSEÑANZA DEL MUESTREO

T E M A R I O

1.- ANTECEDENTES

2.- INTRODUCCION

3.- PRESENTACION DE OBJETIVOS

4.- DESCRIPCION DEL SISTEMA

A) Enfoque del Muestreo :

- a.1) Descripción, usos y diseños.
- a.2) Diseños de Muestreo.
- a.3) Expresión de los estimadores de los diseños.
- a.4) Diseños que se pueden incluir, modificaciones y sugerencias.
- a.5) Consideraciones que se deben tener en cuenta para la interpretación de resultados.

B) Enfoque de Cómputo :

- b.1) Descripción general, usos y diagrama del sistema.
- b.2) Clasificación del Sistema.
- b.3) Capacidades de almacenamiento y recomendaciones de uso.
- b.4) Procedimientos para el "cargado del sistema".
- b.5) ¿Cómo construir archivos? y nociones de CANDE.

5.- APLICACION DEL SISTEMA A UNA POBLACION :

- A) Generada por la computadora, para un diseño muestreo aleatorio simple.

B) Real, proporcionada por el usuario, para un diseño muestreo estratificado.

C) Real, proporcionada por el usuario, para un diseño muestreo de conglomerados.

6.- CONCLUSIONES

7.- BIBLIOGRAFIA

8.- ANEXOS : Programas del sistema.

1.- ANTECEDENTES

El tema desarrollado en esta tesis tiene su origen en un trabajo previo elaborado por Héctor Gutiérrez quien construyó - un sistema similar, como el que aquí se pretende, y fué implementado en una computadora Cyber 70 modelo 2. Dicho trabajo - constituye el inicio de un proyecto que se tiene en mente continuar, conforme se adicionen subsistemas de diseños de muestreo que hasta el momento no se encuentran incluidos.

Es importante mencionar que debido a las características del trabajo realizado y a las limitaciones de la máquina, el sistema desarrollado nunca se pudo implementar para su explotación.

Desde el punto de vista de la enseñanza por computadora el sistema presentado aquí, es una mezcla de los distintos tipos de programas que existen en este campo, no es un tutorial, ni un sistema experto, ni un tipo de instrucción programada, y mucho menos pretende ser un producto de la Inteligencia Artificial. Es un programa de simulación de muestreos que reúne las experiencias del asesor y del autor en el campo de la enseñanza del muestreo. Sin embargo es un programa que pretende participar del apoyo que brinda la computadora en el campo educativo, en especial dentro de los sistemas de simulación de fenómenos.

2.- INTRODUCCION

La experiencia recogida en los salones de clases ha demostrado que una buena parte del tiempo que se tiene para impartir las cátedras se utiliza en la realización de los ejercicios que ayudan a ejemplificar la teoría expuesta, es por esto que el diseñar programas que faciliten la aplicación de los conceptos enseñados en las aulas, de materias como el muestreo, es uno de los objetivos que se pretende al realizar trabajos como el presente, en el que las facilidades de uso y aplicación sean suficientemente atractivas para que el usuario los explote.

Este sistema está dirigido a aquellas personas con conocimientos elementales de estadística, en particular de la teoría del muestreo, que deseen efectuar simulaciones de los diseños básicos de ésta teoría, o que estén interesados en el cálculo de los estimadores para una muestra en particular, o que quieran obtener resultados de una muestra extraída de una población especial.

En esta tesis se desarrollan rutinas especiales para simular cada uno de los siguientes diseños de muestreo :

- Aleatorio Simple
- Estratificado
- De Conglomerados

mismas que se describen de manera general en el siguiente capítulo.

La descripción del sistema se lleva a cabo desde dos puntos de vista; uno el enfoque del muestreo, donde se describe el sistema, se detallan sus usos y facilidades, así como los diseños que lo constituyen, la expresión de los estimadores, modificaciones y sugerencias, y las consideraciones que se deben tener al interpretar los resultados; y dos, el enfoque de cómputo

que aquí se ubica, se presenta una clasificación de éste, se -- detallan sus capacidades de almacenamiento, se describe la forma adecuada de "cargar el sistema", y un punto muy importante, se explica ¿cómo construir archivos? y los comandos elementales del editor CANDE, que es la base de la comunicación del usuario vía una terminal con la computadora.

En el capítulo cinco se presenta una aplicación del sistema, a una población generada y a una población proporcionada por el usuario por medio de un archivo de datos.

En el capítulo seis se detallan las conclusiones a las - que se llegaron después de la construcción de un sistema como - éste.

Para finalizar se presentan los programas que constitu-- yen el sistema, con excepción del programa principal que simula muestreos, debido a que éste último tiene un tamaño considera-- ble.

3.- PRESENTACION DE OBJETIVOS

El objetivo principal de éste trabajo es desarrollar un Sistema de Cómputo que ayude al usuario en el aprendizaje del Muestreo a través de una terminal de teleproceso. Ya que ésta es el único medio de comunicación en línea (directo) con la computadora. Con lo anterior se comprobarán las facilidades y ventajas que se desprenden de la utilización de un sistema automatizado cuando se aplica en la enseñanza del muestreo.

A partir del objetivo principal se desprenden los que se señalan a continuación:

- Un objetivo secundario pretende introducir al alumno de un curso de muestreo al procesamiento de la encuestas mediante el uso de la computadora, así como facilitar al profesor la construcción de ejercicios prácticos, ya que tradicionalmente en los cursos de muestreo no se puede realizar prácticas debido, principalmente, a que las calculadoras de escritorio presentan una seria barrera en su capacidad de almacenamiento.

- El sistema visto como un instrumento de ayuda al aprendizaje del muestreo, cumple el objetivo de enfatizar, aclarar y marcar los pasos que se deben seguir al efectuar el procesamiento de una encuesta por muestreo.

Sentar las bases para el desarrollo de sistemas enfocados a proporcionar ayuda al estudiante en el cálculo de estimadores estadísticos (es otro objetivo que se desea cumplir con la construcción de un sistema como éste).

Construido a base de módulos, y con rutinas generales de fácil acceso, éste sistema permitirá su futuro crecimiento cumpliendo con esto uno de los objetivos más importantes en el desarrollo de sistemas (software) .

4.- DESCRIPCION DEL SISTEMA

En este capítulo se efectúa la descripción del sistema en forma sencilla, resaltando las características generales y específicas de cada diseño incluido.

Como el sistema está enfocado para facilitar el aprendizaje del muestreo, es conveniente describirlo desde dos puntos de vista para su mejor utilización.

A.- ENFOQUE DEL MUESTREO

a.1) Descripción del Diseño y Usos

El sistema esta formado de dos partes: una teórica y una práctica. La parte práctica contiene los tres diseños básicos - de la teoría del muestreo, a saber :

- Muestreo Aleatorio Simple
- Muestreo Estratificado
- Muestreo de Conglomerados

VENTAJAS DEL SISTEMA

El diseño 1/, desde el punto de vista del muestreo, tiene muchas facilidades y ventajas; las principales de ellas son las siguientes :

El sistema tiene incluida la opción de censar, que aun que no se puede catalogar como un diseño de muestreo en algunos casos es necesario trabajar con toda la población.

1/ En la siguiente sección se describe de manera sencilla cada diseño.

Esto suele suceder cuando las poblaciones son pequeñas, o cuando es indispensable para el estudio enumerar la población completa y tener una mayor exactitud en los resultados.

' El sistema permite generar poblaciones ficticias con distribución uniforme y normal dado un conjunto de parámetros. También permite proporcionar la población o la muestra sobre la que se desee efectuar las estimaciones.

' El sistema presenta la facilidad de que, para una población o muestra dada, es posible aplicar los tres diseños básicos del muestreo y la opción del censo para poder efectuar comparaciones.

' El sistema tiene integradas opciones y rutinas que facilitan la comprensión de los conceptos que se manejan.

' El sistema permite efectuar comparaciones para resultados obtenidos de un mismo diseño. Esto se debe básicamente a que en cada paso de la simulación se tiene al menos dos opciones para elegir, de tal forma que de un mismo diseño se puede obtener diferentes resultados según las opciones que se hayan elegido.

' Para cada diseño se contruyó una rutina específica para calcular el tamaño de muestra.

En la parte teórica del sistema se encuentran cinco capítulos a manera de libro, en los cuales debido a la carencia de caracteres especiales como letras griegas, exponente, letras minúsculas, y en general, símbolos matemáticos, fué necesario establecer una simbología con los caracteres de un teclado estándar en una terminal, para poder expresar las fórmulas, ecuaciones y expresiones para los parámetros y sus estimado--

res. En estos cinco capítulos se desarrollan los siguientes temas:

- 1 - Muestreo Aleatorio Simple
- 2 - Muestreo Aleatorio para Proporciones
- 3 - Muestreo Estratificado Aleatorio
- 4 - Muestreo Sistemático
- 5 - Muestreo de Conglomerados de una etapa

Cada tema tiene un conjunto de subtemas que se desglosan al momento de consultar alguno de éstos, de tal forma que el sistema permite consultar tantos subtemas como tenga el capítulo, las veces que se requiera, en una misma sesión de consulta.

USOS DEL SISTEMA

El sistema se puede utilizar como:

1 Un libro de texto. En él se encuentran descritos temas a manera de Introducción de la Teoría del Muestreo, estos fueron obtenidos de resúmenes y transcripciones de los principales libros de muestreo², los cuales se pueden utilizar como referencia o consulta al hacer uso del sistema.

Esta parte del sistema no es propiamente el objetivo de la tesis, ya que, como se mencionó en los antecedentes, se trata de la continuación de un proyecto ya iniciado y el trabajo que aquí se presenta es sólo una etapa más del proyecto original.

2 Un paquete de generación de ejercicios. Que puede utilizar el profesor para agilizar el ritmo de clases y faci-

² Consultar bibliografía anexa.

litar la explicación de los conceptos (tanto básicos como complejos), sin la necesidad de desarrollar tanta "talacha" en -- clases; para ésto, el sistema permite imprimir los resultados obtenidos en cada sesión de consulta o uso del sistema.

' Una herramienta en el cálculo de estimadores. Con la característica de poder proporcionar la población o la muestra según sea el caso, el sistema permite procesar encuestas pequeñas o poblaciones pequeñas. En la segunda parte de éste capítulo sección 3, se explica como proporcionar el archivo de datos con la población o la muestra a procesar.

' Como un instrumento para el aprendizaje del muestreo. Ya que sirve para que el usuario pueda reafirmar o entender -- los pasos o requerimientos que necesita en el procesamiento y cálculo de estimadores en una encuesta. Esto se puede lograr debido a que el sistema, conforme avanza requiere que se seleccione alguna opción de las desplegadas, para poder efectuar -- los cálculos y al mismo tiempo presenta definiciones o aclaraciones que pueden ser útiles en la comprensión de los conceptos, de ésta forma el usuario puede verificar que sus conocimientos del diseño son o no suficientes una vez que se haya efectuado el cálculo de los estimadores.

a.2) DISEÑOS DE MUESTREO

En esta sección se describen en forma general cada diseño de muestreo desarrollado, así como los puntos principales - que se deben considerar al hacer uso de ellos.

a.2.1) Muestreo Aleatorio Simple ("MAS")

El "MAS" es el más sencillo de los diseños de muestreo, - y consiste en la selección de manera aleatoria de n unidades - como muestra tomadas de una población de N elementos, con la - característica de que cada elemento de la población tiene - la misma probabilidad de ser elegido en muestra.

Dado que se efectúan simulaciones para cada diseño, lo - primero que se requiere es la población, ésta como ya se mencionó puede ser generada.

- Con distribución uniforme o normal, en cuyo caso la generará la máquina, requiriendo los parámetros necesarios, o tomando los usuales (dis - tribución normal o uniforme (0,1)).
- Con alguna otra distribución, en el caso de que el usuario la proporcione. En esta situación - se debe tomar en cuenta que en el sistema sólo es posible procesar poblaciones con 6000 elementos como máximo por variable, y para cada elemento únicamente se pueden tomar cinco características o variables a medir. En el punto b.5, se explica cómo crear el archivo que contenga - la población a procesarse.

Si el usuario ya tiene la muestra de la población que desea estudiar en un archivo de datos, entonces sólo utilizará el sistema como un paquete para calcular los estimadores de su di-

seño ("MAS"). El sistema tiene integrada una rutina especial para este caso, en la cuál se le pedirá al usuario la siguiente información:

- El tamaño de la población de la cuál se obtuvo esta muestra. .
- El número de la variable sobre la que se desea efectuar la estimación.
- El estimador que se desea calcular (a escoger -- entre: medias, totales y proporciones); una -- vez efectuados los cálculos, la computadora desplegará el resultado de las estimaciones e invitará al usuario a elegir otra variable de las -- que él proporcionó en la muestra, para hacer otra estimación, o bien dar por concluida la consulta al diseño "MAS".

Para poblaciones de menos de 6001 elementos el sistema -- permite censar la población. Es pertinente aclarar que, si el usuario proporcionó la población, sólomente se podrá efectuar un censo, sobre la variable que él haya elegido, y una vez desplegados los resultados del censo, se dará por concluida la consulta.

A continuación se deberá elegir que parámetros de desea -- estimar:

- Medias
- Totales
- Proporciones

Una vez concluido el paso anterior, es necesario calcular el tamaño de muestra mínimo para efectuar las estimaciones, te

niendo en cuenta las siguientes consideraciones :

- Precisión deseada
- Recursos disponibles (costos involucrados)
- Subdivisiones de la población para las que se requieren estimaciones.
- Características de la variable a medir (número y forma).

Fué necesario diseñar una rutina específica para calcular el tamaño de muestra en el "MAS" , esta rutina permite al usuario:

- 1) Dar el tamaño de muestra (arbitrario)
- 2) Obtenerlo por:
 - ' Precisión y confianza
 - ' Coeficiente de variación
 - ' Varianza fija
- 3) Calcularlo por costos

Si el usuario decide calcular el tamaño de muestra por - costos, la máquina pedirá la estimación del costo total y de - los costos de operación, y dividiendo el primero entre el se - gundo se calculará el tamaño de muestra mínimo.

Si el usuario decide calcular el tamaño de muestra con - la segunda opción, deberá elegir a continuación, si supondrá - o no normalidad del estimador.

Suponiendo normalidad del estimador, el usuario escogerá ahora:

- 1.- El nivel de confianza, entre : 99%, 95%, 90% - 85%, 70%, y 50%.

2.- El nivel de precisión entre: absoluto y relativo (respecto al parámetro).

Cuando el nivel de precisión sea relativo, el usuario - además de proporcionarlo, deberá dar una estimación del parámetro (medias o proporciones).

Como paso siguiente en el cálculo, la máquina solicitará una estimación de la dispersión de la variable (S^2 cuadrada). En caso de que no se tenga tal estimación, el sistema sugerirá efectuar una prueba piloto para tener una estimación de la dispersión y así poder calcular el tamaño de muestra.

Para calcular el tamaño de muestra sin suponer normalidad del estimador, se requiere fijar el valor del coeficiente de variación o el valor de la varianza del estimador.

La opción de fijar el coeficiente de variación, requiere además de una estimación de la esperanza del estimador.

Con los requerimientos y suposiciones anteriores es suficiente para que la máquina calcule el tamaño de muestra mínimo para las estimaciones, pero si se conoce el porcentaje de no - respuesta, es posible aplicar un factor de corrección al tamaño de muestra mínimo para obtener el tamaño de muestra definitivo.

Para efectuar la selección de la muestra se puede elegir entre:

- Selección con reemplazo.- Se extrae un elemento de la población, se efectúa la medición correspondiente a la variable de interés y se -- vuelve a incorporar a la población para darle oportunidad de que pueda ser seleccionado otra vez.

- Selección Sin Reemplazo.- Una vez extraído un elemento de la población este no se reincorpora.

Si se desean estimar proporciones, la máquina solicitará el rango de valores o límites de la característica en estudio. Esto es, se transformará la variable de interés en una variable dicotómica, que tomará el valor de cero o uno según sea el caso, por ejemplo:

Si la variable de interés es la edad, y se desea estimar la proporción de personas con edad entre 17 y 20 años, entonces el rango de valores o límites de la característica en estudio será (17,20), que dan la relación para la variable edad:

$$X = \begin{cases} 1 & \text{si la edad esta entre 17 y 20 años} \\ 0 & \text{si la edad es menor a 17 o mayor a 20 años.} \end{cases}$$

Una vez efectuados los cálculos se escogerá el nivel de precisión para los intervalos de confianza.

Como parte final, la rutina desplegará los datos de identificación de:

- La Población
 - La Muestra
 - Los estimadores
 - El Diseño del Muestro Aplicado
- y los resultados obtenidos del muestreo efectuado.

VENTAJAS DEL "MAS" DENTRO DEL SISTEMA:

- Dada una muestra con 5 variables por elemento-

se puede efectuar estimaciones de medias, totales y proporciones para cada variable.

- Permite efectuar una prueba piloto para calcular el tamaño de muestra.
- Permite calcular el tamaño de muestra para -- cuatro opciones diferentes^{3/}.
- Permite censar la población o la muestra según sea el caso.

LIMITACIONES DEL "MAS" DENTRO DEL SISTEMA:

- Sólo se puede procesar poblaciones con 6000 - elementos como máximo.
- Sólo se puede generar dos tipos de distribución de la población.
- No se puede estimar razones.*

^{3/} En la sección 3 de este capítulo se muestran las expresiones para la obtención del tamaño de muestra en ésta rutina.

a.2.2) MUESTREO ESTRATIFICADO ("MESTRA")

En el "MESTRA", la población se encuentra fraccionada o clasificada en subpoblaciones a las que se les da el nombre de estratos, de ésta forma cada elemento de la población pertenece a un sólo estrato.

Cada estrato o subpoblación es tratado en forma independiente, ésto es, se extrae una muestra y se obtiene un estimador para el estrato.

Los estimadores de los estratos se ponderan de manera adecuada para formar un estimación combinada de la población completa.

Este tipo de diseño es utilizado por varios motivos entre otros :

- Desde el punto de vista de selección, permite trabajar o estudiar a cada estrato por separado.
- Permite derivar estimaciones a nivel de estrato, y cada una de ellas ser estudiada con la precisión solicitada (en el caso de ser el estrato un dominio de estudio).
- Las estimaciones obtenidas resultan ser generalmente más precisas que aquellas derivadas del "MAS".
- Se pueden formar estratos para utilizar diferentes métodos y procedimientos dentro de ellos, - por ejemplo : por conveniencias y distribución física de la población, por deficiencias en los listados del marco, por naturaleza o característica diversa de los elementos, se clasifica la población y se efectúa un procedimiento adecuado para cada subpoblación.

La estratificación se efectúa en base a un conocimiento previo de la población, y se pretende que cada estrato quede -

conformado de elementos lo más homogéneo posibles.

La estratificación efectuada en la población aportará - beneficios substanciales en las estimaciones, dependiendo de - la correlación que exista entre la variable en estudio y la va - riante que se utilice para la estratificación.

En la rutina de "MESTRA", como en la de "MAS", primero hay que proporcionar la población, a continuación se efectuará la estratificación en un máximo de 10 estratos, y para esto se elegirá entre estratos de tamaño aleatorio y estratos de tamaño fijo, en caso de que la población haya sido generada.

Como paso siguiente, se construirán los estratos, dando para cada estrato el rango de valores, estos rangos deberán cu - brir completamente los posibles valores de la población, si lo anterior no se cumpliera entonces la rutina solicitará nueva- - mente datos hasta que la condición se cumpla.

Una vez terminada la estratificación es posible revisar el tamaño y los elementos de cada estrato.

Si la población fué generada, se deberá proporcionar el rango de valores para la variable de interés en cada estrato, y en caso contrario, dar el número asignado a la variable (és- - ta variable debe ser diferente a la utilizada para la estrati- - ficación).

En el caso de que el usuario haya proporcionado la mues - tra, el sistema lo conectará con una rutina diseñada especial- - mente para efectuar las estimaciones correspondientes a éste - diseño.

Como en el "MAS", el "MESTRA" también tiene la opción - de censar la población en estudio.

El paso siguiente es elegir el tipo de afijación que se utilizará para calcular el tamaño de muestra y después repar--tirla en los distintos estratos, se puede escoger entre :

- Afijación Optima.
- Afijación de Neyman.
- Afijación Proporcional.
- Afijación Igual.
- Afijación Arbitraria.

A continuación se deberá elegir que parámetros se desea estimar :

- Medias.
- Totales.
- Proporciones.

Una vez seleccionados los parámetros y el tipo de afijación se procede a calcular el tamaño de muestra mínimo para las estimaciones, como en el caso del "MAS", aquí fué necesario construir una rutina especial para obtener el tamaño de muestra^{4/}.

Para poder calcular el tamaño de muestra en éste diseño se necesita conocer mayor información acerca de la población, y es conveniente que el usuario tenga mucho cuidado en la selección que haga para el cálculo, ya que es muy factible que - la máquina le pida una estimación de :

- Las proporciones de cada estrato.
- Las Sh^2 de cada estrato.
- Los costos Ch de cada estrato.
- El parámetro de cada estrato..

4/ Rutina análoga a la del "MAS", en la que se consideran las suposiciones de Normalidad del Estimador, El cálculo por costos, y el tamaño de muestra sugerido por el usuario.

De igual forma que en el "MAS", en éste diseño es posible efectuar una prueba piloto para conocer o estimar la dispersión de la variable por estrato. En caso de que ésta opción se tome, debido a dificultades en el cálculo es necesario escoger uno de los siguientes tipos de afijación para poder efectuar la prueba piloto:

- Afijación Proporcional
- Afijación Igual
- Afijación Arbitraria

Considerando la complejidad de la rutina para el cálculo del tamaño de muestra, el usuario tiene la opción de revisar el tamaño obtenido en ésta y decidir si lo calcula nuevamente, lo cual dependerá del conocimiento que éste tenga de la variable en estudio.

Una vez determinado el tamaño de muestra total, es necesario repartirla en los diferentes estratos de la población, tomando en consideración el tipo de afijación elegida.

Dependiendo de la forma en que se calculó el tamaño de muestra se requerirá o no proporcionar información, esto es, si para calcular el tamaño fue necesario dar la estimación de:

- Los costos por estrato
 - Los porcentajes por estrato
 - La variabilidad dentro de cada estrato,-
- estos datos ya no será necesario volver a proporcionarlos.

Después de haber repartido la muestra en los diferentes estratos el usuario podrá verificar su distribución, y modificarla, si así lo desea.

Cabe mencionar que en la afijación arbitraria el usuario tiene la opción de fijar los tamaños de muestra para cada

estrato o el porcentaje que se desee respecto a la muestra total.

A continuación se debe seleccionar la muestra en cada estrato, con muestreo aleatorio simple y ésta selección se puede hacer con o sin reemplazo.

En caso de que el tamaño de muestra de un estrato sea mayor a su población éste será censado (utilizando la rutina general de censo).

Este diseño permite elegir uno o más niveles, para los intervalos de confianza de los estimadores de los estratos, esto es, se puede elegir un sólo nivel para todos los estratos, o uno para cada estrato.

Una vez efectuados los cálculos la rutina despliega, tanto los datos de indentificación del diseño, como los resultados obtenidos de la muestra seleccionada. Esta información se observa para cada estrato; y al final, se presenta una estimación de los totales de población, obtenidos de las estimaciones por estrato.

VENTAJAS DEL "MESTRA" DENTRO DEL SISTEMA :

- Dada una muestra de una población con éste diseño se pueden efectuar estimaciones para cuatro variables de : medias, totales, y proporciones, efectuando la estratificación dentro de la rutina.
- Permite efectuar una prueba piloto, para conocer una estimación de la dispersión de la variable dentro de cada estrato, y así poder calcular el tamaño de muestra.
- Si el tamaño de muestra obtenido para un estrato es mayor a su población, la rutina automáticamente censa el estrato.

- Permite elegir entre cinco tipos distintos de afijación.
- Revisa que la estratificación se haya efectuado adecuadamente; esto es, si hay elementos de la población que no se encuentran clasificados dentro de algún estrato, o los rangos se traslaparon, el sistema no permite avanzar en el proceso.
- Vigila que la variable utilizada para la estratificación no sea la misma que la variable de interés.
- Para cada afijación permite escoger entre cuatro maneras distintas de calcular el tamaño de muestra.^{5/}
- Permite censar la población o la muestra según sea el caso.
- Permite efectuar la selección de la muestra con o sin reemplazo para cada estrato.

LIMITACIONES DEL "MESTRA" DENTRO DEL SISTEMA :

- Sólo permite construir 10 estratos como máximo.
- Sólo se pueden procesar poblaciones con 6000 elementos como máximo por variable.
- Sólo se pueden generar poblaciones con dos distribuciones.
- No se pueden estimar razones.

5/ OP. CIT.

a.2.3) MUESTREO DE CONGLOMERADOS ("MCONG")

En muchas ocasiones es difícil contar con una lista completa y confiable de los elementos de la población, y se sabe que dicha población se puede agrupar (o se encuentra agrupada) en conjuntos para su fácil localización. De ésta forma la lista de elementos se ve reducida o transformada en una lista de conjuntos, que simplifica el manejo y la selección. Estos conjuntos contienen una cantidad fija o variable de elementos; - así el problema de elegir un número X de éstos, de una población, en forma aleatoria con un marco poco confiable, se ve - reducido a escoger varios conjuntos de la población y efectuar sobre éstos las estimaciones que se pretendan.

A estos conjuntos o divisiones de la población se les llama conglomerados, y pueden, cómo se dijo anteriormente, estar agrupados por naturaleza o agruparse en base a una característica o cualidad específica (variable de conglomeración).

Las razones principales por las que se efectúa un muestreo de conglomerados son las siguientes :

- Para reducir considerablemente los costos de levantamiento.
- Para facilitar la administración.
- Por dificultades geográficas de la localización de la población.

En este diseño, una vez efectuada la selección de los conglomerados, se procede a censar o submuestrear cada uno de éstos, de ésta forma vamos a poder diferenciar entre un diseño monoetápico o bietápico de conglomerados. Esto, aunado al tamaño fijo o variable de los elementos de los conjuntos (conglomerados), nos da cuatro tipos o clases del diseño de conglomerados que son :

- Conglomerados de tamaños iguales sin submuestreo
- Conglomerados de tamaños diferentes sin submuestreo.
- Conglomerados de tamaños iguales con submuestreo
- Conglomerados de tamaños diferentes con submuestreo.

La rutina que simula el "MCONG" contiene estos cuatro tipos de conglomerados.

Cuando se submuestra un conglomerado, se elige una cantidad de elementos de éste y se efectúa una estimación. Se submuestreará o no, dependiendo de las características de la población y de la precisión de las estimaciones.

EN el "MCONG" se pretende que los conglomerados queden contruidos de elementos lo más heterogéneo posibles.

A diferencia del "MESTRA", el "MCONG" produce menor precisión en las estimaciones que el "MAS", esto es el "precio" - que se paga por reducir los costos.

En esta rutina, como en las anteriores, hay que proporcionar o generar la población con la que se va a simular el díaño.

Como paso siguiente hay que conglomerar la población en un máximo de 3000 conglomerados; esto es, el mínimo número de elementos que conforman un conglomerado debe ser dos.

Si la máquina generó la población se podrá elegir entre conglomerados de tamaño fijo y conglomerados de tamaño aleatorio, en el caso de que el usuario haya proporcionado la población, deberá escoger el número de la variable con la que se va a conglomerar.

Para facilitar la construcción de los conglomerados se desarrollaron dos rutinas para ordenar la población con base en la variable de conglomeración y de ésta forma, permitir al -- usuario construir un número variable de conglomerados (menor o igual a 3000).

Para cada conglomerado el usuario deberá proporcionar - el rango de valores que tendrán los elementos, si éste rango - fuese tal que el conglomerado tuviera tamaño cero, el sistema desplegará un letrero informando al usuario y requiriendo nuevamente el rango de valores para ése conglomerado.

Si la suma de los tamaños de los conglomerados no es igual al total de la población, el sistema informará al usuario y pedirá otra vez los datos suficientes para su nueva construcción.

Una vez conocidos los tamaños, se procede a generar o - direccionar los valores de la variable de interés, según sea - el caso; ésto es, si el usuario proporcionó la población, deberá elegir el número de la variable de interés (diferente a la utilizada para la conglomeración), y si el sistema generó la - población entonces deberá dar el rango de valores para la variable en cada conglomerado.

En caso de que el usuario haya proporcionado la muestra de una población con este diseño, entonces el sistema lo conectará con una rutina diseñada específicamente para el cálculo - de los estimadores. Esta rutina es más compleja que las anteriores ("MAS" y "MESTRA"), ya que el usuario tiene la opción de elegir entre cuatro tipos de Muestreo de Conglomerados y además puede escoger entre observar las estimaciones por conglomerado, o sólo las estimaciones muestrales finales.

Si el usuario lo desea, es posible censar al población.

A continuación hay que elegir el tipo de conglomerados - que se van a trabajar, así como los parámetros que deseamos - estimar.

Debido a que las expresiones para los tamaños de muestra de los tipos :

- Conglomerados de tamaño diferente sin submuestreo
- Conglomerados de tamaño diferente con submuestreo
- Conglomerados de tamaño igual con submuestreo,

involucraban el conocimiento del número de conglomerados, y - esto es precisamente lo que se desea obtener, el usuario debe proporcionar el número de conglomerados.

Sólo para el diseño de conglomerados de tamaños iguales sin submuestreo fué posible adaptar una rutina similar a las construídas para el "MESTRA" y el "MAS", que calculara el número de conglomerados en muestra.

Una vez conocido el número de conglomerados se procede a seleccionar éstos con muestreo aleatorio simple.

Si el muestreo es de una sólo etapa entonces el paso siguiente es revisar la población completa (censar) de cada conglomerado en muestra; o si es de dos etapas, se deberá elegir la forma de obtener los tamaños de muestra dentro de cada conglomerado tomando en consideración las siguientes opciones :

- Dar los tamaños (m_i 's) para cada conglomerado.
- Fijar un factor m_i/M_i para cada conglomerado.
- Fijar un tamaño constante para todos los conglomerados.

Si el muestreo es en conglomerados de tamaños iguales con submuestreo, existe la posibilidad de que éste tamaño se obtenga de manera óptima, para lo cual el usuario deberá proporcionar la siguiente información :

- El costo total del muestreo : C
- El costo de las unidades de primera etapa : C1
- El costo de la entrevista por elemento : C2
- Una estimación de la variabilidad entre los conglomerados : S_b
- Una estimación de la variabilidad dentro de cada conglomerado : S_j

En caso de que el usuario desconozca ésta información, se le permitirá regresar al paso anterior a escoger otra forma de calcular el tamaño de muestra dentro de cada conglomerado.

Una vez obtenidos los tamaños de muestra, la rutina permite revisar la distribución de éstas muestras; y si se desea, corregirlas.

Al igual que en el "MESTRA", se puede efectuar la selección con o sin reemplazo para cada conglomerado.

Si el tamaño de muestra de un conglomerado es mayor que su población, éste se censa automáticamente.

Cuando se han terminado de efectuar los cálculos para cada conglomerado e posible elegir un nivel de precisión para los intervalos de confianza, éste puede ser diferente para cada conglomerado o igual para todos.

Los resultados son desplegados uno a uno para cada conglomerado en muestra, y al final se presentan los datos de identificación del muestreo, tales como :

- El tipo de conglomerados que se utilizó
- La distribución de la población
- El tamaño de la población
- El tamaño de la muestra y su forma de obtención
- Las estimaciones totales.

VENTAJAS DEL "MCONG" DENTRO DEL SISTEMA

- Permite cuatro distintos tipos de muestreo de conglomerados.
- Permite efectuar estimaciones para una muestra de los cuatro tipos de conglomerados.
- En la rutina de estimaciones para una muestra da permite revisar los resultados por conglomerado o solo las estimaciones totales.
- Permite para cada conglomerado en muestra elegir un nivel de precisión para los intervalos de confianza.
- Permite construir un máximo de 3000 conglomerados.
- Permite censar los conglomerados cuyo tamaño de muestra sea mayor a su población.
- Revisa que todos los elementos de la población estén clasificados en los conglomerados.
- Permite para cada conglomerado en muestra efectuar una selección con o sin reemplazo de sus elementos.

LIMITACIONES DEL "MCONG" DENTRO DEL SISTEMA

- No es posible efectuar pruebas piloto para calcular el tamaño de muestra.
- Si el número de conglomerados que se construye es muy grande, resulta ser muy tedioso proporcionar la información, y también produce un gasto considerable de tiempo de proceso, y de tiempo de entrada y salida.
- Sólo se pueden estimar medias, totales y proporciones.
- Sólo para uno de los cuatro tipos es posible estimar mediante una rutina el número de conglomerados en muestra.

a.3) Expresión de los Estimadores de los diseños

Para cada uno de los diseños el sistema permite calcular los estimadores para :

1) La media : \bar{Y}

2) El total : Y

3) La proporción : p .

4) La varianza del total o de la media : $\text{Var} (Y)$,
 $\text{Var} (\bar{Y})$.

Como el sistema permite efectuar la selección de la muestra con o sin reemplazo, las expresiones para uno y otro caso varían un poco. A continuación se dan las expresiones de cada estimador para los diferentes diseños :

a.3.1) "MAS" (Muestreo Aleatorio Simple)

1) Media : $\bar{Y} = \frac{\sum_{i=1}^n X_i}{n}$

2) Total : $Y = N \bar{Y}$

3) Proporción : p ; para el cálculo de este estimador existe una rutina que transforma en ceros y unos los elementos en muestra, según el rango de valores de la característica en estudio; esto es, si el elemento X_i en muestra está dentro del rango de valores de la característica en estudio, entonces a X_i se le asignará el valor 1 y cero en caso contrario, de ésta forma la expresión para el estimador es la siguiente :

$$p = \frac{\sum_{i=1}^n X_i}{n} \quad X_i = \begin{cases} 1 & \text{si } X_i \leq D \\ 0 & \text{si } X_i > D \end{cases}$$

$$4) \text{ Varianza } (\bar{y}) = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n} (1 - F1)$$

donde F1 es igual a n/N si el muestreo es sin reemplazo, y $1/N$ si es con reemplazo. $(1 - F1)$ es el factor de corrección por finitud que se le aplica al estimador cuando la población tiene menos de 30 elementos. o se considera a la población finita

Si se decide efectuar una prueba piloto, la rutina desplegará al final de los resultados, el valor de S^2 , para que sea utilizado en el cálculo del tamaño de muestra.

Para el cálculo del tamaño de muestra se tienen las siguientes expresiones para las cuatro opciones en la rutina correspondiente al "MAS", cuando se supone o no normalidad del estimador.

Sean :

z = el nivel de confianza.

d = el nivel de precisión absoluto

r = el nivel de precisión relativo al parámetro.

n = el tamaño de muestra.

N = el tamaño de la población.

S^2 = la dispersión de la variable.

\bar{y} = un estimador (de la media) del parámetro.

CV = el coeficiente de variación.

V = la varianza fija.

a) Suponiendo normalidad del estimador :

a.1) Para un nivel de precisión absoluto y un nivel de confianza dado :

$$n = \frac{n_0}{1 + \frac{n_0}{N}} \quad \text{donde } n_0 = \frac{z^2 S^2}{d^2}$$

a.2) Para un nivel de precisión relativa al parámetro y un nivel de confianza dado :

$$n = \frac{n_o}{1 + \frac{n_o}{N}} \quad \text{donde } n_o = \frac{Z^2 S^2}{r^2 \bar{y}^2}$$

b) No suponiendo normalidad del estimador :

b.1) Para un coeficiente de variación prefijado :

$$n = \frac{n_o}{CV + \frac{n_o}{N}} \quad \text{donde } n = \frac{S^2}{\bar{y}^2}$$

b.2) Para un varianza fija :

$$n = \frac{S^2}{V + \frac{S^2}{N}}$$

a.3.2) "MESTRA" (Muestreo Estratificado)

En el "MESTRA" se tiene dividida a la población en L estratos, cada uno de ellos tiene una población N_h y se obtiene de ellos una muestra n_h dándose las siguientes identidades :

$$N = \sum_{h=1}^L N_h \quad \text{y} \quad n = \sum_{h=1}^L n_h$$

$$W_h = \frac{N_h}{N} \quad \text{f}_h = \frac{n_h}{n}$$

así las expresiones para los estimadores son :

1) Media Poblacional : $\bar{y}_{st} = \sum_{h=1}^L W_h \bar{y}_h$
 donde \bar{y}_h es la media estimada para el estrato h , de donde :

$$\bar{y}_{st} = \sum_{h=1}^L \frac{N_h}{N} \frac{\sum_{i=1}^{n_h} Y_{hi}}{n_h}$$

2) Total Poblacional : $Y_{st} = N \bar{Y}_{st}$

$$\therefore Y_{st} = \sum_{h=1}^L N_h \frac{\sum_{i=1}^{nh} Y_{hi}}{nh}$$

3) Proporción p : al igual que en el "MAS", en el "MESTRA", cada elemento en muestra dentro de cada estrato es transformado a cero, o uno si está dentro del rango de valores de la característica en estudio, permitiendo con esto utilizar las mismas expresiones que se usaron al calcular la media.

4) Varianza (\bar{Y}_{st}) :

$$V(\bar{Y}_{st}) = \sum_{h=1}^L \frac{W_h^2}{nh} sh^2 (1 - fh)$$

donde :

$$sh^2 = \frac{1}{nh - 1} \sum_{i=1}^{nh} (Y_{hi} - \bar{Y}_h)^2 \quad \text{es la --}$$

varianza del h-ésimo estrato.

Las expresiones para la media, el total, la proporción y la varianza, se verán afectadas dependiendo del Tipo de Afijación que se haya elegido. En cada una de las expresiones anteriores, habrá que sustituir el valor de las nh 's, dependiendo del tipo de afijación, así el valor nh para una afijación :

- Óptima es : $nh = \frac{N_h sh / ch}{\sum_{h=1}^L (N_h sh / ch)} \cdot n$

- De Neyman es : $nh = \frac{N_h sh}{\sum_{h=1}^L N_h sh} \cdot n$

- Proporcional es : $nh = \frac{N_h}{N} \cdot n$

- Igual es : $nh = \frac{n}{L}$

- Arbitraria es $nh = k$ (cte. que depende del estrato h).

Como la rutina que simula el "MESTRA" permite efectuar selecciones con o sin reemplazo, el factor fh cambia para cada caso, siendo fh = nh / Nh si la selección fué sin reemplazo y fh = 1 / Nh en caso contrario.

Si al calcular el tamaño de muestra se decidió efectuar una prueba piloto, entonces para cada estrato la rutina desplegará el valor Sh después de exhibir los resultados de la prueba por estrato.

Para el cálculo del tamaño de muestra se tienen las siguientes expresiones, tomando en cuenta la afijación elegida y la suposición o no de normalidad del estimador, éstas fueron obtenidas de sustituir el valor de nh en la siguiente fórmula general :

$$\frac{d^2}{z^2} = V(\bar{y}) = \sum_{h=1}^L W_h^2 V(\bar{y}_h)$$

Así para una Afijación :

- Optima :

a) Los requerimientos son :

Nh para h = 1, 2, ..., L

sh para h = 1, 2, ..., L

ch para h = 1, 2, ..., L

d², z², N², V² (varianza fija)

CV coeficiente de variación

ȳst una estimación del parámetro.

Sean :

$$T1 = \left(\sum_{h=1}^L \frac{N_h}{N^2} sh \quad ch \right) \left(\sum_{h=1}^L N_h \frac{sh}{ch} \right)$$

$$T2 = \sum_{h=1}^L \frac{N_h}{N^2} Sh^2$$

b) Las expresiones de n suponiendo normalidad del estimador son :

$$b.1) n = \frac{T1}{\frac{d^2}{z^2} + T2}$$

$$b.2) n = \frac{T1}{\frac{r^2 \bar{y}_{st}^2}{z^2} + T2}$$

c) Las expresiones de n no suponiendo normalidad del estimador son :

$$c.1) n = \frac{T1}{CV^2 \bar{y}_{st}^2 + T2}$$

$$c.2) n = \frac{T1}{V^2 + T2}$$

- De Neyman :

a) Los requerimientos son :

Nh para $h = 1, 2, \dots, L$

sh para $h = 1, 2, \dots, L$

$d^2, z^2, N^2, CV^2, V^2, \bar{y}_{st}^2$

Sean :

$$T1 = \frac{1}{N^2} \left(\sum_{h=1}^L Nh sh^2 \right)^2$$

$$T2 = \sum_{h=1}^L \frac{Nh sh^2}{N^2}$$

b) Las expresiones de n suponiendo normalidad - del estimador son :

$$b.1) n = \frac{T1}{\frac{d^2}{z^2} + T2}$$

$$b.2) n = \frac{T1}{r^2 \varphi_{st}^2 + T2}$$

c) Las expresiones de n no suponiendo normalidad del estimador son :

$$c.1) n = \frac{T1}{CV^2 \varphi_{st}^2 + T2}$$

$$c.2) n = \frac{T1}{v^2 + T2}$$

- Proporcional :

a) Los requerimientos son :

N_h para $h = 1, 2, \dots, L$

s_h para $h = 1, 2, \dots, L$

$d^2, z^2, N^2, CV^2, v^2, \varphi_{st}^2$

$$\text{Sean : } T1 = \sum_{h=1}^L \frac{N_h s_h^2}{N}$$

$$T2 = \sum_{h=1}^L \frac{N_h s_h^2}{N^2}$$

b) Las expresiones para n suponiendo normalidad del estimador son :

$$b.1) n = \frac{T1}{\frac{d^2}{z^2} + T2}$$

$$b.2) n = \frac{T1}{\frac{r^2 \gamma_{st}^2}{z^2} + T2}$$

c) Las expresiones para n no suponiendo normalidad del estimador son :

$$c.1) n = \frac{T1}{CV^2 \gamma_{st}^2 + T2}$$

$$c.2) n = \frac{T1}{V^2 + T2}$$

- Igual :

a) Los requerimientos son :

Nh para h = 1,2,...,L

sh para h = 1,2,...,L

d², z², N², CV², V², γ_{st}²

Sean :

$$T1 = L \sum_{h=1}^L \frac{Nh^2 sh^2}{N^2}$$

$$T2 = \sum_{h=1}^L \frac{Nh sh^2}{N^2}$$

b) Las expresiones de n suponiendo normalidad - del estimador son :

$$b.1) n = \frac{T1}{\frac{d^2}{z^2} + T2}$$

$$b.2) n = \frac{T1}{\frac{r^2 \bar{v}_{st}^2}{z^2} + T2}$$

c) Las expresiones de n no suponiendo normalidad del estimador son :

$$c.1) n = \frac{T1}{CV^2 \bar{v}_{st}^2 + T2}$$

$$c.2) n = \frac{T1}{v^2 + T2}$$

- Arbitraria :

a) Los requerimientos son :

Nh para h = 1,2,...,L

sh para h = 1,2,...,L

d², z², N², CV², v², \bar{v}_{st}^2 , kh

donde kh es un porcentaje que depende del estrato.

Sean :

$$T1 = \sum_{h=1}^L \frac{Nh^2 sh^2}{N^2 kh}$$

$$T2 = \sum_{h=1}^L \frac{Nh sh^2}{N^2}$$

b) Las expresiones para n suponiendo normalidad del estimador son :

$$b.1) n = \frac{T1}{\frac{d^2}{z^2} + T2}$$

$$b.2) n = \frac{T1}{\frac{r^2 \bar{Y}_{St}^2}{z^2} + T2}$$

c) Las expresiones para n no suponiendo normalidad del estimador son :

$$c.1) n = \frac{T1}{CV^2 \bar{Y}_{St}^2 + T2}$$

$$c.2) n = \frac{T1}{V^2 + T2}$$

a.3.3) "MCONG" (Muestreo de Conglomerados)

En el muestreo de conglomerados se tiene dividida la población en N conglomerados con tamaños M_i , con $i = 1, 2, \dots, N$, de estos se elegirán n conglomerados para efectuar las estimaciones.

Si el "MCONG" es de una etapa, entonces las expresiones para la media, el total, la varianza y la proporción son las siguientes :

1) Media por unidad o de conglomerados :

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \text{ donde } y_i = \sum_{j=1}^{M_i} y_j \text{ es el total de la característica del } i\text{-ésimo conglomerado.}$$

Nota: El anterior estimador no se despliega en los resultados, pero es muy fácil de obtener con los datos de la población.

2) Media por elemento :

$$\bar{Y} = \frac{1}{n} \frac{1}{\sum_{i=1}^n M_i} \left(\sum_{i=1}^n y_i \right)$$

3) Total :

$$Y = N \bar{Y} = \frac{N}{n} \sum_{i=1}^n y_i$$

4) Varianza de la media, $\text{Var}(\bar{Y})$

$$V(\bar{Y}) = \frac{(1-f)}{n M^2} \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}$$

$$\text{donde } M^2 = \frac{\left(\sum_{i=1}^n M_i \right)^2}{N^2}$$

5) Proporción p , se calcula igual que en el "MESTRA" con una transformación de los elementos de cada conglomerado, y utilizando las expresiones para la media por elemento.

Si el "MCONG" es de dos etapas, las expresiones para la media, el total, la varianza y la proporción son las siguientes:

1) Media por unidad o conglomerado :

$$\bar{y} = \sum_{i=1}^n \sum_{j=1}^{m_i} \frac{y_{ij}}{n} = \frac{Y}{n}$$

Nota : Este estimador no se presenta en los resultados, pero es muy faácil de obtener utilizando la información que presentan estos.

2) Media por elemento :

$$\bar{Y} = \frac{Y}{\sum_{i=1}^n M_i}$$

3) Total :

$$Y = \frac{N}{n} \sum_{i=1}^n Y_i = \frac{N}{n} \sum_{i=1}^n \frac{M_i}{m_i} \sum_{j=1}^{m_i} y_{ij}$$

4) Varianza de la media, $\text{Var}(\bar{Y})$

$$V(\bar{Y}) = \left(\frac{1}{n} - \frac{1}{N} \right) s^2_{b'} + \frac{1}{nN} \sum_{i=1}^n \frac{M_i}{\bar{M}} \left(\frac{1}{m_i} - \frac{1}{\bar{M}} \right) s_i^2$$

donde :

$$s^2_{b'} = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{M_i}{\bar{M}} \bar{y}_i - \frac{1}{n} \sum_{i=1}^n \frac{M_i}{\bar{M}} \bar{y}_i \right)^2$$

$$y \quad s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2$$

5) Proporción p ; para cada elemento en muestra se efectúa una transformación, en ceros y unos, uno si está dentro del rango de los valores de la característica en estudio y cero en otro caso, una vez realizado lo anterior, para el cálculo de la proporción se utilizan las mismas expresiones que para la media por elemento.

Si el "MCONG" es de una sola etapa con tamaños iguales, el tamaño de muestra, en éste caso el número de conglomerados en muestra, se puede obtener utilizando la misma rutina que se usó para el "MAS". Y para los otros tres tipos, hay que proporcionar el número de conglomerados en muestra.

Si el "MCONG" es de tamaños iguales con submuestreo, la

expresión para el tamaño de muestra óptimo en cada conglomerado es :

$$m = \frac{C_1 S^2_W}{C_2 A}$$

y para el número de conglomerados en muestra es :

$$n = \frac{C}{C_1 + C_2 C_1 \cdot \frac{S^2_W}{A}}$$

donde :

C es el costo total del muestreo.

C₁ es el costo de la unidad primaria de selección.

C₂ es el costo de la entrevista por elemento.

S²_W es la variabilidad en los conglomerados.

S²_b es la variabilidad entre los conglomerados

$$S^2_W = \frac{1}{N} \sum_{i=1}^n S_i^2$$

y

$$A = S^2_b - \frac{S^2_W}{\bar{M}}$$

a.4) Diseños que se pueden incluir, modificaciones y sugerencias.

El sistema, como se puede observar, tiene rutinas específicas para cada uno de los diseños de muestreo presentados, y cómo éstas son independientes, trabajadas a manera de módulos, es muy factible integrar otras similares, tales como el muestreo sistemático y el muestreo polietápico, dependiendo de las necesidades del usuario.

De igual forma, es posible incluir nuevas rutinas para la generación de poblaciones con otras distribuciones, por ejemplo : La Poisson, la binomial, la geométrica, la exponencial, la gamma, etc.

Si se desean incluir diseños, rutinas u opciones nuevas para el sistema, se sugiere revisar cuidadosamente la segunda parte de éste capítulo, donde se clasifican los programas y subprogramas, así como los listados de cada rutina.

Para efectuar una modificación a la parte teórica del sistema, se invita al usuario a revisar cuidadosamente el archivo TEORIA/SIN , en el se encuentran los resúmenes de ésta parte; en él se pueden observar una serie de números a la izquierda de cada renglón o línea de información, que corresponden al número de tema, subtema y línea de detalle o desglose, éstos dígitos sirven para que el programa LEE/ESC/APUNT construya los apuntadores, por lo que es importante que el usuario tenga cuidado al efectuar una modificación. Una vez concluida la modificación se sugiere ejecutar los programas :

- ORDENA/MAESTRO/TEORIA
- LEE/ESC/APUNT

en éste último se deberá verificar, que las líneas borradas o incluídas estén tomadas en cuenta en el total de registros leídos

y grabados, esto será muy fácil de verificar, ya que el programa despliega tanto en la pantalla, como en la impresión de la ejecución, la información acerca de los registros leídos y grabados.

a.5) Consideraciones que se deben tener en cuenta para la interpretación de resultados.

Para la interpretación de resultados es conveniente tener en cuenta las siguientes consideraciones :

- El sistema simula muestreos, éstos se efectúan en forma aleatoria con números cuya distribución es uniforme, utilizando la función "Random" de la máquina, y éstas simulaciones como tales tienen un margen de error. Se pretende con éstas simulaciones ayudar al aprendizaje de los diseños simples de la Teoría del Muestreo.
- Cada diseño tiene sus peculiares limitaciones y requerimientos.
- Se escogieron los estimadores insesgados para el cálculo de la media, el total y la varianza.
- En el cálculo del tamaño de muestra se toman en cuenta las situaciones prácticas y teóricas.
- El sistema permite muchas facilidades.

De esta forma la comparación de los resultados que proporcionen los distintos diseños, deberá efectuarse con cuidado y tomando en cuenta las consideraciones anteriores.

Cuando se utilice el sistema para el cálculo de estimadores de una muestra dada, la veracidad o exactitud de éstos dependerá en gran parte de la información complementaria que el usuario proporcione a la máquina, ya que para estos cálculos los estimadores también son insesgados y elegidos cuidadosa---

samente al momento de diseñar las rutinas.

A continuación se presenta un formato de salida, que puede ayudar a explicar los resultados obtenidos al efectuar un muestreo, se trata de la salida de una simulación de "MAS"

MUESTREO_ALEATORIO_SIMPLE

SIN REEMPLAZO

DATOS DE IDENTIFICACION

TIPO DE POBLACION: GENERADA TIPO DE DISTRIBUCION: UNIFORME

CON A = 0.00 Y B = 1.00

TAMAÑO DE LA MUESTRA : 66 LA MUESTRA LA DIO EL USUARIO.

RESULTADO DEL MUESTREO

M U E S T R A L E S .

M E D I A	T O T A L
0,42922	77,2597
VARIANZA DEL ESTIMADOR	
0,0014	43,8716
ERROR ESTANDARD	
0,03680	6,6236
INTERVALOS DE CONFIANZA	
0,50134	90,2418
0,35710	64,2775

TECLEA RETURN :

MUESTREO ALEATORIO SIMPLE

SIN REEMPLAZO

DATOS DE IDENTIFICACION

TIPO DE POBLACION: GENERADA TIPO DE DISTRIBUCION UNIFORME
CON A = 0.00 Y J = 1.00000
TAMANO DE LA MUESTRA : 66 LA MUESTRA LA DIO EL USUARIO.

RESULTADO DEL MUESTREO

M U E S T R A L E S .

M E D I A	T O T A L
0.42922	77.2597
VARIANZA DEL ESTIMADOR	
0.0014	43.3716
ERROR ESTANDAR	
0.03630	6.6236
INTERVALOS DE CONFIANZA	
C.50134	70.2418
C.35710	64.2775
	TECLTA RETURN :

Como se puede observar las primeras líneas se refieren al tipo de :

- Muestreo
- Selección
- Población
- Distribución
- Tamaño de la muestra
- Y forma de obtención del tamaño de muestra.

En las líneas "muestrales" las columnas son los estimadores para la media y el total, que resultaron de la muestra extraída; desglosados por renglón se refieren a :

- 1er renglón : estimadores de la media y el total.
- 2do renglón : varianza estimada de la media y el total.
- 3er renglón : error estándar de la media y el total.
- 4to renglón : límite superior del intervalo de confianza para la media y el total.
- 5to renglón : límite inferior del intervalo de confianza para la media y el total.

Es conveniente mencionar, que este tipo de formato de salida es muy similar al que se muestra en los diseños "MESTRA" y "MCONG".

B.- ENFOQUE DE COMPUTO.

b.1) Descripción general, usos y diagrama del sistema.

Desde este punto de vista, el sistema es un conjunto de programas cuya finalidad es facilitar o ayudar al aprendizaje del muestreo, efectuando simulaciones de los distintos diseños incluidos en él.

El sistema se encuentra cargado en la computadora B7800 Burroughs (propiedad de la UNAM); sin embargo es factible utilizarlo en cualquier máquina de las series B6000 y B7000.

Está escrito casi en su totalidad en FORTRAN IV, un lenguaje muy simple y compatible con muchas máquinas.

El sistema lo constituyen los siguientes programas principales :

- 1.- LEE/ESC/APUNT
- 2.- ORDENA/MAESTRO/TEORIA
- 3.- SISTEMA/FIN
- 4.- IMPRIME
- 5.- IMPRIME/B
- 6.- BAJA/SISTEMA
- 7.- SUBE/SISTEMA

1.- Este programa fue escrito para construir los apuntes a los temas y subtemas de la parte teórica del sistema, y utiliza los siguientes archivos :

- DATOSTEXT (de entrada)
- MAESTRO/TEORIA (de salida)
- APUNTA/TEMAS (de salida)

es un programa completamente interactivo que construye una matriz de apuntadores que es desplegada en la pantalla; conforme se lee el archivo DATOSTEXT, se va llenando ésta matriz y se va creando el archivo MAESTRO/TEORIA, que servirá como un banco de información para el sistema; está escrito en Fortran IV y tiene aproximadamente 80 líneas.

2.- Este programa fue escrito para ordenar el archivo que contiene la parte teórica del sistema (Introducción a la teoría del Muestreo), utiliza los siguientes archivos :

- TEORIA/SIN (de entrada)
- DATOSTEXT (de salida)

está escrito en Algol, por la facilidad que se tiene en este lenguaje para ordenar y tiene 8 líneas. Este programa se sugiere utilizarlo cuando se cargue el sistema, o cuando se efectúen modificaciones a la parte teórica.

3.- Es el programa principal de la tesis, escrito en Fortran IV con aproximadamente 4400 líneas, en él se efectúan tanto las consultas a la parte teórica, como a la parte práctica del sistema y utiliza los siguientes archivos :

- APUNTA/TEMAS (de entrada)
- MAESTRO/TEORIA (de entrada)
- DATOS/B (de entrada)
- RESULTADOS (de salida)

DATOS/B es un archivo definido ficticiamente para poder ligar al sistema un archivo de datos en el que el usuario tenga almacenada la población o la muestra que desee procesar; cuando el sistema pida el archivo de datos, el usuario deberá proporcionar el nombre completo del archivo con un punto al final, por ejemplo :

DATOS/PRUEBA/1.

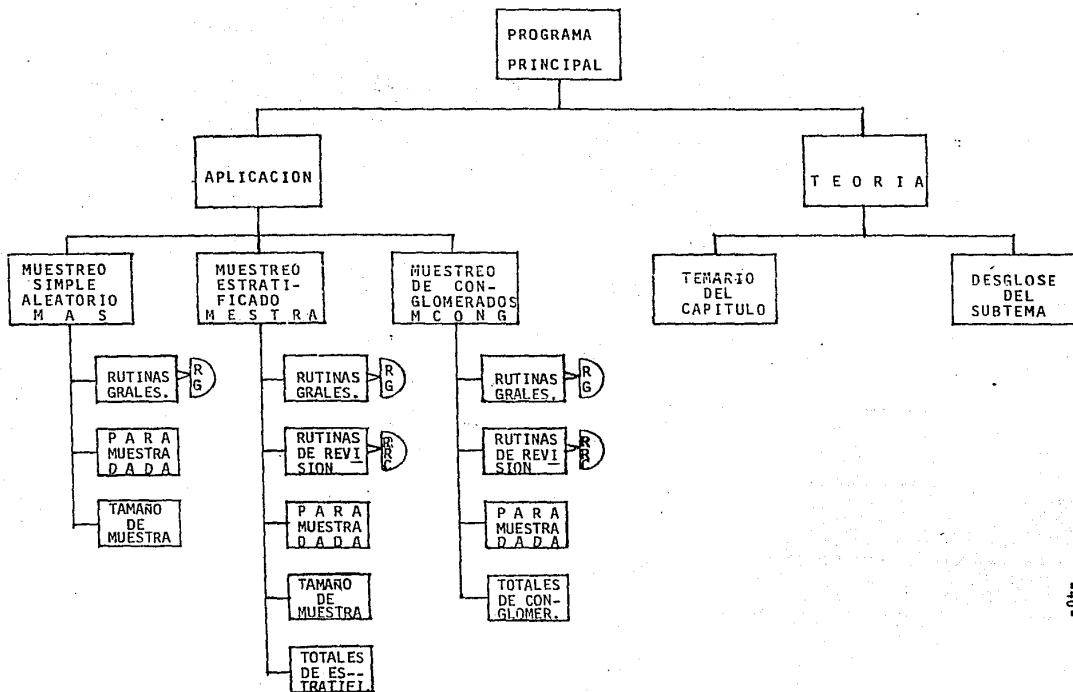
la rutina que permite al usuario proporcionar el archivo de datos, tiene una instrucción para cambiar el nombre del archivo

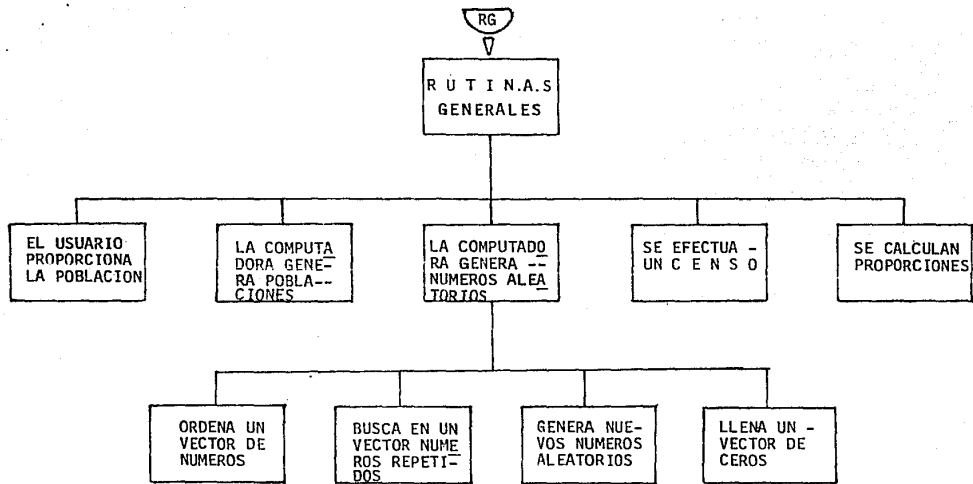
DATOS/B por el nombre que el usuario dió.

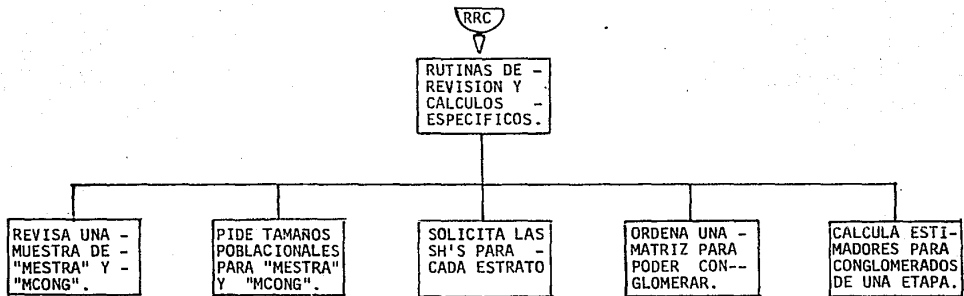
Los programas 4 y 5 los activa el programa SISTEMA/FIN y sirven para imprimir el archivo RESULTADOS, que se produce en una consulta al sistema, este archivo contiene toda la información que se desplegó en una consulta, así como los resultados de las estimaciones efectuadas. La diferencia entre éstos programas es que, el cinco imprime el archivo RESULTADOS y después lo remueve o borra de la memoria.

Los programas 6 y 7 son de utilería, y sirven básicamente para subir y bajar todo el sistema de Cinta a Disco respectivamente, están escritos en WFL (Work Flow Language) y contienen aproximadamente 25 líneas.

A continuación se presenta un Diagrama del programa SISTEMA/FIN visto como un sistema, que sirve para visualizar la estructura de las rutinas que lo constituyen, así como su uso y nivel de desglose.







b.2) Clasificación del Sistema

Como se pudo observar en el diagrama anterior el sistema lo constituyen un programa principal y un total de 28 subrutinas, que son llamadas desde el programa principal o desde las subrutinas que dirigen cada diseño, a continuación se presenta una clasificación de estas subrutinas, haciendo mención a su uso, nombre y parámetros que requiere para su llamado (estos últimos se encuentran dentro de los paréntesis).

Subrutinas Principales : TEORIA Y PRACTICA :

A.- TEORIA : Sirve para las consultas a la parte teórica del sistema, no tiene parámetros y utiliza las subrutinas:

A.1) TEMCAP(RITEMA,RFTEMA,ISUBT) : se utiliza para desplegar los subtemas de un capítulo.

A.2) LISUBT(RISUBT,RFSUBT) : se utiliza para desplegar la información del subtema elegido.

B.- APLICA : No tiene parámetros, se utiliza para elegir el tipo de diseño que se desea simular, y llama a las siguientes subrutinas :

B.1) MAS(A,LPOB,NV,IO) : simula un muestreo aleatorio simple.

B.2) STRA(A,LPOB,NV,IO) : simula un muestreo estratificado.

B.3) CONGLO(A,LPOB,NV,IO) : simula un muestreo de conglomerados.

Las tres rutinas anteriores a su vez utilizan las siguientes subrutinas de uso general, uso específico, y de revisión o cálculo :

C.- De uso general :

- C.1) USUAR(A,NTP,NV,FORTO) : permite al usuario - proporcionar, mediante un archivo, la población o muestra que se desee procesar.
- C.2) GENPOB(DATO,ITOP,XMED,XVAR,RVAR,TIPO) : - genera poblaciones con distribución normal - o uniforme.
- C.3) GENUA(ALE,N,L,RANG,IRC) : genera números - aleatorios con distribución uniforme utilizando la función "Random" de la máquina; como uno de sus parámetros (IRC) se refiere a la selección con o sin reemplazo requiere de - llamar a las siguientes subrutinas :
 - C.3.1) ORDENA(ALE,N) : ordena N numeros.
 - C.3.2) BUSCA(IALE,N,NR,NUEVN) : busca NR numeros repetidos en IALE y los deja en NUEVN.
 - C.3.3) GENUEV(NUEVN,NR,A,FACT,IALE,N) : genera NR nuevos números aleatorios con factor FACT y los deja en IALE para formar N números - aleatorios distintos.
 - C.3.4) LIMPIA(L,N) : llena de ceros el vector L.
- C.4) CENSO(A,ITOP,RMED,RVAR) : efectúa un censo sobre los elementos del vector A.
- C.5) PROPOR(A,ITOP,B,NELEM,P,Q) : transforma en ceros y unos los elementos del vector A y - los deja en el vector B, además calcula las proporciones P y Q.

D.- De Uso Especifico :

- D.1) MUESTR(A,IPOPO,NV) : efectúa el cálculo de los estimadores para una muestra proveniente de un diseño "MAS".
- D.2) TAMMUE(A,IPOPO,NMU,KEST,IPILOT,LMUE) : calcula el tamaño de muestra para un diseño "MAS".
- D.3) MSTRAT(A, IEST, NEST1, NTAMUE, NV) : efectúa el cálculo de los estimadores para una muestra proveniente de un diseño "MSTRA".
- D.4) TMSTRA(A, NTP, NEST1, AFIJA, NMU, KEST, IPILOT, CH, SH, LMUE) : calcula el tamaño de muestra para un diseño "MSTRA".
- D.5) TOTSTR(KEST, NEST1, IEST, NTAE, MEDEST, VAREST) : efectúa el cálculo de las estimaciones para los totales poblacionales, así como la impresión de éstos para un diseño "MSTRA".
- D.6) MUCONG(A, IEST, NEST1, INICON, NTMUE, NV) : efectúa el cálculo de los estimadores para una muestra proveniente de un diseño "MCONG".
- D.7) TOTCONG(KEST, NEST1, IEST, NTAE, MEDEST, VAREST, NTC, NTM) : efectúa el cálculo de los estimadores para los totales poblacionales, así como la impresión de éstos para un diseño "MCONG".

E.- De Revisión o Cálculo :

- E.1) REVMUE(NTAE, NEST1, MTIPO) : permite revisar y

modificar si se desea los tamaños de muestra dentro de los estratos, o dentro de los conglomerados en diseños bietápicos.

- E.2) NHFIJO(NH,NEST1,NTP,MTIPO) : permite que el usuario fije los tamaños poblacionales de los estratos o de los conglomerados.
- E.3) CALCSH(SH,NEST1) : solicita al usuario los valores Sh para cada estrato.
- E.4) ORDTOD(A,NTP,NV,NVORD) : ordena la matriz A en forma ascendente con base a la columna NVORD, esta matriz se utiliza para efectuar la conglomeración.
- E.5) CESTCO(A,ITOP0,RMED,RVAR,KEST) : efectúa el cálculo de estimadores para conglomerados de una sola etapa.

b.3) Capacidades de Almacenamiento y Recomendaciones de Uso

Como ya se mencionó en la sección a.1) de éste capítulo, el sistema sólo es capaz de procesar poblaciones o muestras - con un máximo de 6000 registros o casos por variable leída, - solo se pueden almacenar un máximo de cinco características o variables.

Estas limitaciones se deben a que el sistema utiliza matrices para almacenar y procesar los datos; y la computadora - tiene como capacidad máxima 30,000 posiciones de memoria para una matriz.

Hubo necesidad de hacer uso de las matrices porque, el - efectuar lecturas y escrituras a los archivos en disco, tiene como consecuencia lentitud en la ejecución, y esto aunado a - que el sistema es completamente interactivo con el usuario (- vía una terminal o sistema de tiempo compartido) produce un - gasto considerable de tiempos de : entrada, salida y de proceso.

La decisión de no usar archivos en disco para almacenar la información está basada en la recomendación de dos expertos, que sugirieron utilizar la memoria temporal de las matrices, - en lugar del grabado permanente, procesos de entrada y salida, y de los direccionamientos de archivos en disco.

Las anteriores limitaciones, no quieren decir que el sis tema sólo puede leer registros de cinco variables o campos, ya que como se pide al usuario :

- el nombre del archivo de datos,
- el número de variables a leer,
- el formato de lectura de cada registro,

este tiene plena libertad de construir sus archivos y procesar las variables que desee, esto es, con el formato indicará cuáles y cuántas variables de cada registro se leerán, por ejemplo:

Se tiene un archivo cuyo nombre es DATOS/ENCUESTA que contiene 10 variables en cada registro, 5 de ellas son enteras y 5 son reales y están colocadas en forma alternada dentro del registro, esto es :

R1 E1 R2 E2 R3 E3 R4 E4 R5 E5

el archivo tiene un total de 2000 registros, el usuario requiere efectuar estimaciones para las primeras cuatro variables.

La interacción con la máquina para proporcionar los datos será la siguiente (donde M es la línea desplegada por la computadora y U es la respuesta correcta que el usuario debe proporcionar) :

M ° DAME EL TAMAÑO DE LA POBLACION :
U ° 2000
M ° DAME EL NUMERO DE VARIABLES QUE VOY A LEER POR REGISTRO :
U ° 4
M ° DAME EL FORMATO DE LECTURA^{6/}, POR EJEMPLO :
(3X,3I5,2X,I4)
EL FORMATO ES :
U ° (F6.1,2X,I3,F6.2,5X,I4)
M ° AHORA DAME EL NOMBRE DEL ARCHIVO DONDE SE EN--
CUENTRAN TUS DATOS, CON UN PUNTO AL FINAL, POR
EJEMPLO : DATOS/C.
U ° DATOS/ENCUESTA.
M ° DESEAS QUE TU ARCHIVO SE LEA DESDE EL REGISTRO
UNO :

^{6/} NOTA : Este formato deberá estar escrito en lenguaje Fortran, iniciando y terminando con un paréntesis.

1.- SI

2.- NO

SELECCIONA TU OPCION :

U ° 1

M ° ESPERA SE ESTAN LEYENDO TUS DATOS

M ° DE ACUERDO, TU POBLACION ESTA LISTA PARA
PROCESARSE

de esta forma se han almacenado 4 variables en la matriz de datos del sistema, de las cuales 2 son enteras y 2 son reales, - con 2000 casos cada una.

Si en lugar de usar las 4 primeras variables, se deseara trabajar con las 4 últimas, lo único que hay que cambiar es el formato de lectura, de tal forma que al acceder cada registro - del archivo se lean las posiciones correspondientes a las 4 - variables finales, por ejemplo :

(30X,F6.1,I2,F10.2,I5)

El usuario deberá tener presente cuáles son y el tipo de las variables que están almacenadas, al efectuar sus estimaciones, ya que si alguna de estas se utiliza para clasificar la población, es importante saber si es real o entera, para que los rangos utilizados sean correctos.

Como el sistema permite leer desde el registro número - uno o desde cualquier otro registro, esto hace posible procesar una población de más de 6000 casos por variable, efectuando una partición en la lectura del archivo, esto es, si se tiene un - archivo con 10,000 registros, se sugiere procesar primero 5000 registros, y después leer desde el registro 5001 y almacenar - los siguientes 5000 registros que contiene el archivo: de esta manera es posible procesar (en forma particionada) archivos de cualquier tamaño.

Esta facilidad de procesamiento debe tomarse con mucha - precaución, ya que las divisiones del archivo deben hacerse en forma igual, o siguiendo una metodología previamente definida - procurando que no se afecten las estimaciones.

Este uso del sistema se recomienda solo para aquellas - personas que tienen los suficientes conocimientos de la Teoría del Muestreo y pretendan efectuar estimaciones parciales de una población, ya que los estimadores poblacionales se verán afecta dos por el tipo y forma de segmentación del archivo que contiene los datos.

Cabe mencionar que, debido a la posibilidad de indicarle a la computadora de cual archivo se desea leer, una población - puede estar almacenada en diferentes archivos.

- b.4) Procedimientos para el "cargado del sistema", actualizaciones y modificaciones.

Para poder efectuar el "cargado" y protección del sistema en forma rápida y eficiente se antepuso un nemónico o identificador a los archivos que integran éste, el cual fue "TEMUE", que significa Tésis Muestreo, así a los nombres de los archivos que se mencionaron en la sección b.1) de este capítulo hay que anteponerles este identificador para poderlos utilizar.

El sistema utiliza aproximadamente 12 archivos que contienen programas o datos, y son los siguientes:

- 1.- TEMUE/TEORIA/SIN
- 2.- TEMUE/ORDENA/MAESTRO/TEORIA
- 3.- TEMUE/DATOSTEXT
- 4.- TEMUE/LEE/ESC/APUNT
- 5.- TEMUE/MAESTRO/TEORIA
- 6.- TEMUE/APUNTA/TEMAS
- 7.- TEMUE/SISTEMA/FIN
- 8.- TEMUE/RESULTADOS
- 9.- TEMUE/IMPRIME
- 10.- TEMUE/IMPRIME/B
- 11.- TEMUE/SUBE/SISTEMA
- 12.- TEMUE/BAJA/SISTEMA

En caso de que se desee proteger el sistema en una cinta magnética, se requiere ejecutar el programa TEMUE/SUBE/SISTEMA; teniendo en consideración que se deberá tener una cinta registrada^{7/} lista para su uso; el número registro de la cinta se proporcionará al JOB (entre comillas), para que este automáticamente se encargue de "subir" todos los archivos en la biblioteca del usuario, cuyo nombre inicie con el identificador-TEMUE, la instrucción para la ejecución del JOB, es:

START TEMUE/SUBE/SISTEMA("1541 /ES")

^{7/} Se deberá registrar ésta cinta, en el depto. de Relaciones de la Dirección General de Cómputo Académico (antes PUC), en donde le asignarán un número de registro, por ejemplo: 1541/ES

Es importante mencionar que en esta cinta, sólo se encontrarán archivos del sistema, que está etiquetada y grabada a 1600 BPI (bit's por pulgada).

En caso de que la protección al sistema fallara, el programa desplegará unos mensajes en la consola del operador, - que también se imprimiran en el "JOB SUMMARY" (impresión en - papel) de la ejecución.

Si se requiere "bajar" el sistema en la clave en que se - "cargó" originalmente, se necesita ejecutar el JOB TEMUE/ -- BAJA/SISTEMA, teniendo en cuenta las mismas consideraciones - que se hicieron para "subir" el sistema: éste programa grabará en disco, aquellos programas que no estén presentes en el - "pack" al momento de ejecutarlo.

En el caso de que se desee "bajar" el sistema en una - clave diferente, a la que se utilizó para protegerlo en la - cinta magnética, se deberá pedir al operador, que cambie la - clave de la cinta por la nueva que se utilizará; esto se puede hacer enviando un mensaje vfa la terminal al operador, por ejemplo :

```
? SS SPO BUENOS DIAS/TARDES; POR FAVOR CAMBIE LA -  
CLAVE DE LA CINTA 1541/ES DE SC80 A JB82 , -  
PORQUE NECESITO BAJAR LA INFORMACIÓN DE ESA -  
CINTA, MUCHAS GRACIAS. (WAIT OK? )
```

una vez que el operador conteste, que el cambio ha sido reali-
zado, se podrá correr el JOB TEMUE/BAJA/SISTEMA.

En caso de efectuar modificaciones a alguno de los programas, se sugiere compilarlos nuevamente, y revisar que archivos de entrada o salida se encuentran afectados con las - correcciones realizadas.

Si se llegara a dar la situación de que se hubieran perdido alguno de los siguientes archivos :

- TEMUE/APUNTA/TEMAS
- TEMUE/MAESTRO/TEORIA
- TEMUE/DATOSTEXT

y el usuario no pudiera bajar el sistema completo de cinta, - se recomienda recuperar, o bajar, según sea el caso, el archivo TEMUE/TEORIA/SIN, y ejecutar TEMUE/ORDENA/MAESTRO/TEORIA y TEMUE/LEE/ESC/APUNT, en ese orden, con lo que se volverían a crear los archivos indispensables para el funcionamiento del programa principal TEMUE/SISTEMA/FIN.

Una vez mas se recomienda, revisar cuidadosamente los - programas y subprogramas (así como los archivos que utilizan) antes de efectuar una modificación al sistema.

b.5) Cómo Construir Archivos y Nociones de CANDE.

Para concluir este capítulo se explicará la forma adecuada de construir archivos de datos.

Cómo casi todas las computadoras macros, la máquina B7800 tiene una forma de interactuar con los usuarios, en el caso específico de la Burroughs la forma de trabajo puede ser vía tarjetas perforadas o vía CANDE.

El sistema CANDE (Command and Edit), es un editor que permite al usuario tener acceso al sistema de tiempo compartido, actuando en forma interactiva desde una terminal (remota) de teleproceso con la computadora.

Este sistema de tiempo compartido es particularmente útil para :

- ' Entrada de datos (incluyendo también la entrada de programas fuente, escritos en cualquier lenguaje).
- ' Prueba y ejecución de programas, de los que se obtiene una pequeña cantidad de información en la salida, la cual podrá ser desplegada en la misma terminal.
- ' Edición de datos.

Los archivos creados a través de CANDE, pueden ser editados, protegidos y utilizados por el usuario, por un programa, o por otros usuarios; pudiendo ser transferidos de un dispositivo a otro, tales como : disco fijo, disco removible y cinta magnética; aquellos programas introducidos como archivos pueden ser compilados y ejecutados en el momento deseado.

Para efectuar la conexión con la computadora el usuario deberá encender la terminal y buscar la tecla de transmisión

que pueden ser :

- . La tecla "NEWLINE"
- . La tecla "RETURN"
- . La tecla "ENTER"

una vez localizada ésta, se deberá presionar cada vez que se desee transmitir una línea o conjunto de caracteres a la computadora.

A continuación se presentan los comandos principales o elementales que el usuario de TEMUE/SISTEMA/FIN debe conocer para poder trabajar con la computadora :

- . HELLO (y después presionar la tecla de transmisión).
Es el comando mediante el cual se pide a la máquina el acceso a una sesión de trabajo; en respuesta a éste la computadora despliega las siguientes líneas :

```
# B7800: 356 CANDE 30.140; YOU ARE INVESTIG(48)
```

```
# ENTER USERCODE PLEASE
```

que se refieren a la presentación de la computadora, el número de terminal que se está utilizando (o desde la que se está transmitiendo), y la clave o cuenta en la que se desea trabajar.

En nuestro caso el usuario deberá teclear :

```
JB82/MJ (y después presionar la tecla de transmisión),  
la computadora a su vez desplegará el siguiente mensaje :
```

```
# SESSION 7879 20:35:12 30/12/85
```

que significa la aceptación de la clave JB82 como correcta dentro de la máquina, y el inicio de una sesión de trabajo vía CANDE, a las 20 hrs. con 35 minutos y 12 segundos, del día 30 del mes 12 del año 85.

- . WORKFILE (Archivo de Trabajo).

Un archivo de trabajo, es aquel que acaba de ser

creado a través del comando MAKE, y que no ha sido protegido (o guardado) mediante el comando SAVE, o una copia de un archivo ya protegido en la biblioteca del usuario, que está presente en el disco por medio del comando GET.

. FILE (Archivo de Biblioteca).

Cuando un archivo de trabajo es protegido con el comando SAVE, pasa a ser un FILE - archivo de biblioteca, es decir será almacenado en la biblioteca identificada con el USERCODE del usuario, hasta que sea removido con el comando REMOVE.

. MAKE

Este comando permite crear un archivo de trabajo, el cual se utilizará para almacenamiento interno durante la sesión, grabando registros ya sea de datos o de instrucciones. Un registro puede ser una línea de un programa escrito en algún lenguaje o bien una línea de datos.

En el comando MAKE, debe indicarse el nombre del archivo que se construirá, así como el tipo que tendrá, a elegir entre :

- DATA (para almacenar datos).
- SEQ (para almacenar datos).

o para almacenar algún programa escrito en :

- ALGOL
- BASIC
- COBOL
- FORTRAN
- PASCAL

por ejemplo :

```
MAKE DATOS/PRUEBA DATA
```

crea un archivo llamado DATOS/PRUEBA que contendrá datos;

otro ejemplo :

```
MAKE FORMA/1 FORTRAN
```

crea un archivo cuyo nombre es FORMA/1, que contendrá un programa escrito en Fortran.

Después de enviar el mensaje de MAKE DAT... a la computadora, ésta responde :

```
# WORKFILE DATOS/PRUEBA : DATA
```

que significa que el archivo ha sido identificado, y que no existe otro dentro de la biblioteca del usuario con el mismo nombre.

. SEQ

Una vez creado el archivo de trabajo, para introducir una línea tras otra, considerando cada línea cómo una instrucción o registro de datos, se debe asociar un número de secuencia a ésta. Hay dos maneras de asociar un número de secuencia a los registros de un archivo de trabajo :

1.- Por medio de un generador automático de secuencia, que posee CANDE, el cual es llamado con el comando SEQ.

2.- O bien, tecleando el número de secuencia de un registro, al principio de la línea que se desee introducir.

Si se utiliza la primera opción... cada vez que se introduzca una línea o registro en el archivo de trabajo y se presione la tecla de transmisión, - CANDE proporcionará el siguiente número de secuencia, empezando con el número 100, e incrementándose de 100 en 100. Si se desea una secuencia que empiece con el número 1000 y se incremente de 500 en 500, el comando se deberá escribir :

```
SEQ 1000 + 500
```

y los números de secuencia generados serán 1000, 1500, 2000, 2500, etc.

Puede utilizarse cualquier número entero como base y como incremento en la secuencia. Por ejemplo para la creación de un archivo de datos se debe teclear lo siguiente :

```
MAKE DATOS/PRUEBA/1 DATA (y transmitir)
# WORKFILE DATOS/PRUEBA/1 : DATA
SEQ 50 + 10 (y transmitir)
50 10.75 14.7 23.95 (y transmitir)
60 14.95 19.7 28.95 "
70 19.25 17.1 35.00 "
80 17.35 17.5 19.55 "
90 (presionar la tecla de transmisión)
#
```

como se puede observar en la secuencia 90 se volvió aplicar la tecla de transmisión inmediatamente después de que apareció el número de secuencia, esto se hizo con el fin de salirse de la secuencia de CANDE (o romper la ejecución del comando SEQ). Si se requiere introducir un registro más, únicamente hay que volver a utilizar el comando SEQ, indicando el número de secuencia que se desea, por ejemplo :

```
SEQ 75 (y transmitir)
75 16.25 18.1 20.75
```

En el caso de archivos de trabajo tipo DATA, la secuencia no será introducida en el registro, esto es, los 80 caracteres de los que se dispone en un archivo de este tipo, se utilizarán para los datos.

Si se utiliza la segunda opción del comando SEQ (secuencia no automática), se debe introducir la secuencia al principio de cada línea, tomando en cuenta el número de secuencia en el registro -

anterior del archivo de trabajo. El primer caracter de éstos registros, después de la secuencia no puede ser un número, deber ser un caracter - alfabético, especial, o un blanco.

. SAVE

Una vez que se ha terminado de crear un archivo de trabajo, el cual se desea utilizar posteriormente, es necesario protegerlo en discom, almacenándolo en al biblioteca, bajo la clave del usuario, mediante el comando SAVE.

Si se desea cambiar el nombre del archivo de trabajo, la opción que se utiliza es SAVE AS, por ejemplo :

SAVE AS PRUEBA/2

y el archivo de trabajo que fue creado con otro nombre, ahora por medio de la intrucción anterior será protegido con el nombre PRUEBA/2.

. GET

Un archivo almacenado en disco, puede ser requerido para ser utilizado como un archivo de trabajo, esto se puede lograr con el uso del comando GET, por ejemplo:

GET TEMUE/ORDENA/MAESTRO

El archivo que se llame o recupere, deberá existir en disco, y únicamente un copia de él, será utilizada como un archivo de trabajo, si este es protegido nuevamente, y la antigua versión se encuentra aún almacenada con el mismo nombre, el antiguo archivo será reemplazado (removido) del disco, y se tendrá ahora la nueva versión.

. LIST

Este comando sirve para listar (u observar) completamente un archivo en la terminal. Si sólo se teclaea el comando LIST, se desplegará el archivo que se encuentre presente como un archivo de trabajo.

Si un nombre de archivo se teclaea después del comando List, el contenido de este será listado. - Es importante aclarar que el archivo que se desea observar mediante éste comando, deberá existir - almacenado en disco, bajo la clave del usuario y con ese nombre. por ejemplo :

LIST TEMUE/SUBE/SISTEMA

Quando sólo se quiera listar una parte del archivo de trabajo, el comando LIST deberá contener el rango de la secuencia que se desea observar, por ejemplo :

LIST 2300-2900

. WRITE

A través de este comando, es posible listar un - archivo por impresora.

Para listar el archivo de trabajo, sólo se debe teclrear el comando WRITE, si se desea listar un archivo de biblioteca, se tiene que teclrear el - nombre de éste después del comando WRITE, por - ejemplo :

WRITE TEMUE/IMPRIME/B

. REMOVE

Si ya no se desea utilizar un archivo de trabajo

este puede ser removido por medio del comando -
REMOVE, por ejemplo :

REMOVE

Si se desea remover un archivo de biblioteca el
nombre de este debe escribirse después del coman
do REMOVE, por ejemplo :

REMOVE DATOS/PRUEBA/3

. FIX

Este comando permite, cambiar el contenido de -
una línea en un archivo de trabajo, insertando -
un texto o reemplazando parte de la línea, por -
ejemplo :

FIX 3500/XX/XY

cambiará en la línea 3500 las primeras letras XX
que encuentre en ésta, por las letras XY.

. DELETE

Este comando se utiliza para suprimir :

- una línea específica,
 - un conjunto de líneas de secuencia discon
tinua,
 - parte del archivo de trabajo,
- por medio del número de secuencia, o bien todo -
el archivo de trabajo, por ejemplo :

DELETE 3200

DELETE 3700,3900,4500,7900

DELETE ALL

la primera instrucción borrará la línea 3200, la
segunda las 4 líneas, y la tercera todo el conte-
nido del archivo.

. MOVE

Este comando mueve líneas de un archivo de trabajo, dentro de él mismo, cambiando su secuencia y suprimiéndolas de su posición original, por ejemplo:

```
MOVE 310 TO 410
```

la línea que tenía el número de secuencia 310 se encontrará ahora en la secuencia 410, otro ejemplo:

```
MOVE 235-350 TO 620 + 5
```

las líneas 235 a 350 se encontrarán ahora a partir del número de secuencia 610, incrementándose de 5 en 5 .

. COMPILE

Para compilar un programa en cualquiera de los lenguajes que maneja la BURROUGHS, se utiliza el comando COMPILE. Si se desea compilar un programa que esté presente como un archivo de trabajo solo se tecleará el comando.

En caso de querer compilar un programa que se encuentre en biblioteca, se tecleará el comando y el nombre del programa, por ejemplo :

```
COMPILE TEMUE/IMPRIME
```

compilará el programa TEMUE/IMPRIME, generando su código de máquina.

Si en la compilación se detectan errores de sintaxis, estos serán listados en la terminal hasta un máximo de 10.

. RUN

El comando RUN, causa la compilación, si ésta no se ha efectuado desde la última actualización, y la ejecución del programa.

Si se tecllea el comando únicamente, el programa que esté como archivo de trabajo será ejecutado. En caso de tecllear el comando y el nombre de algún programa que esté en biblioteca, se ejecutará dicho programa, por ejemplo :

RUN TEMUE/SISTEMA/FIN

. SPLIT

A través de este comando es posible terminar la sesión de CANDE y comenzar una nueva sin necesidad de "desconectarse" del sistema.

También se utiliza para liberar la impresión almacenada en el "BACK UP" de la clave.

. BYE

Para "desconectarse" del sistema, se utiliza el comando BYE, el cual causará la aparición de los siguientes mensajes :

#END SESSION 7879 ET=10_18:01 PT=1.6 IO=2.2

#USER=JB82 22:15:01 30/12/85

terminando la sesión.

Nota Aclaratoria.: La mayoría de los comandos que se han presentado, pueden utilizarse en forma abreviada, éstos sólo son unos cuantos de todos los que tiene CANDE, para mayor información acerca de éstos se sugiere al usuario consultar el Manual indicado que se encuentra en la bibliografía.

5.- APLICACION DEL SISTEMA A UNA POBLACION

En este capítulo se presenta una aplicación para ejemplificar el uso del sistema para dos tipos de población, una ficticia generada por la computadora, y la otra real tomada del X Censo Nacional de Población y Vivienda.

En el primer caso se trata de utilizar la computadora - como un instrumento generador de poblaciones para un muestreo aleatorio simple.

En el segundo y tercer ejercicio se trabaja con información real del estado de Tabasco, para las siguientes variables:

- Población Total.
- Población Económicamente Activa (PEA).
- PEA en el sector primario.
- PEA en el sector secundario.
- PEA en el sector terciario.
- Otros Activos.
- Población Económicamente Inactiva (PEI).

Esta información se encuentra grabada en el archivo TEMUE/DATOS /TAB, y contiene 266 registros que corresponden a las 266 AGEBS que conforman el estado.

Las AGEBS - (Area Geoestadística Básica) son una subdivisión de una área mayor denominada área geoestadística municipal. La superficie de esta última corresponde, en general, a la del municipio en donde la AGEB se encuentra localizada, con el objeto de que sus límites sean fácilmente reconocibles en el campo. Este tipo de subdivisión se utilizó por primera vez en el X Censo Nacional de Población y Vivienda.

A) Población generada por la computadora para un -
diseño muestreo aleatorio simple.

Se solicitó al sistema generar una población de 275 elementos con distribución uniforme y parámetros $A=100$ y $B=1500$, el usuario proporcionó el tamaño de muestra, y se efectuó una selección con reemplazo. A continuación se presentan los resu
tados obtenidos para la estimación de los parámetros correspon
dientes.

BIENVENIDO AL SISTEMA
APRENDAMOS
MUESTREO

TECLEA RETURN :

DESEAS VER LA PRESENTACION?

1.-SI

2.-NO

SELECCIONA TU OPCION :

1

PRESENTACION

ESTE SISTEMA TIENE COMO OBJETIVO FUNDAMENTAL EL MOSTRAR QUE LA COMPUTADORA PUEDE SER UTILIZADA PARA LA ENSEÑANZA DEL MUESTREO. POR LO QUE, TENIENDO EN CUENTA LA NATURALEZA DEL TEMA, EL SISTEMA SE DIVIDE EN :

- A) TEORIA : DONDE SE EXPLICAN CONCEPTOS, DEFINICIONES Y FORMULAS.
- B) APLICACION : AQUI SE DESARROLLAN EJEMPLOS SIMULANDO MUESTREOS, QUE FACILITAN LA COMPRESION DE LAS DIFERENTES TECNICAS .

TECLEA RETURN :

QUIERES CONOCER LA NOTACION EMPLEADA EN EL SISTEMA?

1.-SI

2.-NO

SELECCIONA TU OPCION :

1

ACLARACIONES
=====

DADO QUE EN LA COMPUTADORA NO SE PUEDEN REPRESENTAR LETRAS - MINUSCULAS Y SIMBOLOS MATEMATICOS, SE ADOPTO UNA SIMBOLOGIA QUE PERMITE REPRESENTAR LAS ECUACIONES QUE APARECEN EN EL -- DESARROLLO DEL TEMA EN GUESTION, POR LO QUE ES IMPORTANTE LA DESCRIPCION DE LOS SIGUIENTES PUNTOS :

SIMBOLO	DESCRIPCION
Y.	CUANDO APAREZCA UN PUNTO JUNTO A UNA LETRA REPRESENTA UNA LETRA MINUSCULA, ES DECIR, Y MINUSCULA
Y.1	LA I REPRESENTA EL INDICE DE Y.
Ø	SIRVE PARA IDENTIFICAR EL OPERADOR SUMA
YMED	CUANDO EL TERMINO 'MED' ACOMPAÑA A UNA LETRA, SIGNIFICA 'LA MEDIA', ES DECIR, Y MEDIA
YEST	EL TERMINO 'EST' SIGNIFICA 'ESTIMADA'
S**2	VARIANZA CON (N-1) GRADOS DE LIBERTAD.
S.**2(EST)	VARIANZA ESTIMADA (NOTE QUE ES S MINUSCULA)
#	REPRESENTA DERIVADA PARCIAL.
%	REPRESENTA 'SEMEJANTE A'
(1-F.)/N.	CON ESTO SE QUIERE REPRESENTAR QUE TODAS LAS FORMULAS SERAN EXPRESADAS EN NOTACION 'FORTRAN'
[REPRESENTA LA VARIANZA CON N GRADOS DE LIBERTAD
]	REPRESENTA LA LETRA LAMBDA
C(N;M)	REPRESENTA LAS COMBINACIONES DE N TOMADAS DE M EN M.
INT(A;B)	REPRESENTA LA INTEGRAL CON LIMITE SUPERIOR A Y LIMITE INFERIOR B.

TECLÉA RETURN :

NOTA IMPORTANTE
=====

LA TEORIA UTILIZADA PARA LA ELABORACION DE ESTA TESIS FUE CONSULTADA DE LOS LIBROS QUE A CONTINUACION SE MENCIONAN, SI SE DESEA MAYOR INFORMACION ACERCA DE ALGUNO DE LOS DISEÑOS QUE AQUI SE PRESENTAN, SE INVITA AL USUARIO A REVISARLOS.

TECLÉA RETURN :

COCHRAN, W. G. (1981), "TECNICAS DE MUESTREO"
ED. C.E.C.S.A., 2A-EDICION.

RAJ, DES (1977), "TEORIA DEL MUESTREO"
ED. FONDO DE CULTURA ECONOMICA, 2A-EDICION.

RAJ, DES (1979), "LA ESTRUCTURA DE LAS ENCUESTAS POR MUESTREO"
ED. FONDO DE CULTURA ECONOMICA, 1A-EDICION.

KISH, LESLIE (1979), "MUESTREO DE ENCUESTAS"
ED. TRILLAS, 2A-REIMPRESION.

TORRES E ZITRON (1978), "SISTEMAS NUMERICOS"
ED. LIMUSA, 3A-EDICION.

MC CRACKEN DANIEL D. (1980), "PROGRAMACION FORTRAN IV"
ED. LIMUSA, 2A-EDICION.

BURROUGHS CORPORATION (1978), "FORTRAN REFERENCE MANUAL"
SERIES B 7500 / 8000

TECLÉA RETURN :

ACLARACIONES
 =====

DADO QUE EN LA COMPUTADORA NO SE PUEDEN REPRESENTAR LETRAS -
 MINUSCULAS Y SIMBOLOS MATEMATICOS SE ADOPTO UNA SIMBOLOGIA
 QUE PERMITE REPRESENTAR LAS CUACIONES QUE APARECEN EN EL --
 DESARROLLO DEL LEMA EN CUESTION POR LO QUE ES IMPORTANTE LA
 DESCRIPCION DE LOS SIGUIENTES PUNTOS :

SIMBOLO	DESCRIPCION
Y.	CUANDO APAREZCA UN PUNTO JUNTO A UNA LETRA REPRE SENTA UNA LETRA MINUSCULA, ES DECIR, Y MINUSCULA
Y.I	LA I REPRESENTA EL INDICE DE Y.
Ø	SIRVE PARA IDENTIFICAR EL OPERADOR SUMA
YMED	CUANDO EL TERMINO 'MED' ACOMPAÑA A UNA LETRA, SIG NIFICA 'LA MEDIA', ES DECIR, Y MEDIA
YEST	EL TERMINO 'EST' SIGNIFICA 'ESTIMADA'
S**2	VARIANZA CON (N-1) GRADOS DE LIBERTAD.
S.**2(EST)	VARIANZA ESTIMADA (NOTE QUE ES S MINUSCULA)
#	REPRESENTA DERIVADA PARCIAL.
Σ	REPRESENTA 'SEMEJANTE A'
(1-F ₁)/N.	CON ESTO SE QUIERE REPRESENTAR QUE TODAS LAS FOR MULAS SERAN EXPRESADAS EN NOTACION 'FORTRAN'
E	REPRESENTA LA VARIANZA CON N GRADOS DE LIBERTAD REPRESENTA LA LETRA LAMBDA
C(N;M)	REPRESENTA LAS COMBINACIONES DE N TOMADAS DE M EN N.
INT(A;B)	REPRESENTA LA INTEGRAL CON LIMITE SUPERIOR A Y LIMITE INFERIOR B.

TECLEA RETURN :

NOTA IMPORTANTE
 =====

LA TEORIA UTILIZADA PARA LA ELABORACION DE ESTA TESIS
 FUE CONSULTADA DE LOS LIBROS QUE A CONTINUACION SE *
 MENCIONAN, SI SE DESEA MAYOR INFORMACION ACERCA DE *
 ALGUNO DE LOS DISEÑOS QUE AQUI SE PRESENTAN, SE INVI-
 TA AL USUARIO A REVISARLOS.

TECLEA RETURN :

COCHRAN, W. G. (1981). "TECNICAS DE MUESTREO"
 ED. C.E.C.S.A., 2A-EDICION.

RAJ, DES (1977). "TEORIA DEL MUESTREO"
 ED. FONDO DE CULTURA ECONOMICA, 2A-EDICION.

RAJ, DES (1979). "LA ESTRUCTURA DE LAS ENCUESTAS POR MUESTREO"
 ED. FONDO DE CULTURA ECONOMICA, 1A-EDICION.

YISH, LESLIE (1979). "MUESTREO DE ENCUESTAS"
 ED. TRILLAS, 2A-REIMPRESION.

TORRES E ZITRON (1978). "SISTEMAS NUMERICOS"
 ED. LIMUSA, 3A-EDICION.

MC CRACKEN DANIEL D. (1980). "PROGRAMACION FORTRAN IV"
 ED. LIMUSA, 2A-EDICION.

BURROUGHS CORPORATION (1978). "FORTRAN REFERENCE MANUAL"
 SERIES B 7000 / 96000

TECLEA RETURN :

DESEAS CONSULTAR SOBRE :

- 1.-T E O R I A
- 2.-A P L I C A C I O N
- 3.-SE CONCLUYE LA CONSULTA

SELECCIONA LA OPCION :

2

* B I E N V E N I D O *
* V A M O S A J U G A R A L M U E S T R E O *

TECLEA RETURN :

SE PUEDE HACER EL MUESTREO BAJO DOS FORMAS :

- 1.- SIN REEMPLAZO, ES DECIR, TOMADO UN DATO ESTE NO SE PODRA CONSIDERAR NUEVAMENTE PARA SER ELEGIDO, O
- 2.- CON REEMPLAZO

SELECCIONA TU OPCION :

2

A PARTIR DE ESTE MOMENTO COMIENZA EL MUESTREO
ESPERA UN MOMENTO

DAME UN NUMERO DE DIEZ CIFRAS :

45789784.00

GENERACION DE NUMEROS ALEATORIOS UTILIZANDO LA FUNCION RANDOM
 DE LA MAQUINA
 CANTIDAD DE NUMEROS A GENERAR : 47
 RANGO DE VALORES QUE DEBEN TOMAR DE : 1.0 A 275.0
 CON SEMILLA DADA : 45789784.000

TABLA DE NUMEROS ALEATORIOS
 DISTRIBUCION UNIFORME
 ENTRE 1 Y 275

1	24	71	95
2	11	10	10
3	11	10	10
4	11	10	10
5	11	10	10
6	11	10	10
7	11	10	10
8	11	10	10
9	11	10	10
10	11	10	10
11	11	10	10
12	11	10	10
13	11	10	10
14	11	10	10
15	11	10	10
16	11	10	10
17	11	10	10
18	11	10	10
19	11	10	10
20	11	10	10
21	11	10	10
22	11	10	10
23	11	10	10
24	11	10	10
25	11	10	10
26	11	10	10
27	11	10	10
28	11	10	10
29	11	10	10
30	11	10	10
31	11	10	10
32	11	10	10
33	11	10	10
34	11	10	10
35	11	10	10
36	11	10	10
37	11	10	10
38	11	10	10
39	11	10	10
40	11	10	10
41	11	10	10
42	11	10	10
43	11	10	10
44	11	10	10
45	11	10	10
46	11	10	10
47	11	10	10

NUMEROS ALEATORIOS FINALES

08	35	24	268	35	247	169	175	131	397
17	228	103	190	152	108	195	180	287	120
41	85	107	190	152	260	55	184	299	205
22	248	44	142	19	194				

HA TERMINADO LA SELECCION DE LA MUESTRA.
 AHORA COMIENZA EL CALCULO DE LOS ESTIMADORES
 ESPERA UN MOMENTO POR FAVOR

CALCULO DE LA PROPORCION

DAME EL RANGO DE VALORES, SEPARADOS POR COMAS
 QUE SON LOS LIMITES DE LA CARACTERISTICA EN
 ESTUDIO, ES DECIR LOS VALORES PARA LOS CUALES
 LA VARIABLE POSEE LA CARACTERISTICA.
 LOS LIMITES SON :

400.000 , 1050.000.

ELIGE CUALQUIERA DE LOS SIGUIENTES
 NIVELES PARA LOS INTERVALOS DE CONFIANZA :

- 1.- 50 %
- 2.- 30 %
- 3.- 90 %
- 4.- 95 %
- 5.- 99 %

SELECCIONA EL NIVEL DE CONFIANZA :

MUESTREO ALEATORIO SIMPLE

CON REEMPLAZO

DATOS DE IDENTIFICACION

TIPO DE POBLACION: GENERADA TIPO DE DISTRIBUCION UNIFORME
CON A = 100.00 Y B = 1500.00000
TAMANO DE LA MUESTRA : 47 LA MUESTRA LA DIO EL USUARIO.

RESULTADO DEL MUESTREO

M U E S T R A L E S -

<u>P R O P O R C I O N P</u>	<u>T O T A L</u>
0.51064	140.426
VARIANZA DEL ESTIMADOR	
0.0054	409.3255
ERROR ESTANDARD	
0.07357	20.2318
INTERVALOS DE CONFIANZA	
0.63129	173.6057
0.38998	107.2454
TECLEA RETURN :	

DESEAS SEGUIR JUGANDO AL MUESTREO ?? :

1.-SI, PERO CON LA MISMA POBLACION

2.-SI, PERO CON UNA POBLACION DIFERENTE

3.-NO, SE CONCLUYE LA APLICACION

SELECCIONA TU OPCION :

DESEAS CONSULTAR SOBRE :

- 1.-T E O R I A
- 2.-A P L I C A C I O N
- 3.-SE CONCLUYE LA CONSULTA

SELECCIONA LA OPCION :

3

DESEAS QUE SE IMPRIMA UN LISTADO CON LA SESION
QUE ESTAS POR CONCLUIR, ESTE CONTENDRA LOS
DATOS UTILIZADOS, OPCIONES, Y RESULTADOS.

- 1.- SI
- 2.- NO

SELECCIONA TU OPCION :

2

QUE ESTAS POR CONCLUIR, ESTE CONTENDRA LOS
DATOS UTILIZADOS, OPCIONES, Y RESULTADOS.

1.- SI

2.- NO

SELECCIONA TU OPCION :

1

DESEAS QUE PERMANESCA EL ARCHIVO DE IMPRESION
EN EL DISCO ??

1.- SI

2.- NO

SELECCIONA TU OPCION :

1

* ESPERO QUE TE HAYAS DIVERTIDO MUCHO JUGANDO *
* A L M U E S T R E O *

PD. V U E L V E P R O N T O

NUMERO DE REGISTROS LEIDOS 1080

NUMERO DE PAGINAS IMPRESAS 21

B) Población real, proporcionada por el usuario -
para un muestreo estratificado.

Se estratificó la población de Tabasco con base en el -
porcentaje de población por AGEB respecto al total del estado.

La simulación del muestreo para esta población tiene las
siguientes características :

- Se construyeron cinco estratos
- La variable de interés fue la población económicamente activa (PEA) de la AGEB
- El tipo de afijación fue proporcional.
- El tamaño de muestra se calculó con base en los -
costos.
- Se seleccionó la muestra sin reemplazo dentro de
cada estrato.

A continuación se presentan los resultados obtenidos de
las estimaciones, solo se muestran las pantallas finales, ya -
que en ellas se encuentra la información que proporcionó la -
muestra en cada estrato.

MUESTREO ALEATORIO ESTRATIFICADO

SIN REEMPLAZO

ESTRATO 1

ESTRATOS DIFERENTES POBLACION DEL ESTRATO 65

M U E S T R A L E S .

M E D I A	T O T A L
111.93333	29774.26667
V A R I A N Z A	
193.13778	816007.11108
ERROR ESTANDARD	
13.89740	903.33112
INTERVALOS DE CONFIANZA	
134.72507	35836.86924
89.14159	23711.66409

TECLEA RETURN :

MUESTREO ALEATORIO ESTRATIFICADO

SIN REEMPLAZO

ESTRATO 2

ESTRATOS DIFERENTES POBLACION DEL ESTRATO 56

M U E S T R A L E S .

M E D I A	T O T A L
322.76923	85856.61538
V A R I A N Z A	1240977.67261
395.71992	
ERROR ESTANDARD	1113.99177
19.89271	
INTERVALOS DE CONFIANZA	94534.61130
355.39328	77178.61946
290.14519	

TECLEA RETURN :

MUESTREO ALEATORIO ESTRATIFICADO

SIN REEMPLAZO

ESTRATO 3

ESTRATOS DIFERENTES POBLACION DEL ESTRATO 42

M U E S T R A L E S .

M E D I A	T O T A L
539.50000	143507.00000
V A R I A N Z A	
861.78333	1520185.80000
ERROR ESTANDARD	
29.35615	1232.95815
INTERVALOS DE CONFIANZA	
597.03805	158812.12050
481.96195	128201.87950

TECLEA RETURN :

MUESTREO ALEATORIO ESTRATIFICADO

SIN REEMPLAZO

ESTRATO 4

ESTRATOS DIFERENTES POBLACION DEL ESTRATO 43

M U E S T R A L E S .

M E D I A	T O T A L
865.00000	230090.00000
V A R I A N Z A	
2495.16364	4613557.56363
ERROR ESTANDAR	
49.95161	2147.91936
INTERVALOS DE CONFIANZA	
962.90516	256132.77293
767.09484	204047.22707

TECLEA RETURN :

MUESTRO ALEATORIO ESTRATIFICADO

SIN REEMPLAZO

ESTRATO 5

ESTRATOS DIFERENTES POBLACION DEL ESTRATO 60

M U E S T R A L E S -

M E D I A	T O T A L
1985.53846	528153.23077
V A R I A N Z A	
118924.14892	428126936.09375
ERROR ESTANDARD	
344.85381	20691.22848
INTERVALOS DE CONFIANZA	
2551.09871	678592.25599
1419.97822	377714.20555

TECLEA RETURN :

MUESTREO ALEATORIO ESTRATIFICADO

SIN REEMPLAZO

DATOS DE IDENTIFICACION

TIPO DE POBLACION EN ARCHIVO
TAMAÑO DE LA MUESTRA 62
TAMAÑO DE LA POB. 266
TIPO DE AFIJACION PROPORC.

TIPO DE DISTRIBUCION EL USUARIO
LA MUESTRA \$\$\$POR - COSTOS\$\$:

TOTALES DE LA ESTRATIFICACION

ESTRATO	TAMAÑO DE MUESTRA	DE	POBLACION DEL ESTRATO	MEDIA EST. MUESTRAL	VARIANZA EST. MUESTRAL
1	15		65	111.9333	193.1378
2	13		56	322.7692	395.7199
3	10		42	539.5000	861.7833
4	11		43	865.0000	2495.1636
5	13		60	1985.5385	118924.1489
TOTALES	62		266	768.1844	6166.5112

RESULTADO DE LAS ESTIMACIONES TOTALES

MEDIA EST. = 768.184 VARIANZA MEDIA EST. = 6166.511
 TOTAL = 204337.051 VARIANZA DEL TOTAL = 436317664.240

TECLEA RETURN :

C) Población real, proporcionada por el usuario - para un muestreo de conglomerados.

Se conglomeró la población de Tabasco con base en el porcentaje de población económicamente activa (PEA) de la AGEB - respecto al total (de PEA) del estado.

La simulación del muestreo para esta población tiene las siguientes características :

- Se construyeron cinco conglomerados de distinto tamaño.
- El diseño es conglomerados sin submuestreo.
- La variable de interés fue la población por AGEB.
- Tres conglomerados obtenidos con muestreo aleatorio simple conformaron la muestra.

A continuación se presentan las estimaciones obtenidas - para cada conglomerado en muestra, así como los resultados finales del muestreo.

GENERACION DE NUMEROS ALEATORIOS UTILIZANDO LA FUNCION RANDOM DE LA MAQUINA

CANTIDAD DE NUMEROS A GENERAR : 3
RANGO DE VALORES QUE DEBEN TOMAR DE : 1.0 A 5.0
CON SEMILLA DADA : 5848712.000

TABLA DE NUMEROS ALEATORIOS
DISTRIBUCION UNIFORME
ENTRE 1 Y 5

1	4-58	5
2	3-59	4
3	2-25	2

NUMEROS ALEATORIOS ORDENADOS

2
4
5

SE REPITIERON 0 NUMEROS ALEATORIOS

NUM. ALEATORIOS FINALES

2
4
5

MUESTREO DE CONGLOMERADOS

DE TAMA#OS DIFERENTES SIN SUBMUESTREO

ESTIMACIONES PARA EL 2 CONGLOMERADO.
CON TAMA#O DE POBLACION 53

TECLEA RETURN :

EMPIEZA EL CALCULO, ESPERA UN MOMENTO

C A L C U L O

MEDIA ESTIMADA	=	1020.94340	TAMAÑO DE LA POBL.	=	53
VARIANZA ESTIMADA	=	36509.18548	ERROR ESTANDAR	=	237.71661
VALOR MINIMO	=	549.00000	VALOR MAXIMO	=	1477.00000
R A N G O	=	948.00000			

TECLEA RETURN :

ESTIMACIONES PARA EL 4 CONGLOMERADO.
CON TAMAÑO DE POBLACION 42

TECLEA RETURN :

EMPIEZA EL CALCULO, ESPERA UN MOMENTO

C A L C U L O

MEDIA ESTIMADA	=	2594.59524	TAMAÑO DE LA POBL.	=	42
VARIANZA ESTIMADA	=	212756.38379	ERROR ESTANDAR	=	461.25523
VALOR MINIMO	=	1538.00000	VALOR MAXIMO	=	3804.00000
R A N G O	=	2266.00000			

TECLEA RETURN :

ESTIMACIONES PARA EL 5 CONGLOMERADO.
CON TAMAÑO DE POBLACION 63

TECLEA RETURN :

EMPIEZA EL CALCULO, ESPERA UN MOMENTO

C A L C U L O

MEDIA ESTIMADA	=	6742.77778	TAMAÑO DE LA POBL.	=	63
VARIANZA ESTIMADA	=	14909122.99805	ERROR ESTANDAR	=	3861.23335
VALOR MINIMO	=	2891.00000	VALOR MAXIMO	=	20286.00000
R A N G O	=	17395.00000			

TECLEA RETURN :

MUESTREO DE CONGLOMERADOS

SIN REEMPLAZO

DATOS DE IDENTIFICACION

TIPO DE POBLACION EN ARCHIVO	TIPO DE DISTRIBUCION EL USUARIO
TAMAÑO MAXIMO DE LA MUESTRA :	158 LA MUESTRA DE CONGLOMERADOS N°
TAMAÑO DE LA PÖB. 266	
TIPO DE AFIJACION CONG* censo	

TOTALES DE LA CONGLOMERACION

CONGLO.	TAMAÑO DE MUESTRA	POBLACION DEL CONGLO.	MEDIA EST. MUESTRAL	VARIANZA EST. MUESTRAL
1	53	53	1020.9434	56509.1855
2	42	42	2594.5952	212756.3838
3	63	63	6742.7773	114909122.9980
TOTALES	158	158	3697.3459	1899931.5982

TECLEA RETURN :

RESULTADO DE LAS ESTIMACIONES TOTALES
 ::

MEDIA EST. = 3697.346 VARIANZA MEDIA EST. = 1899931.598
 TOTAL = 979796.667 VARIANZA DEL TOTAL = 133422696486.000

TECLEA RETURN :

6.- CONCLUSIONES

El desarrollo de sistemas como el presentado cumple el objetivo de facilitar el aprendizaje, proporcionar al usuario las herramientas indispensables en el cálculo de los estimadores y a la vez revisar los conocimientos del tema en cuestión, sin embargo el desarrollo de éstos debe hacerse con mucho cuidado y siguiendo una metodología previamente definida en forma conjunta por el experto (en la materia) y el constructor, de esta forma los problemas que se presenten durante la construcción podrán fácilmente solucionarse si ambos los afrontan.

El pretender transportar o implementar sistemas como éste en otra computadora diferente a la utilizada para su elaboración, debe evaluarse muy cuidadosamente, ya que, por lo general cuando se realiza el diseño de sistemas como el presente, se hace uso de todos los recursos y facilidades de la máquina y del lenguaje en que se programa, y éstas no es común encontrarlas en la computadora a la que se quiere transportar el sistema, lo anterior tiene como consecuencia que es preferible volver a desarrollar el sistema, que tratar de implementarlo en una nueva máquina.

La participación de este sistema en la enseñanza por -- computadora se ve reflejada en las facilidades de uso, características del procesamiento, sencillez de las explicaciones, y construcción modular, que permiten su explotación, modificación y crecimiento futuro.

La explotación y buen uso que se de a este sistema, dependerá en gran parte del conocimiento previo que tenga el -- usuario de la teoría del muestreo, así como del sistema mismo.

El profesor de un curso de introducción a la teoría del muestreo, podrá construir sus ejemplos y evitar con esto efectuar

la "talacha" en el pizarrón, también podrá utilizar el sistema como un paquete para el cálculo de los estimadores de una población real.

Se invita a los usuarios y/o lectores de esta tesis a - detallar sus comentarios y sugerencias sobre el posible crecimiento de este sistema, así, como de las experiencias obtenidas al hacer uso de él.

7.- BIBLIOGRAFIA

- 1.- COCHRAN, W. G. (1981), "TECNICAS DE MUESTREO", ED. C.E.C.S.A., 2a. edición.
- 2.- RAJ, DES (1977), "TEORIA DEL MUESTREO", ED. FONDO DE CULTURA ECONOMICA, 2a. edición.
- 3.- RAJ, DES (1979), "LA ESTRUCTURA DE LAS ENCUESTAS - POR MUESTREO", ED. FONDO DE CULTURA ECONOMICA, 1a. edición.
- 4.- KISH, LESLIE (1979), "MUESTREO DE ENCUESTAS", ED. TRILLAS, 2a. reimpresión.
- 5.- TORRES E ZITRON (1978), "SISTEMAS NUMERICOS", ED. LIMUSA, 3a. edición.
- 6.- MC CRACKEN, DANIEL D. (1980), "PROGRAMACION FORTRAN IV", ED. LIMUSA, 3a. edición.
- 7.- BURROUGHS CORPORATION (1978), "FORTRAN" REFERENCE MANUAL, SERIES B7000 / B6000.
- 8.- BURROUGHS CORPORATION (1980), "CANDE" REFERENCE - MANUAL, SERIES B7000 / B6000.
- 9.- BURROUGHS CORPORATION (1977), "WORK FLOW LENGUAJE" REFERENCE MANUAL, SERIES B7000/B6000.
- 10.- BURROUGHS CORPORATION (1977), "I / O SUBSYSTEM" - REFERENCE MANUAL, SERIES B7000/B6000.
- 11.- BURROUGHS CORPORATION (1977), "ALGOL" REFERENCE - MANUAL, SERIES B7000 / B6000.

8.- ANEXOS : Programas del Sistema.

TEMUE/ORDENA/MAESTRO/TEORIA (02/03/86):

1019: PA 00100001, LENGHT 10, 1728

100	SET LIST	0000100
200	BEGIN	0000200
300	BEGIN	0000300
400	REMOVEFIL("TEMUE/DATOSTEXT.");	0000400
500	END;	0000500
600	BEGIN	0000600
700	FILE A(KIND=DISK,TITLE="TEAJE/TEORIA/SIN.",FILETYPE=7),	0000700
800	FILE B(KIND=DISK,MAXRECSIZE=14,BLOCKSIZE=420,PROTECTION=1,	0000800
900	TITLE="TEMUE/DATOSTEXT.");	0000900
1000	COOLAN PROCEDURE COMP(X,Y);	0001000
1100	ARRAY X,Y(0);	0001100
1200	COMP:=POINTER(A)<POINTER(Y) FOR 9;	0001200
1300	SORT(3,A,B,COMP,14);	0001300
1400	END;	0001400
1500	END.	0001500

TEMUE/IMPRI ME /: (11/05/85)

9:41 PM TUESDAY, JANUARY 29, 1986

```
100      >SET LIST                                CCCCC100
200     FILE 5(KIND=DISK,TITLE="TEMUE/RESULTADOS",FILETYPE=7)      CCCC0200
300     FILE 6(KIND=PRINTER,MAXRECSIZE=22)                          CCCC0300
400     FILE 9=ENTRA,UNIT=REMOTE,10                                0000400
500           DIMENSION A(17), B(5)                                00000500
600           DATA B/"? REMOVE TEMUE/RESULTADOS  " /            00000600
700           INTEGER PAGINA                                       00000700
800           PAGINA = 1                                           00000800
900           > READ(5,10,END=99) (A(I), I=1,14)                  CCCC08CC
1000        10  FORMAT(14A6)                                       00000900
1100           NR =NR + 1                                          CCCC1000
1200           IF(NR.LT.54)WRITE(6,15) (A(I), I=1,14)            00001100
1300        15  FORMAT(5X,14A6)                                     CCCC1200
1400           IF(NR.LT.54)GO TO 5                                CCCC1300
1500           PAGINA = PAGINA + 1                                00001400
1600           WRITE(6,20)PAGINA                                  00001500
1700        20  FORMAT(1H1, 65X,"PAGINA ",15,/)                  CCCC1600
1800           WRITE(6,15) (A(I), I=1,14)                          00001700
1900           NREG = NREG + NR                                    CCCC1800
2000           NR = 0                                             CCCC1900
2100           GO TO 5                                           CCCC2000
2200        99  WRITE(6,100)NREG,PAGINA                             CCCC2100
2300        100 FORMAT(//,20X,"NUMERO DE REGISTROS LEIDOS ",15    CCCC2200
2400           * //,20X,"NUMERO DE PAGINAS IMPRESAS ",15,/)      CCCC2300
2500           ZIP WITH B                                         CCCC2400
2600        130 CALL EXIT                                         CCCC2500
2700           END                                               CCCC2600
                                           00002700
```


3200		APSR1(TEMAA,ISUBT) = ?I	00003200
3300		APSRF(TEMAA,ISUBT) = REGGB	00003300
3400		RI = REGGB + 1	00003400
3500		SUBTA = SUBT	00003500
3600		GO TO 20	00003600
3700	40	WRITE(7,50)TEMAA,TEMA	00003700
3800	50	FORMAT(//,10X,"ESTO ES TEMA : ",I5,5X,"ESTO ES TEMA : ",I5)	00003800
3900		PRINT*//,TEMAA	00003900
4000		ISUBT = SUBTA + 1	00004000
4100		APSR1(TEMAA,ISUBT) = RI	00004100
4200		APSRF(TEMAA,ISUBT) = REGGB	00004200
4300		APTRF(TEMAA) = REGGB	00004300
4400		RI = REGGB + 1	00004400
4500		TEMAA = YEMA	00004500
4600		SUBTA = SUBT	00004600
4700		GO TO 20	00004700
4800	55	WRITE(6,60)LEIDOS,ESCRIT	00004800
4900		WRITE(7,60)LEIDOS,ESCRIT	00004900
5000	60	FORMAT(//,5X,"EL NUMERO DE REGISTROS LEIDOS FUE : ",I6,//,5X,	00005000
5100	*	"EL NUMERO DE REGISTROS ESCRITOS FUE : ",I6,//)	00005100
5200		DO 80 J=1,5	00005200
5300		WRITE(6,70) (APSR1(J,I),I=1,14)	00005300
5400	70	FORMAT(//,5X,14(2X,F5.0),3X)	00005400
5500	EU	CONTINUE	00005500
5600		DO 100 J=1,5	00005600
5700	100	WRITE(6,70) (APSRF(J,I),I=1,14)	00005700
5800		WRITE(6,140) (APTRF(I),I=1,10)	00005800
5900	140	FORMAT(//,5X,"ESTOS FUERON LOS REGISTROS FINALES DE LOS TEMAS",//	00005900
6000	*	,10(F6.0,2X),//)	00006000
6100		WRITE(6,150)REGGB	00006100
6200	150	FORMAT(//,10X,"ESTOS LLEVO GRABADOS : ",I5,//)	00006200
6300	C*	WRITE(4,160)	00006300
6400	C*160	FORMAT(//,10X,"DAME EL NOMBRE DEL NUEVO ARCHIVO, CON UN PUNTO",//	00006400
6500	C* *	10X," AL FINAL : ")	00006500

6600	C*	READ(4,170)(ARN(J), J = 1,5)	0006600
6700	C*170	FORMAT(5A6)	0006700
6800	C*	CHANGE(9,TITLE=ARN)	0006800
6900		DO 180 I=1,5	0006900
7000	18U	WRITE(9,190)(APSR I(I,J),J=1,14)	0007000
7100	19U	FORMAT(14(F5.0,))	0007100
7200		DO 200 I=1,5	0007200
7300	20U	WRITE(9,190)(APSR F(I,J),J=1,14)	0007300
7400		CLOSE(8,DISP=KEEP)	0007400
7500		CLOSE(9,DISP=KEEP)	0007500
7600		CALL EXIT	0007600
7700		END	0007700