

**IMP**  
**Manual**  
**de**  
**Muestreo**  
**Estadístico**

INSTITUTO MEXICANO DEL PETROLEO

SUBDIRECCION DE ESTUDIOS ECONOMICOS  
Y PLANEACION INDUSTRIAL

**MANUAL DE MUESTREO  
ESTADISTICO**

1 9 7 4

SUBDIRECCION DE ESTUDIOS ECONOMICOS  
Y PLANEACION INDUSTRIAL

LIC. ARTURO DEL CASTILLO  
SUBDIRECTOR

*Publicación No. 74 FN/154*

**COORDINADORES:**

LIC. GENARO SANCHEZ BARAJAS  
LIC. HECTOR AVILA RODRIGUEZ

**INVESTIGADOR:**

LIC. HILDA TERAN BUTRON

# I N D I C E

	<i>Página</i>
1. INTRODUCCION.	1
2. OBJETIVOS DEL MANUAL.	2
3. METODOLOGIA DEL MUESTREO ESTADISTICO.	5
3.1 Métodos de Muestreo.	5
3.1.1 Selección de la Unidad de Muestreo.	5
3.1.1.1 Manejo de las Tablas de Números Aleatorios.	7
3.1.2 Modelos de Muestreo.	8
3.1.3 Diseño del Cuestionario.	60
3.1.4 Trabajo de Campo.	60
3.1.5 Crítica de Cuestionarios.	62
3.1.6 Codificación y Procesamiento de Datos.	63
3.1.7 Evaluación Estadística de los Resultados.	64
3.1.8 Red General de Actividades en una Encuesta de Muestreo.	67

4.	APLICACIONES A LA INDUSTRIA PETROLERA.	75
4.1	Ejemplo Hipotético.	75
4.2	Usos del Muestreo Estadístico en el Análisis Estratégico del Mercado de Lubrificantes en Nuevo León, Querétaro y Sinaloa.	79
4.3	Algunas otras Aplicaciones del Muestreo Estadístico en la Industria Petrolera.	104
5.	BIBLIOGRAFIA.	109

1. INTRODUCCION.

Las diversas actividades que desarrolla una empresa integrada de magnitud de *Petróleos Mexicanos*, requieren, en ciertas ocasiones el diseño de manuales para el buen uso de las técnicas que se utilicen. Esta observación es valedera sobre todo para aquellas disciplinas, que como el muestreo estadístico, tienen un vasto campo de aplicación y por ello continuamente usan en las áreas de trabajo de la Empresa.

El presente manual pretende exponer los métodos de muestreo más importantes, destacar sus características y propiedades relevantes, y tipo de actividades a que debe aplicarse cada uno de ellos. O sea que objetivo es mostrar la combinación de métodos que se adapte al conjunto problemas y que se amolde al temperamento y habilidades del agente que emplea.

Para situarse en la perspectiva adecuada, es necesario considerar que es la estadística. Se puede describir como la tecnología del método científico; la estadística proporciona métodos y técnicas para los investigadores y estos instrumentos pueden ser de aplicación general y útiles cualquier campo de la ciencia. Por otra parte, ciertos métodos pueden ser especialmente indicados en campos restringidos y particulares de la investigación.

Las necesidades apremiantes de nuevas tareas y de problemas nuevos, unidas al creciente conocimiento de procedimientos estadísticos por parte de los funcionarios administrativos y de los investigadores, han contribuido a este desarrollo. Con este progreso las antiguas controversias sobre métodos cualitativos o cuantitativos han sido casi olvidadas.

La última década ha presenciado un notable aumento e interés

hacia los métodos cuantitativos aplicados a las ciencias comerciales y sociales. Parece que han pasado los tiempos en que la intuición era la base primordial en los criterios comerciales y en que hipótesis insostenibles estaban en primer línea en los estudios sociológicos. Siguiendo el camino de los investigadores de las ciencias físicas, que de antiguo se consideraban como las más seguras, sociólogos e investigadores aventajados en el campo de los negocios, trabajan denodadamente con métodos de estudio que se basan en la observación y análisis de hechos.

Aunque los métodos estadísticos en general son prácticamente universales en su aplicación, siempre aparecen problemas especiales en cualquier campo de investigación. Esto es particularmente cierto en el terreno de la economía, donde aparecen muchas dificultades peculiares y problemas característicos. Los métodos que en cierto modo están dedicados a responder a estos requisitos han experimentado amplio desarrollo y exigen ser tratados en forma especial; en segundo lugar, estos métodos se aplican con más efectividad cuando se trata de asuntos concretos; la metodología abstracta carece por completo de interés para la mayoría de las personas.

El objeto que siempre se persigue es estudiar y analizar un conjunto de hechos observados en cualquier campo (una muestra), sintetizar este conjunto en unas cuantas medidas o estadísticas, generalizar estos valores para averiguar los verdaderos parámetros del universo al cual pertenece la muestra estudiada y ver que grado de confianza merecen las estimaciones.

Un método de muestreo es un procedimiento de selección que permite tomar una fracción de la población, de tal manera que la muestra seleccionada represente a la población entera, es decir, que todas las características de la misma, incluyendo la de variabilidad entre sus unidades, se reflejen en la muestra tan aproximadamente como se quiera, para que se puedan hacer inferencias a partir de la muestra, dignas de confianza.

2.

## OBJETIVOS DEL MANUAL.

- a) Que los técnicos de la industria petrolera se familiaricen con la metodología del muestreo estadístico.
- b) Proporcionar la herramienta matemático-estadística de mayor uso en la recabación de información mediante el citado método.
- c) Ilustrar su aplicación con ejemplos prácticos derivados de la actividades desarrolladas en la industria petrolera.
- d) Presentar en un solo documento los conceptos básicos y procedimientos necesarios para la realización de una encuesta, que normalmente aparecen dispersos en la literatura del muestreo estadístico.

### 3. METODOLOGIA DEL MUESTREO ESTADISTICO.

#### 3.1 Métodos de Muestreo.

Existen el muestreo empírico y el probabilista, el primero suele usarse cuando se tiene un amplio conocimiento del fenómeno que se investigará y cuando existen estudios previos al respecto, tal que el estadístico tiene antecedentes y el costo para la investigación es reducido. Este tipo de muestreo se recomienda cuando no se desea un análisis profundo y preciso sobre las características del universo que se estudia. Este método resulta en ocasiones bueno, ya que capta con relativa facilidad las características de la población en estudio. Como podrá notarse no es del todo científico y no permite, por si mismo, llegar a estimaciones precisas, resultando difícil realizar inferencias e imposible cuantificar el error en la estimación.

El método científico, por lo contrario, proporciona una medida de la magnitud del error y de la confianza con que se pueden tomar los resultados. Generalmente suele ser más costoso y quizás tome un poco más de tiempo el realizarlo, en especial cuando hay problemas de información sobre el número de unidades que integran el universo y algunas otras características que permiten un cálculo rápido del tamaño de la muestra, teniendo además que gastarse cierto número de horas-hombre en la recabación de la información requerida.

Es recomendable, sin embargo, usar siempre el método científico para dotar a los estudios de seguridad matemática, aun cuando se tenga que hacer esfuerzos extraordinarios para conseguir los recursos monetarios necesarios.

#### 3.1.1 Selección de la Unidad de Muestreo.

Los métodos de muestreo estadístico tienen por objeto, considerar

algunos elementos del universo que se trata de estudiar, para poder hacer inferencias sobre sus parámetros. La selección de estas unidades se hace a partir de una lista, mapas, croquis, directorios, o una combinación de estos elementos, los que deben contener todas las unidades de interés y permitir determinar la probabilidad de su inclusión, asimismo, en el momento de levantar la encuesta, la identificación de cada unidad en la muestra sea hecha sin ninguna ambigüedad.

Al conjunto de todos estos elementos, se le llama Marco Muestral.

De acuerdo a la forma de seleccionar estas unidades se pueden dar las siguientes definiciones:

#### **Reemplazo.**

Las selecciones sucesivas de una muestra probabilista pueden hacerse con o sin reemplazo de las unidades obtenidas en las selecciones previas. Al primer procedimiento se le llama muestreo con reemplazo y al segundo sin reemplazo.

En el muestreo con reemplazo, al hacerse las estimaciones, cada unidad de la muestra debe considerarse en un número igual al de veces que haya salido en la muestra.

#### **Etapas.**

Las unidades que tengan que investigarse a través del cuestionario, posiblemente convenga agruparlas en el marco muestral y estos grupos a su vez se vuelvan a agrupar y así sucesivamente. Dependiendo del número de agrupamiento de las unidades de interés (o últimas unidades de muestreo), es el nombre que se le dá. Si el marco muestral no presentó agrupamientos, el muestreo se llamará monoetápico (selección directa de las unidades de interés); si el marco muestral presenta agrupamientos de un solo orden se llamará bietápico, o lo que es lo mismo se seleccionarán primero los grupos de unidades (de primera etapa) y finalmente se seleccionarán los de interés o de segunda etapa, y así sucesivamente se tendrá el muestreo trietápico, tetraetápico, etc.

#### **Probabilidad.**

Si las unidades de muestreo en cada etapa son seleccionadas con la misma probabilidad, el muestreo se llamará equiprobable, en el caso contrario se dice que es de probabilidades variables de selección en la o las etapas que correspondan.

#### **Estrato.**

La precisión al hacerse las estimaciones, básicamente depende de dos factores: a) del tamaño de la muestra, b) la variabilidad o heterogeneidad de la población.

Es evidente que mientras más grande sea la muestra, representará más fielmente a la población, tal que se pueden mejorar las estimaciones aumentando el tamaño de la muestra. En cuanto al segundo factor para aumentar la precisión puede dividirse el marco muestral (si es que se dispone de los medios necesarios) en clases homogéneas llamadas estratos y seleccionar separadamente en cada estrato una muestra, garantizando en esta forma cualquier representación deseada de todos los estratos de la población. La denominación de un modelo de muestreo se forma indicando estos conceptos (etapa, probabilidad y reemplazo).

#### **3.1.1.1 Manejo de las Tablas de Números Aleatorios.**

La selección de las unidades de muestreo debe hacerse basándose en las leyes del azar, esto es, debe asignarse a cada unidad del marco muestral una probabilidad de inclusión en la muestra. Con este método la muestra se obtiene en selecciones sucesivas de una unidad, cada una con una probabilidad asignada de antemano, según sea el modelo de muestreo que se utilice, hasta completar el número de unidades que deben incluirse en la muestra para cada etapa. Un procedimiento práctico para seleccionar las unidades, es utilizando una tabla de números aleatorios como la que aparece al final del inciso. El procedimiento es el siguiente (para el muestreo monoetápico y con probabilidades variables de selección):

- a) Identifíquese las N unidades de la población con los números del 1 al N, o lo que es lo mismo preparar una lista de unidades de la población y numerarlas seriadamente.

- b) Encuéntrese el número  $Q$  - Mínimo común múltiplo de  $1, q_1, \dots, q_i, \dots, q_j, \dots, q_{j+1}$  = la probabilidad en notación racional de incluir la  $i$ -ésima unidad en la muestra. }

- c) Fórmense las clases numéricas.

$$Q \sum_{i=1}^j \frac{P_i}{q_i} \leq C_j < Q \sum_{i=1}^{j+1} \frac{P_i}{q_i}$$

- d) Tómese de la tabla un número aleatorio  $R$  menor o igual que  $Q$ , si  $R$  pertenece a la clase  $C_j (R \in C_j)$ , considérese en la muestra la  $j$ -ésima unidad; en esta forma prosigase hasta obtener todas las unidades que deban incluirse en la muestra. Para tomar números aleatorios de la tabla, considere por ejemplo, que se necesitan seleccionar 25 elementos de 500 ( $n = 25, N = 500$ ), lo cual puede hacerse con números de tres cifras de las columnas 11, 12 y 13. Así los números serán el 41, 55, 474, 231, 237, 425, 143, 174, 236, 439, 207, 265, 298, 20 y 469. La regla que se siguió en este caso (considerar las columnas 11, 12 y 13) no es la única posible, en realidad se pueden tomar renglones o combinaciones de renglones con las columnas o seguir cualquier otra regla, la única condición es que no se cambie de criterio en el transcurso de la selección de toda la muestra.

En algunos casos es posible que únicamente se disponga de un marco muestral, en el que sólo se puedan identificar las unidades de las primeras etapas y que por tal motivo la selección de las últimas unidades de muestreo deba hacerse en el momento de levantar la encuesta. Si este es el caso debe fijarse de antemano el procedimiento de selección que deba seguirse.

### 3.1.2 Modelos de Muestreo.

Los modelos de muestreo tienen por objeto indicar el número de unidades que deben incluirse en la muestra, dependiendo de la forma que éstas se seleccionen, de la confianza que se requiera al hacer las inferencias, del error de muestreo que se pueda permitir y del fondo disponible para la

realización de la encuesta.

#### Definiciones Básicas.

#### Error de Muestreo.

Sea  $\mu$  el valor de un parámetro de la población que se estudia mediante el muestreo, y  $\hat{\mu}$  una función definida mediante la muestra, que estima el valor  $\mu$ .

Se define como error de muestreo a:

$$\text{Error de muestreo} = |\mu - \hat{\mu}| \quad (1)$$

que debe ser menor o igual al máximo error de muestreo permisible,  $\epsilon = |\mu| \cdot$

#### Nivel de Confianza.

Es la probabilidad de que el error de muestreo sea menor o igual a  $\epsilon = |\mu|$ , con  $\epsilon > 0$ , o sea:

$$\text{Nivel de confianza} = P ( |\mu - \hat{\mu}| \leq \epsilon = |\mu| ) = 1 - 2\alpha \quad (2)$$

$$\text{Nivel de incertidumbre} = 2\alpha \quad (3)$$

#### Variancia.

Es un valor asociado con la concentración de los valores que pueda tomar una variable aleatoria, en un cierto intervalo que contenga a la media.

$$\text{Coeficiente de variación} = \frac{\text{Variancia}}{\text{Media}}$$

De tal manera que si  $\hat{\mu}$  es un estimador del parámetro  $\mu$  con función de distribución normal, con variancia  $V(\hat{\mu})$ , el error de muestreo satisfice la relación:

$$P \left( \frac{|\mu - \hat{\mu}|}{V(\hat{\mu})} \leq t(\alpha, \infty) \right) = 1 - 2\alpha \quad (4)$$

En donde  $t(\infty, \infty)$  es la variable aleatoria con función de distribución t-student y un infinito número de grados de libertad (llamada también distribución normal) correspondiente al valor  $\infty$  y al nivel de confianza  $1 - 2\alpha$ .

Esta relación puede ponerse en la forma:

$$P(|\mu - \hat{\mu}| \leq t(\infty, \infty) \sqrt{V(\hat{\mu})}) = 1 - 2\alpha \quad (5)$$

Esto es el máximo error de muestreo permisible, al nivel de confianza  $1 - 2\alpha$  es  $t(\infty, \infty) \sqrt{V(\hat{\mu})}$  (6)

Supóngase que  $V(\hat{\mu})$  puede hacerse variar (aumentando o disminuyendo el tamaño de la muestra, por ejemplo) de tal modo que el máximo error de muestreo permisible al nivel de confianza  $1 - 2\alpha$  sea igual a un valor fijado de antemano  $\epsilon$   $|\mu - \hat{\mu}|$  (para fines prácticos de muestreo,  $0 < \epsilon < 1$ ), esto lleva a la ecuación:

$$t(\infty, \infty) \sqrt{V(\hat{\mu})} = \epsilon \quad (7)$$

$$V(\mu) = \frac{\epsilon^2}{t^2(\infty, \infty)} \mu^2 \quad (8)$$

Esta es la ecuación básica para determinar el tamaño de la muestra válida (aproximadamente) para cualquier modelo de muestreo.

### Intervalo de Confianza.

Si  $\hat{\mu}$  es un estimador del parámetro poblacional  $\mu$ , se define como Intervalo de confianza, asociado con el estimador  $\hat{\mu}$  a:

$$(\hat{\mu} - t(\infty, \infty) \sqrt{V(\hat{\mu})}, \hat{\mu} + t(\infty, \infty) \sqrt{V(\hat{\mu})}) \quad (9)$$

de tal manera que:

$$P(\hat{\mu} - t(\infty, \infty) \sqrt{V(\hat{\mu})} \leq \mu \leq \hat{\mu} + t(\infty, \infty) \sqrt{V(\hat{\mu})}) = 1 - 2\alpha \quad (10)$$

Enseguida se presenta una tabla, en la cual se exhiben las fórmulas para calcular las medidas estadísticas indispensables para hacer las estimaciones por muestreo, correspondientes a los modelos más usuales.

En los casos en que el marco muestral esté integrado por agrupamientos distintos de las unidades de interés, éste puede sectionarse en estratos, de tal manera que la información presente el mismo tipo de agrupamientos en cada uno, para que después se seleccione la muestra en forma independiente en cada estrato, reduciéndose el problema a determinar el número de unidades que deban incluirse.

Para cada estrato, el problema de la determinación del número de unidades que deben incluirse en él, se resuelve a través de varios criterios que se conocen como afiliación proporcional, afiliación óptima de Neyman y afiliación arbitraria.

La aplicación de estos criterios puede verse en la siguiente forma:

Si  $\mu$  es la media poblacional y  $\hat{\mu}$  su estimador, puede demostrarse que la variancia de  $\hat{\mu}$  es:

$$V(\hat{\mu}) = \sum_{t=1}^k \left( \frac{1}{n_t} - \frac{1}{N} \right) P_t^2 S_t^2 \quad (11)$$

$$S_t^2 = \frac{1}{N_t - 1} \sum_{i=1}^{N_t} (X_{ti} - \mu_t)^2, \quad P_t = \frac{N_t}{N} \quad (12)$$

en donde:

K = Número de estratos.

$N_t$  = Número total de unidades últimas en el t-ésimo estrato.

N = Número total de unidades últimas en todos los estratos.

$n_t$  = Número de unidades últimas en el t-ésimo estrato para la muestra.

n = Número de unidades de última etapa en la muestra para todos los estratos.

$X_{t_i}$  = Valor de la variable bajo estudio en la i-ésima unidad última del t-ésimo estrato.

$\mu_t$  = Valor medio de la variable bajo estudio en el t-ésimo estrato.

Sustituyendo (11) en (8) resulta:

$$\sum_{t=1}^k \left( \frac{1}{n_t} - \frac{1}{N_t} \right) P_t^2 S_t^2 = \frac{\epsilon^2 \mu^2}{f^2(\alpha, \infty)} \quad (13)$$

a) Afijación proporcional.

$$n_t = \frac{N_t}{\alpha}, \quad n = \frac{N}{\alpha} \quad (14)$$

$$\begin{aligned} \sum_{t=1}^k \left( \frac{1}{n_t} - \frac{1}{N_t} \right) P_t^2 S_t^2 &= \frac{\epsilon^2 \mu^2}{f^2(\alpha, \infty)} \\ \sum_{t=1}^k \left( \frac{1}{\frac{N_t}{\alpha}} - \frac{1}{N_t} \right) P_t^2 S_t^2 &= \frac{\epsilon^2 \mu^2}{f^2(\alpha, \infty)} \\ \frac{1}{N} \sum_{t=1}^k P_t S_t^2 &= \frac{\epsilon^2 \mu^2}{f^2(\alpha, \infty)} + \frac{1}{N} \sum_{t=1}^k P_t S_t^2 \end{aligned} \quad (15)$$

Sustituyendo en (14) resulta:

$$n_t = \frac{\epsilon^2 \mu^2}{P_t K f^2(\alpha, \infty)} + \frac{1}{N P_t K} \sum_{t=1}^k P_t S_t^2$$

lo cual aproximadamente se reduce a:

$$n_t = \frac{P_t S_t^2}{\frac{\epsilon^2 \mu^2}{K f^2(\alpha, \infty)} + \frac{S_t^2}{N}}$$

Y esto es equivalente a:

$$\left( \frac{1}{n_t} - \frac{1}{N_t} \right) S_t^2 = \frac{\epsilon^2 \mu^2}{P_t^2 K f^2(\alpha, \infty)} \quad (16)$$

b) En la afijación de Neyman, puede encontrarse que

$$\left( \frac{1}{n_t} - \frac{1}{N_t} \right) S_t^2 = \frac{\epsilon^2 \mu^2}{P_t^2 S_t^2 K f^2(\alpha, \infty)}$$

Por lo tanto, en los casos en que haya que realizar una investigación por muestreo y deban utilizarse varios modelos (uno por cada estrato), de manera que las estimaciones referidas al total de la población estén sujetas a un error de muestreo y a un nivel de confianza correspondiente a  $\epsilon$  y  $f(\alpha, \infty)$  respectivamente, el error de muestreo y el nivel de confianza en cada estrato para afijación proporcional, deben ser los

correspondientes a  $\frac{N}{N_1} (t_{(\infty, \infty)}^2)^{1/K}$ , en donde  $N_1$  es el número total de unidades últimas de muestreo en el estrato t-ésimo y  $N$  el número total de unidades últimas de muestreo en todos los estratos.

Es importante al calcular el tamaño de la muestra, tener en cuenta las condiciones de no respuesta y no cobertura.

### Consideraciones sobre el Cálculo del Tamaño de la Muestra.

Usualmente el cálculo del tamaño de la muestra se hace suponiendo que es posible conocer con gran exactitud las funciones de costo, lo que en la práctica no es del todo cierto, sobre todo cuando no se tiene una amplia experiencia sobre estas cuestiones, ya que inclusive, no se pueden transferir experiencias de una encuesta a otra con la precisión que se quisiera, excepto en las encuestas continuas sobre un mismo fenómeno. Para resolver este problema en los siguientes desarrollos se encontrarán fórmulas que proporcionen el tamaño de la muestra sin utilizar dichas funciones de costo, sustituyéndolas por alternativas sistemáticas partiendo de un nivel de confianza y precisión fijadas de antemano, para no tener que basarse en supuestos fuera de la realidad.

### Nomenclatura.

- $X$  = Variable aleatoria bajo estudio.
- $\mu$  = Media poblacional de la variable  $X$ .
- $\hat{\mu}$  = Estimador de la media poblacional.
- $K$  = Número de estratos.
- $N$  = Número total de unidades de primera etapa.
- $M$  = Número total de unidades de segunda etapa.
- $L$  = Número total de unidades de tercera etapa.
- $n$  = Número de unidades de primera etapa para la muestra.

$m$  = Número de unidades de segunda etapa para la muestra.

$l$  = Número de unidades de tercera etapa para la muestra.

$t, i, j$  = Subíndices para referirse a estratos, unidades de primera y segunda etapa respectivamente, según sea el caso.

$S_1^2, S_2^2$  = Error cuadrático medio entre unidades de primera etapa en muestreo estratificado y no estratificado, respectivamente.

$S_{11}^2, S_{12}^2$  = Error cuadrático medio entre unidades de segunda etapa en muestreo estratificado y no estratificado, respectivamente.

$s_1^2, s_2^2$  = Error cuadrático medio muestral entre unidades de primera etapa en muestreo estratificado y no estratificado, respectivamente.

$s_{11}^2, s_{12}^2$  = Error cuadrático medio muestral entre unidades de segunda etapa en muestreo estratificado y no estratificado, respectivamente.

$\epsilon | \mu$  = Error de muestreo máximo permisible.

$t_{(\infty, \infty)}$  =  $t$  de student correspondiente al nivel de confianza. 1-2 $\infty$  y con  $\infty$  grados de libertad.

$V(\hat{\mu})$  = Variancia del estimador  $\hat{\mu}$ .

Est.  $V(\hat{\mu})$  = Estimador de la variancia del estimador  $\hat{\mu}$ .

Para distinguir entre variables aleatorias y valores, se procederá como a continuación se indica. Si  $\mu$  es una variable aleatoria,  $\mu$  indicará alguno de sus valores, según sea el caso.

1.- Muestreo Monoetápico Equiprobable y sin Reemplazo.

sea

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{el estimador de la media poblacional}$$

la variancia se define por la siguiente ecuación

$$V(\bar{X}) = E(\bar{X} - E(\bar{X}))^2$$

la cual debido a

$$E(\bar{X}) = E\left\{ \frac{1}{n} \sum_{i=1}^n X_i \right\} = \frac{1}{n} E\left\{ \sum_{i=1}^n X_i \right\} = \frac{1}{n} \sum_{i=1}^n E(X_i) = E(X_1) = \mu \rightarrow$$

se reduce a

$$V(\bar{X}) = E(\bar{X} - \mu)^2$$

cuyo desarrollo es el siguiente:

$$V(\bar{X}) = E\left\{ \left( \frac{1}{n} \sum_{i=1}^n X_i \right) - \mu \right\}^2$$

$$= E\left\{ \frac{1}{n} \sum_{i=1}^n (X_i - \mu) \right\}^2$$

$$= \frac{1}{n^2} E\left\{ \sum_{i=1}^n (X_i - \mu)^2 + \sum_{i \neq j} (X_i - \mu)(X_j - \mu) \right\}$$

$$= \frac{1}{n^2} \left\{ n E(X_i - \mu)^2 + n(n-1) E((X_i - \mu)(X_j - \mu)) \right\}$$

$$= \frac{1}{n} \left\{ \sum_{i=1}^n \frac{1}{N} (X_i - \mu)^2 + (n-1) \sum_{i \neq j} \frac{1}{N} \frac{1}{N-1} (X_i - \mu)(X_j - \mu) \right\}$$

$$= \frac{1}{n} \left\{ \frac{N-1}{N} \frac{1}{N-1} \sum_{i=1}^n (X_i - \mu)^2 + \frac{n-1}{N(N-1)} \sum_{i \neq j} (X_i - \mu)(X_j - \mu) \right\}$$

$$+ \sum_{i=1}^n (X_i - \mu)^2 - \sum_{i=1}^n (X_i - \mu)^2 \left\{ \right\}$$

$$= \frac{N-1}{Nn} S^2 + \frac{n-1}{Nn(N-1)} \left\{ \left( \sum_{i=1}^n (X_i - \mu) \right)^2 - \sum_{i=1}^n (X_i - \mu)^2 \right\}$$

$$= \frac{N-1}{Nn} S^2 + \frac{n-1}{Nn} \frac{1}{N-1} \sum_{i=1}^n (X_i - \mu)^2$$

$$= \frac{N-1}{Nn} S^2 + \frac{n-1}{Nn} S^2$$

$$= \frac{N-n}{Nn} S^2$$

Estimador de la variancia de  $\hat{\mu}$

$$\text{Est. } V(\hat{\mu}) = \frac{N-n}{Nn} s^2$$

este estimador es insesgado, demostración:

$$E \{ \text{Est. } V(\hat{x}) \} = E \left\{ \frac{N-1}{Nn} s^2 \right\}$$

$$= \frac{N-n}{Nn} E \left\{ \frac{1}{n-1} \sum_{i=1}^n (\hat{x}_i - \bar{x})^2 \right\}$$

$$= \frac{N-n}{Nn(n-1)} E \left\{ \sum_{i=1}^n \{ (\hat{x}_i - \mu)^2 + 2(\hat{x}_i - \mu)(\mu - \bar{x}) + (\mu - \bar{x})^2 \} \right\}$$

$$= \frac{N-n}{Nn(n-1)} E \left\{ \sum_{i=1}^n \{ (\hat{x}_i - \mu)^2 + 2(\mu - \bar{x}) \sum_{i=1}^n (\hat{x}_i - \mu) + n(\mu - \bar{x})^2 \} \right\}$$

$$= \frac{N-n}{Nn(n-1)} E \left\{ \sum_{i=1}^n (\hat{x}_i - \mu)^2 + 2(\mu - \bar{x}) n(\bar{x} - \mu) + n(\mu - \bar{x})^2 \right\}$$

$$= \frac{N-n}{Nn(n-1)} E \left\{ \sum_{i=1}^n (\hat{x}_i - \mu)^2 - n(\mu - \bar{x})^2 \right\}$$

obteniéndose:

$$= \frac{N-n}{N(n-1)} \{ E(\hat{x}_i - \mu)^2 - E(\mu - \bar{x})^2 \}$$

lo que da:

$$= \frac{N-n}{N(n-1)} \left\{ \frac{N-1}{N} S^2 - \frac{N-n}{Nn} S^2 \right\}$$

$$= \frac{N-n}{N(n-1)} \left\{ \frac{Nn-n-N+n}{Nn} S^2 \right\}$$

$$= \frac{N-n}{N(n-1)} \frac{N(n-1)}{Nn} S^2$$

$$= \frac{N-n}{Nn} S^2$$

$$= V(\bar{x})$$

Tamaño de la muestra: tomando la ecuación (8)

$$V(\hat{\mu}) = \frac{\epsilon^2 \mu^2}{f^2 (\alpha/2, \infty)}$$

no tiene

$$\frac{N-n}{Nn} S^2 = \frac{\epsilon^2 \mu^2}{f^2 (\alpha, \infty)}$$

$$(N-n) S^2 = \frac{N \epsilon^2 \mu^2 n}{f^2(\alpha, \infty)}$$

$$n \left( \frac{N \epsilon^2 \mu^2}{f^2(\alpha, \infty)} + S^2 \right) = NS^2$$

finalmente:

$$n = \frac{NS^2 f^2(\alpha, \infty)}{N \epsilon^2 \mu^2 + S^2 f^2(\alpha, \infty)}$$

$$n = \frac{S^2 f^2(\alpha, \infty)}{\epsilon^2 \mu^2} \frac{1}{1 + \frac{S^2 f^2(\alpha, \infty)}{\epsilon^2 \mu^2}}$$

## II.- Muestreo Bietápico Equiprobable y sin Reemplazo.

Estimador de la media poblacional.

$$\hat{\mu} = \frac{1}{\sum_{i=1}^M M_i} \sum_{i=1}^M \frac{M_i}{m_i} \sum_{j=1}^{m_i} x_{ij}$$

$$= \frac{1}{\sum_{i=1}^M M_i} \sum_{i=1}^M M_i \bar{X}_i$$

Cálculo de su variancia, al ser insesgado ésta viene dada por:

$$V(\hat{\mu}) = E \left( \frac{1}{\sum_{i=1}^M M_i} \sum_{i=1}^M M_i \bar{X}_i - \mu \right)^2$$

ecuación que puede desarrollarse en la siguiente forma:

$$V(\hat{\mu}) = E \left( \frac{1}{\sum_{i=1}^M M_i} \sum_{i=1}^M M_i (\bar{X}_i - \mu) \right)^2$$

$$= \frac{1}{n^2 M^2} E \left\{ \sum_{i=1}^M [M_i (\bar{X}_i - \mu)]^2 + \sum_{i \neq j} M_i M_j (\bar{X}_i - \mu) (\bar{X}_j - \mu) \right\}$$

$$= \frac{1}{n^2 M^2} E \left\{ \sum_{i=1}^M M_i^2 [(\bar{X}_i - \mu)]^2 + (\mu_i - \mu)^2 + \sum_{i \neq j} M_i M_j (\bar{X}_i - \mu) (\bar{X}_j - \mu) \right\}$$

$$= \frac{1}{n^2 M^2} E \left\{ \sum_{i=1}^M M_i^2 [(\bar{X}_i - \mu_i)^2 + 2(\bar{X}_i - \mu_i)(\mu_i - \mu) + (\mu_i - \mu)^2] + \right.$$

$$\left. + \sum_{i \neq j} M_i M_j (\bar{X}_i - \mu) (\bar{X}_j - \mu) \right\} +$$

$$+ \frac{1}{n^2 M^2} \{ n E [M_i^2 [(\frac{1}{m_i} - \frac{1}{M_i}) S_i^2 + (\mu_i - \mu)^2]] + n(n-1) E [M_i M_j (\mu_i - \mu) (\mu_j - \mu)] \}$$

$$= \frac{1}{n^2 M^2} \{ n \sum_{i=1}^M M_i^2 [(\frac{1}{m_i} - \frac{1}{M_i}) S_i^2 + (\mu_i - \mu)^2] + n(n-1) \sum_{i \neq j} \frac{M_i M_j}{N(N-1)} (\mu_i - \mu) (\mu_j - \mu) \}$$

$$= \frac{1}{n^2 M^2} \sum_{i=1}^M \frac{M_i^2}{M_i} \left( \frac{1}{m_i} - \frac{1}{M_i} \right) S_i^2 + \frac{1}{nN} \sum_{i=1}^M \frac{M_i^2}{M_i^2} (\mu_i - \mu)^2 + \dots$$

$$+ \frac{n(n-1)}{n^2 M^2 (N-1)N} \sum_{i \neq j} M_i M_j (\mu_i - \mu) (\mu_j - \mu)$$

$$+ \sum_{i=1}^M \frac{M_i^2}{n^2 M^2} (\mu_i - \mu)^2 - \sum_{i=1}^M \frac{M_i^2}{n^2 M^2} (\mu_i - \mu)^2 \{$$

$$= \frac{1}{n} \sum_{i=1}^N \frac{M_i^2}{M^2} \left( \frac{1}{m_i} - \frac{1}{M} \right) S_i^2 + \frac{N-1}{nN} S^2 + \frac{(n-1)}{nM^2 N(N-1)} \left\{ \sum_{i=1}^N M_i (\mu_i - \mu)^2 - \sum_{i=1}^N M_i^2 (\mu_i - \mu)^2 \right\}$$

$$= \frac{1}{Nn} \sum_{i=1}^N \frac{M_i}{M} \left( \frac{1}{m_i} - \frac{1}{M} \right) S_i^2 + \frac{N-1}{Nn} S^2 - \frac{n-1}{Nn} \left[ \frac{1}{N-1} \sum_{i=1}^N \frac{M_i}{M} (\mu_i - \mu)^2 \right]$$

$$= \frac{1}{Nn} \sum_{i=1}^N \left( \frac{M_i}{M} \right)^2 \left( \frac{1}{m_i} - \frac{1}{M} \right) S_i^2 + \frac{N-n}{Nn} S^2$$

$$= \left( \frac{1}{n} - \frac{1}{N} \right) S^2 + \frac{1}{Nn} \sum_{i=1}^N \left( \frac{M_i}{M} \right)^2 \left( \frac{1}{m_i} - \frac{1}{M} \right) S_i^2$$

Cálculo del tamaño de la muestra:

$$V(\hat{\mu}) = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2}$$

$$\left( \frac{1}{n} - \frac{1}{N} \right) S^2 + \frac{1}{Nn} \sum_{i=1}^N \left( \frac{M_i}{M} \right)^2 \left( \frac{1}{m_i} - \frac{1}{M} \right) S_i^2 = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2}$$

sea  $m_i = \frac{M_i}{\infty} \rightarrow$  (colocación proporcional)

$$\left( \frac{1}{n} - \frac{1}{N} \right) S^2 + \frac{1}{Nn} \sum_{i=1}^N \left( \frac{M_i}{M} \right)^2 \left( \frac{\infty-1}{M_i} \right) S_i^2 = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2}$$

$$\frac{1}{n} \left\{ S^2 + \frac{\infty-1}{NM^2} \sum_{i=1}^N M_i S_i^2 \right\} = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2} + \frac{S^2}{N}$$

$$\frac{1}{n} \left\{ S^2 + \frac{\infty-1}{NM^2} \sum_{i=1}^N M_i S_i^2 \right\} = \frac{\epsilon^2 \mu^2 N + S^2 t_{(\alpha, \infty)}^2}{N t_{(\alpha, \infty)}^2}$$

$$n = \frac{N t_{(\alpha, \infty)}^2}{\epsilon^2 \mu^2 N + S^2 t_{(\alpha, \infty)}^2} \left\{ S^2 + \frac{\infty-1}{NM^2} \sum_{i=1}^N M_i S_i^2 \right\}$$

definiendo por:

$$a = \frac{N t_{(\alpha, \infty)}^2}{\epsilon^2 \mu^2 N + S^2 t_{(\alpha, \infty)}^2}, \quad b = \frac{1}{NM^2} \sum_{i=1}^N M_i S_i^2$$

Y puesto que:

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N \left( \frac{M_i}{M} \right)^2 (\mu_i - \mu)^2, \quad S_i^2 = \frac{1}{M_i - 1} \sum_{j=1}^{M_i} (X_{ij} - \mu_i)^2$$

$$\mu_i = \frac{1}{M_i} \sum_{j=1}^{M_i} X_{ij}, \quad \mu = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{M_i} X_{ij}$$

la fórmula para n se reduce a:

$$n = a (S^2 + b (\alpha - 1))$$

el número de unidades de segunda etapa (cuestionarios) aproximadamente está dado por:

$$m_0 = n \frac{\bar{M}}{\alpha} = \frac{1}{\alpha} a \bar{M} (S^2 + b) + a b \bar{M}$$

el cual toma sus valores máximo y mínimo en los extremos del siguiente intervalo

$$1 \leq \alpha \leq \left( \frac{N}{ab} + 1 - \frac{S^2}{b} \right)$$

el valor más apropiado de  $\alpha$  debe tomarse considerando el tiempo y costo disponibles para la realización de la encuesta.

### III.- Muestreo Monoetápico, Equiprobable, sin Reemplazo y Estratificado.

El estimador de la media poblacional viene dado por:

$$\hat{\mu} = \frac{1}{N_0} \sum_{t=1}^K \frac{N_t}{n_t} \sum_{i=1}^{n_t} \dot{X}_{ti} \quad , \quad N_0 = \sum_{t=1}^K N_t$$

$$= \frac{1}{N_0} \sum_{t=1}^K N_t \dot{X}_t \quad , \quad \dot{X}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} \dot{X}_{ti}$$

al ser insesgado, su variancia es:

$$V(\hat{\mu}) = E \left\{ \frac{1}{N_0} \sum_{t=1}^K N_t \dot{X}_t - \mu \right\}^2$$

$$V(\hat{\mu}) = E \left\{ \frac{1}{N_0} \sum_{t=1}^K N_t (\dot{X}_t - \mu)^2 \right\}$$

$$= \frac{1}{N_0^2} E \left\{ \sum_{t=1}^K N_t^2 (\dot{X}_t - \mu)^2 + \sum_{t \neq t'}^K N_t N_{t'} (\dot{X}_t - \mu) (\dot{X}_{t'} - \mu) \right\}$$

$$= \frac{1}{N_0^2} \left\{ \sum_{t=1}^K N_t^2 E (\dot{X}_t - \mu)^2 + \sum_{t \neq t'}^K N_t N_{t'} E (\dot{X}_t - \mu) (\dot{X}_{t'} - \mu) \right\}$$

$$= \frac{1}{N_0^2} \sum_{t=1}^K N_t^2 \frac{N_t - n_t}{N_t n_t} S_t^2 + 0$$

$$= \sum_{t=1}^K \left( \frac{N_t}{N_0} \right)^2 \left( \frac{1}{n_t} - \frac{1}{N_t} \right) S_t^2$$

$$\text{En donde } S_t^2 = \frac{1}{N_t - 1} \sum_{i=1}^{N_t} (X_{ti} - \mu_t)^2$$

Estimador de la variancia de  $\hat{\mu}$

$$\text{Est. } V(\hat{\mu}) = \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 s_t^2$$

$$s_t^2 = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} (\dot{X}_{ti} - \dot{X}_t)^2 \quad , \quad \dot{X}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} \dot{X}_{ti}$$

que es insesgado, demostración:

$$\begin{aligned}
 E \{ \text{Est. } V(\hat{\mu}) \} &= E \left\{ \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 \frac{1}{n_t - 1} \sum_{i=1}^{n_t} (\dot{X}_{ti} - \dot{\bar{X}}_t)^2 \right\} \\
 &= E \left\{ \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 \frac{1}{n_t - 1} \sum_{i=1}^{n_t} (\dot{X}_{ti} - \dot{\mu}_t + \dot{\mu}_t - \dot{\bar{X}}_t)^2 \right\} \\
 &= E \left\{ \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 \frac{1}{n_t - 1} \sum_{i=1}^{n_t} [(\dot{X}_{ti} - \dot{\mu}_t)^2 + 2(\dot{X}_{ti} - \dot{\mu}_t)(\dot{\mu}_t - \dot{\bar{X}}_t) + (\dot{\mu}_t - \dot{\bar{X}}_t)^2] \right\} \\
 &= E \left\{ \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 \frac{1}{n_t - 1} \left[ \sum_{i=1}^{n_t} (\dot{X}_{ti} - \dot{\mu}_t)^2 \right] \right. \\
 &\quad \left. + 2(\dot{\mu}_t - \dot{\bar{X}}_t) \sum_{i=1}^{n_t} (\dot{X}_{ti} - \dot{\mu}_t) + n_t (\dot{\bar{X}}_t - \dot{\mu}_t)^2 \right\} \\
 &= E \left\{ \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 \frac{1}{n_t - 1} \left[ \sum_{i=1}^{n_t} (\dot{X}_{ti} - \dot{\mu}_t)^2 + 2(\dot{\mu}_t - \dot{\bar{X}}_t) n_t (\dot{\bar{X}}_t - \dot{\mu}_t) + n_t (\dot{\bar{X}}_t - \dot{\mu}_t)^2 \right] \right\} \\
 &= E \left\{ \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 \frac{1}{n_t - 1} \left[ \sum_{i=1}^{n_t} (\dot{X}_{ti} - \dot{\mu}_t)^2 - n_t (\dot{\bar{X}}_t - \dot{\mu}_t)^2 \right] \right\} \\
 &= \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 \frac{1}{n_t - 1} \left[ n_t (\dot{\bar{X}}_t - \dot{\mu}_t)^2 - n_t (\dot{\bar{X}}_t - \dot{\mu}_t)^2 \right] S_t^2 \\
 &= \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 \frac{1}{n_t - 1} \left[ n_t (\dot{\bar{X}}_t - \dot{\mu}_t)^2 - n_t (\dot{\bar{X}}_t - \dot{\mu}_t)^2 \right] S_t^2 \\
 &= \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 \frac{1}{n_t - 1} \left[ \frac{n_t N_t - n_t - N_t + n_t}{N_t n_t} S_t^2 \right]
 \end{aligned}$$

$$\begin{aligned}
 E \{ \text{Est. } V(\hat{\mu}) \} &= \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 \frac{n_t}{n_t - 1} \frac{N_t (n_t - 1)}{N_t n_t} S_t^2 \\
 &= \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 S_t^2 \\
 &= V(\hat{\mu})
 \end{aligned}$$

Tamaño de la muestra, de la ecuación (8)

$$\begin{aligned}
 V(\hat{\mu}) &= \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2} \\
 \sum_{t=1}^K \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 S_t^2 &= \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2} \\
 \left( \frac{1}{n_t} - \frac{1}{N_t} \right) \left( \frac{N_t}{N_0} \right)^2 S_t^2 &= \frac{\epsilon^2 \mu^2}{K t_{(\alpha, \infty)}^2} \\
 \frac{1}{n_t} \left( \frac{N_t}{N_0} \right)^2 S_t^2 &= \frac{\epsilon^2 \mu^2}{K t_{(\alpha, \infty)}^2} + \frac{1}{N_t} \left( \frac{N_t}{N_0} \right)^2 S_t^2 \\
 n_t &= \frac{\left( \frac{N_t}{N_0} \right)^2 S_t^2 K N_t t_{(\alpha, \infty)}^2}{N_t \epsilon^2 \mu^2 + K \left( \frac{N_t}{N_0} \right)^2 S_t^2 t_{(\alpha, \infty)}^2}
 \end{aligned}$$

$$\frac{S_t^2 f_{(\alpha, \infty)}^2}{\frac{N_o}{N_t} \frac{\epsilon^2}{K} \mu^2}$$

$$n_t = \frac{1}{1 + \frac{S_t^2 f_{(\alpha, \infty)}^2}{N_o \frac{\epsilon^2}{K} \mu^2}}$$

IV.- Muestreo Bietápico, Equiprobable, sin Reemplazo y Estratificado.

Estimador de la media poblacional:

$$\hat{\mu} = \frac{1}{M_o} \sum_{t=1}^K \frac{M_{ot}}{n_t M_t} \sum_{i=1}^{n_t} M_{ti} \dot{X}_{ti}$$

$$M_o = \sum_{t=1}^K M_{ot}, M_{ot} = \sum_{i=1}^{n_t} M_{ti}, \dot{M}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} M_{ti}, \dot{X}_{ti} =$$

$$= \frac{1}{m_{ti}} \sum_{j=1}^{m_{ti}} \dot{X}_{tij}, \dot{X}_t = \frac{1}{n_t \bar{M}_t} \sum_{i=1}^{n_t} M_{ti} \dot{X}_{ti}$$

Variancia del estimador  $\hat{\mu}$

$$V(\hat{\mu}) = E \left\{ \frac{1}{M_o} \sum_{t=1}^K \frac{M_{ot}}{n_t M_t} \sum_{i=1}^{n_t} M_{ti} \dot{X}_{ti} - \mu \right\}^2$$

$$= E \left\{ \frac{1}{M_o} \sum_{t=1}^K M_{ot} \dot{X}_t - \frac{1}{M_o} \sum_{t=1}^K M_{ot} \mu_t \right\}^2$$

$$V(\hat{\mu}) = E \left| \frac{1}{M_o} \sum_{t=1}^K M_{ot} (\dot{X}_t - \mu_t) \right|^2$$

$$= E \left| \sum_{t=1}^K \left( \frac{M_{ot}}{M_o} \right)^2 (\dot{X}_t - \mu_t)^2 + \sum_{t \neq t'}^K \left( \frac{M_{ot}}{M_o} \right) \left( \frac{M_{ot'}}{M_o} \right) (\dot{X}_t - \mu_t) (\dot{X}_{t'} - \mu_{t'}) \right|$$

$$= E \left| \sum_{t=1}^K \left( \frac{M_{ot}}{M_o} \right)^2 (\dot{X}_t - \mu_t)^2 \right|$$

$$= \sum_{t=1}^K \left( \frac{M_{ot}}{M_o} \right)^2 E \left\{ \frac{1}{n_t \bar{M}_t} \sum_{i=1}^{n_t} M_{ti} \dot{X}_{ti} - \mu_t \right\}^2$$

$$= \sum_{t=1}^K \left( \frac{M_{ot}}{M_o} \right)^2 E \left\{ \frac{1}{n_t \bar{M}_t} \sum_{i=1}^{n_t} M_{ti} (\dot{X}_{ti} - \mu_t) \right\}^2$$

$$= \sum_{t=1}^K \left( \frac{M_{ot}}{M_o} \right)^2 E \left\{ \frac{1}{n_t \bar{M}_t} \left\{ \sum_{i=1}^{n_t} M_{ti}^2 (\dot{X}_{ti} - \mu_t)^2 + \right. \right.$$

$$\left. \left. + \sum_{i \neq i'}^{n_t} M_{ti} M_{ti'} (\dot{X}_{ti} - \mu_t) (\dot{X}_{ti'} - \mu_t) \right\} \right\}$$

$$= \sum_{t=1}^K \left( \frac{M_{ot}}{M_o} \right)^2 \left( \frac{1}{n_t \bar{M}_t} \right)^2 E \left\{ \sum_{i=1}^{n_t} M_{ti}^2 (\dot{X}_{ti} - \mu_t)^2 + \mu_{ti}^2 - \mu_t^2 \right\} +$$

$$+ \frac{n_t (n_t - 1)}{n_t (n_t - 1)} \sum_{i \neq i'}^{n_t} M_{ti} M_{ti'} (\mu_{ti} - \mu_t) (\mu_{ti'} - \mu_t) \left\} \right.$$

$$\begin{aligned}
&= \sum_{t=1}^K \left( \frac{M_{0t}}{M_0} \right)^2 \left( \frac{1}{n_t} - \frac{1}{N_t} \right)^2 \left\{ E \sum_{i=1}^{N_t} M_{it}^2 \left\{ (\bar{X}_{it} - \bar{\mu}_{it})^2 - \right. \right. \\
&\quad \left. \left. + 2(\bar{X}_{it} - \bar{\mu}_{it})(\mu_{it} - \mu_t) + (\mu_{it} - \mu_t)^2 \right\} + \right. \\
&\quad \left. + \frac{n_t(n_{t-1})}{N_t(N_{t-1})} \sum_{i \neq j}^{N_t} M_{it} M_{jt} (\mu_{it} - \mu_t) (\mu_{jt} - \mu_t) \right\} \\
&= \sum_{t=1}^K \left( \frac{M_{0t}}{M_0} \right)^2 \left( \frac{1}{n_t} - \frac{1}{N_t} \right)^2 \left\{ \frac{n_t}{N_t} \sum_{i=1}^{N_t} M_{it}^2 \left( \frac{M_{it} - m_{it}}{M_{it} m_{it}} \right) S_{it}^2 + \right. \\
&\quad \left. + n_t \sum_{i=1}^{N_t} \frac{1}{N_t} M_{it}^2 (\mu_{it} - \mu_t)^2 + \right. \\
&\quad \left. + \frac{n_t(n_{t-1})}{N_t(N_{t-1})} \sum_{i \neq j}^{N_t} M_{it} M_{jt} (\mu_{it} - \mu_t) (\mu_{jt} - \mu_t) \right\}
\end{aligned}$$

Tamaño de la muestra:  
de la ecuación (8) se tiene

$$V(\hat{\mu}) = \frac{\epsilon^2 \mu^2}{f^2 (\infty, \infty)}$$

$$\begin{aligned}
&\sum_{t=1}^K \left( \frac{M_{0t}}{M_0} \right)^2 \left\{ \left( \frac{1}{n_t} - \frac{1}{N_t} \right) S_t^2 + \frac{1}{N_t n_t} \sum_{i=1}^{N_t} \frac{M_{it}^2}{M_{it}^2} \left( \frac{1}{m_{it}} - \frac{1}{M_{it}} \right) S_{it}^2 \right\} = \frac{\epsilon^2 \mu^2}{f^2 (\infty, \infty)} \\
&\left( \frac{M_{0t}}{M_0} \right)^2 \left\{ \left( \frac{1}{n_t} - \frac{1}{N_t} \right) S_t^2 + \frac{1}{N_t n_t} \sum_{i=1}^{N_t} \frac{M_{it}^2}{M_{it}^2} \left( \frac{1}{m_{it}} - \frac{1}{M_{it}} \right) S_{it}^2 \right\} = \frac{\epsilon^2 \mu^2}{K f^2 (\infty, \infty)} \\
&\frac{1}{n_t} \left\{ S_t^2 + \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{M_{it}^2}{M_{it}^2} \left( \frac{1}{m_{it}} - \frac{1}{M_{it}} \right) S_{it}^2 \right\} = \frac{M_0 \epsilon^2 \mu^2}{K M_{0t}^2 f^2 (\infty, \infty)} + \frac{S_t^2}{N_t}
\end{aligned}$$

$$\begin{aligned}
&n_t \left\{ \frac{K N_t M_{0t}^2 f^2}{N_t M_0 \epsilon^2 \mu^2 + K M_{0t}^2 S_t^2 f^2 (\infty, \infty)} \right\} S_t^2 + \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{M_{it}^2}{M_{it}^2} \left( \frac{1}{m_{it}} - \frac{1}{M_{it}} \right) S_{it}^2 \\
&\quad \text{sea } m_{it} = \frac{1}{\alpha_t} M_{it}, \alpha_t \geq 1
\end{aligned}$$

$$\frac{K N_t M_{0t}^2 f^2 (\infty, \infty)}{N_t M_0 \epsilon^2 \mu^2 + K M_{0t}^2 S_t^2 f^2 (\infty, \infty)} \left\{ S_t^2 + \frac{(\alpha_t - 1)}{N_t} \sum_{i=1}^{N_t} M_{it} S_{it}^2 \right\}$$

$$\sum_{i=1}^K \left( \frac{M_{0t}}{M_0} \right)^2 \frac{1}{n_t} \left\{ \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{it}}{M_{it}} \right)^2 \left( \frac{M_{it} - m_{it}}{M_{it} m_{it}} \right) S_{it}^2 + \right.$$

$$\left. + \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{it}}{M_{it}} \right)^2 (\mu_{it} - \mu_t)^2 + \frac{n_{t-1}}{N_t(N_{t-1})} \sum_{i=1}^{N_t} \frac{M_{it}}{M_{it}} (\mu_{it} - \mu_t) \right\}^2 -$$

$$\sum_{i=1}^{N_t} \left( \frac{M_{it}}{M_{it}} \right)^2 (\mu_{it} - \mu_t)^2 \left\{ \right.$$

$$\left. \sum_{i=1}^K \left( \frac{M_{0t}}{M_0} \right)^2 \frac{1}{n_t} \left\{ \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{it}}{M_{it}} \right)^2 \left( \frac{M_{it} - m_{it}}{M_{it} m_{it}} \right) S_{it}^2 + \right.$$

$$\left. + \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{it}}{M_{it}} \right)^2 (\mu_{it} - \mu_t)^2 \right\}$$

$$\sum_{i=1}^K \left( \frac{M_{0t}}{M_0} \right)^2 \frac{1}{n_t} \left\{ \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{it}}{M_{it}} \right)^2 \left( \frac{M_{it} - m_{it}}{M_{it} m_{it}} \right) S_{it}^2 + \frac{N_{t-1} - n_{t-1}}{N_t(N_{t-1})} \sum_{i=1}^{N_t} \left( \frac{M_{it}}{M_{it}} \right)^2 (\mu_{it} - \mu_t)^2 \right\}$$

$$= \sum_{t=1}^K \left( \frac{M_{ot}}{M_o} \right)^2 \frac{1}{n_t} \left\{ \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{ti}}{M_{ti}} \right)^2 \left( \frac{M_{ti}}{M_{ti}} - \frac{M_{ti}}{M_{ti}} \right) S_{ti}^2 + \frac{N_t - n_t}{N_t} S_t^2 \right\}$$

$$= \sum_{t=1}^K \left( \frac{M_{ot}}{M_o} \right)^2 \left\{ \left( \frac{1}{n_t} - \frac{1}{N_t} \right) S_{ti}^2 + \frac{1}{N_t n_t} \sum_{i=1}^{N_t} \left( \frac{M_{ti}}{M_{ti}} \right)^2 \left( \frac{1}{m_{ti}} - \frac{1}{M_{ti}} \right) S_{ti}^2 \right\}$$

$$S_{ti}^2 = \frac{1}{N_t - 1} \sum_{i=1}^{N_t} \left( \frac{M_{ti}}{M_{ti}} \right)^2 (\mu_{ti} - \mu_t)^2, \quad S_{ti}^2 = \frac{1}{M_{ti} - 1} \sum_{j=1}^{M_{ti}} (X_{tij} - \mu_{ti})^2, \quad \bar{M} = \frac{M_{ot}}{N_t}$$

$$a_t = \frac{K N_t M_{ot}^2 t^{2(\alpha_t, \infty)}}{N_t M_o \epsilon^2 \mu^2 + K M_{ot}^2 S_t^2 t^{2(\alpha_t, \infty)}}, \quad b_t = \frac{1}{N_t M_{ti}^2} \sum_{i=1}^{N_t} M_{ti} S_{ti}^2$$

$$S_{ti}^2 = \frac{1}{M_{ti} - 1} \sum_{j=1}^{M_{ti}} (X_{tij} - \mu_{ti})^2, \quad S_t^2 = \frac{1}{N_t - 1} \sum_{i=1}^{N_t} \left( \frac{M_{ti}}{M_{ti}} \right)^2 (\mu_{ti} - \mu_t)^2, \quad M_{ot} = \sum_{i=1}^{N_t} M_{ti}$$

$$\mu_{ti} = \frac{1}{M_{ti}} \sum_{j=1}^{M_{ti}} X_{tij}, \quad \mu_t = \frac{1}{M_{ot}} \sum_{i=1}^{N_t} M_{ti} \mu_{ti}, \quad \bar{M}_t = \frac{M_{ot}}{N_t}, \quad M_o = \sum_{t=1}^K M_{ot}$$

$$n_t = a_t \{ S_t^2 + b_t (\alpha_t - 1) \}$$

$$m_{ot} = \frac{n \bar{M}_t}{\alpha_t} = \frac{n \bar{M}_t a_t S_t^2 - \bar{M}_t a_t b_t}{\alpha_t} + \bar{M}_t a_t b_t$$

$$1 \leq \alpha_t \leq \frac{N_t}{a_t b_t} + 1 - \frac{S_t^2}{b_t}$$

V. Muestreo Monoetápico, Equiprobable y con Reemplazo.  
Estimador de la media poblacional.

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Cálculo de la variancia de  $\hat{\mu}$

$$V(\hat{\mu}) = E(\hat{\mu} - \mu)^2$$

$$= E\left(\frac{1}{n} \sum_{i=1}^n \dot{X}_i - \mu\right)^2$$

$$= E\left(\frac{1}{n} \sum_{i=1}^n \dot{X}_i - \mu\right)^2$$

$$= \frac{1}{n^2} E\left\{ \sum_{i=1}^n (\dot{X}_i - \mu)^2 + \sum_{i \neq j} \sum_{j=1}^n (\dot{X}_i - \mu)(\dot{X}_j - \mu) \right\}$$

$$= \frac{1}{n^2} \{ n E(\dot{X}_1 - \mu)^2 + n(n-1) E[(\dot{X}_1 - \mu)(\dot{X}_j - \mu)] \}$$

$$= \frac{1}{n^2} \left\{ \frac{n(N-1)}{N} S^2 + n(n-1) \sum_{i \neq j} \frac{1}{N^2} (X_i - \mu)(X_j - \mu) \right\}$$

$$= \frac{N-1}{Nn} S^2 + \frac{n-1}{n} \frac{1}{N^2} \left( \sum_{i=1}^N (X_i - \mu) \right) \left( \sum_{i=1}^N (X_i - \mu) \right)$$

$$= \frac{N-1}{Nn} S^2$$

Estimador de la variancia de  $\hat{\mu}$

$$\text{Est. } V(\hat{\mu}) = \frac{1}{n} s^2$$

el cual es insesgado, demostración

$$E \{ \text{Est. } V(\hat{\mu}) \} = E \left\{ \frac{1}{n} s^2 \right\}$$

$$= E \left\{ \frac{1}{n} \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2 \right\}$$

$$= \frac{1}{(n-1)n} E \sum_{i=1}^n ((X_i - \mu) + (\mu - \bar{X}))^2$$

$$= \frac{1}{(n-1)n} \left\{ E \sum_{i=1}^n (X_i - \mu)^2 - 2 \sum_{i=1}^n (X_i - \mu)(\bar{X} - \mu) + (X - \mu)^2 \right\}$$

$$= \frac{1}{(n-1)n} \left\{ E \sum_{i=1}^n (X_i - \mu)^2 - 2E[(\bar{X} - \mu) \sum_{i=1}^n (X_i - \mu)] + nE(\mu - \bar{X})^2 \right\}$$

$$= \frac{1}{(n-1)n} \left\{ nE(X_i - \mu)^2 - nE(\mu - \bar{X})^2 \right\}$$

$$= \frac{1}{n-1} \left\{ \frac{N-1}{N} S^2 - \frac{N-1}{Nn} S^2 \right\}$$

$$= \frac{1}{n-1} \frac{N-1}{N} S^2 \left( 1 - \frac{1}{n} \right)$$

$$= \frac{N-1}{Nn} S^2$$

Tamaño de la muestra dados el nivel de confianza y la presión:

$$V(\hat{\mu}) = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2} = \frac{N-1}{Nn} S^2 = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2} \quad n = \frac{(N-1) S^2 t_{(\alpha, \infty)}^2}{N \epsilon^2 \mu^2}$$

VI.- Muestreo Bietápico, Equiprobable y con Reemplazo.

Estimador de la media poblacional:

$$\hat{\mu} = \frac{N}{nM_o} \sum_{i=1}^n \frac{M_i}{m_i} \sum_{j=1}^{m_i} X_{ij}$$

$$= \frac{N}{nM_o} \sum_{i=1}^n M_i \bar{X}_i$$

$$M_o = \sum_{i=1}^N M_i, \quad \bar{X}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} X_{ij}$$

Variación del estimador  $\hat{\mu}$

$$V(\hat{\mu}) = E \left\{ \left( \frac{N}{n M_o} \sum_{i=1}^N M_i \bar{X}_i \right) - \mu \right\}^2$$

$$= E \left\{ \frac{1}{n M_o} \sum_{i=1}^N (N M_i \bar{X}_i - M_o \mu) \right\}^2, \quad \bar{M} = \frac{M_o}{N}$$

$$= \frac{1}{n^2 M_o^2} E \left\{ \sum_{i=1}^N (N M_i \bar{X}_i - M_o \mu)^2 + \sum_{i \neq j} (N M_i \bar{X}_i - M_o \mu)(N M_j \bar{X}_j - M_o \mu) \right\}$$

$$= \frac{1}{n^2 M_o^2} E \left\{ \sum_{i=1}^N ((N M_i \bar{X}_i - N M_i \mu_i)^2 + (N M_i \mu_i - M_o \mu)^2) \right\}$$

$$+ \sum_{i \neq j} (N M_i \bar{X}_i - M_o \mu)(N M_j \bar{X}_j - M_o \mu)$$

$$= \frac{1}{n^2 M_o^2} E \left\{ n \left( \frac{1}{\sum_{j=1}^{M_i} M_i m_j} \sum_{j=1}^{M_i} (N M_i X_{ij} - N M_i \mu_i)^2 + (N M_i \mu_i - M_o \mu)^2 \right) + \right.$$

$$\left. + n(n-1) (N M_i \bar{X}_i - M_o \mu)(N M_j \bar{X}_j - M_o \mu) \right\}$$

$$= \frac{1}{n M_o^2} \left\{ \frac{1}{N} \sum_{i=1}^N \frac{1}{M_i m_i} \sum_{j=1}^{M_i} (N M_i X_{ij} - N M_i \mu_i)^2 + \frac{1}{N} \sum_{i=1}^N (N M_i \mu_i - M_o \mu)^2 + O \right\}$$

$$= \frac{1}{n M_o^2} \sum_{i=1}^N (N M_i \mu_i - M_o \mu)^2 + \frac{1}{n M_o^2} \sum_{i=1}^N \frac{1}{M_i m_i} \sum_{j=1}^{M_i} (N M_i X_{ij} - N M_i \mu_i)^2$$

por lo cual se obtiene:

$$= \frac{N-1}{n M_o^2} S^2 + \frac{1}{n M_o^2} \sum_{i=1}^N \frac{M_i - 1}{M_i m_i} S_i^2$$

en donde:

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (N M_i \mu_i - M_o \mu)^2$$

$$S_i^2 = \frac{1}{M_i - 1} \sum_{j=1}^{M_i} (N M_i X_{ij} - N M_i \mu_i)^2$$

Estimador de la variancia de  $\hat{\mu}$

$$\text{Est. } V(\hat{\mu}) = \frac{s^2}{n M_o^2}$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^N \left\{ \frac{1}{m_i} \sum_{j=1}^{m_i} N M_i X_{ij} \right\} - \left( \frac{1}{n} \sum_{i=1}^N \frac{1}{m_i} \sum_{j=1}^{m_i} N M_i X_{ij} \right)^2$$

$$= \frac{1}{n-1} \sum_{i=1}^N \left\{ N M_i \bar{X}_i - \frac{N}{n} \sum_{i=1}^N M_i \bar{X}_i \right\}^2, \quad \bar{X}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} X_{ij}$$

Este estimador es insesgado, demostración:

$$E \{ \text{Est. } V(\hat{\mu}) \} = E \left\{ \frac{1}{NM_0^2} \frac{1}{n-1} \sum_{i=1}^N \{ NM_i \bar{X}_i - \frac{N}{n} \sum_{i=1}^N M_i \bar{X}_i \}^2 \right\}$$

$$= \frac{1}{nM_0^2} \frac{1}{n-1} E \sum_{i=1}^N \{ (NM_i \bar{X}_i)^2 - \frac{2}{n} NM_i \bar{X}_i \sum_{i=1}^N M_i \bar{X}_i + \left( \frac{1}{n} \sum_{i=1}^N NM_i \bar{X}_i \right)^2 \}$$

$$= \frac{1}{nM_0^2} \frac{1}{n-1} E \left\{ \sum_{i=1}^N (NM_i \bar{X}_i)^2 - \frac{2}{n} \left( \sum_{i=1}^N NM_i \bar{X}_i \right)^2 + \frac{1}{n} \left( \sum_{i=1}^N NM_i \bar{X}_i \right)^2 \right\}$$

$$= \frac{1}{nM_0^2(n-1)} E \left\{ \sum_{i=1}^N (NM_i \bar{X}_i)^2 - \frac{1}{n} \left( \sum_{i=1}^N NM_i \bar{X}_i \right)^2 \right\}$$

$$= \frac{1}{nM_0^2(n-1)} E \left\{ \sum_{i=1}^N (NM_i \bar{X}_i)^2 - \frac{1}{n} \left\{ \sum_{i=1}^N (NM_i \bar{X}_i)^2 + \sum_{i \neq j} (NM_i \bar{X}_i)(NM_j \bar{X}_j) \right\} \right\}$$

$$= \frac{1}{nM_0^2(n-1)} E \left\{ \left(1 - \frac{1}{n}\right) \sum_{i=1}^N (NM_i \bar{X}_i)^2 + \right.$$

$$\left. \sum_{i \neq j} (NM_i \bar{X}_i)(NM_j \bar{X}_j) - \sum_{i \neq j} (NM_i \bar{X}_i)(NM_j \bar{X}_j) \right\}$$

$$= \frac{1}{nM_0^2(n-1)} E \left\{ \frac{n-1}{n} \left( \sum_{i=1}^N NM_i \bar{X}_i \right)^2 - \sum_{i \neq j} (NM_i \bar{X}_i)(NM_j \bar{X}_j) \right\}$$

$$= E \left\{ \left( \frac{1}{nM_0} \sum_{i=1}^N NM_i \bar{X}_i \right)^2 - \frac{1}{M_0^2 n(n-1)} \sum_{i \neq j} (NM_i \bar{X}_i)(NM_j \bar{X}_j) \right\}$$

$$= E \left\{ (\hat{\mu} - \mu)^2 + 2 \mu \hat{\mu} - \mu^2 - \frac{1}{M_0^2 n(n-1)} E (NM_i \bar{X}_i)(NM_j \bar{X}_j) \right\}$$

$$= E \{ (\hat{\mu} - \mu)^2 + 2 \mu \hat{\mu} - \mu^2 - \frac{1}{M_0^2} \left( \frac{1}{N} \sum_{i=1}^N NM_i \bar{X}_i \right)^2 \}$$

$$= V(\hat{\mu}) + 2 \mu E \hat{\mu} - \mu^2 - \mu^2$$

$$= V(\hat{\mu})$$

Tamaño de la muestra dados el nivel de confianza y la precisión:

$$V(\hat{\mu}) = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2}$$

$$\frac{N-1}{nM_0^2} S^2 + \frac{1}{nM_0^2} \sum_{i=1}^N \frac{M_i - 1}{M_i m_i} S_i^2 = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2}$$

Sea  $m_i = \frac{M_i}{\alpha}$ , con  $\alpha \geq 1 \rightarrow$

$$\frac{1}{nM_0^2} \{ (N-1) S^2 + \sum_{i=1}^N \frac{(M_i - 1) \alpha}{M_i^2} S_i^2 \} = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2}$$

$$n = \frac{t_{(\alpha, \infty)}^2}{\epsilon^2 \mu^2 N M_0^2} \left\{ (N-1) S^2 + \infty \sum_{i=1}^N \frac{M_i - 1}{M_i^2} S_i^2 \right\}$$

$$a = \frac{t_{(\alpha, \infty)}^2}{\epsilon^2 \mu^2 N M_0^2}, \quad b = \sum_{i=1}^N \frac{M_i - 1}{M_i} S_i^2$$

$$S^2 = \frac{1}{N-1} \sum_{i=1}^N (N M_i \mu_i - M_0 \mu)^2, \quad S_i^2 = \frac{1}{M_i - 1} \sum_{j=1}^{M_i} (N M_i X_{ij} - N M_i \mu_i)^2$$

$$\mu_i = \frac{1}{M_i} \sum_{j=1}^{M_i} X_{ij}, \quad \mu = \frac{1}{M_0} \sum_{i=1}^N \sum_{j=1}^{M_i} X_{ij}$$

$$n = a \{ (N-1) S^2 + \infty b \}$$

El número total de unidades de segunda etapa (cuestionarios) está dado aproximadamente por:

$$m_0 = \frac{n M}{\alpha} = \frac{\bar{M} a (N-1) S^2}{\alpha} + a b \bar{M}$$

$$\text{en donde: } 1 \leq \alpha \leq \frac{N \cdot a (N-1) S^2}{a b}$$

VII.- Muestreo Monoetápico, Equiprobable, con Reemplazo y Estratificado

Estimador de la media poblacional:

$$\hat{\mu} = \frac{1}{N_0} \sum_{t=1}^K N_t \bar{X}_t$$

$$N_0 = \sum_{t=1}^K N_t, \quad \bar{X}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} X_{ti}$$

$$V(\hat{\mu}) = E \left\{ \frac{1}{N_0} \sum_{t=1}^K N_t \bar{X}_t - \mu \right\}^2$$

$$= E \left\{ \frac{1}{N_0} \sum_{t=1}^K N_t \bar{X}_t - \frac{1}{N_0} \sum_{t=1}^K N_t \mu_t \right\}^2$$

$$= E \left\{ \frac{1}{N_0} \sum_{t=1}^K N_t (\bar{X}_t - \mu_t) \right\}^2$$

$$= \frac{1}{N_0^2} E \left\{ \sum_{t=1}^K N_t^2 (\bar{X}_t - \mu_t)^2 + \sum_{t \neq t'}^K N_t N_{t'} (\bar{X}_t - \mu_t) (\bar{X}_{t'} - \mu_{t'}) \right\}$$

$$= \frac{1}{N_0^2} \sum_{t=1}^K N_t^2 \frac{N_t - 1}{N_t n_t} S_t^2$$

$$= \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \left(\frac{1}{n_t} - \frac{1}{N_t n_t}\right) S_t^2$$

Estimador de la variancia de  $\hat{\mu}$

$$\text{Est. } V(\hat{\mu}) = \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \frac{s_t^2}{n_t}, \quad s_t^2 = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} (\dot{X}_{ti} - \dot{X}_t)^2, \quad \dot{X}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} \dot{X}_{ti}$$

que es insesgado, demostración:

$$E\{\text{Est. } V(\hat{\mu})\} = E\left\{ \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \frac{s_t^2}{n_t} \right\}$$

$$= \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \frac{1}{n_t} E\{s_t^2\}$$

$$= \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \frac{1}{n_t} E\left\{ \frac{1}{n_t - 1} \sum_{i=1}^{n_t} (\dot{X}_{ti} - \dot{X}_t)^2 \right\}$$

$$= \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \frac{1}{n_t (n_t - 1)} E\left\{ \sum_{i=1}^{n_t} (\dot{X}_{ti} - \mu_t + \mu_t - \dot{X}_t)^2 \right\}$$

$$= \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \frac{1}{n_t (n_t - 1)} E\left\{ \sum_{i=1}^{n_t} \{(\dot{X}_{ti} - \mu_t)^2 + 2(\dot{X}_{ti} - \mu_t)(\mu_t - \dot{X}_t) + (\mu_t - \dot{X}_t)^2\} \right\}$$

$$= \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \frac{1}{n_t (n_t - 1)} E\left\{ \sum_{i=1}^{n_t} (X_{ti} - \mu_t)^2 - n_t (\mu_t - \dot{X}_t)^2 \right\}$$

$$= \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \frac{1}{n_t (n_t - 1)} \left\{ n_t \frac{N_t - 1}{N_t} S_t^2 - n_t \frac{N_t - 1}{N_t n_t} S_t^2 \right\}$$

$$= \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \frac{1}{n_t (n_t - 1)} \frac{N_t - 1}{N_t} (n_t - 1) S_t^2$$

$$= \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \frac{N_t - 1}{N_t n_t} S_t^2$$

$$= \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \left(\frac{1}{n_t} - \frac{1}{N_t n_t}\right) S_t^2$$

$$= V(\hat{\mu})$$

Tamaño de la muestra:

$$V(\hat{\mu}) = \frac{\epsilon^2 \mu^2}{t^2 (\infty, \infty)}$$

$$\sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \left(\frac{1}{n_t} - \frac{1}{N_t n_t}\right) S_t^2 = \frac{\epsilon^2 \mu^2}{t^2 (\infty, \infty)}$$

$$\frac{1}{n_t} \left\{ \left( \frac{N_t}{N_o} \right)^2 \left( 1 - \frac{1}{N_t} \right) S_t^2 \right\} = \frac{\epsilon^2 \mu^2}{K t^2 (\infty, \infty)}$$

$$n_t = \left( \frac{N_t}{N_o} \right)^2 \frac{K (N_t - 1) S_t^2 t^2 (\infty, \infty)}{N_t \epsilon^2 \mu^2}$$

### VIII.- Muestreo Bietápico, Equiprobable, con Reemplazo y Estratificado.

Estimador de la media poblacional.

$$\hat{\mu} = \frac{1}{M_o} \sum_{t=1}^K \frac{M_{ot}}{n_t \bar{M}_t} \sum_{i=1}^{n_t} M_{ti} \bar{X}_{ti}$$

$$M_o = \sum_{t=1}^K M_{ot}, \quad M_{ot} = \sum_{i=1}^{N_t} M_{ti}, \quad \bar{M}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} M_{ti}, \quad \bar{X}_{ti} = \frac{1}{m_{ti}} \sum_{j=1}^{m_{ti}} X_{tij}$$

Variancia del estimador  $\hat{\mu}$

$$V(\hat{\mu}) = E \left\{ \frac{1}{M_o} \sum_{t=1}^K \frac{M_{ot}}{n_t \bar{M}_t} \sum_{i=1}^{n_t} M_{ti} \bar{X}_{ti} - \mu \right\}^2$$

$$= E \left\{ \frac{1}{M_o} \sum_{t=1}^K M_{ot} \left\{ \frac{1}{n_t \bar{M}_t} \sum_{i=1}^{n_t} M_{ti} \bar{X}_{ti} - \mu_t \right\}^2 \right.$$

$$\left. + \frac{1}{M_o^2} E \left\{ \sum_{t=1}^K M_{ot}^2 \left\{ \frac{1}{n_t \bar{M}_t} \sum_{i=1}^{n_t} M_{ti} (\bar{X}_{ti} - \mu_t) \right\}^2 + \right.$$

$$\left. + \sum_{t=1}^K M_{ot} M_{ot'} \left\{ \frac{1}{n_t \bar{M}_t} \sum_{i=1}^{n_t} M_{ti} (\bar{X}_{ti} - \mu_t) \right\} \left\{ \frac{1}{n_{t'} \bar{M}_{t'}} \sum_{i=1}^{n_{t'}} M_{t'i} (\bar{X}_{t'i} - \mu_{t'}) \right\} \right\}$$

$$= \frac{1}{M_o^2} \sum_{t=1}^K \left( \frac{M_{ot}}{n_t \bar{M}_t} \right)^2 \left\{ E \left\{ n_t M_{ti}^2 (\bar{X}_{ti} - \mu_t)^2 \right\} + 0 \right\}$$

$$= \frac{1}{M_o^2} \sum_{t=1}^K \frac{M_{ot}^2}{n_t} E \left\{ \left( \frac{M_{ti}}{\bar{M}_t} \right)^2 (\bar{X}_{ti} - \mu_t + \mu_t - \mu_t)^2 \right\}$$

$$= \sum_{t=1}^K \left( \frac{M_{ot}}{M_o} \right)^2 \frac{1}{n_t} \left\{ \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{ti}}{\bar{M}_t} \right)^2 \frac{M_{ti} - 1}{M_{ti} m_{ti}} S_{ti}^2 + \frac{N_t - 1}{N_t} S_t^2 \right\}$$

$$S_t^2 = \frac{1}{N_t - 1} \sum_{i=1}^{N_t} \left\{ \frac{M_{ti}}{\bar{M}_t} (\mu_{ti} - \mu_t) \right\}^2, \quad S_{ti}^2 = \frac{1}{M_{ti} - 1} \sum_{j=1}^{M_{ti}} (X_{tij} - \mu_{ti})^2$$

$$\bar{M}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} M_{ti}, \quad \mu_{ti} = \frac{1}{M_{ti}} \sum_{j=1}^{M_{ti}} X_{tij}, \quad \mu_t = \frac{1}{M_{ot}} \sum_{i=1}^{N_t} M_{ti} \mu_{ti}$$

Estimador de la variancia de  $\hat{\mu}$

$$\text{Est. } V(\hat{\mu}) = \sum_{t=1}^K \left( \frac{M_{ot}}{M_o} \right)^2 \frac{s_t^2}{n_t}$$

$$s_t^2 = \frac{1}{n_t - 1} \sum_{i=1}^{n_t} \left\{ \frac{1}{\bar{M}_t} (M_{ti} \bar{X}_{ti} - \bar{M}_t \bar{X}_t) \right\}^2, \quad \bar{M}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} M_{ti}$$

$$\dot{\bar{X}}_t = \frac{1}{n_t} \sum_{i=1}^{N_t} M_{ti} \dot{X}_{ti}, \quad \dot{\bar{X}}_{tj} = \frac{1}{m_{tj}} \sum_{i=1}^{m_{tj}} \dot{X}_{tj}$$

$$+ \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{ti}}{M_t} \right)^2 \{ \mu_t^2 + 2\mu_{ti}\mu_t - 2\mu_t^2 \} - \mu_t^2 \}$$

$$= V(\dot{\mu}_t)$$

Este estimador es aproximadamente insesgado, demostración:

Tamaño de la muestra:

$$E \{ \text{Est. } V(\dot{\mu}) \} = E \left\{ \sum_{t=1}^K \left( \frac{M_{0t}}{M_0} \right)^2 \frac{1}{n_t(n_t-1)} \sum_{i=1}^{n_t} \left\{ \frac{1}{M_t} (M_{ti} \dot{X}_{ti} - \dot{M}_t \dot{X}_t) \right\}^2 \right.$$

$$= \sum_{t=1}^K \left( \frac{M_{0t}}{M_0} \right)^2 \frac{1}{n_t(n_t-1)} E \left\{ \sum_{i=1}^{n_t} \left( \frac{M_{ti}}{M_t} \right)^2 \dot{X}_{ti}^2 - n_t \dot{X}_t^2 \right\}$$

$$= \sum_{t=1}^K \left( \frac{M_{0t}}{M_0} \right)^2 \frac{1}{n_t(n_t-1)} E \left\{ (n_t-1) \frac{1}{n_t} \sum_{i=1}^{n_t} \left( \frac{M_{ti}}{M_t} \right)^2 (\dot{X}_{ti}^2 - \mu_{ti}^2) + \right.$$

$$\left. + \mu_{ti}^2 - \mu_t^2 \right\} - \frac{1}{n_t M_t^2} \sum_{i \neq j}^{n_t} M_{ti} M_{tj} \dot{X}_{ti} \dot{X}_{tj} \}$$

$$E \{ \text{Est. } V(\dot{\mu}) \} = \sum_{t=1}^K \left( \frac{M_{0t}}{M_0} \right)^2 \frac{1}{n_t(n_t-1)} \{ (n_t-1) \}$$

$$1) \left\{ \frac{1}{n_t} \sum_{i=1}^{N_t} \left( \frac{M_{ti}}{M_t} \right)^2 \frac{M_{ti}-1}{M_{ti} m_{ti}} S_{ti}^2 + \right.$$

$$\left. + \frac{N_t-1}{N_t} S_t^2 + \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{ti}}{M_t} \right)^2 \{ \mu_t^2 + O + 2(\mu_{ti} - \mu_t) \mu_t \} - \mu_t^2 \right.$$

$$= \sum_{t=1}^K \left( \frac{M_{0t}}{M_0} \right)^2 \frac{1}{n_t} \left\{ \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{ti}}{M_t} \right)^2 \frac{M_{ti}-1}{M_{ti} m_{ti}} S_{ti}^2 + \frac{N_t-1}{N_t} S_t^2 + \right.$$

$$V(\dot{\mu}) = \frac{\epsilon^2 \mu^2}{f^2 (\infty, \infty)}$$

$$\sum_{t=1}^K \left( \frac{M_{0t}}{M_0} \right)^2 \frac{1}{n_t} \left\{ \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{ti}}{M_t} \right)^2 \frac{M_{ti}-1}{M_{ti} m_{ti}} S_{ti}^2 + \frac{N_t-1}{N_t} S_t^2 \right\} = \frac{\epsilon^2 \mu^2}{f^2 (\infty, \infty)}$$

$$\left( \frac{M_{0t}}{M_0} \right)^2 \frac{1}{n_t} \left\{ \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{ti}}{M_t} \right)^2 \frac{M_{ti}-1}{M_{ti} m_{ti}} S_{ti}^2 + \frac{N_t-1}{N_t} S_t^2 \right\} = \frac{\epsilon^2 \mu^2}{K f^2 (\infty, \infty)}$$

Sea  $m_{ti} = \frac{M_{ti}}{\infty_t}$  (colocación proporcional)

$$\frac{1}{n_t} \left\{ \frac{1}{N_t} \sum_{i=1}^{N_t} \left( \frac{M_{ti}}{M_t} \right)^2 \frac{\infty (M_{ti}-1)}{M_{ti}^2} S_{ti}^2 + \frac{N_t-1}{N_t} S_t^2 \right\} = \frac{\epsilon^2 \mu^2}{K f^2 (\infty, \infty)}$$

$$\frac{1}{n_t} \left\{ \frac{\infty}{N_t M_t^2} \sum_{i=1}^{N_t} (M_{ti}-1) S_{ti}^2 + \frac{N_t-1}{N_t} S_t^2 \right\} = \frac{\epsilon^2 \mu^2}{K \left( \frac{M_{0t}}{M_0} \right)^2 f^2 (\infty, \infty)}$$

$$n_t = \frac{f^2 (\infty, \infty)}{\epsilon^2} \frac{1}{\mu^2 N_t} \left\{ (N_t-1) S_t^2 + \frac{\infty}{M_t^2} \sum_{i=1}^{N_t} (M_{ti}-1) S_{ti}^2 \right\}$$

$$a_t = \frac{t^2 (\infty, \infty)}{\epsilon^2} \frac{1}{\mu^2 N_t}, \quad b_t = \frac{1}{\bar{M}_t^2} \sum_{i=1}^N (M_{ti} - 1) S_{ti}^2$$

$$K \frac{(M_{ot})^2}{M_o}$$

$$n_t = a_t \{ (N_t - 1) S_t^2 + b_t \infty_t \}$$

$$S_t^2 = \frac{1}{N_t - 1} \sum_{i=1}^{N_t} \left\{ \frac{M_{ti}}{M_t} (\mu_{ti} - \mu_t) \right\}^2, \quad S_{ti}^2 = \frac{1}{M_{ti} - 1} \sum_{j=1}^{M_{ti}} (X_{tij} - \mu_{ti})^2$$

$$\bar{M}_t = \frac{1}{N_t} \sum_{i=1}^{N_t} M_{ti}, \quad M_{ot} = N_t \bar{M}_t, \quad M_o = \sum_{t=1}^K M_{ot}$$

$$\mu_{ti} = \frac{1}{M_{ti}} \sum_{j=1}^{M_{ti}} X_{tij}, \quad \mu_t = \frac{1}{M_{ot}} \sum_{i=1}^{N_t} M_{ti} \mu_{ti}, \quad \mu = \frac{1}{M_o} \sum_{t=1}^K M_{ot} \mu_t$$

$$n_t = a_t \{ (N_t - 1) S_t^2 + b_t \infty_t \}$$

$$1 \leq \infty_t \leq \frac{N_t}{a_t b_t} - \frac{N_t - 1}{b_t} S_t^2$$

IX.- Muestreo Monoetápico con Probabilidad Variable de Selección y con Reemplazo.

Estimador de la media poblacional.

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N \frac{X_i}{P_i}$$

P. Probabilidad de incluir a la i-ésima unidad en la muestra

$$\sum_{i=1}^N P_i = 1$$

Variación del estimador  $\hat{\mu}$

$$V(\hat{\mu}) = E \left\{ \frac{1}{Nn} \sum_{i=1}^n \frac{X_i}{P_i} - \mu \right\}^2$$

$$= E \left\{ \frac{1}{Nn} \sum_{i=1}^n \left( \frac{X_i}{P_i} - N\mu \right) \right\}^2$$

$$= \frac{1}{n} \left\{ \frac{1}{N^2} \sum_{i=1}^n P_i \left( \frac{X_i}{P_i} - N\mu \right)^2 + 0 \right\}$$

$$= \frac{1}{n} \sigma^2$$

$$\sigma^2 = \frac{1}{N^2} \sum_{i=1}^N P_i \left( \frac{X_i}{P_i} - N\mu \right)^2$$

Estimador de la variancia de  $\hat{\mu}$

$$\text{Est. } V(\hat{\mu}) = \frac{s^2}{n}, \quad s^2 = \frac{1}{N^2(n-1)} \sum_{i=1}^N \left( \frac{X_i}{P_i} - N\hat{\mu} \right)^2$$

Estimador insesgado, demostración:

$$E \{ \text{Est. } V(\hat{\mu}) \} = E \left\{ \frac{1}{N^2 n (n-1)} \sum_{i=1}^N \left( \frac{X_i}{P_i} - N\hat{\mu} \right)^2 \right\}$$

$$= \frac{1}{N^2 n (n-1)} E \left\{ \sum_{i=1}^N \left( \frac{X_i}{P_i} \right)^2 - n(N\hat{\mu})^2 \right\}$$

$$= \frac{1}{N^2 n (n-1)} E \left\{ \sum_{i=1}^N \left( \frac{X_i}{P_i} - N\mu + N\mu \right)^2 - nN^2 (\hat{\mu} - \mu + \mu)^2 \right\}$$

$$= \frac{1}{N^2 n (n-1)} \left\{ n E \left( \frac{X_i}{P_i} - N\mu \right)^2 + n(N\mu)^2 - nN^2 V(\hat{\mu}) - nN^2 \mu^2 \right\}$$

resultando:

$$= \frac{1}{N^2 n (n-1)} \left\{ n N^2 \sigma^2 - \frac{n N^2}{n} \sigma^2 \right\}$$

$$= \frac{\sigma^2}{n-1} \left\{ 1 - \frac{1}{n} \right\}$$

$$= \frac{\sigma^2}{n}$$

$$= V(\hat{\mu})$$

Tamaño de la muestra:

$$V(\hat{\mu}) = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2}$$

$$\frac{\sigma^2}{n} = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2}$$

$$n = \frac{\sigma^2 t_{(\alpha, \infty)}^2}{\epsilon^2 \mu^2}$$

$$\sigma^2 = \frac{1}{N^2} \sum_{i=1}^N P_i \left( \frac{X_i}{P_i} - N\mu \right)^2, \quad \mu = \frac{1}{N} \sum_{i=1}^N X_i$$

Muestreo Bietápico con Probabilidades Variables de Selección en cada Etapa y con Reemplazo.

Estimador de la media poblacional:

$$\hat{\mu} = \frac{1}{M_0 n} \sum_{i=1}^R \hat{X}_i$$

donde:

$$\dot{X}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} X_{ij}, \quad \mu = \frac{1}{M_0} \sum_{i=1}^N M_i \mu_i, \quad \mu_i = \frac{1}{M_i} \sum_{j=1}^{M_i} X_{ij}$$

$P_i$  = Probabilidad de incluir en la muestra la  $i$ -ésima unidad de primera etapa.

$P_{ij}$  = Probabilidad de incluir en la muestra la  $i$ -ésima unidad de primera etapa y la  $j$ -ésima unidad de segunda etapa.

$P_{j/1}$  = Probabilidad de incluir en la muestra la  $j$ -ésima unidad de segunda etapa, dado que la  $i$ -ésima unidad de primera etapa ha sido seleccionada.

$$\sum_{i=1}^N P_i = 1, \quad \sum_{j=1}^{M_i} P_{ij} = P_i$$

Variancia del estimador  $\hat{\mu}$

$$V(\hat{\mu}) = E \left\{ \frac{1}{M_0 n} \sum_{i=1}^N \dot{X}_i - \mu \right\}^2$$

$$= E \left\{ \frac{1}{M_0 n} \sum_{i=1}^N (\dot{X}_i - M_0 \mu) \right\}^2$$

$$= \frac{1}{M_0^2 n^2} E \left\{ \sum_{i=1}^N (\dot{X}_i - M_0 \mu)^2 + \sum_{i \neq 1}^N (\dot{X}_i - M_0 \mu) (\dot{X}_1 - M_0 \mu) \right\}$$

$$= \frac{1}{M_0^2 n} \left\{ E (X_1 - \frac{M_1 \mu_1}{P_1} + \frac{M_1 \mu_1}{P_1} - M_0 \mu)^2 + \right.$$

$$\left. + n(n-1) \left\{ E \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{X_{ij}}{P_{ij}} - M_0 \mu \right\}^2 \right\}$$

$$= \frac{1}{M_0^2 n} \left\{ E \left( \dot{X}_1 - \frac{M_1 \mu_1}{P_1} \right)^2 + E \left( \frac{M_1 \mu_1}{P_1} - M_0 \mu \right)^2 \right\}$$

$$= \frac{1}{M_0^2 n} \left\{ E \left\{ \frac{1}{m_i} \sum_{j=1}^{m_i} \dot{X}_{ij} - \frac{M_i \mu_i}{P_i} \right\}^2 + \sum_{i=1}^N P_i \left( \frac{M_i \mu_i}{P_i} - M_0 \mu \right)^2 \right\}$$

$$= \frac{1}{M_0^2 n} \left\{ \sum_{i=1}^N P_i \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{P_{ij}}{P_i} \left( \dot{X}_{ij} - \frac{M_i \mu_i}{P_i} \right)^2 + \sum_{i=1}^N P_i \left( \frac{M_i \mu_i}{P_i} - M_0 \mu \right)^2 \right\}$$

$$= \frac{1}{M_0^2 n} \left\{ \sum_{i=1}^N P_i \frac{1}{m_i} \sigma_i^2 + \sigma^2 \right\}$$

$$\sigma_i^2 = \sum_{j=1}^{m_i} \frac{P_{ij}}{P_i} \left( \frac{X_{ij}}{P_{ij}} - \frac{M_i \mu_i}{P_i} \right)^2, \quad \sigma^2 = \sum_{i=1}^N P_i \left( \frac{M_i \mu_i}{P_i} - M_0 \mu \right)^2$$

Estimador de la variancia de  $\hat{\mu}$

$$\text{Est. } V(\hat{\mu}) = \frac{1}{M_0^2 n} s^2$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^N (\dot{X}_i - M_0 \hat{\mu})^2, \quad \dot{X}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \frac{X_{ij}}{P_{ij}}, \quad \hat{\mu} = \frac{1}{M_0 n} \sum_{i=1}^N \dot{X}_i$$

Este estimador es insesgado, demostración:

$$E \{ \text{Est. } V(\hat{\mu}) \} = E \left\{ \frac{1}{M_0^2 n(n-1)} \sum_{i=1}^N (\bar{X}_i - M_0 \hat{\mu})^2 \right\}$$

$$= \frac{1}{M_0^2 n(n-1)} E \left\{ \sum_{i=1}^N \bar{X}_i^2 - n M_0^2 \hat{\mu}^2 \right\}$$

$$= \frac{1}{M_0^2 n(n-1)} E \left\{ \sum_{i=1}^N \bar{X}_i^2 - \frac{1}{n} \left( \sum_{i=1}^N \bar{X}_i \right)^2 \right\}$$

$$= \frac{1}{M_0^2 n} \left\{ E \left( \bar{X}_i - \frac{M_{i1} \mu_{i1}}{P_{i1}} + \frac{M_{i1} \mu_{i1}}{P_{i1}} \right)^2 - \left\{ \sum_{i=1}^N P_{i1} \left( \frac{M_{i1} \mu_{i1}}{P_{i1}} \right) \right\}^2 \right\}$$

$$= \frac{1}{M_0^2 n} \left\{ E(\bar{X}_i - \frac{M_{i1} \mu_{i1}}{P_{i1}})^2 + \sum_{i=1}^N P_{i1} \left( \frac{M_{i1} \mu_{i1}}{P_{i1}} \right)^2 - (M_0 \mu)^2 \right\}$$

$$E \{ \text{Est. } V(\hat{\mu}) \} = \frac{1}{M_0^2 n} \left\{ E \left\{ \left( \frac{1}{m_i} \sum_{j=1}^{m_i} \bar{X}_{ij} \right) - \right. \right.$$

$$\left. \left. \frac{M_{i1} \mu_{i1}}{P_{i1}} \right\}^2 + \sum_{i=1}^N P_{i1} \left\{ \left( \frac{M_{i1} \mu_{i1}}{P_{i1}} \right)^2 - (M_0 \mu)^2 \right\} \right\}$$

$$= \frac{1}{M_0^2 n} \left\{ E \left\{ \frac{1}{M_{i1}} \sum_{j=1}^{m_i} \left( \bar{X}_{ij} - \frac{M_{i1} \mu_{i1}}{P_{i1}} \right) \right\}^2 + \sum_{i=1}^N P_{i1} \left( \frac{M_{i1} \mu_{i1}}{P_{i1}} - M_0 \mu \right)^2 \right\}$$

$$= \frac{1}{M_0^2 n} \left\{ \left( \frac{1}{m_i} \right)^2 E \left\{ \sum_{j=1}^{m_i} \left( \bar{X}_{ij} - \frac{M_{i1} \mu_{i1}}{P_{i1}} \right)^2 + \sum_{i \neq i'} \left( \frac{\bar{X}_{ij}}{P_{ij}} - \frac{M_{i1} \mu_{i1}}{P_{i1}} \right) \left( \frac{\bar{X}_{i'j}}{P_{i'j}} - \frac{M_{i1} \mu_{i1}}{P_{i1}} \right) \right\} + \sigma^2 \right\}$$

$$= \frac{1}{M_0^2 n} \left\{ \sum_{i=1}^N \frac{1}{m_i} P_{i1} \sum_{j=1}^{m_i} \frac{P_{ij}}{P_{i1}} \left( \bar{X}_{ij} - \frac{M_{i1} \mu_{i1}}{P_{i1}} \right)^2 + \sigma^2 \right\}$$

$$= \frac{1}{M_0^2 n} \left\{ \sum_{i=1}^N \frac{P_{i1}}{m_i} \sigma_i^2 + \sigma^2 \right\}$$

$$= V(\hat{\mu})$$

Tamaño de la muestra:

$$V(\hat{\mu}) = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2}$$

$$\frac{1}{M_0^2 n} \left\{ \sum_{i=1}^N \frac{P_{i1}}{m_i} \sigma_i^2 + \sigma^2 \right\} = \frac{\epsilon^2 \mu^2}{t_{(\alpha, \infty)}^2}$$

$$n = \frac{t_{(\alpha, \infty)}^2}{M_0^2 \epsilon^2 \mu^2} \left\{ \sum_{i=1}^N \frac{P_{i1}}{m_i} \sigma_i^2 + \sigma^2 \right\}$$

Si  $m_i = \frac{M_{i1}}{\alpha}$  con  $1 \leq \alpha$  resulta:

$$n = \frac{t_{(\alpha, \infty)}^2}{M_0^2 \epsilon^2 \mu^2} \left\{ \alpha \sum_{i=1}^N \frac{P_{i1}}{M_{i1}} \sigma_i^2 + \sigma^2 \right\}$$

$$1 < \alpha < \frac{N}{a} - \frac{\sigma^2}{b}$$

$$n = a (\alpha b + \sigma^2)$$

$$a = \frac{t^2_{(\alpha, \infty)}}{M_0^2 \epsilon^2 \mu^2}, \quad b = \sum_{i=1}^N \frac{P_i \sigma_i^2}{M_i}$$

SINTESIS DE LAS CARACTERISTICAS DE ALGUNOS MODELOS DE MUESTREO

DENOMINACION DEL MODELO DE MUESTREO	MEDIA POBLACIONAL	ESTIMADORES		VARIANCIAS DEL ESTIMADOR DE LA MEDIA POBLACIONAL	TAMAÑO DE LA MUESTRA
		MEDIA POBLACIONAL	VARIANCIAS DEL ESTIMADOR DE LA MEDIA POBLACIONAL		
Muestreo monoetápico, equiprobable y sin reemplazo (muestreo irrestricto aleatorio).	$\mu = \frac{1}{N} \sum_{i=1}^N X_i$	$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$	Est. $V(\hat{\mu}) = \frac{N-n}{Nn} s^2$	$V(\hat{\mu}) = \frac{N-n}{Nn} S^2$ $S^2 = \frac{1}{N-1} \sum_{i=1}^N (\bar{X}_i - \mu)^2$	$n = \frac{S^2 t^2_{(\alpha, \infty)}}{\epsilon^2 \mu^2} + \frac{1}{N} \frac{S^2 t^2_{(\alpha, \infty)}}{\epsilon^2 \mu^2}$
Muestreo bietápico, equiprobable y sin reemplazo.	$\mu = \frac{1}{M_0} \sum_{i=1}^N M_i \mu_i$ $\mu_t = \frac{1}{M_i} \sum_{j=1}^{m_i} X_{ij}$ $M_0 = \sum_{i=1}^N M_i$ $\bar{M} = \frac{M_0}{N}$	$\hat{\mu} = \frac{1}{\sum_{i=1}^N M_i} \sum_{i=1}^N M_i \bar{X}_i$ $\bar{X}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \bar{x}_{ij}$	Est. $V(\hat{\mu}) = \left(\frac{1}{n} - \frac{1}{N}\right) \hat{s}^2 + \frac{1}{Nn} \sum_{i=1}^N \left(\frac{M_i}{M}\right)^2 \left(\frac{1}{m_i} - \frac{1}{M_i}\right) \hat{s}_i^2$ $\hat{s}^2 = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{M_i}{M}\right)^2 (\bar{X}_i - \bar{X})^2$ $\hat{s}_i^2 = \frac{1}{m_i-1} \sum_{j=1}^{m_i} (X_{ij} - \mu)^2$	$V(\hat{\mu}) = \left(\frac{1}{n} - \frac{1}{N}\right) S^2 + \frac{1}{Nn} \sum_{i=1}^N \left(\frac{M_i}{M}\right)^2 \left(\frac{1}{m_i} - \frac{1}{M_i}\right) S_i^2$ $S^2 = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{M_i}{M}\right)^2 (\mu_i - \mu)^2$ $S_i^2 = \frac{1}{M_i-1} \sum_{j=1}^{M_i} (X_{ij} - \mu_i)^2$	$n = a (S^2 + b (\alpha - 1))$ $a = \frac{N t^2_{(\alpha, \infty)}}{\epsilon^2 \mu^2 N + S^2 t^2_{(\alpha, \infty)}}$ $b = \frac{1}{N \bar{M}^2} \sum_{i=1}^N M_i S_i^2$ $m_i = \frac{M_i}{\alpha}$ $1 < \alpha < \left(\frac{N}{ab} + 1 - \frac{S^2}{b}\right)$
Muestreo monoetápico, equiprobable, sin reemplazo y estratificado.	$\mu = \frac{1}{N_0} \sum_{t=1}^K N_t \mu_t$ $\mu_t = \frac{1}{N_t} \sum_{i=1}^{N_t} X_{ti}$ $N_0 = \sum_{t=1}^K N_t$	$\hat{\mu} = \frac{1}{N_0} \sum_{t=1}^K N_t \bar{X}_t$ $\bar{X}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} X_{ti}$	Est. $V(\hat{\mu}) = \sum_{t=1}^K \left(\frac{1}{n_t} - \frac{1}{N_t}\right) \left(\frac{N_t}{N_0}\right)^2 \hat{s}_t^2$ $\hat{s}_t^2 = \frac{1}{n_t-1} \sum_{i=1}^{n_t} (X_{ti} - \bar{X}_t)^2$	$V(\hat{\mu}) = \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \left(\frac{1}{n_t} - \frac{1}{N_t}\right) S_t^2$ $S_t^2 = \frac{1}{N_t-1} \sum_{i=1}^{N_t} (X_{ti} - \mu_t)^2$	$n_t = \frac{S_t^2 t^2_{(\alpha, \infty)}}{\left(\frac{N_0}{N_t}\right)^2 \frac{\epsilon^2}{K} \mu^2} + \frac{1}{N_t} \frac{S_t^2 t^2_{(\alpha, \infty)}}{\left(\frac{N_0}{N_t}\right)^2 \frac{\epsilon^2}{K} \mu^2}$

DENOMINACION DEL MODELO DE MUESTREO	ESTIMADORES			VARIANCIA DEL ESTIMADOR DE LA MEDIA POBLACIONAL	TAMAÑO DE LA MUESTRA
	MEDIA POBLACIONAL	MEDIA POBLACIONAL	VARIANCIA DEL ESTIMADOR DE LA MEDIA POBLACIONAL		
Muestreo monoetápico, equiprobable y sin reemplazo (muestreo irrestricto aleatorio).	$\mu = \frac{1}{N} \sum_{i=1}^N X_i$	$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n \dot{X}_i$	$\text{Est. } V(\hat{\mu}) = \frac{N-n}{Nn} s^2$	$V(\hat{\mu}) = \frac{N-n}{Nn} S^2$ $S^2 = \frac{1}{N-1} \sum_{i=1}^N (\dot{X}_i - \mu)^2$	$n = \frac{S^2 f^2(\alpha, \infty)}{\epsilon^2 \mu^2}$ $1 + \frac{1}{N} \frac{S^2 f^2(\alpha, \infty)}{\epsilon^2 \mu^2}$
Muestreo bietápico, equiprobable y sin reemplazo.	$\mu = \frac{1}{M_0} \sum_{i=1}^N M_i \mu_i$ $\mu_i = \frac{1}{M_i} \sum_{j=1}^{m_i} X_{ij}$ $M_0 = \sum_{i=1}^N M_i$ $\bar{M} = \frac{M_0}{N}$	$\hat{\mu} = \frac{1}{\sum_{i=1}^N M_i} \sum_{i=1}^n M_i \dot{X}_i$ $\dot{X}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \dot{X}_{ij}$	$\text{Est. } V(\hat{\mu}) = \left(\frac{1}{n} - \frac{1}{N}\right) s^2 + \frac{1}{Nn} \sum_{i=1}^N \left(\frac{M_i}{M}\right)^2 \left(\frac{1}{m_i} - \frac{1}{M_i}\right) \dot{s}_i^2$ $s^2 = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{M_i}{M}\right)^2 \left(\frac{\dot{X}_i}{M} - \bar{X}\right)^2$ $\dot{s}_i^2 = \frac{1}{m_i-1} \sum_{j=1}^{m_i} (\dot{X}_{ij} - \mu)^2$	$V(\hat{\mu}) = \left(\frac{1}{n} - \frac{1}{N}\right) S^2 + \frac{1}{Nn} \sum_{i=1}^N \left(\frac{M_i}{M}\right)^2 \left(\frac{1}{m_i} - \frac{1}{M_i}\right) S_i^2$ $S^2 = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{M_i}{M}\right)^2 (\mu_i - \mu)^2$ $S_i^2 = \frac{1}{M_i-1} \sum_{j=1}^{m_i} (X_{ij} - \mu_i)^2$	$n = a(S^2 + b(\alpha - 1))$ $a = \frac{N f^2(\alpha, \infty)}{\epsilon^2 \mu^2 N + S^2 f^2(\alpha, \infty)}$ $b = \frac{1}{N \bar{M}^2} \sum_{i=1}^N M_i S_i^2$ $m_i = \frac{M_i}{\alpha}$ $1 \leq \alpha \leq \left(\frac{N}{ab} + 1 - \frac{S^2}{b}\right)$
Muestreo monoetápico, equiprobable, sin reemplazo y estratificado.	$\mu = \frac{1}{N_0} \sum_{t=1}^K N_t \mu_t$ $\mu_t = \frac{1}{N_t} \sum_{i=1}^{n_t} X_{ti}$ $N_0 = \sum_{t=1}^K N_t$	$\hat{\mu} = \frac{1}{N_0} \sum_{t=1}^K N_t \dot{X}_t$ $\dot{X}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} \dot{X}_{ti}$	$\text{Est. } V(\hat{\mu}) = \sum_{t=1}^K \left(\frac{1}{n_t} - \frac{1}{N_t}\right) \left(\frac{N_t}{N_0}\right)^2 \dot{s}_t^2$ $\dot{s}_t^2 = \frac{1}{n_t-1} \sum_{i=1}^{n_t} (\dot{X}_{ti} - \dot{X}_t)^2$	$V(\hat{\mu}) = \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \left(\frac{1}{n_t} - \frac{1}{N_t}\right) S_t^2$ $S_t^2 = \frac{1}{N_t-1} \sum_{i=1}^{n_t} (X_{ti} - \mu_t)^2$	$n_t = \frac{S_t^2 f^2(\alpha, \infty)}{\left(\frac{N_0}{N_t}\right)^2 \epsilon^2 \mu^2}$ $1 + \frac{1}{N_t} \frac{S_t^2 f^2(\alpha, \infty)}{\left(\frac{N_0}{N_t}\right)^2 \epsilon^2 \mu^2}$
Muestreo bietápico, equiprobable, sin reemplazo y estratificado.	$\mu = \frac{1}{M_0} \sum_{t=1}^K M_{0t} \mu_t$ $\mu_t = \frac{1}{M_{0t}} \sum_{i=1}^{n_t} M_{ti} \mu_{ti}$ $\mu_{ti} = \frac{1}{M_{ti}} \sum_{j=1}^{m_{ti}} X_{tij}$ $M_0 = \sum_{t=1}^K M_{0t}$ $M_{0t} = \sum_{i=1}^{n_t} M_{ti}$ $\bar{M}_t = \frac{M_{0t}}{n_t}$	$\hat{\mu} = \frac{1}{M_0} \sum_{t=1}^K M_{0t} \dot{X}_t$ $\dot{X}_t = \frac{1}{n_t M_{0t}} \sum_{i=1}^{n_t} M_{ti} \dot{X}_{ti}$ $\dot{X}_{ti} = \frac{1}{m_{ti}} \sum_{j=1}^{m_{ti}} \dot{X}_{tij}$ $\dot{M}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} M_{ti}$	$\text{Est. } V(\hat{\mu}) = \sum_{t=1}^K \left(\frac{M_{0t}}{M_0}\right)^2 \left\{ \left(\frac{1}{n_t} - \frac{1}{N_t}\right) \dot{s}_t^2 + \frac{1}{N_t n_t} \sum_{i=1}^{n_t} \left(\frac{M_{ti}}{M_{0t}}\right)^2 \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}}\right) \dot{s}_{ti}^2 \right\}$ $\dot{s}_t^2 = \frac{1}{n_t-1} \sum_{i=1}^{n_t} \left(\frac{M_{ti}}{M_{0t}}\right)^2 (\dot{X}_{ti} - \dot{X}_t)^2$ $\dot{s}_{ti}^2 = \frac{1}{m_{ti}-1} \sum_{j=1}^{m_{ti}} (\dot{X}_{tij} - \dot{X}_{ti})^2$	$V(\hat{\mu}) = \sum_{t=1}^K \left(\frac{M_{0t}}{M_0}\right)^2 \left\{ \left(\frac{1}{n_t} - \frac{1}{N_t}\right) S_t^2 + \frac{1}{N_t n_t} \sum_{i=1}^{n_t} \left(\frac{M_{ti}}{M_{0t}}\right)^2 \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}}\right) S_{ti}^2 \right\}$ $S_{ti}^2 = \frac{1}{M_{ti}-1} \sum_{j=1}^{m_{ti}} (X_{tij} - \mu_{ti})^2$ $S_t^2 = \frac{1}{N_t-1} \sum_{i=1}^{n_t} \left(\frac{M_{ti}}{M_{0t}}\right)^2 (\mu_{ti} - \mu_t)^2$	$n_t = a_t \{ S_t^2 + b_t(\alpha_t - 1) \}$ $a_t = \frac{K N_t M_{0t}^2 f^2(\alpha, \infty)}{N_t M_{0t}^2 \epsilon^2 \mu^2 + K M_{0t}^2 S_t^2 f^2(\alpha, \infty)}$ $b_t = \frac{1}{N_t \bar{M}_t^2} \sum_{i=1}^{n_t} M_{ti} S_{ti}^2$ $m_{ti} = \frac{1}{\alpha_t} M_{ti}$ $1 < \alpha_t \leq \frac{N_t}{a_t b_t} + 1 - \frac{S_t^2}{b_t}$
Muestreo monoetápico, equiprobable y con reemplazo (muestreo irrestricto aleatorio con reemplazo).	$\mu = \frac{1}{N} \sum_{i=1}^N X_i$	$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^n \dot{X}_i$	$\text{Est. } V(\hat{\mu}) = \frac{1}{n} s^2$ $s^2 = \frac{1}{n-1} \sum_{i=1}^n (\dot{X}_i - \bar{X})^2$	$V(\hat{\mu}) = \frac{N-1}{Nn} S^2$	$n = \frac{(N-1) S^2 f^2(\alpha, \infty)}{N \epsilon^2 \mu^2}$
Muestreo bietápico, equiprobable y con reemplazo.	$\mu = \frac{1}{M_0} \sum_{i=1}^N M_i \mu_i$ $\mu_i = \frac{1}{M_i} \sum_{j=1}^{m_i} X_{ij}$ $M_0 = \sum_{i=1}^N M_i$ $\bar{M} = \frac{M_0}{N}$	$\hat{\mu} = \frac{1}{n M_0} \sum_{i=1}^n M_i \dot{X}_i$ $\dot{X}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \dot{X}_{ij}$	$\text{Est. } V(\hat{\mu}) = \frac{s^2}{n M_0^2}$ $s^2 = \frac{1}{n-1} \sum_{i=1}^n \left\{ N M_i \dot{X}_i - \frac{N}{n} \sum_{i=1}^n M_i \dot{X}_i \right\}^2$	$V(\hat{\mu}) = \frac{N-1}{Nn M_0^2} S^2 + \frac{1}{Nn M_0^2} \sum_{i=1}^N \frac{M_i-1}{M_i} S_i^2$ $S^2 = \frac{1}{N-1} \sum_{i=1}^N (N M_i \mu_i - M_0 \mu)^2$ $S_i^2 = \frac{1}{M_i-1} \sum_{j=1}^{m_i} (N M_i X_{ij} - N M_i \mu_i)^2$	$n = a \{ (N-1) S^2 + \alpha b \}$ $a = \frac{f^2(\alpha, \infty)}{\epsilon^2 \mu^2 N M_0^2}$ $b = \sum_{i=1}^N \frac{M_i-1}{M_i} S_i^2$ $m_i = \frac{M_i}{\alpha}$ $1 \leq \alpha \leq \frac{N-a(N-1)S^2}{ab}$
Muestreo monoetápico, equiprobable con reemplazo y estratificado.	$\mu = \frac{1}{N_0} \sum_{t=1}^K N_t \mu_t$ $\mu_t = \frac{1}{N_t} \sum_{i=1}^{n_t} \dot{X}_{ti}$	$\hat{\mu} = \frac{1}{N_0} \sum_{t=1}^K N_t \dot{X}_t$ $\dot{X}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} \dot{X}_{ti}$	$\text{Est. } V(\hat{\mu}) = \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \frac{\dot{s}_t^2}{n_t}$ $\dot{s}_t^2 = \frac{1}{n_t-1} \sum_{i=1}^{n_t} (\dot{X}_{ti} - \dot{X}_t)^2$	$V(\hat{\mu}) = \sum_{t=1}^K \left(\frac{N_t}{N_0}\right)^2 \left(\frac{1}{n_t} - \frac{1}{N_t}\right) S_t^2$ $S_t^2 = \frac{1}{N_t-1} \sum_{i=1}^{n_t} (X_{ti} - \mu_t)^2$	$n_t = \left(\frac{N_t}{N_0}\right)^2 \frac{K(N_t-1) S_t^2 f^2(\alpha, \infty)}{N_t \epsilon^2 \mu^2}$
Muestreo bietápico, equiprobable, con reemplazo y estratificado.	$\mu = \frac{1}{M_0} \sum_{t=1}^K M_{0t} \mu_t$ $\mu_t = \frac{1}{M_{0t}} \sum_{i=1}^{n_t} M_{ti} \mu_{ti}$ $\mu_{ti} = \frac{1}{M_{ti}} \sum_{j=1}^{m_{ti}} X_{tij}$ $M_0 = \sum_{t=1}^K M_{0t}$ $M_{0t} = \sum_{i=1}^{n_t} M_{ti}$ $\bar{M}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} M_{ti}$	$\hat{\mu} = \frac{1}{M_0} \sum_{t=1}^K M_{0t} \dot{X}_t$ $\dot{X}_t = \frac{1}{n_t M_{0t}} \sum_{i=1}^{n_t} M_{ti} \dot{X}_{ti}$ $\dot{X}_{ti} = \frac{1}{m_{ti}} \sum_{j=1}^{m_{ti}} \dot{X}_{tij}$ $\dot{M}_t = \frac{1}{n_t} \sum_{i=1}^{n_t} M_{ti}$	$\text{Est. } V(\hat{\mu}) = \sum_{t=1}^K \left(\frac{M_{0t}}{M_0}\right)^2 \frac{\dot{s}_t^2}{n_t}$ $\dot{s}_t^2 = \frac{1}{n_t-1} \sum_{i=1}^{n_t} \left\{ \frac{1}{M_{ti}} (M_{ti} \dot{X}_{ti} - M_{ti} \dot{X}_t) \right\}^2$	$V(\hat{\mu}) = \sum_{t=1}^K \left(\frac{M_{0t}}{M_0}\right)^2 \left\{ \frac{1}{n_t} \sum_{i=1}^{n_t} \left(\frac{M_{ti}}{M_{0t}}\right)^2 \left(\frac{1}{m_{ti}} - \frac{1}{M_{ti}}\right) S_{ti}^2 + \left(\frac{1}{n_t} - \frac{1}{N_t}\right) S_t^2 \right\}$ $S_{ti}^2 = \frac{1}{M_{ti}-1} \sum_{j=1}^{m_{ti}} \left\{ \frac{M_{ti}}{M_{0t}} (\mu_{ti} - \mu_t) \right\}^2$ $S_t^2 = \frac{1}{N_t-1} \sum_{i=1}^{n_t} \sum_{j=1}^{m_{ti}} (X_{tij} - \mu_{ti})^2$	$n_t = a_t \{ (N_t-1) S_t^2 + b_t \alpha_t \}$ $a_t = \frac{f^2(\alpha, \infty)}{\epsilon^2 \mu^2 N_t}$ $b_t = \frac{1}{M_{0t}^2} \sum_{i=1}^{n_t} (M_{ti}-1) S_{ti}^2$ $m_{ti} = \frac{M_{ti}}{\alpha_t}$ $1 \leq \alpha_t \leq \frac{N_t}{a_t b_t} - \frac{N_t-1}{b_t} S_t^2$
Muestreo monoetápico con probabilidades variables de selección y con reemplazo.	$\mu = \frac{1}{N} \sum_{i=1}^N X_i$ $\sum_{i=1}^N P_i = 1$	$\hat{\mu} = \frac{1}{Nn} \sum_{i=1}^n \frac{\dot{X}_i}{P_i}$	$\text{Est. } V(\hat{\mu}) = \frac{s^2}{n}$ $s^2 = \frac{1}{N^2(n-1)} \sum_{i=1}^n \left(\frac{\dot{X}_i}{P_i} - N \hat{\mu}\right)^2$	$V(\hat{\mu}) = \frac{1}{n} \sigma^2$ $\sigma^2 = \frac{1}{N^2} \sum_{i=1}^N P_i \left(\frac{X_i}{P_i} - N \mu\right)^2$	$n = \frac{\sigma^2 f^2(\alpha, \infty)}{\epsilon^2 \mu^2}$
Muestreo bietápico con probabilidades variables de selección en cada etapa con reemplazo.	$\mu = \frac{1}{M_0} \sum_{i=1}^{N_i} M_i \mu_i$ $\mu_i = \frac{1}{M_{ti}} \sum_{j=1}^{m_{ij}} X_{ij}$ $M_0 = \sum_{i=1}^N M_i$ $\bar{M} = \frac{M_0}{N}$	$\hat{\mu} = \frac{1}{M_0 n} \sum_{i=1}^n \dot{X}_i$ $\dot{X}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} \dot{X}_{ij}$	$\text{Est. } V(\hat{\mu}) = \frac{s^2}{n}$ $s^2 = \frac{1}{n-1} \sum_{i=1}^n (\dot{X}_i - M_0 \hat{\mu})^2$	$V(\hat{\mu}) = \frac{1}{M_0^2 n} \left\{ \sum_{i=1}^N \frac{P_i}{m_i} \sigma_i^2 + \sigma^2 \right\}$ $\sigma^2 = \sum_{i=1}^N P_i \left(\frac{M_i \mu_i}{P_i} - M_0 \mu\right)^2$ $\sigma_i^2 = \sum_{j=1}^{m_i} \frac{P_{ij}}{P_i} \left(\frac{X_{ij}}{P_{ij}} - \frac{M_i \mu_i}{P_i}\right)^2$	$n = a(\alpha b + \sigma^2)$ $a = \frac{f^2(\alpha, \infty)}{M_0^2 \epsilon^2 \mu^2}$ $b = \sum_{i=1}^N \frac{P_i}{M_i} \sigma_i^2$ $m_i = \frac{M_i}{\alpha}$ $1 < \alpha < \frac{N}{\sigma^2} - \frac{\sigma^2}{\alpha}$

TABLA DE NUMEROS ALEATORIOS

16408	81899	04153	53381	79401	21438	83035	92350	36693	31238	59649
18629	81953	06520	91962	04739	13092	37662	24822	94730	06496	35090
73115	35101	47498	87637	99016	71060	88824	71013	38735	20286	23153
57491	16703	23167	49323	45021	33132	12544	41035	80780	45393	44812
30405	83946	23792	14422	15059	45799	22716	19792	09983	74353	68666
16631	35006	85900	98275	32388	52390	16815	69298	82732	38480	73817
96773	20206	42559	78985	05300	22164	24369	54224	35083	19687	11062
38935	64202	14349	82674	66523	44133	00697	35552	35970	19124	63318
31624	76384	17403	53363	44167	64486	64758	75366	76554	31601	12614
78919	19474	17403	27889	47914	02584	37680	20801	72152	39339	34805
03931	33309	57047	74211	63445	17361	62825	39908	05607	91284	68833
74426	33278	43972	10119	89917	15665	53872	73823	73144	88662	88970
09066	00903	20795	95452	92648	45454	09552	88815	16553	51125	79375
42238	12426	87025	14267	20979	04508	64535	31355	86064	29472	47689
16153	08002	26504	41744	81959	65642	74240	56302	00033	67107	77510
21454	40742	29820	96783	29400	21840	15035	34537	33310	06116	95240
21581	57802	02050	89728	17937	37621	47075	42080	97403	48626	68995
55612	78095	83197	33732	05810	24813	86902	60397	16489	03264	88325
44657	66999	99324	51281	84463	60563	79312	93454	68876	25471	93911
91340	84979	46949	81973	37949	61023	43997	15263	80644	43942	89203
01227	21199	31935	27022	84067	05462	35216	14486	29891	68607	41867
50001	38140	66321	19924	72163	09538	12151	06878	91903	18749	34405
65390	05224	72958	28609	81406	39147	25549	48542	42627	45233	57202
27504	96131	83944	10573	10573	08619	64482	73923	36152	05182	94143
37169	94851	39117	89632	00959	16487	65536	49071	39782	17095	02330
11508	70225	51111	38351	19444	66499	71945	05422	13442	78675	84081
37449	30362	06694	54690	04052	53115	62757	95348	78662	11163	81651
46515	70331	85922	36329	57015	15765	97161	17869	45349	61796	66345
30986	42223	42416	58353	21532	30502	32305	86482	05174	07901	54339
63798	64995	46583	09785	44160	78126	83991	42865	92520	83531	80377
82486	84846	99254	67632	43218	50076	21361	64816	51202	88124	41870
21885	32906	92431	09060	64297	51674	64126	62570	26123	05155	59194
60336	98782	07408	53458	13564	59089	26445	29789	85205	41001	12535
43937	46891	24010	25560	86355	33941	25786	54990	71899	15475	95434
07656	63175	89303	16275	07100	92063	21942	18611	47348	20203	18534
03299	01221	05418	36982	55756	92267	26759	86367	21216	98442	08343
79626	06486	03574	17668	07785	76020	79924	25651	83325	88428	85076
85636	68335	47539	03129	65651	11977	02510	26118	99447	68645	84016
18039	14367	61337	06177	12143	32989	74014	64768	00533	64768	00533
08362	15656	60627	36478	65648	16764	53412	09013	07832	41574	17639
79556	29060	04142	16268	15387	12856	66227	38358	22478	73373	88732
02608	82674	27072	32534	17075	27698	98204	63863	11951	34643	88732
23968	25835	40055	67006	12293	02753	14827	23235	35071	99704	34422
09915	96306	05908	97901	28395	14166	00821	80703	70426	75647	37353
80037	33300	26695	62247	69927	76123	50842	43834	86654	70959	76315
42488	78077	69882	61657	34136	79180	97526	43092	04098	73571	79725
46764	86273	63003	93017	31204	36692	40202	35275	57306	55543	50799
03237	45430	55417	63282	90816	17349	88298	90183	36600	78406	53265
80591	81482	52647	61582	14972	90053	89534	76036	49199	43716	06215
30834	01715	94964	67288	65680	43772	39560	12918	86537	62738	97546
										19636

### 3.1.3 Diseño del Cuestionario. \*

El diseño de un cuestionario involucra la consideración de un sinnúmero de aspectos diferentes, de los cuales quizás los más importantes son: los objetivos del estudio, la forma que debe tener, si contendrá preguntas abiertas (codificación previa o posterior de las preguntas), la forma como se harán las preguntas, la organización e instrucciones del cuestionario, etc.; lo que también indudablemente determina su diseño, es el tipo de datos que se desean obtener, el método usado para obtenerlos y en última instancia el uso de los resultados. Adicionalmente, podría señalarse que el diseño depende fuertemente de los antecedentes y experiencias del investigador, el tipo de enumeradores disponibles, costo y tiempo.

Así, basándose en los formatos de tabulación del guión de información, los rangos probables de variación tomados de las experiencias anteriores (si las hay) y las posibles respuestas, el cuestionario debe diseñarse en forma simple, fácil de seguir y si es posible atractiva. Lo último es particularmente importante en el caso de los cuestionarios que se envían por correo, donde la decisión de los miembros de la muestra, sobre llenarlo o no depende de la impresión que tengan sobre la apariencia del cuestionario. Al respecto, se aconseja recabar la información a través de entrevistas directas, ya que el enumerador puede inmediatamente captar los datos en forma precisa o corregirlos cuando el caso lo amerite.

#### 3.1.4 Trabajo de Campo.

Es conveniente mencionar que existen diversos métodos para la recolección de datos, de los cuales los principales son:

- a) La selección de la muestra a partir de la información de los archivos de la empresa. Así, una muestra puede ser escogida sin mayor problema y al mismo tiempo los datos pueden ser obtenidos con un alto grado de confianza a un costo relativamente bajo. Además de que la muestra puede mantenerse continuamente sin representar mayores cargos o esfuerzos extraordinarios.
- b) Métodos de observación.- La recolección de los datos por observación es otro instrumento que indirectamente capta la información. Como la información interna, este método no

\* *El desarrollo de los siguientes incisos del capítulo es producto de la experiencia de los autores y de la brillante exposición teórica de R. Farber y S. Hollander en su libro Marketing Research (4).*

requiere contacto directo con los elementos de la muestra. Estos métodos utilizan observadores humanos y/o mecánicos, prefiriendo los primeros en casos donde haya que distinguir, por ejemplo, los adultos de los niños, o las personas por sexo.

- c) Entrevistas telefónicas.- Cuando se puede aplicar este método, resulta altamente eficiente en la recolección directa de la información. Lo anterior se debe a que la población virtualmente está contenida en un directorio y la selección de la muestra se convierte en una actividad de rutina. Las entrevistas son de lo más económico (excepto cuando hayan que hacerse bastantes llamadas de larga distancia) y los datos se obtienen rápidamente. Sin embargo, como los demás métodos también tiene sus limitaciones. Obviamente no es aplicable si las entrevistas comprenden cuestiones visuales (publicidad, pruebas de interpretación, etc.). A la vez, información altamente personal se obtiene con menos éxito por teléfono que, por ejemplo, a través de una entrevista personal.
- d) Entrevistas personales.- Dentro de las formas directas de obtener los datos, este método es sin lugar a duda el más popular, por involucrar una conversación directa "frente a frente" entre un miembro de la muestra y el entrevistador. Como resultado, se puede obtener una gran variedad de información con este método, el cual es flexible en varios sentidos. Por ejemplo, los datos pueden ser registrados en grabadora o en cuestionarios.

La construcción del cuestionario es un arte en sí, requiere numerosas precauciones para evitar respuestas sesgadas.

Desde el punto de vista de la obtención de los datos, puede decirse que existen dos formas de entrevistar. En un extremo se halla la entrevista altamente estructurada, en la cual se prepara un cuestionario formal y las preguntas se hacen bajo instrucciones precisas y el enumerador mantiene un orden estricto para su contestación. Esta forma se usa generalmente para obtener una variedad de información diferente acerca de una materia, siguiendo algún orden particular. Esta forma en cierto modo, evita que la información recabada refleje sesgos debidos a juicios personales de los

enumeradores.

En el otro extremo está la entrevista carente de formalidad para la cual no se requiere un cuestionario, basta una lista de preguntas generales o temas relacionados con la información que se busca.

Dentro de estos extremos existen varias combinaciones. El enumerador puede usar un cuestionario estructurado, pero se le permite hacer las preguntas como él quiera.

Como podrá intuirse, el enumerador es la piedra angular de una entrevista, indistintamente de la forma que adopte para entrevistar o cuál sea la unidad de muestreo. Si está debidamente entrenado (a), no solamente entrevistará a un mayor número de personas, sino que los datos serán más confiables.

Parece que los mejores enumeradores son personas entre 25 y 50 años que tienen una evidente disposición, son inteligentes, poseen cierta cultura, son flexibles y precisos en sus hábitos de trabajo. Indudablemente que la experiencia es útil, pero si se proporciona un buen entrenamiento puede no ser necesaria. En ciertos tipos de nuevas encuestas la experiencia puede ser una limitante, ya que se requiere que el enumerador siga procedimientos contrarios a los acostumbrados en el pasado.

Por lo que se refiere a la organización y control del trabajo de campo, como las demás etapas requiere una programación de tiempos y actividades para asignar al personal correspondiente. Dentro de los aspectos básicos está la fijación de las rutas de trabajo, el plan de trabajo o forma de entrevistar y la supervisión (sobre todo cuando el grupo de trabajo es numeroso o la captación de los datos presente dificultades).

### 3.1.5 Crítica de Cuestionarios.

Los cuestionarios, codificados o no previamente, llegan a la oficina con el orden y presentación de las respuestas dadas por los enumeradores. En algunas ocasiones el trabajo se realiza de acuerdo con las instrucciones establecidas y en seguida pasa al Departamento de Perforación para ser procesado inmediatamente. Sin embargo, en la mayoría de los casos se requiere una crítica o revisión cuidadosa ya que:

a) Pueden traer las respuesta ilegibles.

b) El orden en que aparecen las respuestas no es el indicado.

c) Se contradicen unas respuestas con otras al compararse entre sí.

d) Existen preguntas que vienen en blanco y debían haberse contestado en alguna u otra forma, etc.

e) Se requiere preparar los cuestionarios para la codificación de las respuestas.

f) Se desea verificar la autenticidad de los datos y preliminarmente, comprobar ciertas hipótesis establecidas en la programación inicial de actividades, etc.

Tal que en esta etapa, la información debe quedar depurada y ordenada hasta donde sea posible para su posterior transformación y vaciada en formatos previamente diseñados. En algunos casos se acostumbra utilizar la computadora (filtrado electrónico) para realizar esta etapa.

### 3.1.6 Codificación y Procesamiento de Datos.

Una vez que los datos han sido obtenidos y revisados, deben ser procesados para hacer posible un análisis del fenómeno estudiado. Es generalmente aceptado que esta actividad es un tanto tediosa, pero también que es crítica para asegurar exactitud en los resultados.

Una tabulación hecha sin cuidado puede viciar una buena planeación y el método de obtención de los datos. Asimismo, los peligros de los sesgos aún se presentan en los procesos de preparación, clasificación y tabulación.

Esta etapa está fuertemente ligada a la anterior, ya que, por ejemplo, la preparación consiste en la inspección de cuestionarios (o cualquier otra forma usada para captar los datos), su exactitud, si están completos o no, la inspección del trabajo de campo, arreglos o eliminación de respuestas por su inconsistencia o desconianza, la clasificación o estandarización de los datos en bases comunes y sobre todo su preparación para ser tabulados.

## Clasificación.

Es el arreglo de los datos en clases o categorías para ser manipulados.

## Tabulación.

La tabulación es la etapa que sucede inmediatamente después a la crítica de cuestionarios y es un conjunto de procedimientos que se adoptan para la recopilación o vaciado de los datos en cuadros. Estos últimos comprenden las diferentes relaciones que se establecen entre las variables comprendidas en el estudio, así, habrá cuadros de una sola entrada, doble entrada, etc.

Los datos pueden ser tabulados manual o mecánicamente. La tabulación manual se aconseja cuando las encuestas son pequeñas, existen problemas de presupuesto o no hay ninguna posibilidad de procesar los datos electrónicamente. Por el contrario, cuando la encuesta es grande, la tabulación manual, además de tardada, acarrea el riesgo de cálculos erróneos por lo voluminoso de la información, aconsejándose el uso de las computadoras. Por ello, será necesario que la información sea perforada y se diseñen los programas que calcularán los datos de acuerdo con instrucciones específicas.

Sobre estas cuestiones debe decirse que aun cuando los datos ya están listos para procesarse y la maquinaria realiza sus operaciones a velocidades fantásticas, la realidad es que las horas-máquina gastadas en hacer los programas y/o prepararlos son varias veces los minutos usados en la computación. Así, los más avanzados métodos aún requieren erogaciones fuertes de dinero para poder aplicarse. Para Estudios muy Grandes, estudios repetitivos, donde la programación se hace una sola vez y para investigación de operaciones o problemas de programación lineal, estos aparatos electrónicos si proporcionan economías comparados con otros métodos.

### 3.1.7 Evaluación Estadística de Resultados.

El análisis de los datos recabados con la muestra, incluye indicaciones del valor hasta el cual las estimaciones derivadas de la muestra pueden desviarse de los valores verdaderos de la población. Esta evaluación debe comprender datos sobre la precisión de los estimadores, sobre todo si la

selección ha sido probabilística, así como consideraciones sobre algunos sesgos en la operación de reconocimiento que tienda a distorsionar el valor de los estimadores. Dentro de los sesgos puede considerarse las "no respuestas", cobertura, influencia de los enumeradores sobre la unidad de muestreo entrevistado y lo que anoten en el cuestionario, una codificación de respuestas inadecuada, etc.

Por lo que se refiere a la precisión, ésta se refiere al error de muestreo de un estimador. Mientras más pequeño sea el error, mejor será la precisión. El error de muestreo se mide con la fórmula del error estándar, la cual variará de acuerdo con el tipo de estimador (media, mediana, razón, etc.) y con el diseño muestral.

La exposición de las fórmulas de los errores estándar se presenta en la sección de los métodos de muestreo (3.1) donde se deducen de las variancias de los estimadores (media, total, etc.).

## Descripción de Actividades.

### 1-2 Definición General de Objetivos.

Debe exponerse el problema, con todo el detalle que sea posible, y además establecer el monto total de los fondos disponibles para realizar la encuesta.

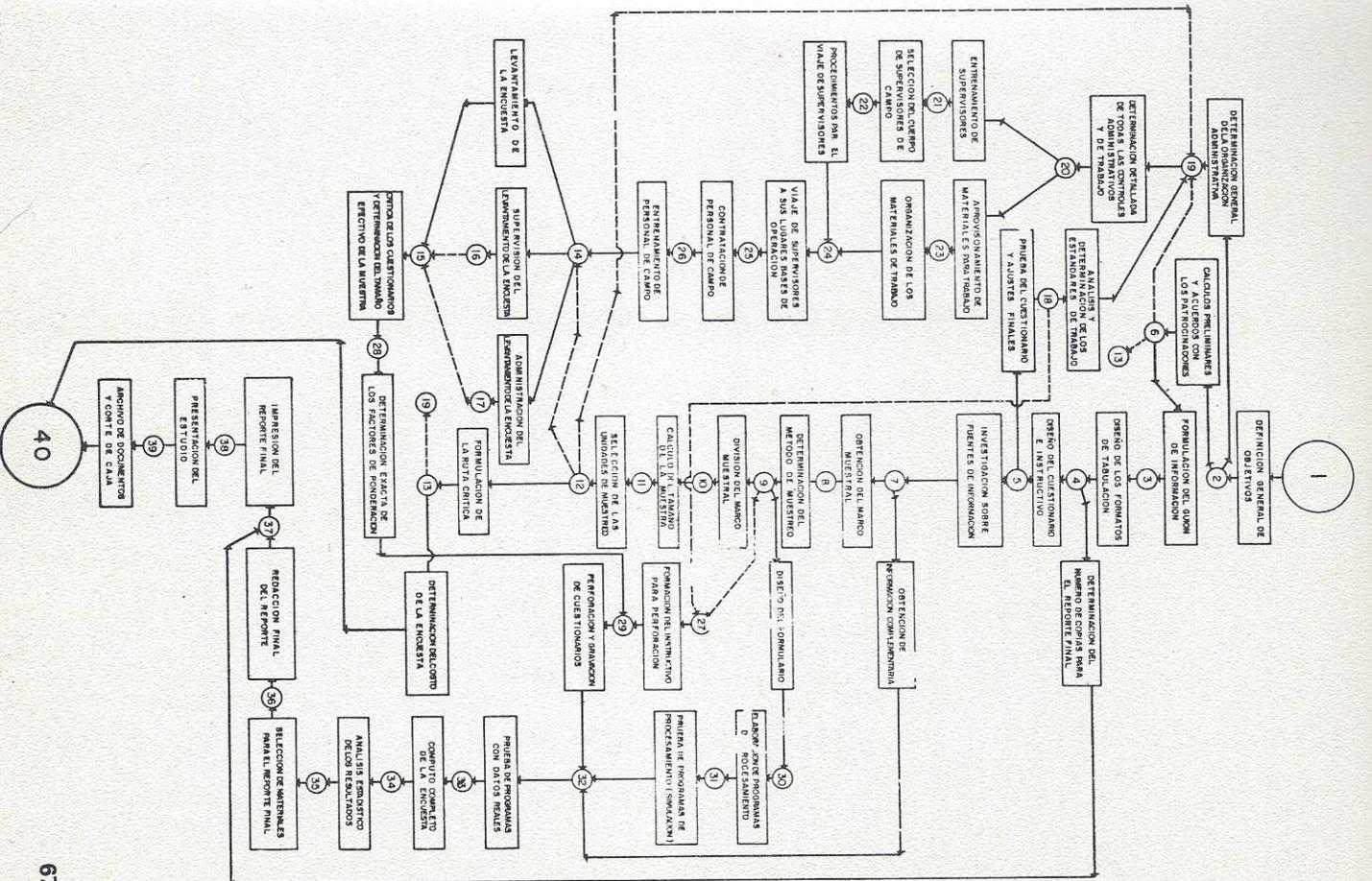
### 2-6 Cálculos Preliminares.

En base de las experiencias de otras encuestas, y de los requerimientos de la presente, háganse las estimaciones en forma preliminar, del tamaño de la muestra y del costo total de la encuesta, para poder llegar a un acuerdo con los patrocinadores, en cuanto a la precisión con que se debe trabajar y de todas las obligaciones que la presente encuesta implique en ambas partes.

### 2-19 Determinación General de la Organización Administrativa.

En esta parte, deben constituirse los departamentos en oficinas centrales de: Personal, Aprovechamiento y Finanzas.

3.1.8 RED GENERAL DE ACTIVIDADES EN UNA ENCUESTA DE MUESTREO



3-4 Diseño de los Formatos de Tabulación.

Los Requerimientos de información y las relaciones significativamente importantes, deben exhibirse en estos formatos con claridad y sencillez, dado que con el éxito que esto se logre, la solución del problema será más convincente y fácil. Deben definirse aquí los títulos de todos los cuadros.

4-5 Diseño del Cuestionario e Instructivo.

Basándose en los formatos de tabulación, del guión de información, de los rangos probables de variación, de las experiencias anteriores y de las posibles respuestas de las preguntas, hágase el diseño de un cuestionario precodificado, procurando y verificando que no se omita ningún concepto, que el llenado del cuestionario, sea lo más sencillo y rápido posible, que el encadenamiento de las preguntas sea el más adecuado, que algunas preguntas sirvan para comprobar las respuestas de otras, etc. Un cuestionario precodificado asigna en cada pregunta un conjunto de claves numéricas, correspondiendo en forma biunívoca, con el conjunto de las posibles respuestas, estas claves se anotan cifra por cifra, en las posiciones (en cuadrícula) que se hayan designado para el caso (véase punto 3.1.3).

4-37 Determinación del Número de Copias para el Reporte Final.

5-7 Investigación sobre Fuentes de Información.

Un marco muestral, es un conjunto de listas o de mapas, o una combinación de estos elementos, de tal manera, que todas las unidades de interés estén contenidas y que al seleccionar la muestra se pueda determinar la probabilidad de su inclusión, asimismo, en el momento de levantar la encuesta, la identificación de cada unidad en la muestra sea posible hacerla sin ninguna ambigüedad.

Para obtener un marco muestral puede recurrirse a ciertas instituciones y recopilar además, datos para: calcular el tamaño de la muestra, confrontar y complementar los resultados de la encuesta, determinar aproximadamente algunos rangos de variación, etc., si es que en los antecedentes (archivos propios) no se tienen.

5-18 Prueba del Cuestionario y Ajustes Finales.

Con objeto de determinar cuáles ajustes deben hacerse al cuestionario, para poder lograr los objetivos en forma satisfactoria, es necesario realizar algunas entrevistas en el campo de estudio, llenar los cuestionarios correspondientes y evaluar los resultados a este nivel.

6-3 Formulación del Guión de Información.

Partiendo de un examen del problema, hágase una relación de todas las variables, cuyos valores puedan ser significativamente relevantes, en la resolución del problema, asimismo, determinense qué relaciones son convenientes: variancias, regresiones, elasticidades, coeficientes de correlación, etc.

7-8 Obtención del Marco Muestral.

Véase actividad 5-6.

7-32 Obtención de Información Complementaria.

Prepárese todo el material que sea necesario, como oficinas debidamente dirigidos y firmados, formas para captar información, etc.

8-9 Determinación del Método de Muestreo.

Los métodos de muestreo tienen por objeto indicar el número de unidades que deben incluirse en la muestra, dependiendo de la forma como éstas se seleccionen, del nivel de confianza que se requiera, del error de muestreo permisible y del fondo disponible para la realización de la encuesta (ver puntos 3.1.1, pág. 5; 3.1.2, pág. 8).

9-10 División del Marco Muestral.

De acuerdo con las formas de realizar un muestreo se han desarrollado modelos, sin embargo, en algunas ocasiones la constitución del marco muestral obliga a utilizar varios modelos de muestreo, cubriendo cada uno, aquella parte en donde sea homogéneo, tanto en agrupamientos de unidades, como en información y/o cualidades entre las unidades de muestreo (ver actividad 13-15).

9-30 Diseño del Formulario.

Incluye fórmulas para estimación de medias poblacionales, totales, regresiones, correlaciones, elasticidades, ajustes de curvas, etc.

10-11 Cálculo del Tamaño de la Muestra.

Ver actividad 8-9.

11-12 Selección de las Unidades de Muestreo.

Ver actividad 8-9.

12-13 Formulación de la Ruta Crítica.

Tómese como base la presente red de actividades.

13-41 Determinación del Costo de la Encuesta.

14-15 Levantamiento de la Encuesta.

El trabajo de los enumeradores debe hacerse exactamente con las unidades de última etapa, determinadas en la selección de la muestra y si ello no fuera posible por deficiencias en el marco muestral, resúlvase el problema con apego a las instrucciones precisas que se hayan hecho para estos casos. Al hacerse las preguntas, téngase cuidado de que las respuestas sean correctas y veraces, considerando los rangos aproximados para los valores que puedan tomar las variables involucradas en el estudio (ver inciso 3.1.4, pág. 60).

14-16 Supervisión del Levantamiento de la Encuesta.

Es conveniente utilizar una forma de reporte, en la cual el supervisor anote cómo se desarrolla el levantamiento de la encuesta, esto es, registrar el material recibido y entregado, folio de los cuestionarios entregados a su grupo, casos de no respuesta y especificación de la resolución tomada, folio de los cuestionarios que fue necesario aclarar, número diario de cuestionarios entregados y de errores por enumerador, porcentaje del avance total del trabajo (llenado de cuestionarios), día y hora para cada reporte a oficinas centrales, números de cuestionarios efectivamente llenados

al terminar la encuesta y registro de los demás documentos recogidos, calificación final de los enumeradores, etc. (ver inciso 3.1.4, pág. 60).

14-17 Administración del Levantamiento de la Encuesta.

Se refiere a todas las actividades como:

Autorización de gastos y obtención de fondos junto con las directrices administrativas para su uso, acuse de lo recibido a oficinas centrales, pago del trabajo de campo, observación del sistema de envíos, tiempos transcurridos entre envío y recepción, condición de llegada del material, retroalimentación de las experiencias de la fase inicial y ajuste en donde ello sea necesario, registro de aquellos procedimientos (o personas) que no funcionaron para referencias futuras y para obtener de ello una experiencia, terminación de obligaciones con el personal eventual, etc.

15-28 Crítica de los Cuestionarios y Determinación del Tamaño Efectivo de la Muestra.

Esto es hacer un filtrado de todos los errores que no hayan sido detectados por los supervisores, así como también verificar y concentrar el número total de cuestionarios encomendados a cada supervisor, para obtener el tamaño efectivo de la muestra (ver punto 3.1.5, pág. 62).

18-19 Análisis y Determinación de los Estándares de Trabajo.

Basándose en el trabajo realizado, al probar el cuestionario y en experiencias anteriores, determinése el número de cuestionarios por individuo y por día como cargo de trabajo, número de visitas antes de declarar la no respuesta, mínimo esencial para considerar un cuestionario completo, definición aproximada de los rangos de variación para algunas variables, etc.

19-20 Determinación Detallada de Todos los Controles Administrativos y de Trabajo.

Elaboración de todos los mecanismos relativos a procedimientos de pago, métodos de retiro de fondos, de adquisición de materiales, de órdenes de trabajo, autorizaciones necesarias, procedimientos para pago de impuestos, definición de obligaciones por ambas partes, procedimientos de envíos, etc.

Elaboración del reglamento de gastos, establecimiento de un calendario de pagos.

Redacción de formas para: reportes del avance del trabajo, registro de personal (con número de clave del registro federal de causantes), credenciales de identificación, contratos, etc.

Establecimiento de los requisitos del personal de campo: escolaridad, disponibilidad de tiempo, sexo, número de ellos en las diversas plazas y zonas de la encuesta, preferencia de contratar maestros, trabajadoras sociales, enfermeras, etc.

20-21 Entrenamiento de Supervisores.

En todas las actividades, tanto administrativas, como de trabajo de campo.

21-22 Selección del Cuerpo de Supervisores de Campo.

22-24 Procedimientos para el Viaje de Supervisores.

Habilitación de supervisores de campo con viáticos, vehículos, boletos, credenciales, cartas de presentación debidamente dirigidas y firmadas, etc.

20-23 Aprovisionamiento de Materiales para Trabajo.

Como cuestionarios, instructivos, fotografías, planos, cintas métricas, etiquetas, ligas, requisiciones para vehículos, gasolina, papelería, tabletas para escribir en el campo, lápices, gomas, plumones, tarjetas con direcciones importantes y teléfonos, instrucciones para casos de emergencia, formas para control y registro del material, para el avance del trabajo, etc.

23-24 Organización de los Materiales de Trabajo.

En paquetes para las diversas zonas y unidades, ¿a quién van dirigidos? controles de entrada y salida de documentos.

24-25 Viaje de Supervisores a sus Lugares Base de Operación y Envío de Direcciones y Teléfonos a Oficinas Centrales.

- 25-26 Contratación de Personal de Campo.
- 26-14 Entrenamiento de Personal de Campo.
- 27-29 Formación del Instructivo para Perforación.  
Esto es, definición de los campos en donde se perforará cada dato del cuestionario.
- 28-29 Determinación Exacta de los Factores de Ponderación.
- 29-32 Perforación y Gravación de los Cuestionarios.
- 30-31 Elaboración de Programas para Cómputo.  
Deben tenerse en cuenta, todas las tabulaciones y los análisis especiales, así como variancias, errores estándar, los formatos de salida, títulos, etc.
- 31-32 Prueba de los Programas para Procesamiento (simulación).  
Con objeto de verificar si los programas funcionan en la forma que se pretende, alimentense con datos ficticios, fáciles de procesar manualmente para comparar resultados.
- 32-33 Prueba de Programas con Datos Reales.  
Antes de procesar todos los datos, es conveniente procesar algunos y hacer comparaciones con cifras de control, para hacer los ajustes finales, tanto de tipo operativo como, de formatos de impresión.
- 33-34 Cómputo Completo de la Encuesta.
- 34-35 Análisis Estadístico de los Resultados.  
Partiendo de una evaluación de la no respuesta y la no cobertura, los intervalos de confianza y los coeficientes de variación se puede obtener un criterio, acerca de la validez de las estimaciones hechas a partir de la muestra, de tal manera que el confrontar algunos resultados de la encuesta, con cifras de otras fuentes, se podrá determinar el origen de las posibles

diferencias significativas, como coberturas distintas, error por no respuesta, error de muestreo, etc.

Este análisis debe tenerse en cuenta para el reporte final (ver punto 3.1.7, pág. 64).

- 35-36 Selección de Materiales para el Reporte Final.
- 36-37 Redacción Final del Reporte.
- 37-38 Impresión del Reporte Final.
- 38-39 Presentación del Estudio.
- 39-40 Archivo de Documentos.

Envío del reporte a la biblioteca para su conservación y clasificación. Destrucción de cuestionarios y otros materiales no utilizables. Control sobre estas operaciones. Felicitaciones al personal involucrado. Cierre de libros y disposición final de fondos. Terminación oficial de la encuesta.

#### 4. APLICACIONES A LA INDUSTRIA PETROLERA.

- 4.1 Ejemplo Hipotético para Determinar el Tamaño de una Muestra y su Respectiva Selección.  
Se desea captar las características relevantes del universo utilizando el muestreo estadístico.

El universo está compuesto por los elementos que aparecen en la siguiente tabla, donde las máquinas se encuentran numeradas en cada estrato y sector.

De acuerdo con la composición del universo se observa que la captación de dichas características puede realizarse a través del modelo Bietápico Equiprobable sin reemplazo y estratificado.

Lo anterior significa que la captación de los datos podría realizarse según los rangos de producción (estratos) donde la unidad de primera etapa está constituida por los sectores y la unidad de segunda etapa está consti-

tuida por las máquinas de cada sector y la selección de cada una de las unidades se realizará sin reemplazo para obtener mejores estimaciones.

RANGO DE PRODUCCION	NUMERO DE MAQUINAS POR SECTOR			PRODUCCION POR SECTOR		
	1	2	3	1	2	3
0 - 100	9	13	3	430	250	181
101 - 200	14	32	5	1,867	4,570	675
201 - 300	6	7	5	1,371	1,525	1,143
301 - 400	9	11	1	3,119	3,553	347
401 - 500	4	4	10	1,721	1,821	4,389
501 - 600	4	3	2	2,082	1,554	1,070
601 - 700	0	6	0	0	3,802	0
701 - 800	5	6	5	3,710	4,480	3,636
801 - 900	5	4	5	4,207	3,278	4,142
901 - 1000	9	5	3	8,472	4,617	2,810

### Cálculo del Tamaño de la Muestra.

En base al desarrollo del Modelo en la tabla, el tamaño de la muestra está dado por:

$$n_t = a_t (S_t^2 + b_t (\alpha_t - 1))$$

$$a_t = \frac{K N_t M_0^2 + t^2 (\alpha, \infty)}{N_t M_0^2 \epsilon^2 \mu^2 + K M_0^2 S_t^2 t^2 (\alpha, \infty)}$$

$$b_t = \frac{1}{N_t M_t} \sum_{i=1}^{N_t} M_{t_i} S_{t_i}^2$$

$$m_{t_i} = \frac{M_{t_i}}{\alpha_t}$$

$$1 \leq \alpha_t \leq \frac{N_t}{a_t b_t} + 1 - \frac{S_t^2}{b_t}$$

$$S_t^2 = \frac{1}{N_t - 1} \sum_{i=1}^{N_t} \left( \frac{M_{t_i}}{M_t} \right)^2 (\mu_{t_i} - \mu_t)^2$$

$$S_{t_i}^2 = \frac{1}{M_{t_i} - 1} \sum_{j=1}^{M_{t_i}} (X_{t_i j} - \mu_{t_i})^2$$

El significado de las literales está contenido en la tabla de nomenclaturas, pág. 14. La aplicación de esta fórmula da los siguientes resultados:

RANGO DE PRODUCCION	VALOR DE $S_{t_i}$ (por sector)	VALOR DE $S_t$	$\frac{N_t}{a_t b_t} + 1 - \frac{S_t^2}{b_t}$	$\alpha_t$	
0 - 100	1,228	829	3	239	11.59
101 - 200	787	961	84	13,780	6.48
201 - 300	387	53	160	1,028	67.15
301 - 400	1,340	426	0	61,714	14.81
401 - 500	840	525	477	63,609	23.60
501 - 600	170	4	375	28,473	161.47
601 - 700	0	846	0	0	0
701 - 800	823	1,605	341	7,671	15.97
801 - 900	1,306	2	437	12,349	27.52
901 - 1000	728	361	162	260,433	27.14

Resultando los siguientes tamaños de muestra para las unidades de primera y segunda etapa.

RANGO DE PRODUCCION	$n_i$	$m_{ij}$		
		1	2	3
0 - 100	1	1	1	1
101 - 200	3	2	4	1
201 - 300	1	1	1	1
301 - 400	3	1	1	1
401 - 500	2	1	1	1
501 - 600	1	1	1	1
601 - 700	1	0	2	0
701 - 800	1	1	1	1
801 - 900	1	1	1	1
901 - 1000	3	1	1	1

El siguiente paso consiste en la selección de las máquinas en cada uno de los sectores en forma aleatoria, para lo cual se utiliza la metodología expuesta en la pág.

El resultado de esta selección está expuesto en la siguiente tabla.

RANGO DE PRODUCCION	No. DE MAQUINA POR SECTOR PARA LA MUESTRA			
		1	2	3
0 - 100	10	6		1
101 - 200	6	4, 7, 11, 20		
201 - 300	9	5		4
301 - 400	1	3		
401 - 500			3	2
501 - 600			6	
601 - 700	4			
701 - 800				2
801 - 900	9	5		3
901 - 1000				

Calculados e identificados los elementos (máquinas) a muestrear, se procede al levantamiento de datos (siguiendo las instrucciones de la red de actividades), los que posteriormente son criticados, perforados, procesados, analizados e interpretados; siendo la fase final la evaluación estadística de las estimaciones.

Con el propósito de abreviar cálculos que suelen llevarse basta tiempo, se presenta a continuación un programa de computadora para el modelo el cual, previas definiciones, sirve también para el cálculo de modo monoatómico estratificado o no estratificado, sin reemplazo.

4.2

Uso del Muestreo Estadístico en el Análisis Estratificado Mercado de Lubricantes en Nuevo León, Querétaro y Sinaloa.

Objetivo:

1. Cuantificar la demanda de los lubricantes automotrices industriales.
2. Identificar las características de la comercialización abastecimiento.
3. Conocer la publicidad desplegada por Petróleos Mexicanos la competencia.

Para ello se diseñaron 3 cuestionarios: uno dirigido a las industrias otro para las estaciones de servicio y talleres autorizados y expendios lubricantes, y el tercero para automóviles, camiones de pasaje, camiones carga, tractores y motocicletas, los que a continuación se presentan.

```

PROGRAM MLESTREC
INTEGER T,GRUPO,C1
REAL M,N,M0,M00,N1, M100
COMMON X(20,20),M(20,20),Y2(20,20),X3(20),N(20),M0(20),SB2(20),S2
1(20,20),A(20),B(20),ALFA(20),ALFA1(20,20),ALFA2(20),N1(20,20)
2,M1(20,20,20),AREA(13),M100(20),B1(20,20),E2(10)
K=10 & L=3 & M00=0. & C=100. & C1=13. & L1=10 & E=0.1 & TE=1.64
C0 MLESTREC BIETAPICO,ESTRATIFICADO CON PROBABILIDADES IGUALES DE SELECCION
C00 SIN REEMPLAZC.
C1 X(T,I)=AREA TOTAL DEL I-ESIMO GRUPO DEL T-ESIMO ESTRATO. M(T,I)=NUM TOTAL UNI
C2 DADES DEL T-ESIMO GRUPO Y DEL T-ESIMO ESTRATO. X2(T,I)= SUMA DE LOS CUADRADO
C3 S DE LAS AREAS DEL T-ESIMO GRUPO Y DEL I-ESIMO ESTRATO. X3(T)= AREA TOTAL EN
C4 EL T-ESIMO ESTRATO. N(T)= NUM TOTAL UNIDADES DE PRIMERA ETAPA DEL T-ESIMO
C5 ESTRATO. M0(T)=NUM TOTAL UNIDADES DE SEGUNDA ETAPA DEL T-ESIMO ESTRATO.
C6 SB2(T)=ERROR CUADRATICO MEDIO ENTRE GRUPOS DEL T-ESIMO ESTRATO.S2(T,I)=ERROR
C7 CUADRATICO MEDIO DENTRO DE LA I-ESIMA UNIDAD DEL T-ESIMO ESTRATO. ALFA1(T,I)
C8 = I-ESIMO VALOR DE ALFA EN EL T-ESIMO ESTRATO. N1(T)=NUMERO DE UNIDADES DE
C9 PRIMERA ETAPA PARA LA MUESTRA EN EL I-ESIMO ESTRATO. M1(T,I)=NUMERO DE UNI-
C10 DADES DE SEGUNDA ETAPA DE LA I-ESIMA UNIDAD DEL ESTRATO T-ESIMO. K=NUM ESTRA
C11 TOS. L= NUM GRUPOS. M00= NUM TOTAL UNIDADES DE SEGUNCA ETAPA. C= AMPLITUD
C12 DE LOS ESTRATOS. C1= NUM DE CAMPOS (PARA AREAS) POR TARJETAS.
10 READ 15,GRUPO,(AREA(I),I=1,C1)
GO TO (40,30)EOFCKF(60)
30 I=GRUPO
DC 39 I1=1,C1
T=AREA(I1)/C+1.
X(T,I)=X(T,I)+AREA(I1)
M(T,I)=M(T,I)+1.
X2(T,I)=X2(T,I)+AREA(I1)**2
X3(T)=X3(T)+AREA(I1)
X4=X4+AREA(I1)
IF (AREA1.LT.AREA(I1))AREA1=AREA(I1)
39 CONTINUE
GO TO 10
40 K=AREA1/C+1.
DC 60 T=1,K
DC 60 I=1,L
IF (M(T,I).EQ.0) GO TO 60
N(T)=N(T)+1.
M0(T)=M0(T)+M(T,I)
M00=M00+M(T,I)
60 CONTINUE
DC 90 T=1,K
E2(T)=(E*M00)**2/(K*M0(T)**2)
J=N(T)
DC 100 I=1,L
IF (M(T,I).EQ.0.OR.M(T,I).EQ.1) 63,64
63 S2(T,I)=0 & GO TO 99
64 S2(T,I)=1./((M(T,I)-1.)*(X2(T,I)-X(T,I)**2/M(T,I)))
B1(T,I)=M(T,I)*S2(T,I)
99 SB2(T)=SB2(T)+(N(T)*X(T,I))**2
100 CONTINUE
IF (J.EQ.0.OR.J.EQ.1) 61,62
61 SB2(T)=0 & GO TO 65
62 SB2(T)=1./(((N(T)-1.)*M0(T)**2)*(SB2(T)-N(T)*X3(T)**2))
65 PRINT 13,SB2(T),N(T),M0(T),X3(T),X4
13 FORMAT (10X,5F10.)
90 CONTINUE
DO 11 T=1,K
11 PRINT 12,(M(T,I),X(T,I),X2(T,I),S2(T,I),B1(T,I),I=1,3),SB2(T),N(T)
12 FORMAT(X,3(F3.,F6.,F12.,F8.,F9.),F10.,F3.)
DO 120 T=1,K
J=N(T)
IF (J.EQ.0.OR.J.EQ.1) 200,201
200 A(T)=B(T)=ALFA(T)=0
GO TO 120
201 DO 130 I=1,L
A(T)=M0(T)**2*K*N(T)*TE**2/(M00**2*N(T)*E**2*(X4/M00)**2+M0(T)**2
1*K*TE**2*SB2(T))
B(T)=M(T,I)*S2(T,I)+B(T)
130 CONTINUE
B(T)=N(T)/M0(T)**2*B(T)
XL1=L1-1
ALFA(T)=(N(T)/(A(T)+B(T))-SB2(T)/B(T))/XL1
DC 120 J=1,L1
IF (ALFA(T).EQ.0) 121,122
121 ALFA1(J,T)=0. & GO TO 120
122 ALFA1(J,T)=ALFA(T)*(J-1)+1

```

```

39 CONTINUE
   GO TO 10
40 K=AREA1/C+1.
   DC 60 T=1,K
   DC 60 I=1,L
   IF (M(T,I).EQ.0) GO TO 60
   N(T)=N(T)+1.
   M0(T)=M0(T)+M(T,I)
   M00=M00+M(T,I)
60 CONTINUE
   DC 90 T=1,K
   E2(T)=(E*M00)**2/(K*M0(T)**2)
   J=N(T)
   DC 100 I=1,L
   IF (M(T,I).EQ.0.OR.M(T,I).EQ.1) 63,64
63 S2(T,I)=0 $ GO TO 99
64 S2(T,I)=1./(M(T,I)-1.)*(X2(T,I)-X(T,I)**2/M(T,I))
   B1(T,I)=M(T,I)*S2(T,I)
99 SB2(T)=SB2(T)+(N(T)*X(T,I))**2
100 CONTINUE
   IF (J.EQ.0.OR.J.EQ.1) 61,62
61 SB2(T)=0 $ GO TO 65
62 SB2(T)=1./((N(T)-1.)*M0(T)**2)*(SB2(T)-N(T)*X3(T)**2)
65 PRINT 13,SB2(T),N(T),M0(T),X3(T),X4
13 FORMAT (10X,5F10.)
90 CONTINUE
   DO 11 T=1,K
11 PRINT 12,(M(T,I),X(T,I),X2(T,I),S2(T,I),B1(T,I),I=1,3),SB2(T),N(T)
12 FORMAT(X,3(F3.,F6.,F12.,F8.,F9.),F10.,F3.)
   DC 120 T=1,K
   J=N(T)
   IF (J.EQ.0.OR.J.EQ.1) 200,201
200 A(T)=B(T)=ALFA(T)=0
   GO TO 120
201 DC 130 I=1,L
   A(T)=M0(T)**2*K*N(T)*TE**2/(M00**2*N(T)*E**2*(X4/M00)**2+M0(T)**2
1*K*TE**2*SB2(T))
   B(T)=M(T,I)*S2(T,I)+B(T)
130 CONTINUE
   B(T)=N(T)/M0(T)**2*B(T)
   XL1=L1-1
   ALFA(T)=(N(T)/(A(T)+B(T))-SB2(T)/B(T))/XL1
   DC 120 J=1,L1
   IF (ALFA(T).EQ.0) 121,122
121 ALFA1(J,T)=0. $ GO TO 120
122 ALFA1(J,T)=ALFA(T)*(J-1)+1
120 CONTINUE
   DC 181 J=1,L1
   DC 181 T=1,K
   DC 180 I=1,L
   IF (N(T).EQ.0) 250,251
250 N1(J,T)=M1(J,T,I)=0
   GO TO 180
251 IF (N(T).EQ.1) 252,253
252 N1(J,T)=1
   IF (S2(T,I).EQ.0) 254,255
254 M1(J,T,I)=0
   GO TO 180
255 M1(J,T,I)=M0(T)*TE **2*S2(T,I)/(M0(T)*E2(T)*X3(T)/M0(T)+TE **2*S
12(T,I))
   GO TO 179
253 N1(J,T)=A(T)*(SB2(T)+B(T)*(ALFA1(J,T)-1))
   IF (N1(J,T).LT.1.)N1(J,T)=1.
   M1(J,T,I)=M(T,I)/ALFA1(J,T)
   IF (M(T,I).LE.ALFA1(J,T))M1(J,T,I)=1
   GO TO 180
179 M100(J)=M100(J)+M1(J,T,I)
   PRINT 951,M100(J),M1(J,T,I)
951 FORMAT (100X,F10.2,I7)
180 CONTINUE
   M100(J)=M100(J)+M0(T)/(A(T)+B(T)*N(T))*N1(J,T)
   PRINT 950,M100(J),M0(T),ALFA1(J,T),N(T),N1(J,T)
950 FORMAT(50X,5F10.2)
181 CONTINUE
   DC 700 J=1,L1
   PRINT 602,M100(J)
   DC 700 T=1,K
   PRINT 603,ALFA1(J,T),N1(J,T),(M1(J,T,I),I=1,L)
700 CONTINUE

```

INDUSTRIAS

INSTITUTO MEXICANO DEL PETROLEO  
SUBDIRECCION DE ESTUDIOS ECONOMICOS Y PLANEACION INDUSTRIAL

"ENCUESTA NACIONAL SOBRE LA DEMANDA DE LUBRICANTES INDUSTRIALES"

No. Folio \_\_\_\_\_

Entidad \_\_\_\_\_ Municipio \_\_\_\_\_ Localidad \_\_\_\_\_

Razón Social \_\_\_\_\_ Domicilio \_\_\_\_\_ Tel. \_\_\_\_\_  
calle, carretera, etc. número

Giro \_\_\_\_\_ No. de empleados \_\_\_\_\_

I. ¿Cuál es el principal tipo de lubricantes que usa?

Marca \_\_\_\_\_ Tamaño(s) \_\_\_\_\_ Lts. Precio \_\_\_\_\_

¿Varía el precio? \_\_\_\_\_ Precio máximo \_\_\_\_\_ Precio mínimo \_\_\_\_\_

¿Obtiene precios especiales? \_\_\_\_\_ Regularmente \_\_\_\_\_ Por temporada \_\_\_\_\_

Cantidad y frecuencia con que los compra. Cantidad \_\_\_\_\_ Frecuencia \_\_\_\_\_  
(Litros) (Semanas)

¿Cuáles son las ventajas y desventajas de su uso?

Ventajas \_\_\_\_\_

Desventajas \_\_\_\_\_

\_\_\_\_\_ Lugar y fecha  
 \_\_\_\_\_ Nombre del Enumerador  
 \_\_\_\_\_  
 \_\_\_\_\_ OBSERVACIONES:

V. Ventas totales de la empresa \$ \_\_\_\_\_  
 ¿Aceptaría probar nuestras marcas?  ¿Por qué? \_\_\_\_\_  
 ¿Su equipo podría utilizar lubricantes de Pemex?   
 IV. ¿Conoce las propiedades y ventajas de los lubricantes de Pemex?   
 ¿Volumen del envase \_\_\_\_\_ Precio \_\_\_\_\_  
 ¿Frecuencia con que lo(s) compra? \_\_\_\_\_ cantidad en litros \_\_\_\_\_  
 ¿Cuándo no hay en existencia el lubricante que usa, cuál(es) es(son) la(s) marca(s) que utiliza? \_\_\_\_\_  
 III. ¿Marca que usaba antes? \_\_\_\_\_ Tamaño(s) \_\_\_\_\_ Lts. Precio(s) \_\_\_\_\_  
 ¿Cuándo cambió? \_\_\_\_\_ ¿Por qué cambió? \_\_\_\_\_  
 II. ¿Siempre ha usado el mismo lubricante? \_\_\_\_\_  
 ¿Cuáles? \_\_\_\_\_  
 ¿Se abastece directamente en la agencia autorizada?  si  no En caso negativo indique \_\_\_\_\_  
 ¿De qué tipo? \_\_\_\_\_  
 ¿Dónde lo compra? \_\_\_\_\_ ¿Tiene problemas de abastecimiento? \_\_\_\_\_

DISTRIBUIDORES

**INSTITUTO MEXICANO DEL PETROLEO**  
**SUBDIRECCION DE ESTUDIOS ECONOMICOS Y PLANEACION INDUSTRIAL**

"ENCUESTA NACIONAL SOBRE DISTRIBUCION Y MERCADEO DE LUBRICANTES AUTOMOTRICES"

No. Folio   
1  
2 3 6

Entidad \_\_\_\_\_ Municipio \_\_\_\_\_ Localidad \_\_\_\_\_  
 Razón Social \_\_\_\_\_ Domicilio \_\_\_\_\_ Calle \_\_\_\_\_ Número \_\_\_\_\_  
 7 8 9 11 12 13

Giro  14

- Clave Giro
- 1 Gasolinera
  - 2 Expendios de aceite
  - 3 Taller Armadora
  - 4 Taller Agencia Autorizada

1. Número de empleados   
15 18

2. Ventas promedio aprox. mensuales \$   
19 26

3. Lo Abastece: 1 2 3 4 5 6  
  
27 32

- 'Clave Progresiva' Tipo
- 1 Agencia de Pemex
  - 2 Agencia de Mobiloil
  - 3 Agencia de Esso
  - 4 Agencia de Texaco
  - 5 Agencia de Quaker State
  - 6 Otros -especifique \_\_\_\_\_

- 1 Camión especial de lubricantes
- 2 Camión de combustibles y lubricantes
- 3 Camiones de agencias autorizadas por Pemex
- 4 Otros -especifique \_\_\_\_\_

Clave  36 Medio de transporte que utiliza Pemex para abastecerlo

- 1 Cada 8 días
- 2 Cada 15 días
- 3 Cada 22 días
- 4 Cada mes
- 5 Cada mes y medio
- 6 Cada dos meses
- 7 Más de dos meses

Clave  35 Frecuencia de abastecimiento de lubricantes de Pemex

- 1 Es regular la entrega del pedido
- 2 No se entrega completo el pedido
- 3 No entregan las marcas y tamaños solicitados
- 4 Constantemente alteran los precios
- 5 Otros -especifique \_\_\_\_\_

Clave  34 ¿De qué tipo? Respuesta

- 1 Sí
- 2 No
- 3 No sabe

Clave  33 ¿Tiene problemas de abastecimiento de Pemex?

8. ¿Varían los precios de los lubricantes de Pemex?  37 Clave

- 1 Sí
- 2 No
- 3 No sabe

9. En particular ¿de qué marca(s)? a.  38 Clave

- 1 Pemex 100
- 2 Faja de Oro
- 3 Ebano
- 4 Diesel - ¿Cuál? \_\_\_\_\_
- 5 Otros -especifique \_\_\_\_\_

b.  39

a. Precio máximo      Precio mínimo

b. Precio máximo      Precio mínimo

Tamaño del envase a.   57 Lts.   52

b.   65 Lts.   64

67

10. ¿Juzga correctos los precios de compra y venta de los lubricantes de Pemex?  68 Clave

- 1 Sí
- 2 No
- 3 No sabe

¿Por qué?       69 71

11. ¿Varían los precios de las marcas de la competencia?  72 Clave

- 1 Sí
- 2 No
- 3 No sabe

14. ¿Tiene problemas para abastecer el área?  28

Clave	1	2	3
Tipo	Sí	No	No sabe

13. Tipo principal de consumidores  27

Clave	1	2	3	4	5	6	7
Tipo	Automóviles	Motocicletas	Camiones de carga	Camiones de pasajeros	Camiones urbanos	Tractores	otros -específique

Clave	09	10	11	12	13	14	15	16
Marca	Esso							

Clave	01	02	03	04	05	06	07	08
Marca	Mobil	Mobil	Delvac	Rubrex	Texaco	Havoline	Havoline	Ursa

Clave	1	2	3
Tipo	Esso	Esso	Esso

Clave	1	2	3
Tipo	Esso	Esso	Esso

12. ¿Qué marcas en particular? a.  73 74 b.  75 76

15. ¿Cuáles son?  11  62

16. ¿Son adecuados los precios de compra y venta de la competencia?  32

Clave	1	2	3
	Sí	No	No sabe

17. ¿Por qué? \_\_\_\_\_

18. En particular ¿Cuáles marcas? \_\_\_\_\_

19. ¿Cuáles son los incentivos o ventajas que obtiene en la compra de lubricantes de la competencia? \_\_\_\_\_

<input type="checkbox"/> 33	<input type="checkbox"/> 35	<input type="checkbox"/> 36	<input type="checkbox"/> 37	<input type="checkbox"/> 38	<input type="checkbox"/> 40
-----------------------------	-----------------------------	-----------------------------	-----------------------------	-----------------------------	-----------------------------

20. ¿Tiene problemas de abastecimiento de la competencia?  41

Clave	1	2	3
	Sí	No	No sabe

21. ¿Le proporciona la competencia ofertas en especie o en dinero?  42

Clave	1	2	3
	Sí	No	No sabe

EN VALOR		EN UNIDADES (Lts.)	
Inventario	1.	Inventario	1.
Compras	2.	Compras	2.
Pemex	1.		
Esso	2.		
Mobiloil	1.		
Texaco	2.		
Quaker St.	1.		
Otros	2.		

25. ¿Estimula usted la venta de alguna marca en particular?  Clave 43  No 1  SI 2

¿Por qué?  44  45  Clave 46  No 2  SI 1

26. ¿Da usted ofertas al consumidor de cierta marca de lubricantes?  Clave 46  No 2  SI 1

22. ¿De qué tipo? \_\_\_\_\_

23. ¿Qué marcas de la competencia hacen las ofertas? \_\_\_\_\_

24. Relación compras-Inventario del \_\_\_\_\_ Día \_\_\_\_\_ Mes \_\_\_\_\_ al \_\_\_\_\_ Día \_\_\_\_\_ Mes \_\_\_\_\_

Marca \_\_\_\_\_  47  48  49  50 ¿En qué consisten? \_\_\_\_\_

27. Tiene publicidad de: Clave Tipo de Publicidad

Pemex	51						56	1	Anaque!, Escaparate, Ventana
Esso	57						62	2	Carteles alusivos a la marca
Mobiloil	63						68	3	Latas arregladas en pirámides, cabeceras, etc., de la marca
Texaco	69						74	4	Carteles de la marca anunciando ofertas
Quaker State	75						80	5	Radio
Otros	3						8	6	Televisión

OBSERVACIONES: \_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

Nombre del Enumerador \_\_\_\_\_

Lugar y Fecha \_\_\_\_\_

1	Pemex - Específfique
2	Mobiloil - Específfique
3	Texaco - Específfique
4	Esso - Específfique
5	Quaker State - Específfique
6	Otros - Específfique

1. ¿Que marca de lubrificante usa?

Clave      Tamaño en litros      Consumo mensual aprox.      Precio/Litro

Entidad      Municipio      Localidad

Nombre      Domicilio      Calle y Número      Placas      Tel.

No. Folio

“PREFERENCIA Y CONSUMO DE LUBRICANTES AUTOMOTRICES”

SUBDIRECCION DE ESTUDIOS ECONOMICOS Y PLANEACION INDUSTRIAL

INSTITUTO MEXICANO DEL PETROLEO

VEHICULOS

2. ¿Por qué prefiere esa marca?  22

- Clave
- 1 Por ser mejor y dar mayor rendimiento
  - 2 Por ser más barato
  - 3 Se la sugieren en el taller de mantenimiento
  - 4 Es la única que se vende en el taller de mantenimiento
  - 5 La sugirió la agencia donde compró la unidad
  - 6 La sugirieron amistades
  - 7 Otras, especifique
  - 8 No sabe

3. Frecuencia con que compra  23

- Clave
- 1 Menos de 15 días
  - 2 Cada 15 días
  - 3 Cada mes
  - 4 Cada dos meses
  - 5 Cada tres meses o más

4. ¿Siempre ha usado esta marca?  24

- Clave
- 1 Sí
  - 2 No
  - 3 No sabe

5. ¿Por qué cambió de marca? \_\_\_\_\_

25  26

6. ¿Qué marca usaba antes?  27

- Clave      Productor
- 1 Pemex
  - 2 Mobiloil
  - 3 Texaco
  - 4 Esso
  - 5 Quaker State
  - 6 Otros

- 1 Pemex-marca \_\_\_\_\_
- 2 Mobiloil-marca \_\_\_\_\_
- 3 Texaco-marca \_\_\_\_\_
- 4 Esso-marca \_\_\_\_\_
- 5 Quaker State-marca \_\_\_\_\_
- 6 Otras marca \_\_\_\_\_

Clave Lubrificantes Específique

31

10. Cuando no hay en existencia el lubricante acostumbrado, ¿Qué otros lubricantes usa?

- 1 Si
- 2 No
- 3 No sabe

Clave

30

9. ¿Varía el precio del lubricante que usa?

- 1 Agencias autorizadas
- 2 Gasolineras
- 3 Talleres mecánicos
- 4 Expendios de aceite
- 5 En cualquier lugar

Clave

29

8. ¿Dónde hace normalmente el cambio de aceite?

- 1 Si
- 2 No

Clave

28

7. ¿Regularmente tiene problemas para abastecerse?

13. Lugar de abastecimiento del lubricante sustituto  37

Clave

Lugar

- 1 Misma agencia
- 2 Misma gasolinera
- 3 Mismo taller mecánico
- 4 Mismo expendio de aceite
- 5 Cualquier parte

14. Modelo de la unidad  38  39 Marca  40 Modelo

Años

Clave

Marca

Especifique

- 01
- 02
- 03
- 04
- 05
- 06
- 07
- 08
- 09
- 10 o más

- 1 General Motors \_\_\_\_\_
- 2 Ford \_\_\_\_\_
- 3 Automex \_\_\_\_\_
- 4 Dina Renault \_\_\_\_\_
- 5 Volkswagen \_\_\_\_\_
- 6 Nissan \_\_\_\_\_
- 7 Renault \_\_\_\_\_
- 8 VAM \_\_\_\_\_
- 9 Otros \_\_\_\_\_

15. Número aproximado de kilómetros que recorre mensualmente

41 44

16. Ingreso mensual global \$      45 49

OBSERVACIONES: \_\_\_\_\_

Nombre del enumerador \_\_\_\_\_

Lugar y fecha \_\_\_\_\_

Paralelamente se hizo un cálculo preliminar del costo de la encuesta, el cual fue presentado a la Subdirección de Estudios Económicos del Instituto Mexicano del Petróleo, quien aprobó un presupuesto para su realización. Como podrá intuirse este cálculo se modificó cuando se conoció posteriormente el tamaño de la muestra.

De acuerdo a las instrucciones del inciso 3.1.8, luego se procedió al diseño de la metodología consistente en:

1. Clasificación de los lubricantes.
  2. Obtención del marco muestral.
  3. Cálculo del tamaño de la muestra.
  4. Desarrollo de las fórmulas para las estimaciones.
- Que se expone a continuación con mayor detalle.

#### METODOLOGIA DE LA INVESTIGACION

##### 1. Clasificación de los Lubricantes.

Para la consecución del estudio se decidió clasificar a los lubricantes en:

	VENTAS INTERIORES
Lubricantes Básicos	\$125,145,000.00
Lubricantes Industriales	111,961,000.00
Lubricantes para Automóviles	297,943,000.00
Lubricantes para Tractores	189,144,000.00
Lubricantes para motores Diesel	28,973,000.00
Lubricantes Aviación	89,000.00
	\$753,255,000.00

FUENTE: Balance 1971. Cuenta 950 de los Estados Financieros de Petróleos Mexicanos.

A través de la clasificación anterior se logró establecer la importancia de cada uno de los subgrupos y su uso concreto dentro de la amplia gama de lubricantes producidos por Petróleos Mexicanos.

Por lo que se refiere a los Lubricantes Básicos, se encontró que son vendidos por Petróleos Mexicanos como insumos a diferentes empresas privadas que los utilizan para la obtención del resto de los lubricantes arriba señalados. Dado que el destino final de ellos, no es propiamente el consumidor final, se optó por no estudiarlos debido a que no reúnen las características de bienes que concurren a un mercado de competencia.

Asimismo, los lubricantes para avión no fueron considerados en el presente estudio por representar un porcentaje reducido dentro de las ventas interiores de la empresa, y por no ser producidos en el país al no reunir nuestros aceites las especificaciones requeridas para su producción.

Decidiéndose en definitiva por estudiar los lubricantes para automóvil, industriales, tractores y para motores diesel, en relación a sus equivalentes producidos por Mobiloil, Texaco, Esso y Quaker State.

En consecuencia, para captar la información correspondiente, se diseñaron tres cuestionarios y fue necesario calcular 3 tamaños de muestra para obtener una visión completa del problema.

##### 2. Marco Muestral.

Con objeto de detectar las causas de los problemas en el mercado nacional de lubricantes que afectan a Pemex, se hizo necesario investigar por una parte, el proceso de distribución y por otra cómo se motivan las preferencias en la demanda. Esto implica que en el presente estudio, se investiguen los centros de abastecimiento inmediatos de los consumidores (gasolineras, expendios de aceite y talleres mecánicos en donde se hacen servicios de lubricación) y los mismos consumidores (automóviles, camiones, tractores, motocicletas e industrias).

Para lograr este objetivo, esta investigación se realizó mediante la técnica estadística del Muestreo Probabilista.

El marco muestral de que se dispuso para la selección de la muestra, estuvo integrado por directorios, unos para distribuidores de

Lubrificantes y otros para consumidores. Por tal razón, se utilizó un muestreo irrestricto aleatorio para el caso de vehículos y distribuidores y en el caso de industrias la selección de la muestra se hizo en una sola etapa, considerando probabilidades de selección, de acuerdo con su importancia económica, entre las industrias con ventas anuales mayores de un millón de pesos, se les asignó una probabilidad tres veces mayor que las que tenían un volumen de ventas menor de un millón de pesos (según datos del Censo Industrial de 1970). Las industrias consideradas en el marco muestral, fueron todas aquellas que en los Censos Industriales del 60 y 65 aparecieron con mayor consumo de lubricantes y combustibles, la lista de las industrias consideradas en el estudio se anexaron al final del capítulo.

En el consumo de lubricantes para las industrias, sólo se incluyeron los lubricantes consumidos en el proceso industrial, quedando excluidos los que utilizan en los vehículos de transporte.

### 3. Tamaño de la Muestra.

Al planear la investigación por muestreo en base a las ventajas de exactitud, costo y tiempo, que esta técnica proporciona, se plantea la necesidad inmediata de obtener el tamaño de la muestra requerido para estimar las características (parámetros) del universo con una precisión y confiabilidad prefijadas.

En este caso, se tomó una confiabilidad del 90% y un error permisible del 10%/o.

Para fijar el tamaño de la muestra en cada entidad se utilizó una variable binomial (porque no se dispuso de datos que permitieran un cálculo más exacto) con la siguiente fórmula:

$$n = \frac{E^2 P}{t^2 (\infty, \infty) q} \left( \frac{1}{1 + \frac{1}{N}} \left( \frac{t^2 (\infty, \infty) q}{E^2 P} - 1 \right) \right)$$

en donde:

- n = Tamaño de la muestra.
- $t_{(\infty, \infty)}$  = t de Student.
- $1 - 2\alpha$  = Confiabilidad.
- q = P · 0.5.
- E = Porcentaje del error permitido.
- N = Número total de unidades de muestreo.

#### A. Estimación.

En la estimación de valores poblacionales se utilizaron las siguientes estadísticas:

#### Muestreo Irrestricto Aleatorio.

Para Consumidores y Distribuidores.

Estimador de la media poblacional.

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

Estimador del total poblacional.

$$X = \frac{N}{n} \sum_{i=1}^n x_i$$

Estimador de la desviación estándar.

$$s(x) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2}$$

Estimador del coeficiente de variación del estimador de la media poblacional.

$$c.v. (\bar{x}) = \frac{s(\bar{x})}{\bar{x}}$$

Intervalo de confianza para la estimación del total de la población al 90% de confianza.

$$(X - 1.64 s \sqrt{n}, X + 1.64 s \sqrt{n})$$

Estimador del coeficiente de correlación.

$$r = \frac{\frac{1}{n} \sum x_i y_i - \bar{x} \bar{y}}{s(x) s(y)}$$

Muestreo monoetápico con probabilidades variables de selección.

Para las Industrias.

Estimador de la media poblacional.

$$\bar{x} = \frac{1}{N} \sum_{i=1}^n \frac{x_i}{p_i}$$

en donde:

$$p_i = \frac{1}{N} \quad \text{si } x_i \text{ se refiere a una industria menor, y}$$

$$p_i = \frac{3}{N} \quad \text{si } x_i \text{ se refiere a una industria mayor}$$

Estimador del total de la población.

$$X = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{p_i}$$

Estimador de la desviación estándar.

$$s(x) = \sqrt{\frac{\frac{1}{n} \sum (x_i / p_i)^2 - \frac{1}{n} (\sum \frac{x_i}{p_i})^2}{n - 1}}$$

Estimador del coeficiente de variación del estimador de la media poblacional.

$$c. v. (\bar{x}) = \frac{s(x)}{\bar{x} \sqrt{n}}$$

Intervalo de confianza para la estimación del total de la población al 95% de confianza.

$$(X - 1.96 s \sqrt{n}, X + 1.96 \sqrt{n})$$

## EVALUACION DE LA ENCUESTA

Los aspectos más relevantes en lo que se refiere a la obtención de información exacta y confiable, pueden reducirse fundamentalmente a los dos siguientes puntos.

- Planeación de la encuesta.
- Obtención de la información o levantamiento de la encuesta.

Por lo que toca a la planeación, ésta debe efectuarse de tal manera, que se asegure suficiente información para poder hacer buenas inferencias sin que esto repercuta en un despilfarro de recursos. En el presente caso, por tratarse de un trabajo en el cual no existían antecedentes que permitieran afinar un tamaño de muestra que respondiera a las exigencias de planeación en forma óptima, debe tomarse como un trabajo piloto a partir del cual sea posible superar todas las limitaciones que en este caso se presentaron; sin embargo, la escasez de información sólo se presentó en algunos conceptos a causa de que los entrevistados opusieron una resistencia sistemática para contestar cuestiones relacionadas con ingresos, tanto en las industrias, como en los usuarios de vehículos, por tal motivo, estos resultados deben tomarse con cierta reserva. La evaluación se presenta en la siguiente serie de indicadores (intervalos de confianza y coeficientes de variación).

### 4.3 Algunas Otras Aplicaciones del Muestreo Estadístico en la Industria Petrolera.

- Conocimiento de las necesidades de capacitación del personal obrero.

De acuerdo con los objetivos de la investigación y siguiendo las instrucciones de la red de actividades, se diseñó el cuestionario, del cual se debe tomar la variable más importante para calcular el tamaño de la muestra.

- Evaluación de la eficiencia de los servicios de apoyo a un sector en particular.

Se definirían como servicios de apoyo el procesamiento electrónico, servicio de biblioteca, información y difusión, algunos talleres, servicio médico, almacén, aprovisionamiento, etc.

# INDICADORES PARA LA EVALUACION DE LA ENCUESTA DE LUBRICANTES

## INDUSTRIAS

CONCEPTO	VALOR DEL INDICADOR			
	INDICADOR	Nuevo León	Sinaloa	Querétaro
Número de industrias que obtienen precios especiales en la compra de lubricantes.	Cota inferior	52	16	1
	Estimación	75	31	4
	Cota superior	94	46	7
	Coefficiente de Variación	17	29	50
Número de industrias opinaron que el lubricante usado es de buena calidad.	Cota inferior	74	67	38
	Estimación	84	85	46
	Cota superior	110	103	54
	Coefficiente de Variación	11	13	10
Número de industrias que manifestaron estar dispuestas a utilizar lubricantes Pemex.	Cota inferior	120	75	42
	Estimación	137	96	50
	Cota superior	162	117	58
	Coefficiente de Variación	9	14	10
Número de industriaa que se abastecen directamente en agencias autorizadas.	Cota inferior	120	75	26
	Estimación	137	93	34
	Cota superior	162	111	42
	Coefficiente de Variación	9	12	15
Consumo de lubricantes (en litros) de las industrias.	Cota inferior	25,586	34,066	..
	Estimación	109,900	59,222	15,763
	Cota superior	194,214	84,378	31,559
	Coefficiente de Variación	76	25	61
Consumo de lubricantes (en pesos) de las industrias.	Cota inferior	449,312	57,331	25
	Estimación	1,313,974	305,420	9,519
	Cota superior	1,313,974	553,509	19,013
	Coefficiente de Variación	117	36	61

## DISTRIBUIDORES

CONCEPTO	VALORES DEL INDICADOR			
	INDICADOR	Nuevo León	Sinaloa	Querétaro
Número de distribuidores sin problemas de abastecimiento de lubricantes de Pemex.	Cota inferior	18	67	17
	Estimación	93	78	21
	Cota superior	105	89	25
	Coefficiente de Variación	8	9	13
Número de distribuidores que manifestaron que los precios de compra de lubricantes de Pemex son adecuados.	Cota inferior	65	68	18
	Estimación	78	79	22
	Cota superior	91	90	26
	Coefficiente de Variación	10	8	13
Número de distribuidores sin problemas de abastecimiento de lubricantes de la competencia.	Cota inferior	44	71	25
	Estimación	56	82	28
	Cota superior	68	93	31
	Coefficiente de Variación	13	8	8
Número de distribuidores que manifestaron que los precios de compra de lubricantes de la competencia son adecuados.	Cota inferior	44	62	11
	Estimación	56	74	15
	Cota superior	68	86	19
	Coefficiente de Variación	13	9	17

## VEHICULOS

### VALOR DEL INDICADOR

CONCEPTO	INDICADOR	Nuevo León	Sinaloa	Querétaro
Consumo promedio de lubricantes automotrices.	Cota inferior Estimación	4.6	13.2	20.7
	Cota superior	4.7	14.0	23.7
	Coefficiente de Variación	4.8	14.8	26.7
Porcentaje de unidades que utilizan los lubricantes de Petróleos Mexicanos.	Cota inferior Estimación	1.3	3.4	5.9
	Cota superior	62.8	74.5	74.2
	Coefficiente de Variación	64.3	78.6	78.1
Porcentaje de unidades que prefieren lubricantes de Petróleos Mexicanos por ser de mejor calidad y de mayor rendimiento.	Cota inferior Estimación	65.8	82.7	82.0
	Cota superior	1.4	3.2	3.1
	Coefficiente de Variación	37.7	48.3	38.3
Porcentaje de unidades que manifestaron no tener problemas en el abastecimiento de lubricantes.	Cota inferior Estimación	38.9	51.6	40.7
	Cota superior	40.2	56.8	44.6
	Coefficiente de Variación	1.9	6.1	5.6
Cota inferior Estimación	Cota inferior Estimación	96.1	94.6	80.8
	Cota superior	96.7	96.5	89.5
	Coefficiente de Variación	97.3	98.3	98.2
Cota superior	Cota superior	0.4	1.8	5.9
	Coefficiente de Variación			

Como en el ejemplo anterior, deberán seguirse las instrucciones descritas en el Manual, o específicamente seguir los pasos recomendados en el inciso 3.1.8.

3. Determinación de recursos mediante ensayos piloto.
4. Selección de yacimientos para determinar los factores de equivalencia (poder calorífico).
5. Determinación de brigadas tipo, cuyo objetivo es fijar "costos predeterminados".
6. Selección de pozos para mejorar la estimación de costos de perforación en distintas condiciones geológicas.
- 6.1 Selección de pozos para determinar la función de costo de extracción, considerando la profundidad, el acondicionamiento y el transporte.
7. Muestreo para control de calidad.
8. Análisis de mercado de los productos.
9. Identificación de las necesidades de vivienda, ingreso, educación, etc., para los trabajadores.
10. Estimación de la demanda regional de energéticos, etc.

1. BIBLIOGRAFIA.

1. W. COCHRAN.- "Técnicas de Muestreo", Editorial C.E.C.S.A.
2. P.V. SUKHATME.- "Teoría de Encuestas por Muestreo con Aplicaciones", Editorial Fondo de Cultura.
3. H. BLALOCK.- "Estadística Social", Editorial Fondo de Cultura Económica.
4. R. FERBER, S. HOLLANDER "Marketing Research", Editorial The Ronald Press Co., New York.

*Impreso en*

**INSTITUTO MEXICANO DEL PETROLEO**

División de Información y Difusión

*Publicaciones*