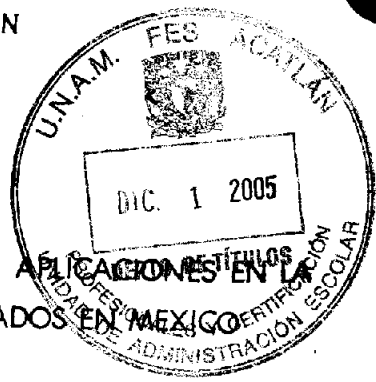




UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO



FACULTAD DE ESTUDIOS SUPERIORES ACATLAN



MUESTREO POLIETAPICO USO Y APLICACIONES EN LA INVESTIGACION DE MERCADOS EN MEXICO

T E S I N A

QUE PARA OBTENER EL TITULO DE:

A C T U A R I O

P R E S E N T A :

CAROLINA SEGOVIA ARREVILLAGA

ASESOR ACT. IVAN MEJIA GUEVARA

NOVIEMBRE DE 2005

m. 350461



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

AGRADECIMIENTOS

A la Universidad Nacional Autónoma de México

Por la gran oportunidad de haber estudiado en sus aulas. Es en verdad un orgullo para mí el pertenecer a esta máxima casa de estudios.

A mi asesor Iván Mejía Guevara

Me animaste a seguir adelante cuando pensé que todo estaba perdido. Sin tu valiosa ayuda y tu orientación no hubiera sido posible terminar este trabajo.

A mis hijas Adriana y Carolina

Porque ustedes son mi mayor tesoro en este mundo. Su presencia me hace seguir adelante en todos los aspectos de mi vida.

A mi esposo Arturo Cortés Guzmán †

Te nos adelantaste en el camino, pero sabemos que eres el ángel que nos cuida desde el cielo.

A mis padres

Gracias por la oportunidad que me dieron para estudiar, y gracias por la larga espera para tener este trabajo en sus manos.

Miguel

Porque cuando he necesitado ayuda, me la has brindado, aún cuando tú no estés consciente de ello.

Laura

Tu ejemplo me dio ánimos para seguir adelante. Tienes una gran cultura y todo lo que te pregunto lo sabes.

Ma. Luisa

Arturo Rodríguez Padilla

Por todo lo que hemos aprendido y porque tu presencia le ha dado otro sentido a mi vida.

CONTENIDO

	Pag.
INTRODUCCIÓN	1
OBJETIVO	3
CAPÍTULO I. CONCEPTOS GENERALES	
A La Encuesta	4
I.A.1. Fundamentos de estadística	4
I.A.2. Concepto de encuesta	7
I.A.3. Recolección de datos	9
I.A.4. Diseño, objetivo y planeación de la encuesta	14
I.A.5. La población y sus valores.....	16
B La Investigación de Mercados	
I.B.1. Antecedentes.....	19
I.B.2. Investigación de Mercados en México.....	23
I.B.3. Servicios.....	25
Capítulo II. EL MUESTREO PROBABILÍSTICO	
II.1. Generalidades	30
II.2. Tipos de Muestreo	33
II.2.1. El Muestreo Aleatorio Simple	34
II.2.2. El Muestreo Estratificado.....	39
II.2.2.1. Notación.....	42
II.2.3. El Muestreo por Conglomerados.....	46
II.2.3.1. Notación para el muestreo por conglomerados del mismo tamaño	48
II.2.3.2. Estimación en el muestreo por conglomerados del mismo tamaño	50
II.2.4. El muestreo Sistemático.....	53

Capítulo III. EL MUESTREO POLIETÁPICO

III.1. Concepto	56
III.2. Etapas del muestreo polietápico.....	59
III.3. Muestreo por conglomerados en dos etapas.....	61
III.4. Determinación del tamaño de muestra en diseños polietápicos.....	63
III.5. Ventajas y desventajas con respecto al muestreo aleatorio simple.....	70

Capítulo IV. APLICACIÓN DEL MUESTREO POLIETÁPICO EN LA INVESTIGACIÓN DE MERCADOS

IV. 1. Generalidades.....	72
IV. 2. Diseño Muestral.....	74
IV. 3. Tamaño de la muestra.....	75
Conclusiones	84
Bibliografía	86

INTRODUCCION

Actualmente una de las prioridades en casi toda actividad humana es la recopilación de información. Se necesita información para planear, evaluar, rectificar, corregir, implementar estrategias, etc. Sin embargo esa recopilación de información no es sencilla. Para que sea confiable debe ser representativa de la población de la cual se extrae, y eso quiere decir que la información obtenida realmente refleje la realidad de la población que se está estudiando.

Y para que ello ocurra es necesario llevar a cabo un proceso que comienza desde el hecho de saber qué es exactamente lo que se quiere investigar hasta saber cuál es la población que quiero estudiar.

El definir ambas situaciones correctamente es prioritario en un buen diseño de investigación. Pero ese es sólo el primer paso. Después hay que establecer los instrumentos de recolección de la información (encuestas, entrevistas, cuestionarios, etc). Esta etapa es muy importante y básica, pues ninguna cantidad de análisis estadísticos puede compensar el uso de una encuesta mal diseñada.

Finalmente y no por ello menos importante es el establecer a quién y de qué manera se le va a aplicar esos instrumentos de recolección de la información, de manera que la información obtenida en la muestra sea tan confiable que reproduzca las características de interés que existan en la población y que van a ser estudiadas.

Este último paso es precisamente el objetivo del muestreo. Y para que ello sea posible, los diseños muestrales deben tener ciertas características y definir

determinados conceptos de los cuales se hablará en el primer capítulo. Asimismo, se hablará en términos generales de lo que es la Investigación de Mercados y su importancia en la actualidad.

En el segundo capítulo se hablará de los diferentes métodos de muestreo en una sola etapa.

El tercer capítulo habla acerca del muestreo polietápico; de sus características, notación, la estimación de algunos parámetros y de las ventajas que tiene con respecto al muestreo aleatorio simple.

Finalmente, en el cuarto capítulo se sugiere a manera de ejemplo la utilización del muestreo polietápico en un problema relativo a la Investigación de Mercados, sobre todo en lo referente a la obtención del tamaño de muestra y la forma en que cambia según los costos y la varianza que se proponga.

OBJETIVO

ESTUDIAR AL MUESTREO POLIETÁPICO COMO UNA ALTERNATIVA DE MUESTREO PARA SER UTILIZADO EN LA INVESTIGACIÓN DE MERCADOS EN MÉXICO, A FIN DE REDUCIR LAS LIMITANTES ECONÓMICAS SIN SACRIFICAR PRECISIÓN EN LOS RESULTADOS.

CAPÍTULO I

CONCEPTOS GENERALES

A La Encuesta

I.A.1. Fundamentos de Estadística

El término estadística se deriva de la palabra Estado. Originalmente, esta rama de la ciencia era aplicada únicamente para asuntos de gobierno y fenómenos sociales. "Los aritméticos políticos introdujeron la estadística con el nombre de *aritmética de estado*, la cual permitía obtener información mediante operaciones con los datos demográficos, sociales y económicos."¹

Actualmente, es posible utilizar la estadística en cualquier área del conocimiento humano para describir, analizar y predecir, haciendo uso de la observación de fenómenos aleatorios, midiendo su frecuencia y sus resultados. Así, "la estadística está ligada con los métodos científicos en la recolección, organización, presentación y análisis de datos, tanto para la deducción de conclusiones como para tomar decisiones razonables de acuerdo con tales análisis."²

Históricamente, la estadística ha contribuido al desarrollo de la ciencia y la técnica en la medida en que facilita la utilización de los datos que pertenecen a un grupo de individuos, objetos o situaciones. Por ejemplo, la altura de jóvenes en una preparatoria, el número de autos que transitan por una calle en un horario determinado o la cantidad de reclamaciones por servicio en una línea aérea. Su

¹ Ríos, Sixto. *Iniciación estadística*. Paraninfo. Madrid. 1997

² Spiegel, Murray R. *Estadística*. McGraw Hill. México. 1984

aplicación abarca asuntos públicos y privados, cuestiones sociales, económicas y políticas, negocios, salud, agricultura, industria, psicología, biología y muchos otros usos.

El objetivo de la estadística es el de hacer inferencias mediante la predicción y la toma de decisiones sobre algunas características de una población con base en la información contenida en una muestra. En lugar de examinar un grupo entero llamado población o universo, se examina una pequeña parte de éste, al cual se le denomina muestra.

A reserva de ampliarlo más adelante, se puede decir que "una población es el conjunto de todas las mediciones de interés de un muestrista. En la mayoría de los casos es imposible observar a todos los elementos de una población, por lo que se selecciona una muestra de la mencionada población, esperando que sea una representación significativa de ella."³ Por lo que respecta a la muestra, ésta puede definirse como una parte de la población. Con ello, el interés por la muestra se basa en la posibilidad de describir y decidir aspectos de la población de la cual fue extraída.

Para lograr el objeto de la estadística es importante que se cubran al menos cinco elementos

- Especificar con claridad una pregunta a contestar tomando en cuenta a la población de estudio. Si se inicia el proceso sin un objeto de investigación comprensible para los participantes, ello provocará contratiempos que afectarán el resultado final.

³ Mendenhal, William y James E. Reinmuth. *Estadística para Administración y Economía*. Wadsworth Internacional Iberoamericana. México. 1981

- Realizar el procedimiento de muestreo u obtención de la muestra. En esta parte, deben considerarse dos variables: el costo o presupuesto con el que se cuenta para obtener precisión en los datos recopilados y el tiempo necesario para iniciar y concluir el proceso. Así, incluir demasiadas observaciones puede ser costoso e inútil. De igual forma, determinar pocas puede resultar insuficiente para contestar la pregunta planteada.
- Recopilar y analizar la información muestral, utilizando el método estadístico apropiado para extraer la información de los datos.
- Inferir acerca de la población, haciendo uso de la información muestral.
- Tomar decisiones sobre la base del análisis, considerando una medida de confiabilidad que permita determinar el grado de certeza de los resultados.

En términos generales, la estadística puede clasificarse en dos: la estadística descriptiva y la estadística inferencial. La primera, utilizando matemáticas elementales, atiende a la descripción cuantitativa de una serie de personas, lugares o cosas, sin tener en cuenta si se trata de una muestra o de una población. De esta forma, la estadística descriptiva "puede definirse como los métodos que aplican la recolección, presentación y caracterización de un conjunto de datos a fin de describir en la forma apropiada las diversas características de ese conjunto de datos."⁴

⁴ Berenson, Mark L. y David M. Levine. *Estadística para Administración y Economía, conceptos y aplicaciones*. Interamericana. México. 1986

Por su parte, la estadística inferencial, también llamada teoría de muestras o inferencia estadística, utiliza el cálculo de probabilidades a fin de obtener conclusiones acerca de un grupo grande de personas, lugares o cosas, por medio de la observación de solo una pequeña parte del conjunto total. Es decir, "tan sólo con base en los resultados de un muestreo se posibilita la estimación de características y la toma de decisiones respecto a una población."⁵

Aunque la estadística descriptiva es importante para caracterizar y presentar información, el desarrollo de la estadística inferencial ha motivado en mayor medida la expansión de la aplicación de los métodos estadísticos. Este hecho, sobre todo, tiene relevancia en el muestreo y el diseño de muestras por que es la estadística inferencial la que proporciona las herramientas matemáticas imprescindibles para concluir y decidir sobre una población estudiada a través de una muestra representativa de ella.

Tomando en cuenta lo anterior, cabe señalar que el proceso de inferencia está ligado al proceso inductivo. En ambos, "nunca se puede estar por completo cierto de los resultados; debe confiarse en el conocimiento e integridad personal de quien presenta los datos. Además, se aplican a casos que no se han observado en realidad, porque si fuera así ya no sería inducción sino enumeración o censo."⁶

I.A.2. Concepto de encuesta

Dentro de las ciencias sociales, la encuesta es un instrumento de investigación muy utilizado, y se puede considerar como la consulta tipificada de

⁵ Berenson, Mark L. y David M. Levine. Op. cit.

⁶ Campbell, Stephen K. *Defectos y errores del pensamiento estadístico*. Limusa. México. 1981

personas que se lleva a cabo con la ayuda de un cuestionario. Las personas a las que se les aplica dicho cuestionario se seleccionan mediante métodos estadísticos.

Cabe mencionar que una de las características de la encuesta es que la información que se obtiene ya está preparada y estructurada de antemano, de acuerdo a los objetivos planteados desde el inicio, a diferencia de la entrevista, en la cual hay una mayor flexibilidad para obtener información.

La encuesta tiene distintas aplicaciones temáticas de acuerdo a lo que se quiera investigar. Hay encuestas por sondeos (determinación de una muestra representativa de una población); encuestas de opinión (encuesta por sondeo para conocer la opinión de un tema en concreto); encuestas urbanas, sociológicas, demográficas, de investigación de mercados, etc.

Con respecto a la estructura de la encuesta, las preguntas que se llevan a cabo pueden ser de varios tipos: ya sea preguntas abiertas (en las cuales el individuo puede responder con unas líneas o frases); cerradas (donde sólo puede responder con un 'sí' o un 'no'); de elección múltiple (es decir, donde puede elegir entre varias respuestas), y de estimación o evaluación (las preguntas presentan grados diferentes de intensidad). Para contestar dichas encuestas, se puede recurrir a una entrevista personal, al correo o por teléfono. Lo importante es que quien aplique dicho cuestionario, debe ser una persona especialmente preparada para llenar correctamente los formatos utilizados a fin de no alterar las respuestas dadas por los participantes de dicha encuesta.

En el siguiente apartado se hablará más ampliamente de la encuesta.

I.A.3. Recolección de datos

Existen tres métodos básicos para obtener datos: recopilar datos que ya fueron publicados por fuentes gubernamentales, industriales o individuales, diseñar un experimento para obtener los datos necesarios y efectuar una encuesta. Para fines de una investigación documental, la primera fuente es la más usada, para fines científicos la segunda requiere de procedimientos estadísticos más complejos y para la realización de una investigación de campo, la encuesta es la más común.

Así, una investigación de campo puede realizarse a través de observaciones y encuestas. Por medio de la observación de documentos, objetos y fenómenos físicos, y conductas, el hombre obtiene datos que traduce en información. "Este es el mejor método de recolección de datos porque está libre de errores de memoria de los que responden, de exageración y de los efectos del prestigio o perjuicio de los entrevistados. Este método es útil para el estudio de pequeñas muestras porque en él se mantiene la objetividad."⁷

Por otro lado, la encuesta representada por cuestionarios y entrevistas también es un medio para obtener datos sobre la muestra de una población. Los cuestionarios por correo, por lo general suelen ser económicos en su implementación. Sin embargo, se depende de la contestación oportuna de los interrogados y del entendimiento de las preguntas. A diferencia de la entrevista, en el cuestionario no puede aclararse la pregunta e incluso se carece de la posibilidad de obtener datos que en principio no se habían contemplado.

⁷Raj, Des. Suárez, Eduardo L. (Traductor). *La estructura de las encuestas por muestreo*. Fondo de Cultura Económica. México. 1979

Así, para incrementar la calidad de un cuestionario es necesario determinar un objetivo, otorgar un valor a cada pregunta, escoger qué tipo de preguntas se harán (abiertas o cerradas, directas o indirectas), adecuar el vocabulario según la persona a cuestionar y el tema, ir de lo general a lo particular, determinar el número de preguntas y fomentar la confianza con los cuestionados.

La entrevista, como medio para realizar encuestas, es diferente al cuestionario. Para incrementar su validez, en la entrevista se requiere que el entrevistador o enumerador sea agudo en la observación y detecte la sinceridad de las respuestas, tenga capacidad para escuchar, transcribir, seleccionar y condensar la información obtenida. A su vez, el entrevistado debe tener la disposición para contestar, debe ser sincero, tener facilidad oral y conocimiento sobre lo que va a responder.

En la práctica, la entrevista es el método más común para la medición en las encuestas, lo cual depende en gran medida de la habilidad del entrevistador para obtener una respuesta válida. Las modalidades de la entrevista pueden ir desde entrevistar con un cuestionario hasta aplicar un guión de temas que abarquen y contribuyan al objeto de la investigación. Por lo que respecta al entrevistador, éste debe ser hábil para manejar situaciones diversas e imprevistas, sin limitarse a seguir instrucciones escritas.

“Una encuesta puede dar información acerca de varias o diversas poblaciones. La misma encuesta puede abarcar contenidos diferentes. Por ejemplo, una encuesta de accidentes puede proporcionar información acerca de accidentes aislados, personas lesionadas, familias que han sufrido accidentes y hogares con

accidentes."⁸ Lo anterior, genera que existan subclases en una porción de muestra previamente diseñada. Por ejemplo, dichas subclases pueden ser periodos de tiempo, tal como una semana, un mes, etcétera.

El muestreo de encuestas -o el muestreo de una población- estudia los métodos para seleccionar y observar una parte (muestra) de la población con el fin de hacer inferencias acerca de toda la población. En comparación con un censo, el muestreo de encuestas ofrece las siguientes ventajas:

- es económico,
- es rápido y oportuno, y
- ofrece precisión y calidad

En términos generales, y dentro del área de estadística, las encuestas de muestreo pueden ser clasificadas en: descriptivas y analíticas. En una encuesta descriptiva, "el objetivo es simplemente el de obtener cierta información acerca de grandes grupos: por ejemplo, el número de personas que vieron un programa de televisión. En una encuesta analítica, se hacen comparaciones entre subgrupos diferentes de la población para describir si las diferencias que existen entre ellos nos permiten formar o verificar hipótesis relacionadas en la población."⁹

Cuando el estadístico desea obtener respuesta a diversos fenómenos, se estará hablando de que obtendrá variables aleatorias o datos observados que pueden diferir entre una respuesta y otra. Una variable es el conjunto de características de las entidades que interesa en una investigación científica. Por ejemplo, un investigador agrícola puede estar interesado en la capacidad de producción de café después de una sequía. "Cuando los valores numéricos que

⁸ Kish, Leslie. *Muestreo de encuestas*. Trillas. México. 1982

⁹ Casas Díaz, Eduardo. *Técnicas de Muestreo*. CECSA. México. 1971

toma una variable provienen de factores fortuitos y si un determinado valor no se puede predecir exactamente con anticipación, esa variable se denomina variable aleatoria."¹⁰

Existen dos tipos de variables aleatorias: las cualitativas y las cuantitativas. Las primeras, arrojan datos categóricos o atributos tales como sexo, estado civil, color de ojos, etcétera, generalmente son aquellos casos donde no es posible hacer medidas numéricas. Las variables cuantitativas están determinadas por algún procedimiento de medición denominado medida numérica, por ejemplo, la talla, el número de hijos, etcétera.

Asimismo, los datos cuantitativos pueden dividirse en discretos o continuos. "Los discretos son respuestas numéricas que surgen de un proceso de conteo y donde existe la presencia de 'vacíos' o 'interrupciones' entre los valores que puede tomar, mientras que los datos cuantitativos continuos son respuestas numéricas que surgen de un proceso de medición."¹¹ Por ejemplo, el número de todos los periódicos en Puebla es un dato cuantitativo discreto. En cambio, el número de periódicos con temas sociales en Puebla es un dato cuantitativo continuo ya que su valor depende de un intervalo donde se incluya qué es un tema social.

El diseño de una encuesta, ya sea por cuestionario o por entrevista, requiere de al menos definir un guión que limite el alcance de los datos que se desean obtener. En el campo de trabajo, la elaboración de una encuesta es un arte que mejora con la experiencia. Para confeccionarlo es necesario considerar la longitud de un documento o la duración de una entrevista que proporcione respuestas significativas y relevantes al objetivo.

¹⁰ Wayne W. Daniel. *Estadística con aplicaciones a las ciencias sociales y a la educación*. McGraw-Hill. México, 1986

¹¹ Berenson, Mark L. y David M. Levine. Op. cit.

El modo de la respuesta, abierta o cerrada, directa o indirecta, está en relación a la recolección y codificación de los datos. Cuando las respuestas son abiertas se da libertad al encuestado para que se exprese en el asunto. Si éste tiene facilidad de palabra y es concreto, el resultado será positivo. Sin embargo, si no expresa adecuadamente lo que piensa, esta forma arrojará resultados limitados. Bajo la modalidad de respuesta directa, el encuestado se concreta a lo preguntado. En cambio, cuando se hace de manera indirecta, el cuidado en la codificación de los datos recibidos debe ser esmerado.

Por otro lado, la entrevista personal es el modo más costoso. Implica un esfuerzo adicional de adiestramiento a fin de tener la seguridad de que el entrevistador podrá extraer con habilidad la información deseada. Su ventaja principal estriba en que puede haber rectificaciones a posibles ambigüedades que se den en las respuestas.

Dado que el muestreo estadístico es el preferido para realizar las encuestas, es conveniente apuntar que es necesario elegir el tipo de muestra para realizar la encuesta. Así, los tipos de encuesta pueden clasificarse en dos grupos: con muestra probabilística y con muestra no probabilística.

La encuesta por muestra no probabilística suele ser más sencilla y económica. Los procedimientos sugeridos pueden ser: muestras de juicio, muestras de cuota y el trozo de pastel. Los muestreos de juicio y de cuota otorgan demasiada discreción al entrevistador en el proceso de selección de los entrevistados, lo cual es una desventaja. En el otro extremo, la muestra con trozo de pastel, se compone por un proceso de auto selección (v. gr. el público en un teatro en particular), donde el enumerador no interviene en el proceso de selección.

Así, el resultado de ellas carece de valor probabilístico para hacer inferencias de una población. "La única forma correcta en la cual el investigador podría hacer inferencias estadísticas de una muestra a una población e interpretar los resultados en forma probabilística, es por medio del uso de una muestra probabilística."¹²

I.A.4. Diseño, objetivo y planeación de la encuesta

Aunque el diseño, el objetivo y la planeación de una encuesta pueden variar en complejidad dependiendo de factores culturales, sociales, económicos y hasta de percepción entre los que hacen la encuesta y quienes responden, los pasos principales son:

- Establecer con claridad, exponer y recordar, desde el inicio hasta el final, el objetivo de la encuesta. Debe haber una necesidad de realizar la investigación y saber exactamente qué se busca, cuánta precisión se requiere y cuándo deben entregarse los resultados. "La enunciación de metas y objetivos resulta más fácil si la encuesta trata de contestar una sola pregunta. Si existieran diversos objetivos, cabría asignarles un valor relativo."¹³
- Hacer que coincida la población muestreada con la población objeto de la cual se desea hacer una inferencia. Ello implica definir conceptos, establecer criterios y capacitar al encuestador para que decida en el campo si el sujeto encuestado pertenece o no a la población de estudio.

¹² Berenson, Mark L. y David M. Levine. Op. cit.

¹³ Raj, Des. Op. cit.

- Rectificar que los datos recolectados son relevantes para el objeto de la encuesta y que no se han omitido datos esenciales. Por lo regular, un cuestionario o entrevista demasiado extensa disminuye la calidad de las respuestas. Además, se debe evitar preguntar cosas que no contribuyen al objetivo.
- Especificar cual es el grado de precisión que se desea. Dado que se recaban datos de una muestra, es necesario considerar que pueden darse errores de muestreo o de respuesta y que éstos deben estimarse.
- Elegir el método de medición acorde con el costo, tiempo y objeto de estudio. Así, la encuesta puede emplear un cuestionario auto administrado, una entrevista, la declaración en un examen, una inspección, entre otros.
- Antes de obtener la muestra, la población debe ser dividida en partes denominadas *unidades de muestreo*. Estas unidades deben cubrir toda la población, evitando sobreponerse de manera que un elemento pertenezca a dos unidades. Por ejemplo, en el muestreo de los integrantes de una universidad, la unidad puede ser un alumno, un grupo o una facultad determinada.
- Seleccionar la muestra en base al objetivo, costo y tiempo disponible.
- Probar la encuesta en una pequeña escala para mejorarla en términos de claridad y extensión. Esto puede ayudar a modificar o identificar cambios que deban ser hechos antes de la ejecución

planeada, tomando en cuenta los comentarios de los entrevistados y la experiencia del investigador.

- Organizar la ejecución basada en la capacitación, supervisión y en verificación anticipada para hacer los ajustes pertinentes.
- Editar la encuesta obtenida para corregir errores y desechar datos equivocados.
- Presentar los resultados junto con los estimadores de errores esperados.
- Hacer de la experiencia, conocimiento para futuras encuestas, evitando fallas anteriores y perfeccionando el proceso.

I.A.5. La población y sus valores

Una población es un grupo de elementos más pequeño que un universo, pero más grande que una muestra. El término universo cubre sucesos o cosas que no tienen límite numérico, como son todas las tiradas posibles de un par de dados. Con ello, los elementos son las unidades individuales que constituyen una población y pertenecen a un universo.

Definir la población en relación a un objetivo requiere de identificar claramente las unidades elementales de la población. Si su definición es operativamente factible, ello determinará que obtener datos de la población sea accesible. En caso de que existan dificultades para acceder a la población debe elaborarse una población de trabajo, también llamada marco o población muestreada. Entre más se

parezca ésta a la población blanco u objetivo, el resultado del estudio estadístico tendrá mayor validez.

El hecho de que cualquier población tenga propiedades particulares y que las variaciones en sus elementos sean claramente limitadas hacen posible escoger al azar una muestra relativamente pequeña que refleje las características de la población. "Aunque la diversidad es una cualidad universal de los datos, no hay virtualmente ninguna población estadística real cuyos elementos varíen entre sí sin límite. Así el arroz varía, en un grado limitado, en longitud, peso, color, etcétera, pero siempre puede ser identificado como arroz."¹⁴

De igual forma, las poblaciones mantienen regularidad o uniformidad en sus características. Las fuerzas que la hacen cambiar están, por lo general equilibradas para generar valores por arriba o por debajo de cierto valor central alrededor del cual tienden a agruparse la mayor parte de los valores de sus elementos. Por ejemplo, son pocos los estudiantes con calificaciones altas o muy bajas, por lo regular en las escuelas el promedio prevalece.

Dada la uniformidad, si se escoge una muestra grande al azar sus características diferirán poco de las de la población. De igual forma, dada la diversidad antes citada, si se toman algunas muestras al azar, las muestras nunca coincidirán completamente unas con otras. Con ello, los promedios son más estables que los valores individuales. Además, entre mayor sea la muestra más estable será el promedio. Por ejemplo, si una moneda se lanza al aire 5 veces es probable que el resultado se oriente hacia un lugar determinado. Sin embargo, si se lanza 30 veces el promedio tenderá a prevalecer.

¹⁴ Chou, Ya-Lun. *Análisis Estadístico*. 2ª Ed. Interamericana. México. 1977

Derivado de lo anterior, "el incremento de la estabilidad de un promedio, hace posible calcular de qué magnitud ha de ser una muestra para que su promedio infunda un grado deseable de estabilidad (un error estándar bajo)."¹⁵

Con todo ello, las poblaciones pueden ser finitas o infinitas. Por ejemplo, las calificaciones en un salón de clases es una población finita, pero si se desean las calificaciones en el pasado, presente y futuro de una escuela, la población es infinita. Es decir, definiendo el objeto a estudiar se define la población que pertenece a un universo que contiene todo.

Las estimaciones básicas de una población son:

- la media o el promedio de la población; por ejemplo, el promedio de computadoras utilizadas en las pequeñas empresas,
- el total de la población; por ejemplo, el total de perros vacunados en una colonia determinada,
- la proporción de la población respecto a un universo; por ejemplo, el porcentaje de niños que aprobaron el año,
- la razón de la población en base a otra población; por ejemplo, la razón de hombres a mujeres.

Tomando en cuenta el concepto de población y elemento, el valor de población es aquel que resulta de la medición de cada elemento de la población. Es una expresión numérica que sintetiza los valores de una o varias características de los elementos de una población completa. Algunos valores de la población relacionados con la media son las proporciones, medianas y el valor agregado total de la población. Otros valores de la población miden relaciones. Los más comunes

¹⁵ Chou, Ya-Lun. *Análisis Estadístico*. 2ª Ed. Interamericana. México. 1977

de estos son la diferencia entre dos medias, el coeficiente de regresión y el coeficiente de correlación

Sin embargo, cabe reconocer que los elementos de la población pueden mentir o no conocer sobre sus propios datos. Por ejemplo, si se pregunta la edad a una persona, ésta puede engañar, puede equivocarse o simplemente decir la verdad. Con ello estamos ante un valor real y otro valor de estimación de la población que surge a partir de un censo.

La diferencia entre el valor real y el estimado es lo que da origen a los sesgos o errores. "Los sesgos pueden resultar de errores no aleatorios en la medición, la estimación u otras fuentes. Su calidad crítica reside en que es un error sistemático (conduce a la estimación en una u otra dirección). Por ejemplo, una estimación de la edad puede ser sesgada en sentido descendente si existe una tendencia en los entrevistados a aparecer más jóvenes de lo que realmente son."¹⁶

Comparado el error derivado de la observación de la población con el error de la muestra, el primero es generalmente pequeño comparado con la varianza de muestreo. "De ahí que el valor de la población, por lo general se considere una constante desconocida y se desprece el error de medición."¹⁷

B La Investigación de Mercados

I.B.1. Antecedentes

¹⁶ Lininger, Charles A. y Donald P. Warwick. *La Encuesta por Muestreo: Teoría y Práctica*. CECSA. México. 1984

¹⁷ Kish, Leslie. *Muestreo de encuestas*. Trillas. México. 1982

“Para hablar de la investigación de mercados, es necesario mencionar que una empresa realiza en general tres actividades principales que son: la inteligencia (actividades que recopilan información para tomar decisiones operativas ó estratégicas), las operaciones (aspectos básicos de un negocio como manufactura, ventas, finanzas, etc.) y el desarrollo de estrategia.”¹⁸ Estas actividades forman parte del concepto de mercadotecnia, que es una orientación de la administración que afirma que para que una empresa alcance sus metas, es necesario averiguar las necesidades y deseos del mercado. De esta manera dicha empresa puede ofrecer de una manera más eficiente que la competencia, las satisfacciones deseadas por el mercado.

La investigación de mercados se ha definido como “La recopilación, procesamiento y análisis de información, respecto a temas relacionados con la mercadotecnia. Comienza con la definición del problema y termina con un informe y recomendaciones de acción.”¹⁹

Una definición más completa es la siguiente: “La investigación de mercados es la función que enlaza al consumidor, al cliente y público con el experto en mercadotecnia, a través de la información utilizada para identificar y definir oportunidades y problemas de mercadotecnia; generar, depurar y evaluar las acciones de mercadotecnia, verificar el desempeño de la misma, y mejorar la comprensión del concepto de mercadotecnia como un proceso.

La investigación de mercados especifica la información requerida para afrontar estos problemas, diseña el método para recopilar información; maneja e

¹⁸ Donald R. Lehmann. Investigación y análisis de mercados. Compañía editorial continental S.A. de C.V. México, 1998

¹⁹ Donald R. Lehmann. Investigación y análisis de mercado. Compañía editorial continental S.A. de C.V. México, 1998.

implementa el proceso de recopilación de datos; analiza los resultados y comunica los hallazgos y sus implicaciones."²⁰

Estas definiciones, llevan a deducir que la función central de la investigación de mercados consiste en ayudar a los gerentes en la toma de decisiones, pues proporciona información sumamente útil y constituye un componente indispensable del sistema de información de mercadotecnia.

Este sistema de información proporciona datos diversos dirigidos principalmente a entender las cambiantes necesidades y deseos de los consumidores (gustos y preferencias); detectar nuevas aplicaciones de los productos existentes; desarrollar nuevos productos y medir su potencial de éxito antes de su lanzamiento al mercado; medir la eficiencia de la publicidad que se está manejando; conocer cuáles son las tendencias de los mercados; evaluar la factibilidad de éxito de nuevos proyectos de inversión; conocer a la competencia. Con todo ello, se tienen elementos de información claves para el desarrollo de estrategias adecuadas que aseguren una posición competitiva en los mercados

Considerando lo anterior, se puede decir que los problemas que atiende la investigación de mercados son principalmente:

- Pronóstico. Una de las tareas más obvias de la investigación de mercados es pronosticar las ventas, tanto de bienes de consumo como de servicios.
- Análisis del comprador: Estudiar a los compradores para encontrar las características de los usuarios de diferentes marcas, a fin de asignar recursos de una manera más eficiente.

²⁰ Donald R. Lehmann. Op. cit.

- **Procesos de elección y de información:** Esta es la forma más común de investigación que se realiza dentro de la mercadotecnia, y se refiere a la forma en que los compradores obtienen información de un producto y sobre cómo hacen su elección.
- **Pruebas de factor:** La base de la mayor parte de la investigación de mercados es seleccionar entre diferentes combinaciones y niveles los diversos factores que forman parte de la mercadotecnia, como son precio, nivel de publicidad, promoción, envase, etc.

Para llevar a cabo estas actividades, la investigación de mercados abarca desde la encuesta y el análisis de los datos, hasta la elaboración de estadísticas que permitan analizar las tendencias en el consumo. De esta manera, se podrá prever la cantidad de productos y la localización de los mercados más rentables para un determinado tipo de bien o servicio.

Asimismo, dado que la investigación de mercados abarca diversos aspectos, tiene algunas ramas asociadas, como son:

- **Psicología y Sociología,** que permiten identificar elementos clave de las inclinaciones de las personas, de sus necesidades, sus actividades, circunstancias y motivaciones generales, todos ellos factores clave para entender los distintos patrones de comportamiento de los consumidores.
- **Microeconomía,** de donde surgió la teoría de la utilidad y de algunos conceptos afines.
- **Estadística,** de donde se toman la mayor parte de los procedimientos analíticos que permiten analizar los datos recabados.

- Diseño experimental, de donde se han tomado los conceptos fundamentales de pruebas y del diseño de investigaciones.

Todas estas disciplinas se conjuntan para llevar a cabo la investigación de mercados y, en ocasiones, los estudios son tan complejos, que a través del tiempo han surgido agencias especializadas en esta actividad. Además, dado que la investigación de mercados ha llegado a ser una prioridad para gran parte de empresas a nivel mundial, existen agencias internacionales que surgieron para dedicarse en especial a esta actividad. En el caso de México, también se ha dado este fenómeno, sin embargo eso se tratará más ampliamente en el siguiente apartado.

I.B.2. Investigación de Mercados en México

A raíz de los cambios económicos y a la globalización que enfrenta nuestro país, aumentó la necesidad de contar con información mercadológica relevante y útil para la Toma de Decisiones, pues resulta imprescindible y prioritario asegurar la permanencia de las empresas en el mercado, debido sobre todo a la mayor competencia que existe, no sólo en México, sino también a nivel mundial. Para ello, es esencial entender las cambiantes necesidades y deseos de los consumidores (gustos y preferencias); detectar nuevas aplicaciones de los productos existentes; desarrollar nuevos productos y medir su potencial de éxito antes de su lanzamiento al mercado; medir la eficiencia de la publicidad que se está manejando; conocer

cuáles son las tendencias de los mercados; evaluar la factibilidad de éxito de nuevos proyectos de inversión; y conocer a la competencia.

"Ante esta realidad, el entendimiento del consumidor /usuario en nuestros días es, sin duda alguna, una ventaja competitiva que en el largo plazo puede ser fundamental en el desenvolvimiento de nuestro producto o servicio en el mercado, no importando en que categoría participemos"²¹

Ante estas necesidades y ante lo competitivo del mercado, fue necesario establecer lineamientos generales que unificaran los criterios relacionados con la investigación de mercados. Fue por ello que en septiembre de 1992 se creó la AMAI (Asociación Mexicana de Agencias de Investigación de Mercado y Opinión Pública), "con la finalidad de establecer normas de calidad y estándares comunes en métodos, técnicas, terminología, criterios de análisis, etc., que estén relacionados con la investigación de mercados en nuestro país".²²

Esta asociación agrupa a las compañías de investigación de mercados con mayor prestigio en México, aunque también existen compañías que sin pertenecer a esta asociación, también proporcionan estudios de calidad en los servicios que prestan.

Algunas de las empresas más importantes son:

- BIMSA (Buró de investigación de mercados, S.A.)
- Pearson research

²¹ Tomado de www.merc.com.mx/empresa.html

²² www.amai.org/espanol.shtml

- IDM (Investigación de mercados)
- IEDSA (Investigación Estadística y Demográfica S.A. de C.V.)
- MERC Analistas de Mercados S.A. de C.V.
- Énfasis
- AC Nielsen

I.B.3. Servicios

Actualmente, la investigación de mercados se utiliza en prácticamente toda empresa que produzca determinados artículos o que proporcione algún tipo de servicio. Numerosas empresas solicitan estudios de agencias dedicadas a la investigación de mercados. En este apartado se mencionarán en forma general los servicios que ofrecen algunas de las agencias más importantes del país, ya que ofrecen un amplio panorama de las actividades que se están llevando a cabo en el área de la investigación de mercados a nivel nacional.

En lo que respecta a "Investigación Estadística y Demográfica, S.A. de C.V.", ésta "tiene como misión asesorar, desarrollar estudios y capacitar en las áreas de investigación de mercados, estadística y matemática aplicada a instituciones privadas y de gobierno mediante la utilización de metodologías y procedimientos de carácter científico, con el propósito de coadyuvar al desarrollo de sus clientes, promover el continuo mejoramiento de sus empleados, el beneficio de sus

accionistas y el crecimiento de la empresa.²³ Los servicios que proporciona son los siguientes:

- 1) Investigación de mercados, que comprende:
 - Estudios de mercado
 - Pruebas de calidad
 - Investigación de nuevos productos
 - Perfiles y estilos de vida de clientes
 - Análisis de publicidad (penetración, información, acción)
 - Pruebas de presentación del producto (empaquete, visibilidad, imagen, función)
 - Investigación de imagen y posicionamiento
 - Evaluación de la satisfacción de servicios y productos

Para efectuar estos estudios, se auxilia de otros servicios como son:

- 1) Información Estadística
- 2) Demografía
- 3) Matemáticas financieras
- 4) Capacitación y desarrollo
- 5) Captura de datos

En PEARSON, se llevó a cabo un acuerdo comercial con la empresa Envirosell de Estados Unidos. Con ese acuerdo, la empresa ha cubierto tres nuevas áreas de investigación:

²³ www.iedsa.com.mx

1) *Category Management*, con la cual se pretende entender la dinámica de cómo se compra una categoría determinada de productos y cómo se compra en los diferentes canales de venta al menudeo, ayudando al cliente en el desarrollo de estrategias de mercadeo en el interior de las tiendas, basado en el comportamiento real del cliente, conociendo la psicología del comprador.

2) Planificación de tiendas al menudeo. Que consiste en trabajar directamente para minoristas, normalmente para analizar un nuevo diseño de la tienda o desarrollar pautas para nuevos diseños, basándose en los patrones actuales de compra de las tiendas existentes.

3) Comportamiento de las tiendas de servicio. Se pretende mejorar la eficiencia del servicio para una mejor relación con el cliente, aportando ideas en el diseño y funcionalidad de espacios, basado en el comportamiento real del cliente.

En general los estudios de mercado se realizan en empresas dedicadas a diversas áreas, como por ejemplo:

- Alimentos
- Alimentos para animales
- Bebidas
- Cuidado personal
- Industria de medios de comunicación
- Industria farmacéutica (productos éticos y OTC)
- Limpieza en el hogar

- Maquinaria industrial
- Licores
- Telefonía
- Servicios
- Industria textil
- Industria automotriz
- Hábitos de navegación y publicidad en Internet
- Tiendas departamentales
- Pinturas
- Análisis de publicidad (penetración, información, acción)
- Pruebas de presentación del producto (empaquete, visibilidad, imagen, Función).
- Investigación de imagen y posicionamiento
- Evaluación de la satisfacción de servicios y productos

Existen estudios que están más enfocados al ámbito social y político, y en PEARSON, se efectúan los siguientes:

Ambito Social	Ambito Político
Servicios Públicos	Evaluación a Campanías Políticas
Planificación Familiar	Evaluación a Candidatos
Problemas Ecológicos	Votación Electoral y Tendencias
Alcoholismo y Drogadicción	Simulación del Voto
Impartición de Justicia	<i>Exit Polls</i>
Seguridad Social	
Atención a Problemas de Salud Pública	

CAPITULO II

EL MUESTREO PROBABILÍSTICO

II.1. Generalidades

En la actualidad, a menudo nos encontramos con información en prácticamente todas las esferas de la actividad humana. Hay estudios de población para saber el número de personas en un área, su distribución por sexo y edad; el número de nacimientos y de muertes, así como el porcentaje de migración. Hay estudios de problemas laborales como el número de horas trabajadas y el monto de los salarios pagados; la estructura ocupacional de la población y el número de desempleados. Asimismo, hay datos en el sector industrial: el número de establecimientos por clase, el valor de la producción y la cantidad de artículos defectuosos en un determinado lote. En fin, la lista es mucho más amplia.

Se sabe que toda esta información se obtiene mediante la estadística, pues uno de sus objetivos es el de hacer inferencias para auxiliar en la toma de decisiones, de ahí que sea una de las ciencias de las cuales se auxilia la investigación de mercados para llevar a cabo sus análisis. Sin embargo, se puede preguntar ¿Cómo se obtiene toda esta información?, ya que es prácticamente imposible analizar a cada momento la totalidad de la población dados los costos y el trabajo tan complejo que ello implica. La respuesta es el muestreo.

El muestreo permite conocer las características de una población de personas u objetos con sólo examinar a unos cuantos elementos de dicho grupo. A primera vista, la idea puede parecer simple o incluso absurda, lo que causa que muchas personas desconfíen del muestreo o bien que estén en contra de él. Y hay que reconocer que en muchas ocasiones tienen razón, pues hay estudios mal diseñados que conducen a resultados erróneos.

Un estudio de muestreo bien realizado puede ser tan útil y exacto como puede serlo un censo. Ahora la pregunta es: ¿Cómo sabemos que los elementos seleccionados realmente nos darán una idea precisa de las características de toda la población?

Existen diversas formas de elegir a dichos elementos, lo que nos conduce necesariamente a hablar sobre la técnica del muestreo. Para ello es necesario definir algunos conceptos básicos.

Población: Es un grupo de elementos más pequeño que un universo, pero más grande que una muestra. El término universo cubre sucesos o cosas que no tienen límite numérico, como son todas las tiradas posibles de un par de dados.

Población objetivo: Es el conjunto completo de objetos que se desea estudiar. La población depende del planteamiento inicial de la investigación. "Así, si el objetivo es describir el uso que hacen los niños de la televisión, lo más factible es que debemos interrogar a un grupo de niños. También serviría entrevistar a las mamás de los niños. Escoger entre los niños, sus mamás o ambos, dependería no sólo del objetivo de la investigación, sino del diseño de la misma"²⁴

Muestra: Es un subconjunto de la población.

²⁴ Sampieri Hernández R., Collado Fernández C. Metodología de la Investigación. McGraw-Hill. 1991.

Marco: Es la lista de donde se selecciona la muestra. Puede ser un listado de personas, un directorio telefónico o una lista de votantes registrados. Sin embargo, pueden tener como inconveniente el que no estén actualizados.

Para algunos diseños muestrales se requieren marcos múltiples.

Unidad elemental o unidad de análisis: Es toda aquella unidad contenida en la población cuyas características serán medidas, es decir, es el sujeto o sujetos a quienes en última instancia se les va a aplicar el instrumento de medición (encuesta o entrevista).

Unidades de muestreo: Son las unidades donde se realiza la muestra. Se definen de acuerdo al diseño que se utilice. Con todo lo anterior, podemos decir que una muestra se define como un conjunto de unidades de muestreo elegidas como el microcosmos representativo para realizar inferencias sobre una población.

Así, existen dos formas de elegir una muestra:

- al azar
- por juicio

La muestra al azar se escoge por medio de la aplicación de métodos de probabilidades, según los cuales el investigador no puede decidir qué unidades serán incluidas en la muestra. Particularmente, en la muestra al azar puede medirse el error de muestreo y controlarse éste a través de la teoría de probabilidades. De igual forma, los errores sistemáticos de selección y la falta de respuesta pueden ser eliminados o por lo menos controlados dentro de ciertos límites.

Por otra parte, una muestra por juicio es tomada de acuerdo con una percepción personal. "Los criterios que intervienen para seleccionarla son resultado de la experiencia del investigador para detectar, según él, la representatividad de la muestra. En consecuencia, la probabilidad de que cada elemento individual sea extraído de la muestra es desconocida y la fidelidad de sus resultados no pueden ser objeto de un análisis científico."²⁵

Cuando la muestra sea realmente representativa, la opinión sobre el resto de la población será tan válida como si se tuviera tiempo y dinero para revisar toda la población.

II.2. Tipos de muestreo

Con base en lo anterior, se deduce que, en general, existen dos tipos de muestreo: el muestreo no probabilístico y el muestreo probabilístico. Para efectos de este trabajo, sólo vamos a considerar el segundo.

Entre los tipos de muestreo probabilístico, existen:

- el muestreo aleatorio simple
- el muestreo estratificado
- el muestreo por conglomerados
- el muestreo sistemático

²⁵ Chou, Ya-Lun. Análisis Estadístico. 2° Ed. Interamericana. México. 1977

II.2.1. El muestreo aleatorio simple

Este tipo de muestreo es la forma más sencilla de realizar un muestreo. En este caso no solo todas las unidades tienen la misma probabilidad de ser incluidas en la muestra, sino que también todas las combinaciones han de ser igualmente probables.

“La elección al azar que asegura la aleatoriedad de un procedimiento de muestreo solo puede ser mantenida por métodos ordenados y deliberados. Esto equivale a decir que, para asegurar la elección al azar, el proceso de muestreo debe ser correctamente controlado.”²⁶

El procedimiento correcto para elegir una muestra al azar consiste en asignar un número a cada elemento de la población, para después elegir a los elementos de acuerdo a una tabla de números aleatorios. Los números seleccionados corresponderán a los elementos que van a formar parte de la muestra.

El muestreo aleatorio se puede plantear bajo dos puntos de vista:

- sin reemplazo
- con reemplazo

En el muestreo aleatorio sin reemplazo, si se considera una población formada por N elementos y se observa un elemento particular e , sucede que: la probabilidad de que e sea elegido en primer lugar es $1/N$.

²⁶ Chou, Ya-Lun. Op. Cit.

Si no ha sido elegido en primer lugar, lo que ocurre con una probabilidad de $(N-1)/N$, la probabilidad de que sea elegido en el segundo intento es de $1/(N-1)$. En el $(i + 1)$ ésimo intento, la población consta de $N - i$ elementos, con lo cual si e no ha sido seleccionado previamente, la probabilidad de que lo sea en este momento es de $1/(N-i)$.

Ahora, si se considera una muestra de $n \leq N$ elementos, donde el orden en la elección de los mismos tiene importancia, la probabilidad de elección de una muestra $M = (e_1, e_2, \dots, e_n)$ es:

$$P(M) = \frac{1}{\binom{N}{n}} = \frac{n!(N-n)!}{N!}$$

donde $\binom{N}{n}$

es el total de muestras posibles y cada una igualmente probable.

Por otra parte, en el muestreo aleatorio con reposición sobre una población E de tamaño N podemos realizar extracciones de n elementos, de modo que cada elemento extraído es repuesto al total de la población.

De esta forma, el elemento n puede ser extraído varias veces. Si el orden en la extracción de la muestra interviene, la probabilidad de una cualquiera de ellas, formada por n elementos es:

$$\frac{1}{N^n}$$

El muestreo aleatorio con reposición es también denominado muestreo aleatorio simple, el cual se caracteriza porque:

- cada elemento de la población tiene la misma probabilidad de ser elegido,
- las observaciones se eligen con reemplazamiento. De este modo, cada observación es realizada sobre la misma población (no disminuye con las extracciones sucesivas)

El muestreo aleatorio simple es útil cuando no se dispone de información adicional que permita diseñar la encuesta. Por ejemplo, puede ser que sólo se disponga de una lista de nombres de personas, pero que no se tenga más información acerca del área donde viven o del trabajo que desempeñen.

También es posible que la investigación a realizar no requiera considerar el dividir a la población en conjuntos más pequeños como estratos o conglomerados. De cualquier manera, el muestreo simple aleatorio o el muestreo sistemático quizá sean las mejores estrategias del muestreo de probabilidad, además de ser los más sencillos de analizar.

Existen principalmente tres estimaciones a realizar cuando se analiza una población: la estimación de la media poblacional, el total poblacional y las proporciones en que se dan ciertas características a estudiar. Estas estimaciones en el muestreo simple aleatorio se calculan de la siguiente manera:

Media muestral

$$\bar{y} = \frac{1}{n} \sum_{i \in S} y_i$$

Donde S es el conjunto de valores seleccionados para la muestra.

La varianza de la población S^2 se estima mediante la varianza de la muestra

$$s^2 = \frac{1}{n-1} \sum_{i \in S} (y_i - \bar{y})^2$$

Un estimador insesgado de la varianza de y es:

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

El error estándar es la raíz cuadrada de la varianza

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

El coeficiente de variación de una estimación proporciona una medida de la variabilidad relativa de una estimación. Se calcula al dividir el error estándar entre la media.

$$CV(y) = \frac{EE(y)}{y}$$

Para el total de la población se tiene lo siguiente:

$$\hat{t} = Ny$$

La varianza del total de la población es:

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

El error estándar es

$$EE(\hat{t}) = N \cdot EE(y)$$

Para el coeficiente de variación:

$$CV(\hat{t}) = CV(y)$$

Si lo que se desea estimar es la proporción de las unidades en la población que tengan cierta característica, se tiene:

$$p = \sum_{i=1}^N \frac{y_i}{N} = \bar{y}$$

$$\hat{p} = \bar{y}$$

\hat{p} es un estimador insesgado de p

La varianza de la muestra se calcula de la siguiente manera:

$$S^2 = \frac{1}{n-1} \sum_{i \in S} (y_i - \hat{p})^2 = \frac{n}{n-1} p(1-p)$$

La varianza de la distribución de muestreo para proporciones es:

$$V(\hat{p}) = \left(1 - \frac{n}{N}\right) \frac{p(1-p)}{n-1}$$

$$EE(\hat{p}) = \sqrt{\left(1 - \frac{n}{N}\right) \frac{p(1-p)}{n-1}}$$

II.2.2. El muestreo estratificado

En este caso la población se divide en subgrupos llamados estratos. "Los estratos no se traslapan y conforman la población completa, de modo que cada

unidad de muestreo pertenece exactamente a un estrato²⁷. Después de llevar a cabo esta división, se utiliza el muestreo simple aleatorio para obtener una muestra de cada estrato. Dichos subgrupos se eligen de acuerdo a los objetivos de la investigación.

Existen diversas razones por las cuales se recurre al muestreo estratificado. Una de ellas es para disminuir las varianzas de las estimaciones de la muestra. Por ejemplo, en el muestreo proporcional, el tamaño de la muestra se elige en forma proporcional al tamaño del estrato. "La varianza disminuye según el grado en que difieran las medias por estrato y el grado de homogeneidad existente dentro de los estratos"²⁸

Por otra parte, en un muestreo no proporcional, se eligen diferentes razones de muestreo en los diferentes estratos. Para disminuir la varianza se elige una fracción de muestreo mayor en aquellos estratos en los que hay más variación, y una fracción de muestreo menor en los estratos que presenten más homogeneidad en sus elementos.

Otra de las razones por las que se utiliza este tipo de muestreo es para obtener una muestra más representativa que la que se podría obtener con el muestreo simple aleatorio, pues éste último puede tener la desventaja de dejar fuera de la muestra a una parte de la población que sea importante para el estudio a realizar.

Asimismo, es posible que se quieran obtener datos de un subgrupo en particular, por ejemplo, puede interesarnos saber lo que piensen los adolescentes

²⁷ Lohr, Sharon I. Muestreo: Diseño y Análisis. Thomson. 1999

²⁸ Kish, Leslie. Muestreo de encuestas. Trillas. 1972

acerca de un determinado tema, por una parte, y por otra saber lo que piensan los adultos de treinta a cuarenta años.

Si se hubiera utilizado el muestreo simple aleatorio, es posible que se hubiera dejado fuera de la muestra una gran parte de la población adolescente o bien de la población adulta. Mediante el muestreo estratificado, se pueden considerar dos estratos para asegurarse de que ambos grupos estarán representados.

Finalmente, otra de las ventajas de este tipo de muestreo es que puede administrarse a un menor costo, pues se pueden aplicar diferentes esquemas de muestreo en los diferentes estratos.

En general, el muestreo estratificado consta de los siguientes pasos:

- a) La población completa se divide en subpoblaciones llamadas estratos.
- b) Dentro de cada estrato se selecciona una muestra a partir de todas las unidades de muestreo que componen ese estrato
- c) Se calcula una media para cada estrato. Posteriormente se ponderan adecuadamente para formar una estimación combinada de la población completa.
- d) Las varianzas también se calculan en forma separada para cada estrato y después se ponderan. De esta forma se obtiene una estimación combinada para la población.

Por supuesto que estos pasos pueden modificarse según el diseño de muestreo que se elija.

II.2.2.1. Notación

Dado que en el muestreo estratificado la población se divide en subgrupos o estratos, cada estrato tendrá N_h elementos, de manera que

$$N_1 + N_2 + N_3 + \dots + N_H = N$$

Donde N es el total de elementos de la población, H es el total de estratos en que se dividió a la población y h es el estrato a considerar.

El caso más sencillo del muestreo estratificado es el muestreo aleatorio estratificado, en el cual se elige una muestra aleatoria en cada uno de los estratos en que se ha dividido a la población. De esta manera, en el estrato h los elementos que pertenecen a la muestra elegida se denotan por n_h , de manera que:

$$n = \sum n_h = n_1 + n_2 + \dots + n_h + \dots + n_H$$

n_h es el número de elementos que integran la muestra en el estrato h

S_h es el conjunto de las n_h unidades en la muestra aleatoria simple para el estrato h .

y_{hj} es el valor de la unidad j en el estrato h

Al utilizar las estimaciones de la muestra aleatoria simple dentro de cada estrato, se tiene:

Para la media y para el total muestral

$$y_h = \frac{\sum_{j \in S} y_{hj}}{n_h}$$

En este caso la media de la muestra se obtiene de forma separada e independiente para cada estrato.

El total de elementos se obtiene al multiplicar la media del estrato por el total de elementos del estrato. Después se suman los productos de todos los estratos, de la siguiente manera:

$$\hat{t}_h = \frac{N_h}{n_h} \sum_{j \in S_h} y_{hj} = N_h \bar{y}_h$$

La varianza se obtiene de la siguiente forma:

$$S_h^2 = \sum_{j \in S_h} \frac{(y_{hj} - \bar{y}_h)^2}{n_h - 1}$$

Para proporciones se tiene lo siguiente:

$$\hat{p}_{est} = \sum_{h=1}^H \frac{N_h}{N} \hat{p}_h$$

El factor N_h / N es la razón entre el total de elementos del estrato h con respecto al total de elementos de la población.

Para calcular la varianza de la distribución de proporciones, también se utiliza el factor de corrección para poblaciones finitas n_h / N_h , solo que esta vez es por estrato. De ahí que aparezca el subíndice h .

Con respecto al número total estimado de unidades que pertenecen a la población con la característica, es la suma de los totales estimados en cada estrato. Así tenemos que

$$\hat{t}_{est} = \sum_{h=1}^H N_h \hat{p}_h$$

La varianza es:

$$\hat{V}(\hat{t}_{est}) = N^2 \hat{V}(\hat{p}_{est})$$

Dentro del muestreo estratificado, existe el muestreo proporcional de elementos. En este caso, la fracción de muestreo en cada estrato se hace igual a la fracción de muestreo para la población completa. Es decir, n_h / N_h se hace igual a n/N para toda h . Una vez que se ha decidido cuál será el tamaño de la muestra total, se procede a calcular el tamaño de la muestra para cada estrato, es decir $n_h = f N_h$, donde f es la fracción de muestreo. El redondeo de n_h al entero más próximo introduce ligeras desviaciones en los valores de las fracciones de muestreo, sin embargo estas desviaciones son tan comunes que pueden despreciarse.

Un aspecto importante del muestreo proporcional es que representa a todos los estratos de la muestra con las mismas razones de los estratos en la población: es decir: $n_h / n = N_h / N = W_h$ para toda h .

Este diseño produce una muestra autoponderada, y las estimaciones muestrales son las siguientes:

Media muestral:

$$y_{prop} = \frac{y}{n} = \frac{1}{n} \sum_j y_j$$

Varianza:

$$Var(y_{prop}) = \frac{1-f}{n} (\sum w_h s_h^2) = \frac{1-f}{n} s_w^2$$

La expresión que está dentro del paréntesis es la varianza por elemento dentro del estrato.

Varianza del total de la muestra:

$$Var(y) = (1-f) \sum_h \frac{n_h}{n} \frac{1}{n_h - 1} \left[\sum_i^{n_h} y_{hi}^2 - \frac{y_h^2}{n_h} \right]$$

II.2.3. El muestreo por conglomerados

El muestreo por conglomerados es una técnica en la que se divide la población en grupos o "conglomerados" y luego se procede a extraer una muestra probabilística de ellos.

Cuando los grupos, que son las unidades primarias son extraídos, se pueden incluir en la muestra todas las unidades elementales de los grupos elegidos. Cuando se han observado todas las unidades elementales de dichos grupos, tenemos lo que se conoce como muestreo en una sola etapa. Cuando se extrae una muestra de unidades elementales de los grupos escogidos, se tiene un tipo de diseño de muestreo que se llama muestreo en dos etapas. Cuando el muestreo por conglomerados incluye más de dos etapas para elegir la muestra final, se llama muestreo en múltiples etapas.

El muestreo por conglomerados presenta algunas ventajas sobre otros diseños muestrales. En primer lugar, puede ser que el obtener un marco muestral que liste a todos los elementos de la población sea muy difícil de obtener, si no es que imposible; sin embargo el conseguir una lista de todos los conglomerados puede ser más fácil. Por ejemplo, es más fácil obtener una lista de todas las universidades del país que el tener una lista de todos los estudiantes universitarios a nivel nacional.

En segundo lugar, este procedimiento de muestreo tiene la ventaja de reducir los costos y simplificar el trabajo de campo, pues la selección de los elementos de los grupos por lo general simplifica el proceso de entrevistas, pues es más fácil

muestrear conglomerados que se encuentran juntos físicamente. Esto es particularmente útil cuando se realizan estudios a nivel estatal o nacional.

Existe a veces confusión al hablar de muestreo estratificado y de muestreo por conglomerados.

En la muestra estratificada se seleccionan individuos dentro de cada estrato, por lo que existe la seguridad de que cada estrato está representado por cierto número de casos. Los errores de selección se dan por consiguiente dentro de los estratos. Entonces se requiere que éstos sean en sí mismos lo más homogéneos posible, mientras que entre estratos sean lo más diferentes que se pueda unos de otros.

En el muestreo por conglomerados (de fase única), en cambio, no se tiene fuente de error dentro del conglomerado, pues se utilizan todos los casos que existen dentro de él, pero como se toma una muestra de conglomerados, el error muestra variabilidad entre los conglomerados. "Si las medias de los conglomerados difieren considerablemente en comparación con la variabilidad dentro de los mismos, corremos el riesgo de obtener un conglomerado muy poco usual en la muestra de conglomerados. Si esto ocurriera efectivamente y si los conglomerados fueran homogéneos, nuestro error de muestra podría ser considerable. Pero si los conglomerados son heterogéneos en sí mismos en comparación con las diferencias entre ellos, podemos salir adelante con pocos conglomerados relativamente grandes"²⁹.

²⁹ Blabek, Jr. Humbert M. Estadística Social. Fondo de Cultura Económica. 1983

De esta manera, se trata de obtener estratos homogéneos, pero en cambio, conglomerados heterogéneos, pues la razón de la variancia en la estrategia es la diferencia en cuanto al punto de extracción de la muestra.

Por otra parte, se puede comparar al muestreo por conglomerados con el muestreo simple aleatorio. En casi todos los ejemplos que se tienen, las muestras por conglomerados tendrán mayores errores de selección, lo que indica que son menos eficaces que las muestras simples aleatorias de igual tamaño. Sin embargo, en cuanto a costos, éstos serán menores al obtener muestras por conglomerados. El problema es entonces equilibrar los costos y la eficiencia. Esta última se mide en términos del error estándar de la estimación. Un error pequeño indica una eficiencia alta.

Si un grupo es perfectamente homogéneo, no se necesitan muchos casos para obtener una estimación precisa de su media. En tal caso, se puede tomar una muestra muy pequeña del conglomerado, con lo que se da un ahorro de dinero, el cual podrá destinarse al estudio de otros conglomerados. Así pues, son dos factores los que determinan la eficiencia relativa del diseño de conglomerados. Por una parte el grado de homogeneidad dentro del conglomerado y el tamaño del mismo: si el conglomerado es homogéneo, se seleccionarán sólo unos cuantos casos. Si el conglomerado es heterogéneo, se pueden tomar más casos de cada uno de ellos sin una disminución significativa de la eficiencia.

II.2.3.1. Notación para el muestreo por conglomerados del mismo tamaño.

a) Una etapa.

N .- es el número de conglomerados existentes en la población (unidades primarias)

n .- es el número de conglomerados considerados para la muestra, con igual probabilidad de ser seleccionados

M_i .- son todos los elementos del conglomerado i . Todos estos elementos se incluyen en la muestra cuando se tiene este tipo de muestreo en una etapa.

$K = \sum_{i=1}^N M_i$: Es la suma de todas las unidades de los N conglomerados. En este caso,

todas las unidades se incluyen en la muestra.

$t_i = \sum_{j=1}^{M_i} y_{ij}$: Es el total de elementos en la unidad primaria i . La y_{ij} es el valor del

elemento j del conglomerado i .

$t = \sum_{i=1}^N t_i = \sum_{i=1}^N \sum_{j=1}^{M_i} y_{ij}$: Es el total de la población. En este caso se tiene la suma de los

elementos de los N conglomerados que conforman la población completa.

$$S_i^2 = \sum_{i=1}^N \frac{\left(t_i - \frac{t}{N}\right)^2}{N-1} : \text{Es la varianza en la población}$$

“En la población de N unidades primarias, la unidad i contiene M_i unidades secundarias (elementos). Extraemos de la población una muestra aleatoria simple de n unidades primarias y medimos nuestra variable de interés en cada elemento de la unidad primaria elegida. Así para el muestreo por conglomerados en una etapa, $M_i = m_i$ ³⁰. Es decir, el número de elementos del conglomerado es igual al número de elementos de la muestra.

II.2.3.2. Estimación en el muestreo por conglomerados del mismo tamaño.

Este tipo de muestreo por lo regular no se encuentra en poblaciones que existen de manera natural, como puede ser un conjunto de personas, animales o regiones naturales. Normalmente los conglomerados que tienen el mismo tamaño se encuentran en poblaciones ya diseñadas de antemano, como puede ser el proceso de producción en una fábrica, en la cual los artículos producidos se empaquetan en cajas que contienen el mismo número de elementos. Cada caja se puede considerar como un conglomerado. Sin embargo, el considerar este caso es útil, pues ayuda a entender diversos conceptos, así como a generalizar resultados.

Si los conglomerados tienen el mismo tamaño, entonces $M_i = m_i = M$. Para estimar las medias o totales de la población, sólo se consideran las medias o totales

³⁰ Lohr, Sharon. Op. Cit.

del conglomerado como si fueran las observaciones y se ignoran los elementos individuales.

Para estimar el total, se puede usar el siguiente estimador:

$$\hat{t} = \frac{N}{n} \sum_{i \in S} t_i$$

En este caso, el estimador se asemeja a aquel que se tiene en el muestreo aleatorio simple de n unidades, que ha sido extraída de una población de N unidades. Como resultado, \hat{t} es un estimador insesgado de t , con varianza dada por:

$$V(\hat{t}) = N^2 \left(1 - \frac{n}{N}\right) \frac{S_t^2}{n}$$

y con error estándar dado por

$$EE(\hat{t}) = N \sqrt{\left(1 - \frac{n}{N}\right) \frac{S_t^2}{n}}$$

Donde S_t^2 y s_t^2 son la varianza de la población y de la muestra de los totales de la unidad primaria, respectivamente.

La varianza de la población está dada por:

$$S_t^2 = \frac{1}{N-1} \sum_{i=1}^N \left(t_i - \frac{t}{N} \right)^2$$

La varianza muestral de los totales de los n conglomerados se obtiene de la siguiente manera:

$$s_t^2 = \frac{1}{n-1} \sum_{i \in S} \left(t_i - \frac{t}{N} \right)^2$$

Para estimar la media de la población, se divide el total estimado entre el número de personas, con lo que se obtiene:

$$\hat{y} = \frac{t}{NM}$$

La varianza y el error estándar son:

$$V(\hat{y}) = \left(1 - \frac{n}{N} \right) \frac{S_t^2}{nM^2}$$

$$EE(\hat{y}) = \frac{1}{M} \sqrt{\left(1 - \frac{n}{N} \right) \frac{s_t^2}{n}}$$

II.2.4. El muestreo sistemático

Este tipo de muestreo se puede utilizar cuando los elementos de la población están ordenados en fichas o en una lista, como puede ser una lista de alumnos de una escuela, una lista de posibles votantes o listas de viviendas.

El procedimiento para llevar a cabo este muestreo se da de la siguiente manera:

Sea N el tamaño de la población y n el tamaño de muestra deseado.

Hacemos $k = N/n$. Posteriormente se elige un número aleatorio entre 1 y k . Una vez que se ha seleccionado este número, los elementos de la muestra serán los siguientes:

$$\{ e_m, e_m + k, e_m + 2k, \dots, e_m + (n-1)k \}$$

Por ejemplo, si se tiene una población $N = 2500$ y el tamaño de la muestra es $n = 50$, entonces el intervalo de muestreo $k = 2500/50$, por lo que $k = 50$. Supóngase que el número aleatorio elegido es 34. Entonces los elementos de la muestra serán 34, 84, 134, 184, etc.

Una de las ventajas de este método de muestreo es su simplicidad y fácil administración. Otra de las ventajas se da cuando el criterio de ordenación en la lista es tal que los elementos que la forman tienen un orden aleatorio, pues la muestra será muy similar a la que se obtenga mediante el muestreo aleatorio simple, ya que

se recorre la población de un modo más uniforme. Por otro lado, es más difícil cometer errores con un muestreo sistemático que con el aleatorio simple.

Otra ventaja del muestreo sistemático es que produce fácilmente una muestra proporcional, ya que la muestra se reparte uniformemente en la población mediante un ordenamiento correspondiente de esta última.

Sin embargo, a pesar de las ventajas, el muestreo sistemático también presenta algunas desventajas, pues el método tal como se ha definido es sesgado si k no es entero, ya que los últimos elementos de la lista nunca podrán ser escogidos. Un modo de evitar este problema consiste en considerar la lista como si fuese circular (el elemento $N+1$ coincide con el primero). De este modo se toma como muestra los elementos de la lista que consisten en ir saltando de k elementos en k , a partir de m , teniendo en cuenta que la lista es circular.

Otra de las posibles desventajas es que la muestra obtenida por este método no necesariamente es representativa de la población, sobre todo si los elementos de la lista tienen algún orden periódico o cíclico, pues en este caso se tomarán los elementos que presenten una determinada característica y se dejarán fuera los otros elementos. Por ejemplo, si k es par y si en la lista están alternados hombres y mujeres, entonces la muestra sólo contendrá hombres o solo mujeres.

Ahora bien, en el muestreo sistemático, las unidades no se seleccionan de forma independiente, ya que sólo se elegirán ciertas combinaciones, dependiendo del intervalo de selección y del orden que tienen las unidades de muestreo en la lista. Sin embargo, esto no impide que cada unidad tenga una probabilidad de selección conocida y que pueda llevarse a cabo un análisis de los errores de muestreo.

Se considera que el muestreo sistemático es un caso particular del muestreo por conglomerados, pues al elegir los elementos de k en k , obtenemos determinadas combinaciones, cada una de las cuales se considera como un conglomerado.

CAPITULO III

EL MUESTREO POLIETAPICO

III.1. Concepto

Los diferentes tipos de muestreo probabilístico que se expusieron en el capítulo anterior, con excepción del muestreo por conglomerados, son modelos monoetápicos, es decir, que se aplican directamente a la población a estudiar. Sin embargo, existen encuestas que, por el tipo de investigación a realizar, se llevan a cabo en poblaciones muy grandes, como son aquellas a nivel nacional.

En estos casos, el aplicar un diseño monoetápico, representa grandes inconvenientes, y uno de los más importantes es que eleva demasiado los costos, pues en muchas ocasiones la población está muy dispersa, por lo que es necesario que los entrevistadores se desplacen grandes distancias para realizar las encuestas. Esto representa desperdicio de tiempo y de recursos económicos, ya que se tiene que capacitar a más personas, o se gasta más en desplazar a las ya existentes.

Otro inconveniente del muestreo monoetápico se da cuando no existe, o no se dispone de una lista práctica o fidedigna de los elementos de la población. En este caso es conveniente utilizar el muestreo polietápico. Por ejemplo, es más fácil tener una lista de las universidades del país, que una lista con los nombres de todos los alumnos de dichas universidades.

El muestreo polietápico es una variación del muestreo simple aleatorio, y se caracteriza por operar en sucesivas etapas, empleando en cada una de ellas el método de muestreo probabilístico más adecuado.

El procedimiento básico está delineado por dos etapas, aunque puede extenderse a otras. La primera etapa consiste en dividir a la población objetivo en varios conglomerados de elementos, que se llaman unidades primarias de muestreo (UPM's), y utilizar después procedimientos aleatorios para seleccionar a una gran cantidad de estas unidades para submuestrearlas después.

La primera etapa se diseña de tal manera que las UPM's seleccionadas contengan más elementos que los que se necesitan para la muestra final.

La segunda etapa comprende un submuestreo dentro de las unidades primarias de muestreo seleccionadas para extraer la muestra de elementos, ya sea individuos, viviendas, etc.

Este submuestreo se realiza con cualquiera de las técnicas de muestreo ya vistas anteriormente.

Además, para que el muestreo polietápico se lleve a cabo de la mejor manera, es necesario considerar algunos puntos prácticos:

- "Las UPM's deben cubrir todos los elementos de la población que se va a muestrear. Cada elemento debe identificarse con una u otra UPM, pero ninguno debe aparecer en más de una"³¹. Así se asegura que todo elemento va a tener una probabilidad conocida de ser seleccionado para la muestra en la primera etapa.

³¹ La encuesta por muestreo: Teoría y práctica

- Es necesario que las definiciones y límites de las UPM sean precisos, sobre todo cuando se refieran a áreas geográficas. En éste caso hay que especificar exactamente sus límites, precisando qué es lo que se incluye y qué es lo que no se incluye en dicha área.
- El éxito de una muestra depende de la observación de reglas predeterminadas para la última etapa del muestreo: la selección de entrevistados. Las instrucciones que se imparten a los entrevistadores deben especificar exactamente a quién debe entrevistar y cómo se le debe elegir.
- Una muestra polietápica bien realizada, puede convertirse en la base para muestras posteriores, sobre todo si el trabajo de campo y el marco muestral está bien estructurado. "Esto significa que la definición y formación de los conglomerados, la sistematización de la información de base, la elaboración del material cartográfico y su actualización, son insumos que se pueden aprovechar para desarrollar varias rondas de una misma encuesta o para realizar investigaciones independientes sobre distintos temas, utilizando mecanismos de estratificación apropiados que mantengan una correlación significativa con la variable de diseño."³²

³² Lja encuesta por muestreo: Teoría y práctica

III.2. Etapas del muestreo polietápico

El procedimiento del muestreo polietápico se lleva a cabo en dos o más etapas sucesivas y dependientes, ya que los listados de las unidades de una etapa se limitan a aquellas unidades de muestreo seleccionadas en la etapa anterior.

En general se siguen los siguientes pasos:

- Se forman unidades primarias de muestreo dividiendo a la población en varios grupos. Estos grupos pueden obtenerse ya sea mediante estratificación o mediante formación de conglomerados. La primera opción se refiere a la subdivisión de la población en subconjuntos con características propias.

Este procedimiento tiene como objetivo reducir la varianza del parámetro de interés, por lo que es conveniente que las variables de estratificación estén altamente correlacionadas con aquella(s) utilizada(s) para la determinación del tamaño de muestra. De modo que "los estratos son, por definición homogéneos en su interior, lo que, de paso, establece una diferencia fundamental respecto a las características de los conglomerados."³³

Sin embargo, a pesar de lo anterior, en general los estratos no suelen utilizarse como unidades de selección, ya que todos ellos están representados en la muestra, mientras que los conglomerados se definen como las "típicas unidades de selección en las primeras etapas, ya que su

³³ Planificación y Desarrollo de encuestas de Hogares. www.eclac.cl/deype/noticias/

heterogeneidad garantiza que en su interior queden incluidos diferentes elementos de la población³⁴. Esto quiere decir que los conglomerados son como pequeñas poblaciones con diversas características en su interior.

Ahora bien, cabe mencionar que la formación de conglomerados tiende a incrementar la varianza muestral. Sin embargo se utilizan en el diseño de la muestra, "pues permiten que exista flexibilidad en el número de etapas de selección y reducen substancialmente los costos asociados a la construcción del marco y a la realización del trabajo de campo."³⁵

Una vez que se han definido los conglomerados, se seleccionan aleatoriamente una gran cantidad de los mismos que representen a toda la población.

Supóngase ahora que cada unidad de muestreo se puede dividir en cierto número de unidades más pequeñas ó subunidades.

- En la segunda etapa, se selecciona y mide una muestra de subunidades de cada conglomerado, pues en este tipo de muestreo, la unidad no se mide completamente, sino que a su vez es objeto de un muestreo. A ésta técnica también se le llama submuestreo, ó muestreo en dos etapas.

El submuestreo tiene una gran variedad de aplicaciones, por lo que es muy utilizado al realizar encuestas.

³⁴ Planificación y Desarrollo de encuestas de Hogares. www.eclac.cl/deype/noticias/

³⁵ Planificación y Desarrollo de encuestas de hogares. www.

- En algunos casos, es necesario aplicar una tercera etapa al muestrear las subunidades que se eligieron en la segunda etapa. En este punto es necesario aclarar que en cada etapa se emplea el método de muestreo probabilístico más adecuado.

III.3. Muestreo por conglomerados en dos etapas

A este tipo de muestreo también se le llama muestreo por conglomerados en dos etapas. En este caso, primero se elige una muestra aleatoria simple de unidades primarias (conglomerados), y posteriormente se elige una muestra aleatoria simple de m_i elementos del conglomerado i se denota como S_i .

En este caso, la etapa adicional hace que las estimaciones sean más complicadas, especialmente en lo que se refiere a la varianza, pues se debe tomar en cuenta la variabilidad que existe entre ambas etapas de recolección de datos.

Para estimar el total en el muestreo en dos etapas, se debe considerar que no se observan todas las unidades secundarias en la unidad primaria de la muestra, por lo que hay que estimar los totales individuales de la unidad primaria de la siguiente manera:

$$t_i = \sum_{j \in S_i} \frac{M_i}{m_i} y_{ij} = M_i \bar{y}_i$$

El estimador insesgado del total de la población es:

$$\hat{t}_{ins} = \frac{N}{n} \sum_{i \in S} \hat{t}_i = \frac{N}{n} \sum_{i \in S} M_i y_i$$

En este tipo de muestreo, los totales por conglomerado son variables aleatorias.

Por lo tanto la varianza de t tiene dos componentes:

- La variabilidad entre las unidades primarias y
- La variabilidad de las unidades secundarias dentro de las unidades primarias.

En el caso del muestreo por conglomerados en una etapa, esta variabilidad no se toma en cuenta, ya que se consideran a todos los elementos de cada una de las unidades primarias.

Para el muestreo por conglomerados en dos etapas, la varianza del total es la siguiente:

$$V(\hat{t}_{ins}) = N^2 \left(1 - \frac{n}{N} \right) \frac{s_t^2}{n} + \frac{N}{n} \sum_{i=1}^N \left(1 - \frac{m_i}{M_i} \right) M_i^2 \frac{s_i^2}{m_i}$$

En donde el primer término es la varianza del muestreo por conglomerados en una etapa, y el segundo término es la varianza adicional debida al submuestreo.

El estimador insesgado de la varianza del total es el siguiente

$$\hat{v}(\hat{t}) = N^2 \left(1 - \frac{n}{N} \right) \frac{s_t^2}{n} + \frac{N}{n} \sum_{i \in S} \left(1 - \frac{m_i}{M_i} \right) M_i^2 \frac{s_i^2}{m_i}$$

La media de la población se puede estimar de la siguiente manera:

$$\hat{y}_{ins} = \frac{f_{ins}}{k}$$

y el error estándar es:

$$EE(\hat{y}_{ins}) = \frac{EE(f_{ins})}{k}$$

III.4. Determinación del tamaño de muestra en diseños polietápicos

Las muestras pueden ser de una sola etapa o de varias, según el número de niveles que se hayan determinado al diseñar la encuesta.

Para determinar el tamaño muestral, existe un procedimiento bien definido si el diseño consiste en una sola etapa. En este caso, dicho tamaño está determinado al considerar:

- a) El tamaño del error de la muestra que se desea
- b) El nivel de confianza elegido. Los más comunes son el 90%, el 95% y el 99%.

Así, para un nivel de confianza dado, un error muestral más pequeño implica un tamaño de muestra más grande. De forma similar, para un error de la muestra

dado, un nivel de confianza más alto implica que el tamaño muestral sea más grande.

Para estimar la media de la población, el tamaño de la muestra se calcula de la siguiente manera:

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

donde:

n = tamaño de la muestra

Z = número de unidades de desviación estándar en la distribución normal que producirá el nivel deseado de confianza (para el 95%, $Z=1.96$; para una confianza del 99%, $Z= 2.58$)

σ = desviación estándar de la población (conocida o estimada a partir de estudios anteriores)

E = error, o diferencia máxima entre la media muestral y la media de la población que se está dispuesto a aceptar en el nivel de confianza que se ha elegido.

La mayor dificultad para el cálculo del tamaño muestral para estimar la media de la población consiste en calcular la desviación estándar de la población. En ocasiones se toma dicho dato de trabajos anteriores, pero cuando no son confiables, se emplean estudios con muestras pequeñas para conocer su valor.

Si lo que se quiere es estimar una proporción en la población, el tamaño de la muestra se calcula de la siguiente forma:

$$n = \frac{Z^2 P(1-P)}{E^2}$$

donde:

n = tamaño de la muestra

Z = número de unidades de desviación estándar en la distribución normal, que producirá el grado deseado de confianza (para una confianza del 95%, $Z = 1.96$; para una confianza del 99%, $Z = 2.58$)

P = proporción de la población que posee la característica de interés (en caso de que no se conozca dicho valor, se utiliza $P = 0.5$ en la fórmula)

E = error, o máxima diferencia entre la proporción muestral y la proporción de la población que se está dispuesto a aceptar en el nivel de confianza que se ha elegido.

Generalmente se supone que la muestra es pequeña en comparación a la población total, pero en algunos casos, la muestra es del 5% o más de la población. En este caso el procedimiento varía un poco, pues a medida que el tamaño muestral se acerca al tamaño de la población, desaparece el error muestral.

Para este caso, se tiene lo siguiente al estimar la media de una población finita

$$n = \frac{\sigma^2}{\frac{E^2}{Z^2} + \frac{\sigma^2}{N}}$$

Donde:

N es el tamaño de la población, y los otros términos son los mismos que se mencionaron anteriormente.

Y para estimar la proporción de una población finita, se tiene

$$n = \frac{P(1-P)}{\frac{E^2}{Z^2} + \frac{P(1-P)}{N}}$$

Cuando la realización de la encuesta se lleva a cabo mediante un muestreo estratificado, la decisión es ahora cuántas unidades habrá que elegir en cada estrato.

Si el muestreo estratificado es proporcional, simplemente el tamaño total de la muestra se distribuye proporcionalmente en cada estrato. El problema es cuando el muestreo estratificado es no proporcional.

En este caso, se puede obtener el error muestral más pequeño posible, aplicando la siguiente fórmula de la asignación óptima del tamaño de la muestra total:

$$n_A = \frac{nN_A\sigma_A}{(N_A\sigma_A + N_B\sigma_B + N_C\sigma_C + \dots)}$$

Donde:

n_A = tamaño de la muestra que se extrae del estrato A

n = tamaño total de la muestra

N_A = número de elementos en el estrato A

σ_A = desviación estándar de elementos en el estrato A

N_B = número de elementos en el estrato B

σ_B = desviación estándar de los elementos en el estrato B

Todo lo anterior se refiere al cálculo del tamaño muestral en diseños monoetápicos, sin embargo el cálculo del tamaño de muestra en un diseño polietápico es un problema para el cual no existe un procedimiento totalmente establecido, pues los supuestos que se establecen en los diferentes métodos, consideran que la selección de las unidades de observación se realiza en una sola etapa y con sólo una característica a estudiar. "Sin embargo en la práctica no es

común que se presenten estas situaciones, ya que en la mayoría de los diseños de muestra al menos se requieren dos etapas para seleccionar a las unidades de observación, y prácticamente todas las encuestas se pueden considerar de propósitos múltiples³⁶

Así, el problema del cálculo del tamaño de muestra se resuelve de diferentes maneras, pero principalmente se considera si se va a trabajar como si se tratara de una variable, ó más de una (propósitos múltiples).

En general, la forma tradicional de proceder cuando se estudian varias variables consiste en estimar el tamaño óptimo para cada variable, utilizando la expresión propuesta por Cochran. Posteriormente se elige el valor mayor, ya que se considera que este valor satisface los requerimientos de todas las variables consideradas. Sin embargo, esta solución no siempre es la más acertada, pues en ocasiones dicha cifra supera en mucho el presupuesto destinado para llevar a cabo la encuesta.

Hay que mencionar que otro punto importante a considerar es la varianza, pues además de minimizar los costos, también se trata de minimizar la varianza de los estimadores en forma simultánea.

Para resolver este tipo de problemas, se han propuesto algunos algoritmos de optimización no lineal que permiten calcular el tamaño de muestra óptimo. Algunos algoritmos de este tipo los han propuesto diversos autores, como Kokan (1963), Kokan y Khan, (1967) y Medina (1998).

Otra de las formas en que se resuelve el problema de seleccionar la muestra en un diseño en varias etapas, es "corrigiendo el tamaño de muestra obtenido

³⁶ Medina, Fernando. Tamaño óptimo de muestra en encuestas de propósitos múltiples. CEPAL

mediante muestreo aleatorio simple por un factor denominado efecto de diseño (efd), que relaciona el coeficiente de correlación intraconglomerados con el tamaño promedio de las unidades de segunda etapa³⁷.

Este factor fue propuesto por Kish (1979), y permite aproximarse al número de selecciones necesarias para que un diseño de conglomerados proporcione la misma varianza. El efecto de diseño se define de la siguiente manera:

$$efd_k = \frac{Var_c(\bar{y})}{(1-f)\frac{S^2}{n}}$$

en donde:

$$Var_c(\bar{y})$$

representa la varianza de la variable de interés en el diseño de conglomerados y

$$(1-f)\frac{S^2}{n}$$

la del esquema de selección aleatoria, de manera que

$$Var_c(\bar{y}) = (1-f)\frac{S^2}{n} \cdot efd_k^{37}$$

³⁷ Fernando Medina H. Los errores de muestreo en las encuestas complejas: Usos y abusos de la información. Trabajo presentado en la reunión de la Asociación Internacional de Estadísticos de Encuestas y la Asociación Internacional de Estadísticas Oficiales, en la ciudad de Aguascalientes, México del 1 al 4 de septiembre de 1998.

Existen otras propuestas para determinar el tamaño de muestra en una investigación. Una de ellas es el "Criterio de optimalidad", el cual parte del hecho de que para realizar una encuesta, se tiene de antemano un determinado presupuesto. Entonces de lo que se trata es de minimizar dicho costo, sujeto a la condición de que la varianza de la j -ésima característica de interés no sea mayor a una cantidad prefijada, para un nivel de confianza determinado.

III.5. Ventajas y desventajas con respecto al muestreo aleatorio simple

Se ha visto que el muestreo polietápico se utiliza sobre todo cuando la población a considerar es muy grande, pues en éste caso, este tipo de muestreo es muy eficiente, pero es necesario mencionar que así como tiene ventajas, también llega a presentar inconvenientes.

En comparación con el muestreo aleatorio simple, sus mayores ventajas son la comunidad y la economía, pues los costos de preparación de la muestra se reducen considerablemente, ya que este tipo de muestreo no requiere listas completas de elementos individuales de la población. Por ejemplo, para efectuar las selecciones de la primera etapa, se necesita una lista de las unidades primarias de muestreo y una medida del tamaño de cada una, a diferencia de otros métodos de muestreo como el aleatorio simple o el sistemático, en los cuales se necesitan listas de elementos individuales de la población.

Otra de las ventajas de este tipo de muestreo es que existe una gran economía en el proceso de entrevistas y en la administración del trabajo de campo.

Finalmente, el muestreo polietápico contribuye a una mejor calidad de las selecciones finales porque tiene una mayor flexibilidad y control en la conformación de listas de elementos de la población.

Por otra parte, es necesario mencionar que así como existen ventajas, también hay algunas desventajas, y la más importante de ellas es que el muestreo polietápico aumenta el error de muestreo que se origina en la selección de casos en grupos y no en forma independiente. Sin embargo esto se puede controlar si se selecciona una cantidad relativamente grande de unidades primarias de muestreo en la primera etapa. De cualquier manera, en este caso, el error estándar es mayor que en el muestreo simple ó estratificado.

CAPITULO CUATRO

EJEMPLO DE MUESTREO POLIETÁPICO

IV.1. Generalidades

La toma de decisiones en mercadotecnia es una actividad frecuente y muy importante, pero implica incertidumbre, tanto en la información sobre la cual se basan dichas decisiones como en los pronósticos a realizar. De aquí se deriva que una buena decisión se basa en una investigación de mercados exitosa.

Es por ello que la investigación de mercados debe llevarse a cabo solo cuando los resultados reduzcan la incertidumbre e influyan sobre las decisiones y por supuesto, para que esto suceda, debe ser oportuna, eficiente y exacta.

Es oportuna cuando se realiza en forma anticipada para ayudar a tomar una decisión. Por ejemplo, si se va a lanzar al mercado un producto la próxima primavera, entonces se deben realizar con bastante anticipación todas las decisiones basadas en la investigación, como son el precio, la formulación del producto, el nombre, la atracción en cuanto a publicidad, etc.

Es eficiente si la calidad de la información que se proporciona ha implicado un mínimo de gastos en cuanto a tiempo y dinero, aun cuando algunas veces se justifica un estudio costoso y complicado cuando la decisión es importante y la información es útil y oportuna.

Es exacta cuando se han llevado a cabo una variedad de métodos y procedimientos estadísticos que ayudarán a garantizar la exactitud en los resultados. Y es precisamente en este punto donde es importante un buen diseño de investigación, que muchas veces incluye estudios que se basan en el muestreo. Por

ejemplo, a las empresas les es de utilidad contar con información del nivel socioeconómico del público al cual están dirigidos sus productos.

Es por ello que en éste capítulo se desarrollará como caso de estudio un diseño basado en el muestreo polietápico para una empresa de telefonía celular. En este caso, se supondrá que la empresa desea investigar qué tanta proporción de población estaría dispuesta a gastar un 5 % de su sueldo mensual en servicios de telefonía celular.

Para ello se considerarán algunas variables de nivel socioeconómico que permitan calcular el tamaño de muestra suficiente para llevar a cabo un estudio de esta magnitud.

Los datos que se consideran para este ejemplo están tomados del XII Censo General de Población y Vivienda 2000 y del Estudio de Regiones Socioeconómicas de México, elaborado por INEGI³⁸.

El diseño propuesto es el siguiente:

VARIABLES

Las variables a considerar son las siguientes:

- a) Porcentaje de población económicamente activa entre 20 y 49 años
- b) Porcentaje de población ocupada que percibe más de dos y medio salarios mínimos
- c) Porcentaje de población ocupada que percibe más de cinco salarios mínimos

³⁸ <http://jweb.inegi.gob.mx/niveles/jsp/index.jsp>.

POBLACIÓN OBJETIVO: Personas entre 20 y 49 años que residan permanentemente en viviendas particulares ubicadas en el D.F.

MARCO MUESTRAL: Lista de población de la base de Regiones Socioeconómicas de México.

FORMACIÓN DE UNIDADES DE MUESTREO:

La formación de las unidades de muestreo se realiza como a continuación se describe:

- Unidad primaria de muestreo (UPM).- Las unidades primarias de muestreo están constituidas por delegaciones
- Unidad secundaria de muestreo (USM).- Las unidades secundarias de muestreo están constituidas por AGEB's

IV.2. DISEÑO MUESTRAL

El ejercicio a realizar se va a considerar probabilístico, polietápico y por conglomerados

- a) Probabilístico: Esto significa que cada miembro de la población tiene una probabilidad conocida y distinta de cero de entrar en la muestra, lo cual permite conocer la precisión de los resultados muestrales.

- b) Polietápico: Porque la unidad última de selección (persona) es seleccionada después de varias etapas.
- c) Por conglomerados: Porque previamente se conforman conjuntos de unidades muestrales de los cuales se obtiene la muestra.

IV.3. TAMAÑO DE LA MUESTRA

Para garantizar que las estimaciones obtenidas tengan una buena calidad, es necesario, entre otros elementos, que el tamaño de muestra que se defina para cada parámetro sea suficiente, y como generalmente se consideran varios indicadores a estudiar, se toma como variable fundamental para el cálculo, el indicador de menor frecuencia en la población objetivo. Esto permite obtener un tamaño de muestra que garantice que el resto de los indicadores queden cubiertos.

En este ejercicio se van a calcular tres tamaños de muestra. una para cada uno de los indicadores considerados. De esta manera se podrá elegir el tamaño que garantice que todos los indicadores queden considerados.

Para calcular el tamaño de muestra se va a utilizar la siguiente notación:

N = el número de conglomerados (delegaciones) en la población

n = el número de conglomerados seleccionados en una muestra irrestricta aleatoria

M_i = el número de elementos (AGEB's) en el conglomerado i

m_i = el número de elementos (AGEB's) seleccionados en una muestra aleatoria del conglomerado i

$M = \sum_{i=1}^N M_i$ = el número de elementos en la población

$\bar{M} = \frac{M}{N}$ = el tamaño de conglomerado promedio para la población

y_{ij} = la j -ésima observación en la muestra del i -ésimo conglomerado

$\bar{y}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} y_{ij}$ = la media muestral para el i -ésimo conglomerado

El procedimiento a seguir será el siguiente:

- 1) En primer lugar se va a calcular σ_w^2 , es decir la varianza entre los elementos dentro de los conglomerados, pero

$$\sigma_w^2 = s_w^2$$

Y

$$s_w^2 = \frac{1}{n} \sum_{i=1}^n s_i^2$$

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

2) Calcular $\sigma_b^2 = s_1^2 - \frac{s_w^2}{m}$, es decir la varianza entre las medias de conglomerados

3) Calcular m, es decir, el número de AGEB's a seleccionar en cada delegación

$$m = \sqrt{\frac{\sigma_w^2}{\sigma_b^2} \left(\frac{c_1}{c_2} \right)}$$

Donde, c_1 y c_2 son costos de selección de delegaciones y AGEB's

4) Finalmente, se va a calcular n, es decir el número de delegaciones a considerar

$$V(\rho) = \frac{\sigma_b^2}{n} + \frac{\sigma_w^2}{nm}$$

Para el caso de c_1 y c_2 , se estimó conveniente que el costo de trasladarse de un conglomerado a otro es 3 veces el trasladarse dentro del conglomerado, considerando las tarifas de transporte público vigentes en el D. F., por lo que el cociente $c_1 / c_2 = 3$.

El tamaño de muestra para cada uno de los indicadores se calculó de la siguiente manera:

- a) Indicador: "Porcentaje de población económicamente activa entre 20 y 49 años"

$$\text{Var (b)} = 0.220402$$

$$\text{Var (m)} = 0.001184$$

$$m = \sqrt{\frac{0.220402}{0.001184}}(3) = 23.63$$

Por lo tanto se deben seleccionar 24 AGEB's de cada delegación seleccionada.

Para encontrar n, se elige una $V(\rho) = 0.05$, de manera que

$$0.005 = \frac{0.001184}{n} + \frac{0.220402}{24n}$$

Si se despeja n, se obtiene $n = 2.07$, es decir, se toman dos delegaciones para la muestra, y en cada delegación se deben tomar 24 AGEB's.

b) Indicador: "Porcentaje de población ocupada que percibe más de dos y medio salarios mínimos

$$\text{Var (b)} = 0.249235$$

$$\text{Var (m)} = 0.031511$$

$$m = \sqrt{\frac{0.249235}{0.031511}}(3) = 4.5 \approx 5$$

En este caso se deben seleccionar 5 AGEB's de cada delegación seleccionada.

Para encontrar n , se elige también una $V(\mu) = 0.005$, de manera que

$$0.005 = \frac{0.031511}{n} + \frac{0.249235}{5n}$$

Entonces $n = 16.27$, por lo que en este caso hay que considerar las 16 delegaciones para la muestra, pero en este caso, se deben elegir 5 AGEBS en cada una.

- d) Indicador "Porcentaje de población ocupada que percibe mas de cinco salarios minimos"

$$\text{Var}(b) = 0.164698$$

$$\text{Var}(m) = 0.024023$$

$$m = \sqrt{\frac{0.164698}{0.024023}}(3) = 4.5 \approx 5$$

En este caso también se deben elegir 5 AGEBS de cada delegación seleccionada.

Ahora se va a calcular n , con $V(\mu) = 0.005$

$$0.005 = \frac{0.024023}{n} + \frac{0.164698}{5n}$$

Entonces $n = 11.39$, por lo que deben considerarse 11 delegaciones para la muestra, y en cada delegación, se van a considerar 5 AGEBS.

De acuerdo a lo anterior, la opción que asegura obtener un tamaño de muestra que garantice que el resto de los indicadores queden cubiertos es la b.

El ejercicio anterior se elaboró considerando $V(\hat{\mu}) = 0.005$, pero ¿qué sucede si este valor aumenta?. Digamos que ahora se considera $V(\hat{\mu}) = 0.009$. Desde luego lo que va a cambiar es el número de delegaciones a considerar ya que el cálculo de m no se altera.

Para el indicador: "Porcentaje de población económicamente activa entre 20 y 49 años", $m = 24$, pero el valor de n , se obtiene de la siguiente manera:

$$0.009 = \frac{0.001184}{n} + \frac{0.220402}{24n}$$

Entonces $n = 1.15$, es decir, se va a elegir una delegación en lugar de 2, y de esa delegación se elegirán 24 AGEB's.

Para el indicador: "Porcentaje de población ocupada que percibe más de dos y medio salarios mínimos" se tiene lo siguiente: $m = 5$, es decir, se van a considerar 5 AGEB's por delegación, pero el número de delegaciones será:

$$0.009 = \frac{0.031511}{n} + \frac{0.249235}{5n}$$

De donde $n = 9.04$, lo que quiere decir que se van a elegir nueve delegaciones en este caso en lugar de las 16 que se calcularon para la varianza de 0.005.

c) Con respecto al indicador "Porcentaje de la población ocupada que percibe mas de cinco salarios mínimos", $m = 5$. El valor de n , será:

$$0.009 = \frac{0.024023}{n} + \frac{0.164698}{5n}$$

De donde $n = 6.33$, lo que significa elegir seis delegaciones en lugar de las 11 que se obtuvieron en el caso anterior.

En este caso también se observa que la opción b es la que asegura que los tres indicadores queden incluidos. La siguiente tabla muestra los resultados obtenidos para cada indicador y para las varianzas consideradas.

	a) Porcentaje de PEA entre 20 y 49 años	b) Porcentaje de población ocupada que percibe más de 2 y medio salarios mínimos.	c) Porcentaje de pobl. ocupada que percibe mas de cinco salarios mínimos.
	$m = 24$	$m = 5$	$m = 5$
Var = 0.005	$n = 2$	$n = 16$	$n = 11$
Var = 0.009	$n = 1$	$n = 9$	$n = 6$

Como se puede observar, el tamaño de muestra en la primera etapa (delegación) se modifica cuando cambia la varianza, y es de esperar que cuando la varianza aumenta, el número de delegaciones a considerar disminuye. Estos resultados se presentan como opciones para que se tome una decisión.

Desde luego que la decisión final la toman los responsables de llevar a cabo el estudio, y para ello un factor importante, y a veces esencial es el considerar los recursos económicos de que se dispone para un proyecto de este tipo. A este respecto cabe mencionar que precisamente se eligió un diseño polietápico debido a que es más económico que otro tipo de muestreo de una sola etapa, como se consideró en el apartado correspondiente en que se habló de las ventajas de este tipo de muestreo con respecto a otros.

Finalmente, y una vez que se ha tomado la decisión pertinente, las delegaciones y las AGEB's a considerar se eligen en forma aleatoria.

CONCLUSIONES

En el presente trabajo se ha analizado la importancia que el muestreo tiene en una actividad tan relevante como lo es la investigación de mercados.

Entre los puntos más sobresalientes se pueden mencionar los siguientes:

- La investigación de mercados es una actividad que en nuestros días ha presentado un gran desarrollo debido a la necesidad de recopilar información para la toma de decisiones, tanto en la comercialización de bienes como en la prestación de servicios. Actualmente en México han surgido diversas empresas dedicadas a esta actividad.
- Existen disciplinas de las cuales se auxilia la investigación de mercados para realizar sus estudios. Una de estas disciplinas es el muestreo, cuyo objetivo es controlar y medir la confiabilidad de la información estadística obtenida a través de la teoría de probabilidades.
- Existen varios tipos de muestreo, como son el muestreo simple aleatorio, el muestreo sistemático, el muestreo estratificado y el muestreo por conglomerados, que forman parte de lo que es el muestreo probabilístico y que se realizan en una sola etapa.

- Dentro del muestreo, existe el llamado muestreo polietápico, que combina dos o más técnicas de las enunciadas anteriormente y que se realiza en dos o tres etapas según lo requiera el diseño de investigación elegido. Este muestreo es importante como herramienta debido a la complejidad de las poblaciones que se analizan, ya que permite "fraccionar" a la población total para llevar a cabo un análisis más exacto de dicha población.

- Uno de los puntos fundamentales a considerar en un diseño polietápico es la elección del tamaño de muestra, que fue precisamente el ejemplo que se desarrolló en el presente trabajo.

- Un diseño muestral de tipo polietápico permite considerar para elegir el tamaño de muestra dos factores que son fundamentales: En primer lugar los costos y en segundo lugar la precisión deseada. El ejemplo que se desarrolló ejemplifica la forma en que se consideran ambos factores para la obtención del tamaño de muestra, con lo que se puede decir que se consideran costos sin sacrificar precisión en los resultados.

BIBLIOGRAFIA

- Berenson, Mark L. y David M. Levine. (1986). Estadística para Administración y Economía, conceptos y aplicaciones. Interamericana. México.
- Campbell, Stephen K. (1981). Defectos y errores del pensamiento estadístico. Limusa, México.
- Cochran, W. G. (1953). Casas Díaz, Eduardo (Traductor). (1971). Técnicas de Muestreo. CECSA. México.
- Chou Ya-Lun. (1977). Análisis Estadístico. 2a. Ed. Interamericana. México
- Kish, Leslie. (1982). Muestreo de encuestas. Trillas. México
- Lininger, Charles A. y Donald P. Warwick. (1984). La encuesta por Muestreo: Teoría y Práctica. CECSA. México
- Subdirección de Estudios Económicos y Planeación Industrial. (1974). Manual de Muestreo Estadístico. Instituto Mexicano del Petróleo. México
- Medina, Fernando. (1998) Tamaño óptimo de muestra en encuestas de propósitos múltiples. CEPAL
- Mendenhall, William y James E. Reinmuth. (1981). Estadística para Administración y Economía. Wadsworth Internacional Iberoamericana. México
- Ríos García, Sixto. (1997). Iniciación Estadística. 9ª. Ed. Paraninfo. Madrid
- Scheaffer, Richard L., Mendenhall, William y Lyman OTT. Elementos de Muestreo. Iberoamericana. México
- Spiegel, Murray R. (1970). Estadística. McGraw Hill. México
- Sukhatme, Pandurang V. (1962). Teoría de encuestas por muestreo con aplicaciones. Fondo de Cultura Económica. México
- Wayne W. Daniel. (1981). Estadística con aplicaciones a las ciencias sociales y a la educación. McGraw Hill. México
- Raj, Des. (1979). Suárez Eduardo L. (Traductor) . La estructura de las encuestas por muestreo. Fondo de Cultura Económica. México

- Blalock Jr. Humbert M. (1983). Naves, Juan (Traductor). Estadística social. 2ª. Ed. Fondo de Cultura Económica. México
- Arturo Núñez del Prado Benavente. (1992). Estadística básica para planificación. 16ª. Ed. Siglo XXI Editores. México
- Fernando Medina H. Los errores de muestreo en las encuestas complejas: usos y abusos de la información. Documento presentado en la reunión de la Asociación Internacional de Estadísticos de Encuestas (IASS) en Aguascalientes, Méx. del 1 al 4 de septiembre de 1998.
- Roberto Hernández Sampieri et al.(1998) Metodología de la Investigación. 2ª. Ed. McGraw Hill. México
- Sharon L. Lohr. (1999). Oscar Alfredo Palmas Velasco (Traductor). Muestreo: Diseño y Análisis. International Thomson Editores.
- Aaker, David A. Investigación de Mercados. McGraw Hill, 1989.
- Lehmann, Donald R. Investigación y análisis de mercado. CECSA, 1993. Traductor Alberto García Mendoza
- David Edwin John, Investigación de Mercados. El Ateneo. México, 1975
- Fischer de la Vega, Laura. Introducción a la Investigación de Mercados. McGraw Hill, 1996.