



Universidad de Carabobo  
Facultad de Ingeniería.  
Escuela de Ingeniería Industrial



## **Técnicas de Muestreo aplicadas a la Ingeniería Industrial**

**Tutor Académico:**

**Prof. Angel Carnevali**

**Autores:**

**Br. Adriana Rodríguez C.I: 17.904.839**

**Br. Manuel Yépez C.I: 17.355.199**

**Bárbula, Junio 2011**



Universidad de Carabobo  
Facultad de Ingeniería.  
Escuela de Ingeniería Industrial



## **Técnicas de Muestreo aplicadas a la Ingeniería Industrial**

Trabajo Especial de Grado presentado ante la Ilustre Universidad de Carabobo, para optar al Título de Ingeniero Industrial

**Tutor Académico:**

**Prof. Angel Carnevali**

**Autores:**

**Br. Adriana Rodríguez C.I: 17.904.839**

**Br. Manuel Yépez C.I: 17.355.199**

**Bárbula, Junio 2011**



Universidad de Carabobo  
Facultad de Ingeniería.  
Escuela de Ingeniería Industrial



### CERTIFICADO DE APROBACIÓN

Quienes suscriben, Miembros del Jurado designado por el Consejo de Escuela de la Facultad de Ingeniería de la Universidad de Carabobo, para examinar el Trabajo Especial de Grado titulado “**Técnicas de Muestreo aplicadas a la Ingeniería Industrial**”, el cual está adscrito a la Línea de Investigación “Probabilidad y Estadística” del Departamento de Investigación de Operaciones, presentado por los Bachilleres **Adriana Rodríguez C.I: 17.904.839** y **Manuel Yépez C.I: 17.355.199**, a los fines de cumplir con el requisito académico exigido para optar al Título de Ingeniero Industrial, dejan constancia de lo siguiente:

1. Leído como fue dicho Trabajo Especial de Grado, por cada uno de los Miembros del Jurado, éste fijó el día martes 7 de junio de 2011, a las 2:00 pm, para que los autores lo defendieran en forma pública, lo que éste hizo, en el Salón de Reuniones, mediante un resumen oral de su contenido, luego de lo cual respondió satisfactoriamente a las preguntas que le fueron formuladas por el Jurado, todo ello conforme a lo dispuesto en el Reglamento del Trabajo Especial de Grado de la Universidad de Carabobo y a las Normas de elaboración de Trabajo Especial de Grado de la Facultad de Ingeniería de la misma Universidad.
2. Finalizada la defensa pública del Trabajo Especial de Grado, el Jurado decidió aprobarlo por considerar que se ajusta a lo dispuesto y exigido por el Reglamento de Estudios de Pregrado.

En fe de lo cual se levanta la presente acta, a día, mes y año, dejándose también constancia de que actuó como Coordinador del Jurado el Tutor, Prof. Ángel Carnevali

Firma del Jurado Examinador

Prof. Ángel Carnevali

Tutor

Prof. Teodoro García

Jurado

Prof. Jorge Piña

Jurado

## **Agradecimientos**

En primer lugar agradezco a Dios por la vida y por permitirme lograr esta meta. A mi compañero Manuel por haber hecho un buen trabajo de equipo. Al profesor Ángel Carnevali, por guiarnos en la realización de este proyecto. A mi familia, por ellos son quien soy hoy y por siempre estar pendiente de mi. A mi jefa Maribel C. por ser tan comprensiva y a mis compañeros de trabajo. A mis amigos Gabi, Adriana, Isabella, Karina, Karem, Leonardo, Jesús, y en especial Juan Carlos.

## **Dedicatoria**

Esto se lo dedico al que todo lo puede, Dios. A mis abuelos Norberto y Glicería, a mi papá Ramón y a mis tíos Elba y Vito porque siempre han estado ahí para mí, los quiero mucho. Mis tíos Orlando, Yolanda, Norma, Annerys y mi madrina Mariela. Mis hermanos Julio César, Anna, Gliseth, José, Miguelangel y Fiorella. También a mis primos y el resto de mi familia a la que no cambio por nada. Gracias por todo el apoyo que me han dado siempre.

Adriana Rodríguez Felice

## **Agradecimiento**

Ante todo le agradezco a Dios, por darme esta oportunidad de cumplir estas metas, por darme la luz y la constancia para llegar al final.

También quiero agradecer a todas las personas que de una u otra manera han dado su aporte y su colaboración para este trabajo se hiciera realidad.

A mi mama, mi papa y mi hermana, por ser siempre un apoyo, no solo en la realización de este trabajo, sino en toda mi carrera.

A María Laura por el apoyo incondicional para realizar este trabajo y así poder cumplir este sueño.

A todos mis amigos en especial al Sr Edgar Navas por todo su apoyo y colaboración incondicional.

Al Profesor Ángel Carnevali por siempre estar dispuesto a ayudarnos y apoyarnos en la elaboración del trabajo.

## **Dedicatoria**

A Dios por ser quien guía los pasos de mi vida, y porque todo lo que tengo y he conseguido se lo debo a Él (El tiempo de Dios es perfecto).

A mi mama, mi papa y mi hermana, por ser mi familia querida y base fundamental de lo que soy hoy en día.

A mis abuelos, los que están aquí en la tierra y los que están en el cielo, porque se que sienten orgullo de mi logro.

A María Laura por ser mi compañera y mi apoyo en todo momento.

Manuel Yépez

## Tabla de Contenido

Resumen .....	7
Introducción .....	8
Capítulo I .....	11
I.1 Planteamiento del problema .....	11
I.2 Objetivos .....	12
I.2.1 Objetivo General .....	12
I.2.2 Objetivos Específicos.....	12
I.3 Alcance .....	13
I.4 Limitaciones .....	13
I.5 Justificación .....	13
Capitulo II .....	15
II.1 Diseño de la Investigación .....	15
Capitulo III .....	17
III.1 Conceptos Básicos.....	17
III.2 Tipos de muestreo.....	30
III.3 Aplicaciones del muestreo .....	56
III.3.1 Aplicaciones en Ingeniería Industrial .....	56
III.3.2 Aplicaciones en otras áreas .....	101
Conclusiones .....	104
Recomendaciones.....	105
Bibliografía.....	106



UNIVERSIDAD DE CARABOBO  
FACULTAD DE INGENIERÍA  
ESCUELA DE INGENIERÍA INDUSTRIAL



## Técnicas de Muestreo aplicadas a la Ingeniería Industrial

Tutor Académico:

Prof. Ángel Carnevali

Autores:

Adriana Rodríguez

Manuel Yépez

### Resumen

Lo que se quiere lograr con esta investigación es demostrar la importancia que tiene la Teoría del Muestreo en la formación profesional del un Ingeniero Industrial. La teoría elemental del muestreo, estudia la relación que existe o puede existir entre una población y las muestras tomadas de la misma población. Es usada para estimar magnitudes desconocidas de una población, tales como la media, la desviación estándar o la varianza, en situaciones donde es imposible físicamente el registro completo de la población. Es útil para hacer comparaciones de poblaciones o de muestras de esas poblaciones, a través de lo que se conoce como contrastes de hipótesis y de significación, muy importante en la teoría de las decisiones. Es de carácter amplio por lo que tiene infinidad de usos y aplicaciones en áreas como física, ciencias sociales, ciencias de la salud, control de calidad, entre otros. En este trabajo se hace mención de algunas de ellas, pero se hace énfasis en las relacionadas con las asignaturas contenidas en el pensum de estudio de la escuela de Ingeniería Industrial de la Universidad de Carabobo.

**Palabras clave:** teoría de muestreo, muestra, población, desviación estándar, varianza, media, pensum.

## Introducción

Los orígenes del uso del muestreo como herramienta de estudio se remontan a finales del siglo XIX cuando estadístico A.N Kaier, de nacionalidad noruega, demostró empíricamente que tomando muestras estratificadas se obtienen mejores resultados para la estimación de medias y totales. Su trabajo fue presentado formalmente en una reunión del Instituto Internacional de Estadística (ISI, por sus siglas en inglés), en Berna, Suiza, bajo el nombre de “Método Representativo” aunque despertó interés fue rechazado. En 1906, Arthur L. Bowley propone la formula para la estimación de la varianza para muestreos estratificados basándose en aproximaciones a la Distribución Normal.

Ya para la década de 1920 el método representativo era usado en Estados Unidos y el mundo, es por eso que en 1924 el ISI decide crear una comisión para la discusión del método. El informe resultante de esa comisión junto con los trabajos de Bowley hicieron que Jerzy Neyman, un matemático, examinara las mismas bases de inferencia de poblaciones finitas lo que dio origen a un trabajo en el cual se desarrolló el muestreo de poblaciones finitas y la estimación por intervalo de confianza, considerado uno de los pilares en donde descansan los fundamentos del muestreo como se le conoce hoy en día, eliminando paradigmas y brindando las ideas de eficiencia, asignación óptima y muestreo por conglomerados. As adelante propuso el muestreo en dos fases.

En el año 1931 fue fundado el instituto India de Estadística por Prasanta Mahalanobis, donde se hacen grandes aportes formulando expresiones de la varianza de estimadores en función del costo de una encuesta. Mahalanobis también introdujo el concepto de encuesta piloto y defendió la utilidad de los métodos de muestreo.

En la década de los 40 se presta mayor atención a los errores del muestreo, William Deming incluye su listado de errores que afectan la calidad de la encuesta los errores debidos al entrevistador. En 1946, Mahalanobis publica uno de los primeros estudios experimentales sobre la contribución del entrevistador a los

errores de la encuesta. A partir de los años 50, se consolida el proceso iniciado con la publicación del artículo de Neyman con la edición de textos considerados a la fecha como clásicos.

En Venezuela, se promulga la Ley de Estadística y Censos Nacionales mostrando interés por parte del gobierno nacional sobre la materia estadística. Dicha ley reforzó la obligación de todos los residentes del país, organismos públicos, de remitir los datos que se les exigiera y colaborar con el Ejecutivo Nacional.

La ley asignó a la Dirección General de Estadística la recolección, elaboración y publicación de los datos estadísticos de la nación. De la misma manera, la ley renovó la obligación de organizar oficinas estadísticas y mandó a levantar un censo de población cada diez años a partir de 1950.

El 21 de diciembre de 1977 se crea la Oficina Central de Estadística e Informática (OCEI) como organismo auxiliar del Presidente de la República y del consejo de ministros. De esta manera, se produjo una integración de los organismos que generaban y procesaban información estadística en el país. La creación de la Oficina de Promoción de Servicios de Información (OPSI) en 1995 pone a disposición del público la información recopilada y procesada por la OCEI, permitiendo la divulgación estadística mediante una plataforma tecnológica moderna.

En 1997 se inició la elaboración de una nueva ley de estadística y censos sociales con el fin de establecer congruencia con los nuevos esquemas políticos, nuevas concepciones de la información estadística y la revolución tecnológica experimentada en el mundo del procesamiento y la difusión estadística. Fue aprobada en 2001 bajo el nombre de Ley de la Función Pública Estadística. Esta nueva ley define y regula la función estadística en general y las relaciones del Estado con el sector privado en el desempeño de esa función, sin dejar a un lado los cambios experimentados por la sociedad desde 1944.

El siguiente trabajo se estructuró en tres capítulos. El capítulo I presenta el planteamiento del problema, los objetivos, alcance, limitaciones y justificación. El capítulo II se refiere al marco metodológico el cual comprende el tipo de investigación y sus fases. El capítulo III contempla el desarrollo teórico del muestreo lo que se desglosa en conceptos básicos, tipos de muestro, aplicaciones del muestreo en Ingeniería Industrial y en otras áreas. Por último se presentan las conclusiones, recomendaciones y un listado de referencia bibliográficas.

## **Capítulo I: El Problema**

### **Capítulo I El Problema**

#### **I.1 Planteamiento del problema**

En la actualidad, existe en el país un mercado laboral muy competitivo, es por ello que es muy importante, que para que un Ingeniero pueda sobresalir, estar preparado para la resolución de problemas con la mayor eficiencia posible. Para lograr esto es necesario que cuente con una serie de conocimientos clave para poder desempeñar su labor. La responsabilidad de la correcta formación recae sobre la universidad. Para ello se diseñan los planes de estudios.

En el pensum de estudio de la Escuela de Ingeniería Industrial están contenidas las materias que garantizan la formación integral de un profesional que es una pieza clave en las organizaciones, debido a que éste se encarga, entre otras cosas, del control de los procesos y la mejora constante de los mismos. Ahora bien, aun cuando el contenido es bastante completo, en general hay temas en específico que no se imparten adecuadamente, ya que no se le da la importancia que éstos tienen, como ejemplo, y a lo cual se hace referencia, se encuentran las Técnicas de Muestreo.

El muestreo es un proceso en el cual se hace un estudio a un conjunto de unidades (pueden ser personas, animales, procesos, entre otros) sobre un tema en específico del cual se desea obtener información, para luego ser analizada y posteriormente tomar decisiones, lo cual es importante para los profesionales de la investigación de mercado, el comercio y las industrias. Aunque parezca muy simple, es algo mucho más complejo, ya que en la mayoría de los casos el estudiar a la población completa requiere de mucho tiempo e implica costos elevados. Es por eso que se toma una muestra de la población, la cual debe ser

## **Capítulo I: El Problema**

significativa, para ello se realizan cálculos que dependen de factores como el nivel de precisión deseado, el tiempo disponible, los costos, entre otros.

Entre las materias comprendidas en el pensum se encuentra Mercados, en la cual se hace uso de encuestas para la elaboración de un proyecto, mediante el cual se evalúa el contenido impartido en la asignatura. Este proyecto consiste en el estudio de mercado y estrategias del lanzamiento de un nuevo producto. Otra de las materias es Control Estadístico de Calidad, cuyo contenido está relacionado con el control estadístico de procesos y la conformidad o no-conformidad de éstos. Para ello también es importante el uso del muestreo. Por último, en la materia Ingeniería de Métodos II, se realizan muestreos para medir tiempos.

Tomando en cuenta lo mencionado anteriormente se puede establecer que el propósito de esta investigación es determinar la importancia del muestreo, y mas específicamente, cuales son las Técnicas de Muestreo que son útiles para la Ingeniería Industrial.

### **I.2 Objetivos**

#### **I.2.1 Objetivo General**

- Demostrar que el muestreo es una herramienta importante para la carrera de un ingeniero industrial

#### **I.2.2 Objetivos Específicos**

- Examinar la bibliografía relacionada con las técnicas de Muestreo.
- Determinar en cuales áreas de la Ingeniería Industrial es importante la aplicación del Muestreo.
- Plantear ejemplos de uso del muestreo en áreas de Ingeniería Industrial de la Universidad de Carabobo
- Relacionar la aplicación teórica y práctica de Técnicas de Muestreo para la formación profesional del Ingeniero Industrial.

## **Capítulo I: El Problema**

### **I.3 Alcance**

El alcance de esta investigación abarca la población estudiantil y docente de la Escuela de Ingeniería Industrial de la Universidad de Carabobo, con el fin de proporcionar las bases teóricas de las Técnicas de Muestreo y así como también establecer la importancia de este tema para el ejercicio profesional.

Es importante resaltar que no se pretende hacer una modificación en el actual pensum, pero si hacer un llamado de conciencia para que se le de a este tópico la importancia que tiene en la carrera.

### **I.4 Limitaciones**

Dentro de las limitaciones de esta investigación se puede decir que es un trabajo sin precedentes ya que no existen trabajos de grados en los que se haga mención el tema del muestreo y la importancia que este tiene para un Ingeniero Industrial.

Así como tampoco existen trabajos de grados de esta índole, los textos para consulta disponible son escasos y, en algunos casos, de difícil acceso.

### **I.5 Justificación**

Este trabajo de investigación busca el desarrollo de un material teórico y práctico, en el cual se mostrará cuándo y cómo un Ingeniero Industrial puede hacer uso de las Técnicas de Muestreo, cabe destacar que este tema no pertenece a ninguna materia de la carrera de ingeniería industrial lo cual esto es una gran desventaja que tienen los profesionales egresados de la Universidad de Carabobo con respecto a la competencia. Lo que se quiere con esta investigación es que sea de ayuda para el personal docente y estudiantil de esta universidad y que de algún modo las autoridades de la escuela evalúen la posibilidad de tomar con más importancia este tema de tal punto de que esta pueda ser incorporada en el pensum de estudio de la Escuela de Ingeniería Industrial de la Universidad de Carabobo.

## **Capítulo I: El Problema**

Otros de los motivos de la realización de esta investigación es que esta nos permitirá obtener el título de ingeniero industrial egresado de la Universidad de Carabobo.

## Capítulo II: Marco Metodológico

### Capítulo II Marco Metodológico

#### II.1 Diseño de la Investigación

##### Metodología

Para llevar a cabo una investigación es primordial contar con un diseño metodológico, ya que brinda una estructura sistemática para el análisis de datos y poder alcanzar así los objetivos planteados.

“La metodología representa la manera de organizar el proceso de la investigación, de controlar los resultados y de presentar posibles soluciones al problema que nos llevará a la toma de decisiones” Zorrilla y Torres (1992)

##### Tipo de Investigación

En este trabajo se estableció, de acuerdo al carácter del estudio, como tipo descriptivo, ya que el fin de esta investigación es demostrar la importancia que tiene el uso del muestreo en la Ingeniería Industrial.

“Los estudios descriptivos buscan especificar las propiedades importantes de personas, grupos, comunidades o cualquier otro fenómeno que sea sometido a análisis. La investigación descriptiva, en comparación con la naturaleza poco estructurada de los estudios exploratorios, requiere considerable conocimiento del área que se investiga para formular las preguntas específicas que busca responder” (Dankhe, 1986).

##### Fases de la Investigación

- **Primera Fase:** Revisar Bibliografía.

Se hace una revisión exhaustiva de toda la bibliografía que se encuentre al alcance relacionada con el tema del muestreo, con el propósito de obtener los

## Capítulo II: Marco Metodológico

conocimientos necesarios para desarrollar este trabajo de grado. Organizar la información encontrada colocando solo la de mayor importancia o utilidad.

- **Segunda Fase:** Identificación de las áreas en las cuales es importante el muestreo.

Una vez hecha la revisión de la bibliografía, se identifica las áreas de la Ingeniería Industrial en las cuales se aplica el muestreo como herramienta de investigación.

- **Tercera Fase:** Desarrollar ejemplos útiles de uso del muestreo en áreas de Ingeniería Industrial de la Universidad de Carabobo.

Ahora bien, se establecen las relaciones del muestreo con las asignaturas contenidas en el pensum de la Escuela de Ingeniería Industrial de la Universidad de Carabobo a fin de resaltar la importancia del muestreo y hacer una ejemplificación.

- **Cuarta Fase:** Demostrar la importancia del muestreo en su formación profesional y laboral para el Ingeniero Industrial.

## **Capítulo III: Desarrollo Teórico Del Muestreo**

### **Capitulo III Desarrollo Teórico Del Muestreo**

#### **III.1 Conceptos Básicos**

##### **Teoría Elemental del Muestreo**

La teoría elemental del muestreo, estudia la relación que existe o puede existir entre una población y las muestras tomadas de la misma población. Esta teoría se considera de gran utilidad en muchos campos modernos de labor profesional. Es usada para estimar magnitudes desconocidas de una población, tales como la media, la desviación estándar o la varianza, en situaciones donde es imposible físicamente el registro completo de la población, o por motivos de la necesidad de estimar magnitudes de poblaciones de carácter infinito. Estas magnitudes desconocidas se conocen con el nombre de parámetros de la población en contraposición de las magnitudes resultantes de las muestras que se denominan estadísticos o estadígrafos.

La teoría del muestreo, es asimismo útil para hacer comparaciones de poblaciones o de muestras de esas poblaciones, a través de lo que se conoce como contrastes de hipótesis y de significación, muy importante en la teoría de las decisiones.

##### **Orígenes del muestreo.**

Algunas fuentes que se consideran como predecesoras del muestreo son:

- Los censos y recuentos de la antigüedad.
- Los procesos inferenciales con base en datos empíricos.
- Los juegos de azar.
- El estudio de los errores de medición.

### **Capítulo III: Desarrollo Teórico Del Muestreo**

El muestreo se ha alimentado de estas cuatro fuentes, algunas han contribuido desde hace muchos siglos y otras hace relativamente poco tiempo, al desarrollo de la teoría de muestreos.

- **Censos y Recuentos de la antigüedad.**

Se pueden mencionar algunos trabajos censales que han sido realizados desde los albores de la humanidad. Desde sus mismo inicios, las sociedades a medida que se fueron organizando, se iban preocupando por la recolección de datos y evidencias, por ejemplo La Biblia cita; Moisés realizó un censo de los israelitas en su travesía del desierto, también reseña un censo realizado por David. En la china Confucio describe como un Rey llamado Yao, unos 3000, años a.C. hizo realizar un recuento agrícola, “Industrial y comercial”, en todo el país.

Se conoce que hubo estudios que llevaron por nombre renta vitalicia, que fueron hechos durante la vigencia del imperio romano, y bajo el mandato del emperador César Augusto, quien decretó que todos los súbditos tenían que pagar tributo, por lo tanto se hizo necesario un censo, para de esta forma, poder saber el número de ellos, así como, su grado de riqueza, este censo lo llevó a cabo el estadístico de la época, el recaudador de impuesto.

También hubo antecedentes de censos o especies de muestreo, en la época de Guillermo el Conquistador, (1027-1087), quien ordenó un censo de tierras y de personas con fines de tributación y de servicio militar, este censo se llamó Domesday Book.

En los Estados Unidos, se tienen trabajos realizados por Carrol Wright, en Massachussets;(1895), en Inglaterra el de Arthur Bowley; de la escuela de economía de Londres (1897), en Rusia; A. Kaufmann. (1897). Todos estos hombres fueron precursores de los estudios y revisiones estadísticas realizados a través de muestreos.

- **Procesos Inferenciales con base en datos Empíricos**

### **Capítulo III: Desarrollo Teórico Del Muestreo**

Ha habido antecedentes de estudios inferenciales en la ciencia de la naturaleza, estudiando muestras procedentes de poblaciones infinitas, al menos tratadas como tales, con modelos más o menos aproximados a la realidad.

Asimismo se puede mencionar la escuela inglesa encabezada por dos de los más grandes estadísticos del siglo pasado, como lo fueron Galton y Pearson, que fueron los precursores, en el estudio de grandes muestras, y que fueron continuados por William Gosset, quien se hizo famoso con el seudónimo de Student, para el caso de muestras menores a 30 datos. Un tiempo mas tarde Ronald Fisher, también hizo grandes contribuciones al estudio de los procesos inferenciales basándose en datos empíricos. Todos estos precursores, dieron origen con sus experimentos tanto biológicos como matemáticos a las teorías modernas de muestreo. La cual pasó del uso de muestras grandes al uso de muestras pequeñas con o sin reposición, o del estudio de poblaciones infinitas al de poblaciones finitas.

- **Los Juegos de Azar.**

Las actividades de juego donde el carácter de aleatoriedad es el mismo hace impredecible el resultado, fue lo que hizo inicialmente que el hombre se interesara por tratar de controlar o predecir los mismos. En el caso de los juegos se considera que cada partida es una muestra de una población de infinitas probabilidades. Es conocidos por todos el intercambio de las famosas cartas entre el Conde de Meré y el matemático Pierre de Fermat, con relación al juego y entre el mismo Conde y otro jugador científico como Blaise Pascal, donde el jugador rechazaba las consideraciones hechas por Pascal y que el conde argüía que no eran ciertas en el juego físico.

- **Estudio de los errores de medición.**

De la misma forma que se considera a una partida como una muestra de una población de infinitas probabilidades, en el caso de una medición, cada una de ellas es un resultado dentro de una gama infinita de las mismas. De esta manera las personas encargadas de medir, observadores, geómetras, y

### **Capítulo III: Desarrollo Teórico Del Muestreo**

agrimensores, se fueron dando cuenta que había una dispersión alrededor de las medidas, cuestión que ellos al principio tomaron como una perturbación molesta, pero aceptada, así ocurrió hasta el mismo en el cual Laplace y Gauss desarrollaron la teoría de errores de observación a principios del siglo XIX, que guarda mucha relación con las llamadas muestras grandes de poblaciones infinitas. Posteriormente las complejidades y obstáculos producto de la selección mismo, sobre todo al trabajar con poblaciones finitas, dieron origen a numerosas publicaciones donde se trató el tema, hasta llegar al primer texto de este tipo que desarrolló Yates, (1949), seguido por otros como Sukhatame(1959) y Murthy, (1962), los cuales han contribuido grandemente al desarrollo de la teoría de muestreos.

#### **Población y Muestra.**

Son dos conceptos fundamentales para la comprensión de este trabajo, el primero de ellos, el de población. Comúnmente es definido como: Aquella concentración de datos formadas por todas las unidades o elementos correspondiente a un tipo homogéneo o agrupado en función de una o varias características previamente definidas. La muestra se define como: conjunto de elementos que constituyendo una pequeña proporción de la población, que es el todo, son seleccionados de la misma con un carácter de homogeneidad, con fines de analizarlos. Es apropiado definir el concepto de universo: es un conjunto de datos, valores o atributos en su totalidad con un carácter uní, o pluridimensional, que puede contener al menos una población, de acuerdo al interés o perspectiva del observador o analista.

#### **Introducción al Muestreo.**

El propósito de un muestreo es realizar una inferencia, e inferir, de hecho significa que se está haciendo una afirmación contando con una información limitada, proveniente de una muestra, acerca de al menos una característica de un cuerpo más grande de datos, que se llama población. Esa colección en su totalidad de los entes observados, que conforman el cuerpo mayor se llama

### Capítulo III: Desarrollo Teórico Del Muestreo

universo, y cada característica de nuestro interés, también, en su totalidad dentro de ese universo la se llama población. Es decir cada variable que conforma al universo es una población, y genera una distribución poblacional. Cada característica poblacional se llama parámetro, y es sobre estas características poblacionales que se trata de inferir cuando se toma una muestra.

Todos los componentes que conforma el universo del cual una muestra es seleccionada se conoce con el nombre de unidades de muestreo. En cambio los datos que generan las cuantificaciones o medidas muestrales dentro de una característica poblacional ya definida, se conoce con el nombre de elementos de muestreo. En otras palabras, las unidades de muestreo al ser seleccionadas conforman una muestra y los elementos de muestreo proveen las cuantificaciones y mediciones.

Una muestra entonces es; cualquier número de componentes seleccionados o extraídos de un cuerpo mayor de datos, que no incluyen a la totalidad de los mismo o mejor dicho, a las unidades de muestreos, y la medida de la característica, se llama estadístico obtenido o estimador. Los estimadores o (medidas de una muestra) estiman a los parámetros (características de una población completa).

Una de las preguntas que se realiza frecuentemente es ¿por qué tomar una muestra? En ese sentido una muestra es buscar información, acerca de las características de poblaciones cuando tiempo y/o costo, son factores de importancia. Muy a menudo en situaciones de la vida real una enumeración completa de todas las unidades de muestreo, en una población es una operación demasiado engorrosa y que consume mucho dinero, esta situación generalmente ocurre en poblaciones demasiado grandes o de carácter infinito, donde el proceso enumerativo total es imposible.

También hay situaciones donde la población no siendo tan grande o de carácter finito, implica su registro o prueba, la destrucción de los elementos de muestreo. Porque la cuantificación o medida de cada uno de ellos los destruye,

### Capítulo III: Desarrollo Teórico Del Muestreo

por ejemplo cuantificar la duración de un bombillo equivale a su destrucción ya que el medir su duración conlleva a la destrucción de la capacidad del mismo de emitir luz. Entonces la pregunta fundamental de principio, debería ser ¿por qué hacer un muestreo? ¿Da suficiente garantía de precisión y eficacia con respecto a lo que se quiere medir? La respuesta es no, ya que la muestra por si misma no es garantía de ello, ya que debe reunir ciertos requisitos para garantizar la precisión y la eficacia.

#### **Precisión y Eficacia.**

Usualmente se mide la precisión de una muestra por medio del error estándar, es decir a menor error estándar más precisión, (menos varianza en la distribución muestral), pero cuando se usa al error de muestreo envuelto de precisión tiene que ver solo con la condición del error de muestreo envuelto en el proceso de muestrear, esto significa que se reconoce que en cualquier proceso de muestreo siempre hay implícita la condición de un error interno. Y este error es consecuencia del mismo proceso de hacer una muestra. (Error de muestreo).

Por esta misma razón es que se espera realizar un muestreo se obtenga una medida, llamada estimador, que presenta una desviación con respecto a la medida poblacional estudiada (parámetro). Sin embargo se debe aclarar que existen otros dos factores que pueden ser causa que un estimador se desvíe del parámetro a estimar, y estos dos factores son subestimados cuando se hace un diseño muestral, lo cual debería evitarse, ya que tienen mucha importancia su consideración en el diseño de la muestra. Entonces estos dos últimos factores junto con el error de muestreo conforman los tres factores básicos de desviación en una muestra:

1. Error de muestreo.
2. Error externo.
3. Sesgo.

Estos tres factores forman la causa de la diferencia entre un estimador y su parámetro.

### Capítulo III: Desarrollo Teórico Del Muestreo

1. **Error de muestreo o error interno.** Es aquel error que se mide a través del error estándar y es causado por el simple hecho de tomar una muestra, este error es proporcional al tamaño de la muestra ya que al crecer el tamaño de la muestra decrece el error y viceversa.
2. **Error externo.** Este error está relacionado con la práctica de tomar una muestra y es causado casi siempre por las personas que conducen el muestreo y/o procesan los datos, y se conocen como errores de registro y errores de proceso, los primeros son cometidos cuando se hacen los registros de los datos y los segundos cuando tiende a crecer a medida que la muestra crece, estos errores se pueden minimizar cuando se reduce el número de unidades de muestreo o cuando se contrata personal idóneo para manejo y cálculo de la data, y se usan medios electrónicos.
3. **Sesgo.** Mientras que los errores internos y externos, tienen cualquier dirección, el sesgo es un factor que causa un desvío en el estimador en una sola dirección, una forma sencilla de visualizar. Generalmente el sesgo es el más problemático de los tres factores que estamos considerando, por una simple razón, es el más difícil de detectar. Siempre se espera el error interno, se puede recalcular los procesos estadísticos, pero muchas veces se puede hacer estimaciones sin reconocer el sesgo. Incluso hay veces que se usa con conocimiento de causa a un estimador sesgado, como en el caso de la desviación estándar ( $S \rightarrow \sigma$ ) que se sabe que es un estimador sesgado, pero consistente de la varianza ( $\sigma_x^2$ ). En este caso obviamos la condición sesgada del estimador. Pero hay casos donde esto no es posible, a saber:
  - Un universo pobremente definido.
  - Inadecuado diseño muestral.
  - Preguntas o encuestas mal redactadas.

### Capítulo III: Desarrollo Teórico Del Muestreo

- Cualquier fenómeno que instigue a una respuesta errónea o distorsionada.

#### Técnicas de Muestreo.

Todas las técnicas de muestreo caen dentro de dos categorías: estadísticas y no estadísticas. El muestreo estadístico incluye todos aquellos muestreos que se realizan al azar o probabilísticamente. El muestreo no-probabilístico, es aquel que se realiza en situaciones en las cuales el muestreo al azar no es posible o no es deseado, y supone que se tiene control directo o indirecto sobre que unidades seleccionar para conformar la muestra. Entre los muestreos no-probabilístico se tiene:

1. Muestreo de Juicio; En este tipo de muestreo el diseñador del mismo, cree tener la convicción que algunos miembros de la población tienen mejor información acerca de los parámetros que otros o son mas fáciles de acceder.
2. Muestreo por cuotas; En este muestreo el tomador de decisiones requiere que la muestra contenga un cierto número de elementos con una cierta característica, establecidas para completar la muestra, como tamaño, localización, estructura, como cuando la muestra es de casas.
3. Muestreo de conveniencia; En este caso el tomador de decisiones selecciona una muestra de cualquier población de tal forma que esta es hecha de fácil y convenientemente, según su criterio.

#### Conceptos Básicos del Muestreo Probabilísticos.

**Distribución poblacional.** Desde el punto de vista de una población como una colección de elementos que presentan un carácter mensurable, el conjunto ordenado de las diferentes medidas de estos elementos, en función de la frecuencia de cada medida constituye una distribución frecuencial, en el sentido que una proporción de elementos tienen un cierto valor. Por lo tanto la distribución

### Capítulo III: Desarrollo Teórico Del Muestreo

poblacional será la expresión mensurable o cognoscible resumida como una distribución de frecuencias, de las diferentes medidas de la característica.

El espacio muestral (conjunto de todos los resultados posibles de un experimento aleatorio) puede considerarse una distribución probabilística  $n$ -dimensional. Constituida por los puntos  $(x_1, x_2, \dots, x_n)$ , representativo de todas las muestras posibles, con las limitaciones naturales de diseño y las correspondientes probabilidades. Se puede considerar al muestreo como un experimento aleatorio (es una actividad cuyos resultados distintos y posibles no se puede predecir). Si para cada muestra se pudiera calcular el valor numérico de una cierta función  $h$ , de  $n$  datos,  $h(x_1, x_2, \dots, x_n)$ , las probabilidades de obtener los distintos valores posibles, se llama una distribución de muestreo del estadístico  $h$  que conforman de hecho, una variable aleatoria.

Generalmente se trata de estimar con  $h$  alguna característica poblacional o valor paramétrico  $H$ , en este caso se dice que  $h$  es un estimador de  $H$  esto significa que cada muestra se tiene un valor numérico parcial o estimación. En este sentido un estimador es una variable aleatoria en el muestreo, cuyos valores numéricos constituyen las estimaciones posibles.

Algunas propiedades de los estimadores:

1. Insesgado; Se dice que un estimador es insesgado cuando su valor esperado es igual al parámetro estimado.
2. Consistente; Se dice que es un estimador consistente si proporciona estimaciones convergentes probabilísticamente hacia el parámetro poblacional.
3. Eficiente; Se dice que un estimador es eficiente dentro de un conjunto de estimadores si posee el menor error cuadrático medio.
4. Suficiente; Un estimador es suficiente, si incluye toda la información que puede proporcionar la muestra, en función de un parámetro.

## Capítulo III: Desarrollo Teórico Del Muestreo

### Tipo de Estimaciones.

En la mayoría de los estudios estadísticos, los parámetros poblacionales son desconocidos, y es necesario estimarlos mediante una o varias muestras, ya que en una gran mayoría de las veces es imposible o muy costoso, o impráctico, analizar y estudiar poblaciones, con mayor precisión posible, es una parte muy importante de la estadística. Las variables aleatorias utilizadas para estimar los parámetros poblacionales reciben el nombre de estimadores, mientras que los valores específicos de esas variables se llaman estimaciones.

Así las variables  $\bar{X}$  y  $S^2$  son estimadores poblacionales de  $\mu$  y  $\sigma^2$ . La estimación de un parámetro no es siempre un único valor, sino que es posible que sea un conjunto de valores. Esto significa que las estimaciones que implican un solo valor se llaman estimaciones puntuales y las que especifican un recorrido de valores se denominan estimaciones por intervalos. Por lo general los muestreos con raíz estadística, con fines de hacer estimaciones, son del tipo por intervalos.

### Intervalos de Confianza.

Usualmente los problemas estadísticos tienen un basamento de construir intervalos de frecuencias probabilísticas para  $\mu$  basándose en  $\bar{X}$ , esto se hace con la finalidad de superar los problemas que acarrea el hacer estimaciones puntuales. La relación que existe entre las estimaciones puntuales y las estimaciones por intervalos se podrá entender mejor a través de la siguiente analogía.

Todos conocen el juego de los dardos, que consiste en lanzar un dardo puntiagudo desde una línea situada a unos cuantos metros de un blanco circular, con anillos concéntricos, donde el éxito consiste en clavar el dardo en el centro de los mismos, como este anillo es el más pequeño, es muy difícil acertar, por un lado por pequeño del centro del blanco y por el otro lado, lo fino de la punta del dardo. Obviamente en determinados momentos que se lanza el dardo, va a ser muy difícil que se acierte con el centro del blanco, lo más probable es que con la práctica se llegue muchas veces cerca y algunas veces demos en el centro del blanco. El dardo

### Capítulo III: Desarrollo Teórico Del Muestreo

es similar a una estimación puntual, el objetivo de la estimación (el dardo) es dar en el blanco (acertar con el parámetro poblacional).

Así como se acepta de buena manera que es difícil dar en el blanco cuando se lanza el dardo, de esa misma forma se tiene que aceptar que es difícil con una estimación puntual dar en el blanco del parámetro poblacional. Se sabe también como se planteó en las nociones elementales de muestreo, que a medida que se aumenta el tamaño de la muestra se reduce el error muestral, (práctica del lanzamiento del dado) similarmente cuando se realiza práctica tras práctica del lanzamiento del dardo, se va teniendo mejor puntería y se puede acercarse más y más a la diana. Posiblemente se llegue a la frustración lanzando el dardo al blanco y nunca acertar, pero si se coloca una tablilla de madera en forma horizontal de extremo a extremo del centro del blanco de tal manera que toque el centro concéntrico y se extienda a los lados por unos centímetros, casi con seguridad que se acertará con mayor facilidad con la tablilla y esta a su vez toda el blanco. Esta situación es similar a la estimación por intervalos, el blanco es, por ejemplo, la media que equivale al centro del blanco, y el intervalo es la tablilla que equivale al rango de valores.

Esta tablilla hablando en términos estadísticos, se puede construir en diferentes tamaños y cada tamaño equivale a un porcentaje de confianza, en estadística se trabaja usualmente con porcentajes de 90, 95, y 99% de confianza. Puede interesar construir un intervalo de un 90%, basado en  $\bar{X}$ , tal que  $\mu$  se encuentre en ese intervalo el 90% de las veces, esto significa que en promedio 90 de cada 100 intervalos calculados a partir de la media muestral, con muestras de tamaño  $n$ , incluirá a la media poblacional.

#### **Probabilidad de error alfa ( $\alpha$ ).**

Tal como se explicó en el segmento anterior, con respecto a la probabilidad que el parámetro esté en el intervalo, es mucho mejor expresar la probabilidad que no esté en él, es decir que en un intervalo de confianza no incluya el parámetro, esta probabilidad se conoce como el error alfa ( $\alpha$ ). Indica la proporción de veces

### Capítulo III: Desarrollo Teórico Del Muestreo

que uno se equivoca, o está en error al suponer que el intervalo incluye al parámetro.

Dado que  $\alpha$  considera como la probabilidad que el intervalo no contenga al parámetro, por lo tanto,  $(1-\alpha)$  será la probabilidad que el parámetro poblacional esté en dicho intervalo. Por supuesto que existe una relación obvia entre el valor de  $\alpha$  y el tamaño del intervalo de confianza: a menor valor de  $\alpha$ , mayor será el intervalo de confianza, por lo tanto, si no se requiere mucha seguridad que el parámetro este en el intervalo diseñado, entonces no será necesario un intervalo muy ancho, por el contrario si se requiere mucha seguridad entonces el intervalo tiene que ser muy ancho.

En general los intervalos se construyen a partir de información muestral, de manera que tantos los cambios en  $\alpha$ , como lo del tamaño muestral ( $n$ ), afectan a la amplitud de intervalo, ya que mientras mayor sea el número de elementos de muestreo considerados, más confianza habrá en la estimación del parámetro poblacional y bastaría un intervalo pequeño para garantizar la inclusión del parámetro dentro del intervalo, normalmente es deseable un intervalo de confianza lo más pequeño posible, además , siempre es necesario en función de determinar un tamaño óptimo del intervalo, considerar el costo del muestreo y el riesgo que estemos dispuestos a asumir con relación a cometer errores.

#### **Precisión de la estimación.**

Como se explicó anteriormente, una forma de incrementar la precisión es reducir el ancho del intervalo, por ejemplo, si se desea buscar más precisión en un intervalo con un  $\alpha=0,05$  se podría el intervalo de 0,95 a 0,80. Pero esa reducción implica que ahora se corre un riesgo mayor de no encontrar al parámetro en el intervalo, ya que  $\alpha$  aumento de 0,05 a 0,20 y esto quiere decir que solo se espera que se encuentre al parámetro en un 80% de las veces, es cierto que se ha enfocado mucho más estrictamente el intervalo haciéndolo más angosto, pero, también se ha decrecido en la confianza. Esto reduce la probabilidad de dar con el parámetro. ¿Hasta qué punto un tomador de decisiones está dispuesto a reducir el

### Capítulo III: Desarrollo Teórico Del Muestreo

nivel de confianza?, esto depende del costo de un intervalo asociado a un intervalo que no contenga al parámetro poblacional. Mientras más alto sea el costo, menor será la disposición del decisor de reducir el intervalo de confianza, dependiendo del costo asociado con el intervalo que no contiene al parámetro. Obviamente mientras mayor sea el costo menor será la disposición para reducir el intervalo.

Si se hace un símil, con la relación a la reducción del intervalo. Suponiendo que se tiene un recipiente 100 metras, de ellas 99 son rojas y una verde, esto es el equivalente a un intervalo de 99% de confianza, si el experimento consiste, en extraer al azar una metra, si resulta que esta es de color rojo gana Bs 1.000, pero si se selecciona una metra de color verde se debe pagar, Bs1.000, es seguro que con toda la confianza mete la mano y extrae una metra con la seguridad que se va a ganar mil bolívares fuertes. Pero qué ocurre si aumentan las metras verdes a 5 y se reducen las rojas a 95, quizás usted todavía se sienta seguro y confiado, que va a ganar mil bolívares fuertes. Y acepte el reto. Pero imagine usted que se reduce el número de metras rojas a 80, y se aumentan las verdes a 20, lo más probable es que usted no esté tan confiado ni seguro porque solo tiene el equivalente a un intervalo de confianza del 80%. Y tendrá que considerar el costo de no sacar la metra roja, que es el equivalente de dar con la media poblacional. De esta manera se entiende que si se tiene suficiente confianza en el intervalo que se está manejando, lo más probable es que no se intente reducirlo.

De esta misma manera puede preguntarse usted. ¿Por qué no se usa siempre un intervalo del 99% para tener una gran seguridad de dar con el parámetro? La razón es que, por supuesto usted gana confianza, pero pierde precisión. Hay alternativas para ganar un poco más de precisión en vez de reducir el intervalo de confianza, y esta alternativa es el aumentar el tamaño de la muestra, ya que se sabe que a mayor tamaño muestral menor error estándar y al tener una muestra mayor automáticamente se reducen los límites del intervalo.

## Capítulo III: Desarrollo Teórico Del Muestreo

### III.2 Tipos de muestreo

#### Muestreo aleatorio simple sin reposición.

Consiste en seleccionar  $n$  elementos entre  $N$  que conforma la población de tal manera que todas las muestras posibles de tamaño  $n$ , tengan la misma probabilidad de ser seleccionada.  $\frac{1}{\binom{N}{n}}$ , bajo este supuesto, es posible calcular la probabilidad que una muestra como elemento poblacional, forme parte de una muestra.  $\frac{\text{Numero de muestras favorables}}{\text{Numero de muestras posibles}}$  Ya que se supone que todos los elementos son equiprobables, por esta razón la posibilidad de una muestra dada, constituida por  $n$  elementos es:  $\left(\frac{1}{N} \cdot \frac{1}{N-1} \cdots \frac{1}{N-n+1}\right) n! = \frac{1}{\binom{N}{n}}$  se observa que el resultado de cada extracción, solo es independiente del resultado de las otras extracciones anteriores.

Se usa el muestreo irrestricto aleatorio, para realizar estimación, para realizar estimaciones acerca de los parámetros poblacionales. Dentro de este esquema el investigador se puede enfrentar a dos problemas;

1. Como seleccionar la muestra irrestricta aleatoria.
2. Como estimar el parámetro de interés.

Para seleccionar una muestra irrestricta aleatoria se puede usar a nuestro criterio para seleccionar la muestra "aleatoriamente", en lo que se denomina muestreo casual. Existe una segunda técnica que se denomina muestreo representativo, consiste en seleccionar una muestra que se considere típica representativa de la población, tanto el primero como el segundo esta sujetos al sesgo del investigador, esto significa que las dos técnicas mencionadas conducen a estimadores cuyas propiedades no pueden ser evaluadas. La manera más práctica de realizar un muestreo irrestricto aleatorio es usando una tabla de números aleatorios.

### Capítulo III: Desarrollo Teórico Del Muestreo

#### Estimación de una Media y un total poblacionales.

Se ha mencionado anteriormente que el objetivo de la muestra es hacer inferencias acerca de una población, partiendo de la información contenida en ésta. La manera más usual de estas inferencias es estimar parámetros de la población. Y generalmente la intención es estimar la media poblacional simbolizada por la letra griega mu ( $\mu$ ) o un total poblacional llamado tau ( $\tau$ ), los cuales se considerarán a continuación:

Se tiene una muestra aleatoria de  $n$  cuentas y se desea estimar el valor medio por cuenta para la población total. Obviamente se emplearía de inmediato el promedio muestral para estimar ( $\mu$ ), a través del cálculo de la media muestral  $\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$ , es de suponer que se entiende que el valor de  $\bar{X}$  revelará muy poco acerca del parámetro  $\mu$ , sea necesario fijar un límite sobre el error de estimación, no es la intención de este trabajo exponer que  $\bar{X}$  es un estimador insesgado de  $\mu$ , pero se sabe que posee muchas propiedades para estimar al parámetro.

- Es estimador insesgado de;  $\mu$ ,  $E(\bar{X}) = \mu$
- Tiene una varianza que decrece a medida que se incrementa el tamaño de la muestra;  $V(\bar{X}) = \frac{\sigma^2}{n} \left( \frac{N-n}{N-1} \right)$

Aunque en este trabajo no se hará demostración alguna con respecto a esto, se hará una ejemplificación que mostrará su validez.

Asumiendo que se tiene una población de  $N = 4$ , dado por  $\{1; 2; 3; 4\}$ . Si una sola observación  $x$ , es tomada al azar de esta población, entonces  $x$  puede tomar cualquiera de los cuatro valores posible, cada uno con una probabilidad de  $\frac{1}{4}$ , por esa razón,

$$\mu = E(x) = 1 \left( \frac{1}{4} \right) + 2 \left( \frac{1}{4} \right) + 3 \left( \frac{1}{4} \right) + 4 \left( \frac{1}{4} \right) = \frac{1}{4} (1 + 2 + 3 + 4) = \frac{10}{4} = 2,50$$

### Capítulo III: Desarrollo Teórico Del Muestreo

$$\sigma^2 = V(x) = E(x - \mu)^2 = \sum (x - \mu)^2 p(y) = (1 - 2,5)^2 \cdot \left(\frac{1}{4}\right) + \dots + (4 - 2,5)^2 = \frac{5}{4}$$

Suponiendo ahora que se selecciona una muestra aleatoria de tamaño  $n = 2$  sin reposición, de esta población de cuatro valores existe la posibilidad de seleccionar combinaciones de cuatro elementos tomados 2 a 2. Es decir aplicando la siguiente fórmula;

$$\frac{n!}{r!(n-r)!} = \frac{4!}{2!(4-2)!} = \frac{4 \cdot 3 \cdot 2!}{2!(2!)} = 6$$

(1,2), (1,3), (1,4), (2,3), (2,4), (3,4) todas estas muestras son igualmente probables, cada una de ellas tendrá por lo tanto una probabilidad de  $\frac{1}{6}$  conociendo la distribución de muestreo de  $\bar{X}$ , sería;

Muestra	$\bar{X}$
(1,2)	1,5
(1,3)	2
(1,4)	2,5
(2,3)	2,5
(2,4)	3
(3,4)	3,5

**Tabla 1. Media**

Entonces, de acuerdo con:

$$E(\bar{X}) = \sum_{i=1}^n \bar{X}_i \cdot p_{\bar{X}_i} = 1,5 \left(\frac{1}{6}\right) + \dots + 3,5 \left(\frac{1}{6}\right) = 2,50 = \mu$$

Para la varianza:

$$V(\bar{X}) = E(\bar{X} - \mu)^2 = \sum_{i=1}^n (\bar{X}_i - \mu)^2 \cdot p_{\bar{X}_i} = \frac{1}{6} \cdot (1,5 - 2,5)^2 + \dots + \frac{1}{6} (3,5 - 2,5)^2 = \frac{5}{12}$$

### Capítulo III: Desarrollo Teórico Del Muestreo

Recordando que:

$$V(\bar{X}) = \frac{\sigma^2}{n} \left( \frac{N-n}{N-1} \right) = \frac{5}{4} \left( \frac{4-2}{3} \right) = \frac{2,5}{6} = 0,42 \quad \text{y} \quad \frac{5}{12} = 0,42$$

Cuando se trabaja con la desviación estándar la fórmula es.

$$\frac{\sigma_x^2}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Tanto en la varianza como en la desviación estándar, la fórmula usada corresponde a la varianza a la desviación estándar del estimador, multiplicada por el factor de corrección por usar una población finita (cuando  $\frac{n}{N} > 0,05$ ).

Límites para el error de estimación =  $2\sqrt{V(\bar{X})} = 2\sqrt{\frac{s^2}{n} \left( \frac{N-n}{N} \right)}$  en el caso de la distribución muestral de las medidas, suponiendo que se toma la muestra 2,4 cuya medida es 3 y la desviación estándar = 1,41, la varianza = 1,99, entonces el límite para el error de estimación sería;  $2\sqrt{\frac{1,99}{2} \left( \frac{4-2}{4} \right)} = 0,71$ , en el caso de la estimación de  $\mu = 3$ , sería  $\bar{X} \pm 0,71$ ;  $3 \pm 0,71$  se tendría que el intervalo iría desde 2,93 hasta 3,71

#### **Selección del tamaño muestral apropiado para estimar medias totales poblacionales.**

En alguna de las etapas del proceso, el investigador debe tomar decisiones acerca del tamaño de la muestra que será seleccionada de la población. Hasta este momento se ha visto el muestreo irrestricto aleatorio, pero no se ha mencionado nada acerca del tamaño de las muestras a tomar. Las observaciones,

### Capítulo III: Desarrollo Teórico Del Muestreo

las entrevistas y los análisis de los datos, cuestan mucho dinero, por lo tanto si la muestra es muy grande, tiempo y dinero es desperdiciado, por el contrario si la muestra es pequeña, se encuentra con una información donde se pagó con dinero y esfuerzo que se desperdició pues la información resultante es inadecuada.

El número de observaciones necesarias para estimar la media poblacional, con un límite para el error de estimación de magnitud B, viene dado por dos desviaciones estándar del estimador,  $\bar{X}$ , es igual a B, es decir,  $2\sqrt{V_{\bar{X}}} = B$  (se tomo 2 por aproximación a 1,96) para n. Recordando que la varianza estimada para  $\bar{X} = V(\bar{X}) = \frac{s^2}{n} \left(\frac{N-n}{N}\right)$  y también,  $V(\bar{X}) = \frac{S^2}{n} \left(\frac{N-n}{N-1}\right)$ , esta última es la conocida varianza de  $\bar{X}$ , es decir, el error estándar de estimación multiplicada por el factor de corrección (c.p.f) que es  $\left(\frac{N-n}{N-1}\right)$ , si se despeja n de la siguiente ecuación, se tiene el tamaño de la muestra requerida.

La solución del despeje de la ecuación se presenta a continuación.

$$2\sqrt{V(\bar{X})} = 2\sqrt{\frac{\sigma^2}{n} \left(\frac{N-n}{N-1}\right)} = B$$

Tamaño de muestra requerida a  $\mu_x$ , con límite para el error estándar = B:

$$n = \frac{N\sigma_x^2}{(N-1)D + \sigma^2}, \text{ donde } D = \frac{B^2}{4}$$

En una situación práctica, la solución para n generalmente presenta un problema por desconocer la varianza ( $\sigma_x^2$ ), pero como la varianza muestral  $S_x^2$  casi siempre se encuentra disponible por experimentos anteriores, o realizando pruebas piloto, se puede tomar a esta ultima como reemplazo de la varianza poblacional  $s_x^2 \approx \sigma_x^2$ .

#### Muestreo aleatorio estratificado

Se conoce que el muestreo irrestricto aleatorio suministra buenas estimaciones de cantidades poblacionales a un costo bajo, así mismo existe otro

### **Capítulo III: Desarrollo Teórico Del Muestreo**

tipo de muestreo aleatorio ya reseñado en los tipos de muestreo que suministra en muchas ocasiones un incremento de la información suministrada, a un costo dado, este tipo de muestreo se conoce con el nombre de Muestreo estratificado.

Un muestreo aleatorio estratificado es aquel que obtiene la muestra mediante la separación de los elementos de la población en grupos no trasladados, llamados estratos, en conjunto con una selección irrestricta aleatoria posterior dentro de cada estrato.

El muestreo estratificado se emplea bajo las siguientes condiciones:

1. Tomando en cuenta que el objetivo al diseñar una encuesta es maximizar la información obtenida, o minimizar el límite del error de estimación. Bajo estas condiciones si existen muestras que tiene poca variabilidad dentro de si, entonces una muestra pequeña, producirá pequeños límites para el error de estimación.
2. Cuando existen áreas compactas, el costo de la muestra puede ser minimizado si se toma una muestra estratificada.
3. Si se tienen estratos identificables, entonces es factible obtener estimaciones de parámetros poblacionales, para subgrupos de la población.

#### **Como seleccionar una muestra aleatoria estratificada.**

El primer paso es definir claramente los estratos, así cada unidad de muestreo se ubica en el estrato apropiado, ese paso no es tan fácil como parece; puesto que esa etapa implica especificar los significados relativos de cada estrato. El segundo paso, consiste en que después que las unidades de muestreo han sido definidas en estratos, se selecciona una muestra irrestricta aleatoria de cada estrato, tal como se mostró en el caso de muestreo anterior.

El tercer paso radica en asegurarse que las muestras en los estratos sean independientes, esto significa que se deben aplicar distintos esquemas de muestreo aleatorio dentro de cada estrato a fin de asegurarse que las muestras

### Capítulo III: Desarrollo Teórico Del Muestreo

que se seleccionan en un estrato no dependan de las seleccionadas en el estrato anterior. El muestreo estratificado requiere una notación adicional:

L: número de estratos

N<sub>i</sub>: número de unidades muestrales en el estrato i

N: número de unidades muestrales en la población.  $N = \sum_{i=1}^L N_i$

Ejemplo ilustrativo:

Una empresa de mercadeo por TV, está interesada en establecer que tanto debe enfatizar los spots comerciales en los diversos sectores del municipio. Esto implica determinar el tiempo promedio que la gente ve televisión en el municipio. El cual está formado por; dos ciudades y un área rural. El pueblo A, esta conformado por personas que en su mayoría, trabajan en una inmensa fabrica de zapatos ubicada dentro de la misma ciudad, y tienen niños que en su mayoría están en edad escolar. El pueblo B es una ciudad satélite conformado por personas que se han jubilado y que en su mayoría no tienen niños. El pueblo A está formado por 155 casa, en el pueblo B, hay 62 y en área rural hay 93.

La media poblacional será estimada según la siguiente fórmula:

$$\bar{X}_{st} = \left( \frac{N_1 X_1 + N_2 X_2 + N_3 X_3}{N} \right) = \text{Varianza estimada de } \bar{X}_{st} = V(\bar{X}_{st}) = \frac{1}{N^2} \sum N_i^2 \left( \frac{N_i - n_i}{N_i} \right) \left( \frac{S_i^2}{n_i} \right)$$

$$\text{el límite para error o estimación sería } 2\sqrt{V(\bar{X}_{st})} = 2\sqrt{\frac{1}{N^2} \sum_{i=1}^L N_i^2 \left[ \frac{N_i - n_i}{N_i} \left( \frac{S_i^2}{n_i} \right) \right]}.$$

Solución, la empresa tiene presupuesto para entrevistar a cuarenta hogares, como hace para determinar una muestra que sea proporcional a cada estrato, lo primero sería aplicar la fórmula de estratos proporcionales;  $N_a = \frac{N_i}{\sum_{i=1}^n (N_i)} n = \frac{155}{155+62+93} 40 = \frac{155.40}{310} = \frac{6200}{310} = 20$  tamaño proporcional al estrato A. Para el pueblo B, se hace la misma operación  $N = \frac{62.40}{310} = 8$ ; para el estrato rural

### Capítulo III: Desarrollo Teórico Del Muestreo

$= N_R = \frac{93,40}{310} = 12$  de esta manera quedarían conformados los estratos por los siguientes elementos de muestreo **A = 20**; **B = 8**; **R = 12**. Ahora se realiza el muestreo irrestricto aleatorio dentro de cada estrato, los resultados fueron:

Estrato I pueblo A;	Estrato II, pueblo	Estrato III Zona Rural
32 28 26 41	27 4 49 10	8 15 21 7
43 29 32 37	15 41 25 30	14 30 20 11
36 25 29 31		12 32 34 24
39 38 40 45		
28 27 35 34		

Cálculos: Fórmulas a usar para todos los estratos:

La media;  $\bar{X}_{ABR} = \frac{\sum x}{n}$  la desviación estándar  $S_x = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}}$ .

Resultados:

$\bar{X}_A = 33,90$	$s_A=5,946$	$s_A^2 = 35,354$
$\bar{X}_B = 25,125$	$s_B=15,245$	$s_B^2 = 232,41$
$\bar{X}_R = 19$	$s_R=9,361$	$s_R^2 = 87,628$

**Tabla 2. Media y Desviación Estándar**

Solución media estimada.

$$\mu_{\bar{X}} = \frac{(155)(33,90) + (62)(25,125) + (93)(19)}{310} = \frac{8579,25}{310} = 27,68$$

### Capítulo III: Desarrollo Teórico Del Muestreo

Para la varianza se usará la siguiente fórmula.

$$V(\bar{X}) = \frac{1}{N^2} \left[ \frac{N_1^2(V_{\bar{X}_A})}{N_A} + \frac{N_2^2(V_{\bar{X}_B})}{N_B} + \frac{N_3^2(V_{\bar{X}_R})}{N_R} \right]$$

Como se trabaja con muestras pequeñas menores a 30, entonces se debe usar el factor de correcciones de población finita, que es igual a;  $\sqrt{\frac{(N-n)}{(N-1)}}$ , entonces la fórmula a usar para la estimación de la varianza total queda así;  $V(\bar{X}_{est}) = \frac{1}{N_2} \sum_1^3 N_i^2 \left[ \left( \frac{N-n}{N-1} \right) \left( \frac{S_i^2}{n_i} \right) \right]$ , el factor de corrección sería;  $\left( \frac{310-40}{310-1} \right) = \frac{270}{309} = 0,874$  volviendo la fórmula de estimación.

$$\begin{aligned} & \frac{1}{310^2} \left[ \frac{155^2(0,874)(35,358)}{20} + \frac{62^2(0,874)(232,411)}{8} + \frac{93^2(0,874)(87,636)}{12} \right] \\ & = \frac{1}{96100} \left[ \frac{742441,98}{20} + \frac{780821,01}{8} + \frac{662460,32}{12} \right] = 1,975 \end{aligned}$$

Por lo tanto el error estándar de estimación será,  $\sqrt{1,975} = 1,405$  luego la estimación de la media poblacional con un límite para el error de estimación de aproximadamente dos desviaciones estándar es,  $2(1,405)=2,81$  por lo cual se puede decir que la media poblacional estará entre: 25, 66 y 28,48 con un 95% de confianza.

#### **Tamaño de la muestra apropiada para estimar medidas y poblaciones totales.**

La cantidad de información que se puede extraer de una muestra, está en función del tamaño de la muestra, ya se explicó anteriormente que el error de muestreo disminuye a medida que aumenta el tamaño de la muestra.

### Capítulo III: Desarrollo Teórico Del Muestreo

Existe un método que permite seleccionar el tamaño apropiado de una muestra, de tal manera que se pueda obtener una cantidad fija de información, para estimar un parámetro poblacional.

Si se quiere que una estimación este dentro L unidades de la media poblacional, una probabilidad aproximada de 0,95, en términos de fórmula se tendrá;  $2\sqrt{V_{\bar{X}_{est}}} = B$  es decir que  $V(\bar{X}_{est}) = \frac{B^2}{4}$ : la misma ecuación, pero en función de estimar el total poblacional  $\tau$  (tau) se tendría;  $2\sqrt{V_{\bar{X}_{est}}} = B$  por lo tanto  $V(\bar{X}_{est}) = \frac{B^2}{4N^2}$ .

**Tamaño de muestra aproximada para estimar  $\mu$  (mu) o  $\tau$  (tau) con un límite B, para el error de estimación.**

$$n = \frac{\sum_1^R N_i^2 \frac{\sigma_i^2}{R_i}}{N^2 D + \sum_{i=1}^R N_i^2 \sigma_i^2} \text{ Donde } R_i \text{ es la observación de fracciones que se le}$$

asignan al estrato i, así para estimar;

$$\text{Para } \mu \text{ seria } D = \frac{B^2}{4}$$

$$\text{Para } \tau \text{ seria } D = \frac{B^2}{4N^2}$$

Para usar la ecuación es necesario obtener aproximaciones de las varianzas poblacionales  $\sigma_1^2; \sigma_2^2; \dots; \sigma_R^2$  con  $S_1^2; S_2^2; \dots; S_R^2$  de algún experimento previo o quizás de una prueba piloto. Existe otro método que requiere conocer la amplitud de variación de las observaciones de cada estrato, este método requiere la utilización del teorema de Tchebycheff y la distribución normal. La amplitud de la variación debe ser entre 4 a 6 variaciones estándar. Este esquema esta fuera del alcance de esta investigación.

Muestras estratificadas, cuando se conoce la desviación estándar.

### Capítulo III: Desarrollo Teórico Del Muestreo

Cuando en el ejemplo anterior se seleccionó una muestra estratificada proporcional a cada estrato asumiendo que cada estrato tiene la misma varianza, o simplemente ignorando la diferencia que puede existir entre cada estrato. Podría ser que en determinados momentos haya interés en no solo determinar los estratos proporcionales sino que al mismo tiempo, considerar la desviación estándar de cada estrato. Esto permite aumentar el tamaño del estrato que tenga mayor variabilidad, de esta manera obtener mayor precisión en cada estrato. La fórmula para calcular los estratos proporcionales en consideración a la variabilidad de cada uno es:

$$N_i = \frac{N_i \sigma_i}{\sum_{i=1}^r (N_i \sigma_i)} \cdot n \text{ Donde,}$$

$N_i$  : tamaño del estrato

Tamaño de la muestra

$N_i$ : tamaño del estrato buscado.

La aplicación de esta fórmula es similar a la aplicación anterior, la diferencia estriba en que esta fórmula se usa la desviación estándar, de cada estrato. El problema que presenta esta fórmula es que muy pocas veces la desviación estándar es conocida o está disponible o es accesible.

#### **Muestreo estratificado usando la variabilidad y el costo.**

Además de la consideración de la variabilidad de cada estrato en función de calcular su proporción, es factible también considerar el costo de cada estrato, pues en la fórmula inicial de los estratos proporcionales así como la segunda donde se consideran las variables de cada estrato, es también posible que se considere el costo de cada elemento de muestreo para cada estrato puesto que es bastante probable que cada estrato tenga costos diferentes.

### Capítulo III: Desarrollo Teórico Del Muestreo

Se debe recordar que el diseño de un muestreo, es proporcionar estimadores con varianzas pequeñas y al menor costo posible, es decir, en términos del objetivo de un muestreo, el mejor diseño o esquema, esta influido por tres factores:

1. El número total de elementos de cada estrato
2. La dispersión de cada estrato
3. El costo de cada observación para cada estrato

En cuanto al número de elementos, es obvio que afecta la información que se puede obtener en cada estrato. Por ejemplo una muestra de 10 elementos de una población de 200, necesariamente debe contener más información que una muestra de 10 elementos de una población de 15000.

En cuanto a la dispersión, debe de considerarse ineludiblemente, porque se necesita una muestra mayor en un estrato menos homogéneo, que en un estrato más homogéneo.

Con relación al costo, si una muestra tiene costos distintos para cada estrato, entonces forzosamente se debe tomar muestras más pequeñas donde los estratos tengan altos costos, ya que la intención en el diseño original es minimizar los costos del muestreo. La asignación aproximada que disminuye el costo para un valor fijo de la varianza de la media estimada o que minimiza la misma para un

costo fijo es; 
$$N_i = \left( \frac{\frac{N_i \sigma_i}{\sqrt{c_i}}}{\sum_{i=1}^r \frac{N_i \sigma_i}{\sqrt{c_i}}} \right) \cdot n$$
 donde  $N_i$  denota el tamaño del i-esimo estrato.  $\sigma_i$

es la desviación estándar del i-esimo estrato y  $c_i$  es el costo para lograr el i-esimo estrato. Generalmente para poder utilizar esta formula es necesario obtener una aproximación de la varianza de cada estrato, a través de mediciones internas de cada uno de los estratos que conforman la muestra, o por aproximaciones logradas a través de trabajos anteriores.

## Capítulo III: Desarrollo Teórico Del Muestreo

### Muestreo sistemático

Tanto el muestreo irrestricto aleatorio, como el estratificado, requieren un trabajo laborioso para la selección de la muestra. Habiendo una forma de identificar las unidades de muestreo dentro de un marco adecuado y diferente al que se ha reseñado hasta ahora, de tal manera que un proceso aleatorio que permita la selección de las unidades específicas para la muestra, se conoce como muestreo sistemático. Ya que este proceso simplifica la selección de las unidades que van a conformar la muestra. La idea básica del muestreo sistemático es utilizar un intervalo fijo a lo largo de una lista, población, o cualquier variable, y elegir los elementos de muestreo a intervalos iguales.

Definición: un muestreo sistemático es aquel que permite seleccionar aleatoriamente un elemento de los primeros  $K$  elementos en el marco, y luego sucesivamente cada  $K$ -ésimo elemento, el resto que conformará la muestra.

El muestreo sistemático proporciona una excelente alternativa para realizar un muestreo irrestricto aleatorio, por las razones que siguen:

- 1.- Es más fácil de realizar, por lo que ha diferencia de los muestreos irrestricto aleatorio y estratificado, esta menos expuesto a los errores de selección, que usualmente cometen los investigadores a la hora de seleccionar elementos en los trabajos de campo.
- 2.- Puede proporcionar mayor información que la que puede proporcionar el muestreo irrestricto aleatorio, por unidad de costo.

### Como seleccionar una muestra sistemática

Ya se ha visto que tanto, el muestreo irrestricto aleatorio, como el estratificado, proporcionan alternativas útiles, usan métodos diferentes. El irrestricto aleatorio usa una tabla de números aleatorios para el proceso de selección de la muestra, el estratificado por el contrario usa una combinación de conformación de estratos y una posterior selección aleatoria post estrato.

### Capítulo III: Desarrollo Teórico Del Muestreo

En el muestreo sistemático, se tienen diversos métodos disponibles. El investigador puede seleccionar diversos métodos accesibles, puede escoger entre una muestra por intervalos de 1 en 3, de 1 en 4 o en general de 1 en K. Por ejemplo, en una población de 15000 personas un investigador en vez de seleccionar una muestra de 1000, elementos a través de una tabla de números aleatorios, trabajo por ciento, bastante engorroso, o puede hacer por medio de un muestreo sistemático con saltos de 10 en 10, y seleccionar al azar un número entre los primeros diez, y de ahí en adelante ir saltando de diez en diez, hasta completar los 1000 elementos.

La selección de los K-elementos, se puede hacer por ejemplo, si el tamaño de la población N es conocido, se puede determinar el tamaño aproximado para la muestra n, a través de la formulación ya conocida y empleada para el muestreo irrestricto aleatorio:

$$n = \frac{N\sigma^2}{(N-1)D + \sigma^2} \text{ Donde } D = \frac{B^2}{4}.$$

Luego de este proceso seleccionar K para obtener el tamaño de la muestra, en la población que se planteó anteriormente hay 15.000 personas y suponiendo que desea hacer una selección de 100 personas ( $n = 100$ ), entonces se debe seleccionar a K como  $\leq 150$ , si se elige  $K = 150$  entonces se obtiene exactamente 100 observaciones, en cambio para  $K < 150$  se obtendría un tamaño muestral mayor a 100. En resumen para una muestra sistemática de n elementos de una población N, K debe ser menor o igual a  $\frac{n}{N}$  es decir  $K \leq \frac{n}{N}$ . Cuando el tamaño de la población no es conocido, no podemos determinar exactamente a K, pero se puede en ese caso establecer un tamaño para n aproximado. Y conjuntamente un valor de K, necesario para lograr un tamaño de muestra n. Si se selecciona un tamaño de K muy grande, el tamaño de muestra K no se lograra usando una muestra sistemática de 1 = en K de la población. Este resultado no tendrá problemas si el investigador puede volver sobre la misma y obtener una nueva

### Capítulo III: Desarrollo Teórico Del Muestreo

muestra de 1 = en K. Esto a veces no es posible, como en el caso de muestras que se les hacen a compradores, no se puede volver sobre la misma población.

Como se explico en el caso de muestreo irrestricto aleatorio, el propósito principal de un muestreo es estimar al menos un parámetro de la población. Esas estimaciones se pueden hacer a través de los estadísticos obtenidos:

Estimador de la media poblacional;  $\hat{\mu}_{sx} = \frac{\sum_{i=1}^n x}{n}$  donde el subíndice "sx" indica que se uso el muestreo sistemático.

$$\text{Varianza estimada de } \bar{X}; \hat{V}_{SX} = \frac{S^2}{n} \left( \frac{N-n}{N} \right)$$

Límite para el error de estimación =  $2\sqrt{\hat{V}} = 2\sqrt{\frac{S^2}{n} \left( \frac{N-n}{N} \right)}$  si N es desconocida se elimina el c.p.f (corrección de población finita)

Se sabe que la varianza estimada  $\bar{X}$  es idéntica a la calculada en el muestreo irrestricto aleatorio, pero esta similitud no implica que las varianzas poblacionales sean iguales, la varianza de  $\bar{X}$  esta dada por  $V(\bar{X}) = \frac{S^2}{n} \left( \frac{N-n}{N} \right)$  en cambio la varianza de  $\bar{X}_S = \frac{\sigma^2}{n} [1 + (n-1)\rho]$  donde  $\rho$  es una medida de la correlación entre los pares de elementos dentro de la misma muestra sistémica. Si  $\rho$  esta cerca de uno, entonces los elementos dentro de la muestra son muy similares con relación a la característica que se esta midiendo, y el muestreo sistemático arrojará una varianza muestral mayor que la obtenida en el muestreo irrestricto aleatorio. En cambio si  $\rho$  es negativo entonces el muestreo sistémico puede ser mejor que el irrestricto aleatorio. (Esto implica que los elementos dentro de la muestra son extremadamente diferentes).

Cuando  $\rho$  cercano a 0, y n es lo bastante grande, entonces el muestreo sistemático es análogo al irrestricto aleatorio.

Para cuales poblaciones ocurre esta relación, se debe considerar primero los tres tipos de poblaciones.

### Capítulo III: Desarrollo Teórico Del Muestreo

1. Población aleatoria
2. Población ordenada
3. Población periódica

1.- Una población es aleatoria si sus elementos siguen un orden azaroso, dentro de una población de este tipo, si se hace un muestreo sistemático, los elementos seleccionados se esperan que sean heterogéneos, con  $\rho \approx 0$ . En este caso el muestreo sistemático es equivalente al irrestricto aleatoria, y si  $n$  es grande, entonces la  $V(\bar{X}_s) = V(\bar{X})$ .

2.- Una población es ordenada si los elementos que conforman a la población están ordenados en relación con una magnitud acordada. Una muestra sistemática extraída de una población con esta característica, es usualmente heterogénea con una. Por lo tanto, una muestra sistemática obtenida de una población ordenada, proporciona mas información que una irrestricta aleatoria por unidad de costo, debido a que la  $V(\bar{X}_{sx}) < V(\bar{X})$  y se puede utilizar una estimación conservadora de  $V(\bar{X}_{sx})$  con  $\hat{V} = \frac{s^2}{n} \left( \frac{N-n}{N} \right)$ .

3.- Una población periódica, si los elementos que conforman la población tienen variación cíclica. Los elementos extraídos at través de una muestra sistemática de una población que presente variación cíclica, es decir, sea una población periódica, pueden ser homogéneos, esto significa que  $\rho > 0$ , esto implica que la varianza estimada de  $\bar{X}_{sx}$ , es mayor que la varianza de  $\bar{X}$  es decir, que  $V(\bar{X}_{sx}) > V(\bar{X})$ .

En este caso el muestreo sistemático proporciona menos información que el irrestricto aleatorio, por la unidad de costo. Como  $V(\bar{X})$ , no puede ser estimada directamente por medio de una sola muestra sistemática, podemos aproximar su valor usando  $V(\bar{X})$ , aunque sabemos que este estimador subestimara la verdadera varianza de  $\bar{X}_{sx}$ .

## Capítulo III: Desarrollo Teórico Del Muestreo

### Muestreo sistemático replicado

Un muestreo sistemático replicado, como su nombre lo indica, requiere de replicas, es decir, repeticiones, esto quiere decir que seleccionan más de una muestra sistémica. Por ejemplo diez muestras sistémicas de 1 en 50, cada una conteniendo 6 mediciones. Que puede ser obtenida aproximadamente en el mismo tiempo en el que se obtiene una muestra de 60 mediciones de 1 en 5. Ambos procesos producen 60 mediciones para estimar la media poblacional  $\mu$ , con la diferencia del muestreo replicado nos permite estimar  $V(\bar{X}_{Sx})$  utilizando el cuadro de las desviaciones de las  $n= 10$  medidas muéstrales individuales, alrededor de su media. Para seleccionar  $n$  muestras sistémicas replicadas, debemos separar más los elementos de cada muestra. Por lo tanto 10 muestras de 1 en 50, de 6 mediciones cada una, contienen el mismo número de mediciones que una sola de 1 en 5. Conteniendo 60 mediciones. La selección se hace eligiendo aleatoriamente 10 puntos de inicio e irle adicionando el elemento  $K$  a cada uno de ellos hasta completar la muestra.

### Muestreo por conglomerado

Una muestra por conglomerado es una muestra aleatoria, en la cual cada unidad de muestreo es una colección, o conglomerado de elementos.

En términos económicos, el muestreo por conglomerado es menos costoso que los muestreos, irrestricto aleatorio y estratificado, si el costo de las entrevistas u observaciones aumenta con la distancia, que separa a los elementos. Para el caso 1, se refiere al muestreo irrestricto aleatorio, donde se requiere de un marco que liste a todos los elementos, y este marco puede ser o muy costoso o imposible de obtener. Para el segundo caso, refiriéndose al muestreo estratificado puede que sea necesario un marco para cada estrato. En los casos el muestreo por conglomerados es más efectivo en cuanto al costo ya que el muestreo por conglomerado por principio requiere que los elementos a entrevistar estén geográficamente cerca el uno del otro.

### **Capítulo III: Desarrollo Teórico Del Muestreo**

En pocas palabras el muestreo por conglomerado es muy útil y económico para obtener información, cuando se dan estas condiciones:

1. No está disponible o es muy costoso un marco que liste o enumere los elementos de la población. En cambio es muy fácil en un marco que liste o enumere los conglomerados
2. El costo de las observaciones es proporcional a la distancia de los elementos.

#### **Como seleccionar una muestra por conglomerados**

El primer paso es especificar los conglomerados apropiados, los elementos de un conglomerado usualmente están muy cerca unos de otros. Por lo que tienden a presentar características similares, en otras palabras puede que los elementos que conforman un conglomerado estén altamente correlacionados, en términos de las mediciones. Esto implica que la información acerca de un parámetro no se incrementa sustancialmente al tomar nuevas mediciones dentro del mismo. Obviamente se podría desperdiciar dinero si se selecciona un conglomerado de gran tamaño.

Sin embargo, esto no sucede, siempre hay ocasiones en las cuales los elementos de un conglomerado son muy diferentes entre si, en estos casos una muestra que contenga pocos conglomerados grandes, puede producir una estimación muy buena sobre todo de la media. El problema de elegir un tamaño apropiado para un conglomerado, es a veces complicado, cuando se dispone de infinitos posibles tamaños de conglomerados. En comparación de los estratos, nótese la diferencia, los estratos deben ser lo mas homogéneos posible dentro de cada estrato pero diferir en cuanto sea posible entre estrato y estrato, con relación a la característica a medir. En cambio los conglomerados deben ser intrínsecamente diferentes como sea posible, y muy similar a otro para poder aprovechar las ventajas económicas del muestreo por conglomerado.

### Capítulo III: Desarrollo Teórico Del Muestreo

Una vez que los conglomerados sea han definido, se debe conformar un marco que liste a todos los conglomerados de la población. Una vez realizado esto se selecciona una muestra irrestricta aleatoria de estos conglomerados, en la forma ya conocida.

#### Estimación de una media y un total poblacional

El muestreo por conglomerados, es un muestreo irrestricto aleatorio con cada unidad de muestreo conteniendo un número de elementos. Por esta razón los estimadores de la media  $\mu$  y de  $\tau$  son semejantes al del muestreo irrestricto aleatorio.

El estimador de la media poblacional  $\mu$  es la media  $\bar{X}$  la cual es dada por,  $\frac{\sum_{x=1}^n x}{n}$  entonces la media  $\bar{X}$  conforma un estimador de razón, con  $m$ , tomando el lugar de  $\bar{X}_i$  por lo tanto la varianza estimada de  $\bar{X}$  toma la forma de la varianza de un estimador de razón.

Estimador de la media poblacional  $\mu$

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

Varianza estimada de  $\bar{X}$

$$V = \left( \frac{N - n}{Nn \bar{M}^2} \right) \frac{\sum_{i=1}^n (y_i - m_i Y)^2}{n - 1}$$

Donde  $\bar{M} = \frac{M}{N}$  representa el tamaño promedio del conglomerado de la población.

N: número de conglomerados en la población.

n : número de conglomerados seleccionados en una muestra irrestricta aleatoria.

$m_i$ : número de elementos en el conglomerado i.

### Capítulo III: Desarrollo Teórico Del Muestreo

$\bar{m} = \frac{1}{n} \sum_{i=1}^n m_i$  : Tamaño promedio del conglomerado en la muestra.

$x_i$  : totalidad de las observaciones en el i-esimo conglomerado.

Límite para error de estimación:

$$2\sqrt{V(\bar{X})} = 2\sqrt{\left(\frac{N-n}{Nn\bar{M}}\right) \frac{\sum_{i=1}^n (x_i - m_i Y)^2}{n-1}}$$

Selección del tamaño de la muestra para la estimación de la media.

$$n = \frac{N\sigma_c^2}{ND + \sigma_c^2} \text{ Donde } \sigma_c^2 \text{ se estima por } S_c^2 \text{ y } D = \frac{B^2\bar{M}}{4}.$$

#### Muestreo por conglomerados en dos etapas

Una muestra por conglomerados en dos etapas es una extensión del concepto de muestreo por conglomerados. Un conglomerado usualmente tiene demasiados elementos para lograr una media de cada uno de ellos o estos son tan semejantes que la medición de unos cuantos provee información sobre ese conglomerado completo. Cuando alguna de las dos situaciones ocurre es factible que se realice una muestra aleatoria de conglomerados y después tomar una muestra aleatoria de elementos dentro de cada conglomerado.

Esto se logra seleccionando primero una muestra aleatoria de conglomerados y posteriormente una muestra aleatoria de los elementos dentro de los conglomerados muestreados.

Este tipo de muestreo se usa habitualmente en las encuestas grandes que tienen que ver con las investigaciones que se realizan con unidades habitacionales. También el muestreo que se realiza en control de calidad. Las ventajas que presenta el muestreo por conglomerado en dos etapas, son las mismas ventajas que se manifestaron en el muestreo por conglomerados.

### Capítulo III: Desarrollo Teórico Del Muestreo

#### Como seleccionar muestras por conglomerados en dos etapas.

El problema básico consiste en la selección de los conglomerados apropiados, bajo dos condiciones básicas;

1. Proximidad geográfica de los elementos que conforman los conglomerados
2. Tamaño apropiado para su manejo de conglomerados.

Notación:  $M = \sum_{i=1}^N M_i$ ; numero de elementos de la población

$N$ = numero de conglomerados en la población

$n$  = numero de conglomerados seleccionados en una población

$M_i$  = número de elementos en el conglomerado  $i$

$m_i$  = número de elementos del conglomerado  $i$  seleccionados

$\bar{M} = \frac{M}{N}$  = el tamaño promedio, de conglomerado por población

$X_{ij}$  = la  $j$ -iesima observación en la muestra del  $i$ -esimo conglomerado.

$$X_i = \frac{1}{m_i} \sum_{j=1}^{m_i} X_{ij}$$

Estimador insesgado de la media poblacional =  $\mu = \left(\frac{N}{M}\right) \frac{\sum_{i=1}^n M_i \bar{X}_i}{n}$

Varianza estimada de  $M = \left(\frac{N-n}{N}\right) \left(\frac{1}{n\bar{M}^2}\right) S_b^2 + \frac{1}{nM\bar{M}^2} \sum_{i=1}^n M^2 \left(\frac{M_i - m_i}{M_i}\right) \left(\frac{S_i^2}{m_i}\right)$

Donde,  $S_b^2 = \frac{\sum_{i=1}^n (M_i \bar{X}_i - \bar{M} \hat{\mu}_X)^2}{n-1}$  y  $S_i^2 = \frac{\sum_{j=1}^{m_i} (X_{ij} - \bar{X}_i)^2}{m_i - 1}$  donde  $i = 1, 2, 3, \dots, n$

Con un límite para el error de estimación igual a  $2\sqrt{\hat{V}(\hat{\mu})}$

El estimador  $\mu$  que se encuentra en la ecuación como estimador insesgado, depende de  $M$ , es decir, el número de elementos de la población.

### Capítulo III: Desarrollo Teórico Del Muestreo

Un método para estimar  $\mu$  cuando se desconoce  $M$  es;  $\frac{N}{Nn} \sum_{i=1}^n M_i \bar{X}_i$

Suponiendo que desea realizar una auditoría en un lote de 1000 facturas, con la condición que al seleccionar una muestra aleatoria de  $n$  unidades, se garantice que se encuentren con un 90% de confianza, facturas defectuosas  $\leq 0,02$ .

Pasos previos a la selección del tamaño de la muestra:

1. Seleccionar la tabla apropiada al tamaño del lote, en este caso  $N=1000$
2. Se busca en la columna tasa de error en la población  $mud = 2\%$
3. Se busca en forma vertical debajo de la tasa de error (2%) el valor mas cercano a  $(100-90)=10\%$ .
4. Como no se encuentra exactamente 10% se busca el, o los mas próximos, encontramos a varios valores cercanos 11,9;7,3;6,7;3,7.
5. A cada uno, le corresponde un tamaño muestral diferente, de ellos el tamaño muestral mas pequeño es el correspondiente a 11,9, que es 100, por lo tanto ese se tomara.
6. Se busca una forma horizontal del  $mud=$  máximas unidades defectuosas, 0 como numero de aceptación y 100 el tamaño muestral.

¿Qué significan los valores encontrados?

Estos fueron; 100; 0, significando que si en una población analizada de 1000 facturas, se toma una muestra aleatoria de 100 y no se encuentra ningún error ( $mud=0$ ) entonces puede tener la seguridad de un  $(100-11,9) = 88,10$  que la tasa de error es menor a 2%. Si hubiese tomado otro plan cercano a 10% por ejemplo, la de 7,3, es decir  $100-7,3 = 92,75$ , el plan de muestreo seria (1000;125;0) esto representa que de una población de 1000, facturas tomando una muestra de tamaño 125, y no encuentra ninguna factura defectuosa hay un 92,75 de confianza que la tasa de errores, es menor a 2%.

### Capítulo III: Desarrollo Teórico Del Muestreo

n	mud	La Población contiene una tasa de error de				
		0,50%	1%	2%	3%	10%
Probabilidad de aceptar el lote es						
25	0	88,1	77,5	60	46,6	6,9
25	1	99,4	97,6	91,3	82,9	26,7
25	2	99,9	99,9	98,8	96,4	53,6
25	3	99,9	99,9	99,9	99,5	77
50	0	77,3	59,7	35,5	21	0,4
50	1	97,8	91,5	7,6	55,1	3,1
50	2	99,9	98,9	92,6	81,4	10,6
50	3	99,9	99,9	98,5	94,2	24,3
.	.	.	.	.	.	.
.	.	.	.	.	.	.
.	.	.	.	.	.	.
<b>100</b>	<b>0</b>	59	34,7	<b>11,9</b>	4	.....
100	1	91,9	73,6	38,9	17,9	.....
100	2	99,2	93,1	67,7	40,8	0,1
100	3	99,9	98,8	86,9	64,8	0,6
100	4	99,9	99,8	95,8	82,7	1,9
125	0	51,2	26,1	7,3	1,7	.....
125	1	88	68,9	26,4	9,3	.....
125	2	98,4	88,2	53,4	25,4	.....
125	3	99,9	97,3	76,7	47,1	0,1
125	4	99,9	99,6	90,7	68,2	0,2

Tasa de error = 2%(mun) 11,9 lo más cerca de 10%

**Tabla 3. Muestreo de aceptación. Para una N = 1000 (Arkin, 1963)**

### Capítulo III: Desarrollo Teórico Del Muestreo

Ejemplos ilustrativos de tamaños muestrales apropiados para  $N = 1000$ , bajo las siguientes condiciones:

Mínimo% de Error inaceptable	% nivel de confianza	Tamaño muestral apropiado
2	80	(150)
3	90	(75)
10	99	( )
2	95	( )

**Tabla 4. Tamaños muestrales para  $N=1000$**

Para el primer paso se busca lo mas cerca posible a  $100-80=20\%$  por debajo del 1% se encontró;

Dos valores próximos: a) 17,3 y b) 20,3

Para el primer ejemplo; el muestreo apropiado es, 1000; 150; 1 es decir, que en una población de 1000, si se toma una muestra de 150 ítems y se encuentra como máximo una (1) unidad defectuosa. Entonces se tendrá un 82,75 de confianza que la máxima tasa de error será de un 2%.

La segunda alternativa, seria el muestreo 1000; 200; 2, es decir, que para una población de 1000 elementos si se toma una muestra  $n = 200$ , si se encuentra un máximo de 2 unidades defectuosas se tiene un 79,705 de confianza que la tasa máxima de error será de un 2%. De las dos alternativas la que tiene mayor nivel de confianza y es más pequeña es la primera.

2do. Ejemplo, tasa de 3%,90% de confianza y  $N= 1000$ ; se busca debajo de 3% el 5 mas cercano a  $1000- 100 = 10\%$  se encuentra 9,3, es decir, si se muestrean 75 unidades y se encuentran 0 unidades defectuosas, habría la seguridad de un 90,70% de que la tasa de error  $\leq 3\%$

### Capítulo III: Desarrollo Teórico Del Muestreo

3er ejemplo; 10% mínima tasa de error inaceptable y 99% de confianza, se busca lo mas cerca posible de  $100-99=1\%$  y debajo de 10%. Se encuentra; 1,2 y 1,3 con 60,1 y 75; 2 obviamente la muestra mas pequeña de las dos es 60,1 por lo tanto es la escogida.

4to ejemplo;  $N=1000$ ; 2% 95% de nivel de confianza. El mas cerca es 1,1 que corresponde a una muestra de tamaño  $n=200$  y si no se encuentra ninguna unidad defectuosa habrá la confianza de un 95% de que no será mayor a un 2% la tasa de error-

En este tipo de muestreo de selección hay un problema que se presenta muy a menudo, y consiste en que siendo muy fácil usando las tablas Arkin determinar el plan de muestreo, bien sea al escoger la mayoría de los lotes con una tasa que exceda al  $X\%$  o que no rehace lotes con una tasa de error menor del  $(X-Y)\%$ . Es muy difícil en auditorias cumplir con ambas metas simultáneamente.

**Ejemplo ilustrativo.** El departamento de cobranzas de una compañía revisa lotes de 1000 contratos de compras a crédito, que se han realizado este año, el auditor externo esta chequeado los que presentan morosidad, y desea asegurarse con un 90% de confianza que los contratos procesados con morosidad, no excede un 3% por lote. Al mismo tiempo desea estar seguro en un 95% que no dejara pasar lotes que contengan contratos morosos con 1% o menos de contratos morosos.

Se busca en la tabla de Arkin, las dos opciones, para  $N=1000$

Minima tasa de error inaceptable 3% (90% de confianza)

Máxima tasa de error inaceptable 1% (95% de confianza)

Para el primer caso debajo de 3% lo mas cerca de 10% corresponden a 9,3% es decir, ese es el valor mas cercano a 10, y equivale a un muestreo de  $n=75$ ; 0, esto expresa, que si de una población de 1000 contratos, se toma una muestra de 75 de ellas y no se encuentra ninguna morosa, entonces estaremos

### Capítulo III: Desarrollo Teórico Del Muestreo

seguros en un 90,70% de que rechazaremos todos los lotes que contengan un 3% o más de contratos morosos.

Para el caso 2. Pero al mismo tiempo rechaza un 45,7% de lotes con solo un 1% de morosidad, o menos. Esto parece excesivo rechazar un 45,7% de lotes que deberían ser aceptables. (Ver debajo 1%, el 45,7% tabla de Arkin)

Que aparece así: acepta lotes con  $\leq 1\%$  rechaza lotes con  $\geq 3\%$

Tamaño	Mud	0,5%	1%	2%	3%	10%
75	0	67,7	45,7	20,7	9,3	-----
75	1	95,2	83,1	55,0	32,7	0,3
75	2	99,6	96,7	81,6	60,6	1,3
75	3	99,9	99,6	94,4	81,8	4,4

**Tabla 5. Fragmento tabla Arkin**

En este tipo de muestreo se presenta a menudo este problema, que implica tomar decisiones con respecto a la aceptación o rechazo de una población, empleando como patrón el que se cumpla o no, alguna norma preestablecida. Llevando implícitos dos riesgos.

1. Rechazar una población siendo aceptable
2. Aceptar una población que es inaceptable

Por esta razón este tipo de muestreo se utiliza más en labores de control de calidad, más que en las auditorías, ya que el auditor no está en posición de aceptar o rechazar una población.

#### **Muestreo de descubrimiento**

Este tipo de muestreo es usado habitualmente, por los auditores que efectúan análisis preliminares, acerca de la calidad o el tipo de algunos datos, por esta razón se conoce también como muestreo exploratorio. Como se recordara en el caso anterior de muestreo de aceptación, el auditor está interesado no solo en

## Capítulo III: Desarrollo Teórico Del Muestreo

la mínima tasa del error inaceptable (METI), sino, en la máxima tasa de error rechazable (MTRI).

Cuando se uso la tabla de Arkin es notorio que la aceptación de un lote tiene una connotación diferente cuando se consideran las dos vertientes, es decir, las dos tasas antes mencionada. En cambio cuando se trata del muestreo de descubrimiento, el interés se centro solo en la tasa mínima de error inaceptable, (METI). Por lo que de acuerdo a lo que se trabajo anteriormente siempre ocurre en este tipo de muestreo que, el número aceptable de unidades defectuosas, o las unidades que poseen el atributo buscado es igual a cero. Esto significa que en realidad el muestreo de descubrimiento es un tipo o una variante del muestreo de atributos.

### III.3 Aplicaciones del muestreo

#### III.3.1 Aplicaciones en Ingeniería Industrial

##### Control estadístico de calidad

La palabra *Calidad* significa la capacidad que tiene un producto o servicio para satisfacer las exigencias y expectativas del cliente, es decir, lo que esperan de éste. También se puede definir como una herramienta para competir en un mundo cambiante.

La gran importancia del control de calidad puede vislumbrarse si se considera que ha pasado históricamente por tres etapas distintas. En una primera etapa, el énfasis se centraba en la labor de inspección y en el establecimiento de tolerancias para los productos. Esta etapa comienza en los años 30 y se extiende hasta comienzos de los 60. El control típico, en esta concepción, es el control de recepción para materiales y el control de auditoría del producto final. Las limitaciones de este enfoque son claras: no evita los defectos de fabricación, sino, únicamente, que se envíen al mercado unidades defectuosas.

### Capítulo III: Desarrollo Teórico Del Muestreo

La segunda etapa del control de calidad se propone evitar las causas de los problemas de calidad durante la fabricación. Se extiende, por accidente, a finales de los años 50, con los estudios de capacidad de procesos y diseño de procesos. El control más importante es el control de calidad en curso de fabricación; las tolerancias comienzan a contemplarse como estándares a superar y no como objetivos a conseguir. Las ventajas de este enfoque radican en su capacidad para mejorar procesos y prevenir la aparición de problemas.

Finalmente, como consecuencia de la intensa competencia internacional, la tercera etapa, desarrollada especialmente en Japón, prosigue la dirección de evitar los problemas antes de que aparezcan, y pone el énfasis en el diseño de productos para que cumplan altas cotas de calidad.

En cualquier proceso de producción, por muy bien que se diseñe y por muy cuidadosamente que se controle, siempre existirá una cierta variabilidad que no se puede evitar. Esta variabilidad natural, también conocida como ruido de fondo, es el efecto acumulado de muchas pequeñas causas de carácter, esencialmente, incontrolable. Cuando el ruido de fondo sea relativamente pequeño se considera aceptable el nivel de funcionamiento del proceso y se dice que la variabilidad natural es originada por un sistema estable de causas de azar. Un proceso sobre el que solo actúan causas de azar se dice que está bajo control estadístico.

Por el contrario, existen otras causas de variabilidad que pueden estar, ocasionalmente, presentes y que actuarán sobre el proceso. Estas causas se derivan, fundamentalmente, de tres fuentes:

- Ajuste inadecuado de las máquinas
- Errores de las personas que manejan las máquinas
- Materia prima defectuosa.

La variabilidad producida por estas causas suele ser grande en comparación con el ruido de fondo y habitualmente sitúa al proceso en un nivel

### Capítulo III: Desarrollo Teórico Del Muestreo

inaceptable de funcionamiento. Estas se denominan causas asignables y se dice que un proceso funcionando bajo causas asignables está fuera de control.

Un objetivo fundamental del Control Estadístico de Procesos es detectar rápidamente la presencia de causas asignables para emprender acciones correctoras que eviten la fabricación de productos defectuosos.

Alcanzar un estado de control estadístico de proceso puede requerir un gran esfuerzo pero es sólo el primer paso. Una vez alcanzado, se puede utilizar la información de dicho control como base para estudiar el efecto de cambios planificados en el proceso de producción con el objetivo de mejorar la calidad del mismo.

Muchas características de calidad pueden expresarse en términos de medida numérica. Una característica cualitativa que sea medible tal como un volumen, un peso, o cualquier dimensión, en general, es una variable.

Cuando se refiere a una variable, es una práctica normal el controlar tanto el valor medio como la dispersión. El control del valor medio se realiza, habitualmente, con el gráfico de control para medias, o gráfico  $\bar{X}$ . El control de la dispersión puede efectuarse bien con el gráfico de control de la desviación típica (gráfico S) o con el gráfico de control de rangos (gráfico R). El uso del gráfico R está más extendido que el del gráfico S.

Establecer una gráfica de control requiere los siguientes pasos:

1. Elegir la característica que debe graficarse.
2. Elegir el tipo de gráfica de control
3. Decidir la línea central que deben usarse y la base para calcular los límites superior e inferior. La línea central puede ser el promedio de los datos históricos o puede ser el promedio deseado, dada por:  $\bar{X} = \frac{X_1+X_2+X_3...+X_n}{n}$

Para el gráfico de R, teniendo en cuenta que  $d_2\sigma = \mu_R$  y  $d_3\sigma = \sigma_{Rse}$

tienen los siguientes parámetros  $LC = d_2\sigma \pm 3d_3\sigma$ .

### Capítulo III: Desarrollo Teórico Del Muestreo

4. Seleccionar el subgrupo racional. Cada punto en una gráfica de control representa un subgrupo que consiste en varias unidades de producto.
5. Proporcionar un sistema de recolección de datos si la gráfica de control ha de servir como una herramienta cotidiana en la planta.
6. Calcular los límites de control, usando la siguiente fórmula  $LC = \bar{x} \pm 3 \frac{\sigma}{\sqrt{n}}$ .  
Los valores  $3 \frac{\sigma}{\sqrt{n}}$  suelen encontrarse tabulados para diferentes valores de  $n$  y se representan por el símbolo A; y proporcionar instrucciones específicas sobre la interpretación de los resultados y las acciones que debe tomar cada persona en producción.
7. Graficar los datos e interpretar los resultados.

Una vez construido el gráfico, se van tomando muestras sucesivas y se representa sobre el mismo cada estimación calculada a partir de los datos muestrales, para ver si queda entre los límites establecidos. Cuando una observación cae fuera se rechaza la hipótesis nula, suponiendo que algo ha ocurrido en el proceso de producción que ha alterado sus condiciones satisfactorias.

Cuando no se dispone de un patrón dado previamente que permita tener un valor de  $\sigma$  se deben tomar  $m$  muestras del mismo tamaño  $n$ , para cada una de las cuales se estima la media, desviación estándar y recorrido. Como media teórica se tomará la media de todas las calculadas anteriormente representada por  $\bar{\bar{x}}$ .

Análogamente suelen calcularse la desviación estándar y recorridos teóricos a partir de las medias de las  $\sigma$  y  $R$  calculadas en cada muestra.

Habiendo mencionado todo esto, también es importante señalar que algunas características de calidad no pueden ser representadas convenientemente por medio de variables cuantitativas por lo que se recurre a los llamados gráficos por atributos. En estos casos, las unidades de producto se clasifican en “conformes” o en “no conformes” según la característica o

### Capítulo III: Desarrollo Teórico Del Muestreo

características cualitativas sean o no conformes con las especificaciones. Las características de calidad de este tipo se denominan atributos. Los datos de tipo atributo tienen solamente dos valores: Conforme / no conforme, pasa / no pasa, funciona / no funciona, presente / ausente. También se consideran atributos aquellas características cuantitativas que se registran en términos.

Se considera una muestra de tamaño  $n$  de unidades obtenidas del proceso de fabricación. Se denota por  $p$  la probabilidad de que una unidad sea defectuosa, siendo  $p$  la fracción de defectuosos que indica la calidad de la producción. Entonces, se considera el estadístico

$$W = \frac{\sum_{i=1}^n X_i}{n}$$

donde  $X_i$  vale uno si la unidad  $i$  es defectuosa y vale cero en caso contrario. Es decir,  $X_i, i = 1, \dots, n$ ; se distribuyen según Bernoulli independientes con parámetro  $p$ . Se tiene entonces

$$EW = p \quad \sigma_w = \sqrt{\frac{p(1-p)}{n}}$$

donde  $\sigma_w$  denota la desviación típica de  $W$ : Por tanto, el gráfico de control para la fracción de unidades defectuosas viene dado por los límites de control:

$$3 \pm p \sqrt{\frac{p(1-p)}{n}}$$

Existen cuatro tipos de gráficos por atributos:

- Gráficos P: Fracción o Porcentaje de unidades defectuosas.
- Gráficos NP: Número de Unidades defectuosas en la producción.
- Gráficos C: Número de defectos de todas las unidades producidas.
- Gráficos U: Número de defectos por unidad producida.

#### Muestreo de trabajo

La productividad es la relación cuantitativa entre lo producido y los recursos utilizados, es decir, no mide la producción sino como he combinado la producción

### Capítulo III: Desarrollo Teórico Del Muestreo

con los recursos. En la actualidad, toda organización realiza estudios en pro de aumentar su productividad, siendo uno de ellos el muestreo de trabajo.

El muestreo de trabajo tiene por objetivo establecer el porcentaje que con respecto al período total de tiempo se dedica a ciertas actividades. Es una técnica en la cual se realiza un gran número de observaciones a un grupo de máquinas, procesos u es operarios durante un periodo de tiempo.

“El muestreo de trabajo es una técnica en lo cual se realiza un gran número de observaciones en un grupo de máquina, procesos u operarios durante un período de tiempo. Cada observación registra lo que está ocurriendo en ese instante, y el porcentaje de observaciones registradas para una actividad particular o demora es una medida del porcentaje de tiempo durante el cual esa actividad o demora ocurren. El porcentaje de tiempo dedicado a una actividad particular se establece a partir de un número de observaciones realizadas al azar.” Burgos (1995)

El Muestreo de Trabajo se basa en las leyes de la probabilidad, las cuales dicen: Una muestra extraída aleatoriamente de un universo dado tiende a tener la misma distribución que dicho universo. Si la muestra es grande, los resultados difieren muy poco de la composición de dicho universo. El muestreo aleatorio requiere que cada una de las partes del universo tenga la misma probabilidad de ser extraída. El concepto de aleatoriedad es fundamental ya que se asocia conceptualmente con el azar.

La probabilidad de x ocurrencias de un evento en n observaciones:

$$(p + q)^n = 1$$

p: probabilidad de una ocurrencia

q: probabilidad de que no haya ocurrencia

### Capítulo III: Desarrollo Teórico Del Muestreo

n: número de observaciones

La distribución de estas probabilidades se conoce como distribución binomial, la estadística elemental dice que a medida que aumenta n la distribución binomial o binómica tiende a la distribución normal. En estudios de muestreo de trabajo se toma una muestra de tamaño n en un intento de estimar p.

#### Fases para su aplicación:

##### 1) Definir el problema

Lo que realmente se desea "medir" es la inactividad y sus causas, por lo tanto se deberá clasificar las inactividades consultando al capataz, operarios, etc. (Las inactividades más frecuentes son: necesidades personales, esperar material y/o su retiro, interferencias varias, rotura de máquina, falta de elementos de trabajo, esperar instrucciones o recibirlas, etc. Cada una de estas inactividades tiene una causa que la origina y es lo que se buscará resolver al concluir el estudio).

##### 2) Determinar la precisión que se desea obtener en los resultados:

Para un nivel de precisión de  $1 \sigma$  (equivalente al 68%) se tiene:

$$S.p = \sqrt{\frac{p(1-p)}{N}}$$

Para un nivel de  $2 \sigma$  (equivalente al 95%):

$$S.p = 2 \sqrt{\frac{p(1-p)}{N}}$$

etc., donde:

S: es la precisión deseada

p: porcentaje de actividad

(1 - p): porcentaje de inactividad

N: n° total de observaciones aleatorias (variable que hay que despejar).

### Capítulo III: Desarrollo Teórico Del Muestreo

**3)** Hacer una estimación preliminar del porcentaje de ocurrencia del evento que se va a medir:

Para ello es necesario establecer un Estudio Piloto, a partir del cual se decidirá respecto a sobre cual actividad, para la precisión deseada, se basará el número total de observaciones.

Proyectar el estudio:

- a) Determinar el número de observaciones a realizar.
  - b) Determinar el número de observados (máquinas u operarios)
  - c) Determinar el número de observadores que se necesitarán (depende de la cantidad de observados y la cantidad de observaciones por día que se quiera realizar)
  - d) Determinar el número de días o turnos que requerirá el estudio.
  - e) Establecer procedimientos detallados para efectuar las observaciones y la ruta a seguir por el observador.
  - f) Proyectar la tabla con los intervalos de observación
  - g) Proyectar la hoja de observación
- 4)** Hacer las observaciones de acuerdo con el plan. Analizar y resumir los datos.
- a) Hacer las observaciones y anotar los datos.
  - b) Resumir los datos al final de cada día.
  - c) Determinar los límites de control.
  - d) Pasar los datos al gráfico de control al final de cada día.
  - e) Preparar el informe y establecer conclusiones. Hacer las recomendaciones correspondientes.

El método de muestreo de trabajo tiene varias ventajas sobre el de obtención de datos por el procedimiento usual de estudios de tiempos. Tales ventajas son:

### **Capítulo III: Desarrollo Teórico Del Muestreo**

- No requiere observación continua por parte de un analista durante un período de tiempo largo.
- El tiempo de trabajo de oficina disminuye
- El total de horas-trabajo a desarrollar por el analista es generalmente mucho menor
- El operario no está expuesto a largos períodos de observaciones cronométricas
- Las operaciones de grupos de operarios pueden ser estudiadas fácilmente por un solo analista.

#### **Mercados**

Las estadísticas de mercados constituyen uno de los elementos básicos del análisis de la demanda. En el análisis e investigación de mercados, en el llamado marketing, en los estudios de motivación, predicción de la conducta de los consumidores, etc. es imposible extender el estudio a poblaciones completas, por razones de tiempo y costo, hay que recurrir a procedimientos de muestreo.

Para tal propósito, los investigadores se basan en muestras de la población de consumidores. El diseño de dicha muestra implica tomar tres decisiones importantes ¿quién será el entrevistado? ¿Cuántas personas? ¿Cómo se elegirá a las personas de la muestra?

El quién corresponde a determinar a quien va dirigido el estudio que proporcionará la información que se necesita. En este caso, se debe conocer la composición de la población por sexo, edades, clases socio-económicas, zona rural y urbana

Seguido de esto es importante establecer el tamaño de la muestra el cual debe ser suficiente para poder obtener información confiable sobre la población

### Capítulo III: Desarrollo Teórico Del Muestreo

total. Por último, se debe disponer como elegir a las personas de la muestra. La siguiente tabla explica los diferentes tipos de muestras:

<b>Tipos de Muestras</b>	
<b>Muestras Probabilísticas</b>	
Muestra de muestra al azar	Cada miembro de la población tiene una probabilidad conocida e igual de ser escogido.
Muestra fortuita estratificada	Se divide la población en grupos excluyentes entre sí y se toma una muestra al azar de cada uno de los grupos.
Muestras de racimo (área)	Se divide la población en grupos excluyentes entre sí y el entrevistador saca una muestra de los grupos para entrevistarlos.
<b>Muestras no probabilísticas</b>	
Muestras de correspondencia cómoda	El investigador selecciona a los miembros de la población que le proporcionarán información con mas facilidad.
Muestras según juicio	El investigador aplica su juicio para seleccionar a los miembros de la población que pueden ser buenos prospectos para brindar información exacta.
Muestreo por cuotas	El investigador encuentra y entrevista a un número determinado de personas, para cada una de varias categorías.

**Tabla 6. Tipos de Muestras**

## **Capítulo III: Desarrollo Teórico Del Muestreo**

Seguido de esto es importante la selección del instrumento a utilizar. Existen dos tipos de instrumentos básicos que son los cuestionarios y los aparatos mecánicos.

Los cuestionarios son más populares, consisten en una serie de preguntas a ser contestadas por el entrevistado. Se deben poner mucha atención al momento de decidir lo que se va a preguntar, cómo, las palabras a usar y el orden de presentación. Es necesario asegurarse que cada una de las preguntas sirve para alcanzar el objetivo del cuestionario.

### **Metodología para la investigación por muestreo**

Una vez identificado el problema y tomada la decisión de investigarlo aplicando las técnicas del muestreo, la ejecución óptima de la misma requiere cubrir las siguientes etapas:

Etapa I.- Concepción general de la encuesta

Etapa II.- Elaboración de instrumentos básicos.

Etapa III.- Administración del cuestionario.

Etapa IV.- Tratamiento y análisis de las respuestas.

Seguidamente se desarrollan cada una de estas etapas:

#### **ETAPA I: Concepción general de la encuesta**

La etapa I comprende ocho fases, las cuales se detallan a continuación:

##### **1.1. Determinación de las necesidades a satisfacer, es decir la fijación de los objetivos a alcanzar.**

### **Capítulo III: Desarrollo Teórico Del Muestreo**

La primera actividad de un estudio estadístico es la definición de su objeto. Para ello hay que hacer un inventario de las necesidades que el estudio debe satisfacer, incluso precisando de antemano la naturaleza de los resultados que se pretende alcanzar. El objeto no se debe extender más de lo necesario, porque se corre el riesgo de entorpecer la encuesta hasta el punto de comprometer sus resultados.

Una encuesta será tanto más complicada cuanto más numerosa sean las dificultades que encuentre por parte de las personas interrogadas. Una encuesta demasiado larga provocara por en ellas reacciones de cansancio, de impaciencia, que se traducirá en un incremento de negativas a responder, en respuestas voluntarias inexactas y no respondidas en su totalidad.

Es necesario detallar claramente que se desea obtener, que persigue la investigación, es decir, establecer los objetivos que permiten determinar el alcance del estudio. Esto objetivos deben plantearse como generales y específicos.

Los objetivos generales indican la idea global o los propósitos del estudio, señalado el problema sin profundizar en la sección del conjunto de variables que se investigan. Establecen el fin que espera obtener o por el cual se realiza el estudio.

Los objetivos específicos son susceptibles de cuantificación; establecen instrucciones o enunciados para la obtención de los logros parciales. Definen todas las variables de estudio, justificación su inclusión, y son base para la preparación de los modelos de tablas en las que se presentarán los resultados y para la elaboración de los instrumentos de recolección de la información. Por ello se plantea una correspondencia entre los objetivos específicos y el nivel de desagregación de la información.

Las variables de estudios que van a hacer cuantificadas en la muestra, deben ser enumeradas claramente y a partir de las mismas hacer las inferencias acerca de los parámetros.

### **Capítulo III: Desarrollo Teórico Del Muestreo**

Con frecuencia se formulan hipótesis. Estas representan un conjunto de proposiciones donde, temporalmente, se recogen posibles soluciones a los problemas planteados. Las hipótesis se plantean de forma tal que facilitan la recolección de la información que finalmente permitirá tomar una decisión sobre su veracidad. Deben presentarse con claridad, como una sentencia declarativa y con frecuencia indican suposición de relación entre las variables.

Es necesario destacar que el estudio de algunas necesidades de información puede determinar que no es necesaria la ejecución de una encuesta, porque existen antecedentes que la contienen, hay que asegurarse de que la información que se desea no exista en otras fuentes. Por tanto hay que investigar materiales relacionados con el tema de estudio, para conocer experiencias previas, disponer de información estadística relacionada, además es necesario evaluar los marcos muestrales y paquetes estadísticos usados en los procedimientos.

#### **1.2. Definición de la población objeto de estudio.**

Una vez fijado en sus detalles el objeto de estudio estadístico, se debe determinar y delimitar la colectividad sobre la cual recaerá el estudio. Hay que circunscribir la colectividad a estudiar, mediante una definición precisa de las unidades estadísticas que la constituyen, así como, llegando al caso, mediante una limitación espacial y temporal, para que las observaciones se realicen realmente sobre la población capaz de dar los resultados, es decir la población objeto de estudio, y que se aplique enteramente a esa colectividad sin omitir nada y sin sobrepasarla.

Es necesario:

- a) Elegir una definición para el conjunto a estudiar o, mejor aun, para las unidades estadísticas que son los elementos constitutivos de este conjunto.
- b) Hacer una delimitación geográfica o espacial.
- c) Hacer una delimitación Temporal.

### **Capítulo III: Desarrollo Teórico Del Muestreo**

La definición de la unidad estadística debe ser precisa y clara y corresponder muy de cerca al término definido con el fin de evitar equivocaciones.

#### **1.3. Información que se desea obtener.**

Se debe verificar todos los datos a recolectar son importantes para los propósitos de la investigación; datos esenciales no deben ser omitidos. Hay cierta tendencia a recopilar información para muchas variables y, luego, no son analizadas totalmente.

Hay que definir con claridad los conceptos y características que serán contados o medidos. Igualmente los promedio o totales a ser determinados.

#### **1.4. Periodicidad de la Información.**

Es necesario establecer la frecuencia con que se va a recopilar la información cuando se trate de investigaciones continuas en el tiempo, excepto cuando se realicen estudios puntuales.

#### **1.5. Nivel de desagregación de la información**

Se denomina cobertura vertical de la encuesta y corresponde a los niveles de detalle en que se requiere la información de acuerdo a los objetivos planteados. La cobertura vertical es prioritaria para la elaboración del diseño muestral que se aplicara en el estudio.

#### **1.6. Marco muestral o base de muestreo.**

Constituye la base fundamental sobre el cual se diseñan los procesos de selección de la muestra. Está conformado por listado, mapa, orden o imagen en que pueden presentarse los elementos de la población o unidades elementales de donde se extrae la muestra. Por lo tanto se considera como uno de los instrumentos más importantes en el trabajo estadístico.

El marco muestral no debe presentar unidades solapadas y ser colectivamente exhaustivo, es decir, contener todos los elementos de la población

### **Capítulo III: Desarrollo Teórico Del Muestreo**

objeto de estudio identificados claramente. Las unidades que conforman el marco muestral se denominan unidades de muestreo.

El muestrista debe considerar la composición interna del marco muestral para elegir el proceso de selección más conveniente.

#### **1.7. Técnicas de recolección de la información.**

Es importante la forma como se recogen los datos. Es necesario definir si se realizan observaciones, entrevistas personales encuestas telefónicas, uso del correo o algún otro medio. Cada forma presenta ventajas y desventajas y debe utilizarse el más adecuado, de acuerdo a los propósitos de la investigación.

#### **1.8. Diseño Muestral.**

Implica determinar el tamaño de la muestra, modo de entrevista, tipo de cuestionario y tomar decisión sobre el tipo de muestra probabilística o no probabilística.

En las muestras no probabilísticas, tanto el tamaño de la muestra como la selección se realizan utilizando el criterio u opinión del investigador. Pueden ser de utilidad cuando la opinión utilizada es de un experto en el tópico investigado o cuando se hace difícil la planificación de una muestra, no hay forma de conocer la exactitud de los estimadores y no puede hacerse inferencia válida de los parámetros. La única forma de evaluar las muestras no probabilísticas es mediante un estudio exhaustivo o censo, porque no proporcionan ningún mecanismo para la evaluación del error que se introduce por investigar únicamente una parte de la población.

En las muestras probabilísticas, también denominadas al azar o aleatorias, cada miembro de la población tiene una probabilidad conocida y diferente de cero de pertenecer a la muestra; la muestra es seleccionada por algún método de selección aleatorio consistente con esas probabilidades; las estimaciones de la

### **Capítulo III: Desarrollo Teórico Del Muestreo**

población a partir de los resultados muestrales, toman en cuenta dichas probabilidades.

El muestreo es parte integral del diseño de una encuesta, es una de sus partes más técnicas. Siempre es posible reducir el tamaño y el costo de una encuesta mediante el uso de diseños muestrales complejos. Hay diferentes clases de diseños que cubren los objetivos de la encuesta. La selección de uno de ellos esta generalmente relacionado con los costos y simplicidad.

El diseño muestral mas convenientes es aquel que se adapte a los recursos disponibles. El muestrista debe considerar cual es el tamaño de una muestra más apropiado de acuerdo a los recursos disponibles y la precisión dada o exigida.

La determinación del tamaño de la muestra y cálculo de estimadores y sus precisiones depende de la clase de muestra a utilizar. Las clases de muestreo probabilísticos más utilizados son: muestreo aleatorio simple, muestreo estratificado aleatorio, muestreo sistemático y muestreo por conglomerados.

#### **ETAPA II: Elaboración de Instrumentos básicos**

Una vez cubierta la etapa I, cuando se tiene la concepción general de la encuesta, se deben elaborar los instrumentos básicos que son el plan de tabulaciones básicas, el cuestionario estadístico y el instructivo del cuestionario.

##### **2.1. Plan de Tabulaciones Básicas.**

Cosiste en la preparación de tablas o cuadros y gráficos estadísticos que recojan los objetivos de la investigación. En la elaboración de los formatos donde representar la información procesada.

### **Capítulo III: Desarrollo Teórico Del Muestreo**

El plan de tabulaciones básicas, señalan Dussaix y Grosbras (1993), Seijas (1981) y Villan (1985), es fundamental para elaborar en forma precisa los instrumentos de recopilación de la información (cuestionario estadístico), porque permite comprobar que la información se recopilara de acuerdo a todos los objetivos planteados y orienta las labores de procesamiento de esa información.

Cuando la investigación requiere entrevistas estructuradas, es indispensable que los instrumentos de recolección de información sean diseñados con posterioridad al plan de tabulación.

Después de elaborado el plan de tabulaciones básico, hay que diseñar los documentos que permiten las observaciones para cada una de las unidades estadísticas interrogadas u observadas. Es la preparación de los instrumentos de recolección de información, porque la observación estadística requiere un documento destinado a recibir el extracto de las características observadas en cada una de las unidades de colectividad estudiada. Este documento se denomina genéricamente cuestionario estadístico, incluso cuando las unidades observadas son elementos inanimados incapaces de reponer a preguntas.

El cuestionario se puede considerar como uno de los más importantes instrumentos del trabajo estadístico. Chevy (1987), Ramos (1984), Seijas (1981) destacan que, en general, cualquier encuesta estadística estará basada en algún cuestionario. El éxito de una encuesta depende en gran parte de la calidad del cuestionario; por esta razón debe insistirse en la importancia de estudiar los aspectos y detalles de este, antes de usarlo en una encuesta.

El cuestionario es un medio para informar a los respondientes acerca del tipo de datos que se necesitan en una encuesta, así como la forma en que deberán ser presentados.

El diseñar el cuestionario se considera, a veces, una tarea difícil que puede realizarse pronta y satisfactoriamente en cualquier momento durante los preparativos de una encuesta, pero requieren un cuidadoso estudio. La

### Capítulo III: Desarrollo Teórico Del Muestreo

consideración de todos los problemas que aparecen en la preparación de un cuestionario es, por tanto, una tarea de mucha responsabilidad.

El cuestionario estadístico debe favorecer la recolección de la información necesaria y facilitar al máximo el tratamiento que se le hará posteriormente. Allí radica la importancia de su implementación y el cuidado que debe ponerse en su elaboración.

Aunque el cuestionario no lo es todo en la encuesta, no es exagerado decir que la calidad de este documento depende, en gran medida, el valor de la encuesta y el de sus resultados, una encuesta estadística puede fracasar por diversas razones, aun si se utiliza un buen cuestionario, pero no se podrá conseguir una buena encuesta con un mal cuestionario.

Dussaix y Grosbras (1993), destacan que en la construcción del cuestionario, se deben seguir tres pasos fundamentales, a saber:

- a. Redacción Preliminar. Antes de hacer una primera redacción del cuestionario, es preciso determinar de antemano cuales son los resultados esperados y hacer el cuestionario en función de estos resultados, o al menos de los que, de entre ellos, parecen accesibles, se debe consultar a los usuarios de los resultados. Hay que aprovechar la ocasión para recoger el máximo de información posible de forma tal que satisfaga al mayor número de necesidades. Pero no se puede sobrepasar el límite de los que es razonable pedir a la buena voluntad del público interrogado. Hay que incluir solo las preguntas esenciales. Se debe, además, atender la necesidad de conocimientos especializados. La preparación del cuestionario exige un conocimiento suficiente de la esfera general abarcada por la encuesta, para determinar de antemano los puntos particulares sobre los que deben plantearse preguntas para recopilar resultados significativos y para que el cuestionario hable de por si el lenguaje del campo a que se refiere. Finalmente, no hay que perder de vista

### Capítulo III: Desarrollo Teórico Del Muestreo

que resulta aventurado querer explorar un campo relativamente complejo con una pregunta única. Es necesario someter a las personas a una batería de preguntas que permitan aclarar, además de su posición fundamental sobre el tema, matrices y reservas que pudieran tener.

La relación preliminar del cuestionario, permite asegurar que la preguntas y las expresiones utilizadas son comprendidas correctamente o que el cuestionario no es excesivamente largo.

- b. Prueba Piloto. No debe utilizarse un cuestionario que no se haya ensayado antes sobre el terreno. Estos ensayos se realizan con una pequeña muestra, en las condiciones más parecidas a las encuestas verdaderas. Siempre se logran observaciones y enseñanzas útiles sobre la encuesta misma o ante algunas preguntas formuladas y sobre el tiempo promedio que haría falta para completar un cuestionario. Este paso es indispensable, ya que permite probar la factibilidad de las técnicas proyectadas, perfeccionar la fraseología y conceptos, asegurar que las preguntas y los cuestionarios no es excesivamente largo, porque un cuestionario muy largo produce fatiga, cansancio, rechazo y pérdida de interés de los encuestados.
- c. Versión Definitiva. Esta se realiza considerando las observaciones tanto de los encuestadores como de los encuestados en la prueba piloto. Se recomienda que las preguntas estén recopiladas para facilitar el tratamiento de la información. Las preguntas que se formulan debe ser perfectamente claras y no presentar ambigüedad, ni ninguna dificultad de interpretación.

Según Chevy (1967) se recomienda no hacer preguntas que exijan largas investigaciones o trabajo de cálculo del respondiente. Deben solicitarse los datos

### **Capítulo III: Desarrollo Teórico Del Muestreo**

en el estado mismo en que los tienen a su disposición. Es preciso desconfiar de las palabras que parecen perfectamente claras cuando se redacta el cuestionario, pero que corren el riesgo de serlo mucho menos para el público en general y dan lugar a diferentes interpretaciones.

Para que las respuestas sean precisas y exactas, se requiere que también lo sea el cuestionario, deben evitarse preguntas que dejen demasiadas iniciativas a los encuestadores, o incluir aquellas que conllevan al menos trabajo posible de investigación y hasta de reflexión. El cuestionario no debe omitir nada esencial y deberá adaptarse tanto a las necesidades como a las características que presentan las unidades estadísticas para las que se hace.

Seijas (1981) recomienda entrevistas de una hora de duración, aunque no hay una idea sistemática de la duración ideal de una entrevista. La duración depende tanto de los objetivos como del interés y motivación de los entrevistados en el estudio. Se observa un mayor número de no respuestas al final de los cuestionarios. La fraseología de las preguntas es crítica, pero no se considera el hecho de que el respondente puede conocer el significado de las palabras usadas.

#### **2.2. Construcción del cuestionario.**

Aunque el cuestionario estadístico es esencial para la calidad final de los resultados de la encuesta, es difícil dar reglas precisas para construir un buen cuestionario; el saber hacerlo y la experiencia juegan el rol determinante. Es necesaria una sensibilización sobre la influencia de la formulación de las preguntas y de la construcción del cuestionario en los resultados obtenidos.

#### **2.3. Formato del Cuestionario.**

En un cuestionario se distinguen esencialmente preguntas abiertas y cerradas. La pregunta abierta no presenta las posibles respuestas para que el

### **Capítulo III: Desarrollo Teórico Del Muestreo**

entrevistado pudiera elegir la que considere correcta o adecuada, es decir, no delimita de antemano las alternativas de respuesta. El entrevistado responde de la forma como quiere y generalmente se registra todo lo que dice. La pregunta cerrada presenta las posibles respuestas más frecuentes y cuando se esperan que estas no sean numerosas.

Las preguntas abiertas poseen como ventajas de no encasillar las diversas contestaciones en respuestas señaladas previamente. Se deja por lo tanto una mayor libertad para que el encuestado conteste de manera más amplia, su desventaja estriba en las dificultades que se enfrentan al momento de cerrar las preguntas, porque existen diversos criterios para ellos.

La pregunta cerrada presenta como ventaja fundamental, la facilidad para trabajar con la información obtenida, ya que no se requiere cerrarla; se desventaja reside en que no permiten proporcionar información amplia y profunda, sobre todo cuando se refiere a temas complejos.

#### **2.4. Redacción de las preguntas.**

Se han establecido tres reglas fundamentales para las preguntas:

Primera regla. La pregunta debe ser bien comprendida por el entrevistado. Hay que evitar listas que no sean exhaustivas. Deben omitirse palabras técnicas, es mejor utilizar el vocabulario empleado en encuestas de estudios cualitativos previos, evitar palabras con significados imprecisos, deben evitarse las dobles negaciones. No se debe utilizar preguntas que se refieren a varios aspectos a la vez, es preferible presentar cada aspecto separadamente.

Segunda Regla. Los entrevistados deben ser capaces de responder las preguntas. Hay que asegurarse que, cuando se presente una lista de alternativas de repuestas, se han previsto todas las situaciones posibles y que la respuesta no

### **Capítulo III: Desarrollo Teórico Del Muestreo**

sea muy difícil para la memoria. En ciertos estudios, los cuestionarios autoadministrados pueden ser mejores que las entrevistas cara a cara, porque el entrevistado puede consultar documentos.

Tercera Regla. La formulación de las preguntas debe permitir la obtención de respuestas sinceras. Deben evitarse las preguntas sesgadas cuya redacción induce las respuestas. Es necesario considerar los deseos de pertenencia del entrevistado a una clase social diferente a la que corresponde, el miedo al cambio y tendencia al conformismo, atracción a las respuestas positivas o tendencia del entrevistado a responder “sí”, “verdad”, “de acuerdo”.

#### **2.5. Orden de las preguntas en el cuestionario.**

El orden de las preguntas en el cuestionario tiene una influencia sobre las respuestas. Dussaix y Grosbras (1993) describen dos efectos producidos por el orden que presentan las preguntas de un cuestionario.

Efecto de Halo: Cuando las preguntas tienen escala orientadas en el mismo sentido, el entrevistador tiende a responder de manera similar.

Efecto de contaminación: la sucesión de preguntas planteadas influye sobre las respuestas de las preguntas sucesivas.

#### **2.6. Longitud del cuestionario.**

La duración aceptable de cuestionario depende de muchos factores similares tales como:

- a. El modo de administración. Un cuestionario telefónico debe ser más corto que uno realizado cara a cara.

### **Capítulo III: Desarrollo Teórico Del Muestreo**

- b. El lugar donde es administrativo. Un cuestionario realizado en la calle no debe tomar más de 10 a 15 minutos y no debe tardarse más de una hora cuando se realiza a domicilio.
- c. El interés del tema por parte de entrevistado.

#### **2.7. Tipos de preguntas.**

- a. Preguntas a las que conviene respuestas numéricas. Estas con las preguntas que el público responde con mayor facilidad y correctamente. Chevy (1967) recomienda que el número dado como respuesta no debe ser el resultado de un cálculo complicado.
- b. Alternativas. Entre las que se encuentran las dicotómicas y las de varias alternativas excluyentes. En las preguntas referentes a hechos o a conocimientos, es preciso prever la ausencia de respuesta o de las respuesta “no se “y tenerlas en cuenta en la elaboración del cuestionario. Este tipo de preguntas son las más adecuadas por su simplicidad.
- c. Preguntas de selección múltiple. Aquellas que presentan más de dos posibles, figurando cada una de las soluciones posibles en el mismo cuestionario y que el respondiente puede señalar varias de estas alternativas como respuesta.
- d. Preguntas abiertas. Cuando el numero de respuestas posibles es tal que su lista no puede figurara en el cuestionario, hay que dejar el espacio para recoger las respuestas que indique el entrevistado con sus propias palabras, este tipo de preguntas es la menos recomendable y se debe evitar, a menos que sea imposible hacerlo

### **Capítulo III: Desarrollo Teórico Del Muestreo**

de otro modo, aunque en encuestas de opinión puede rendir magníficos servicios.

- e. Preguntas trampas. Para descubrir a las encuestados que responde al azar. En general son preguntas de elección múltiple donde una de las alternativas es ilógica o totalmente falsa.

Es oportuno destacar que hay que prestar mucha atención a la elaboración del cuestionario. Nunca se podrá considerar como excesivo el tiempo invertido en su preparación. Es común el comentario según el cual los especialistas en muestreo se esfuerzan mucho sobre los aspectos teóricos, como disminuir algunas decimas el error de muestreo que afecta los resultados de las encuestas, pero descuida puntos importantes de errores atribuibles solamente a la manera de exponer las preguntas, porque no se le ha dado la importancia que merece a la redacción de las preguntas y a su presentación material.

#### **2.8. Presentación del material del cuestionario.**

Es necesario prestar atención a la presentación del cuestionario. Hay que cuidar la calidad del papel, especialmente si quien responde es quien debe llenarlo. Se recomienda una impresión con composición la más grata posible a la vista y un formato no muy grande, que no sea incomodo para el encuestador ni para el entrevistado.

Cuando una encuesta comprende cuestionarios de varios tipos que se refieren a unidades estadísticas o a materiales diferentes, hay que diferenciar bien los distintos documentos para que el público no se imagine que sean ejemplares repetidos de un mismo cuestionario. Se recomienda en estos casos, el uso de papeles de colores diferentes.

Después de elaborado el cuestionario estadístico, hay que prepara un documento explicativo de todas las preguntas que este contiene. La explicación de las preguntas debe, además, incluir ejemplos prácticos. Este documento se llama

### **Capítulo III: Desarrollo Teórico Del Muestreo**

instructivo de cuestionario y generalmente se añade al cuestionario o se presenta a continuación de cada pregunta cuando el cuestionario debe ser llenado directamente por el respondiente.

#### **ETAPA III: Administrador del Cuestionario.**

Corresponde la llamada ejecución de la encuesta, es la observación sobre el terreno de cada una de las unidades estadísticas que conforman la muestra estudiada, comúnmente se denomina trabajo u operaciones de campo.

La administración del cuestionario comprende una serie de actividades para lograr que la información se recopile en el tiempo estipulado y con la calidad requerida, y finaliza con la obtención en la oficina correspondiente de todo el material diligenciado y revisado.

Para realizar una adecuada ejecución de la encuesta, Seijas (1981) señala la necesidad de diseñar controles de la operación y usar formatos de control para el personal, para entrega y recepción del material y para la verificación y chequeo de los datos. Es decir, hay que diseñar organigrama de ejecución, cronograma de actividades, flujograma del material de la encuesta, formato de controles de personal y de entrega y recepción de material de encuestas y hojas de inconsistencias y validación de la información en el terreno, manuales de instrucciones del encuestador y del supervisor.

De la investigación, se familiaricen con el cuestionario, se aclaren dudas y se precisen las actividades de supervisión y control de las labores de estos encuestadores.

En una encuesta por muestreo, los cuestionarios pueden ser administrados por un encuestador cara a cara o por teléfono o autoadministrados, sin la asistencia de un encuestador. Entre éstos se distinguen las encuestas postales y los cuestionarios por video a través de un terminal de computadora. Los modos de recogerlos pueden ser combinados. La calidad de las encuestas cara a cara, depende de la calidad de trabajo del encuestador.

### Capítulo III: Desarrollo Teórico Del Muestreo

La clase y tipo de encuestadores, tanto para las entrevistas personales como para las telefónicas, puede tener un efecto profundo en los datos. Se supone que encuestadores experimentados realizan mejores entrevistas y obtienen un menor número de no respuestas que los menos experimentados. Muchas organizaciones cuentan con encuestadores como empleados fijos o poseen un archivo con la información de los encuestadores para contactarlos oportunamente, especialmente los que han tenido más experiencias según las evaluaciones de los supervisores de las encuestas que han realizado.

Los encuestadores, generalmente, son entrenados para cada encuesta específicamente. Con frecuencia, este entrenamiento o lo realizan los supervisores y otras personas del equipo de investigación y se hace en grupo. El tiempo de entrenamiento varía de acuerdo a la complejidad de las entrevistas. En el entrenamiento deben discutirse los objetivos del estudio y las técnicas generales de entrevistas, simular entrevistas, etc. La evaluación del entrenamiento comprende conocimiento, habilidad y comportamiento de los posibles supervisores y encuestadores.

Los instructivos constituyen una pieza fundamental en el entrenamiento de los encuestadores, por lo que se incluirán en un documento que se llama "Manual del Encuestador". Este manual contiene formas de presentarse al entrevistado, conducción de la entrevista, definición de términos básicos y hojas de inconsistencias. Estas hojas de inconsistencias son las que sirven para detectar, al final de las entrevistas, los errores relevantes que pudieran existir en la misma y así poderlos corregir en esa instancia.

Es necesaria una motivación previa. Las personas que van a ser entrevistadas, son informadas, previamente al levantamiento de la información, que serán entrevistadas y del tipo de información que les será requerida, Zarkovich (1968). Por supuesto que cuando se hace por correo, esta notificación

### **Capítulo III: Desarrollo Teórico Del Muestreo**

previa no es necesaria. Hay que ingeniárselas con el fin de crear por anticipado un clima favorable a la realización de la encuesta, haciendo llegar al público información y argumentos adecuados para combatir los prejuicios, vencer sus resistencias y, si es posible, convencerlos del interés y utilidad de la encuesta. Cada encuesta plantea, entonces, un problema de información e incluso de -propaganda, que debe ser estudiado cuidadosamente.

Una encuesta con aceptación y popularidad entre las unidades de la población que serán investigadas, tendrá mayor éxito que una encuesta que sea rechazada, lo cual implica la imposibilidad de realizarla.

El proceso de las entrevistas debe ser controlado para verificar la calidad y exactitud de las respuestas, además de detectar posibles contradicciones, y disminuir las no respuestas. Hay que tener cuidado sobre los efectos de las entrevistas. Muchos encuestadores no hacen las preguntas ni registran las respuestas de una manera uniforme. Si se reduce el volumen de trabajo de los encuestadores, puede disminuir el efecto de sesgo personal, mientras que los datos basados en el trabajo de sólo unos pocos encuestadores puede estar seriamente afectado por tales sesgos.

Es necesario establecer controles para asegurar la calidad de los datos que se recogen. Es una característica importante evaluar la recolección de datos en cada entrevista, especialmente en entrevistas largas, y los datos con frecuencia deban ser reorganizados. Debe completarse un reporte escrito, apenas finalice cada entrevista, sobre cada una de las fases de ésta. De esta manera, el encuestador puede destacar información que no pudo obtener, señalar discrepancias aparentes y organizar la información de manera sistemática. Los reportes para diferentes entrevistas sobre la misma encuesta determinarían inconsistencias entre las entrevistas. Se establecen procedimientos de insistencias, reentrevistas y seguimiento. Algunas veces, la insistencia de

### **Capítulo III: Desarrollo Teórico Del Muestreo**

entrevista puede ser de solo dos, y otras, hasta de cuatro e incluso más, según los objetivos del estudio.

La preparación de los reportes escritos, puede llevar a la planificación de llamadas telefónicas sobre puntos específicos, cuando todavía la encuesta está fresca en la mente de los entrevistados. Esto dará la oportunidad a otras personas trabajando en el proyecto hacer preguntas que no fueron respondidas en la entrevista.

Algunas veces se le dan incentivos a los entrevistados para contar con su colaboración en las respuestas, pero esto debe verse con mucho cuidado, porque a la larga puede ser más perjudicial que ventajoso, y puede llevar a grupos de la población a considerar la necesidad de los incentivos ante cualquier circunstancia.

Debe considerarse la conveniencia de identificación del entrevistado a través de su nombre completo y dirección. Comúnmente se considera que hay una tasa de respuesta más elevada, si al encuestado no se le pregunta su nombre y dirección, aunque debe tenerse presente que hay estudios donde tal identificación es necesaria.

Hay que calcular la tasa de respuestas. La cual debe ser manejada adecuadamente. Es el número de entrevistas cubiertas dividido por el tamaño de muestra, Ballar y Lanphier (1978). Hay que definir si las no respuestas de los que no pueden responder, por problemas de idioma u otro índole, son elegibles para el cálculo de la tasa de respuestas. Hay dificultad en comparar la tasa de respuesta entre encuestas cuando se hacen, sustituciones de las unidades que no responden, y no se incluye la unidad sustituida en el total de no respuestas; así que la tasa de respuesta es más alta que la realidad.

#### **ETAPA IV: Tratamiento y análisis de las respuestas.**

### Capítulo III: Desarrollo Teórico Del Muestreo

Esta etapa se inicia con la revisión de los cuestionarios, la codificación de las preguntas abiertas, la verificación y depuración de los códigos, y finaliza con la publicación del informe de investigación.

Las observaciones realizadas, y que han sido registradas en los cuestionarios, son procesadas de acuerdo al plan de tabulaciones. El procedimiento de la información es el proceso mediante el cual los datos recopilados en los cuestionarios se organizan de acuerdo a determinados lineamientos o hipótesis, a fin de alcanzar los objetivos planteados en la investigación. Los datos se ordenan y concentran en tablas y gráficos de tal forma que puedan manipularse estadísticamente con el objeto de facilitar el análisis de la información empírica. Por lo tanto se tendrán clasificaciones y cálculos para los elementos observados, y se presentarán los resultados numéricos de esas clasificaciones y cómputos en tablas y gráficos acompañados de los comentarios indispensables para su comprensión y precisión de su alcance. El procesamiento de los datos puede ser manual, mecánico, mecánico-manual y electrónico. Finalizado el procesamiento de la información y después de realizar los análisis correspondientes, los resultados analizados se presentan resumidos en un Informe final.

Para realizar el procesamiento de la información recopilada, es necesario ejecutar las siguientes operaciones:

**a) Edición Preliminar:** La primera etapa de un procesamiento de datos es la revisión de los cuestionarios. Cuando se realizan entrevistas personales, ella es realizada por los supervisores en las revisiones de campo para completar totalmente los cuestionarios. Con frecuencia se observan no respuestas y datos inconsistentes en todo el cuestionario o en partes de él, llenado intencional de las preguntas, desidia en la anotación de las respuestas, sustitución arbitraria de las unidades muestrales, etc. Los cuestionarios deben editarse en forma preliminar cuando se reciben en el lugar de procesamiento. Si esto se hace desde un principio permite corregir errores de los encuestadores y es posible realizar

### Capítulo III: Desarrollo Teórico Del Muestreo

reentrenamiento o despido de los mismos. Muchas veces es necesario regresar los cuestionarios al campo para completarlos o para resolver inconsistencias.

**b) Codificación:** Es el proceso mediante el cual se asigna un código (letras o preferiblemente números) a cada una de las respuestas previstas para las preguntas de un cuestionario. Dicho proceso facilita el procesamiento de la información. Una vez que se asigna un código para cada alternativa de respuesta, éste se utilizará en lo sucesivo para tabular la información.

Es recomendable el uso de preguntas pre codificadas, pero esto no siempre es posible. Cuando el cuestionario tiene preguntas abiertas es necesaria la codificación de las mismas antes del procesamiento, labor que debe ser estrictamente evaluada y supervisada para verificar la veracidad de la codificación.

El cierre de las preguntas abiertas significa establecer categorías, es decir, agrupar las respuestas similares que se dan a una misma pregunta estructurada en forma abierta. Teóricamente existen tantas posibilidades de respuestas a las preguntas abiertas como individuos se encuestan o entrevistan. Es necesario, por lo tanto, agrupar las respuestas comunes, o que se expresan en distintos términos pero que significan lo mismo, a fin de poder manejar la información considerando un número reducido de categorías.

Al cerrar las preguntas abiertas cabe siempre la posibilidad de no contar con las categorías necesarias para incluir todas las respuestas. Por ello, se recomienda que se efectúen las pruebas pertinentes, para tener seguridad de que las categorías establecidas cubran la generalidad de las respuestas obtenidas en una pregunta abierta.

**c) Transcripción de Datos:** Esta es una labor muy importante en la planificación y ejecución del procesamiento de los datos, que debe ser controlada y supervisada adecuadamente, porque en la transcripción se genera la materia

### Capítulo III: Desarrollo Teórico Del Muestreo

prima para la tabulación de los resultados finales. Aunque se considera una tarea de rutina y a veces se subcontrata, hay que establecer los controles correspondientes cuando se hace la transcripción de los datos a medios magnéticos como discos o cintas.

Para garantizar la calidad de la transcripción de los datos hay que hacer una inspección del trabajo de los operadores, Zarkovich (1968) recomienda realizarla mediante los siguientes procedimientos:

**Inspección normal:** Se utiliza cuando la tasa de error de un operador es sensiblemente la misma que la norma de calidad deseada en el procedimiento de transcripción de los datos. Constituye el fundamento del método de control estadístico de la calidad de la transcripción y se apoya en un plan de muestreo de las unidades de trabajo del operador.

**Inspección al 100%:** Algunas veces se inspecciona la totalidad de todas las unidades de trabajo de los operadores, con lo que un personal especializado crea un doble registro de la información. El doble registro permite la corrección, por parte del personal de inspección, de los valores erróneos introducidos y la orientación en la formación de nuevos operadores, y de aquellos que presentan una tasa de error constantemente muy alta. No se recomienda esta inspección exhaustiva por ser muy costosa. Por lo general, se implementa en el periodo de aprendizaje, cuando el operador se inicia en el trabajo y cuando la tasa de error de algún operador es muy elevada.

**Inspección reducida:** Si la tasa de error de un operador es mucho más baja que la norma de calidad establecida para la transcripción de los datos, la optimización del plan de control de calidad de ese operador conduce a determinar muestras de inspección de tamaño más pequeño, en la cual el umbral de aceptación de los lotes será mucho menos exigente que en el caso de inspección normal. Cuando la calidad esperada del trabajo de un operador es considerada

### Capítulo III: Desarrollo Teórico Del Muestreo

"buena", respecto a la norma establecida, se podrá garantizar una buena calidad media final de su trabajo, con una baja carga de trabajo de inspección del mismo.

**Inspección Ajustada:** Se aplica cuando la tasa de error de un operador, es sistemáticamente más elevada que la normas de calidad deseada en la transcripción de los datos. Esta inspección se distingue de la inspección normal en el empleo de una muestra de tamaño más grande y un umbral de aceptación de una muestra de inspección más bajo.

**d) Edición Definitiva:** Se establecen restricciones en los valores que puede tomar una variable, las cuales pueden ser especificadas como reglas de aceptación o rechazo. Los errores en los registros se detectan según dichas restricciones. Hay máquinas y programas especiales que chequean el rango de cada pregunta para corroborar que los valores de los códigos son válidos y consistentes.

El programador debe participar activamente, incluso desde el diseño del cuestionario. La programación se efectúa antes que se inicien las labores de campo, y los programas son probados con los primeros cuestionarios que se recogen, Seijas (1981), Villan (1985). El programa muestra cada campo que falla en la edición.

**e) Ajuste da datos Erróneos y Preguntas no Respondidas:** En muchas encuestas, aun si todas las personas seleccionadas participan, algunas dejan ciertas preguntas sin responder, bien porque no entendieron la pregunta, la pasaron por encima o, simplemente, no quisieron responder. El resultado puede ser una tasa de no respuesta mayor, para algunas preguntas, que la tasa total de no respuesta. Además, mediante la edición se detectan datos que se consideran erróneos. En estos casos se hacen ajustes mediante programas de computación, para estimar el dato faltante o corregir el erróneo, tomando en cuenta los valores limites que puede tomar la variable, los valores promedio, la información de otras

### **Capítulo III: Desarrollo Teórico Del Muestreo**

variables correlacionadas con aquella " los datos que suministran las otras unidades muestrales. Este proceso se denomina "depuración".

**f) Ponderaciones:** Con frecuencia hay que establecer ponderaciones para los datos recopilados. Estas son necesarias cuando:

- 1) Se seleccionan las unidades muestrales con probabilidades variables, según el diseño muestral considerado.
- 2) Se hacen ajustes por no respuestas.
- 3) Se hacen ajustes por sesgos de otras fuentes.

Algunos ajustes dependen de estimaciones de población recientes, por raza, sexo, religión u otra variable. El uso de ponderaciones depende de la sofisticación estadística de los conductores de la encuesta. Cuando se utilicen, deben justificarse y describirse detalladamente.

Se debe tener siempre presente que la encuesta se hace, no para conocer las características de las unidades muestrales, sino para estimar los parámetros correspondientes. Generalmente se hacen estimaciones de totales, promedio y porcentajes simples. Algunas veces se requieren análisis más complejos, como regresiones, estimación de correlaciones, análisis factoriales, de correspondencia, etc. Ahora es fácil hacer tales análisis con la ayuda de paquetes estadísticos de computación.

Los planes de estimación deben incluir las estimaciones de las varianzas, para lo cual hay que considerar el diseño muestral empleado, sobre todo si hay ponderaciones. No deben hacerse como si se considerara, un muestreo aleatorio simple que es un error cometido con mucha frecuencia.

#### **4.1. -Análisis de los Resultados.**

### **Capítulo III: Desarrollo Teórico Del Muestreo**

Consiste en examinar la información tabulada, presentada gráficamente u obtenida de fuentes documentales de acuerdo con una hipótesis o idea rectora, y considerando los objetivos de la investigación, con el propósito de responder a las distintas interrogantes planteadas. El análisis permite penetrar en el conocimiento exhaustivo de los diversos componentes del objeto de estudio con el fin de llegar a una comprensión profunda de sus diferentes aspectos y relaciones. Hecho el análisis se procede a la reconstrucción, vía el pensamiento conceptual, de la realidad concreta. El resultado de esta operación es una síntesis que permite tener una visión de totalidad del fenómeno que se estudia.

El análisis de la información se realiza, por lo regular, partiendo de los aspectos más generales del tema, a fin de introducirse paulatinamente en las cuestiones específicas y de mayor profundidad.

En un primer momento el análisis es básicamente descriptivo, y tiene por objeto conocer la forma como se manifiestan los diferentes componentes de la situación que se estudia, teniendo presente la determinación de la confiabilidad; de las estimaciones. En un segundo momento, se busca conocer el comportamiento de las variables en estudio para interrelacionarlas en cuestiones que se refieren a un mismo aspecto del fenómeno, y lograr una visión de conjunto del problema investigado. El análisis de datos debe permitir la consecución de los objetivos y la comprobación de las hipótesis planteadas.

La mayor dificultad estadística está en la interpretación de los datos presentados. Informes sobre cosas controversiales, o en situaciones en que los resultados de la encuesta afectan políticas públicas, deben, necesariamente, incluir suficientes evidencias para soportar los hallazgos. Así que cualquier inferencia debe hacerse con suficiente exactitud. Las comparaciones entre diferentes grupos deben acompañarse por las varianzas muestrales y las estimaciones de los sesgos. Es necesaria una advertencia cuando no se utiliza un muestreo probabilístico.

### **Capítulo III: Desarrollo Teórico Del Muestreo**

Cuando se tiene una tasa alta de no respuesta, los usuarios deben ser cuidadosos con su influencia y no considerar una forma fácil de comprender el problema, los resultados obtenidos deben soportar el peso de cualquier análisis a que sean sometidos en relación a la confiabilidad estadística de las cifras.

Una vez analizada la información, se elaboran las conclusiones del estudio. Esta es la parte del trabajo en la que el investigador resume, de acuerdo con las hipótesis y objetivos planteados, los principales hallazgos encontrados en la investigación. Las conclusiones pueden hacerse en términos que abarquen los distintos aspectos de la misma, destacando aquellos -según el criterio del investigador- que son básicos para mostrar los resultados teóricos o empíricos del estudio. Las conclusiones pueden también referirse a cada uno de los capítulos que componen el trabajo, para alcanzar una mayor comprensión del texto.

Es importante destacar que en las conclusiones del estudio pueden plantearse aquellas hipótesis que surgen como resultado final del trabajo y que servirán como base para iniciar nuevos estudios. De esta manera, las conclusiones cierran un proceso concreto de investigación, lo que no implica que se llegue a verdades absolutas, acabadas, sino que se arriba a un determinado conocimiento que puede enriquecerse a medida que se lleven a cabo otras investigaciones sobre el objeto de estudio.

#### **4.2. Informes de los Resultados de la Encuesta.**

En una encuesta por muestreo, se pueden considerar varios tipos de informes: puede ser uno simple con los resultados muestrales y la metodología, puede ser un informe separado de la metodología muestral empleada. Un contratante externo del estudio puede requerir sólo los resultados y la metodología, mientras que la oficina que lo realiza se encarga de la interpretación correspondiente.

### Capítulo III: Desarrollo Teórico Del Muestreo

El informe de investigación es el resumen que se realiza de los diferentes aspectos de la investigación para difundir los resultados del trabajo o para entregarlo a los patrocinadores del proyecto de investigación. El informe debe contener los elementos más importantes que permitan la comprensión del proceso que llevó a las conclusiones y sugerencias.

En los informes no deben omitirse los aspectos siguientes:

- 1) Discusión de los factores del diseño.
- 2) Posibles influencias de las correlaciones Intraclásicas.
- 3) Costos de los diferentes procedimientos.
- 4) Implicaciones de una tasa alta de no respuesta.
- 5) información detallada de los cálculos realizados.

Los cuestionarios con frecuencia los mantiene quien realiza la encuesta. Los datos generalmente se suministran en cintas o discos, copia de los cuales se dejan en la oficina que realiza la encuesta y se suministra una copia a los contratantes.

**Confiabilidad de los resultados:** Es una consecuencia de todo el proceso de investigación, así como de los distintos procedimientos utilizados para recopilar, clasificar y analizar la información. La confiabilidad está relacionada con la precisión. Seijas (1981), señala que el grado de confiabilidad indica la calidad del dato o estimación obtenida en la investigación. Se expresa a través del error total, el cual depende tanto de los errores de muestreo como de errores ajenos al muestreo (no de muestreo). Los errores que no son de muestreo se determinan en forma aproximada, mediante el estudio y análisis del sistema de control de operaciones.

Una vez discutidas cada una de las etapas para la ejecución óptima de una investigación por muestreo, es de importancia capital tener presente que no debe subestimarse el tiempo de cada una de las etapas, y que el tiempo total varía.

### **Capítulo III: Desarrollo Teórico Del Muestreo**

Según la encuesta, de 2 a 3 semana hasta más de 1 año. Hay que cumplir con todas las etapas. No abordar algún proceso en una de las etapas puede ser catastrófico. En las operaciones de campo nunca serán exagerados los procedimientos de control de material y personal involucrados, y los controles de calidad establecida.

Aunque los cuestionarios son los más usados también existen instrumentos mecánicos como el galvanómetro, el cual detecta los grados diminutos de sudor que produce la excitación de las emociones. El taquiscopio sirve para proyectarle un anuncio a un sujeto, con un rango de exposición de entre menos de una centésima de segundo a varios segundos, luego describe todo lo que recuerda. Las cámaras de ojos se usan para estudiar los movimientos oculares de los sujetos para determinar hacia donde enfoca la vista primero y que tanto tiempo permanece ésta en un artículo dado cualquiera.

#### **Ejemplo del uso del Muestreo en Ingeniería Industrial**

Se tomó el Trabajo de Grado de Fagúndez y Goya (2009). “Evaluación de la calidad de servicio percibido por los usuarios de las unidades sectoriales de recursos humanos de la Universidad de Carabobo”

La gestión de recursos humanos de una organización es fundamental para el desarrollo de las actividades diarias de la misma. Además, se evidencia como tiene cada vez mayor peso, debido a la importancia en la aplicación de medidas como la formación continua del Personal o el ambiente laboral; es por ello que la calidad de servicio ha tomado un papel primordial con relación a dicha gestión.

De esta realidad no escapan las Unidades Sectoriales de Recursos Humanos de la Universidad de Carabobo, en las mismas se observa como el usuario al acudir a ellas percibe retraso en el servicio y emisión de algunos Trámites administrativos, así como poca información sobre talleres desactualización, entre otros. Es por ellos que el objetivo de realizar una encuesta

### Capítulo III: Desarrollo Teórico Del Muestreo

es evaluar la calidad del servicio percibido por los usuarios de las Unidades Sectoriales de Recursos Humanos de la Universidad de Carabobo con el uso del modelo SERVQUAL para evaluar la calidad percibida plantean estas dimensiones generales y definen que dicha percepción es consecuencia de la diferencia para el consumidor entre lo esperado y lo percibido. Sin embargo, también parten del planteamiento del hallazgo de una serie de desajustes o *gaps* en el proceso. Éstos influyen en la percepción del cliente y son el objeto de análisis cuando se desea mejorar la calidad percibida. Así, las percepciones generales de la calidad de servicio están influidas por estos *gaps* que tienen lugar en las organizaciones prestadoras de servicios. Parasuraman, Zeithaml y Berry (1985), definen *gap* como una serie de discrepancias o deficiencias existentes respecto a las percepciones de la calidad de servicio de los ejecutivos y las tareas asociadas con el servicio prestado a los consumidores. Estas deficiencias son los factores que afectan a la posibilidad de ofrecer un servicio percibido por los clientes como de alta calidad. El modelo SERVQUAL, con el estudio de los cinco *gaps*, analiza los principales motivos de la diferencia que llevaban a un fallo en las políticas de calidad de las organizaciones. El resultado es el modelo del *gráfico 1*, que presenta cuatro *vacíos* identificados por los autores como el origen de los problemas de calidad del servicio.

#### **Población**

Para efectos de la investigación: la población o universo a estudiar estuvo constituida por los empleados y obreros fijos de las Unidades Sectoriales de las Facultades de Ingeniería y Ciencias y Tecnología, ya que se sabe gracias a la Dirección de Recursos Humanos que la primera Facultad podría estar más crítica que la segunda en cuanto a la calidad del servicio que ofrecen, y esto servirá como punto de comparación a la hora de realizar la evaluación y validación del instrumento.

#### **Muestra**

La muestra de estudio corresponde a las denominadas muestras probabilísticas estratificadas proporcionalmente, para los diferentes tipos de personal (Administrativo y Obrero), en la medida que se consideraron como

### Capítulo III: Desarrollo Teórico Del Muestreo

estratos tanto la Facultad de Ingeniería como la Facultad Experimental de Ciencias y Tecnología.

#### Tamaño de la muestra

Tomando en cuenta los conocimientos adquiridos por las autoras de estadística para el cálculo de la muestra se parte con una distribución binomial de parámetros  $(n, p)$ , que mide el número de éxitos en una secuencia de  $n$  ensayos independientes de Bernoulli, con una probabilidad fija  $p$  de ocurrencia del éxito entre los ensayos, pero como se está estudiando una población finita mayor que 30 por el teorema central del límite, la distribución binomial se aproxima a una distribución normal con parámetros  $\mu = n.p$  y  $\sigma = n.p.q$ . y en ésta, se establece el error "e" de donde se podrá despejar el valor de n.

Tomando en cuenta el error e y despejando n se obtiene la siguiente ecuación para el cálculo del tamaño de la muestra:

$$n = \frac{p \cdot q \cdot z^2}{e^2}$$

Como la población es conocida esta ecuación se multiplica por el factor  $(N - n)/(N - 1)$ , generando la siguiente ecuación:

$$n = \frac{p \cdot q \cdot Z^2 \cdot N}{(N - 1)e^2 + p \cdot q \cdot Z^2}$$

Donde:

n: Tamaño muestral

N: Población

p: Probabilidad de éxito

q: Probabilidad de fracaso

Z: Estadístico de la distribución normal.

E: Error que se prevé cometer

### Capítulo III: Desarrollo Teórico Del Muestreo

La población es de 280 trabajadores de las facultades de Ingeniería y Ciencia y Tecnología.

Para la probabilidad de éxito se aplicó la opción más desfavorable  $p=0,5$ , que hace mayor el tamaño muestral por lo tanto, la probabilidad de fracaso será  $q=0,5$ . El estadístico de la distribución normal estandarizada con un nivel de confianza de 95% genera un fractil de 1,96. El Error que se previó cometer fue de 10%. Aplicando la fórmula para el Personal Administrativo.

$$n = \frac{(0,50)(0,50)(1,96)^2(220)}{(220 - 1)(0,1)^2 + (0,50)(0,50)(1,96)^2}$$

$$n = 67,06 \approx 67$$

Aplicando la fórmula para el Personal Obrero.

$$n = \frac{(0,50)(0,50)(1,96)^2(60)}{(60 - 1)(0,1)^2 + (0,50)(0,50)(1,96)^2}$$

$$n = 37,16 \approx 37$$

Finalmente se realizaron 120 encuestas, pues hubo algunas que se eliminaron del estudio, porque no fueron contestadas objetivamente.

#### Tipo de muestreo

Después de calcular el tamaño de la muestra, se utilizó el tipo de muestreo probabilístico estratificado proporcional. Estratificado, en la medida que la población está dividida en sub-grupos de acuerdo a la Facultad a la que pertenezca (Ingeniería y Ciencias y Tecnología) y Proporcional, ya que, el tamaño

### Capítulo III: Desarrollo Teórico Del Muestreo

de cada estrato se fijo teniendo en cuenta la población da personal (Administrativo y Obrero) que labora en ellas.

El cálculo del Factor proporción para el Personal Administrativo.

$$F = \frac{n}{N} = \frac{67}{220} = 0,304$$

Estrato del Personal Administrativo	Tamaño de la población de cada estrato	Determinación del tamaño de los estratos muestrales	Estratos de la muestra
F. Ingeniería	148	148*0,304	45
F. Ciencias y Tecnología	72	72*0,304	22
Total	220	-	67

**Tabla 7. Muestra Estratificada Proporcional del Personal Administrativo.**

El cálculo del Factor proporción para el personal Obrero.

$$F = n/N = 37/60 = 0,616$$

Estrato del Personal Administrativo	Tamaño de la población de cada estrato	Determinación del tamaño de los estratos muestrales	Estratos de la muestra
F. Ingeniería	54	54*0,616	33
F. Ciencias y Tecnología	6	6*0,616	4
Total	60	-	37

**Tabla 8. Muestra Estratificada Proporcional del Personal Obrero.**

## Capítulo III: Desarrollo Teórico Del Muestreo

### Fuentes y técnicas de recolección de datos

Para el proceso de investigación se requirió del uso de diversas técnicas e instrumentos que permitieran a las investigadoras obtener toda la información o datos necesarios para el desarrollo del mismo.

### Fuentes

#### Fuentes Primarias

La Observación Directa: para este trabajo de investigación se aplicó la observación directa, ya que las investigadoras pudieron evidenciar y visualizar de cerca la problemática de las Unidades Sectoriales de Recursos Humanos de las diferentes Facultades.

Mediante la encuesta se reunió de manera sistemática los datos necesarios de los sujetos que conforman la población de estudio, en este caso, los empleados y obreros fijos de la Facultad de Ingeniería y la Facultad Experimental de Ciencias y Tecnología de la Universidad de Carabobo. Se utilizó, una modificación del instrumento *SERVQUAL* llamado *SERVQUAL<sub>Ing</sub>* (Mejías, 2004), que fue usado por éste para conocer la percepción acerca de la calidad del servicio de los estudiantes del Postgrado en Ingeniería de la Universidad de Carabobo. Este cuestionario ha sido validado anteriormente en diferentes entornos, por estudios previos y bajo características similares, todo con el propósito de conocer sus puntos de vista y opiniones, para lograr una mayor objetividad al definir la situación actual.

#### Fuentes Secundarias

Se revisó la información escrita disponible, como libros, documentos electrónicos, artículos de revistas y otros documentos de trabajo; para consulta y aclarar dudas teóricas en el contexto de la Gestión de Recursos Humanos en Universidades y también, acerca de la Calidad de Servicio.

### Técnicas

### **Capítulo III: Desarrollo Teórico Del Muestreo**

La calidad de las investigaciones realizadas por estudios estadísticos se fundamenta en la utilización de las más diversas y actualizadas técnicas de recolección y análisis de datos, que permiten extraer el mayor provecho a los fenómenos estudiados.

#### **Análisis estadístico para evaluar la calidad de servicio**

El modelo SERVQUALing posee una amplia utilización en todo tipo de servicios y ha sido aplicado en la Universidad de Carabobo en investigaciones similares a ésta, tales como la de Gual y Miquilena (2004), Villegas (2008); entre otras, lo que permitió establecer comparaciones entre los resultados obtenidos.

Para evaluar la calidad de servicio percibido por los usuarios de las Unidades Sectoriales de Recursos Humanos de la Universidad de Carabobo se diseñó una encuesta basada en las cinco dimensiones establecidas por el modelo *SERVQUALing*, compuestas por un total de veintidós ítems que fueron adaptados al servicio evaluado, además se agregaron ocho ítems, para realizar las evaluaciones necesarias para validar el instrumento. Por lo tanto, la encuesta está compuesta por 30 preguntas que sirvieron para medir la calidad de servicio, a través de una escala de Likert que oscila entre 1 y 7 puntos, donde 1 es totalmente en desacuerdo, 2 bastante en desacuerdo, 3 en desacuerdo, 4 ni en desacuerdo ni de acuerdo, 5 de acuerdo, 6 bastante de acuerdo y 7 totalmente de acuerdo.

#### **Análisis de las medias de las variables**

Con este análisis se busca determinar la percepción de los trabajadores (Administrativos y Obreros) con respecto al servicio que reciben de las Unidades Sectoriales de Recursos Humanos de la Universidad de Carabobo. Este estudio se realizó calculando la media de cada una de las variables aplicadas en el instrumento, tanto para la Facultad de Ingeniería como para la Facultad Experimental de Ciencias y Tecnología, con el fin de comparar si existen discrepancias o no con respecto al servicio percibido por sus trabajadores.

### Capítulo III: Desarrollo Teórico Del Muestreo

#### Análisis de las Medias para la Facultad de Ingeniería

En la Tabla se puede observar que la variable con media más alta es V3, ésta corresponde a la pregunta "El personal de la U.S.RRHH que le atiende a usted tiene apariencia pulcra" con un valor de 5,45 esto quiere decir que los trabajadores (Administrativos y Obreros) de las Unidades Sectoriales de la Facultad de Ingeniería, según la escala, lo perciben como un poco de acuerdo. La variable V16 "El personal es siempre amable con usted" tiene una media de 5,01, lo cual indica que los empleados de las Unidades Sectoriales le dan un buen trato a algunos de sus trabajadores, y la variable V19 "La U.S.RR.HH tiene horarios de trabajo convenientes para todos ustedes" tiene un valor de 4,40 esto quiere decir que los trabajadores no están ni de acuerdo ni E:3n desacuerdo con los horarios de atención que presta la Unidad Sectorial para esta Facultad.

Las variables con menores medias son V21, V14 y V11. La Variable V21 "La U.S.RRHH se preocupa por sus intereses" tiene una media de 3,19 esto afirma que la Unidad Sectorial de Ingeniería no está realizando estrategias para conocer las necesidades de sus trabajadores porque en su mayoría lo perciben así. Con una media de 3,31 se encuentra la variable V14 "El comportamiento del personal de la U.S.RRHH le inspira confianza" este valor indica que los trabajadores no están de acuerdo con el trato suministrado por los empleados de esta Unidad. La variable V11 "El personal le ofrece un servicio rápido" tiene una media de 3,42 quiere decir que no se efectúan los tramites en el tiempo establecido.

Variable	Media	Desviación Estándar	Variable	Media	Desviación Estándar
V1	4,96	1,743	V12	4,52	1,762
V2	5,37	1,573	V13	4,56	4,649
V3	6,04	1,3156	V14	4,07	1,92
V4	3,67	1,544	V15	3,89	1,826
V5	3,56	1,528	V16	5,33	1,861
V6	4,15	1,61	V17	4,37	1,334
V7	3,89	1,423	V18	4,74	1,678
V8	3,85	1,61	V19	4,85	1,875
V9	4	1,732	V20	4,26	1,678

### Capítulo III: Desarrollo Teórico Del Muestreo

V10	4,3	1,815	V21	3,85	1,791
V11	3,74	1,559	V22	4,07	1,466

**Tabla 9. Media de las variables del modelo teórico aplicado SERVQUAL para la facultad experimental de Ciencias y Tecnología**

#### **Análisis de las Medias para la Facultad Experimental de Ciencias y Tecnología**

Las medias obtenidas de las encuestas realizadas al personal Administrativo y Obrero de la Facultad de Ciencias y Tecnología se pueden observar en la Tabla, donde la variable V3 "El personal de la U.S.RR.HH que le atiende a usted tiene apariencia pulcra", es la de mayor media con un valor de 6,04, estos resultados son similares a los obtenidos en la Facultad de Ingeniería, es decir, que indudablemente el personal de las Unidades Sectoriales no tiene problemas con respecto a este ítems. La variable V2 "Las Instalaciones físicas de la U.S.RRHH son visualmente atractivas" tiene una media de 5,37, resultado esperado ya que las instalaciones son nuevas, por lo tanto los trabajadores se sienten a gusto con éstas. La variable V16 "El personal es siempre amable con usted" tiene una media de 5,01, esto quiere decir, que los trabajadores están de acuerdo con el trato ofrecido por los empleados de las Unidades Sectoriales. Estas variables antes mencionadas, son la que menos tienen problema en esta Facultad por lo tanto es recomendable enfocarse en mejorar aquellas cuyas medias sean más bajas.

La variable V5 "Cuando el personal de la U.S.RRHH promete hacer algo en cierto tiempo, lo hace" es la de menor media con una valor de 3,56, esto implica, que los empleados no están cumpliendo los servicios en el tiempo establecido. La variable V4 "Los documentos emitidos (folletos, reportes, similares) son visualmente atractivos" tiene una media de 3,67, es decir, los empleados de la Unidad Sectorial] de esta Facultad deben ocuparse un poco más en realizar sus carteleras y folletos atractivos para poder llamar la atención de los trabajadores. La variable V11 "El personal le ofrece un servicio rápido" tiene una media de 3,74,

### Capítulo III: Desarrollo Teórico Del Muestreo

esto confirma que los empleados se están tardando más de lo debido en prestar los servicios.

Variable	Media	Desviación Estándar	Variable	Media	Desviación Estándar
V1	4,31	1,699	V12	3,97	1,623
V2	4,27	1,847	V13	4,33	1,744
V3	5,45	1,726	V14	3,31	1,868
V4	3,58	1,733	V15	3,49	1,77
V5	3,69	1,751	V16	5,01	1,674
V6	3,64	1,756	V17	4,16	1,88
V7	3,94	1,783	V18	4,15	1,681
V8	3,73	1,638	V19	4,4	1,923
V9	3,66	1,763	V20	3,94	1,766
V10	3,54	1,925	V21	3,19	1,644
V11	3,42	1,751	V22	3,48	1,82

**Tabla 10. Media de las variables del modelo teórico aplicado SERVQUAL para la facultad de Ingeniería**

#### III.3.2 Aplicaciones en otras áreas

Podría resultar arduo enumerar todas las otras aplicaciones del muestreo ya que prácticamente puede usarse de manera ilimitada.

A continuación se hace mención de aplicaciones, que aunque son importantes, no son parte de esta investigación.

#### Demografía

La demografía estudia fenómenos tales como natalidad, nupcialidad, fecundidad, mortalidad, migraciones, actividad y educación con el fin de hacer inferencia y proyecciones. Para tal fin, se realizan censos los cuales no son más que muestreos por medio de los cuales se obtiene toda la información necesaria para estudiar aquellos procesos que determinan la formación, la conservación y la desaparición de las poblaciones.

## **Capítulo III: Desarrollo Teórico Del Muestreo**

### **Política**

El concepto de la encuesta política tiene su origen en 1960 previo a las elecciones presidenciales de John F. Kennedy, realizando un sondeo el cual sirvió para diseñar la campaña que logró derrotar a Richard Nixon.

Existen dos tipos de encuesta política: las públicas que son de conocimiento general ya que se publican en periódicos, revistas, radio y televisión; y las privadas cuyos resultados son secretos bien guardados por el candidato o partido político ya que la información que proporcionan es usada para el desarrollo de las estrategias y mensajes de la campaña.

### **Contabilidad**

No siempre es posible que un interventor público realice un examen de un gran número de transacciones por lo que emplea muestras intencionales u opináticas para juzgar las aseveraciones sobre cuestiones financieras.

Ahora bien, el problema no radica en la recolección, propiamente dicha, sino en establecer un diseño eficiente pueda proporcionar estimación de grado de precisión y seguridad que se considere adecuado para el caso.

### **Economía**

Una de las aplicaciones relevantes del muestreo es en la que se obtiene información de los presupuestos familiares, entendiéndose por familia a un grupo de individuos que habitan bajo un mismo techo aun cuando no tengan ningún parentesco. El objetivo principal es el estudio del nivel de ingresos y egresos con la finalidad de hacer un análisis de la demanda en sus diversos aspectos, por ejemplo, el aumento del consumo de ciertos artículos al aumentar los ingresos.

### **Biología**

Cuando se trata de poblaciones zoológicas y botánicas que no sean domésticas o cultivadas, puede que no solo no se conozca el tamaño de la población sino que resulte difícil calcularlo.

### **Capítulo III: Desarrollo Teórico Del Muestreo**

El conocimiento del tamaño de la población puede ser importante para estimar los recursos biológicos con fines alimenticios, industriales, etc. así como también por la preocupación por conservar el equilibrio ecológico, cuya destrucción puede ser bastante perjudicial para el entorno. Se plantea así el óptimo de animales que pueden cazarse o vegetales que pueden recogerse sin perjudicar peligrosamente a la población.

Pero el recuento o enumeración total ya es bastante dificultoso y antieconómico por lo que todas las ventajas del muestreo son aplicables para este tipo de poblaciones.

## **Conclusiones**

La teoría del muestreo es el estudio de las relaciones existente entre una población y muestras extraídas de la misma. Las muestras se obtienen con la intención de inferir propiedades de la totalidad de la población y significan lograr una versión simplificada de ésta, que reproduzca de algún modo sus rasgos básicos, para lo cual deben ser representativas de la misma. Para cumplir esta característica la inclusión de sujetos en la muestra debe seguir una técnica de muestreo.

Cada estudio tiene un tamaño muestral idóneo que permite comprobar lo que se pretende con una seguridad aceptable y el mínimo esfuerzo posible, para ello existen fórmulas estadísticas apropiadas. El punto de partida para afrontar un diseño muestral es definir cuál es el universo o población de estudio, de acuerdo con los objetivos planteados ya que de ello dependerán los resultados.

El muestreo como herramienta de estudio tiene muchas ventajas, tales como costos reducidos, mayor rapidez para obtener resultados, mayor exactitud o calidad de información ya que se maneja un volumen reducido de trabajo, puede existir mayor supervisión del trabajo y existe una probabilidad baja de cometer errores durante el procesamiento de la información.

Actualmente en la escuela de Industrial al tema de la teoría del muestreo no se le ha dado la importancia que merece en lo que respecta a la formación de los profesionales que egresan de ésta, lo que puede ser visto como una desventaja cuando salen al mercado laboral.

En la práctica profesional el Ingeniero Industrial, en su incesante labor para la mejora de los procesos, se enfrenta a situaciones en las que el uso de la teoría del muestreo es fundamental para lo cual es importante que tenga conocimientos sólidos del tema, de lo contrario, puede conducir a la formación de conclusiones erróneas y opiniones en ese mismo nivel, lo que puede limitar la validez de los resultados y en ocasiones llegar a invalidar la investigación completa.

## **Recomendaciones**

Al Ingeniero Industrial en su formación profesional y educación continuada, deben exigírsele conocimientos sólidos del muestreo estadístico, por lo que es necesario reforzar los cursos de la carrera universitaria, es por ello que se invita al estudio de factibilidad para concretar la apertura de una materia de índole al menos optativa para los estudiantes que estén interesados en formarse en esta área, y que sea un profesional especializado en el área quien se encargue de impartirla.

Así como también, hacer especial énfasis del tema en asignaturas tales como Mercados, Control Estadístico de Calidad, Ingeniería de Métodos II, Métodos Estadísticos I y Métodos Estadísticos II.

Asimismo, se debe dotar de material bibliográfico más completo relacionado con el tema de la Teoría del Muestreo a la biblioteca rental de la escuela ya que lo que se encuentra en existencia es escaso

## Bibliografía

- Cochram, William G. "Técnicas de Muestreo" Editorial Continental. México 1972.
- Kish, Leslie "Muestreo de Encuestas" Editorial Trillas. México 1979
- Malhotra, Naresh K. "Investigación de mercados" Pearson Educación. México 2004.
- Azorin Poch, Francisco. "Curso de muestreo y aplicaciones". Editorial UCV. Caracas 1970.
- Canavos, George. "Probabilidad y Estadística - Aplicaciones y Métodos". McGraw-Hill 1988.
- Levin, Richard I. "Estadística para administradores". Prentice Hall. 2004
- Sabino, Carlos A. "Cómo hacer una tesis y elaborar todo tipo de escritos". Editorial Panapo. Caracas 1994.
- Burgos, Fernando. "Ingeniería de Métodos". Universidad de Carabobo. 2005.
- Claret Véliz, Arnold. "Cómo hacer y defender una tesis". Editorial Texto. Venezuela 2010.
- Seijas, Félix. "Investigación electoral. Encuesta por muestreo". Editorial UCV. Caracas 1998.
- Gutiérrez, Hugo Andrés. "Estrategias de muestreo. Diseño de encuestas y estimación de parámetros" Editorial Universidad Santo Tomás. Facultad de Estadística. Bogotá, Colombia. 2009.
- Walpole, Ronald E.; Myers, Raymond H., Myers, Sharon L. "Probabilidad y Estadística para Ingenieros". Prentice Hall. México. 1999.
- Vivanco, Manuel. "Muestreo Estadístico: diseño y aplicaciones" Editorial Universitaria. 2005.
- Farfán, Ramón. "Manejo de técnicas estadísticas de muestreo como factor importante en la carrera de contaduría pública". Facultad de Ciencias Económicas y Sociales (FACES). Universidad de Carabobo. 2002.

- Martínez de Tortolero, Evelyn. “Metodología para la realización de investigaciones por muestreo en el área socioeconómica y aplicaciones”. Facultad de Ciencias Económicas y Sociales (FACES). Universidad de Carabobo. 1996.
- Vinuesa, Julio; Zamora, Francisco; Génova, Ricardo; Serrano, Pedro; Recaño, Joaquín. “Demografía. Análisis y proyecciones”. Editorial Síntesis. 1997
- Peña, Daniel. “Análisis de datos multivariantes”. Editorial McGraw Hills. 2002
- Greenacre, Michael J. “Correspondence Analysis in Practice”. Academic Press. 1993.
- Lebart, Ludovic; Morineau, Alain; Warwick, Kenneth M. “Multivariate Descriptive Statistical Analysis”. John Wiley & Sons. 1984
- Cuadras, C.M. “Métodos de Análisis Multivariante”. Eunibar. 1981.
- Everitt, Brian S. “Cluster Analysis”. Halsted Press. 1993
- Revista Española de Investigaciones Sociológicas (REIS) número 111. Centro de Investigaciones Sociológicas (CIS). Julio-Septiembre 2005.
- Glosario – Hiru.com [http://www.hiru.com/gizarte\\_zientziak/glosarioa#M1](http://www.hiru.com/gizarte_zientziak/glosarioa#M1)
- Water Monitoring Alliance: Glosario <http://194.242.113.59/index.php?id=783&L=2>
- Definición de Muestreo <http://www.mitecnologico.com/Main/DefinicionDeMuestreo>
- Diccionario/ Federación de Municipios del Istmo Centroamericano <http://www.femica.org/diccionario/index2.php?strSearch=p>
- Diccionario del cambio climático <http://elguanche.net/Ficheros/diccionariocambioclimatico.htm>
- <http://oceanologia.ens.uabc.mx/~chelo/estadistica/doc-pdf/muestr-1.pdf>
- Muestreo no probabilístico [http://www.angelfire.com/sc/matasc/EyD/bioesta/muest\\_noprob.htm](http://www.angelfire.com/sc/matasc/EyD/bioesta/muest_noprob.htm)