



**Universidad Nacional Autónoma de México**

**FACULTAD DE CONTADURIA Y ADMINISTRACION**

**ELEMENTOS DE ESTADISTICA Y MUESTREO  
EN LA MERCADOTECNIA**

**Seminario de Investigación Administrativa**

**LORENZO ALVAREZ LOPEZ**

**MEXICO, D. F.**

**1975**

T

HF101

A4

C.1



1080080477

AL DE CANJE  
BIBLIOTECA - F. C. A.  
U. N. A. M.

**BIBLIOTECA**  
**GRUPO CYDSA FID. A-1-73 FIFOSA**



**Universidad Nacional Autónoma de México**

**FACULTAD DE CONTADURIA Y ADMINISTRACION**

**ELEMENTOS DE ESTADISTICA Y MUESTREO  
EN LA MERCADOTECNIA**

**Seminario de Investigación Administrativa**

**QUE PARA OBTENER EL TITULO DE  
Licenciado en Administración de Empresas**

**P R E S E N T A**

**LORENZO ALVAREZ LOPEZ**

**11863**

**MEXICO, D. F.**

**1 9 7 5**

HF1017  
A4

+  
A4722



**▲ MIS PADRES:**

**▲ VIRGINIA:**

**▲ MIS FAMILIARES Y AMIGOS:**

**QUE EN UNA FORMA Y OTRA ME HAN BRINDADO SU APOYO Y AMISTAD.**

**...GRACIAS.**

## INDICE.

### I.- METODOS ESTADISTICOS:

1.- RECOPIACION DE DATOS ESTADISTICOS.

2.- ORGANIZACION DE DATOS ESTADISTICOS.

3.- PRESENTACION DE DATOS ESTADISTICOS.

a).- Tablas estadísticas.

b).- Gráficas estadísticas.

4.- ANALISIS DE DATOS ESTADISTICOS.

a).- PROMEDIO.

Media.

Mediana.

Moda.

b).- DISPERSION.

Desviación cuartílica.

Desviación media.

Desviación estándar.

5.- INTERPRETACION DE LOS RESULTADOS DEL ANALISIS.

### II.- TEORIA DEL MUESTREO.

1.- DETERMINACION DEL TAMAÑO ADECUADO DE LA MUESTRA.

2.- TECNICAS BASICAS DEL MUESTREO.

a).- METODOS DE SELECCION DE MUESTRAS.

MUESTREO ALEATORIO SIMPLE.

MUESTREO ALEATORIO.

Aleatorio simple.

Aleatorio modificado.

Sistemático.

Estratificado.

De conglomerados.

### III.- HIPOTESIS.



1.-ERRORES TIPO I ( $\alpha$ ) Y TIPO II ( $\beta$ ).

2.- NIVEL DE SIGNIFICACION.

3.- PRUEBAS BASICAS DE HIPOTESIS.

a).- ENSAYO UNILATERAL.

b).- ENSAYO BILATERAL.

IV.- BIBLIOGRAFIA.

1.- BIBLIOGRAFIA PARA EL ALUMNO.

V.- PRACTICA.

1.- SOLUCION DE LA PRACTICA.

VI.- TABLAS. Y CUADROS SINOPTICOS.

VII.- CONCLUSIONES

1.) METODOS ESTADISTICOS.

## I.- METODO ESTADISTICO.

La Estadística es una técnica desarrollada para tratar metódica y sistemáticamente la información cuantitativa o datos numéricos que es necesario conocer para la solución de un problema determinado, para lo cual se recopilarán los datos importantes y significativos que hagan posible el análisis de la situación y los resultados obtenidos sean interpretados de acuerdo a las necesidades de ese problema determinado.

La información que se relacione con el estudio de cualquier situación particular puede registrarse o encontrarse expresada en números, es decir, puede ser cuantitativa o numérica y deberán obtenerse de ella conclusiones válidas y objetivas, bases para obtener las mejores soluciones.

La información cuantitativa utilizada para el estudio estadístico se llamará Datos Estadísticos y se define como: Un conjunto o conjuntos de números que representen cosas mensurables que muestren relaciones significativas para un asunto determinado y que sean susceptibles a comparaciones, análisis e interpretaciones válidas de acuerdo a los objetivos del estudio.

Al área de la cual los datos estadísticos son obtenidos para determinar las características que se estime es necesario conocer se llama población o Universo y es llamada Finita o Infinita, dependiendo de si el número de sus elementos es limitado o no.

La obtención de datos estadísticos de una población finita pero grande o de una población infinita constituye una tarea costosa que puede consumir una gran cantidad de tiempo y a veces es imposible, como es el caso de las poblaciones infinitas. Entonces es posible obtenerlos por medio de una Muestra O Muestras de elementos representativos que se extraen de la población en estudio. La muestra o muestras convenientemente seleccionadas nos dan resultados válidos para inferir y sacar conclusiones para toda la población.

Los datos estadísticos generalmente son abundantes o de un volumen tal que se dificulta su procesamiento, haciéndose necesario el uso de

métodos sistemáticos para organizarlos, presentarlos, analizarlos e interpretarlos con mayor efectividad. Para satisfacer esta necesidad se desarrollaron los llamados Métodos Estadísticos<sup>1</sup>, los cuales, de acuerdo con el orden de su aplicación en el proceso lógico de un estudio estadístico se dividen en cinco pasos básicos que son:

1.-RECOPIACION DE DATOS ESTADISTICOS.

2.-ORGANIZACION DE DATOS ESTADISTICOS.

3.-PRESENTACION DE DATOS ESTADISTICOS.

4.-ANALISIS DE DATOS ESTADISTICOS.

5.-INTERPRETACION DE LOS DATOS ESTADISTICOS.

1.-RECOPIACION DE DATOS ESTADISTICOS.

Los datos estadísticos necesarios para el conocimiento de una situación determinada de acuerdo con su localización pueden clasificarse en :

A).- Externos.

B).- Internos.

2.- ORGANIZACION DE DATOS ESTADISTICOS.

El primer paso para corregir cada uno de los datos recopilados es organizarlos. El siguiente paso es decidir las clasificaciones adecuadas para incluir todos los elementos recopilados. El último paso es tabular.

3.- PRESENTACION DE DATOS ESTADISTICOS.

La presentación de los datos estadísticos como primer requisito; debe estar presentado en una forma de fácil lectura y comprensión. Hay tres modos de presentar un conjunto de datos recopilados; a) Mediante enunciados, b) Tablas estadísticas, c) Gráficas estadísticas.

---

<sup>1</sup> SHAO P. STEPHEN. Estadística Para Economistas y Administradores. México, 1971 Cap. I.

#### 4.- ANALISIS DE DATOS ESTADISTICOS.

Los métodos de análisis estadístico más usados son:

- A).- **Análisis estadístico simple.** Es el método básico, hace uso de las medidas de tendencia central, de dispersión, etc.
- B).- **Inducción estadística.** Analiza una población o universo en base a un estudio muestral.
- C).- **Análisis de series de tiempos.** Esta parte analiza los cambios en las actividades de los negocios.
- D).- **Análisis de relación.** Analiza las relaciones entre dos o más conjuntos correspondientes de datos estadísticos.

#### 5.- INTERPRETACION DE DATOS ESTADISTICOS.

Una conclusión válida solo puede ser alcanzada después que los resultados del análisis son interpretados. Estos frecuentemente son mal interpretados causas como las siguientes; datos estadísticos inadecuados, sesgos personales supuestos falsos, indicación falsa de relación, comparaciones impropias y errores en operaciones matemáticas.

## 1.- LA RECOPIACION DE DATOS ESTADISTICOS.

Los datos pueden ser registros internos de la organización que llevan los diversos departamentos de la organización, los tipos más comunes que se conservan y que pueden ser utilizados son los registros contables, tales, como los balances y estados de resultados, mismos que pueden ser comparados con los de años anteriores.

Estos datos estadísticos también se pueden obtener de publicaciones; al obtener los datos de publicaciones es importante saber:

1.- La clase de fuente de datos; que pueden ser primarias cuando los datos se obtienen de la publicación editada por el recopilador original de los datos. La fuente se llama secundaria cuando los datos son obtenidos de una reimpresión, la cual es publicada por una organización distinta del recopilador original. En la realidad es más fácil la obtención de datos por fuentes secundarias pues es más el volumen de éstas que se publican que de fuentes primarias.

2.- Guías para localizar datos publicados. Siendo tan grande la cantidad de datos publicados en el área de los negocios es casi imposible para cualquier estadístico conocer todas las fuentes de la vasta cantidad de datos publicados, por lo que es conveniente hacer uso de las guías que están disponibles en bibliotecas públicas y universitarias.

3.- Las fuentes de datos publicados. Entre las principales fuentes de datos publicados encontramos a: Agencias gubernamentales, Asociaciones empresariales. Organizaciones privadas de servicios de estadística. Revistas y publicaciones periódicas de empresas. Periódicos y almanaques. Y organizaciones internacionales.

Por encuesta es la otra forma de obtener los datos deseados, pero a está es conveniente llegar hasta haber agotado las anteriores fuentes de datos. El trabajo de elaborar una encuesta esta limitado por los factores tiempo, dinero y mano de obra disponible para el estudio.

## 2.-LA ORGANIZACION DE DATOS ESTADISTICOS.

Para la organización de los datos estadísticos obtenidos se requiere efectuar tres pasos que son:

1.- Crítica y corrección de datos recopilados. El trabajo de corregir datos recopilados de registros internos y fuentes externas es relativamente simple. Sin embargo, los datos recopilados en una encuesta pueden necesitar corrección extensiva. Las respuestas de una encuesta registrada por observadores, entrevistadores o respondientes no son por lo regular presentados en una forma que faciliten la clasificación y su tabulación rápidamente. El corrector deberá hacer entonces una crítica y corregir cada respuesta que a la luz de los objetivos, lo requiera.

2.- Clasificación de datos corregidos. La clasificación de los datos obtenidos en una encuesta y los obtenidos de fuentes internas es básicamente el mismo. La colección de datos publicados están usualmente en forma clasificada, aunque algunas veces es necesario reclasificarla para nuestro estudio particular. La clasificación de los datos se hace con el fin de que se facilite la lectura e interpretación de los datos. Así, los datos se agruparan en clasificaciones con base al TIEMPO, LUGAR, CALIDAD, Y CANTIDAD. Por ejemplo: Las ventas mensuales de una gasolinera con base cronológica, puede clasificarse así:

MES (clase)	VENTAS (cantidades)
ENERO	\$ 189.000.00
FEBRERO	175.000.00
MARZO	200.000.00

Un grupo de gasolineras de acuerdo a su localización geográfica, pueden clasificarse de la siguiente manera:

LOCALIZACION DE LA GASOLINERA (CLASE)	NUMERO DE GASOLINERAS (CANTIDAD)
DISTRITO FEDERAL	110
GUADALAJARA	69
MONTERREY	85

Las clasificaciones también pueden tener como base una marca o sesgo característico o sea, una cualidad. En una gasolinera las ventas pueden clasificarse:

GASOLINA	LITROS VENDIDOS
ROJA	90.000
AMARILLA	89.000
VERDE	74.344

Sin embargo, todos los tipos de información cuantitativa pueden clasificarse de acuerdo al orden de magnitudes de las cifras incluidas en la información.

3.- Tabulación de datos clasificados. Esto es, arreglar en una resumida el conjunto de hechos cuantitativos, resumen basado en las clasificaciones que se hicieron. Los principales métodos para la tabulación son los que utilizan:

TARJETAS DE ESCRITURA MANUAL. (tabla A)

HOJAS DE REGISTRO. (tabla B)

TARJETAS DE PERFORACION MANUAL.

TARJETAS PERFORADAS.

PROCESAMIENTO ELECTRONICO DE DATOS.

Los tres primeros se usan para la tabulación de datos obtenidos en estudios a pequeña escala, en cuanto al uso de los dos últimos se refiere a estudios a mayor escala, pues se obtienen una serie de ventajas por la velocidad de procesamiento de los datos que es muy alta.

TABLA A .) ILUSTRACION DE UNA TARJETA DE ESCRITURA MANUAL.

CLASE.	PRIMERA CLASIFICACION NUMERO DE CLIENTES	SEGUNDA CLASIFICACION	
		HOMBRES	MUJERES
1 y menos de 20	20	19	1
21 y menos de 40	180	110	70
41 y menos de 60	60	80	10
61 y menos de 80	18	6	4



TABLA B.- ILUSTRACION DE UNA HOJA DE REGISTRO.

Una hoja de registro contendrá espacios para clasificaciones, y cada hecho se registrará conforme vaya ocurriendo mediante una marca (/) en el espacio correspondiente, después se marcaran los totales de cada clasificación.

CLASE	NUMERO DE HECHOS		VERDE	TOTAL	AZUL	TOTAL	TOTAL FINAL
	ROJA	TOTAL					
0 y menos de 20	/////	////	////////	8	///	3	20
20 y menos de 40	/////		////	4	////////	6	15
40 y menos de 60	////		////	5	////////	8	17
60 y menos de 80	/////	///	////////	9	////	4	20
80 y menos de 100	/////	///	////	4	////////	9	21
TOTAL		33		30		30	93

### 3.- PRESENTACION DE DATOS ESTADISTICOS.

Los datos estadísticos organizados para que puedan ser leídos con mayor facilidad y simplificar su análisis deberán ser presentados en una forma adecuada. En general hay tres formas de presentación de los datos estadísticos que son:

- 1.- Presentación en forma oral.
- 2.- Presentación por tablas estadísticas. (tabla C)
- 3.- Presentación por medio de gráficas estadísticas. (t.D)

De las cuales las últimas dos formas de presentación adquieren mayor importancia ya que permiten resumir un conjunto considerable de información en una forma clara e interesante.

**LAS TABLAS ESTADISTICAS.-** Las tablas estadísticas pueden ser clasificadas en dos grupos de acuerdo con los propósitos para los cuales sirven: 1) tablas generales, también llamadas tablas de referencia o repositorias y 2) tablas especiales llamadas también de resumen, de texto o tablas analíticas. Las tablas estadísticas dependiendo del uso para el cual serán utilizadas tienen un número de partes para que su empleo sea útil y práctico, en general estas son las partes principales que constituyen una tabla estadística.

- 1.- Título.- Es la descripción general del contenido de la tabla, deberá ser compacto y completo.
- 2.- Encabezado.- Es el título de la parte superior de cada columna.
- 3.-Concepto.- También llamada columna matriz, son las descripciones en las hileras de la tabla.
- 4.- Cuerpo.- Es el contenido de los datos estadísticos, es decir las cifras de la tabla.
- 5.- Nota de encabezado.- Son explicaciones de puntos relacionados con la tabla completa que no han sido incluidos en el título, en el encabezado ni en los conceptos.

6.- Nota de pie.- son usualmente colocadas abajo de los conceptos y se usan para aclarar partes que no se han explicado antes.

7.- Fuente.- Para señalar el origen de los datos.

TABLA C. ILUSTRACION DE UNA TABLA ESTADISTICA.

TITULO. NUMERO DE ALUMNOS INSCRITOS EN EL CURSO DE ESTADISTICA.<sup>1</sup>

NOTA ENCABEZADO. (EN MILES)

ENCABEZADO.	FACULTAD.	TOTAL	HOMBRES	MUJERES	
CONCEPTO.	F.C.A.	58	30	28	} CUERPO
	ECONOMIA.	30	16	14	
	CIENCIAS.	46	22	24	
	INGENIERIA.	52	50	2	

NOTA DE PIE. 1 Inscripciones realizadas en el primer semestre.

FUENTE. Hipotetica.

Para la construccion de una tabla estadística deben seguirse ciertos principios generales para que se haga más efectiva su utilización, estos son:

- 1.- Simplificación de la presentación de la tabla.
- 2.- Tratar un solo tema en la tabla.
- 3.- Un arreglo ordenado de la tabla.
- 4.- Hacer fáciles las comparaciones.
- 5.- Enfatizar cifras importantes.
- 6.- Redondear los detalles innecesarios.

LAS GRAFICAS ESTADISTICAS. La gráfica o diagrama estadística es una expresión plástica de la exposición de una información que permite presentar los datos y mostrar relaciones entre varios grupos de ellos. para dibujar una gráfica básicamente se usan los sistemas de barras o líneas.

Las partes importantes de una gráfica son:

- 1.- Título.- Es la descripción del contenido de la gráfica.
- 2.- Los diagramas.- Estas son líneas, barras, símbolos, etc. que representan datos estadísticos.
- 3.- Las escalas.- son las magnitudes de los símbolos usados

y sus clasificaciones.

4.- Fuente.- Donde se acredita la información al recopilador original.

Es posible hacer uso de muchos tipos de gráficas, pero las más comunes son: a) GRAFICA DE LINEA. b) GRAFICAS DE BARRA. c) PICTOGRAMAS.

TABLA. D. ILUSTRACION DE UNA GRAFICA.

TITULO. NUMERO DE ALUMNOS INSCRITOS EN EL CURSO DE ESTADISTICA.

ESCALAS.

ADMINISTRACION. [ HOMBRES | MUJERES ]

ECONOMIA. [ H | M ]

CIENCIAS. [ H | H ]

INGENIERIA. [ H | H ]

DIAGRAMAS.

FUENTE DE DATOS: TABLA. C.

#### 4.-ANALISIS DE DATOS ESTADISTICOS.

Después de que los datos estadísticos son recopilados, organizados y presentados en forma de fácil lectura, principalmente en tablas estadísticas, éstos están listos para ser analizados. Este análisis es generalmente por PROMEDIOS (media, mediana, moda.) y dispersión.

##### Promedio y dispersión.

Los valores incluidos en un grupo de datos usualmente varían en magnitud; algunos de ellos son pequeños y otros son grandes.

Un PROMEDIO es un valor simple, el cual es considerado como el valor más representativo o típico para un grupo de datos o números, Obviamente el valor más representativo para un un grupo de números no es el valor más pequeño ni el más grande, sino un número cuyo valor está en algún punto intermedio del grupo. Así, un promedio es frecuentemente referido como una medida de tendencia central.

Los promedios se emplean con frecuencia como mecanismos para resumir un conjunto de cantidades o números, sobre todo si es grande, a fin de describirlos datos estadísticos. También son usados para comparar un grupo de datos con otro. Los promedios estadísticos más comunes son:

- 1).- LA MEDIA ARITMETICA.
- 2).- LA MEDIANA.
- 3).- LA MODA.
- 4).- LA MEDIA GEOMETRICA.
- 5).- LA MEDIA ARMONICA.

Cada promedio tiene sus características particulares. La determinación de cuál de los diferentes tipos de promedios deberá usarse bajo diferentes circunstancias depende grandemente de las características de los promedios.

DISTRIBUCION DE FRECUENCIA (datos agrupados). Cuando un grupo grande de datos es recopilado, los valores de los mismos deberán ser puestos en orden para facilitar el análisis estadístico. Los valores pueden ser a-

rreglados primero al orden de magnitud ascendente o descendente. Los datos ordenados de esta manera constituye un arreglo. Si hay valores repetidos en el arreglo deberán ser indicados y entonces el arreglo será llamado; un arreglo de frecuencia y el número que indica las veces que un valor está repetido se llamará frecuencia. El arreglo de frecuencia puede ser constituido mediante el uso de marcas; Como sigue:

ILUSTRACION DE UN ARREGLO DE FRECUENCIAS.

VALOR.	MARKA.	FRECUENCIA.
2	////////	8
4	/////	6
6	//////////	10
8	///	3
10	/	1
TOTAL		28 VALORES.

Más aún cuando los valores son agrupados en varias clases son bases cuantitativas y el número de valores dentro de cada clase es indicado, puede obtenerse una presentación tabular más compacta de los datos.

La tabla que muestra datos agrupados cuantitativamente se llama distribución de frecuencia. Ejemplo:

Si los valores 4, 6, 2, 9, 8, 4, 8, 8. son arreglados de la siguiente manera: 2, 4, 4, 6, 8, 8, 8, 9. Una distribución de frecuencias de ellos quedaría así:

ILUSTRACION DE UNA DISTRIBUCION DE FRECUENCIA.

INTERVALO DE CLASE.	FRECUENCIA.
1-3	1
4-6	3
7-9	4
TOTAL	8 VALORES

El arreglo de una distribución de frecuencias tiene gran efecto en el cálculo de los distintos promedios, lo mismo que en otras fases del análisis estadístico. Sin embargo, no hay reglas precisas que puedan ser usadas para construir una tabla perfecta de distribución de frecuencias. Pero deberán ser examinados ciertos principios referentes al número de clases, q deberán ser usualmente entre cinco y veinte dependiendo del número de valores agrupados y el tipo de información que le pueda interesar al investigador, los límites de clases que indican la cuota o frontera de cada clase de distribución, el punto medio o centro de clase es el que más se usa para representar cada valor original, y tamaños de los intervalos de clase, que es la diferencia entre los límites de clase verdadero superior e inferior y depende también del tipo de información deseada, el número de clases y el grado de variación de los datos originales.

#### LA MEDIA ARITMETICA.

El cociente de la suma de los valores divididos por el número de valores en el conjunto de datos es llamado media aritmetica y es el tipo de promedio más usado en el análisis estadístico.

$$\text{MEDIA ARITMETICA.} = \frac{\text{SUMA DE VALORES.}}{\text{NUMERO DE VALORES.}}$$

FORMULA PARA SU CALCULO:

$$\bar{X} = \frac{\sum X}{n}$$

DONDE:

X= Al conjunto de valores, o la variable X.

n= Al número de valores en el conjunto.

$\sum$  = "la suma de"

$\bar{X}$  = La media de la variable X, llamada X barra.

La media ponderada; es llamada así a la media calculada cuando a cada uno de los valores en un conjunto de datos les es asignado una pondera-

ción de acuerdo con la importancia relativa en el grupo y es obtenida así;

1.- Multiplicar cada valor por la ponderación asignada al valor correspondiente.

2.- Sumar esos productos.

3.- Dividir la suma de los productos por la suma de la ponderación.

Sea "W" la ponderación asignada a cada valor de la variable "x" ; entonces:

$$\bar{X} = \frac{\sum WX}{\sum Wn}$$

Cálculo de la media para datos agrupados.- Para calcular la media de datos agrupados, el centro o punto medio es tomado para representar el valor de cada elemento incluido en la clase. La media para datos agrupados es obtenida; primero, multiplicando cada centro de clase por la frecuencia de la clase; segundo, sumar estos productos; y tercero, dividir la suma de los productos por la suma de las frecuencias.

Si; X= al centro de las clases individuales.

f= la frecuencia de las clases individuales.

n= a la suma de las frecuencias ó =  $\sum f$ .

Entonces:

$$\bar{X} = \frac{\sum (fx)}{n}$$

Método abreviado para calcular la media.- El método abreviado en el cálculo de la media tanto para datos agrupados como para datos no agrupados permitirá ahorrar una considerable cantidad de tiempo cuando un grupo muy grande de datos esta involucrado en este cálculo.

Este método consiste; primero, en seleccionar cualquiera de los valores arbitrariamente el cual sera la media supuesta. Segundo, obtener las desviaciones de los valores individuales respecto a la media supuesta. (X- $\bar{X}$ ). Tercero, hacer la suma algebraica de las desviaciones que sera dividida por el número de valores en los datos. Y cuarto, el cociente obte-



nido será sumado al valor arbitrariamente seleccionado o media supuesta, esta suma nos dará la media real.

Si;

A= La media supuesta o el valor seleccionado arbitrariamente.

V= La desviación de cada valor respecto a la media supuesta ó  
= X - A.

Entonces:

$$\bar{X} = A + \frac{\sum (X-A)}{n}$$

ó

$$\bar{X} = A + \frac{\sum V}{n}$$

EJEMPLOS:

Los kilómetros recorridos por 20 clientes al venir a nuestra empresa, están dados en seguida:

0.8, 1.2, 2.6, 2.8, 3.3, 3.4, 3.7, 4.0, 4.5, 5.3, 5.8, 6.1, 6.2, 6.5, 7.1, 7.3, 7.4, 7.6, 7.8, 9.2. TOTAL. 102.6 Km.

La media de los datos brutos es:

$$\text{MEDIA} = \frac{102.6}{20} = 5.13 \text{ Km.}$$

Si los datos son clasificados en cinco grupos en la siguiente tabla:

AGRUPACION DE DATOS.

Km. Recorridos.	Marcas.	Número de clientes.
0 Y menos de 2	// // // // //	2
2 y menos de 4		5
4 y menos de 6		4
6 y menos de 8		8
8 y menos de 10		1
TOTAL.		20

Esta tabla puede ser arreglada para calcular la media de estos datos así.

CALCULOS PARA EL EJEMPLO.

Km. Recorridos. (intervalo de clase)	Km. Promedio. (punto medio)	Número de clientes. (frec. de clase)	Total de Km. Recorridos.
	X	f	fX
2 y menos de 4	3	5	15
4 y menos de 6	5	4	20
6 y menos de 8	7	8	56
8 y menos de 10	9	1	9
TOTAL.		f = n = 20	fX = 102

$$\text{MEDIA} = \bar{X} = \frac{\sum(fX)}{n} = \frac{102}{20} = 5.1 \text{ Km.}$$

La diferencia entre las dos medias es muy pequeña.

Método abreviado.- Para el cálculo de la media por medio del método abreviado consideraremos como "A" a 5Km, que será la media supuesta o valor arbitrariamente seleccionado para obtener las desviaciones correspondientes.

Km. recorridos. (intervalo de clase)	Km. promedio. (punto medio)	N° de clientes. (frecuencia)	DEsviaciones respecto a la media supuesta en Km. v = X - A	Total desv. en Km. fv
	X	f	v	fv
0 y menos de 2	1	2	-4	-8
2 y menos de 4	3	5	-2	-10
4 y menos de 6	5	4	0	0
6 y menos de 8	7	8	2	16
8 y menos de 10	9	1	4	4
TOTAL.		n=20		$\sum(fv) = 2$

$$\text{MEDIA} = \bar{X} = A + \frac{\sum(fv)}{n} = 5 + \frac{2}{20} = 5 + 0.1 = 5.01 \text{ Km.}$$

### LA MEDIANA.

La mediana de un conjunto de valores es el valor del elemento central del conjunto. Para encontrar la mediana, primero se arreglarán los valores en el conjunto de acuerdo con la magnitud, es decir del "m" más pequeño al más grande o del más grande al más pequeño, segundo, localizar el valor central; es decir, el número de valores sobre la mediana es el mismo que el número de valores bajo la mediana.

Para calcular la mediana de una tabla de frecuencias serán dados los siguientes pasos:

1.- Poner una columna en la tabla para registrar las frecuencias acumuladas. La frecuencia acumulada de cada clase es la suma de las frecuencias de esta clase más las anteriores.

2.- Encontrar la clase mediana. La clase mediana es la clase que contiene a la mediana.

3.- Interpolar la mediana a partir de los valores incluidos en la clase mediana. Para la interpolación todos los tipos de datos son tratados como datos continuos. El procedimiento es como sigue:

#### EJEMPLO:

Los kilómetros recorridos por 20 clientes al venir de sus casas hasta nuestra empresa están dados en la siguiente tabla:

#### LOCALIZACION DE LA MEDIANA.

Km. recorridos. (intervalo de clase)	Número de clientes. (frecuencia)	Frecuencia acumulada.
0 y menos de 2	2	2
2 y menos de 4	5	7
<u>4 y menos de 6</u>	<u>4</u>	<u>11 clase mediana.</u>
<u>6 y menos de 8</u>	8	19
8 y menos de 10	1	20

CALCULO DE LA MEDIANA.- Hay 20 elementos (clientes) en la distribución la mediana debe ser el valor al final del decimo ( $20 \text{ sobre } 2 = 10$ ) elemento

en la distribución. La columna de frecuencias acumuladas muestra que la frecuencia acumulada la cual es justamente abajo de 10, es decir 7, por tanto, la mediana debe caer dentro de la clase la cual es justamente arriba del séptimo elemento o dentro de la clase "4 y menos de 6".

SEA:

Md = la mediana.

L = límite inferior verdadero de la clase mediana.

n = el número de la frecuencia total de los datos dados.

C = la frecuencia acumulada precisamente hasta la clase anterior a la clase mediana.

i = amplitud del intervalo de la clase mediana.

Entonces.

$$Md = L + \frac{\frac{n}{2} - C}{f} i$$

Substituyendo valores;

$$Md = 4 + \frac{\frac{20}{2} - 7}{4} (2) = 9.5 \text{ Km.}$$

#### LA MODA.

La moda o promedio típico de un conjunto de valores es el valor el cual ocurre mas frecuentemente en el conjunto. Si un valor es seleccionado al azar de un conjunto dado, un valor modal es el valor más probable a ser seleccionado. Así, la moda es considerada generalmente como el valor más típico en una serie de datos.

Cuando ha más de dos modas en un conjunto de datos, los datos son llamados bimodales o multimodales según sea el caso.

La moda cruda es el valor del punto medio de la clase modal. La clase modal es la clase con la mas alta frecuencia en la distribución.

Formula para su cálculo:

$$\text{MODA CRUDA (CMo)} = \frac{\text{Límite inferior} + \text{límite superior.}}{2}$$

Ejemplo; Tomando el ejemplo dado para la mediana, en la tabla se ve claramente que la clase modal es:

$$\frac{6 \text{ y menos de } 8}{\text{CLASE MODAL}} \quad 19$$

Por ser la clase que mayor frecuencia tiene.

Cálculo de la moda, la clase modal es "6 y menos de 8". El punto medio de la clase modal es:

$$\text{MODA CRUDA (CMo)} = \frac{6+8}{2} = 7 \text{ Km.}$$

#### DISPERSION.

Los valores incluidos en un grupo de datos generalmente varían en magnitud, algunos de los valores son pequeños y algunos son grandes. La variación de los valores es llamada dispersión y puede ser medida por muchos medios diferentes.

Una medida de dispersión es importante por dos aspectos que son: Primero, si por ejemplo hay una muy baja dispersión de los ingresos en un sector del mercado de una empresa, indicara que sus ingresos son aproximadamente iguales. Segundo, la medida de dispersión puede ser usada para suplementar un promedio, para describir un grupo de datos o para comparar un grupo con otro. Cuando la dispersión es alta, el promedio se vuelve de poca o ninguna significación.

Los tipos más comunes de dispersión que son expresados en valores absolutos son:

- 1.- El recorrido que es el más simple respecto a su concepto y

cálculo.

2.- La desviación cuartílica.

3.- La desviación media.

4.- La desviación estándar, que es la más importante puesto que es matemáticamente lógica y puede ser usada en cálculos adicionales.

1.- El recorrido de un grupo de datos o valores es la diferencia entre los valores más altos y más bajos, la fórmula para su cálculo será:

$$R = XN - Xi$$

Donde:

R = recorrido.

XN = el valor más alto.

Xi = el valor más bajo.

2.- DESVIACION CUARTILICA.- la desviación cuartílica de un grupo de datos está basada solamente en dos valores que son la primera y la tercera cuartillas del grupo. Para encontrar las cuartillas primero dividimos los elementos del grupo en cuatro partes iguales de acuerdo a sus valores

La primera cuartilla ( $Q_1$ ) es el punto sobre la escala de los elementos bajo la cual se encuentra la cuarta parte de éstos.

La segunda cuartilla ( $Q_2$ ) es el punto sobre o bajo del cual hay la mitad de los elementos y corresponde a la mediana.

La tercera cuartilla ( $Q_3$ ) es el punto abajo del cual hay las tres cuartas partes de los elementos.

La diferencia entre la primera y la tercera cuartilla es llamado el recorrido intercuartílico. Cuando esta diferencia es dividida por dos el cociente es la desviación cuartílica (Q.D.) o semirecorrido intercuartílico.

$$Q.D. = \frac{Q_3 - Q_1}{2}$$

Las cuartillas para datos agrupados son obtenidos por medio de las fórmulas:

$$Q_1 = L + \frac{\frac{n}{4} - C}{f} (i)$$

$$Q_3 = L + \frac{\frac{3n}{4} - C}{f} (i)$$

Donde:

L = al límite inferior verdadero de la clase aplicable a la cuartilla; es decir, la clase de la primera cuartilla, o de la tercera cuartilla, cualquiera que se aplique.

n = al número total de las frecuencias de los datos dados.

C = La frecuencia acumulada justamente hasta la clase anterior (o más pequeña) a la clase de la cuartilla aplicable.

f = la frecuencia de la clase de la cuartilla aplicable.

i = la amplitud del intervalo de la clase aplicable.

3.- LA DESVIACION MEDIA. El recorrido y la desviación cuartílica son medidas posicionales de dispersión. Están basadas en las posiciones de ciertos elementos en una distribución. La desviación media y la desviación estándar están basadas en todos los elementos y están diseñadas para medir la dispersión alrededor de un promedio.

La desviación media es la media aritmética de las desviaciones de los valores individuales con respecto al promedio de los datos dados. El promedio que se usa para calcular la desviación media puede ser la media aritmética o la mediana.

Al calcular la desviación media, los valores, absolutos de las desviaciones son usados; es decir se ignoran los signos negativos o positivos de los valores.

Método para calcular la desviación media para datos agrupados.

Procedimiento:

1.- Encontrar la media aritmética  $\bar{X}$ .

2.- Encontrar la desviación de cada valor medio o punto medio de cada clase con respecto a la media aritmética. Se usarán los valores absolutos,  $|x| = |X - \bar{X}|$ , y multiplicar la desviación en cada clase por la frecuencia de la clase.  $(f|x|)$ .

3.- Encontrar la suma de los valores absolutos de las desviaciones.  $\sum (f|x|)$ .

4.- Encontrar la desviación media dividiendo la suma por el número de valores  $(n \text{ ó } \sum f)$  en los datos.

$$\text{DESVIACION MEDIA = A.D.} = \frac{\sum (f|x|)}{n}$$

4.- DESVIACION ESTANDAR ( $s$ ) Y VARIANZA ( $s^2$ ). La desviación estándar se calcula de la misma manera que la desviación media, excepto que los signos positivos y negativos de las desviaciones individuales son tomados en cuenta. Los métodos para calcular la desviación estándar de datos agrupados son métodos: a) básicos y b) abreviado.

Procedimiento para el método básico:

1.- Encontrar la media aritmética de los datos dados ( $\bar{X}$ ).

2.- Encontrar la desviación de los puntos medios de cada clase con respecto a la media aritmética,  $(X - \bar{X}) = x$ .

3.- Encontrar el total de las desviaciones al cuadrado para cada clase  $fx^2$ . Esta total puede ser obtenido de dos maneras:

a) Primero elevar al cuadrado la desviación  $X^2$ , después, multiplicar la desviación por la frecuencia de la clase  $f(X^2)$ .

b) primero, multiplicar la desviación por la frecuencia de la clase  $fx$ . después, multiplicar el producto de la desviación,  $(fx)x = fx^2$ . La segunda forma es usualmente más aplicada.

4.- Encontrar la suma de la desviación al cuadrado,  $\sum fx^2$ .

5.- Encontrar la varianza ( $s^2$ ) dividiendo la suma por el número



número de valores ( n ó  $\sum f$  ) de los datos.

$$S^2 = \frac{\sum fx^2}{n}$$

6.- Extraer la raíz cuadrada de la varianza para encontrar la d  
desviación estándar (s).

$$S = \sqrt{\frac{\sum fx^2}{n}}$$

La desviación estándar para datos agrupados se calcula como sigue:

Supongamos que hemos determinado que los ingresos de 20 empleados de  
una compañía están representados en la tabla a contin acción; calcular la  
desviación estándar de sus ingresos.

SUELDO SEMANAL DE 20 EMPLEADOS.

Sueldo que perciben. (cientos de \$)	sueldo promedio. X	número de empleado f	$x-X-\bar{X}$	fx	$fx^2$
0 y menos de 2	1	2	-4.1	-8.2	33.62
2 y menos de 4	3	5	-2.1	-10.5	22.05
4 y menos de 6	5	4	-0.1	-0.4	0.04
6 y menos de 8	7	8	1.9	15.2	22.88
8 y menos de 10	9	1	3.9	3.9	15.21
TOTAL		20		0	99.80

$\bar{X} = 5.1$  (cientos de pesos).

$$S^2 = \frac{\sum fx^2}{n} = \frac{99.80}{20} = 4.99$$

$$S = \sqrt{4.99} = 2.23$$

Este dato obtenido (2.23) es la desviación estándar de los ingresos de los 20 empleados con respecto a la media de los mismos.

Lo cual nos da idea de :

1.- El promedio de ingreso de los 20 empleados es de \$510.00 semanales.

2.- Ocho de los empleados tienen ingresos semanales entre \$300.00 y \$800.00. Que es la mayor frecuencia.

3.- La desviación estándar respecto de la media es relativamente alta es decir, que es muy grande la dispersión de ingresos.

#### 5.- INTERPRETACION DE LOS RESULTADOS DEL ANALISIS.

Después de que el análisis de los datos estadísticos está completo, los resultados del análisis deben ser interpretados. La interpretación correcta guiará a una conclusión válida del estudio y así puede ayudar a uno a tomar decisiones. Por tanto el análisis estadístico corresponde a la persona que realiza el estudio.

II.) TEORIA DEL MUESTREO.

## II. TEORIA DEL MUESTREO.

La teoría del muestreo es un estudio de las relaciones existentes entre una población y muestras extraídas de la misma. Tiene gran interés en aspectos de la estadística y sobre todo, en cualquier tipo de investigación. Por ejemplo, de una población grande aunque finita o infinita de la que resultaría impráctico y a veces imposible, estimar cantidades tales como promedios o desviaciones.

La teoría del muestreo es también útil para determinar las diferencias que se pueden observar entre dos o más muestras y si éstas diferencias son debidas a la aleatoriedad de las mismas o si por lo contrario son realmente significativas.

Se conoce como inferencia estadística al estudio de las diferencias que se realiza sobre una población mediante muestras consistentes extraídas de la misma, junto sobre la exactitud de tales inferencias aplicando la probabilidad.

Un concepto fundamental del muestreo es el que se conoce como distribución normal, que aparece en situaciones donde una variable medida tiene una tendencia a tomar un determinado valor fijo con relación a la media, las variaciones son probabilísticas.

La curva normal adquiere la apariencia de una campana que se adelgaza en forma gradual hacia ambos lados de la parte central, -ésto se debe a que las variaciones al azar pueden ocurrir en cualquier dirección de la media y es más factibles que estas variaciones al azar sean pequeñas que grandes. Muchas medidas económicas, biológicas, y físicas tienden a caer dentro de las distribuciones normales. Así mismo, cuando una distribución no es conocida con exactitud pero se sabe que está afectada por variaciones al azar, se supone que la distribución es normal. Las muestras generalmente son seleccionadas al azar y los valores en una muestra están sujetos a variaciones al azar, por tanto, la distribución normal es básica en el muestreo.

La configuración de la curva normal depende del valor de las observaciones individuales diseminadas o dispersas ( $x$ ) en torno a la media ( $\bar{X}$ ) que por supuesto, siendo el punto de mayor frecuencia es el punto medio de la curva. Un análisis matemático revelará los siguientes puntos:

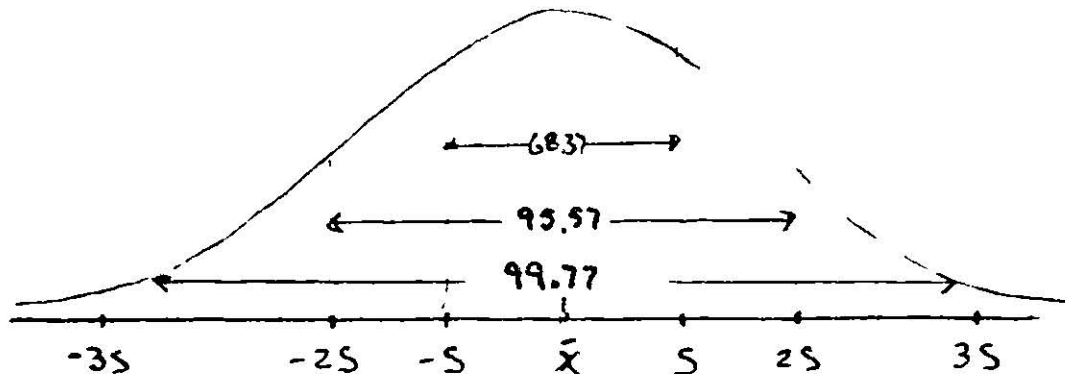
a).- Dentro de la escala comúnmente se utilizará la desviación estándar ( $s$ ) y para el lugar convenido entre  $\bar{X} \pm s$  está determinada una probabilidad del 68.3% del total de frecuencias

b).- Dentro de las  $2s$  de la media ocurren el 95% de las frecuencias, esto es ;  $\bar{X} \pm 2s = 95.5\%$ .

c).- En la escala de variaciones  $\bar{X} \pm 3s$  caen el 99.7% de las frecuencias.

d).- Puesto que la área total de la curva es igual a una probabilidad ( $p$ ) de 1.000 la probabilidad de que una muestra sacada al azar caiga fuera de la escala de variación de  $\bar{X} \pm 2s$  sera de  $1.000 - 0.955 = 0.045$  (ó 4.5%).

GRAFICA DE UNA CURVA NORMAL.



Cuando la selección de una muestra ha sido apropiada aún queda el problema de estudiar el universo o población. Una muestra puede proporcionar información acerca de la población pero no acerca de elementos específicos, las inferencias basadas en muestras se refieren sólo a grupos de elementos, más específicamente a valores de la población.

Los valores de las poblaciones y de la muestra deberán diferenciarse para que al aplicar las formulas correspondientes no haya confusión.

SÍMBOLOS PARA ESTADÍSTICOS Y PARÁMETROS CORRESPONDIENTES.

MEDIDA.	SÍMBOLO PARA EL ESTADÍSTICO. (MUESTRA)	SÍMBOLO PARA EL PARÁMETRO. (POBLACION)
MEDIA.	$\bar{X}$	$\mu$
DESVIACION ESTANDAR.	s	$\sigma$
NÚMERO DE ELEMENTOS.	n	N
PROPORCIÓN.	p	P

Un valor de la población es una expresión numérica que sintetiza los valores de una o varias características de los N elementos de una población completa, es una medida resumida de una cualidad de la variable o variables de la población definida.<sup>1</sup>

L. KISH, LESLIE. MUESTREO DE ENCUESTAS.  
MEXICO 1972.

Generalmente las inferencias se obtienen de la media de la muestra, pero lo que se expone en seguida tambien se aplica a la mediana, desviación estándar y otras medidas de la muestra. Así cuando de una población son extraídas un número de muestras al azar y se calcula la media para cada muestra, estas medidas tienden a distribuirse normalmente y la media de esta distribución tiende a la media del universo o población. Por ejemplo se extrae una muestra al azar de los ingresos de los residentes en cierta comunidad, y se obtiene el ingreso medio correspondiente a esta muestra, pero si en seguida se extrae otra muestra de esta misma comunidad, el ingreso medio de esta última muestra probablemente diferirá de la media de la primera y así sucederá con el ingreso medio que se calcule de las demás muestras que se extraigan de esa comunidad, pero si las diferentes medias son arregladas en una distribución de frecuencias, esta distribución tenderá a ser normal y la media de esta distribución tenderá a ser igual a la media de la comunidad.

En el muestreo estadístico se trabajan tres tipos de distribución de frecuencias.

1.- Distribución de frecuencias en el universo. Aquí se trabaja con la población total que comprende todas las medidas. Puesto que la distribución de frecuencias de la población no es conocida, se aproxima a ella mediante la distribución de muestras al azar.

2.- Distribución de muestras al azar. La cual es una representación en pequeña escala de la población, contando aproximadamente con las mismas características que tiene la población.

3.- Distribución de medias de la muestra. Esta distribución tiende a ser normal, aun cuando las otras no estan normalmente distribuidas, y tiene una dispersión menor que las otras dos.

La desviación estándar indica la dispersión de los valores individuales en una distribución en torno alas medias. La desviaciones estándar de las distribuciones de medias de la muestra es diferente y para distri-

guirla, esta desviación estándar de las medias de las muestras se le llama EL ERROR ESTANDAR ( $\sigma$ ).

EL ERROR ESTANDAR. Se usa para estimar las probabilidades vinculadas con las inferencias obtenidas respecto a la población. El método para determinar el error estándar es diferente si la muestra consiste en a).- Datos enumerados que son expresados en forma de proporciones. y b).- Datos medidos o continuos que se refieren al valor de una variable.

a).- Los datos enumerados o discretos se refieren a la presencia o ausencia de alguna característica, "son típicos de aquellos que dan origen a la distribución binomial. Estas proporciones se expresan como  $p + q = 1$ . Donde "p" es la proporción portadora del atributo y "q" es la proporción sin atributo. La fórmula para calcular el error estándar o error estándar de la proporción :

$$\sigma_p = \sqrt{\frac{pq}{n}}$$

Donde:  $\sigma_p$ .- es el error estándar de la proporción.

p .- es la muestra portadora de alguna característica.

q .- es igual a;  $1-p$

n .- es el número de elementos de la muestra.

b).- Los datos medidos o continuos dan origen a una distribución continua por frecuencias. La fórmula para calcular el error estándar de este tipo de datos es:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Donde;  $\sigma_{\bar{x}}$ .- es el error estándar de la distribución de las medias de la muestra.

$\sigma$ .- es la desviación estándar de todos los conceptos comprendidos en la población.



n.- es el tamaño de la muestra.

Debido a que la desviación estándar del universo se desconoce ( $\sigma$ ) generalmente se usa la desviación estándar de la muestra ( $s$ ) como una aproximación a ella.

El error muestral o error del muestreo es la diferencia entre el resultado obtenido de una muestra y el resultado, el cual deberá ser obtenido de la población y es medido por el error estándar del estadístico en terminos de probabilidad bajo la curva normal. El resultado de la media indica la precisión de la estimación de la población basada en el estudio de la muestra.

Procedimiento de obtención de las medias de las muestras.- En general, un valor de la muestra, o estadístico, es una estimación que se calcula a partir de los "n" elementos de la muestra. La media de los elementos de la muestra es:

$$\bar{X} = \sum X_i / n$$

Donde :  $X_i$  es el valor observado del elemento i-ésimo de la muestra.  $X_i$  es la variable aleatoria, que depende del diseño de la muestra y de la combinación de los elementos que se seleccionarán.

La media de la muestra, como los otros estadísticos, están sujetos a errores de muestreo que ocurren cuando solamente una parte de la población total se observa en la muestra. Y errores no de muestreo se deben a las imperfecciones de los procedimientos de observación.

Ahora bien, el conocimiento de la media muestral no tiene un valor práctico en sí. Su valor consiste en lo que podemos saber, a partir de ella, de la población de la cual se seleccionó.

Se sabe que la media de la distribución de muestreo es el valor esperado o la esperanza matemática de la media, entonces, el valor medio de la distribución de las medias puede o no ser igual al valor medio de la población. La esperanza matemática se formuló así:

$$E(\bar{x}) = \sum p \bar{x}$$

Donde:

$E$ .- es la media de la distribución de muestreo de la media  $\bar{X}$

$p$ .- es la probabilidad de materialización de cada media.

$\bar{X}$ .- es cada una de las medias posibles encontradas de cada muestra.

La diferencia entre el valor medio  $E(\bar{X})$  y el valor de la población  $\bar{X}$  se llama sesgo de muestreo.

El muestreo desempeña un papel importante en el diseño de investigaciones enfocadas a poblaciones humanas y la aplicación más práctica se encuentra en la investigación por medio de encuestas. La selección de muestras válidas de tamaño tal que equilibre factores de costo y eficiencia y cuyos resultados puedan ser evaluados y determinar con precisión y grado de confianza serán los problemas que en la investigación por medio de encuestas tratará de solucionar el estudio del muestreo.

El diseño de encuestas tiene dos aspectos básicos y son:

a.- Un proceso de selección.- Que consiste en reglas y operaciones mediante las cuales se tratará de incluir en muestra aquellos elementos de la población que proporcionarán la información deseada.

b.- Un proceso de estimación.- En el cual se calculan valores de la muestra, llamados estadísticos de la muestra, como son; la media el error estándar, etc. Estos son estimaciones de valores de las características de la población o sean, expresiones numéricas que sintetizan los valores de estas características de la población.

En la investigación por encuestas, el diseño completo de encuestas incluye otros aspectos importantes que pueden ser llamados conjuntamente objetivos de la encuesta:

1.- La definición de las variables de la encuesta que debe especificar la naturaleza de las clasificaciones y de las unidades para expresarlas. Debe determinarse también el alcance y contenido de la población de encuesta.

2.- Los métodos de observación, que incluye tanto la recolección como el procesamiento de los datos dan significado operacional a las variables de la encuesta y determinan la naturaleza de los datos de la encuesta.

3.- Los métodos de análisis estadístico y sustanciales reducen los datos de la encuesta a resultados que puedan comprenderse y utilizarse.

4.- La utilización de los resultados de la encuesta puede, en ocasiones, tomar la forma de decisiones concretas que se basen en los resultados y en información proveniente de otras fuentes, que deben relacionarse con el contenido de la encuesta.

5.- La precisión que se desea que tengan los resultados de la encuesta puede enunciarse claramente en las muestras diseñadas para llegar a una decisión estadística especificada. Es más frecuente que las pretensiones sean muchas y se enuncien vagamente, a pesar de lo cual el investigador puede encontrar límites amplios de la precisión que se quiere obtener.

Los objetivos de la encuesta deben determinar el tamaño de la muestra, así como su selección y grado de precisión.

1.- DETERMINACION DEL TAMAÑO ADECUADO DE LA MUESTRA ( n ).

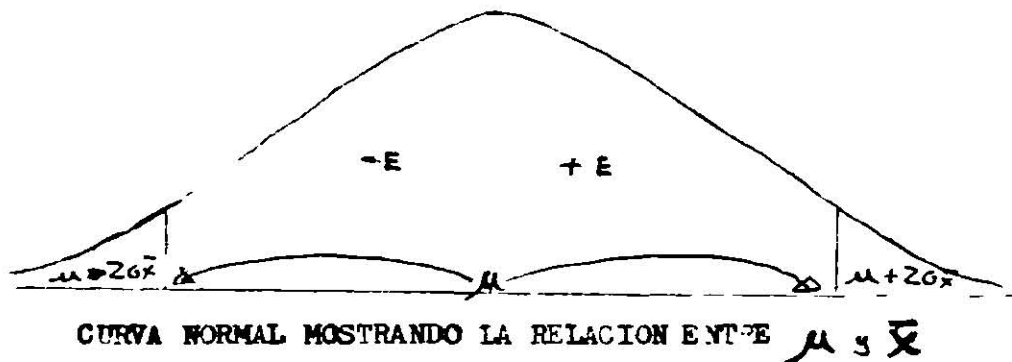
Un punto importante del diseño en la teoría del muestro es la determinación del tamaño de la muestra, punto que constituye un problema que en la práctica es indispensable sea adecuado. Si el tamaño de la muestra es demasiado grande la información que se obtenga de ella será más costosa pero el resultado es más exacto, lo contrario sucede si el tamaño de la muestra es pequeño.

Existen dos casos para determinar el tamaño de la muestra:

- a).- Tamaño de muestra para estimar una media poblacional.
- b).- Tamaño de la muestra para estimar una proporción poblacional.

Procedimiento para determinar el tamaño adecuado de la muestra para estimar una media poblacional.

La relación entre la media de la población y la distribución en el muestro puede ser comprendida bajo la curva normal<sup>1</sup>



La curva normal muestra el intervalo de confianza;

$$\mu \pm 2\sigma \bar{x} = \mu \pm E, \text{ y, } E = 2\sigma \bar{x},$$

<sup>1</sup> STEPHEN P. SHAO.- ESTADÍSTICA PARA ECONOMISTAS Y ADMINISTRADORES DE EMPRESAS. cap. 13 PAG. 367.

Donde :

$E$  = el error muestral, o la diferencia entre una media muestral  $\bar{X}$  y la media poblacional  $\mu$ ; es decir,  $E = \bar{X} - \mu$ .

También muestra que el recorrido del intervalo de confianza es  $2(E)$ .

Cuando  $\sigma_x$  es reemplazada por  $\frac{G}{\sqrt{n}}$ , la fórmula, la ecuación de arriba se vuelve:

$$E = z \frac{G}{\sqrt{n}} \quad \text{y} \quad \sqrt{n} = \frac{zG}{E}$$

Despejando "n" de ambos lados nos queda:

$$n = \left( \frac{zG}{E} \right)^2 \quad \text{o} \quad n = \frac{z^2 G^2}{E^2}$$

Que es la fórmula para obtener el tamaño de la muestra ( n ).

Donde:

$E$  : es el error muestral máximo que esté dispuesto a aceptar cada investigador en un estudio muestral.

$z$  : es establecida mediante el nivel de confianza, según la siguiente tabla:

COEFICIENTES DE CONFIANZA COMUNES Y SUS VALORES DE  $z$

coeficientes de confianza	50%	68.27%	90%	95%	95.45%	99%	99.73%
$z$	0.6745	1.00	1.645	1.96	2.00	2.58	3.00

$G$ : es el valor de la desviación estándar de la población. puede ser real o estimado de experiencias pasadas, o estimado mediante  $\hat{s}$ , el cual es la desviación estándar, ya sea de una muestra previa o un estudio piloto.

Después de que se toma una muestra del tamaño calculado, el resultado de la muestra deberá ser evaluado. Esto puede ser hecho encontrando el error estándar de la media  $s_{\bar{x}}$ , de acuerdo con la desviación estándar de la muestra  $\hat{s}$ . Si el producto de  $z(s_{\bar{x}})$  es menor que el error de muestreo especificado, la estimación de la muestra es considerada satisfactoria. Si el producto es mayor, el tamaño de la muestra deberá ser revisado e incrementado.

EJEMPLO:

El gerente de una estación de radio quiere saber la proporción del auditorio que prefiere un programa en particular. El especifica que:

a).- El error muestral no deberá ser más de 2% de la verdadera proporción.

b).- El nivel de intervalo de confianza deberá ser del 95%.

c).- La proporción de la gente a quienes les gusta el programa es de aproximadamente del 60%.

¿Cuál es el tamaño máximo de la muestra de acuerdo con sus especificaciones?

¿Cuál es el tamaño máximo de la muestra si solamente especifica a y b. ?

SOLUCION.

La respuesta se calcula así:

1.- el intervalo de confianza es:

$$p \pm 2\% \text{ por tanto } E = 5\% = 0.05$$

2.- para el intervalo de confianza de 95%,  $z = 1.96$

3.- la estimación de la población es:

$$p = 60\% = 0.60 \quad \text{y} \quad q = 1-p = 1-0.60 = 0.40$$

Substituyendo los valores en la fórmula:

$$n = \frac{z^2 pq}{E^2} \quad \text{Donde; } pq = 0^2$$

$$n = \frac{(1.96)^2 (0.60) (0.40)}{(0.02)^2}$$

$$n = \frac{3.8416 (.24)}{.0004}$$

$$n = \frac{.921984}{.0004}$$

$$n = 2,305$$

El tamaño de la muestra para la primera pregunta es de 2,305 .  
El tamaño máximo de la muestra si solamente se especifican el error muestral máximo y el nivel del intervalo de confianza será.

- a).-  $p = 50\% = 0.5$
- b).-  $q = 1 - p = 1 - 0.5 = 0.50$
- c).-  $z = 1.96$
- d).-  $E = 2\% = 0.02$

Substituyendo en la fórmula encontramos:

$$n = \frac{3.8416 (.50)(.50)}{(0.02)^2}$$

$$n = \frac{.860400}{.0004}$$

$$n = 2,151 \text{ elementos.}$$

## 3. → TECNICAS BASICAS DE MUESTREO.

### TERMINOLOGIA:

Los nuevos términos, los cuales son los más frecuentemente usados en estas técnicas estadísticas son:

#### ESTADISTICO:

Un estadístico es una medida usada para describir alguna característica de una muestra, tal como una media aritmética, una mediana, o una desviación estándar de una muestra.

#### PARAMETRO:

Un parametro es una medida usada para describir alguna característica de una población, tal como una media aritmética, una mediana, o una desviación estándar de una población.

### METODOS DE SELECCION DE MUESTRAS.

Una muestra debe ser representativa ya que será usada para estimar las características de la población. Los métodos para seleccionar una muestra representativa son numerosos, dependiendo del tiempo, dinero y habilidad disponibles para tomar una muestra y la naturaleza de los elementos individuales de la población.

Los métodos de selección de muestras pueden ser clasificados de acuerdo a:

1.- El número de muestras tomadas de una población dada para un estudio.

2.- La manera usada para seleccionar los elementos necesarios en la muestra.

Métodos de muestreo clasificados de acuerdo con el número de muestras tomadas de una población.

Muestras básicas de este tipo de muestreo.

1).- Simple.

2).- Doble.

3).- Múltiple.



### MUESTREO SIMPLE.

Este tipo de muestreo toma solamente una muestra de la población dada, para el propósito de inferencia estadística. Puesto que solamente es una la muestra tomada el tamaño de la muestra debe ser lo suficientemente grande para extraer una conclusión. Una muestra grande muchas veces cuesta demasiado dinero y tiempo.

### MUESTREO DOBLE.

Bajo este tipo de muestreo, cuando el resultado del estudio de la primera muestra no es decisivo, una segunda muestra es extraída de la población. Las muestras son combinadas para analizar los resultados. Este método permite a una persona principiar con una muestra relativamente pequeña para ahorrar costos y tiempo. Si la primera muestra arroja resultados definitivos, la segunda ya no es necesaria.

### MUESTREO MULTIPLE.

El procedimiento bajo este método es similar al muestreo doble, excepto que el número de muestras sucesivas requeridas para llegar a una decisión es más de dos.

### MÉTODOS DE SELECCION DE LA MUESTRA.

Métodos de muestreo clasificados de acuerdo con las maneras usadas en seleccionar los elementos de una muestra:

- a).- Muestreo de juicio.
- b).- Muestreo aleatorio.

#### a).- MUESTREO DE JUICIO.

Una muestra es llamada de juicio cuando sus elementos son seleccionados mediante juicio personal, la persona que selecciona los elementos de la muestra, usualmente es un experto en la materia. Una muestra de juicio es también llamada una muestra no probabilística, puesto que este método está basado en los puntos subjetivos de una persona y la teoría de la probabilidad no puede ser empleada para medir el error del muestreo. Las principales ventajas de una muestra de juicio son la facilidad de ob-

tenerlas y el costo es usualmente bajo.

b).- MUESTREO ALEATORIO.

Una muestra es extraída al azar cuando la manera de seleccionar es tal, que cada elemento de la población tiene igual oportunidad de ser seleccionado. Una muestra aleatoria es también llamada una muestra probabilística, puesto que cada elemento tiene una probabilidad conocida. Los muestreos probabilísticos son frecuentemente preferidos por los estadísticos porque la selección de la muestra es objetiva y el error muestral puede ser medido en términos de probabilidad bajo la curva normal. Los tipos más conocidos de muestreo aleatorio son:

MUESTREO ALEATORIO SIMPLE.

MUESTREO SISTEMÁTICO.

MUESTREO ESTRATIFICADO.

MUESTREO DE CONGLOMERADOS.

MUESTREO ALEATORIO SIMPLE.

Una muestra aleatoria simple es seleccionada de tal manera que cada elemento (muestra) posible del mismo tamaño tiene igual probabilidad de ser seleccionada de la población.

Por ejemplo; en una población de tres elementos (A.B.C.). Es posible extraer tres muestras de dos elementos cada una (AB. BC. AC.). Si cada una de las tres muestras tienen la misma probabilidad de ser seleccionada ( $1/3$ ), una muestra seleccionada, es una muestra aleatoria simple.

Una técnica para seleccionar una muestra aleatoria simple es asignar números a cada uno de los miembros de la población, escritos estos números en pequeños papeles, se introducen en una urna y después se extraen los números de la urna, teniendo cuidado de mezclarlos bien antes de cada extracción, este muestreo puede ser con o sin reemplazamiento, con reemplazamiento cuando se vuelven a meter a la urna después de cada extracción y sin reemplazamiento cuando se elimina el número extraído.

Esto puede ser substituidos por el empleo de una tabla de números aleatorios.

Obtener una muestra aleatoria simple no es una tarea fácil o práctica bajo muchas circunstancias o puede ser una tarea tardada o costosa y algunas veces imposible. Por ejemplo, si deseamos tomar una muestra aleatoria simple de una población finita pero grande de un millón de familias, aunque es posible, no es una tarea fácil asignar un número a cada una de las familias y luego extraer una muestra al azar de los números. Cuando una población es finita o más bien infinita, es obvio que la tarea de numerar cada uno de los elementos de la población es imposible. Por tanto, son necesarias ciertas modificaciones del muestreo aleatorio simple.

Los tipos más comunes de muestreo aleatorio modificado son:

MUESTREO SISTEMÁTICO.

MUESTREO ESTRATIFICADO.

MUESTREO DE CONGLOMERADOS.

MUESTREO SISTEMÁTICO.

Una muestra sistemática es obtenida cuando los elementos son seleccionados en una manera ordenada. La manera de selección depende del número de elementos incluidos en la población y el tamaño de la muestra. El número de elementos en la población es, primero, dividido por el número deseado en la muestra. El cociente indicará cuál es el elemento en la población que va a ser seleccionado.

El primer elemento de la muestra es seleccionado al azar. Por lo tanto, una muestra sistemática puede dar la misma precisión de estimación acerca de la población, que una muestra aleatoria simple cuando los elementos en la población están ordenados al azar. Por ejemplo, suponga que una tienda tiene un registro de 10,000 clientes ordenados en orden alfabético. Si se desea una muestra de las compras de 200 clientes, estos pueden seleccionarse del archivo cada quincuagésimo cliente ( $10,000 / 200 = 50$ ), puesto que el orden del archivo no tiene nada que ver con las compras de los

clientes, el monto de la primera compra puede ser seleccionada al azar del primer grupo de 50 clientes del archivo. Si el primer elemento es el dieciochoavo en la población, el segundo será el 68 ( = 18 + 50-), el tercero será el 118, el cuarto será el 168, y así sucesivamente.

#### MUESTREO ESTRATIFICADO.

Definición.- El muestreo estratificado es aquel en que la población se divide en cierto número de subpoblaciones, cada una de las cuales se muestrea independientemente.<sup>1</sup>

Para obtener una muestra aleatoria estratificada primero se divide la población en estratos, que son más homogéneos que la población como un todo. Los elementos de muestra son entonces seleccionados al azar o por un método sistemático en cada estrato. Las estimaciones de la población, basadas en las muestras estratificadas usualmente tienen mayor precisión (o menor error muestral) que si la población entera fuera muestreada mediante muestreo aleatorio simple.

El número de elementos seleccionado de cada estrato puede ser proporcional o desproporcional al tamaño del estrato en relación con la población. Bajo el método proporcional por ejemplo, si el tamaño del estrato A es, 40% de la población, el mismo porcentaje es seleccionado para elegir el número de elementos del estrato. Por lo tanto, si el tamaño de la muestra es de 200 elementos, 40% del tamaño de muestra, o sea 80 elementos, van a ser seleccionados del estrato A. Cuando la selección es desproporcional, es relativamente difícil ponderar adecuadamente los resultados de los estratos individuales.

#### MUESTREO DE CONGLOMERADOS.

Para obtener una muestra de conglomerados, primeramente debe dividirse la población en grupos que sean convenientes para el muestreo.

En seguida, se seleccionará una porción de los grupos al azar o por medio de un método sistemático. Finalmente se debe tomar todos los elemen-

<sup>1</sup> Neter, John. fundamentos de estadística, aplicada a los negocios y la economía. México, 1972. p. 508

tos o parte de ellos al azar o por un método, sistemático de los grupos seleccionados para obtener una muestra. Bajo este método, aunque no todos los grupos son muestreados, cada grupo tiene igual oportunidad de ser seleccionados, por tanto es una muestra aleatoria también.

Por ejemplo, cuando no es práctico tomar una muestra simple de una población de un millón de familias en una ciudad, la ciudad puede ser dividida en pequeñas áreas de acuerdo con un mapa de la misma. Un número es asignado a cada área. Una parte de las áreas es seleccionada al azar o por un método sistemático de los números que representan las áreas. Las familias dentro de las áreas seleccionadas son entrevistadas, ya sea todas las familias del área seleccionada o parte de ellas, seleccionadas al azar o por un método sistemático. Esta muestra de conglomerados es también llamada muestra de áreas.

Una muestra de conglomerados, usualmente produce un mayor error muestral (por tanto, da menor precisión en las estimaciones sobre la población) que una muestra aleatoria simple del mismo tamaño. Los elementos individuales dentro de cada "conglomerado" tienden usualmente a ser iguales. Por ejemplo, gente de una misma condición económica suele vivir en la misma zona, no todas las áreas son muestreadas en un muestreo de conglomerados. La variación entre los elementos obtenidos de las áreas seleccionadas es, por lo tanto, frecuentemente mayor que la obtenida si la población entera es muestreada mediante muestreo aleatorio simple. Esta debilidad puede ser reducida cuando se incrementa el tamaño de la muestra de área. El incremento del tamaño de la muestra puede fácilmente ser cubierto puesto que los entrevistadores no tienen que caminar demasiado lejos para entrevistar a más familias. Por tanto, una muestra grande de área puede ser obtenida dentro de un corto período de tiempo y a bajo costo.

Por otra parte una muestra de conglomerados puede producir la misma precisión en la estimación que una muestra aleatoria simple, si la varia-

ción de los elementos individuales dentro de cada conglomerado es tan grande como dentro de la población.

MUESTREO ALEATORIO.

EJEMPLO<sup>1</sup>

(A).- Indicar cómo seleccionar 30 muestras al azar de 4 estudiantes, cada una (con remplazamiento) de la siguiente tabla, mediante números aleatorios.

TABLA DE 100 ESTUDIANTES DE LA UNAM.

ALTURA. (pulgadas)	NUMERO DE ESTUDIANTES.
60 - 62	5
63 - 65	18
66 - 68	42
69 - 71	27
72 - 74	8
TOTAL.	100

(B).- Hallar la media y la desviación típica de la distribución muestral de medias de (A).

SOLUCION:

(A).- Utilizar dos dígitos para numerar cada uno de los 100 estudiantes: 00, 01, 02, 03, .....99. Así los 5 estudiantes con alturas 60 - 62 pulgadas son numerados del 00 al 04, los 18 estudiantes con alturas 63 - 65 pulgadas se numeran 05 al 22, etc. El número de cada estudiante se llama número muestral.

TABLA DE NUMEROS MUESTRALES.

ALTURA. (pulgadas)	FRECUENCIA.	NUMERO MUESTRAL.
60 - 62	5	00 - 04
63 - 65	18	05 - 22
66 - 68	42	23 - 64
69 - 71	27	65 - 91
72 - 74	8	92 - 99

<sup>1</sup> Murray R. Spiegel. Estadística (Schaum,s). Colombia, 1971. p. 147.

Ahora se extraen los números muestrales de la tabla de números aleatorios.<sup>1</sup> En la primera línea se encuentra la sucesión 51, 77, 27, 46, 40, etc. que se toman como números del muestreo aleatorio, cada uno de los cuales da la altura de un estudiante determinado. Así el 51 corresponde a un estudiante que tenga una altura entre 66 - 68 pulgadas, que se toma como 67 (la marca de clase). Análogamente, 77, 27, 46 dan las alturas 70, 67, 67 pulgadas, respectivamente.

Mediante este proceso se obtiene la siguiente tabla. Que muestrales números muestrales extraídos, las correspondientes alturas y la altura media para cada una de las 30 muestras.

NÚMERO MUESTRAL EXTRAÍDO.	ALTURAS CORRESPONDIENTES.	ALTURA MEDIA.
1. 51,77,27,46.	67,70,67,67.	67.75
2. 40,42,33,12.	67,67,67,64.	66.25
3. 90,44,46,62.	70,67,67,67.	67.75
4. 16,28,98,93.	64,67,73,73.	69.25
5. 58,20,41,96.	67,64,67,70.	67.00
6. 17,64,08,70.	64,67,64,70.	66.25
7. 56,74,06,32.	67,67,61,67.	65.50
8. 34,21,83,58.	67,70,70,67.	68.50
9. 70,65,6,21.	70,70,70,64.	68.50
10. 96,02,13,87.	73,61,64,70.	67.00
11. 76,10,51,08.	70,64,67,64.	66.25
12. 63,97,45,39.	67,73,67,67.	68.50
13. 05,81,45,93.	64,70,67,73.	68.50
14. 96,01,73,52.	73,61,70,67.	67.75
15. 07,82,54,24.	64,70,67,67.	67.00
16. 11,64,55,58.	64,67,67,67.	66.25
17. 70,56,97,43.	70,67,73,67.	69.25
18. 74,28,93,50.	70,67,73,67.	69.25
19. 79,42,71,30.	70,67,70,67.	68.50
20. 58,60,21,33.	67,67,64,67.	66.25
21. 75,79,74,54.	70,70,70,67.	69.25
22. 06,31,04,18.	64,67,61,64.	64.00
23. 67,07,12,97.	70,64,64,73.	67.75
24. 31,71,69,88.	67,70,70,70.	69.25
25. 31,71,60,88.	64,67,64,70.	66.25
26. 03,58,57,93.	61,67,67,73.	67.00
27. 53,21,93,88.	67,70,73,70.	70.00
28. 23,22,96,79.	67,64,73,70.	68.50
29. 08,56,50,36.	73,67,67,67.	68.50
30. 08,15,08,84.	64,64,64,70.	65.50

1. - Murray F. Spiegel, *op. cit.*, p. 340.



NUMEROS ALEATORIOS

51772	74640	42331	29044	46621	62898	93582	04186	19640	87056
24033	23491	83587	06568	21960	21387	76105	10863	97453	90581
45939	60173	52078	25424	11645	55870	56974	37428	93507	94271
30586	02133	75797	45406	31041	86707	12973	17169	88116	42187
03585	79353	81938	82322	96799	85659	36081	50884	14070	74950
64937	03355	95863	20790	65304	55189	00745	65253	11822	15804
15630	64759	51135	98527	62586	41889	25439	88036	24034	67283
09448	56301	57683	30277	94623	85418	68829	06652	41982	49159
21631	91157	77331	60710	52290	16835	48653	71590	16159	14676
91097	17480	29414	06829	87843	28195	27279	47152	35683	47280
50532	25496	95652	42457	73547	76552	50020	24819	52984	76168
07136	40876	79971	54195	25708	51817	36732	72484	94923	75936
27989	64728	10744	08396	56242	90985	28868	99431	50995	20507
85184	73949	36601	46253	00477	25234	09908	36574	72139	70185
54398	21154	97810	36764	32869	11785	55261	59009	38714	38723
65544	34371	09591	07839	58892	92843	72828	91341	84821	63886
08263	65952	85762	64286	39238	18776	84303	99247	46149	03229
39817	67906	48236	16057	81812	15815	63700	85915	19219	45943
62257	04077	79443	95203	02479	30763	92486	54083	23631	05825
53298	90276	62545	21944	16530	09878	07516	95715	02526	33537

(2)

(2) Murray R. Spiegel. Op. cit., P.349.

Para elegir las 30 muestras de cuatro elementos cada una se podría haber entrado por cualquier parte de la tabla de números aleatorios.

(B).- La siguiente tablada la distribución de frecuencias de la altura media de las muestras obtenidas en (A). Esto es una distribución muestral de medias. La media y desviación típica se obtienen de esta tabla<sup>1</sup>

MEDIA MUESTRAL.	CUEN-TEO.	f	u	fu	fu <sup>2</sup>
64.00	1	1	-4	-4	16
64.75	0	0	-3	0	0
65.50	2	2	-2	-4	8
66.25	6	6	-1	-6	6
A--- 67.00	4	4	0	0	0
67.75	4	4	1	4	4
68.50	7	7	2	14	28
69.25	5	5	3	15	45
70.00	1	1	4	4	16
		$\Sigma f = N = 30$		$\Sigma fu = 23$	$\Sigma fu^2 = 123$

$$\text{MEDIA} = A + c\bar{u} = A + \frac{c \Sigma fu}{N} = 67.00 + \frac{(0.75)(23)}{30}$$

MEDIA = 67.58 pulgadas.

$$\text{DESVIACION TIPICA} = c \sqrt{\bar{u}^2 - \bar{u}^2} = c \sqrt{\frac{\Sigma fu^2}{N} - \left(\frac{\Sigma fu}{N}\right)^2} = 0.75 \sqrt{\frac{123}{30} - \left(\frac{23}{30}\right)^2}$$

DESVIACION TIPICA = 1.41 pulgadas.

<sup>1</sup> Para formulas ver p.55 murray R. Spiegel. Op. cit.

III.) HIPOTESIS.

### III. HIPOTESIS.

#### HIPOTESIS ESTADISTICA.

Una hipótesis estadística, o simplemente hipótesis, es una suposición o conjetura concerniente a la población, para tomar una decisión estadística con el fin de aceptarla o rechazarla. Antes de aceptar o rechazar una hipótesis, todo investigador deberá probar la validez o nulidad de la misma, puesto que puede ser o no verdadera. Claramente se ve que un medio seguro de probar la hipótesis será un examen de la población entera. Sin embargo, este examen puede llegar a ser impráctico o imposible. Un modo práctico es probar la hipótesis usando una muestra de acuerdo con la teoría de la probabilidad. El resultado de la prueba conducirá a un estadístico ya sea para aceptar o rechazar la hipótesis. La aceptación o rechazo conducirá al investigador a tomar una decisión.

Se hace la hipótesis o suposición de acuerdo con la experiencia pasada, la cual puede o no ser verdadera, afín de encontrar la validez hay que tomar una muestra representativa de la población. La prueba para saber si la hipótesis es buena se hace de acuerdo con la probabilidad de la distribución en el muestreo de la media.

Si la prueba muestra una alta probabilidad, entonces aceptamos la hipótesis propuesta. A menos que la información adicional muestre otra cosa.

Por otro lado, si la prueba muestra una baja probabilidad, podemos establecer que el resultado encontrado es inconsistente con la hipótesis, y por lo tanto rechazamos o nulificamos la hipótesis. En tal caso, se llega a una decisión diferente que puede o no estar basada en la hipótesis.

Otro modo común de establecer una hipótesis es que no hay diferencia entre los valores, tal como una media poblacional que es comparada con una media muestral.

El termino "no diferencia" no significa que el valor real de la diferencia sea cero. Más bien el termino significá que la diferencia es me-

ramente debida a fluctuaciones en el muestreo, por lo tanto si el resultado de la prueba muestra un valor no significativo ( que la variación es solamente debida a fluctuaciones en el muestreo ) de la hipótesis, a esta la tomamos como buena. Si por el contrario la prueba estadística muestra que la diferencia es significativa, la hipótesis es rechazada o nulificada.

Las hipótesis estadísticas que son establecidas con el propósito de posible rechazo o nulificación son llamadas hipótesis nulas. Una hipótesis nula es usualmente denotada por el símbolo  $H_0$ .

Cualquier hipótesis, que difiere de la hipótesis nula, es llamada hipótesis alternativa y es denotada por  $H_1$ . En una prueba dada hay usualmente sólo una hipótesis nula, pero puede haber muchas hipótesis alternativas.

Los métodos estadísticos, los cuales son usados para decidir si se acepta o rechaza una hipótesis son llamadas pruebas de hipótesis, pruebas de significación o métodos para formular reglas de decisión.

#### 1.) ERRORES TIPO I Y TIPO II

Cuando alguien toma una decisión concerniente a una hipótesis basada en la teoría de la probabilidad, debe saber que esta decisión puede ser correcta o incorrecta de la siguiente manera.

##### 1.- Dos tipos de decisión correcta:

I.- Rechaza una hipótesis cuando no es verdadera.

II.- Acepta una hipótesis cuando es verdadera.

##### 2.- Dos tipos de decisiones incorrectas o errores:

ERROR TIPO I.- Rechaza una hipótesis cuando realmente es verdadera.

ERROR TIPO II.) Acepta una hipótesis cuando realmente NO es verdadera.<sup>1</sup>

---

1.- Shao. Op. cit. p. 380.

La probabilidad de cometer un error tipo I es usualmente denotada por  $(\alpha)$ , y la probabilidad de cometer un error tipo II es denotada por  $(\beta)$ .

## 2.- NIVEL DE SIGNIFICACION.

La probabilidad máxima, con la que el ensayo de una hipótesis se puede cometer un error tipo I se llama nivel de significación del ensayo. Esta probabilidad se denota como ya hemos enunciado anteriormente por  $\alpha$ ; generalmente se fija antes de la extracción de las muestras, de modo que los resultados obtenidos no influyen en la elección.

En la práctica se acostumbra a utilizar niveles de significación entre el 0.05 al 0.01, aunque igualmente pueden emplearse otros valores. Si, por ejemplo, se elige un nivel de significación del 0.05 ó 5% al diseñar un ensayo de hipótesis, entonces hay aproximadamente 5 ocasiones en 100, en que se rechazaría la hipótesis cuando debería ser aceptada, es decir, es está con un 95% de confianza de que se toma la decisión adecuada. En tal caso se dice que la hipótesis ha sido rechazada al nivel de significación de 0.05 lo que significa que se puede cometer error con una probabilidad de 0.05 ó 5%.

## 3.) PRUEBAS BASICAS DE HIPOTESIS.

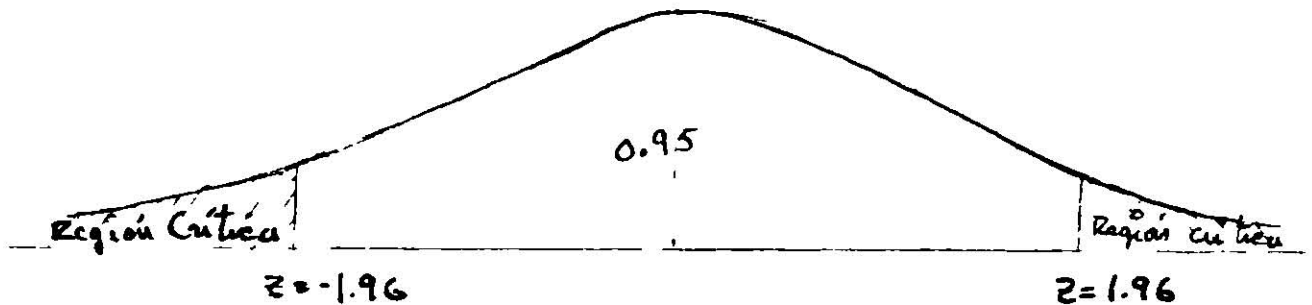
Ensayos referentes a la distribución normal.

Para aclarar las ideas anteriores, supóngamos que con una hipótesis dada, la distribución muestral de un estadístico S es una distribución normal con media  $\mu_s$  y desviación típica  $\sigma_s$ . Entonces la distribución de la variable tipificada (representada por Z) dada por  $Z = \frac{(s - \mu_s)}{\sigma_s}$  es normal tipificada (media 0, varianza 1) y se muestra en la siguiente figura.

Como se indicará en la figura, se puede estar con el 95% de confianza de que, si la hipótesis es cierta, el valor de Z obtenido de una muestra real para el estadístico S se encontrará entre - 1.96 y 1.96 (puesto que el área bajo la curva normal entre estos valores es 0.95).

Sin embargo, si al elegir una muestra al azar se encuentra que Z pa-

ra este estadístico se encuentra fuera de rango  $-1.96$  y  $1.96$ , lo que significa que es un suceso con probabilidad de solamente  $0.05$  (área sombreada en la figura) si la hipótesis fuera verdadera. Entonces puede decirse que esta  $Z$  difiere significativamente de lo que cabía esperar de esta hipótesis y se estaría inclinando a rechazar la hipótesis.



El área total sombreada  $0.05$  es el nivel de significación del ensayo. Representa la probabilidad de cometer error tipo I. Así pues, se dice que la hipótesis se rechaza al nivel de significación del  $0.05$  ó que la  $Z$  obtenida del estadístico muestral dado, es significativa al nivel de significación del  $0.05$ .

El conjunto de las  $Z$  que se encuentran fuera del rango  $-1.96$  y  $1.96$  constituyen lo que se llama región crítica o rechazo de la hipótesis, ó región de significación. El conjunto de las  $Z$  que se encuentran dentro del rango  $-1.96$  y  $1.96$  podía entonces llamarse región de aceptación de la hipótesis o región de no significación.

De acuerdo con lo dicho hasta ahora, se puede formular la siguiente regla de decisión o ensayo de hipótesis o significación.

a).- Se rechaza la hipótesis al nivel de significación del  $0.05$  si la  $Z$  obtenida para el estadístico  $S$  se encuentra fuera del rango  $-1.96$  y  $1.96$  (es decir,  $Z \geq 1.96$  ó  $Z < -1.96$ ). Esto equivale a decir

que el estadístico muestral observado es significativo al nivel del 5%.

b).- Se acepta la hipótesis (o si se desea no se toma decisión alguna) en caso contrario.

A causa de su importante papel en los ensayos de hipótesis y significación, Z recibe también el nombre de ensayo estadístico.

Debe ponerse de manifiesto que pueden igualmente emplearse otros niveles de significación. Por ejemplo, si se utilizase el nivel del 0.01 se sustituiría 1.96 en todo lo visto por 2.58.

TABLA DE VALORES DE "Z" PARA ENSAYOS BILATERALES.<sup>1</sup>

---

NIVELES DE CONFIANZA%	99.73	99.00	98.00	96.00	95.45	95.00	90.00	80.00	50.00
Z.	3.00	2.58	2.33	2.05	2.00	1.96	1.645	1.28	0.6745

---

#### ENSAYOS DE UNA Y DOS COLAS.

En el ensayo anterior se atendía a los valores extremos del estadístico S o su correspondiente Z a ambos lados de la media, es decir, en las dos "colas" de la distribución. Por esta razón, tales ensayos se llaman ensayos de dos colas o bilaterales.

Sin embargo, con frecuencia, se puede estar solamente interesado en los valores extremos a un solo lado de la media, es decir en una cola de la distribución, como, por ejemplo, cuando se están ensayando la hipótesis de que un proceso es mejor que otro (que es diferente a ensayar si un proceso es mejor o peor que otro). Tales ensayos se llaman ensayos de una cola o ensayos unilaterales. En tales casos, la región crítica es una región a un lado de la distribución, con área igual al nivel de significación.

---

1.- Murray R. Spiegel. Op. cit. p.157.



TABLA DE VALORES DE "Z" PARA ENSAYOS UNILATERALES.<sup>1</sup>

NIVEL DE SIGNIFICACION.	10%	5%	1%	.5%	.2%
VALORES DE "Z" PARA UNA COLA.	-1.28 ó1.28	-1.645 ó1.645	-2.33 ó2.33	-2.58 ó2.58	-2.88 ó2.88
VALORES DE "Z" PARA DOS COLAS.	-1.645 y1.645	-1.96 y1.96	-2.58 y2.58	-2.81 y2.81	-3.08 y3.08

Para otros niveles de significación se pueden usar las áreas bajo la curva normal.

HIPOTESIS (EJEMPLOS).

1.- La duración media de una muestra de 100 tubos fluorescentes producidos por una compañía resulta ser 1,570 horas, con una desviación típica de 120 horas, si  $\mu$  es la duración media de todos los tubos producidos por la compañía, comprobar la hipótesis  $\mu = 1,600$  horas contra la hipótesis alternativa  $\mu \neq 1,600$  horas con un nivel de significación de (a) 0.05 y (b) 0.01<sup>2</sup>

SOLUCION:

Se tiene que decidir entre las hipótesis:

$$H_0 : \mu = 1,600 \text{ horas.}$$

$$H_1 : \mu \neq 1,600 \text{ horas.}$$

Un ensayo bilateral debe ser aquí utilizado, puesto que  $\mu \neq 1,600$  incluye valores mayores y menores de 1,600.

(a).- Para un ensayo bilateral al nivel de significación de 0.05 se tiene la siguiente regla de decisión:

(1).- Se rechaza  $H_0$  si la Z de la media muestral está fuera del rango -1.96 y 1.96.

(2).- Se acepta  $H_0$  (o no se toma decisión alguna) en caso contrario.

1.- Murray R. Spiegel. Op. cit. p.169

2.- Ibid., p. 175.

El estadístico bajo consideración es la media muestral  $\bar{X}$ . La distribución muestral de  $\bar{X}$  tiene una media  $\mu_{\bar{X}} = \mu$  y una desviación típica  $\sigma_{\bar{X}} = \sigma/\sqrt{n}$ , donde  $\mu$  y  $\sigma$  son la media y la desviación típica de la población de todos los tubos producidos por la compañía.

Bajo la hipótesis  $H_0$ , se tiene  $\mu = 1,600$  y  $\sigma_{\bar{X}} = \sigma/\sqrt{n} = 120/\sqrt{100} = 12$ , utilizando la desviación típica muestral como una estimación de  $\sigma$ . Puesto que  $Z = (\bar{X} - 1600) / 12 = (1570 - 1600) / 12 = -2.50$  se encuentra fuera del rango  $-1.96$  a  $1.96$ , se rechaza  $H_0$  al nivel de significación del  $0.05$ .

(b).- Si el nivel de significación es  $0.01$ , el rango  $-1.96$  a  $1.96$  en la regla de decisión del apartado (a) se sustituye por  $-2.58$  a  $2.58$ . Entonces, puesto que  $Z = -2.50$  se encuentra dentro de este rango, se acepta  $H_0$  (o no se toma ninguna decisión) al nivel de significación del  $0.01$ .

2.- En el problema anterior, ensayar la hipótesis  $\mu = 1,600$  frente a la alternativa  $\mu < 1,600$  horas con los niveles de significación de (a)  $0.05$  y (b)  $0.01$ .

#### SOLUCION.

Se debe decidir entre las dos hipótesis:

$$H_0 : \mu = 1,600 \text{ horas.}$$

$$H_1 ; \mu < 1,600 \text{ horas.}$$

Debe aquí utilizarse un ensayo unilateral siendo la figura correspondiente idéntica a la del problema anterior.

(a).- Si el nivel de significación es  $0.05$ , se tiene que  $Z = -1.645$  (según la tabla de ensayos unilaterales), por tanto, será la regla de decisión la siguiente:

(1). Se rechaza  $H_0$  si  $Z$  es menor de  $-1.645$ .

(2). Se acepta  $H_0$  (o no se toma ninguna decisión) en caso contrario.

Como en el problema anterior apartado (a), el valor de  $Z$  es  $-2.50$ , que es menor de  $-1.645$ , se rechaza, pues,  $H_0$  al nivel de significación

del 0.05. Adviértase que esta decisión es idéntica a la obtenida en el problema anterior mediante un ensayo bilateral.

(b).- Si el nivel de significación es de 0.01, el valor de  $Z_1$  es -2.33. De aquí que se adopte la regla de decisión:

(1).- Se rechaza  $H_0$  si  $Z$  es menor de -2.33.

(2).- Se acepta  $H_0$  (o no se toma decisión alguna) en caso contrario.

Como en el problema anterior apartado (a), el valor de  $Z$  es -2.50, que es menor de -2.33, se rechaza, pues,  $H_0$  al nivel de significación del 0.01. Adviértase que esta decisión es la misma que la obtenida en el problema anterior apartado (b) mediante un ensayo bilateral.

Las decisiones concernientes a hipótesis  $H_0$  dadas, basadas en ensayos de una o dos colas, no están siempre en concordancia. Esto cabe, naturalmente, esperarse, puesto que se ensaya  $H_0$  contra una alternativa diferente en cada caso.

3.- El fabricante de una patente médica sostiene que las mismas tienen un 90% de efectividad en el alivio de una alergia, por un período de 8 horas. En una muestra de 200 individuos que tenían la alergia, la medicina suministrada alivió a 160 personas. Determinar si la aseveración del fabricante es cierta.

SOLUCION:

Denótese por la " $p$ " la probabilidad de obtener alivio de la alergia utilizando la medicina. Entonces se debe decidir entre las dos hipótesis:

$H_0$ :  $p = 0.9$  y la aseveración es correcta.

$H_1$ :  $p < 0.9$  y la aseveración es falsa.

Se elige un ensayo unilateral, puesto que se está interesado en determinar si la proporción de gente aliviada por la medicina es demasiado baja.

Si se toma el nivel de significación del 0.01, que corresponde a  $Z_1 =$

a -2.33, como puede comprobarse en la tabla para ensayos unilaterales.

Se toma como regla de decisión:

(1).- La aseveración no es legítima si Z es menor de -2.33 (en cuyo caso se rechaza  $H_0$ ).

(2).- En caso contrario, la aseveración es legítima y los resultados obtenidos se deben a la casualidad (en cuyo caso se acepta  $H_0$ ).

Si  $H_0$  es cierta  $\mu = Np = 200(0.9) = 180$  y  $\sigma = \sqrt{Npq} = \sqrt{200 \times 0.9 \times 0.1} = 4.23$ .

Ahora, 160 en unidades tipificadas =  $(160 - 180)/4.23 = -4.73$ , que es mucho menor de -2.33. Así, pues, por la regla de decisión se deduce que la aseveración no es legítima y que los resultados muestrales son altamente significativos.

**IV.) BIBLIOGRAFIA.**

IV. BIBLIOGRAFIA.

- ESTADISTICA PARA ECONOMISTAS Y ADMINISTRADORES DE EMPRESAS.  
STEPHEN P. SHAO.  
MEXICO, 1970.
  
- ESTADISTICA (SCHAUM'S).  
MURRAY R. SPIEGEL PH. D.  
COLOMBIA, 1971.
  
- FUNDAMENTOS DE ESTADISTICA, APLICADA A LOS NEGOCIOS Y LA ECONOMIA.  
NETER, JOHN.  
MEXICO, 1972.
  
- INTRODUCCION A LA ESTADISTICA DESCRIPTIVA.  
RASCON, CH. OCTAVIO.  
MEXICO, 1970.
  
- MUESTREO DE ENCUESTAS.  
KISH, LESLIE.  
MEXICO, 1972.
  
- INVESTIGACION Y ANALISIS DEL MERCADO.  
LUCAS - WALES - TAYLOR.  
BARCELONA, 1971.

1.) BIBLIOGRAFIA PARA EL ALUMNO.

ELEMENTOS BASICOS DE ESTADISTICA.

SHAO. CAP. I. PP. 3 a la 20.

MURRAY. CAP. II. PP. 27 a la 44.

PRESENTACION DE DATOS ESTADISTICOS.

SHAO. CAP. IV. PP. 58 a la 92.

PROMEDIOS: MEDIA MEDIANA MODA.

SHAO. CAP. VI. PP. 152 a la 191.

MURRAY. CAP. III. PP. 45 a la 69.

MEDIDAS DE DISPERSION.

SHAO. CAP. VIII. PP. 217 a la 234.

MURRAY. CAP. III. PP. 69 a la 89.

DISTRIBUCION NORMAL.

SHAO. CAP. XI. PP. 300 a la 319.

MURRAY. CAP. VII. PP. 122 a la 141.

METODOS DE MUESTREO.

SHAO. CAP. XII. PP. 322 a la 349.

MURRAY. CAP. VIII. Y IX. PP. 156 a la 166.

NETEP. CAP. XI. PP. 495 a la 541.

KISH. CAP. II. PP. 58 a la 100.

RASCON. VOL. I. CAP. I. PP. 1 a la 32.

HIPOTESIS.

SHAO. CAP. XIV. PP. 376 a la 398.

MURRAY. CAP. X. PP. 167 a la 187.

NETEP. CAP. XV. PP. 543 a la 628.

**PRUEBAS DE HIPOTESIS.**

**SHAO. CAP. XV. PP. 398 a la 450.**

**MURRAY. CAP. X. PP. 167 a la 187.**



V.) PRACTICA.

V.- PRACTICA.

1.- Los valores que se dan enseguida, representan las distancias en kilómetros de 30 distribuidores, hasta el almacén de una compañía.; 12, 50, 9, 11, 5, 35, 50, 5, 12, 8, 41, 32, 11, 5, 9, 50, 32, 43, 5, 12, 21, 25, 8, 12, 50, 42, 22, 21, 33, 5.

(a).- Haga una tabla de distribución de frecuencias usando los siguientes intervalos de clase:

0 y menos de 10.

10 y menos de 20.

20 y menos de 30.

30 y menos de 40.

40 y menos de 50.

(b).- Con base a la distribución hecha, obtenga:

1.- La media.

2.- La moda.

3.- La desviación media.

4.- La desviación estandar.

(c).- Interprete los resultados obtenidos.

2.- El gerente de una estación de servicio desea muestrear las notas de venta a fin de encontrar la cantidad promedio (media) por venta durante un período dado . El indica que;

(a).- El máximo de error muestral no deberá ser mayor que  $\$0.20$  por arriba o por abajo de la verdadera media.

(b).- El nivel de confianza deberá ser 99.73%.

(c).- La desviación estándar de la población basada en su experiencia pasada es estimada en  $\$0.80$ .

Encontrar el tamaño de muestra adecuado de acuerdo con sus indicaciones.

3.- Si se desea obtener datos de las edades de los <sup>353</sup>empleados de una



10.- Si se tienen una población de 1,300 calificaciones de geografía de una escuela de bachillerato, ¿qué tipo de población es?

FINITA.                      ó                      INFINITA.

11.- Las poblaciones finitas tienen un número; \_\_\_\_\_ de elementos.

LIMITADO.                      ó                      ILIMITADO.

12.- Las poblaciones infinitas tienen un número \_\_\_\_\_ de elementos.

LIMITADO.                      ó                      ILIMITADO.

14.- La mayoría de las veces no es posible o práctico observar todos los elementos de una población, en tal caso se toma sólo una parte de ella llamada muestra.

¿Tiene una muestra más elementos que la población?

SI.                      ó                      NO.

15.- Un grupo de datos extraídos de una población constituye una:

MUESTRA.                      ú                      OBSERVACION.

16.- Al ingresar los estudiantes de una universidad, se les somete a un examen de admisión . Todas las calificaciones de dicho examen constituyen una población de observación numérica. Una parte de estas observaciones constituirá una \_\_\_\_\_.

OBSERVACION.                      o                      MUESTRA.

17.- Si se quisiera muestrear el grado de inteligencia de los estudiantes de una escuela, y dicha muestra se seleccionara extrayendo de una urna los nombres de cada alumno, escritos en una tira de papel, reponiendo cada tira extraída, todos los alumnos tendrían igual oportunidad de ser extraídos y los elementos serían independientes, por lo cual, se obtendría una muestra \_\_\_\_\_.

NO REPRESENTATIVA.

REPRESENTATIVA. (aleatoria)

18. Cuando algunos elementos de la población no tiene la misma probabilidad que los demás de ser incluido en una muestra, dicha muestra no es \_\_\_\_\_.

NO PROBABILISTICA.

ALEATORIA.

SESGADA.

19.- Cuando se extrae una muestra con datos independientes, pero con diferente oportunidad de ser seleccionados, la muestra es \_\_\_\_\_

SESGADA.

ALEATORIA.

20.- Es conveniente, para ahorrar tiempo durante la selección de la muestra, que la numeración sea progresiva, comenzando con el número uno.

Si se procede en esta forma para una población de 500 elementos, los números asignados a cada elemento irán del 1 al \_\_\_\_\_.

0.	1,000	570.	500.	50.
----	-------	------	------	-----

21.- Si se utiliza el procedimiento de la urna para seleccionar una muestra aleatoria, cada bola de la urna representa un elemento de la población y cada bola extraída identifica a un elemento de \_\_\_\_\_.

LA MUESTRA.

LA POBLACION.

22.- Existen tablas de números aleatorios para facilitar la selección de una muestra aleatoria. Estos números se calculan en tal forma que todos ellos tengan la misma oportunidad de ser seleccionados y sean independientes. Entonces, si se utiliza una tabla de números aleatorios para obtener una muestra, ésta resulta, una muestra \_\_\_\_\_.

SESGADA.

ALEATORIA.

23.- Si a cada elemento de una población se le asigna un número diferente, y se seleccionan diez números de una tabla de números aleatorios se obtiene una muestra \_\_\_\_\_ la cual contiene \_\_\_\_\_ datos.  
                  segsada/aleatoria  cuántos?

24.- Mediante el uso de una tabla de números aleatorios, se obtiene una muestra representativa (o aleatoria) porque todos los elementos de la población tienen                      oportunidad de ser seleccionados y además son independientes entre sí.  
igual/diferente.

25.- Si la población que se va a muestrear consta de 6,600 elementos y se utiliza numeración corrida para identificarlos, se requieren números aleatorios con cuatro dígitos para seleccionar la muestra aleatoria. Si la población constara de 50 elementos se necesitarían números aleatorios con                      dígitos para seleccionar los elementos de una muestra aleatoria.  
1/2/5/7/8/9

26.- El número 25,436 tiene cinco dígitos, el 356, 789 tiene seis dígitos. ¿Cuántos dígitos tienen los números 4,768 y 26?                     

27.- Si los números de la tabla tienen cinco dígitos y solo necesitamos cuatro, hay que excluir una de ellos. Por ejemplo, si los números aleatorios son:

16408.

18629.

73115.

Entonces, para seleccionar la muestra representativa, se excluye el dígito del extremo izquierdo y sólo se necesitarán usar los números 6408, 8629, 3115. Si sólo se necesitarán dos dígitos usaríamos los números                       
408

115, 629/08, 29, 15.

28.- El grupo de todos los resultados posibles al realizar una secuencia exhaustiva de experimentos, se denomina                     .  
poblacion/muestra.

29.- Un grupo de datos constituyen una                     .  
poblacion/ muestra.

30.- Una hacienda exportará 3,200 cabezas de ganado vacuno; para obtener una idea de la ganancia que se obtendrá, se seleccionan al azar 10 animales y se pesan; para lo cual, se numeran en orden progresivo empezando por el uno.

- a).- ¿Cuántos elementos tiene la población?
- b).- ¿La población es finita o infinita?
- c).- ¿La muestra de cuántas observaciones consta?
- d).- ¿Cuáles son las observaciones que interesan, los animales o sus pesos?
- e).- ¿Con una tabla de números aleatorios determine la muestra.

31.- Supongase que un gerente desea seleccionar 10 elementos de un grupo de 200 elementos para un muestreo de inventario. Para un muestreo de inventario. ¿Usted cómo lo haría?

32.- (hipótesis). Una urna contiene bolas rojas y azules. Para ensayar la hipótesis de proporciones iguales de estos colores, se acuerda de tomar una muestra de 64 bolas con reemplazamiento, anotando los colores extraídos y adoptando la siguiente regla de decisión: (1) Se acepta la hipótesis si se extraen entre 26 y 36 bolas rojas; (2) Se rechaza en caso contrario. Hallar la probabilidad de rechazar la hipótesis cuando en realidad sea correcta.

33.- Supongase que en el anterior problema se desea ensayar la hipótesis de que hay mayor proporción de bolas rojas que de azules.

- (a).- ¿Cuál sería la hipótesis nula y cuál la alternativa.
- (b).- ¿Se utilizaría un ensayo de una o dos colas?
- (c).- ¿Qué regla de decisión se adoptaría si el nivel de significación fuese de 0.5?
- (d).- ¿Cuál sería la regla de decisión si el nivel de significación fuese del 0.01?

34.- La resistencia a la rotura de los cables producidos por un fabricante tienen una media de 1,800 libras y una desviación típica de 100 libras. Mediante una nueva técnica en el proceso de fabricación se aspira a que esta resistencia pueda ser incrementada. Para ensayar esta aspiración, se ensaya una muestra de 50 cables y se encuentra que su resistencia media es de 1,850 libras. ¿ Puede mantenerse que, en efecto, hay un aumento de resistencia al nivel de significación del 0.01?.



1.- SOLUCION DE LA PRACTICA.

1.- TABLA DE DISTRIBUCION DE FRECUENCIAS.

DISTANCIAS EN Km. clase.	MARCA.	N° DE FRECUENCIAS. f
0 y menos de 10.	11111 1111	9
10 y menos de 20.	11111 1	6
20 y menos de 30.	1111	4
30 y menos de 40.	1111	4
40 y menos de 50.	11111 11	7
TOTAL.		n = 30

(b). 1.- MEDIA = 23 kilómetros.

2.- MODA CRUDA = 5 kilómetros.

3.- DESVIACION MEDIA = 14 kilómetros.

4.- DESVIACION ESTANLAR = 15.5 kilómetros.

(c).- INTERPRETACIONES:

1.- La media de 23 Km. Es la distancia promedio entre el almacén y los distribuidores.

2.- La moda de 5 Km. Si queremos visitar a un distribuidor cualquiera, al azar, podemos estimar que visitaremos a uno que se encuentre a una distancia entre 0 y menos de 10, que su punto medio es 5.

3.- Las dispersiones, quieren decir que esas distancias son las desviaciones más probables con respecto a la media.

2.- El tamaño de la muestra de acuerdo a las indicaciones dadas es de "144 elementos".

3.- "353"

4.- POBLACION.

- 5.- 10,000.
- 6.- POBLACION.
- 7.- SI.
- 8.- FINITA.
- 9.- SI.
- 10.- FINITA.
- 11.- LIMITADO.
- 12.- ILIMITADO.
- 13.- FINITAS E INFINITAS.
- 14.- NO.
- 15.- MUESTRA.
- 16.- MUESTRA.
- 17.- ALEATORIA (representativa).
- 18.- ALEATORIA.
- 19.- SEEGADA.
- 20.- 500.
- 21.- LA POBLACION.
- 22.- ALEATORIA.
- 23.- ALEATORIA. Y 10.
- 24.- IGUAL.

25.- DOS.

26.- CUATRO Y DOS RESPECTIVAMENTE.

27.- 08, 29, 15.

28.- POBLACION.

29.- UNA MUESTRA.

30.- 3200, PINITA, 10, SUS PESOS, Y CUALQUIERA.

31.- POR NÚMEROS ALEATORIOS.

32.- LA PROBABILIDAD DE RECHAZAR LA HIPOTESIS CUANDO EN REALIDAD SEA CORRECTA ES DE 0.2606.

33.- (a).-  $H_0 : p = 0.5$  y  $H_1 : p > 0.5$

(b).- ENSAYO UNILATERAL.

(c).- RECHAZAR  $H_0$  SI SE EXTRAEN MAS DE 39 BOLAS ROJAS Y ACEPTARLO EN CASO CONTRARIO.

(d).- RECHAZAR  $H_0$  SI SE EXTRAEN MAS DE 41 BOLAS ROJAS, Y ACEPTARLO EN CASO CONTRARIO.

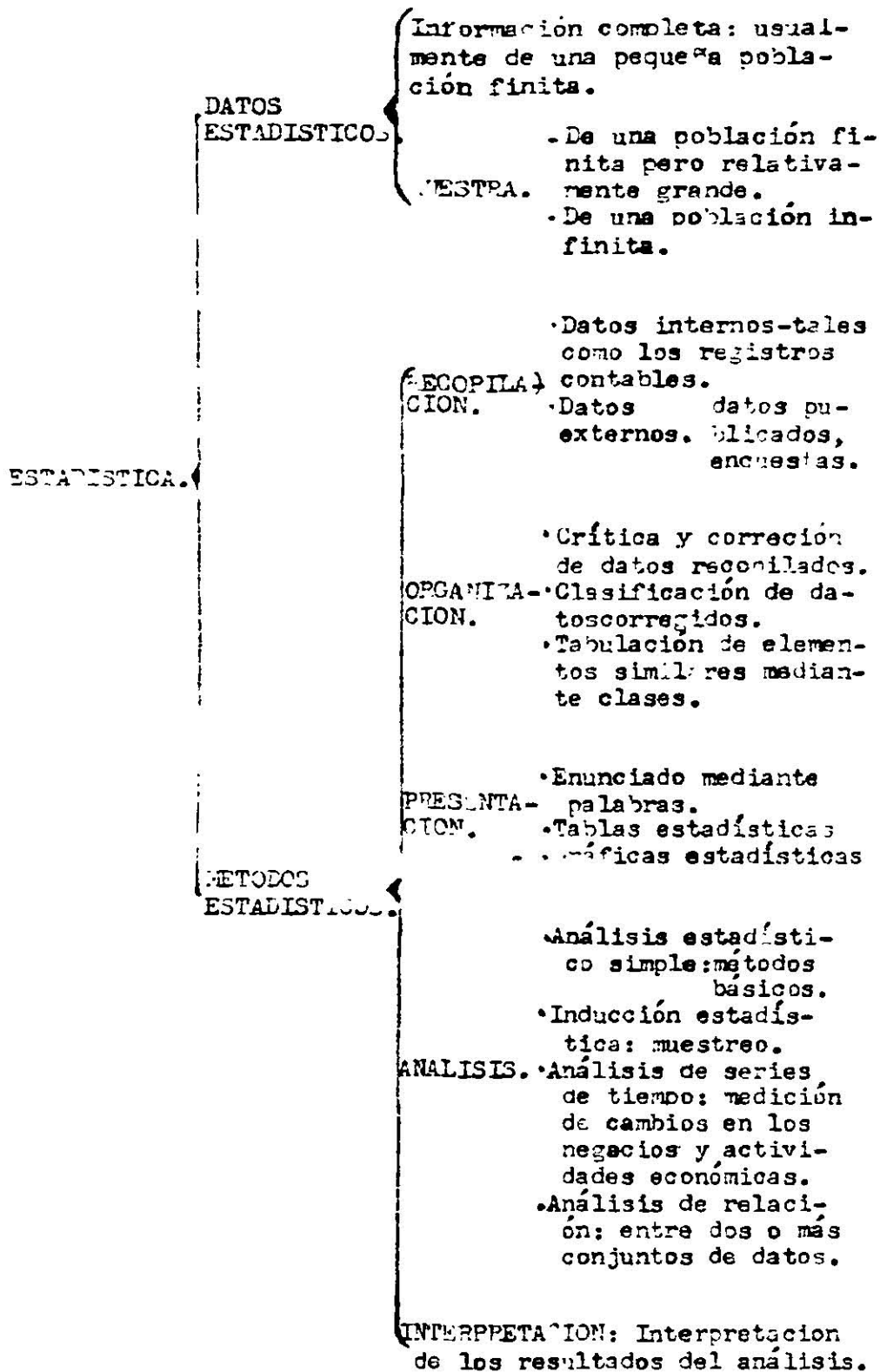
34.- SE USA UN ENSAYO UNILATERAL.

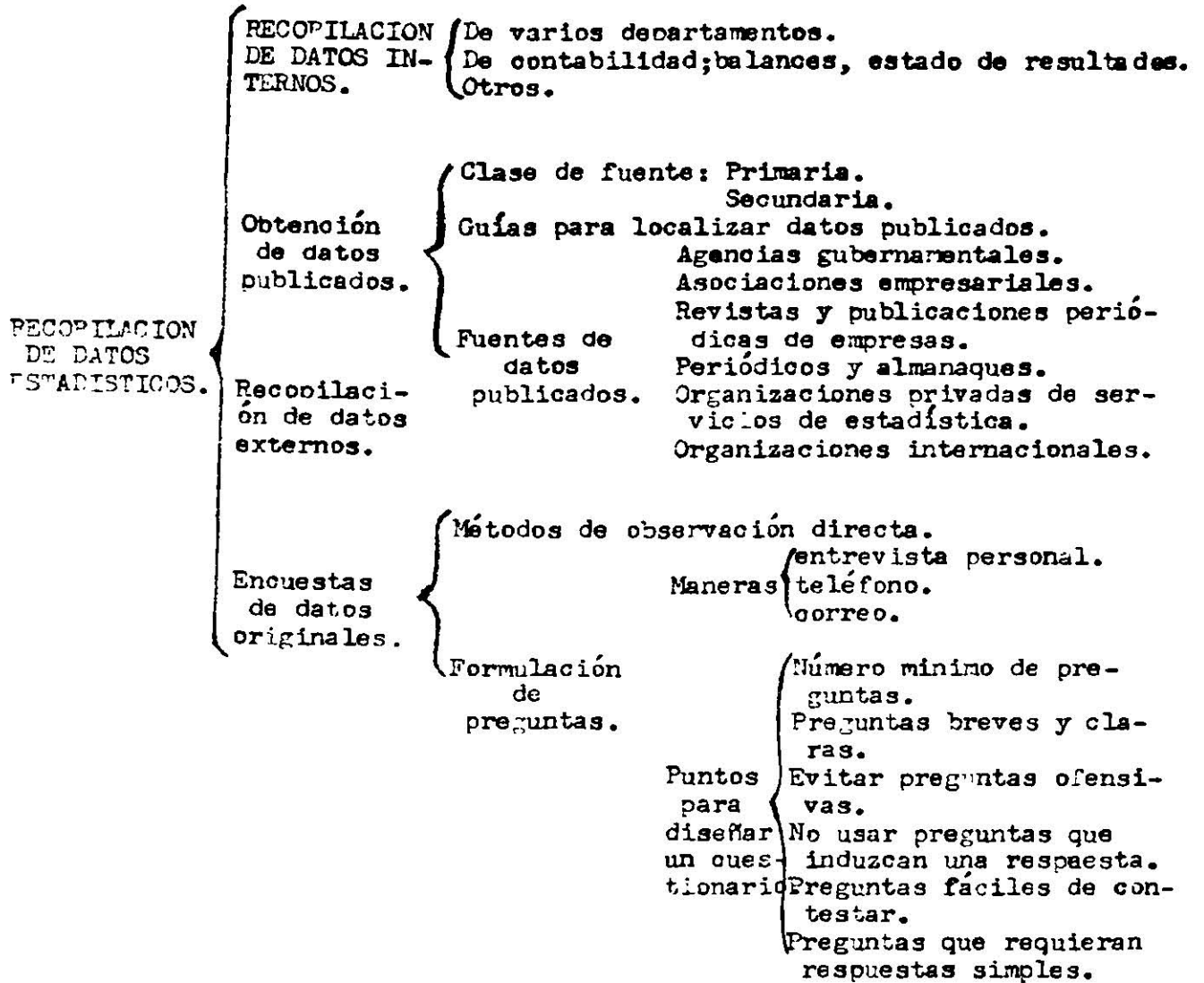
$$Z = 3.55$$

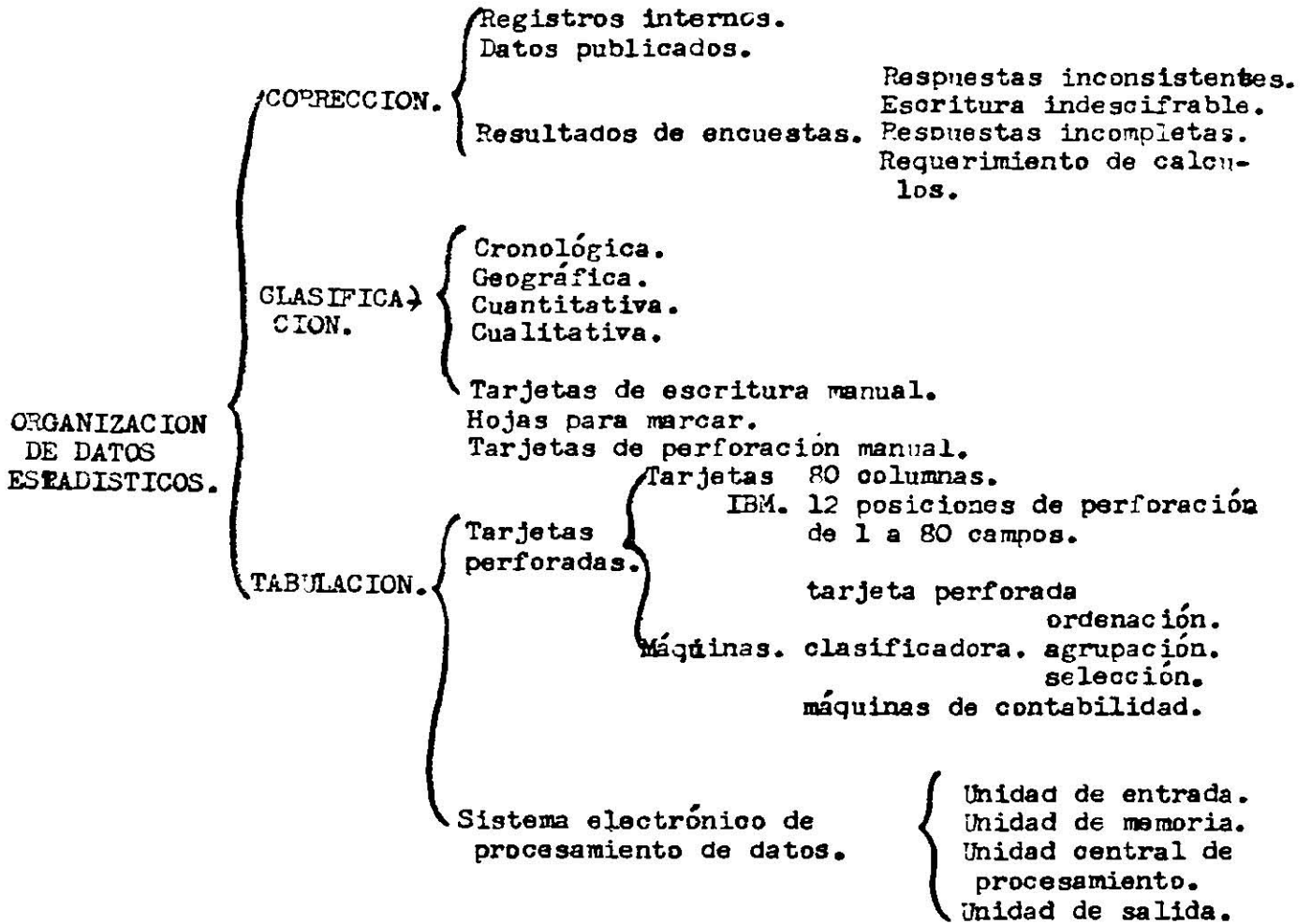
QUE ES MAYOR DE 2.33 DE AQUI SE DEDUCE QUE LOS RESULTADOS SON ALTAMENTE SIGNIFICATIVOS Y LA ASPIRACION DE MEJORA DEBE SER ADMITIDA.

**VI.) TABLAS Y CUADROS SINOPTICOS.**

VI.- TABLAS Y CUADROS SINOPTICOS.







V  
E  
S  
T  
I  
G  
I  
O  
S  
A  
I  
C  
O  
S  
E  
C  
O  
N  
O  
M  
I  
C  
O  
S  
E  
C  
O  
N  
O  
M  
I  
C  
O  
S

**TABLAS.**

Mediante palabras: usado para presentar datos concretos.

Propósitos. Tablas para propósitos generales.  
Tablas para propósitos especiales.

Partes principales.
 

- Básicas. Título.  
Encabezado.  
Conceptos ó columna matriz.  
Cuerpo.
- Adicionales. nota de encabezado  
nota de pie.  
fuente de datos.

Puntos importantes en la construcción de una tabla.
 

- Simplificar la presentación de la tabla.
- Tratar un tema por tabla.
- Arreglar las clasificaciones en orden.
- Hacer clasificaciones fáciles.
- Enfatizar cifras importantes.
- Redondear en los dígitos los detalles innecesarios.
- Mejorar la apariencia en la tabla.

**Gráficas.**

Principio de construcción de gráficas: Coordenadas rectangulares.

Partes principales
 

- Básicas. Título.  
Escalas. (vertical, y Horizontal).  
Gráficas.
- Adicionales. nota de encabezado.  
nota de pie.  
fuente de datos y gráfica.

Tipos comunes.
 

- Gráfica de línea. Verticales.
- Gráficas de barras. Horizontales.  
De línea.
- Gráfica de partes componentes. De barras.  
Gráficas de pastel.
- Gráfica de dimensiones. Gráfica de áreas. (2 dimensiones).  
Gráficas de volumen. (3).
- Pictogramas: usa dibujos de símbolos de igual tamaño.
- Mapas.



TIPOS DE MUESTREO.

DE ACUERDO AL NUMERO DE MUESTRAS TOMADAS DE LA POBLACION.

MUESTREO SIMPLE.

MUESTREO DOBLE.

MUESTREO MULTIPLE.

A LA MANERA USADA A SELECCIONAR LA MUESTRA.

MUESTREO DE JUICIO. (no probabilistico).

MUESTREO ALATORIO. (probabilistico).

MUESTREO ALATORIO SIMPLE.

MUESTREO ALATORIO MODIFICADO.

MUESTREO SISTEMATICO.

MUESTREO ESTRATIFICADO.

MUESTREO DE CONJUNTERADOS.

**HIPOTESIS.**

**TIPOS DE ERRORES:**

**ERROR TIPO I ( $\alpha$ ).**

**ERROR TIPO II ( $\beta$ ).**

**PRUEBAS DE HIPOTESIS.**

**PRUEBA DE UNA COLA  
(UNILATERAL)**

**PRUEBA DE DOS COLAS.  
(BILATERAL)**

## VII.- CONCLUSION.

La estadística es una herramienta de la cual la mercadotecnia hace uso y gracias a ella se logran resultados más confiables principalmente en lo que respecta a la investigación de mercados que es donde los elementos de probabilidad y muestreo son más aplicados.

Los elementos de probabilidad y muestreo que son aplicados en la investigación de mercados sirven entre otras cosas para elegir el tipo más apropiado de muestreo para cada caso particular, para hacer la elección de un tamaño de muestra adecuado así como para poder predecir que grado de confiabilidad se obtiene o se podría obtener de una muestra.

El seminario de investigación que aquí se presenta está realizado con el fin de que sirva de guía y motive a los alumnos de los nuevos planes de estudios (modulares), en el curso correspondiente al módulo producto-mercado (mercadotecnia).

Corresponde este trabajo al tema número tres del programa del módulo de investigación producto-mercado y es por tanto sólo una parte del programa modular correspondiente a la mercadotecnia que se pretende cubrir con los seminarios de otros compañeros.

