

# DISTRIBUCIÓN DE FRECUENCIA

## CAPITULO 2

### METODOLOGÍA DE LA ESTADÍSTICA

La estadística es una ciencia que se encarga de recopilar, organizar, presentar, analizar e interpretar un conjunto de datos naturales, político, económico, o sociales, para tomar una decisión de incertidumbre en un momento determinado, mediante una investigación determinada.

Tomando en cuenta lo anterior para realizar un estudio estadístico es necesario presentar el métodos para su estudio, los elementos básicos de mismo son los siguientes: 1) Recopilación, 2) Organización, 3) Presentación, 4) Análisis e 5), Interpretación.

Sí un investigador desea realizar un estudio de la edad de un grupo estudiantes de una universidad determinada, lo primero que deberá hacer es, recopilar los datos estadísticos de una muestra de las edades de los estudiantes de la universidad procurando que la misma sea representativa de la población objeto de la investigación. El tamaño apropiado de una muestra debe estar acorde con las técnicas del muestreo. Segundo puede organizar las edades recopiladas clasificándolos en diferentes grupos de edades tercero, puede presentar los datos organizados en forma tabular para su mejor estudio, 4º, puede analizar las edades presentadas en la tabla para obtener la información deseada. Por ejemplo el puede encontrar que el grupo de edades típica de los estudiantes de esa universidad está comprendida entre 18 y veinte años puesto que allí se encuentra el mayor número de estudiantes. Finalmente el investigador puede interpretar los resultados analizados de la muestra, diciendo que la edad crítica de todos los estudiantes en ese centro de estudios se encuentran comprendidos entre dieciocho y veinte años, ya que la edad de ese grupo de estudiantes, es la que mas se repite.

Para la aplicación de los métodos estadísticos es necesario comenzar a reconocer la existencia de algunas herramientas y conceptos que son de gran importancia. Tal es el caso del concepto de variable, clasificación y su medición, pues esto puede ser considerado al momento de seleccionar las herramientas que le pueden ser aplicadas. Por otro lado, el uso de instrumentos para recopilar la información resulta una parte fundamental para la obtención de datos en el área social, y sus características y validez se deben tomar en cuenta al momento de trabajar en la obtención y recopilación de la información.

### VARIABLES

Las variables son magnitud que puede tener un valor cualquiera de los comprendidos en un conjunto de valores de un estudio o investigación determinada. Son todos aquellos datos u observaciones que pueden ser expresados mediante números, es decir, son características de una población determinada, susceptible de medición. Son características que pueden ser observadas en determinado fenómeno

natural, social, económico, político etc. Las mismas son susceptibles de adoptar distintos valores o ser expresadas en varias categorías.

En los estudios estadísticos que se realizan se busca investigar acerca de una o varias características de la población observada. Para un correcto manejo de la información, estas características deben ser tomadas en cuenta de acuerdo a su tipo para poder aplicar algunas de las operaciones que son necesarios llevar a cabo. Existen muchas definiciones de variables, entre las cuales tenemos:

Son aquellos datos u observaciones que pueden ser expresados cuantitativamente, es decir, son características de una población específica, en las cuales se realiza una investigación en un momento dado. Las variables por lo general se representan con letras mayúsculas y sus valores particulares con minúsculas, es decir, si se hace referencia a los salarios devengados por un grupo de trabajadores la variable **salario** estaría representado por una letra mayúscula, en este caso  $X_i$  y varios salarios de diferentes trabajadores en particular, estarían representados con la letra minúscula correspondiente, así:

$$x_1 = 180.000, x_2 = 190.000, x_3 = 480.00, \text{ etcétera.}$$

Es aquella característica de una población que puede tomar diferentes valores en un estudio determinado. Son símbolos tal como X, Y, Z, A, B, etc., que puede tomar un valor cualquiera de una característica especificada de un estudio determinado. Por lo tanto, son características que pueden ser observadas en determinado fenómeno natural, social, político, económico, etc. Las mismas son susceptibles de adoptar distintos valores o ser expresadas en varias categorías.

Por ejemplo, la estatura de las personas, la talla de un grupo de trabajadores petroleros, la edad de un conjunto de estudiantes universitarios, el índice académico de los estudiantes, son variable. En otras palabras, una variable es una función que asocia a cada elemento de la población la medición de una característica, particularmente de aquella que se desea observar.

## Clasificación de las Variables

Existen muchas clasificaciones de variables, pero en este curso estudiaremos las siguientes:

**Variables Numéricas:** Son aquellas variables que toman valores numéricos. A estas variables le corresponde la escala de medición de intervalo y razón o proporción.

**Variable Aleatoria o probabilística:** Es aquel valor que asume la variable que han sido antecedidas por una selección aleatoria de los objetos medidos o son resultados de algún proceso al azar, es decir, es la variable que toma diferentes valores como resultado de un experimento aleatorio. A las variables aleatorias usualmente se les denota por letras X, Y, Z; y a los valores por las respectivas minúsculas.

Por ejemplo, si en un Banco los números de las cuentas corrientes de los clientes que presentan un saldo superior a 1.000.000 de pesos se escogen veinte al azar, en un día determinado para darles un premio, la variable **número de cuenta corriente** de cada cliente constituye una variable aleatoria que podemos designar  **$X_i$** . Por lo tanto, la variable aleatoria es una función de valor real que tiene como dominio el espacio muestral asociado a un experimento aleatorio.

**Variabes Cualitativas o categóricas:** Son aquellas variables cuyos valores son del tipo categórico, es decir; que indican categorías o son etiquetadas *numéricamente* o con *nombres*. Son las que se refieren a clasificaciones, como: estado civil, profesión, color de los ojos, preferencia por una marca etc., es decir, son aquellas que no aparecen en forma numérica, sino como categorías o atributos. Esta a su vez, se clasifica en:

**Variabes Categóricas Nominales:** Son las variables categóricas que, además de que sus posibles valores son mutuamente excluyentes entre sí, no tienen alguna forma “natural” de ordenación. Por ejemplo, cuando sus posibles valores son: “Sí” y “No”. A este tipo de variable le corresponde las escalas de medición nominal.

**Variabes Categóricas Ordinales:** Son las variables categóricas que tienen algún orden. Por ejemplo, cuando sus posibles valores son: “siempre”, “casi siempre” y “nunca”. A estos tipos de variables le corresponden las escalas de medición ordinal.

**Variable Continua:** Es aquella que puede tomar cualquier valor dentro de la escala de los números reales, es decir, es aquella que teóricamente puede tomar cualquier valor, bien sean valores enteros o fraccionados. Los valores que puede tomar esa es cualquiera e incluso valores fraccionados por ejemplo, un alumno A mide 1,68 m y otro alumno B mide 1,69 y otro C mide 2,00 m. Entre estos tres alumnos están los que miden 1.682 m, 1.693 m, 2,125 m, , entre otros. Otros ejemplos de variables continuas, pueden ser la talla de las personas o el tiempo necesario para realizar una transacción bancaria por parte de un conjunto de clientes de una entidad bancaria.

**Variable Discreta:** Es aquella variable que solo puede tomar valores enteros en la escala de los números naturales, es decir, la variable no puede tomar valores fraccionarios. Por ejemplo, el número de hijos en un matrimonio puede ser: 0,1, 2, 3 ó 4; pero ningún matrimonio tiene 0,5 ó 3,89 hijos. También, el número de personas que llegan a una entidad bancaria a solicitar un servicio. Estos ejemplos, muestran que las variables discretas sólo asumen valores enteros, es decir, no admiten fraccionamientos.

**Variable Dependiente:** Es aquella que se presenta como consecuencia de una variable antecedente, que por lo general es la variable independientes.

**Variable Independiente:**

Es aquella que se presenta como *causa y condición* de la variable dependiente. Es la manipulada por el investigador; recibe el nombre de *variable experimental*.

**UNIVERSO:** En estadística es el nombre específico que recibe particularmente en la investigación social la operación dentro de la delimitación del campo de investigación que tienen por objeto la determinación del conjunto de unidades de observaciones del conjunto de unidades de observación que van a ser investigadas. Para muchos investigadores el término universo y población son sinónima. En general, el universo es la totalidad de elementos o características que conforman el ámbito de un estudio o investigación.

**POBLACIÓN:** En estadística el concepto de población va más allá de lo que comúnmente se conoce como tal. En términos estadísticos, población es un conjunto finito o infinito de personas, animales o cosas que presentan características comunes, sobre los cuales se quiere efectuar un estudio

determinado. En otras palabras, la población se define como la totalidad de los valores posibles (mediciones o conteos) de una característica particular de un grupo especificado de personas, animales o cosas que se desean estudiar en un momento determinado. Así, se puede hablar de la población de habitantes de un país, de la población de estudiantes universitarios de la zona sur del Estado Anzoátegui, de la población de casas de la Urbanización Los Ríos de la ciudad de Ibagué, el rendimiento académico de los estudiantes, el número de carros marca Mazda de la ciudad de Ibagué, la estatura de un grupo alumnos, la talla, etc.

La población es el conjunto formado por todos los valores posibles que puede asumir la variable objeto de estudio, ya que constituye la totalidad del grupo que se quiere estudiar los que van a poseer una característica de ese grupo específico de individuos, animales o cosas. Es la colección de todos los elementos que se están estudiando, acerca de los cuales se intenta sacar conclusiones, el cual puede ser un conjunto finito o infinito de personas, animales o cosas que presentan características comunes.

Así por ejemplo, en un estudio sobre la preferencia de los electores que participaran en una elección presidencial, la población consiste en todos los participantes registrados para votar en ese proceso. Pero el término no sólo está asociado a la colección de seres humanos y organismos, también pueden ser cosas no vivientes tales como: el estudio de mercado que se realiza para determinar las ventas anuales de los supermercados de una zona determinada de una ciudad, luego, las ventas anuales de todos los supermercados constituyen así mismo la población.

**MUESTRA:** La muestra es un subconjunto de la población, seleccionado de tal forma, que sea representativo de la población en estudio, obteniéndose con el fin de investigar alguna o algunas de las propiedades de la población de la cual procede. En otras palabras es una parte de la población que sirve para representarla. Es una parte o porción extraída de un conjunto por métodos que permiten considerarla como representativa del mismo. Entonces, una muestra no es más que una parte de la población que sirve para representarla. La muestra debe obtenerse de la población que se desea estudiar; una muestra debe ser definida sobre la base de la población determinada, y las conclusiones que se obtengan de dicha muestra sólo podrán referirse a la población en referencia.

La muestra debe ser representativa ya que debe contener las características relevantes de la población en las mismas proporciones en que están incluidas en tal población, es decir, contiene las características más importantes de esa colección de elementos que representan la población bajo estudio, a fin de investigar alguna o algunas de las propiedades de la población de la cual procede; y las conclusiones que se obtengan de dicha muestra sólo podrán referirse a la población en referencia.

La muestra es el elemento básico sobre el cual se fundamenta la posterior inferencia acerca de la población de donde se ha tomado. Por ello, su escogencia y selección debe hacerse siguiendo ciertos procedimientos o parámetros que son indispensables, es decir, se selecciona de acuerdo con una regla o plan definido. En estadística, en vez de estudiar las poblaciones en su totalidad, se acude al recurso de considerar solamente una parte de ella, a la cual se le denomina muestra. Por lo tanto, una muestra es una parte de la población, seleccionada de acuerdo con una regla o plan definido.

**Muestreo:** Es el procedimiento mediante el cual se obtiene una o más muestras de una población determinada. Existen dos tipos de muestreos a saber:

**Muestreo no Probabilístico:** Es aquel en el cual se toma la muestra según el criterio del investigador, estos pueden ser: *muestreo intencional u opinático* y el *muestreo sin norma o circunstancial*.

**Muestreo Probabilístico:** Es aquel que se selecciona utilizando métodos aleatorios en los que se utilizan las probabilidades matemáticas. Entre estos se pueden mencionar: *Muestreo aleatorio con reemplazamiento*, *muestreo aleatorio sin reemplazamiento*, *muestreo estratificado*, *muestreo por conglomerado o por área* y *muestreo aleatorio simple*.

**Los Parámetros.-** Son cualquiera característica que se pueda medir y cuya medición se lleve a cabo sobre todos los elementos que integran una población determinada, los mismos suelen representarse con letras griegas. El valor de un parámetro poblacional es un valor fijo en un momento dado. Ejemplo: La media Aritmética =  $\mu$  (miu), La desviación Típica =  $\sigma$ , (Sigma) etcétera.

**Los Estadígrafos (Estadísticos o Estimadores).-** Son aquellas características medibles, cuya medición se realiza sobre los componentes de una muestra, los mismos se representan con letras del alfabeto castellano. Los estadígrafos no tienen un valor único, sino que pueden tomar distintos valores al ser calculados a partir de muestras diferentes. Ejemplos: la media aritmética =  $\bar{X}$ , La desviación Típica = S.

**MEDICIÓN:** La asignación o magnitud que se aplica a las categorías o clases de acuerdo a ciertas reglas o símbolos. Consiste en la recopilación de datos y su utilización mediante el empleo de una serie de normas de tipo estadístico; es la representación simbólica de un dato o serie de datos obtenidos por algún tipo de observación.

## ESCALAS DE MEDICIÓN DE LAS VARIABLES

Una escala de medición es una asignación de valores numéricos a las características de una muestra o una población, se mide básicamente en proporción de escala. Las escalas de medición son el conjunto de los posibles valores que determinada variable puede tomar. Por tal razón, los tipos de escala de medición están íntimamente ligadas con los tipos de variables a estudiar. Las magnitudes de las observaciones cuantitativas se conocen como los valores que una variable puede asumir. Consiste entonces, en una serie de graduaciones que permiten darle un valor numérico a las características que estamos midiendo; para hacer más comprensible y que adquiera un significado mediante un arreglo ordenado para establecer un análisis estadístico. Son denominaciones o clasificaciones de individuos o características. Las escalas de medición es el alma fundamental de toda investigación Científica, puesto que, sólo a través de ellas es como se pueden calibrar los fenómenos, sus relaciones, entre otros.

Se refieren habitualmente a las asignaciones de números a observaciones, de una forma tal que los números sean susceptibles de análisis por medio de manipulaciones u operaciones; estas escalas permiten asignarle un valor numérico a las características que se están midiendo. Por lo general proceden de las medidas de una o más variables. Dependiendo de la medición y de la esencia de las variables, se obtienen diversas clases de datos que originan diferentes escalas. Resulta intensamente conocer el tipo de escala que representan los datos, debido a que, de su esencia dependen las técnicas estadísticas que se deberán aplicar para su análisis.

Para lograr estadísticas confiables hay que manipular cuantiosos datos estadísticos, los cuales poseen determinadas características. Por ejemplo, si los datos son alumnos, entre algunas de ellos se puede señalar el peso, la estatura, el sexo, el rendimiento académico, entre otros. Al elaborar estadísticas con

datos y su característica es necesario contarlas, jerarquizarlas y medirlas, es por ello que, se utilizan las escalas de medición como el proceso de asignar números o establecer una correspondencia uno a uno entre objetivos y observaciones. Las escalas de medición se clasifican de la siguiente forma: *Escala Nominal, Escala Ordinal, Escala de Intervalos y Escalas de razón o Proporción.*

**Escala de Medición Nominal:** Es aquella en la que los números sólo se emplean para diferenciar los objetos de distintas categorías o cuando se emplean nombres. Se dice que los datos que se obtienen para una variable cualitativa se miden en una escala nominal. Si los datos observados simplemente se clasifican en distintas categorías que no implican orden, se tiene un nivel de medición nominal. Ejemplos de números, esta característica son las que usan los jugadores de béisbol, los números telefónicos, los números de las Cédulas de Identidad, etcétera. Se usa una escala nominal cuando se distribuyen conjuntos de objetos, personas o características entre dos o más categorías. La Escala Nominal se utiliza como medida de identidad. Los números pueden servir como indicativos o etiquetas para identificar objetos o clases, pues se usa cuando un objeto se diferencia de otro solamente por la nominación con que se conoce. La escala nominal es la forma más débil de medición porque no se puede intentar el conteo de las diferencias dentro de una categoría determinada o especificar cualquier orden o dirección a lo largo de las diversas categorías. Sin embargo, no se intenta medir diferencias entre los valores clasificados dentro de una categoría determinada.

### Propiedades de la escala Nominal

1. No intervienen mediciones, ni escala, en vez de esto solo hay cuentas o conteos. Esta escala es considerada excluyente, es decir que la persona u objeto se incluye solamente en una categoría.
2. No existe un orden específico para esta categoría.
3. No presentan el cero.
4. No se basa en diferencia cuantitativa.
5. Los elementos de una categoría deben de ser equivalentes, idénticos.

**Ejemplos:** Una muestra de personas puede clasificarse con base en la religión profesada: (1) Cristianos; (2) Judíos; (3) Musulmanes; (4) Otros; y (5) Sin Creencia alguna. O bien podrían clasificarse según el sexo, el color de los ojos, algún partido político, etcétera. Otros ejemplos de escala nominal puede ser el número de placa de los vehículos, los números de los teléfonos de una ciudad, la Cédula de Identidad de los habitantes de un país, etcétera.

El tipo de operación estadística más utilizada en la escala nominal es el conteo de las frecuencias con que se presentan las características en las unidades de las respectivas subclases. Estas frecuencias pueden ser presentadas con números absolutos, porcentajes y proporciones. Además, puede calcularse razones, tasas de incremento, y el coeficiente de contingencia.

**Escala de Medición Ordinal:** Es aquella en la que los números se utilizan para diferenciar en orden de supremacía de acuerdo con cierto criterio jerárquico, como son los números que se emplean para clasificar los distintos estratos socio-económicos o para designar preferencias. Si los datos observados se clasifican en categorías distintas en las que existe algún orden, se obtiene un nivel de medición ordinal. Cuando los objetos son medidos en escala ordinal los que tengan la misma asignación se consideran iguales; pero los que tengan asignaciones diferentes pertenecen a categorías distintas. La

diferencia entre dos números ordinales no tiene significado cuantitativo, sólo expresan, por ejemplo, que una situación es mejor que otra, pero no cuanto. La escala ordinal es una forma un tanto más fuerte de medición que la nominal, porque se dice que un valor observado que se clasifica en una categoría posee más la propiedad que se mide que algún valor observado que se clasifica en otra categoría.

También, la escala ordinal sigue siendo una forma de medición débil porque no se pueden hacer planteamientos numéricos significativos con respecto a las diferencias entre las categorías. Es decir, la ordenación establece sólo cuál categoría es “mayor”, “mejor” o “preferida”; y no se habla cuánto es “mayor”, “mejor” o más “preferida”. Esta escala se emplea, cuando un estudio está basado en ciertas normas que se asignan a un conjunto de objetos, personas o características o a un conjunto de categorías ordenadas. Las categorías de la escala se ordenan de acuerdo con la cantidad de rasgos o características que representan cada una de ellas ya que la escala ordinal distingue los diferentes valores de la variable, ubicando las características en orden, desde la más alta hasta la más baja.

### Propiedades de la escala Ordinal

1. Las observaciones o elementos se les ordena en rangos o categorías diferentes.
2. Cada categoría o rango mantiene una relación entre sí, estas relaciones se expresan en términos algebraicos de desigualdades (mayor que o menor que).
3. No es posible definir unidades de mediciones iguales en todos los puntos de la escala.
4. Las categorías son mayores o menores que otras categorías, es decir, que existe una clasificación de mayor a menor (jerarquía).
5. Las categorías son mutuamente excluyentes y exhaustivas.
6. No presentan el cero.

### Ejemplo de escala Ordinal:

---

Calificaciones de Estudiantes de Estadística en el Semestre 2003

---

Calificaciones	Nº de calificaciones
Excelente	3
Sobresaliente	6
Distinguido	10
Bueno	38
Satisfactorio	25
Deficiente	50
Muy Deficiente	0

Se pueden clasificar los habitantes de una ciudad de acuerdo a su situación económica, a los estudiantes tomando en cuenta a su rendimiento académico, etcétera. Aunque la escala ordinal resulta en cierta forma más precisa que la nominal, no alcanza el grado de precisión deseado en una investigación.

Otro ejemplo, sería el caso de cuatro estudiantes de administración de empresas, que presentan una prueba Extraordinaria de estadística y obtienen las siguientes calificaciones: 19,16, 12 y 10. El estudiante que obtuvo 19 puntos se le denomina como Excelente, al que obtuvo 16 puntos como

Distinguido, el que logró 12 tendrá una categoría de Bueno y al que obtuvo 10 puntos se le asigna la categoría de Satisfactorio. Es decir, los alumnos han sido clasificados de acuerdo con su rendimiento académico.

El orden jerárquico de los militares (Subteniente, Teniente, Capitán, Mayor, Teniente Coronel, Coronel, General) y la clasificación académica de los profesores universitarios( Instructor, Asistente, Agregado, Asociado, Titular) son ejemplos de escala ordinal.

En la escala ordinal las unidades de las subclases guardan una cierta relación entre sí, esto se pone de manifiesto cuando existe la posibilidad de establecer la relación menor que o mayor que, respecto a las características de las unidades escaladas. Por ejemplo, El grado militar de A es el de Subteniente y el de B es el de Teniente, luego el grado de B es mayor que el de A ( $B > A$ )

Otro ejemplo de medición con la escala ordinal es el referido al de la escala de dureza de los minerales, es decir, la resistencia que oponen los minerales al ser rayados, los cuales van del uno al diez. El talco se asocia con el valor uno porque no raya a ninguno, el diamante se asocia con el valor diez porque no es rayado por ningún otro, pero sí puede rayar al resto de los minerales, etcétera.

**Escala de Medición de Intervalos:** Es una escala más especializada que las dos anteriores, puesto que es posible ordenar las mediciones y expresar además en cuánto difiere una situación de la otra. Por Ejemplo, en las mediciones de temperatura ambiental no sólo se puede afirmar “*hoy hace más calor que ayer*”, si no que de la misma forma se puede expresar “*hoy la temperatura es cuatro grados Centígrados más alta que la de ayer a la misma hora*”. Esta escala se caracteriza por tener una unidad de medida y un origen (cero) arbitrario. La distancia entre dos mediciones tiene un significado preciso. La escala de intervalos a diferencia de la nominal y ordinal, es una escala efectivamente cuantitativamente. Una escala de intervalo es una escala ordenada en la cual la diferencia entre las mediciones es una cantidad significativa.

La escala de intervalos posee además, de las características de la escala nominal y ordinal, la propiedad de que la distancia entre dos valores es de una magnitud conocida, lo que le permite a esta escala un mayor grado de perfección, ya que proporciona números que manifiestan diferencias palpables entre individuos, objetos o cosas. Por tal razón, la escala de intervalo revela que un individuo u objeto es tantas unidades más grandes o más pequeño, más pesado ó más ligero, más rápido o más lento que otro, es decir, muestra la cantidad en la que un objeto se diferencia de otro cuantitativamente. En esta escala el punto cero y la unidad de medición son arbitrarios. La razón entre dos intervalos es siempre independiente del punto cero y de la unidad que se emplee en la medición. En el caso de las escalas de intervalos las unidades de medición son iguales.

#### **Propiedades de la escala de Intervalo:**

1. Esta escala implica la cuantificación de los datos
2. En estas medidas se utilizan unidades constantes de medición (capacidad, peso, Céntimos, grados fahrenheit o centígrados) los cuales producen intervalos iguales entre puntos de la escala.
3. Proporcionan números que manifiestan diferencias palpables entre individuos, objetos o cosas.
4. En esta escala de intervalos el punto cero (0) y la unidad de medida es arbitrario.



5. Se pueden aplicar todas las medidas estadísticas más conocidas, con excepción del **coeficiente de variación**.
6. Son mutuamente exclusivas y exhaustivas.

### Ejemplos de las escalas de intervalo.

Calificación de una prueba de Estadísticas realizada en la Universidad

Puntuación	Nº de Estudiantes
90-99	2
80-89	6
70-79	15
60-69	30
Menos de 60	60

Otro ejemplo de esta escala lo constituyen las escalas utilizadas para medir temperatura, bien sea en grados Centígrados o Fahrenheit. En estas escalas la diferencia entre 80° y 85° es igual a la que existe entre 90° y 95° o entre dos puntos cualesquiera de la escala. La escala de intervalo tiene carácter cuantitativo y esto le permite el cálculo de las medidas estadísticas más comunes (medias, desviaciones típicas o estándar, coeficientes de correlación de Pearson, entre otros), esto confirma el por que muchos valores estadísticos se utilizan con las escalas de intervalos.

**Escala de Medición de Razón o Proporción:** Esta constituye el nivel más alto de medición, posee todas las características de las escalas nominales, ordinales y de intervalos; además tiene un cero absoluto o natural que tiene significado físico. Si en ella la medición es cero, significa ausencia o inexistencia total de la propiedad considerada. Son posibles todas las operaciones aritméticas. Los números indican los valores concretos de la propiedad que se está midiendo; peso, estatura, ingresos monetarios y gastos directos, son ejemplos de medidas con una escala de razón.

La distancia entre dos valores de la escala es conocida en el sentido cuantitativo y su razón es independiente de las unidades empleadas. Por ejemplo, en las unidades de longitud, peso y capacidad el valor cero indica ausencia de medida, mientras que en la escala de intervalo si se tiene cero grado centígrado no se puede afirmar que hay ausencia de temperatura. La escala de razón permite establecer ciertas comparaciones entre los valores que no son permitidos en la escala de intervalo. Por lo tanto, la proporción de un punto cualquiera de la escala a otro es independiente de la unidad de medida. Si una persona mide 2.00 m puede afirmarse que duplica en estatura a otro que mide 1.00 m. Las escalas de razón más comunes corresponden a medidas de longitud, peso, capacidad, sonido, entre otros. Al medir la temperatura absoluta la escala de Kelvin, que es de este tipo, tiene su punto cero a 273°, este valor es el más bajo posible. Cuando se emplea este tipo de escala, los números indican razones o cocientes entre ciertas magnitudes de los objetos, y los datos obtenidos con tales escalas pueden ser sometidos a los tratamientos más elaborados.

### Propiedades de la escala a razón:

1. La distancia entre los números es un tamaño conocido y constante.
2. Los datos tienen un punto cero significativo.

3. Puede utilizarse cualquier prueba de tipo estadístico, incluyendo el **coeficiente de variación**.
4. Permite hacer comparaciones entre los números verdaderos con un cero aritmético siendo arbitrario únicamente la unidad de medida.

### **Ejemplo de escala a razón:**

Número de computadores vendidos en el último trimestre del año 1999	
Meses	Nº de computadores
Octubre	4.000
Noviembre	6.000
Diciembre	15.000

Generalmente, se supone que los datos que se obtienen para una variable cuantitativa se miden en escalas de intervalo o de razón. Estas escalas constituyen los niveles más elevados de medición. Son formas más fuertes de medición que la escala nominal y ordinal, porque permiten comprender no sólo cuál de los valores de un estudio es mayor o menor, sino por cuántas unidades de medida. Las escalas de razón son medidas de Longitud, peso, capacidad, etc., los números reflejan razones entre particularidades y los datos obtenidos según tales escalas pueden ser sometidas a cualquier tratamiento estadístico

## **INSTRUMENTOS DE MEDICIÓN DE LAS VARIABLES**

Son las herramientas que se manipulan para obtener información y para llevar a cabo las observaciones de una investigación o estudio determinado. Conforme a lo que se desea estudiar o investigar, la característica a observar, sus propiedades y factores relacionados con aspectos naturales, económicos, políticos, sociales, etc., cuando se selecciona uno de estos instrumentos. En otras palabras, estos son los que permiten efectuar observaciones, de uno u otro fenómeno, en una forma más despejada y precisa de la descripción de los hechos a estudiar.

**La encuesta, el cuestionario, la entrevista, la observación y las escalas de actitudes y opiniones,** constituyen técnicas o herramientas que se utilizan para medir las variables, las mismas deben reunir dos características fundamentales:

### **a) Validez:**

Esta característica se refiere a que la calificación o resultado obtenido mediante la aplicación del instrumento, mida lo que realmente se desea medir. La validez de contenido puede definirse como la propiedad que posee el instrumento para medir todos los factores de la variable que se están estudiando. Para determinar los parámetros de la validez de contenido de un instrumento, es necesario:

Definir operativa y teóricamente las variables que se van a medir.

Programar todas las formas en que esta variable se puede presentar para establecer los indicadores más adecuados, para ello es necesario haber realizado una amplia revisión bibliográfica y consultado a especialistas en la materia.

Efectuar una prueba piloto que contribuya a mejorar la validez del instrumento. La validez de predicción esta relacionada con la eficacia que alcanza la técnica para predecir el comportamiento de los fenómenos ante determinadas circunstancias. Se puede verificar comparando el resultado obtenido a

través de la aplicación del instrumento, si predice el comportamiento del fenómeno en estudio, entonces, es conveniente compararlo con el estándar o prueba piloto con los resultados obtenidos en las pruebas aplicadas a la muestra objeto del estudio. con los resultados obtenidos en el desempeño del trabajo.

## **b) Confiabilidad:**

La confiabilidad está referida a la estabilidad, consistencia y exactitud de los resultados, es decir, que los resultados obtenidos por el instrumento serán confiables siempre y cuando sean similares a los resultados que se obtengan si se vuelven a aplicar el instrumento sobre la misma muestra en igualdad de condición.

Dentro de la clasificación de los instrumentos de medición se pueden considerar básicamente los siguientes: la **observación**, la **encuesta** (que utiliza **cuestionarios**) y la **entrevista**. Para utilizar alguno de estos instrumentos de medición es indispensable que se cumplan con las siguientes condiciones, en la que, por lo general coinciden los tres:

1. **Definir el objeto de la encuesta:** Formulando con precisión los objetivos a conseguir, describiendo el problema a investigar, eliminando lo superfluo y centrando el contenido de la encuesta, y diseñando la muestra. Se incluye la forma de presentación de resultados así como los costos de la investigación.
2. **La formulación del cuestionario** que se utilizará o **de los puntos a observar** es fundamental en el desarrollo de una investigación, debiendo ser realizado meticulosamente y comprobado antes de pasarlo a la muestra representativa de la población.
3. **El trabajo de campo**, consistente en la obtención de los datos. Para ello será preciso seleccionar a los entrevistadores, formarlos y distribuirles el trabajo a realizar de forma homogénea.
4. **Obtener los resultados**, o sea, procesar, codificar y tabular los datos obtenidos para que luego sean presentados en el informe y que sirvan para posteriores análisis.

## **LA OBSERVACIÓN**

Es la técnica de recolección de información por excelencia y se utiliza en todas las ramas de la ciencia. Su uso está regido por alguna teoría y éstas determinan los aspectos que se van a observar. Hay que tener presente que para que sea válido este instrumento de observación, se deben cumplir las siguientes sugerencias:

1. Con respecto a las condiciones previas a la observación:
  - El observador debe estar familiarizado con el medio.
  - Se deben realizar ensayos de la observación, previos a la observación definitiva.
  - El observador debe memorizar lo que se va a observar.
2. Con respecto al procedimiento en la observación:

- Las notas deben ser registradas con prontitud (en minutos).
  - Las notas deben incluir las acciones realizadas por el observador.
3. Con respecto al contenido de las notas:
- Las notas deben contener todos los datos que permitan identificar el día, el lugar y la hora de la observación, así como las circunstancias, los actores, etcétera, que estuvieron involucrados.
  - Se deben eliminar apreciaciones subjetivas sobre el carácter o personalidad de los sujetos. En su lugar se debe incluir la descripción de los hechos.
  - Las conversaciones van transcritas en estilo directo.
  - Las opiniones y deducciones del observador se deben hacer aparte, de preferencia al margen para así no perder la relación entre la opinión del observador y la parte de las notas a que le corresponde.
4. Con respecto a la ordenación de las notas:
- Las notas deben ser revisadas y corregidas a la brevedad posible.
  - Asimismo, las notas deben ser clasificadas y ordenadas para permitir su manejo más ágil, además de evitar que se pierdan, se confundan con otras partes de la observación, se traspapelen, etcétera.

## LA ENCUESTA

Esta es una de las herramientas más utilizada en la investigación de ciencias sociales. Para su implementación, la encuesta utiliza los cuestionarios como medio principal para obtener información. De esta manera, las encuestas pueden realizarse para que el individuo encuestado procese por sí mismo las respuestas en el papel.

Es trascendente que el investigador en los cuestionarios sólo solicite la información indispensable, la mínima para que sean comprendidas las preguntas. Más información, o información innecesaria, puede derivar en respuestas no veraces.

De la misma forma, al diseñar la encuesta y confeccionar el cuestionario hay que tomar en cuenta los recursos (tanto humanos como materiales) de los que se disponen, tanto para la recopilación como para la lectura de la información, para así lograr un diseño funcionalmente eficaz.

Según M. García Ferrando, "prácticamente todo fenómeno social puede ser estudiado a través de las encuestas", y podemos considerar las siguientes cuatro razones para sustentar éstos:

1. Las encuestas son una de las escasas técnicas de que se dispone para el estudio de las actitudes, valores, creencias y motivos.
2. Las técnicas de encuesta se adaptan a todo tipo de información y a cualquier población.
3. Las encuestas permiten recuperar información sobre sucesos acontecidos a los entrevistados.
4. Las encuestas permiten estandarizar los datos para un análisis posterior, obteniendo gran cantidad de datos a un precio bajo y en un período de tiempo corto.

Según Cadoche y sus colaboradores, las encuestas se pueden clasificar atendiendo al ámbito que abarcan, a la forma de obtener los datos y al contenido, así:

**Encuestas exhaustivas y parciales:** Se denomina **exhaustiva** cuando abarca a todas las unidades estadísticas que componen el colectivo, universo, población o conjunto estudiado. Cuando una encuesta no es exhaustiva, se denomina **parcial**.

**Encuestas directas e indirectas:** Una encuesta es **directa** cuando la unidad estadística se observa a través de la investigación propuesta registrándose en el cuestionario. Será **indirecta** cuando los datos obtenidos no corresponden al objetivo principal de la encuesta pretendiendo averiguar algo distinto o bien son deducidos de los resultados de anteriores investigaciones estadísticas.

**Encuestas sobre hechos y encuestas de opinión:** Las **encuestas de opinión** tienen por objetivo averiguar lo que el público en general piensa acerca de una determinada materia o lo que considera debe hacerse en una circunstancia concreta. Se realizan con un procedimiento de muestreo y son aplicadas a una parte de la población ya que una de sus ventajas es la enorme rapidez con que se obtienen sus resultados. No obstante, las encuestas de opinión no indican necesariamente lo que el público piensa del tema, sino lo que pensaría si le plantease una pregunta a ese respecto, ya que hay personas que no tienen una opinión formada sobre lo que se les pregunta y contestan con lo que dicen los periódicos y las revistas. A veces las personas encuestadas tienen más de una respuesta a una misma pregunta dependiendo del marco en que se le haga la encuesta y por consecuencia las respuestas que se dan no tienen por qué ser sinceras.

Las **encuestas sobre hechos** se realizan sobre acontecimientos ya ocurridos, hechos materiales.

## EL CUESTIONARIO

El cuestionario es un formato redactado en forma de interrogatorio con el mismo se obtiene información relacionada con las variables objeto de la investigación. Pueden ser aplicados personalmente o por correo y en forma individual o colectiva. Esta formado por un conjunto de preguntas elaboradas cuidadosamente sobre los hechos y aspectos que se desean conocer sobre una población o parte de ella; este instrumento es respondido por el participante sin la intervención directa del entrevistador. En el cuestionario simple el encuestado contesta, previa lectura del escrito, sin intervención directa de ninguna de las personas que participa en la investigación.

En la entrevista, el cuestionario es aplicado a los sujetos investigados, por personas especializadas en esa tarea. Estas hacen a los encuestados las preguntas del cuestionario y anotan en el las respuestas.

*Las escalas* son una forma especial de cuestionario; se caracteriza porque las preguntas y sus diferentes respuestas tienen atribuido un valor numérico, lo que permite cifrar cuantitativamente y en cierta forma medir el nivel que alcanza en cada caso la actitud o aspecto investigado.

El diseño del cuestionario habrá de fundamentarse en el marco teórico, la hipótesis, sus variables y los objetivos de la investigación. Cada pregunta que se incluya deberá estar relacionada con las variables indicadoras. Es muy conveniente que cuando se elabore el cuestionario se tenga a la mano la operatividad de las variables, para asegurarse de que todos los indicadores están siendo investigados.

### Los cuestionarios pueden ser:

**Cuestionario individual:** Es el que el encuestado contesta de forma individual por escrito y sin que intervenga para nada el encuestador.

**Cuestionario-lista:** Es aquel que es preguntado al encuestado en una entrevista por uno de los especialistas de la investigación.

Como los cuestionarios están formados por preguntas, se considera que una de las características básicas que estos deben reunir es la ser **excluyentes** y **exhaustivas**, lo que permitirá que una pregunta no produzca dos respuestas simultáneamente, por lo tanto, a cada pregunta solamente le corresponderá una y sólo una respuesta.

Por otro lado, una manera de clasificar las preguntas, es por la forma de su respuesta, así:

**Preguntas cerradas:** que consiste en proporcionar al sujeto observado una serie de opciones para que escoja una como respuesta. Tienen la ventaja de que pueden ser procesadas más fácilmente y su codificación se facilita; pero también tienen la desventaja de que si están mal diseñadas las opciones, el sujeto encuestado no encontrará la opción que él desearía y la información se viciaría. Una forma de evitarlo es realizar primero un estudio piloto y así obtener las posibles opciones para las respuestas de una manera más confiable.

También se consideran cerradas las preguntas que contienen una **lista de preferencias** u **ordenación de opciones**, que consiste en proporcionar una lista de opciones al encuestado y éste las ordenará de acuerdo a su interés, gustos, etcétera.

**Preguntas abiertas:** que consisten en dejar totalmente libre al sujeto observado para expresarse, según convenga. Tiene la ventaja de proporcionar una mayor riqueza en las respuestas; mas, por lo mismo, puede llegar a complicar el proceso de tratamiento y codificación de la información. Una posible manera de manipular las preguntas abiertas es llevando a cabo un proceso de **categorización**, el cual consiste en estudiar el total de respuestas abiertas obtenidas y clasificarlas en categorías de tal forma que respuestas semejantes entre sí queden en la misma categoría.

Es importante señalar que *es el objetivo de la investigación la que determina el tipo de preguntas a utilizar*. Según Cadoche y sus colaboradores, las preguntas pueden ser clasificadas de acuerdo a su contenido:

**Preguntas de identificación:** edad, sexo, profesión, nacionalidad, etcétera.

**Preguntas de hecho:** referidas a acontecimientos concretos. Por ejemplo: ¿terminó la educación básica?

**Preguntas de acción:** referidas a actividades de los encuestados. Por ejemplo: ¿ha tomado algún curso de capacitación?

**Preguntas de información:** para conocer los conocimientos del encuestado. Por ejemplo: ¿sabe qué es un hipertexto?

**Preguntas de intención:** para conocer la intención del encuestado. Por ejemplo: ¿utilizará algún programa de computación para su próxima clase?

**Preguntas de opinión:** para conocer la opinión del encuestado. Por ejemplo: ¿qué carrera cursarás después del bachillerato?

Existe otra clasificación de los cuestionarios que toma en la función que las preguntas desarrollaran dentro del cuestionario. De esta manera tenemos:

**Preguntas filtro:** son aquellas que se realizan previamente a otras para eliminar a los que no les afecte. Por ejemplo: ¿Tiene usted coche? ¿Piensa comprarse uno?

**Preguntas trampa o de control:** son las que se utilizan para descubrir la intención con que se responde. Para ello se incluyen preguntas en diversos puntos del cuestionario que parecen independientes entre sí, pero en realidad buscan determinar la intencionalidad del encuestado al forzarlo a que las conteste coherentemente (ambas y por separado) en el caso de que sea honesto, pues de lo contrario «caería» en contradicciones.

**Preguntas de introducción o rompehielos:** utilizadas para comenzar el cuestionario o para enlazar un tema con otro.

**Preguntas muelle, colchón o amortiguadoras:** son preguntas sobre temas peligrosos o inconvenientes, formuladas suavemente.

**Preguntas en batería:** conjunto de preguntas encadenadas unas con otras complementándose.

**Preguntas embudo:** se empieza por cuestiones generales hasta llegar a los puntos más esenciales.

Para la elaboración de un cuestionario eficaz y útil, Cadoche y su equipo proponen **17 reglas fundamentales** para su confección:

1. Las preguntas han de ser pocas (no más de 30).
2. Las preguntas preferentemente cerradas y numéricas.
3. Redactar las preguntas con lenguaje sencillo.
4. Formular las preguntas de forma concreta y precisa.
5. Evitar utilizar palabras abstractas y ambiguas.
6. Formular las preguntas de forma neutral.
7. En las preguntas abiertas no dar ninguna opción alternativa.
8. No hacer preguntas que obliguen a esfuerzos de memoria.
9. No hacer preguntas que obliguen a consultar archivos.
10. No hacer preguntas que obliguen a cálculos numéricos complicados.
11. No hacer preguntas indiscretas.
12. Redactar las preguntas de forma personal y directa. }
13. Redactar las preguntas para que se contesten de forma directa e inequívoca.
14. Que no levanten prejuicios en los encuestados.
15. Redactar las preguntas limitadas a una sola idea o referencia.
16. Evitar preguntas condicionantes que conlleven una carga emocional grande.
17. Evitar estimular una respuesta condicionada. Es el caso de preguntas que presentan varias respuestas alternativas y una de ellas va unida a un objetivo tan altruista que difícilmente puede uno negarse.

Asimismo, hay que considerar que no todas las preguntas, o todas las formulaciones, son posibles de utilizar, algunas de las preguntas que *no* deben hacerse son:

**Preguntas de intelectuales:** Por ejemplo: ¿Qué aspectos particulares del actual debate positivista-interpretativo le gustaría ver reflejados en un curso de psicología del desarrollo dirigido a una audiencia de maestros?

**Preguntas complejas:** Por ejemplo: ¿Cuando prepara sus clase prefiere consultar un libro determinado incorporando la terminología que este propone o escoge varios libros de los que extrae un poco de cada uno pero que explica con sus propias palabras para hacerlos más accesibles a sus alumnos y no confundirlos?

**Preguntas o instrucciones irritantes:** Por ejemplo: ¿Ha asistido alguna vez en tiempo de servicio a un curso de cualquier clase durante su carrera entera de maestro?. Si tiene mas de 40 años y nunca ha asistido a un curso, ponga una marca en la casilla rotulada NUNCA y otra en la casilla rotulada VIEJO.

**Preguntas que emplean negaciones:** Por ejemplo: ¿Cuál es su sincera opinión sobre que ningún maestro debería dejar de realizar cursos de perfeccionamiento durante su ejercicio profesional?

**Preguntas demasiado abiertas:** Por ejemplo: Use las pág. 5, 6 y 7 respectivamente para responder a cada una de las cuestiones a cerca de sus actitudes respecto a los cursos de perfeccionamiento en general y a sus opiniones a cerca de su valor en la vida profesional del maestro.

Tomando en cuenta lo ante mencionados se puede comprender la importancia del planteamiento y la elaboración del cuestionario, ya que de la forma en que se elabore este instrumento y se redacten las formulaciones, dependerán los resultados de la investigación que se desea realizar. Una encuesta no puede obtener buenos resultados con un cuestionario mal elaborado, puesto que si el cuestionario es oscuro, ambiguo o impreciso, los resultados obtenidos también lo serán y no se podrá confiar en los resultados que se obtengan para emprender la investigación.

Castañeda Jiménez sugiere que se tomen en cuenta las siguientes previsiones en la elaboración de un cuestionario:

1. En la **elaboración** o **construcción** del instrumento:
  - Hay que determinar los reactivos de acuerdo a lo que se necesita observar.
  - Hay que determinar el orden de los reactivos de acuerdo a los aspectos que se mencionan más adelante.
  - Se debe tener cuidado en la formulación de los reactivos. Una formulación incorrecta o diferente puede dar lugar a interpretaciones diferentes por parte del entrevistado a las que el observador desea.
2. Respecto al **orden** de los reactivos:
  - Es conveniente situar los reactivos que sean más difíciles de ser contestado honestamente al final, de esta manera no se desanimará de antemano el entrevistado.
  - Otra opción es repetir dos o tres reactivos que posean la misma información pero con diferente redacción. Estos reactivos de control permitirán detectar cuándo el entrevistado está contestando honestamente.
3. Respecto a la **redacción** de los reactivos:
  - La redacción, y el vocabulario, debe estar acorde a la persona observada, tomando en cuenta su edad, nivel cultural, nivel escolar, nivel socio-económico, etcétera.
  - Cada reactivo debe contener una y sólo una pregunta.



- En la redacción de la pregunta no debe estar sugerida alguna de las respuestas.
  - Tampoco conviene apoyarse o mencionar opiniones o sugerencias ya existentes, como son posiciones de instituciones, de personas, etcétera.
- Asimismo, Cadoche y sus colegas proponen una guía para preparar un cuestionario:

### **Decisiones sobre el contenido de las preguntas:**

- ¿Es necesaria la pregunta? ¿Será útil?
- ¿Se necesitan varias preguntas sobre esta cuestión?
- ¿Cuentan los informantes con los datos necesarios para contestar la pregunta?
- ¿Necesita la pregunta ser más concreta, específica e íntimamente ligada con la experiencia personal del informante?
- ¿Es el contenido de la pregunta lo suficientemente general y está libre de concreciones y especificidades falsas?
- ¿Expresan las preguntas actitudes generales y son tan específicas como suenan?
- ¿Está el contenido de la pregunta polarizado o cargado en una dirección sin preguntas acompañantes que equilibren el énfasis?
- ¿Darán los informantes la información que se les pide?

### **Decisiones sobre la redacción de las preguntas:**

- ¿Se puede malinterpretar la pregunta? ¿Contiene fraseología difícil o poco clara?
- ¿Expresa la pregunta adecuadamente la alternativa con respecto al punto?
- ¿Es engañosa la pregunta por culpa de asunciones no establecidas o de implicaciones que no se ven?
- ¿Está polarizada la redacción? ¿Está cargada emocionalmente o inclinada hacia un tipo particular de contestación?
- ¿Puede ser objetable por el informante la redacción de la pregunta?
- ¿Produciría mejores resultados una redacción más personalizada de la pregunta?
- ¿Puede preguntarse mejor la cuestión, de manera más directa o más indirecta?

### **Decisiones sobre la forma de respuesta de la pregunta:**

- ¿Puede contestarse mejor la pregunta con un impreso que exija la contestación por una marca (o contestación corta de una o dos palabras, o un número), de respuesta libre o por una marca con contestación ampliatoria?
- Si se usa la contestación por una marca, ¿cuál es el mejor tipo de cuestión: dicotómica, de elección múltiple, o de escala?
- Si se usa una lista de comprobación, ¿cubre adecuadamente todas las alternativas significativas sin solaparse y en un orden definible? ¿Es de una longitud razonable? ¿Es la redacción de los ítems imparcial y equilibrada?
- ¿Es fácil, definida, uniforme y adecuada para la finalidad, la forma de respuesta?

## Decisiones sobre la ubicación de la pregunta en la secuencia:

- ¿Puede verse influida por el contenido de las cuestiones precedentes la contestación a la pregunta?
  - ¿Está dirigida la pregunta en una forma natural? ¿Está en correcto orden psicológico?
- ¿Aparece la pregunta demasiado pronto o demasiado tarde desde el punto de vista de despertar interés y recibir la atención suficiente?

## LA ENTREVISTA

La entrevista es muy utilizada en investigación social, y sus características son similares a las del cuestionario, siendo la principal diferencia el hecho de que es el encuestador u observador quien anota las respuestas a las preguntas.

La utilización de este instrumento requiere de una mayor habilidad por parte del encuestador u observador para llevar el tema de la entrevista, debido a que las respuestas son por lo general abiertas y admiten implementar nuevas preguntas no vislumbradas por el encuestador inicialmente. Esto facilita la ventaja de explotar temas no contemplados inicialmente o ahondar en algunos de los contemplados. No obstante, tiene la desventaja de que, si no se tiene la suficiente habilidad para mantener el tema, la entrevista se "pierde" e, incluso, puede invalidarse; por lo tanto, el entrevistador debe poseer aptitudes específicas para utilizar el arte de la entrevista

Las recomendaciones y características sugeridas, son las mismas que se utilizan para el caso del cuestionario, sin embargo, se debe utilizar una grabadora (de audio o de vídeo) para la posterior copia de los diálogos.

Los tipos fundamentales son:

**-La entrevista cerrada:** Es aquel tipo en la que las alternativas de contestación a que debe someterse el encuestado están predeterminadas.

**-La entrevista con profundidad:** Es aquella que se hace conscientes los contenidos mentales transformándolos en profundos.

**-La entrevista semiestructurada:** Es aquella en la que, si bien hay una guía para las preguntas, las respuestas son libres, y su ventaja radica en que permiten obtener información complementaria.

La entrevista es una de las técnicas más utilizada en la investigación. Mediante ésta, una persona, el entrevistador solicita información a otra, el entrevistado. La entrevista puede ser uno de los instrumentos más valiosos para obtener información y aunque aparentemente no necesita estar muy preparada, es posible definirla como el arte de escuchar y captar información, esta habilidad requiere de capacitación, ya que no toda persona puede ser un buen entrevistador. Además, es una manera de interactuar socialmente puesto que es a través del diálogo como el investigador obtiene los datos que requiere para su estudio.

## ESCALA DE ACTITUDES Y OPINIÓN

**La escala de actitudes y opinión,** son instrumentos que se utilizan para medir la intensidad de las actitudes y opiniones de una población hacia un fenómeno determinado. Se llaman escalas porque se forman de un continuo de valores que tienen diversos puntos intermedios. Una actitud se define como el grado de afecto positivo o negativo asociado a un objeto psicológico. Existen infinidad de escalas de

actitud y de opinión ya estandarizadas o que el mismo investigador puede diseñar. Entre estas escalas las más utilizadas son: la escala de Thurstone, la de Likert, la de Guttman, las escalas de ordenación, entre otros.

**CENSO:** Es la medición o análisis de cada componente de la población. En algunos casos es necesario inspeccionar a cada persona o elemento de la población que deseamos descubrir, a esto lo llamamos enumeración completa o censo. Este sirve para evaluar el estado de la población de un país en un momento dado, generalmente cada diez años. Es un registro donde se concentra toda la información referente a la población o riqueza de una nación o localidad, es un estudio exhaustivo ya que se realiza en toda la población y los resultados son ratificados en un lapso de tiempo largo. El Censo constituye una lista o padrón de la población o riqueza de un país, con fines estadísticos. Sirve para evaluar el Estado de la población de un país en un momento dado.

Aunque en la antigüedad tenía una finalidad estrictamente impositiva, hoy en día constituye el punto de partida para la elaboración de las políticas demográficas y sociales. El documento básico del censo es un cuestionario en el que figuran datos como el lugar de residencia, la edad, el sexo, el estado civil, la lengua materna, el nivel de estudio y la profesión. A partir de este cuestionario, el censo indica el tamaño de la población, su distribución en el territorio y su estructura o composición. Entre las variables que pueden extraerse de los datos censales cabe mencionar la densidad de la población, relación entre el número de habitantes y la superficie; la distribución por edad y sexo, que permite elaborar las denominadas pirámides de edad, representación gráfica del grado de juventud o envejecimiento de la población, y la estructura socioeconómica de la población considerada, población activa y reparto de la primera en los diferentes sectores productivos.

## TIPOS DE CENSO

**Censo Electoral:** Es un registro que contiene la inscripción de todas aquellas personas que reúnen los requisitos para ser electores. Su elaboración y mantenimiento corresponde a la Oficina del Censo Electoral de las registradurías. El censo electoral se compone de: censo de colombianos residentes en Colombia, censo de colombianos residentes en el extranjero y censo de extranjeros residentes en Colombia con derecho a voto en las elecciones.

**Censo de Población:** Es el conjunto de operaciones de recopilación, resumen, valoración, análisis y publicación de los datos de carácter demográfico, cultural, económico y social de todos los habitantes del país (residentes tanto en viviendas como en establecimientos colectivos) y de sus divisiones político-administrativas, referidos a un momento o periodo determinado.

**Censo de Viviendas:** Es el conjunto de operaciones que consiste en la recopilación, resumen, valoración, análisis y publicación de los datos relativos a viviendas. Clasificando éstas en viviendas familiares, alojamientos y establecimientos colectivos.

## DATOS ESTADÍSTICOS

**Dato estadístico.-** Es un conjunto de valores numéricos que tienen relación significativa entre sí. Los mismos pueden ser comparados, analizados e interpretados en una investigación cualquiera. Se puede afirmar que son las expresiones numéricas obtenidas como consecuencia de observar un individuo de la

población; por lo tanto, son las características que se han tomado en cuenta de cualquiera población para una investigación determinada.

La información cuantitativa o numérica puede encontrarse casi donde quiera: en negocios, economía y muchas otras áreas. Por ejemplo el precio marcado en un sombrero en mostrado en un cierto número de pesos, la situación de empleo en una nación es expresada en un número de personas, la inscripción en una universidad es registrada mediante un número de estudiantes, la distancia recorrida por un agente de ventas es reportada en número de kilómetros, y la edad de una persona es representada por el número de años. Sin embargo, no toda la información cuantitativa es considerada como dato estadístico. La información cuantitativa apropiada para análisis estadístico debe ser un conjunto (conjuntos) de números que muestren relaciones significativas entre si. En otras palabras, los datos estadísticos son números que pueden ser comparados, analizados e interpretados. Un número aislado que no se compara o que no muestra relación significativa con otro número no es un dato estadístico.

En el ejemplo de arriba la edad de Luis a solas no constituye dato estadístico si no hay otro disponible para comparación. Sin embargo, las edades de mil estudiantes son datos estadísticos, puesto que las edades pueden ser comparadas y analizadas, y los resultados del análisis pueden ser interpretados. También, las llamadas *estadísticas* de un paciente tal como son medidas por un doctor no son datos estadísticos, puesto que cada medida, tal como la estatura no muestra relación significativa con otras medidas, tal como el número de pulsaciones por minuto o la medida de la vista del paciente. Sin embargo, la información relativa a la estatura de los pacientes dentro de un centro hospitalario en un cierto periodo de tiempo si son datos estadísticos, puesto que las estaturas pueden ser comparadas, analizadas e interpretadas de acuerdo con sus relaciones.

El área de la cual los datos estadísticos son recopilados es generalmente referida como la *población* o *universo*. Una población puede ser finita o infinita. Una población finita tiene un número limitado de individuos u objetos, mientras que una población infinita tiene un número ilimitado. Por ejemplo una clase de estadística de cuarenta estudiantes es una población finita. El número de estudiantes universitarios de América del Sur durante el año 99 es ilimitado; por lo tanto, tales estudiantes forman una población infinita. La tarea de recopilar un conjunto completo de datos de una población finita pequeña es relativamente simple. Si deseamos obtener la edad de 25 estudiantes en una clase de estadística, se puede simplemente preguntar a cada estudiante la edad. Así tendríamos un conjunto completo de datos. Sin embargo, recopilar tales datos de una población finita pero grande, es algunas veces imposible o impracticable. Recopilar un conjunto completo de la concierne a la edad de todos los estudiantes de la escuela primaria de Colombia en 1999, por ejemplo, puede ser impracticable, aunque es posible, no se debería realizar debido al tiempo y los costos consumidos. La recopilación de datos completos de una población infinita es infinitamente imposible.

Al fin de evitar la tarea imposible o impráctica de recolectar la totalidad de los datos de poblaciones infinitas, usualmente se extrae una muestra de elementos representativa de la población. La muestra, es entonces, utilizada para el estudio estadístico y los resultados de la muestra son usados como las bases para describir, estimar o predecir las características de la población. Supongamos quinientos estudiantes son representativos de los educandos de la Universidad y que los mismos son seleccionados del total de cinco mil estudiantes que posee esa institución en 1999 (población). El conjunto de datos recopilados concierne a las edades de los quinientos estudiantes es una muestra, luego, un investigador puede usar estos resultados para estimar o predecir las edades de todos los estudiantes la Universidad en el 2002..

## **CUADRO O TABLAS ESTADÍSTICAS**

**Cuadros estadísticos.-** Son esquemas organizados en los que se registran los datos estadísticos en forma organizada con la frecuencia de cada uno de estos, los mismos se observan en columnas y filas con la finalidad de presentar la información recopilada de una investigación o estudio determinado. Por lo tanto, los cuadros estadísticos es una ordenación de datos numéricos en filas y columnas con las especificaciones correspondientes acerca de la naturaleza de los datos. Constituye una forma útil de presentar los datos estadísticos obtenidos en una investigación a través de cuadros, tablas y gráficos. Esta puede presentar la información para referencias generales o para un uso específico o particular.

La ordenación de datos en cuadros estadísticos, denominada forma tabular o tabulación, están constituidos por datos cuantitativos y éstos a su vez están en filas y columnas de acuerdo con las especificaciones de los datos. La tabulación es una presentación sistemática de los datos estadísticos de una investigación determinada, estos se presentan en forma resumida a través de las tablas o cuadros estadísticos.

## **TIPOS DE CUADROS ESTADÍSTICOS**

### **CUADROS ESTADÍSTICOS GENERALES O DE REFERENCIA**

Son aquellos que proporcionan información en forma bastante detallada y de fácil referencia, se utilizan como referencia de uso general. No se elaboran para una exposición específica. Los cuadros presentados por las organizaciones gubernamentales casi siempre son de este tipo. Son construidos de una forma breve y simple. Están referidos a observaciones independientes entre sí.

### **CUADROS ESTADÍSTICOS ANALÍTICOS O DE RESUMEN**

Las tablas o cuadros analíticos o de fines especiales, son aquellos que proporcionan un resumen o un análisis de los datos, en otras palabras; tiene por objeto el de exponer los resúmenes finales de una serie de cuadros de investigación, o los datos de éstos que más interesan. Estos se refieren a observaciones dependientes entre sí.

## **PARTES QUE INTEGRAN UN CUADRO ESTADÍSTICO**

Los cuadros estadísticos están compuestos por las siguientes partes:

1. Título.
2. Encabezamiento.
3. Columna Matriz. o Concepto
4. Cuerpo.
5. Notas de Encabezado
6. Nota de Pie
7. Fuente de Datos

1) **Título:** Es una descripción del contenido de la tabla. Debe ser compacta y completa. Este comprende las siguientes partes:

A) *Numeración del Cuadro:* cuando los cuadros forman parte de un texto o de un grupo deben ser numerados en la parte superior central de la hoja.

B) *Título Propiamente Dicho:* se debe seguir los siguientes puntos:

1) Se debe ubicar centrado en la parte superior del cuadro sin subrayar, y usando letras mayúsculas para todo el enunciado.

2) Se debe redactarse con precisión y que exprese brevemente los datos que se presentan en el cuadro.

3) En general el orden del enunciado será así:

a) Referencia Geográfica.

b) Naturaleza de los Datos.

c) Referencia Cronológica.

d) Detalle de las clasificaciones o unidades. Estas deberán colocarse entre paréntesis y utilizando mayúsculas únicamente al iniciar la palabra.

Ejemplo:

### **Un título completo indica:**

¿Qué son los datos incluidos en el cuerpo de la tabla?

¿Dónde está el área representada por los datos?

¿Cómo están los datos clasificados?

¿Cuándo ocurrieron los datos?

**Encabezado:** Es el título de la parte de una columna o columnas. Las tablas más simples pueden consistir solamente de dos columnas y dos encabezados: Una para los conceptos y otra para los datos. Debe disponerse en la parte superior del cuadro y las designaciones que comprenden deberán escribirse en lo posible horizontalmente, debiendo ser preciso y breves, así mismos se dispondrá en un orden lógico de izquierda a derecha. Otra observación para la elaboración del encabezamiento, es que tanto él como las diversas columnas deben separarse con rayas, cerrando el cuadrado por la parte superior e inferior con una raya gruesa o una doble raya, en la actualidad existe la preferencia de no rayar verticalmente el encabezamiento.

**Concepto o Columnas Matriz:** La descripción en hilera de la tabla son llamados conceptos; y estos son colocados al lado izquierdo de la tabla. La naturaleza de las clasificaciones es indicada por los encabezados de las columnas, incluyendo la columna matriz

Es bueno Recordar que los datos estadísticos pueden referirse a clasificaciones cualitativas, cuantitativas, cronológicas o geográficas; recordar esto es importante puesto que la naturaleza de los datos tomando en cuenta esta clasificación determinará en parte el arreglo en que se lleven éstos a la columna matriz. Existen variadas formas de arreglo de los datos en la columna matriz. Es permisible disponerlos en orden alfabético, método que se usa habitualmente cuando los datos se clasifican geográfica o cualitativamente. Pueden ordenarse también según clases fijadas por la costumbre: casado,

soltero, divorciado, viudo. Es factible observar que si las diferentes nominaciones son ordenadas alfabéticamente, se hace muy sencillo localizarlas.

Cuando se trata de clasificaciones cuantitativas, el arreglo puede hacerse en orden ascendente o descendente. Cuando se refiere a clasificaciones cronológicas, se ordenan los años en sentido ascendente: 1995, 1996, 1997, 1998, y cuando se trate de meses se comienza por el mes de Enero.

**Cuerpo del cuadro:** El cuerpo del cuadro es la parte que contiene los datos estadísticos presentados en éste. Cada dato individual ocupa en el cuadro un lugar que corresponde a la intersección de una fila y una columna dada; por tanto, el significado de los datos en un lugar está indicado por las especificaciones o partidas combinadas de la columna y la fila que se interceptan. Cuando el valor de uno de los lugares del cuerpo del cuadro sea cero es conveniente marcar ese lugar con un guión, si no existe el dato, si es estimado, o si la cifra indica alguna consideración distinta a la del resto de los otros, debe indicarse con una llamada y su respectiva aclaratoria al pie del cuadro. La representación efectiva de los datos en la tabla depende de los arreglos de las columnas en hileras.

**Nota de Encabezado:** Son usualmente escritas justamente arriba de los encabezados y debajo de los títulos. Son usados para explicar ciertos puntos relacionados con la tabla completa que no han sido incluidos en el título ni en los encabezados ni en los conceptos.

**Nota de Pie:** Las notas de pie son usualmente colocadas debajo de los conceptos. Son usados para clarificar algunas partes incluidas en la tabla que no son explicadas en otras partes, tal como las notas de pie en la tabla. Las notas al pie de los cuadros se utilizan para hacer aclaratorias sobre uno o varios elementos en particular. La nota sobre la fuente de los datos debe indicar el origen de la información presentada en el cuadro.

**Fuentes:** Las fuentes de datos o simplemente fuentes, es usualmente escrita debajo de las notas de pie. Si los datos fueron recopilados y presentados por la misma persona, es costumbre no establecer la fuente en la tabla. El objeto de la indicación de las fuentes de los datos es el de proporcionar el debido reconocimiento a la persona u organismo que recopiló y /o publicó los datos, además de indicar, a quienes deseen ampliar la información, el origen de la misma

### **Otras consideraciones sobre la Construcción de Cuadros**

En cuanto a los totales, cuando se les quiera recalcar, por tener gran relevancia, se colocarán en la parte superior de la columna matriz y a la izquierda del encabezamiento, aunque algunos autores lo suelen poner en la parte inferior de la columna matriz y a la derecha del encabezamiento.

Los cuadros estadísticos que llamamos de resumen o analíticos, deben construirse de tal forma que se destaquen las comparaciones importantes, Esto puede lograrse colocando las cifras que se van a comparar en columnas o filas contiguas.

**Uso de porcentajes:** En los cuadros estadísticos, normalmente se utilizan porcentajes. Existen cuadros que sólo poseen datos en forma de porcentajes, mientras que otros vienen expresados tanto en valores absolutos como en porcentajes. La finalidad del uso de los porcentajes en los cuadros es facilitar la comparación, de tal manera que las relaciones que puedan existir se perciban. Cuando se usan estos es necesario recalcar las bases sobre las cuales se han establecido dichos porcentajes; esto tiene como propósito indicar al lector cual es la base que se está utilizando para determinar el mismo.

## PROPIEDADES DE LOS CUADROS ESTADÍSTICOS

- \* Deben simplificar la presentación de las tablas.
- \* Tratar un solo tema en ese.
- \* Elaborar un arreglo apropiado de clasificación.
- \* El tamaño del cuadro debe crearse de tal manera que no sea ni muy largo y angosto, ni muy ancho o corto.
- \* Cada signo de presentación que se va a utilizar debe estar plenamente identificado.
- \* Las notas que se encuentran al pie de los cuadros deben incluir las descripciones en forma precisa.

**Tabulación.-** Es una presentación sistemática de los datos estadísticos de una investigación determinada, estos se presentan en forma resumida a través de las tablas o cuadros estadísticos.

**Cuadros estadísticos.-** Son esquemas organizados en los que se registran los datos estadísticos en forma organizada con la frecuencia de cada uno de estos, los mismos se observan en columnas y filas con la finalidad de presentar la información recopilada de una investigación o estudio determinado.

Sean los siguientes datos el tiempo de servicio de un grupo de trabajadores de la empresa FATEXTOL:

4	2	4	5	3	3	5	3	2	2
2	5	2	4	6	6	2	3	3	6
3	5	3	4	6	3	4	2	4	6

Estos datos se tabularan y se presentaran en un cuadro estadístico de la siguiente manera:

**Cuadro correspondientes a los años de servicios de los obreros de la empresa FATEXTOL**

<b>Años de servicios</b> <b>(Variable)</b>	<b>N° de obreros</b> <b><math>f_i</math></b>
2	7
3	8
4	6
5	4
6	5
<b>TOTAL</b>	<b>30</b>



**Gráficas o Diagramas.-** Son expresiones en forma de figura, de información originada de un conjunto de datos estadísticos, que explican un fenómeno determinado. Son descripciones de operaciones y demostraciones que se representan por medio de figuras o signos, los mismos se realizan con los valores de los cuadros estadísticos. En otras palabras, es una representación de la relación entre variables, que se realiza en un plano determinado.

### **Gráficos estadísticos más importantes**

El fin que persigue todo gráfico es el de dar una idea rápida de la situación que en ese momento se está investigando. Por tal motivo, la presentación de los datos por medio de gráficos debe ser de una forma simple y de una comprensión fácil. Es preferible construir un conjunto de gráficos en donde cada uno de ellos presente un aspecto sencillo de una situación determinada, que presentar un solo gráfico en el cual se observen demasiadas relaciones que se haga difícil estudiar de una forma efectiva. Por lo tanto, no debe sobrecargarse un gráfico para tratar de mostrar demasiadas categorías, ya que, la simplicidad es una de la característica básica de estos.

Existe una gran variedad de tipos de gráficos entre los que se pueden mencionar los pictogramas, cartogramas, de cuadrados, de triángulos y círculos proporcionales, de sectores circulares, de barras, lineales, estereogramas, polares, etc., pero los más utilizados y de interpretación sencilla son los: Los *gráficos de barras*, los *de sectores circulares* y los *lineales*. En este curso solo se estudiarán las siguientes gráficas:

- 1.- Diagrama de Líneas.
- 2.- Diagrama de Barras.
- 3.- Diagrama Circular o de Pastel.
- 4.- Histograma.
- 5.- Polígono de Frecuencia.
- 6.- Polígono Acumulativo (OJIVA).

Los diagramas de líneas, el histogramas, el polígono de frecuencia y la ojiva son gráficos cartesianos por que para su construcción requieren del plano cartesiano, a estos se le denominan en términos generales gráficos de líneas. El diagrama de barras y el de pastel se les denomina gráficos de sectores, puesto que, no requieren del plano cartesiano para su construcción.

### **Diagrama de Línea**

El diagrama de línea es una gráfica que se representa en el plano cartesiano, con los datos de un fenómeno determinado para el cual se ha elaborado un cuadro estadístico. En términos generales se puede decir que son aquellas líneas que se dibujan en los ejes cartesianos, siguiendo algunos criterios.

### **Criterios para elaborar un diagrama de Líneas**

- 1.- La utilización de la escala que se utilizará en el plano cartesiano puede variar tomando en cuenta el fenómeno que se va graficar. No es necesario que las abscisas (ejes  $x$ ) y las ordenadas (eje  $y$ ) del plano cartesiano lleven la misma escala; sin embargo, cuando las magnitudes de las variables no se

diferencian sustancialmente, es recomendable utilizar escalas iguales para obtener un gráfico de mayor precisión.

2.- Cuando un a de las variables en estudio se inicia con valores muy altos es recomendable no comenzar el eje por el origen cartesiano sino por un valor próximo o por el mismo valor por donde comienza la variable.

3.- Es costumbre representar en el eje de las  $x$  del plano cartesiano la variable independiente del estudio que se realiza y en el eje de las  $y$  la variable dependiente. En aquellos casos que se dificulta distinguir el tipo de variable se recomienda colocar en la ordenada del plano cartesiano las frecuencias de las variables en estudio y sobre la abscisa la variable cronológica (años, meses, semanas, días, horas, etc.).

Ejemplo: Los datos que se presenta a continuación corresponden a los años de servicios de 60 empleados de la empresa FATEXTOL:

4	3	4	5	6	7	8	9	10	8	4	8	6	3	8	10	7	10	9	10
8	3	5	7	8	6	10	9	7	8	5	3	8	7	8	10	8	10	8	7
7	9	8	7	6	5	7	10	8	9	8	10	7	6	7	8	6	7	6	8

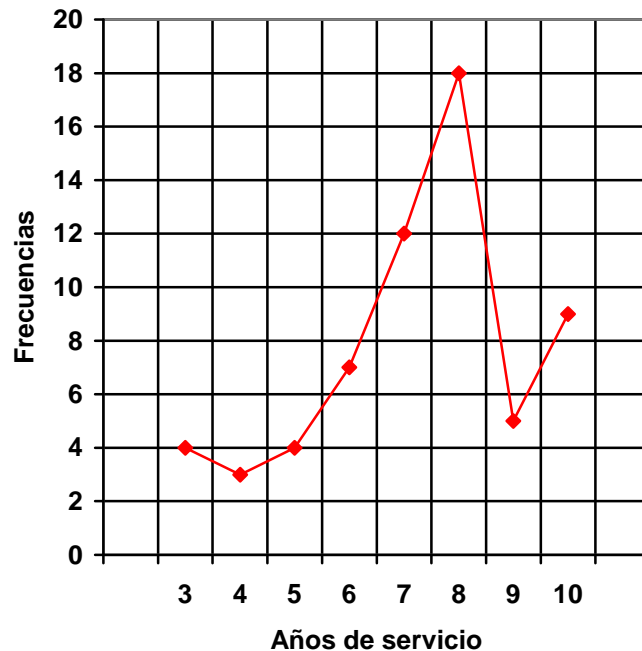
**Procedimiento.-** Con los datos se procedió a elaborar un cuadro estadístico y se obtuvo el siguiente:

1.- Con los datos que fueron suministrados se elaboró un cuadro estadístico con la frecuencia de cada variable y se organizaron las mismas en una forma ascendentes de la siguiente manera:

**Cuadro resumen de los años de servicio de los Empleados de la Empresa FATEXTOL.**

<b>Años de Servicio (Variable Independiente)</b>	<b>N° de Empleados (<math>f_i</math>)</b>
3	4
4	3
5	4
6	7
7	12
8	16
9	5
10	9
<b>TOTAL</b>	<b>60</b>

**Grafica de Linea correspondiente a los años de servicio de los empleados de la empresa FATEXTOL.**



2.- Se marcó en el eje de las “x” los años de servicio con la frecuencia correspondiente en el eje de las “y”. Luego, esos puntos se unieron mediante líneas y el resultado fue la gráfica de línea de los años de servicio de los empleados de la empresa FATEXTOL.

### Diagrama de Barras

Los diagramas de barras son gráficas que se utilizan con mucha frecuencia para representar datos de una investigación determinada, son de fácil interpretación para cualquier lector. Estos gráficos están constituidos por una serie de rectángulos o barras. La longitud y anchura de cada barra representa un fenómeno.

La forma de elaborar los mismos es la siguiente: se utiliza un sistema de coordenadas rectangulares y se llevan al eje de las “x” los valores que toma la variable en estudio y en el eje de las “y” se colocan las frecuencias de cada barra. Luego se construyen los rectángulos, tomando como base al eje de las abscisas, cuya altura será igual a cada una de las diferentes frecuencias que presentan las variables en estudio. La magnitud con que viene expresada la variable se observa en la longitud de las barras (rectángulos). Es importante destacar que solamente la longitud de las barras y no su anchura es lo que denota la diferencia de magnitud entre los valores de la variable. Todas las barras tienen que tener una anchura igual, separadas entre sí, preferiblemente por una longitud igual a la mitad del ancho de estas

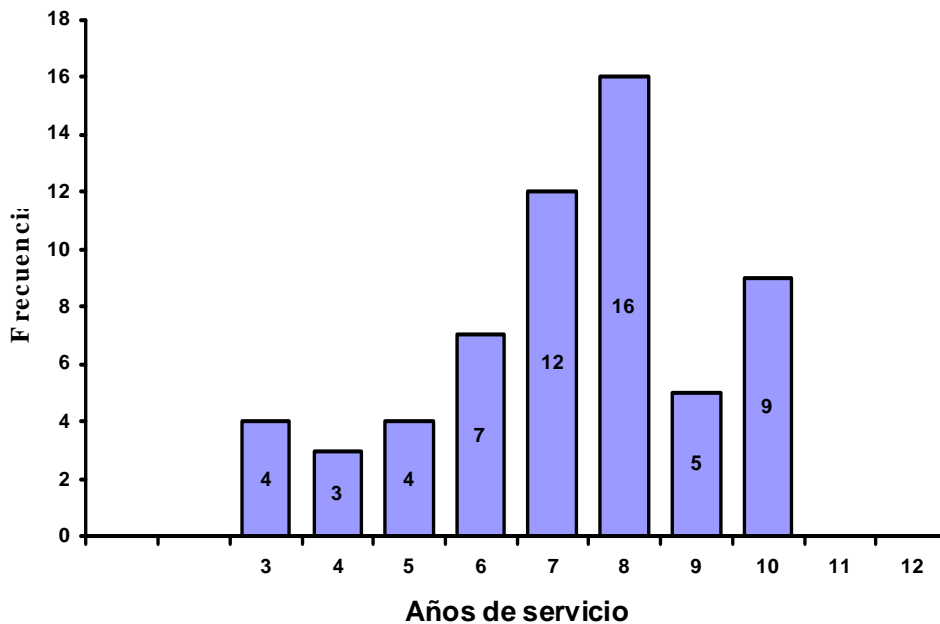
o distancias iguales entre barras. Es recomendable, que las barras no sean ni excesivamente cortas y anchas, ni demasiado largas y angostas, esto es con el objeto de dar una visión objetiva de la investigación en estudio. Las barras se pueden graficar tanto verticalmente como horizontalmente. Se pueden elaborar barras compuestas y barras agrupadas. Ejemplo:

Sea el siguiente cuadro resumen los años de servicio de los empleados de la empresa FATEXTOL, con el mismo elabore un diagrama de barras.

Años de servicio (Variable Independiente)	N° de Empleados (fi)
3	4
4	3
5	4
6	7
7	12
8	16
9	5
10	9
<b>TOTAL</b>	<b>60</b>

### Gráfico Circular o de Pastel

Diagra de barra correspondiente a los años de servicio de los empleados de la Empresa FATEXTOL



La gráfica de pastel (gráfico de sectores) es un tipo de gráfica que consiste en representar por medio de la circunferencia o un círculo las magnitudes que expresan los datos de un estudio determinado. Este

tipo de gráfica considera la circunferencia como representante de los datos estadísticos de una investigación cualquiera. Por tal motivo, se dividirá en tantos sectores como variables tenga la investigación en estudio; la magnitud de cada sector se encontrará en relación directa con la magnitud de la variable a representar, tomando en cuenta que toda la investigación se representa con  $360^\circ$ . En general, los datos que se representan por medio de este diagrama son partes componentes de un total.

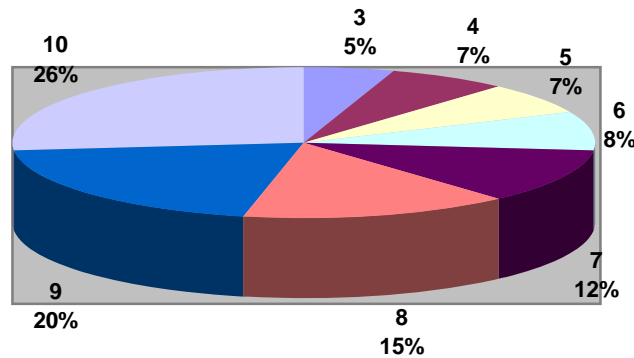
Para su elaboración se procede de la siguiente forma: se considera la circunferencia como representación del total de la investigación en estudio, por tal motivo, se dividirá toda su superficie en tantas secciones como variables tenga la investigación en estudio, las superficies de las sesiones de la circunferencia deben ser proporcionales a la magnitud de cada componente representado por las variables.

Ejemplo: Sea el siguiente cuadro resumen los años de servicio de los empleados de la empresa FATEXTOL, elabore una gráfica de Círculo.

Años de servicio	$f_i$	Grados	%
3	4	24	7.0
4	3	18	5.0
5	4	24	7.0
6	7	42	12.0
7	12	72	20.0
8	16	96	26.0
9	5	30	8.0
10	9	54	15.0
<b>TOTALES</b>	<b>60</b>	<b><math>360^\circ</math></b>	<b>100.0</b>

En la gráfica de pastel se observa la variable y el % correspondiente de la misma.

Grafica de pastel correspondiente a los años de servicio de los empleados de la empresa FATEXTOL



Si se considera la circunferencia como representación del total de los datos que en el estudio anterior referente a la empresa FATEXTOL es de 60, entonces, se debe igualar 60 a los 360° de la circunferencia y por medio de una simple regla de tres, distribuir esos 360° proporcionalmente entre las frecuencias de las diferentes variables del estudio en cuestión, para así obtener las magnitudes de los diferentes sectores que representaran el número de empleados con los diferentes años de servicio como se observa en el cuadro anterior. De la misma forma se obtienen los porcentajes de cada variable, planteando una sencilla regla de tres. Los diferentes grados que formaran el diagrama de círculo en este caso se obtienen aplicando una regla de tres de la siguiente forma:

$$\begin{array}{l}
 60 \text{ ——— } 360^\circ \\
 4 \text{ ——— } X \\
 X = 24^\circ
 \end{array}
 , \quad
 \begin{array}{l}
 60 \text{ ——— } 360^\circ \\
 3 \text{ ——— } X \\
 X = 18^\circ
 \end{array}
 , \quad
 \begin{array}{l}
 60 \text{ ——— } 360^\circ \\
 4 \text{ ——— } X \\
 X = 24^\circ
 \end{array}
 , \quad
 \begin{array}{l}
 60 \text{ ——— } 360^\circ \\
 7 \text{ ——— } X \\
 X = 42^\circ
 \end{array}$$

$$\begin{array}{l}
 60 \text{ ——— } 360^\circ \\
 12 \text{ ——— } X \\
 X = 72^\circ
 \end{array}
 , \quad
 \begin{array}{l}
 60 \text{ ——— } 360^\circ \\
 16 \text{ ——— } X \\
 X = 96^\circ
 \end{array}
 , \quad
 \begin{array}{l}
 60 \text{ ——— } 360^\circ \\
 5 \text{ ——— } X \\
 X = 30^\circ
 \end{array}
 , \quad
 \begin{array}{l}
 60 \text{ ——— } 360^\circ \\
 9 \text{ ——— } X \\
 X = 54^\circ
 \end{array}$$

Los resultados obtenidos en estos cálculos se encuentran ubicados en el cuadro anterior. Los porcentajes de cada una de las variables se encuentran en el mismo cuadro y los mismos se calcularon así:

$$\begin{array}{l}
 60 \text{ ——— } 100 \\
 4 \text{ ——— } X \\
 X = 7.0
 \end{array}
 , \quad
 \begin{array}{l}
 60 \text{ ——— } 100 \\
 3 \text{ ——— } X \\
 X = 5.0
 \end{array}
 , \quad
 \begin{array}{l}
 60 \text{ ——— } 100 \\
 4 \text{ ——— } X \\
 X = 7.0
 \end{array}
 , \quad
 \begin{array}{l}
 60 \text{ ——— } 100 \\
 7 \text{ ——— } X \\
 X = 11.0
 \end{array}$$

$$\begin{array}{cccc}
 60 \frac{\quad}{12} 100 & , & 60 \frac{\quad}{16} 100 & , & 60 \frac{\quad}{5} 100 & , & 60 \frac{\quad}{9} 100 \\
 X = 20.0 & & X = 26.0 & & X = 8.0 & & X = 15.0
 \end{array}$$

Una vez realizado los cálculos de los grados que corresponden a cada variable, se llevan a una circunferencia utilizando para ello un transportador. Luego se iniciará el marcaje de los grados considerando el 0° lo que corresponde a las doce de un reloj y el marcaje se realizará tomando en cuenta el orden lógico del cuadro y con el sentido de dirección que sigue las agujas de un reloj.

## Representación de tronco y hoja

Un método para iniciar el análisis exploratorio de los datos, previo al uso de los métodos estadísticos tradicionales, y que además proporciona información rápida, visual y es relativamente nueva, es la **representación gráfica de tronco y hoja**. Esta representación se basa en la ordenación de los datos a manera de gráfico, pero sin llegar a ello, utilizando las decenas y las unidades.

Esta técnica se puede encontrar en el libro de [Freund y Simon](#), pero comentaremos su uso a través del siguiente ejemplo que contiene las calificaciones obtenidas en una prueba de matemáticas:

78	93	61	100	70	83	88	74	97	72
66	73	76	81	83	64	91	70	77	86

Ahora se analizarán cada uno de los datos separando las decenas de las unidades, es decir, el número 51 se verá como 5 | 1. De esta manera las decenas se pondrán en una columna, en forma vertical, y las unidades a su derecha:

6		1 6 4
7		8 0 4 2 3 6 0 7
8		3 8 1 3 6
9		3 7 1
10		0

Para entenderle un poco más, se ha de decir que el primer renglón que dice 6 | 1 6 4 quiere decir que entre la lista de datos se encuentran los valores 61, 66 y 64.

Esta es la representación gráfica **tronco y hoja**, donde cada renglón es una posición de **tronco** y cada dígito de la derecha es una **hoja**.

El procedimiento para realizarla es primero empezar con los troncos, es decir la columna de la izquierda, y después dato por dato ir llenando las hojas a la derecha de la línea vertical, en el tronco correspondiente.

Además, si se desean tener los datos ordenados, y hay gente que lo prefiere así, se pueden ordenar las hojas en cada renglón para que la representación quede como sigue:

6	1 4 6
7	0 0 2 3 4 6 7 8
8	1 3 3 6 8
9	1 3 7
10	0

En realidad una representación de tronco y hojas presenta la misma información que la lista original de datos, pero de una manera mucho más compacta (especialmente si la lista de datos es más grande) y manejable.

Sin embargo, información más compleja resulta un poco más difícil de manejar, por lo que en ocasiones conviene redondear los datos, ignorar sus partes decimales o utilizar las centenas u otras posiciones de los números para las troncos. En cada uno de esos casos conviene hacer alguna anotación, o poner una nota, a fin que los lectores puedan identificar las adecuaciones realizadas y así poder interpretar lo que se quiere transmitir.

Para mostrar la información de manera más clara, es posible modificar el número de posiciones del tronco, aumentándola o disminuyéndola de acuerdo a las necesidades particulares de cada problema. Por ejemplo, con los datos del examen anterior, se pueden dividir en dos cada posición del tronco, utilizando la primera posición para disponer las hojas 0, 1, 2, 3 y 4, y la segunda posición para las hojas restantes. De esta manera, se obtiene la **representación gráfica de doble tronco**:

6-	1 4
6+	6
7-	0 0 2 3 4
7+	6 7 8
8-	1 3 3
8+	6 8
9-	1 3
9+	7
10-	0

Con esto se han duplicado el número de posiciones del tronco, con la intención de buscar una mayor claridad en la presentación.

Esta manera de representación inicial de los datos no la profundizaremos más, sino que la utilizaremos más adelante en algunos casos para, precisamente, presentar una representación inicial de la información obtenida.

**Frecuencia.-** La frecuencia es el número de veces que se repite (aparece) el mismo dato estadístico en un conjunto de observaciones de una investigación determinada, las frecuencias se les designan con las letras  $f_i$ , y por lo general se les llaman frecuencias absolutas.

**Distribución de Frecuencia.-** En estadística existe una relación con cantidades, números agrupados o no, los cuales poseen entre sí características similares. Existen investigaciones relacionadas con los precios de los productos de la dieta diaria, la estatura y el peso de un grupo de individuos, los salarios de los empleados, los grados de temperatura del medio ambiente, las calificaciones de los estudiantes, etc., que pueden adquirir diferentes valores gracias a una unidad apropiada, que recibe el nombre de variable. La representación numérica de las variables se denomina dato estadístico.



La distribución de frecuencia es una disposición tabular de datos estadísticos, ordenados ascendente o descendientemente, con la frecuencia ( $f_i$ ) de cada dato. Las distribuciones de frecuencias pueden ser para *datos no agrupados* y para *datos agrupados* o de intervalos de clase.

**Distribución de frecuencia para datos no Agrupados.-** Es aquella distribución que indica las frecuencias con que aparecen los datos estadísticos, desde el menor de ellos hasta el mayor de ese conjunto sin que se haya hecho ninguna modificación al tamaño de las unidades originales. En estas distribuciones cada dato mantiene su propia identidad después que la distribución de frecuencia se ha elaborado. En estas distribuciones los valores de cada variable han sido solamente reagrupados, siguiendo un orden lógico con sus respectivas frecuencias.

Ejemplo: Los datos que se presenta a continuación corresponden a los años de servicios de 60 empleados de la empresa FATEXTOL, con los mismos, elabore una distribución de frecuencia para datos no agrupados:

3	5	4	5	6	7	8	9	10	8	4	8	6	3	8	10	7	10	9	10
8	3	5	7	8	6	10	9	7	8	5	3	8	7	8	10	8	10	8	7
7	9	8	7	6	5	7	8	8	9	8	10	7	6	7	8	6	7	6	10

**Procedimiento.-** Con los datos se procedió a elaborar un cuadro estadístico con la frecuencia de cada variable y se organizaron las mismas en una forma ascendentes y se obtuvo *la siguiente distribución de frecuencia para datos no agrupados:*

Años de Servicio	( $f_i$ )
3	4
4	3
5	4
6	7
7	12
8	16
9	5
10	9
<b>TOTAL (N)</b>	<b>60</b>

En la distribución se observa que “N” (número total de datos) es de 60 pero el rango (número de variables diferentes) de esta serie de valores es de 8, por lo tanto, la distribución más conveniente es la que se utiliza *para datos no agrupados*.

**Distribución de frecuencia de clase o de datos Agrupados.-** Es aquella distribución en la que la disposición tabular de los datos estadísticos se encuentran ordenados en clases y con la frecuencia de cada clase; es decir, los datos originales de varios valores adyacentes del conjunto se combinan para

formar un intervalo de clase. No existen normas establecidas para determinar cuándo es apropiado utilizar datos agrupados o datos no agrupados; sin embargo, se sugiere que cuando el número total de datos ( $N$ ) es igual o superior 50 y además el rango o recorrido de la serie de datos es mayor de 20, entonces, se utilizará *la distribución de frecuencia para datos agrupados*, también se utilizará este tipo de distribución cuando se requiera elaborar gráficos lineales como el histograma, el polígono de frecuencia o la ojiva.

La razón fundamental para utilizar la distribución de frecuencia de clases es proporcionar mejor comunicación acerca del patrón establecido en los datos y facilitar la manipulación de los mismos. Los datos se agrupan en clases con el fin de sintetizar, resumir, condensar o hacer que la información obtenida de una investigación sea manejable con mayor facilidad.

Este tipo de distribución se basa en el principio de que una observación no puede considerarse diferente de otra por presentar pequeñas diferencias cuantitativas, como por ejemplo el sueldo mensual de dos empleados que difieran en 500 pesos, de dos edades de personas adultas que difieran en un año, dos alturas de un edificio que difieran en un metro, el costo de 2 autos nuevos que difieran en 5000 pesos, etc.

Al agrupar los datos en una distribución de frecuencia de clase se pierde parte de la información. La reducción o agrupamiento a que son sometidos los datos de una serie de valores cuando existen muchos valores diferentes, originan los denominados *errores de agrupamiento*; sin embargo, estos errores son en general muy pequeños, razón por la cual la distribución de frecuencia de clase tiene una validez estadística práctica.

Cuando se dispone de una serie de datos que sea igual o mayor que 50 y, además, el rango de esa serie de valores sea mayor de 20, lo más recomendable es utilizar una distribución de frecuencia de clase.

## Componentes de una distribución de frecuencia de clase

**1.- Rango o Amplitud total (recorrido).**- Es el límite dentro del cual están comprendidos todos los valores de la serie de datos, en otras palabras, es el número de diferentes valores que toma la variable en un estudio o investigación dada. Es la diferencia entre el valor máximo de una variable y el valor mínimo que ésta toma en una investigación cualquiera. El rango es el tamaño del intervalo en el cual se ubican todos los valores que pueden tomar los diferentes datos de la serie de valores, desde el menor de ellos hasta el valor mayor estando incluidos ambos extremos. El rango de una distribución de frecuencia se designa con la letra  $R$ .

Para calcular el rango de una distribución de frecuencia de clase se calcula la diferencia entre el dato mayor ( $X_M$ ) y el dato menor ( $X_m$ ), y se le agrega una Unidad de Medida (UM), que por lo general es la unidad. La unidad de medida en una distribución de frecuencia se encuentra al obtener la diferencia de dos datos consecutivos de la serie de valores. En algunos casos, los valores de los datos de la serie de observaciones pueden estar expresadas con números decimales, o ser múltiplos de algunos otros números, cuando esto sucede, la unidad de medida adquiere un valor diferente a la unidad. Ver ejemplos:

EJEMPLOS	A	B	C
OBSERVACIONES	6, 9, 11, 12, 19, 20, 26, 27, 32, 33, 39	0.5, 0.6, 0.10, 0.11 0.19, 0.21, 0.22,	6, 9, 12, 21, 33, 39, 48,
UNIDAD DE MEDIDA	1	0.1	3

Los datos que se presenta a continuación corresponden a los años de servicios de 60 empleados de la empresa FATEXTOL, con los mismos calcule el rango de la distribución de frecuencia.

3	5	4	5	6	7	8	9	10	8	4	8	6	3	8	10	7	10	9	10
8	3	5	7	8	6	10	9	7	8	5	3	8	7	8	10	8	10	8	7
7	9	8	7	6	5	7	8	8	9	8	10	7	6	7	8	6	7	6	10

Para calcular el rango lo primero que se hace es ubicar el  $X_M$ , el  $X_m$  y la UM.

$X_M = 10$ ,  $X_m = 3$ ,  $UM = 5 - 4 = 1$  (diferencia entre dos valores consecutivos), luego

$R = X_M - X_m + UM$ , entonces,  $R = 10 - 3 + 1 = 8$

**R = 8.**

**2.- Clase o Intervalo de clase.-** Son divisiones o categorías en las cuales se agrupan un conjunto de datos ordenados con características comunes. En otras palabras, son fraccionamientos del rango o recorrido de la serie de valores para reunir los datos que presentan valores comprendidos entre dos límites.

Para organizar los valores de la serie de datos hay que determinar un número de clases que sea conveniente. En otras palabras, que ese número de intervalos no origine un número pequeño de clases ni muy grande. Un número de clases pequeño puede ocultar la naturaleza natural de los valores y un número muy alto puede provocar demasiados detalles como para observar alguna información de gran utilidad en la investigación.

Uno de los problemas que se presentan al elaborar una distribución de frecuencia de clase es el de fijar el número de clases a utilizar, puesto que no existe un criterio general para determinar el número de clases a elegir; sin embargo, algunos especialistas en la materia creen que un buen criterio es considerar de 7 a 20 clase, dependiendo esto, de las características del estudio que se realiza; por tal motivo el número de clase a utilizar en una investigación determinada dependerá de la persona que realice la investigación. En este curso se utilizará el criterio anteriormente descrito.

Las clases de una distribución de frecuencia indican las cotas o fronteras de cada clase en la distribución, las clases están formadas por dos números, denominados límites aparentes (LA), ejemplo: 32—37, el primero de estos (32) se le llama límite inferior aparente (LIA) y al segundo(37) se le denomina límite superior aparente(LSA).

Los Límites Reales(LR) o verdaderos de una clase son aquellos que se obtienen restándole *media unidad de medida* al límite aparente inferior de una clase y sumándole *media unidad de medida* al

límite superior aparente de las diferentes clases, es decir, son valores no observables de la variable en estudio, puesto que no lo registra *la unidad de medida* utilizada; hay que tener cuidado de que los límites reales de clase no coincidan con valores observables de la variable, para evitar ambigüedades sobre la clase a la que corresponde una observación. Si se toma como ejemplo: 32 ———37 (32 a 37), se puede observar que estos son los límites aparentes inferior 32 y superior 37 de esa clase, si se aplica el concepto de límite real se tendrán los siguientes límites verdaderos:

31.5———37.5, como se puede observar el límite inferior aparente disminuyó en *media unidad de medida* y se convirtió en (LRI) Límite Real Inferior (31.5) y el límite superior aparente aumentó *media unidad* y se convirtió en (LRS) Límite Real Superior (37.5).

### Tamaño de los Intervalos de Clase

Los intervalos de clase pueden ser de tres tipos, según el tamaño que estos presenten en una distribución de frecuencia: a) *Clases de igual tamaño*, b) *clases desiguales de tamaño* y c) *clases abiertas*.

### Clases de igual tamaño

Este tipo de clases es el más utilizado en los cálculos estadísticos; cuando todas las clases son del mismo tamaño, los cálculos relacionados con la distribución de frecuencia son simplificados grandemente. En términos generales, este tipo de distribución, es el que se utiliza comúnmente en casi todas las investigaciones. Ejemplo:

CLASES	fi
5——7	5
8——10	10
11——13	15
14——16	18
17——19	11
20——22	5
<b>TOTALES</b>	<b>64</b>

En esa distribución de frecuencia de clase se puede observar que cada *clase* posee tres variables diferentes, por lo tanto, los intervalos de clases son de igual tamaño.

### Clases desiguales de tamaño

Los intervalos de clases desiguales no son frecuentes en los análisis estadísticos, la utilización de los mismos se debe evitar; sin embargo, en algunas investigaciones es indispensable su utilización; tal es el caso de aquellas investigaciones que tienen como propósito particular el de analizar valores que varían en un amplio recorrido de la variable. Cuando se utiliza este tipo de clase los intervalos de clases deberían ser incrementados de una forma ordenada si es posible. Este tipo de clases se utiliza algunas veces para reportar datos relacionados con valuaciones de activos o ingresos personales. Ejemplo:

CLASES	AMPLITUD
100—499	399
500—999	499
1000—4999	3999
5000—9999	4999
10000—24999	14999
25000—50000	25000

Como se puede observar en la anterior distribución cada clase tiene un tamaño diferente, es decir, sus amplitudes o tamaño son diferentes para cada clase.

### Clases Abiertas

Las clases abiertas son aquellas en las que uno de sus dos límites de clases no está definido numéricamente. Este tipo de clase se utiliza cuando las distribuciones poseen algunos datos u observaciones que son mucho mayores o mucho más pequeños que los demás y se quiere condensar en un sólo. En lo posible es conveniente evitar este tipo de clase ya que en estas no es posible definir el punto medio de la distribución, por lo cual se hace difícil la representación gráfica y en realizar otros cálculos con los datos que presentan los cuadros estadísticos. Sin embargo, existen investigaciones donde la aplicación de clases abiertas es conveniente, por cuanto, la existencia de valores de la serie de datos son mucho menores o mucho mayores que el resto de la serie. Ejemplo:

Sea la siguiente distribución de frecuencia, las observaciones correspondientes al salario que devenga un grupo de personas que viven en determinado barrio de una ciudad.

CLASES	fi	$\dot{X}$
Menos de 150000	67	?
150000—239000	36	194500
340000—429000	10	384500
430000—519000	8	474500
520000—609000	7	564500
610000—699000	8	654500
700000—789000	7	744500
790000 y Más	7	?
<b>TOTALES</b>	<b>135</b>	

Como se puede observar en ese cuadro estadístico los valores de las variables que conforman la primera y la última clase de la distribución no tienen valores definidos.

### 3.-Amplitud de Clase, Longitud o Ancho de una Clase

La amplitud o longitud de una clase es el número de valores o variables que concurren a una clase determinada. La amplitud de clase se designa con las letras  $I_c$ . Existen diversos criterios para determinar la amplitud de clases, ante esa diversidad de criterios, se ha considerado que lo más importante es dar un ancho o longitud de clase a todos los intervalos de tal manera que respondan a la naturaleza de los datos y al objetivo que se persigue y esto se logra con la práctica. Existe una fórmula para determinar el  $I_c$  y la misma se expresa así:

$$I_c = \frac{R}{NC} \dots, \text{Donde } R = \text{Rango } \dots, NC = \text{Numero } \dots \text{de } \dots \text{clases } \dots, I_c = \text{Amplitud } \dots \text{de } \dots \text{clase } \dots$$

Con la fórmula anterior se puede determinar el  $I_c$ , conociendo el rango y el número de clases. Cuando se tenga dudas en determinar la amplitud de clase de una serie de valores, es de gran utilidad utilizar el método sugerido por Hebert A. Sturges el cual establece que:

En esta fórmula  $1 + 3,322 \log N = NC$  (Número de clases), en la gran mayoría de los casos el resultado

$$I_c = \frac{\text{Rango}}{1 + 3,322 \cdot \log N} \dots, \text{donde } N = \text{numero } \dots \text{total } \dots \text{de } \dots \text{datos } \dots$$

final es un número fraccionario, el cual no es adecuado en la práctica, sin embargo, se puede aplicar las técnicas de redondeo para convertirlo en un número entero. En este curso se utilizará el método de Sturges para determinar el  $I_c$  de una distribución de frecuencia de clase siempre y cuando el mismo sea aplicable. Algunos investigadores consideran que el método de Sturges pierde eficacia cuando el número total de datos de una serie de valores es muy extenso, considerando como extenso un  $N > 500$ . En una distribución de frecuencia de clase el  $I_c$  se puede determinar aplicando la siguiente fórmula:

$$I_c = LRS - LRI$$

Es recomendable que el  $I_c$  sea un número impar para que el punto medio o marca de clase de una distribución coincida con un número entero lo cual facilitará cálculos posteriores.

### Como se inicia la primera Clase de una distribución de frecuencia de clase

El límite inferior de la primera clase de una distribución de frecuencia debería comenzar por un múltiplo del  $I_c$  (Ancho o tamaño del intervalo), ejemplo:

Si el  $I_c$  de una distribución de frecuencia de clase = 3, y el dato menor es 49, entonces el límite inferior de la primera clase se ubicara así:  $3 \times 1 = 3$ ,  $3 \times 2 = 6$ ,  $3 \times 3 = 9$ ,  $3 \times 4 = 12$ ,  $3 \times 5 = 15$ ,  $3 \times 6 = 18$ , .....  $3 \times 13 = 39$ ,  $3 \times 14 = 42$ ,  $3 \times 15 = 45$ ,  $3 \times 16 = 48$ ,  $3 \times 17 = 51$ , en fin por un número que sea múltiplo de 3, siempre y cuando no deje fuera de la primera clase al menor de los datos de la serie de valores, en este caso

se iniciaría la primera clase en: 48—50, no se tomo 51 por ser este valor mayor que valor menor de la serie de datos y si tomáramos 51 los valores correspondientes a 48 y 50 quedarían fuera de la clase.

Si el valor más pequeño de los datos de una serie de valores es inferior al  $I_c$ , entonces, el límite inferior de la primera clase tiene que iniciarse con ese valor más pequeño.

#### 4.-Punto medio o Marca de clase

El centro de la clase, es el valor de los datos que se ubica en la posición central de la clase y representa todos los demás valores de esa clase. Este valor se utiliza para el calculo de la media aritmética. El punto medio se representa por:

$$\dot{X} = \text{Punto ..Medio} .$$

El punto medio de una clase se determina por la semisuma del límite inferior y superior de una clase, tal como lo indica la formula siguiente:

$$\dot{X} = \frac{LAI + LAS}{2}$$

#### 5.-Frecuencia de clase

La frecuencia de clase se le denomina frecuencia absoluta y se le designa con las letras  $f_i$ . Es el número total de valores de las variables que se encuentran presente en una clase determinada, de una distribución de frecuencia de clase.

#### 6.- Frecuencia Relativa

La frecuencia relativa es aquella que resulta de dividir cada uno de los  $f_i$  de las clases de una distribución de frecuencia de clase entre el número total de datos(N) de la serie de valores. Estas frecuencias se designan con las letras  $fr$ ; si cada  $fr$  se multiplica por 100 se obtiene la frecuencia relativa porcentual ( $fr\%$ ).

#### 7.-Frecuencias acumuladas

Las frecuencias acumuladas de una distribución de frecuencias son aquellas que se obtienen de las sumas sucesivas de las  $f_i$  que integran cada una de las clases de una distribución de frecuencia de clase, esto se logra cuando la acumulación de las frecuencias se realiza tomando en cuenta la primera clase hasta alcanzar la ultima. Las frecuencias acumuladas se designan con las letras  $fa$ . Las frecuencias acumuladas pueden ser *menor que* ( $fa < que$ ) y frecuencias acumuladas mayor que ( $fa > que$ ).

#### 8.-Frecuencia acumulada menor que

Las frecuencias acumuladas menor que ( $fa < que$ ) son aquellas frecuencias acumuladas que se forman con el  $f_i$  de los valores más pequeños de las variables de cada clase hacia los valores mayores de la misma. Para graficar los polígonos de frecuencias acumuladas (ojiva)  $fa < que$ , se utilizan como

variables independientes los límites superiores de cada clase y como ordenada los diferentes valores de la  $f_a < \text{que}$ .

### 9.-Frecuencia acumulada mayor que

Las frecuencias acumuladas mayor que ( $f_a > \text{que}$ ) son aquellas frecuencias acumuladas que se forman de las  $f_i$  de los valores mayores de las variables de cada clase hacia los valores menores de la misma. Para graficar los polígonos de frecuencias acumuladas (ojiva)  $f_a > \text{que}$ , se utiliza como variable independiente los límites inferiores de cada clase con los valores de  $f_a > \text{que}$  como ordenada en el plano cartesiano.

### 10.- Frecuencia acumulada relativa

La frecuencia acumulada relativa es aquella que resulta de dividir cada una de las  $f_a$  de las diferentes clases que integran una distribución de frecuencia de clase entre el número total de datos (N) de la serie de valores, estas frecuencias se designan con las letras  $far$ . Si las  $far$  se multiplican por 100 se obtienen las frecuencias acumuladas relativas porcentuales y las mismas se designan así:  $far \%$ .

### Problema resuelto

1.- Sean los siguientes datos las horas extras trabajadas por un grupo de obreros petroleros de la zona durante un mes. Con esos datos elabore una distribución de frecuencia de clase utilizando el método de Sturges.

22	39	37	28	23	39	24	38	31	35	36	28	23	27	38	40	22	23	36	27
32	33	26	60	39	33	40	27	34	22	30	31	37	33	41	39	58	59	56	41
54	56	57	58	39	40	34	45	53	52	52	28	36	37	40	26	34	25	23	32
56	33	58	40	36	25	42	33	45	55	29	52	38	28	38	38	32	42	53	58
45	43	40	28	60	41	37	42	31	45	30	28	40	37	28	44	40	39	57	60

Para elaborar la distribución de frecuencia hay que realizar los siguientes cálculos:

1.- Calcular el rango R así:

$$R = X_M - X_m + 1 \text{ UM}, \quad X_M = 60, \quad X_m = 22, \quad \text{UM} = 23 - 22 = 1. \text{ Luego:}$$

$$R = 60 - 22 + 1$$

$$R = 39.$$

2.- Se calcula el Ic de la serie de valores aplicando el método de Sturges así:

$$Ic = \frac{R}{1 + 3,322 \cdot \log .N}$$



R = 39, N = 100, Log. 100 = 2.0.

$$Ic = \frac{39}{1 + 3,322 \times 2.0} = \frac{39}{7.64} = 5.10.$$

Como el  $Ic = 5.10$  se redondea al impar más cercano que en este caso es  $Ic = 5.0$ . Ahora se procede a buscar el límite inferior de la primera clase de la distribución, para ello se busca un múltiplo del  $Ic$  que no sea superior al menor valor de los datos que este caso es 22. Tomando en cuenta este criterio el límite inferior de la primera clase será entonces:  $5 \times 4 = 20$ , que es un múltiplo del  $Ic$  y no es mayor que el menor valor de la serie de datos. Se procede ahora a elaborar las diferentes clases que integraran la distribución de frecuencia. La primera clase se forma así: 20—24, el resto de las clases y las demás columnas que integran la distribución se formaran así (se recomienda al estudiante que realice todos los cálculos necesarios para completar la distribución):

Clases	$f_i$	$fa < que$	$fa > que$	$\dot{X}$	$fr$	$fr \%$	$far < que$	$far \%$
20—24	8	8	100	22	0.08	8.0	0.08	8.0
25—29	15	23	92	27	0.15	15.0	0.23	23.0
30—34	16	39	77	32	0.16	16.0	0.39	39.0
35—39	21	60	61	37	0.21	21.0	0.60	60.0
40—44	16	76	40	42	0.16	16.0	0.76	79.0
45—49	4	80	24	47	0.04	4.0	0.80	80.0
50—54	6	86	20	52	0.06	6.0	0.86	86.0
55—59	11	97	14	57	0.11	11.0	0.97	97.0
60—64	3	100	3	62	0.03	3.0	1.00	100.0
<b>Total</b>	<b>100</b>				<b>1.00</b>			

Para calcular las  $fa > que$  de la distribución, se inicia la acumulación de los  $f_i$  desde la última clase de la distribución, que es donde se encuentran los valores mayores de las variables, hasta llegar a la primera clase que es donde se ubican los valores menores de los datos. Los cálculos de las demás columnas de la distribución se explicaran ampliamente en la teoría.

2.- Los datos que se presentan a continuación corresponden al consumo de carne de ganado, en un trimestre, de un grupo de familia de un barrio de la ciudad de Ibagué. Con los mismos elabore una distribución de frecuencia de clase utilizando para ello el método de Sturges.

25	3	5	3	8	10	12	14	3	8	10	12	15	27	30	28	25	30	24	26
29	28	24	28	27	2	30	22	21	20	26	4	8	10	12	15	16	12	10	8
10	12	15	16	12	10	8	5	5	8	11	25	30	17	18	13	11	9	6	7
13	17	17	13	11	9	7	7	9	22	13	18	19	14	15	11	9	7	7	8
10	12	14	20	19	17	16	27	15	16	14	10	23	18	19	22	17	12	9	30

Para resolver el problema planteado se inician los siguientes pasos:

1.- Se calcula el rango R de la distribución aplicando la formula:  $R = X_M - X_m + 1UM$ .

$X_M = 30, X_m = 2, UM = 3 - 2 = 1; R = 30 - 2 + 1 = 29, R = 29, N = 100, \text{Log}.100 = 2.0$ .

2.- Se calcula el Ic aplicando el método de sturges:

$$Ic = \frac{R}{1 + 3,322 \cdot \log \cdot N}$$

$$Ic = \frac{29}{1 + 3,322 \times 2.0} = \frac{29}{7.64} = 3.8.$$

Como se puede observar el  $Ic = 3.8$ , hay que redondear este Ic al impar más cercano, que en este caso sería 3.0. Ahora se procede a buscar el límite inferior de la primera clase de la distribución esta sería un múltiplo del Ic que sea menor o igual al menor de los datos de la serie de valores.

En este caso el mínimo múltiplo del Ic es 3, pero el límite inferior de la serie de valores no se puede iniciar con 3 ya que de ser así quedarían valores fuera de la clase como es el caso de 2 que no sería incluido en la clase, en este caso se tiene que utilizar como límite inferior de la primera clase *el menor valor de la serie de datos, es decir, 2*. Luego la primera clase sería: 2 — 4, y así sucesivamente hasta completar todas las clases. Se procederá ahora a completar la distribución de frecuencia, se recomienda al estudiante realizar los diferentes cálculos necesarios para completar la misma.

CLASES	fi	fa<que	fa>que	$\dot{X}$	fr	fr %	far<que	far %<que
2—4	5	5	100	3	.005	5.0	0.05	5.0
5—7	9	14	95	6	0.09	9.0	0.14	14.0
8—10	20	34	86	9	0.20	20.0	0.34	34.0
11—13	16	50	66	12	0.16	16.0	0.50	.50.0
14—16	13	63	50	15	0.13	13.0	0.63	63.0

17— 19	11	74	37	18	0.11	11.0	0.74	74.0
20— 22	6	80	26	21	0.06	6.0	0.80	80.0
23— 25	6	86	20	24	0.06	6.0	0.86	86.0
26— 28	8	94	14	27	0.08	8.0	0.94	94.0
29— 31	6	100	6	30	0.06	6.0	1.00	100.0
<b>TOTAL</b>	<b>100</b>				<b>1.00</b>	<b>100.0</b>		

Se recomienda al estudiante que realice todos los cálculos necesarios para completar los datos que conforman el cuadro de la distribución de frecuencia anterior.

3.- Los datos que a continuación se presentan corresponden al consumo de azúcar trimestralmente por un grupo de familias de una Urbanización de Ibagué. Con los mismos elabore una distribución de frecuencia de clase, para ello utilice el método de Sturges.

14	24	32	38	40	30	26	16	18	28	30	40	42	32	28	18	20	28	34	42
44	34	28	20	22	30	34	44	46	36	30	22	24	30	36	46	48	36	30	24
22	30	38	46	48	38	30	24	26	32	38	50	52	48	46	16	18	20	24	22
22	24	26	28	30	16	14	16	20	18	14	18	34	36	40	42	44	50	14	38
20	22	24	26	18	16	24	28	30	32	26	24	22	20	18	16	14	28	32	52

Para elaborar la distribución de frecuencia de clase lo primero que se hace es calcular el rango de la serie de valores así:

$$X_M = 52, X_m = 14, UM = 16 - 14 = 2, UM = 2; R = X_M - X_m + 1 UM.$$

En este caso se puede observar que la unidad de medida es puesto que la serie de valores sigue una secuencia de números múltiplos de dos.

$$R = 52 - 14 + 2; R = 40. \text{ El número total de datos es } 100 \text{ y el } \text{Log. } 100 = 2.0.$$

Después de calculado R se calcula el Ic utilizando el método de Sturges así:

$$Ic = \frac{R}{1 + 3,322 \cdot \log N} = \frac{40}{1 + 3,322 \times 2.0} = \frac{40}{7.64} = 5.24 \therefore Ic = 5.24.$$

El Ic en este caso es 5.24 y si se hace el redondeo al impar más cercano, entonces el Ic = 5.0. Ahora se buscare el limite inferior de la primera clase que integrará la distribución, esta deberá ser un múltiplo del Ic y que sea menor o igual al valor más pequeño de la serie de valores, en este caso el limite

inferior será 10 que es un múltiplo de 5 y no es mayor que el menor valor de la serie de datos, entonces la primera clase se iniciara así: 10 —14 y el resto de las mismas seguirán la secuencia lógica.

La siguiente distribución presenta los diferentes componentes que la integran, se recomienda al estudiante realizar todos esos cálculos necesarios para conformar la misma

Clases	$f_i$	$fa<que$	$fa>que$	$\dot{X}$	$fr$	$fr \%$	$far<que$	$far \%<que$
10—14	5	5	100	12	0.05	5.0	0.05	5.0
15—19	13	18	95	17	0.13	13.0	0.18	18.0
20—24	22	40	82	22	0.22	22.0	0.40	40.0
25—29	12	52	60	27	0.12	12.0	0.52	52.0
30—34	19	71	48	32	0.19	19.0	0.71	71.0
35—39	9	80	29	37	0.09	9.0	0.80	80.0
40—44	9	89	20	42	0.09	9.0	0.89	89.0
45—49	7	96	11	47	0.07	7.0	0.96	96.0
50—54	4	100	4	52	0.04	4.0	1.00	100.0
<b>Total</b>	<b>100</b>							

Se recomienda al estudiante realizar todos los cálculos necesarios para completar el cuadro.

### Gráficos de la distribución de frecuencias de clases

Con los cuadros de las distribuciones de frecuencias se pueden elaborar varios tipos de gráficos, los más utilizados son: Los *histogramas*, los *polígonos de frecuencias* y la *ojiva* o *polígono de frecuencias acumuladas*. El método más utilizado para graficar los datos de una distribución de frecuencia es el histograma.

### Histograma

El histograma es un diagrama en forma de columna, muy parecido a los gráficos de barras. Se define como un conjunto de rectángulos paralelos, en el que la base representa la clase de la distribución y su altura la magnitud que alcanza la frecuencia de la clase correspondiente. Son barras rectangulares levantadas sobre el eje de las abscisas del plano cartesiano utilizando escalas adecuadas para los valores que asume la variable en la distribución de frecuencia. El ancho de la base de los rectángulos es proporcional a cada clase de la distribución, de tal manera que, cuando la distribución tiene clases de igual el tamaño de todos los rectángulos tendrán bases iguales. Los lados del rectángulo se levantan sobre los puntos del eje de las  $x$  que corresponden a los límites de cada clase y la longitud de los mismos será igual a la frecuencia que tenga esa clase, los lados por lo tanto corresponden a la frecuencia de cada clase de la distribución de frecuencia.

Cuando se elaboran gráficas estadísticas en el plano cartesiano es recomendable que en el eje de las *ordenadas* se representen las frecuencias y en el eje de las *abscisas* las variables independiente. El eje de las *y* que representa las frecuencias debe empezar siempre en cero. Es importante señalar que la longitud del eje de las *y* que representa la altura tenga el 75 % de la longitud del eje de las *x*, es decir, si las variable independientes ocupan en el eje *x* 8 cm, la máxima altura que ocuparan las frecuencias en el eje *y* tendrá que ser de 6 cm.

### Pasos para construir un histograma

1.- Se trazan dos ejes de coordenadas, el de abscisas y el de ordenada.. Se coloca sobre el eje de las *x* los límites inferiores de cada clase y el último límite superior de la distribución, y sobre el eje de las *y* se coloca la magnitud de la frecuencia de cada clase.

3.- Se trazan perpendiculares por los límites de cada clase, la altura de las perpendiculares será igual a la frecuencia de cada clase; y para finalizar se unen las dos perpendiculares que representan a cada clase, el resultado final será el histograma.

### Problemas

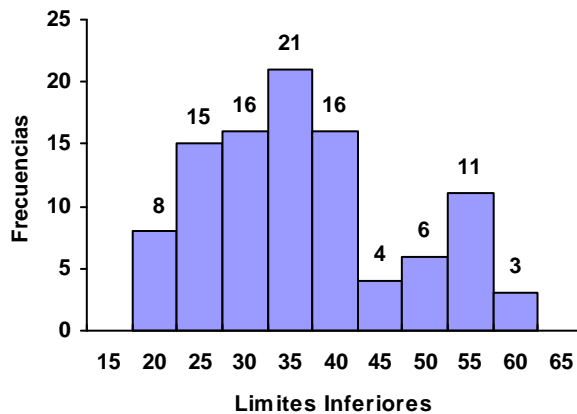
1.- Dada la siguiente distribución de frecuencia correspondiente al consumo de arroz durante un trimestre por un grupo de familias de una Urbanización de Ibagué. Elabore un histograma. Se recomienda al estudiante elaborar los cálculos respectivos.

Clases	<i>f<sub>i</sub></i>
20—24	8
25—29	15
30—34	16
35—39	21
40—44	16
45—49	4
50—54	6
55—59	11
60—64	3
<b>Total</b>	<b>100</b>

Cuadro resumen:

Clases	$f_i$	$fa < que$	$fa > que$	$\dot{X}$	$fr$	$fr \%$	$far < que$	$far \%$
20—24	8	8	100	22	0.08	8.0	0.08	8.0
25—29	15	23	92	27	0.15	15.0	0.23	23.0
30—34	16	39	77	32	0.16	16.0	0.39	39.0
35—39	21	60	61	37	0.21	21.0	0.60	60.0
40—44	16	76	40	42	0.16	16.0	0.76	79.0
45—49	4	80	24	47	0.04	4.0	0.80	80.0
50—54	6	86	20	52	0.06	6.0	0.86	86.0
55—59	11	97	14	57	0.11	11.0	0.97	97.0
60—64	3	100	3	62	0.03	3.0	1.00	100.0
<b>Total</b>	<b>100</b>							

Histograma correspondiente a las horas extras laboradas por un grupo de obreros petroleros.



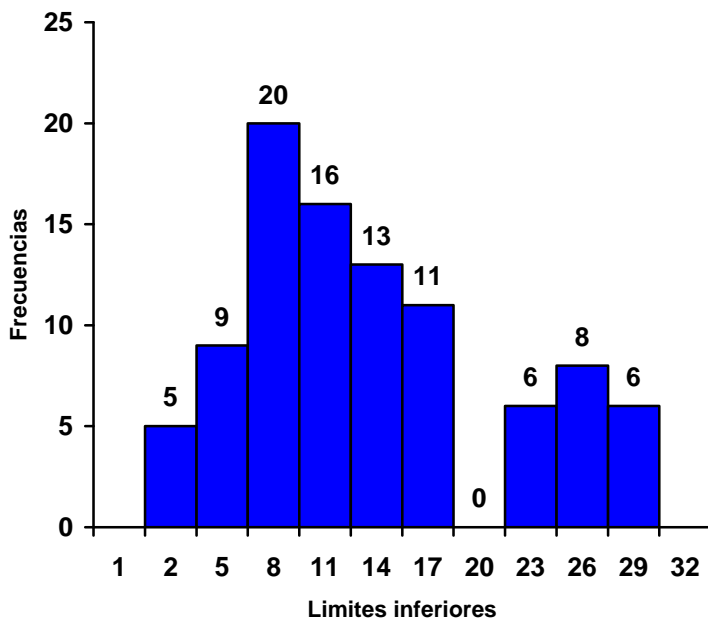
2.- La Distribución e frecuencia que se presentan a continuación corresponden al consumo de carne de res en kg.(en un trimestre) de un grupo de familia de un barrio de la ciudad de Ibagué. Con los mismos elabore un Histograma. Elabore los cálculos.

Cuadro resumen:

Clases	fi
2—4	5
5—7	9
8—10	20
11—13	16
14—16	13
17—19	11
20—22	0
23—25	6
26—28	8
29—31	6
<b>TOTAL</b>	<b>100</b>

Clases	fi	$\dot{X}$	fa<que	fa>que
2—4	5	3	5	100
5—7	9	6	14	95
8—10	20	9	34	86
11—13	16	12	50	66
14—16	13	15	63	50
17—19	11	18	74	37
20—22	0	21	80	26
23—25	6	24	86	20
26—28	8	27	94	14
29—31	6	30	100	6
<b>TOTAL</b>	<b>100</b>			

**Histograma correspondiente al consumo de carne de res, en kg., por un grupo de familias de un barrio de Ibagué.**



### Polígono de frecuencia

Es un diagrama de líneas que representa los puntos medios y las respectivas frecuencias de una distribución de frecuencia de clase. Es una representación gráfica cerrada de una distribución de frecuencia. Es otra de las formas de graficar los valores de una distribución de frecuencia de clase.

No existe ninguna razón estadística para seleccionar los polígonos de frecuencia en vez de los histogramas o viceversa, los histogramas simplemente representan una manera de graficar y los polígonos de frecuencia otra; la diferencia entre ambos radica en que una barra vertical rectangular representa una clase y su frecuencia en el histograma y un punto cumple la misma función en el polígono de frecuencia.

### Pasos para elaborar un polígono de frecuencia

- 1.- Se dibuja un plano cartesiano.
- 2.- Se traza sobre el eje de las abscisas, a distancias iguales, los puntos medios de las diferentes clases de la distribución de frecuencia.



3.- Se levantan perpendiculares por cada una de las marcas de clase, con una longitud igual a la frecuencia de cada una de las clases que integran la distribución de frecuencia. Al final de cada perpendicular se marca un punto.

4.- Los puntos resultantes se unen por medio de una línea recta obteniéndose una línea poligonal.

5.- Con la finalidad de cerrar la línea poligonal se agrega una clase imaginaria con frecuencia cero a cada extremo de la distribución de frecuencia, por tal motivo ambos extremos del polígono se cortan con el eje de las abscisas.

También se puede elaborar un polígono de frecuencia después de haber graficado un histograma; si se determina el punto medio de cada rectángulo de un histograma y esos puntos medios se unen por medio de segmentos de recta dan como resultado el polígono de frecuencia.

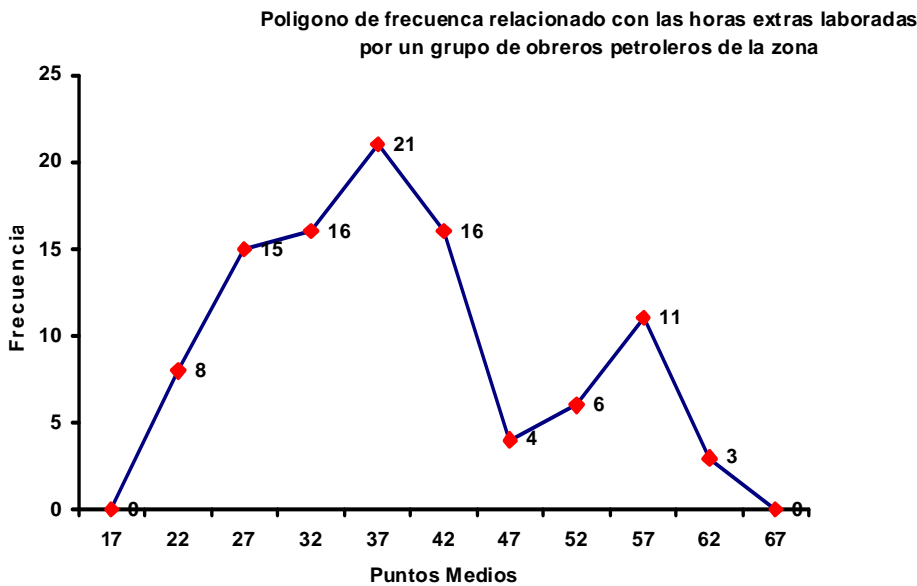
## Problemas resueltos

1.- Sea la siguiente distribución correspondiente a las horas extras trabajadas por un grupo de obreros petroleros de la zona durante un mes. Con esos datos elabore un polígono de frecuencia. Elabore los cálculos respectivos.

Clases	$f_i$
20—24	8
25—29	15
30—34	16
35—39	21
40—44	16
45—49	4
50—54	6
55—59	11
60—64	3
Total	100

Cuadro resumen:

Clases	$f_i$	$\dot{X}$	$fa<que$	$fa>que$
20—24	8	22	8	10
25—29	15	27	23	92
30—34	16	32	39	77
35—39	21	37	60	61
40—44	16	42	76	40
45—49	4	47	80	24
50—54	6	52	86	20
55—59	11	57	97	14
60—64	3	62	100	3
<b>Total</b>	<b>100</b>			



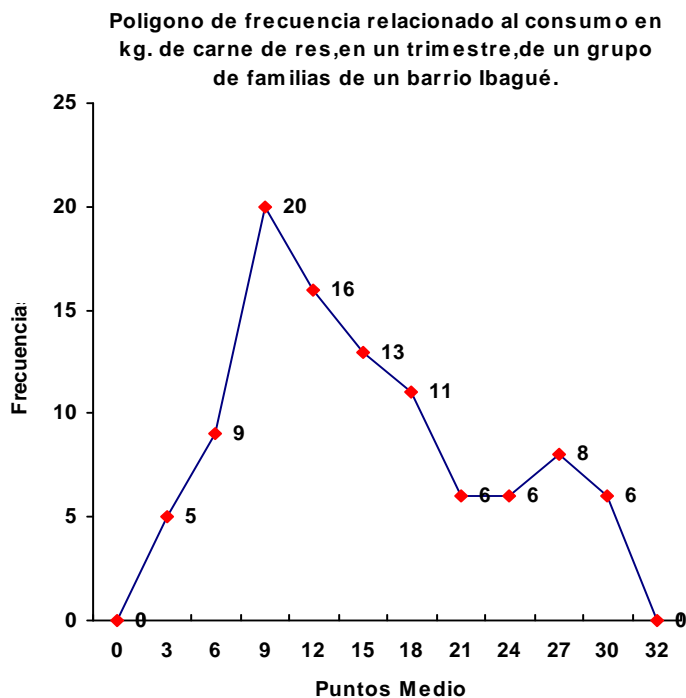
Observe que los puntos medios, 17 y 67 del polígono son imaginarios, se utilizan para cerrar la línea poligonal, lo que da origen al polígono de frecuencia.

2.- La Distribución de frecuencia que se presentan a continuación corresponden al consumo de carne de res en kg., en un trimestre, de un grupo de familia de un barrio de la ciudad de Ibagué. Con los mismos elabore un polígono de frecuencia. Realice los cálculos respectivos.

Clases	$f_i$
2—4	5
5—7	9
8—10	20
11—13	16
14—16	13
17—19	11
20—22	6
23—25	6
26—28	8
29—31	6
<b>TOTAL</b>	<b>100</b>

Cuadro resumen:

Clases	$f_i$	$\dot{X}$	$fa < que$	$fa > que$
2—4	5	3	5	100
5—7	9	6	14	95
8—10	20	9	34	86
11—13	16	12	50	66
14—16	13	15	63	50
17—19	11	18	74	37
20—22	6	21	80	26
23—25	6	24	86	20
26—28	8	27	94	14
29—31	6	30	100	6
<b>TOTAL</b>	<b>100</b>			



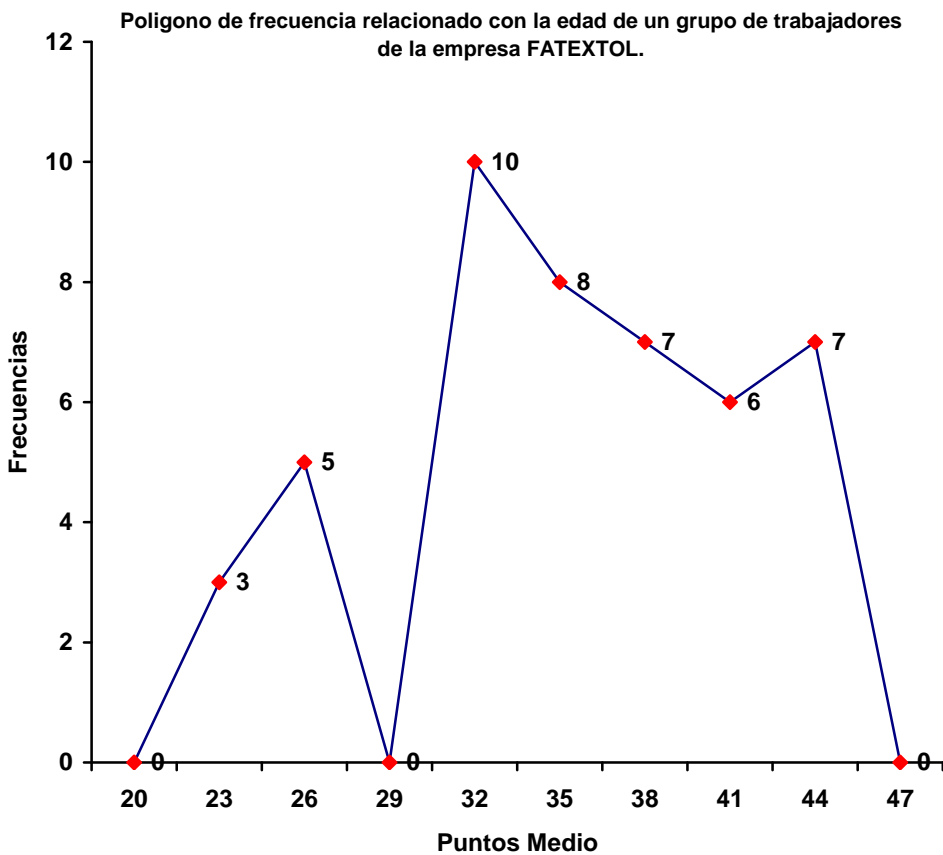
Los puntos medios primero y último del polígono son imaginarios, se puede observar que tienen como frecuencia cero, los mismos se utilizan para cerrar la línea poligonal y el área que se ubica debajo de esta es la correspondiente al polígono de frecuencia.

3.- La siguiente distribución de frecuencia corresponde a la edad de un grupo de trabajadores de la empresa FATEXTOL. Elabore un polígono de frecuencia. Realice los cálculos respectivos.

Clases	$f_i$
22 —24	3
25— 27	5
28 —30	0
31— 33	10
34 —36	8
37 —39	7
40— 42	6
43 —45	7
<b>Total</b>	<b>46</b>

Cuadro resumen:

Clases	$f_i$	$\dot{X}$	$fa<que$	$fa>que$
22 —24	3	23	3	46
25— 27	5	26	8	43
28 —30	0	29	8	38
31— 33	10	32	18	38
34 —36	8	35	26	28
37 —39	7	38	33	20
40— 42	6	41	39	13
43 —45	7	44	46	7
<b>Total</b>				



Se puede observar que los puntos medios 20 y 47 pertenecen a clases imaginarias y su función es la de cerrar la línea poligonal, dando origen al polígono de frecuencia.

## Polígono de frecuencia acumulativa u ojiva

Es una gráfica que se elabora con los valores de las frecuencias acumulados (menor que y mayor que) y los límites de las clases de una distribución de frecuencia. El polígono de frecuencia acumulada se le conoce comúnmente como *ojiva*. La ojiva es una representación gráfica que consiste en una línea, que puede ser *ascendente* o *descendente* y se utiliza para representar las distribuciones de frecuencias acumuladas *menor que* y *mayor que*, según los datos utilizados. En los estudios de análisis estadísticos la ojiva es de gran utilidad porque permite obtener con gran aproximación cierta información requerida, en un momento determinado.

### Pasos para elaborar una ojiva

- 1.- Se trazan los ejes de abscisa y ordenada del plano cartesiano.
- 2.- Se marca sobre el eje de las  $x$  los límites superiores de cada clase, si se trata de la *ojiva menor que* o los límites inferiores de la misma si se desea graficar la *ojiva mayor que*, curva descendente, y sobre el eje de las ordenadas se marcan las magnitudes de las frecuencias acumuladas *menor que* (curva ascendente) de cada clase, si se quiere graficar la *ojiva menor que* o las frecuencias acumuladas *mayor que* de cada clase, si se desea graficar la *ojiva mayor que*.
- 3.- Se trazan perpendiculares por los límites superiores o inferiores de cada clase, según la ojiva que se desea graficar, la altura de la perpendicular tiene que ser igual a la frecuencia acumulada *menor que* o *mayor que* de la clase respectiva y al final de la misma se marca un punto.
- 4.- por último se unen todos los puntos por medio de segmentos de recta, dando origen a la ojiva.

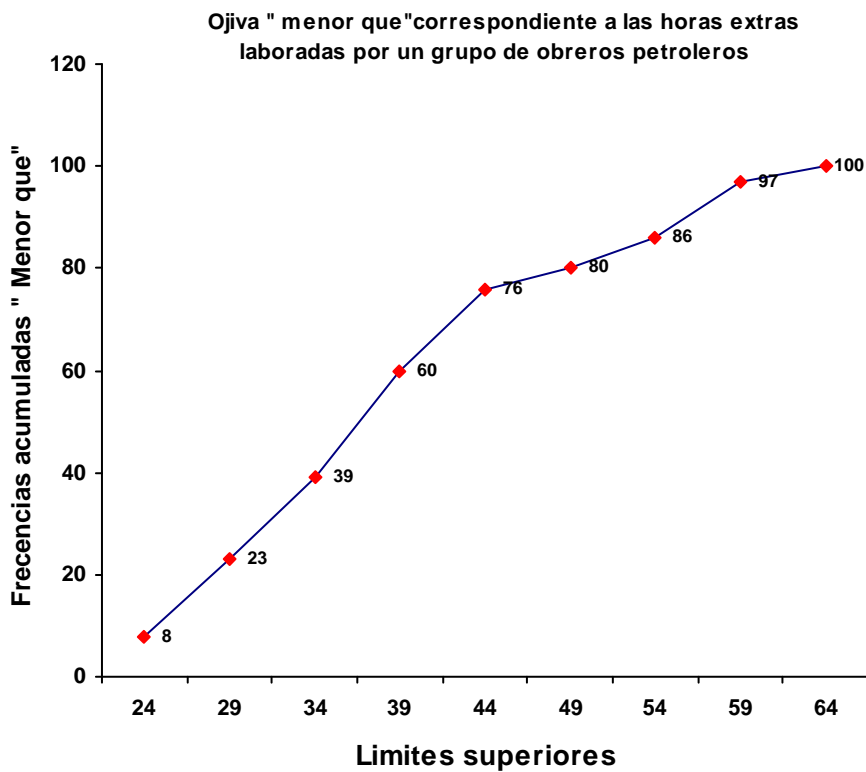
Nota.- algunos investigadores consideran que la *ojiva menor que* y la *mayor que* se deberían graficar con los límites inferiores de clase y al final el último límite de la distribución.

## Problemas resueltos

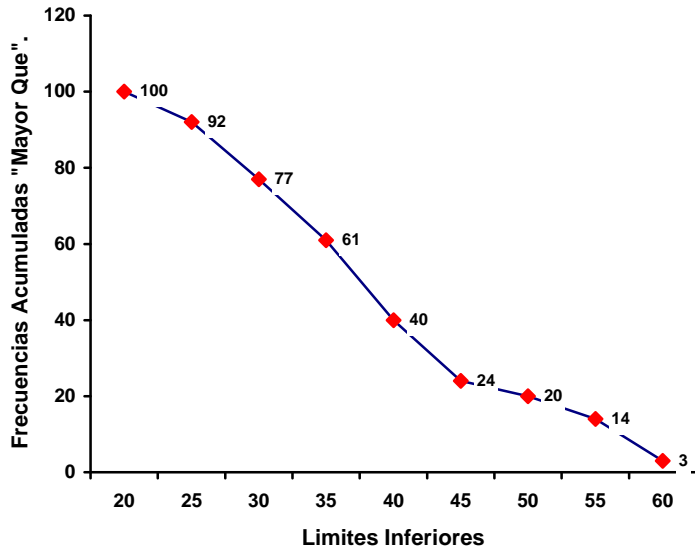
1.- Sea la siguiente distribución correspondiente a las horas extras laboradas por un grupo de obreros petroleros de la zona durante un mes. Con esos datos elabore un polígono de frecuencia acumulada *menor que* y otro *mayor que*. Realice los cálculos respectivos para completar el siguiente cuadro.

Clases	$f_i$
20—24	8
25—29	15
30—34	16
35—39	21
40—44	16
45—49	4
50—54	6
55—59	11
60—64	3
Total	100

Clases	$f_i$	PM	$fa<que$	$fa>que$
20—24	8	22	8	10
25—29	15	27	23	92
30—34	16	32	39	77
35—39	21	37	60	61
40—44	16	42	76	40
45—49	4	47	80	24
50—54	6	52	86	20
55—59	11	57	97	14
60—64	3	62	100	3
Total	100			



Ojiva "mayor que" correspondiente a las horas extraslaboradas por un grupo de obreros petroleros.

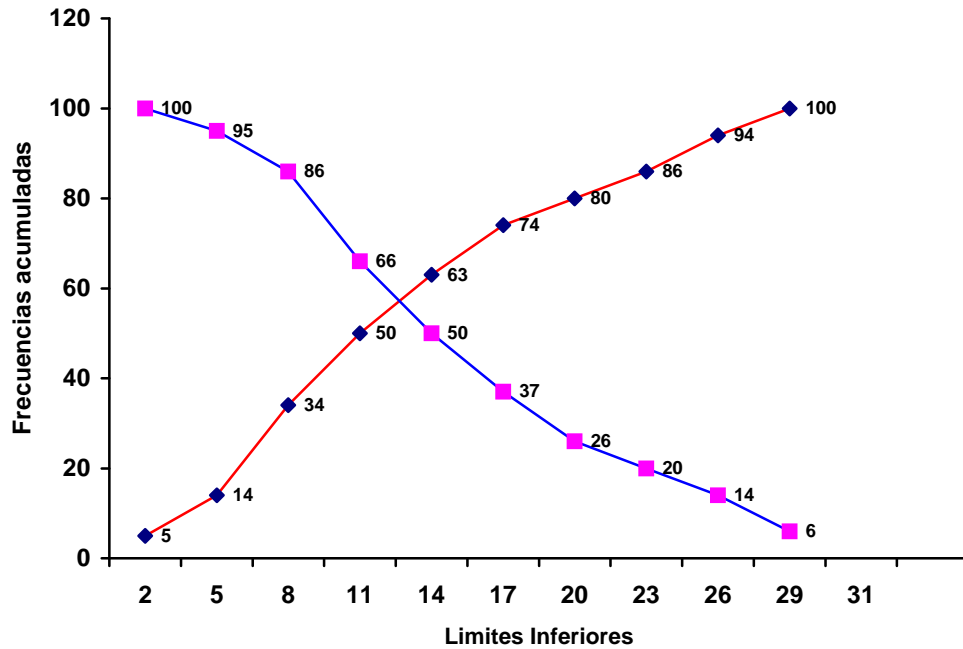


2.- La Distribución de frecuencia que se presentan a continuación corresponden al consumo de carne de res en kg., en un trimestre, de un grupo de familia de un barrio de la ciudad de Ibagué. Con los mismos elabore un polígono de frecuencia Acumulado, *menor que* y otro *mayor que*. Realice los cálculos respectivos para completar el siguiente cuadro.

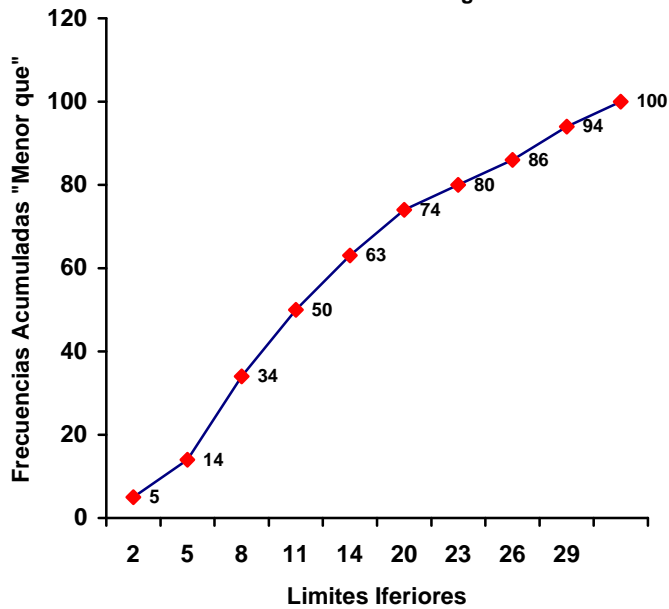
Clases	$f_i$		Clases	$f_i$	PM	$fa < que$	$fa > que$
2—4	5		2—4	5	3	5	100
5—7	9		5—7	9	6	14	95
8—10	20		8—10	20	9	34	86
11—13	16		11—13	16	12	50	66
14—16	13		14—16	13	15	63	50
17—19	11		17—19	11	18	74	37
20—22	6		20—22	6	21	80	26
23—25	6		23—25	6	24	86	20
26—28	8		26—28	8	27	94	14
29—31	6		29—31	6	30	100	6
<b>TOTAL</b>	<b>100</b>		<b>TOTAL</b>	<b>100</b>			



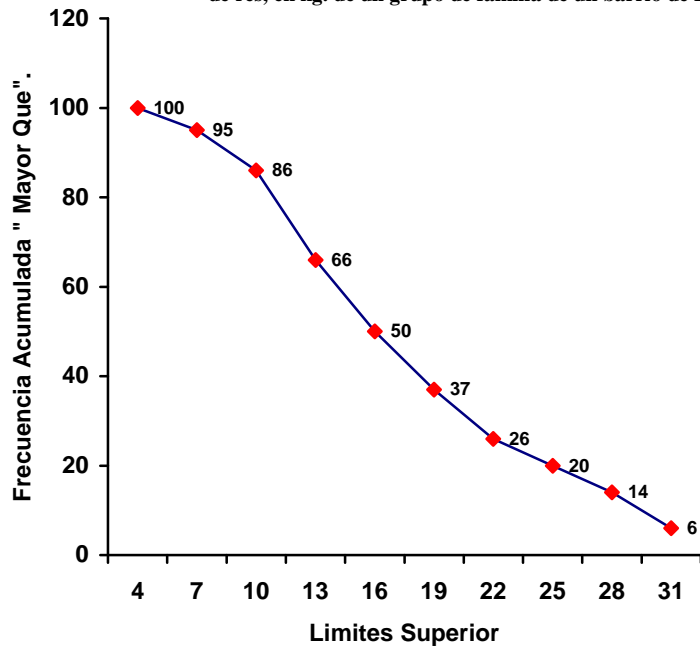
Ojiva "Menor Que" y "Mayor que" correspondiente al consumo de carne de res, en kg. de un grupo de familias de un barrio de bagué.



Ojiva "Menor Que" correspondiente al consumo de carne de res, en kg., de un grupo de familia de un barrio de Ibagué.



Ojiva "Mayor Que" correspondiente al consumo de carne de res, en kg. de un grupo de familia de un barrio de Ibagué.



## BIBLIOGRAFÍA

- Benavente del Prado, Arturo Núñez(1992): **Estadística Básica par Planificación**.Editorial Interamericana. 6ª. Edición. México.
- Berenso, Mark.(1.992): **Estadística Básica en Administración**. Editorial. Harla. Cuarta Edición. México.
- Best,J. W. (1987): *Como Investigar en Educación*. Editorial Morata. Madrid – España.
- Budnick Frank S. (1992): *Matemáticas Aplicadas para Administración, Economía y Ciencias Sociales*. Tercera Edición. Editorial McGaw-Hill Interamericana de México, S.A de C.V. México.
- Caballero, Wilfredo (1975): *Introducción a la Estadística*. Editorial ICA. Costa Rica.
- Cadoche, L. S.; G. Stegmayer, J. P. Burioni y M. De Bernardez (1998). Material del Seminario de *Encuestas en Educación*, impartido vía internet por parte de la Universidad Nacional del Litoral, en Santa Fe, y de la Universidad Tecnológica Nacional, Regional Santa Fe, en la República de Argentina.
- Castañeda J., J.(1991): *Métodos de Investigación 2*. Editorial McGraw-Hill. México.
- Carono, R., Minujin, A. y Vera, G.(1982): **Manual de técnicas de evaluación yajuste de información Estadísticas**. Fondo de cultura económica. México.
- Chao, L.(1993): **Estadística para la Ciencia Administrativa**. Editorial McGraw –Hill. 4<sup>ta</sup> Edición. Colombia
- CHOU, YA-LUN (1972): *Análisis Estadístico*. Editorial Interamericana. México
- DANIEL WAYNE, W. y Otros (1993): *Estadística con Aplicación a las Ciencias Sociales y a la Educación* Editorial McGraw-Hill Interamericana de México, S.A. de C.V. México.
- De Oteyza de O., E; Emma Lam O., Carlos Hernández G. y Ángel M. Carrillo H. (1998). *Temas Selectos de Matemáticas*. Prentice Hall. México
- Enciclopedia Microsoft Encarta 2003 (2003): *Censo- Cuestionario- Encuesta. Estadística*. Editorial Microsoft corporation. USA.
- ERKIN KREYSZIA (1978): *Introducción a la Estadística Matemática*. Editorial Limusa, S.A. México.

- FREUD J: E. y Otros (1990): *Estadística para la Administración con Enfoque Moderno*. Editorial, S.A. México.
- Gomes Rondón, Francisco (1985): *Estadística Metodologica*: Ediciones Frigor. Caracas.
- González, Nijad H. (1986): *Métodos estadísticos en Educación*. Editorial Bourgeón, Caracas.
- Guilford, J. Y Fruchter, B. (1984): *Estadística aplicada a la Psicología y la Educación*. Editorial McGraw-Hill Latinoamericana, S. A., Bogotá.
- Hamdan González, Nijad (1986): *Métodos Estadísticos en Educación*. Editorial BourgeónC.A. Caracas – Colombia.
- KEVIN, RICHARD I. (1988): *Estadística para Administradores*. Editorial Hispanoamericana. México.
- LARSON HAROLD, J. (1985): *Introducción a la Teoría de Probabilidades e inferencia Estadística*. Editorial Limusa. México.
- LEHMANN, CHARLES H. (1995): *ÁLGEBRA*. Editorial limusa, S.A. DE C.V. Grupo Noriega Editores. México.
- LEITHOLD, LOUIS (1992): *El Cálculo con Geometría Analítica*. Editorial HARLA México.
- LINCON L., CHAO (1996): *Estadística para Ciencias Administrativas*. Cuarta edición. Editorial McGraw-Hill. Usa.
- Lenin, R.y Kubin, D.(1992): **Estadística para Administradores**. Editorial Hispanoamérica. VI edición. México.
- LOPEZ CASUSO, R. (1984): *Introducción al Cálculo de Probabilidades e Inferencia Estadística*. Editorial Instituto de Investigaciones Económicas, UCAB. Caracas- Colombia.
- Mason, Robert (1.992): **Estadística para la Administración y Economía**. Ediciones Alfaomega S.A.N. México.
- MENDENNAF, W. y OTROS (1981): *Estadística para Administradores y Economía*. Editorial Iberoamericana. México.
- Mode, Elmer B. (1988): *Elementos de Probabilidades y Estadística* Editorial Reverte Mejicana. México.
- Murria, R.(1993): *Estadística. Edición Interamericana*. 2<sup>da</sup> Edición. México.
- PARZEN, E. (1986): *Teoría Moderna de Probabilidades y sus Aplicaciones* Editorial Limusa: México
- PUGACHEV, V. S. (1973): *Introducción a la Teoría de Probabilidades* Editorial Mir. Moscú.

- Rivas González, Ernesto(1980): *Estadística General*. Ediciones de la Biblioteca UCV. Caracas – Colombia.
- Soto Negrin, Armando (1982): *Iniciación a la estadística*. Editorial José Martí. Caracas – Colombia.
- Stephen P., Shao (1986): *Estadística para Economistas y Administradores de Empresa*. Editorial Herreros Hermanos, Sucs., S.A., México.
- Stevenson, William(1991): *Estadística para la Administración y Económica*. Editorial Harla. México.
- Universidad Nacional Experimental “Simón Rodríguez” (1983): *Estadística 1*. Ediciones UNESR, Caracas.
- WALPOLE, R. y Myers, R. (1987): *Probabilidad y Estadística para Ingenieros*. Editorial Interamericana. México.
- Webster, Allen L. (1996): *Estadística Aplicada a la Empresa y la Economía*. Editorial Irwin. Segunda edición. Barcelona – España.
- Weimer, Richard C. (1996) *Estadística*. Compañía Editorial Continental, SA de CV. México.
- Wonnacott, T. H. y Wonnacott, R: J. (1989): *Fundamentos de Estadística para Administración y Economía*. Editorial LIMUSA. México.