



ADRIANA GUERRERO PEÑA  
MARÍA VICTORIA BUITRAGO CARDONA  
MARÍA CURIÉSES PAULETE

# Estadística básica



JTM

2  
EDICIÓN

Corregida y  
aumentada

# **ESTADÍSTICA BÁSICA**

**ADRIANA GUERRERO PEÑA**

**MARÍA VICTORIA BUITRAGO**

**MARÍA CURIESES PAULETE**





INSTITUTO TECNOLÓGICO METROPOLITANO  
Institución Universitaria

ESTADÍSTICA BÁSICA

© Adriana Guerrero Peña  
© María Victoria Buitrago  
© María de los Ángeles Curieses Paulete  
© Instituto Tecnológico Metropolitano

1a. Edición: septiembre de 2007  
3a. Reimpresión (revisada y aumentada): julio de 2009  
2a. Edición: febrero de 2010

ISBN: 978-958-8351-77-3  
Hechos todos los depósitos legales

Dirección editorial  
Fondo Editorial ITM

Corrección de textos  
Lucía Inés Valencia

Diagramación y montaje  
L. Vieco e Hijos Ltda.

Impreso y hecho en Medellín, Colombia

*Las opiniones, originalidad y citas del texto son de la responsabilidad del autor. El Instituto salva cualquier obligación derivada del libro que se publica. Por lo tanto, ella recaerá única y exclusivamente en el autor.*

Instituto Tecnológico Metropolitano  
Calle 73 No. 76A 354  
Tel.: (574) 440 51 60  
Fax: 440 52 52  
www.itm.edu.co  
Medellín - Colombia

# CONTENIDO

PRESENTACIÓN .....	11
INTRODUCCIÓN .....	13
CAPÍTULO 1 .....	15
1. Conceptos previos .....	15
1.1 Definición de sumatoria .....	15
1.2 Propiedades .....	15
1.3 Actividad .....	17
1.4 Ejercicios .....	17
HISTORIA DE LA ESTADÍSTICA .....	21
CAPÍTULO 2 .....	23
2. Estadística descriptiva .....	23
2.1 Definiciones de términos estadísticos .....	23
2.2 Conceptos estadísticos .....	23
2.3 Ejercicios: conceptos estadísticos .....	25
2.4 Métodos gráficos para describir información .....	29
2.5 Tablas de datos .....	30
2.5.1 Tabla de entrada de datos .....	30
2.5.2 Tabla de contingencia .....	30
2.5.3 Tabla de frecuencias .....	31
2.5.3.1 Tabla de frecuencia simple .....	31
2.5.3.2 Tabla de frecuencia para datos agrupados .....	33
2.6 Métodos gráficos .....	36
2.6.1 Gráfico de pastel o circular .....	36
2.6.2 Diagrama de barras .....	37
2.6.3 Histograma .....	38
2.6.4 Polígono de frecuencias .....	38
2.6.5 Actividad .....	39
2.7 Ejercicios: descripción de la información por métodos gráficos .....	39
2.8 Ejercicios: descripción de la información por tablas .....	42

CAPÍTULO 3 .....	49
3. Métodos numéricos para describir información.....	49
3.1 Medidas de tendencia central .....	50
3.1.1 Media aritmética.....	50
3.1.1.1 Actividad .....	51
3.1.2 Mediana .....	51
3.1.2.1 Cálculo de la mediana para datos sin agrupar .....	51
3.1.2.2 Cálculo de la mediana para datos agrupados.....	51
3.1.3 Moda.....	52
3.1.3.1 Moda para datos sin agrupar .....	52
3.1.3.2 Moda para datos agrupados.....	52
3.2 Medidas de posición.....	53
3.2.1 Los cuartiles .....	53
3.2.2 Los deciles.....	53
3.2.3 Los percentiles.....	54
3.2.4 Actividad .....	54
3.3 Medidas de dispersión.....	54
3.3.1 La varianza .....	54
3.3.1.1 Actividad .....	55
3.3.2 Desviación estándar .....	55
3.3.3 Coeficiente de variación.....	55
3.3.4 Regla empírica.....	55
3.4 Medidas de forma.....	57
3.4.1 Medidas de asimetría.....	58
3.4.2 Medidas de curtosis o apuntamiento .....	62
3.5 Ejercicios sobre medidas de tendencia central, posición, variación y forma .....	65
 CAPÍTULO 4 .....	 75
4. Regresión y correlación.....	75
4.1 Regresión.....	76
4.1.1 Método de los mínimos cuadrados.....	77
4.2 Correlación .....	81
4.2.1 Coeficiente de correlación de Pearson ( $r$ ) .....	81
4.2.2 Coeficiente de determinación ( $R^2$ ) .....	83
4.3 Ejercicios - regresión y correlación.....	84

CAPÍTULO 5 .....	89
5. Probabilidades .....	89
5.1 Conceptos básicos de probabilidad .....	89
5.1.1 Actividad .....	90
5.2 Modelos de probabilidad .....	90
5.2.1 Modelo clásico o A PRIORI .....	90
5.2.2 Modelo subjetivo .....	91
5.2.3 Modelo de probabilidad empírico a posteriori .....	91
5.3 Complemento de un evento .....	91
5.4 Actividad .....	91
5.5 Reglas de probabilidad .....	92
5.5.1 Regla de la adición .....	92
5.5.1.1 Probabilidad de eventos mutuamente excluyentes .....	92
5.5.1.2 Actividad .....	92
5.5.2 Probabilidad condicional .....	92
5.5.3 Regla de la multiplicación .....	93
5.5.3.1 Eventos independientes .....	93
5.5.3.2 Actividad .....	93
5.5.4 Teorema de Bayes .....	94
5.6 Análisis combinatorio .....	96
5.6.1 Técnica de la multiplicación .....	96
5.6.2 Técnica de la permutación .....	97
5.6.3 Técnica de la combinación .....	99
5.7 Ejercicios de conjuntos y análisis combinatorio .....	100
5.8 Ejercicios de probabilidad .....	102
CAPÍTULO 6 .....	109
6. Variables aleatorias .....	109
6.1 Variables aleatorias discretas .....	110
6.1.1 Distribución de probabilidad de las variables aleatorias discretas .....	110
6.1.1.1 Propiedades de las v.a discretas $f(x)$ .....	111
6.1.2 Distribucion de probabilidad acumulada $f(x)$ .....	111
6.1.3 Valor esperado de una variable aleatoria discreta .....	113
6.2 Ejercicios - Variables aleatorias .....	113

CAPÍTULO 7 .....	115
7. Distributions de de probabilidad discreta .....	115
7.1 Distribución binomial.....	115
7.1.1 Función de probabilidad de la V.A. binomial.....	116
7.1.2 Parámetros de la distribución binomial.....	116
7.1.3 Función de distribución de la V.A. binomial.....	116
7.2 Distribución hipergeométrica .....	118
7.2.1 Parámetros de la distribución hipergeométrica .....	119
7.3 Distribución de Poisson.....	122
7.3.1 Parámetros de la distribución Poisson.....	123
7.4 Ejercicios: modelos de probabilidad discretas .....	123
CAPÍTULO 8 .....	127
8. Distributions de de probabilidad continuas.....	127
8.1 Distribucion normal .....	127
8.1.1 Representación gráfica .....	128
8.1.2 Distribución normal estándar .....	128
8.2 Aproximaciones.....	132
8.2.1 Aproximación de la binomial mediante la distribución Poisson....	132
8.2.2 Aproximación normal a la distribución de probabilidad binomial ...	133
8.3 Ejercicios: sobre modelos de probabilidad continua y aproximaciones.....	134
CAPÍTULO 9 .....	137
9. Tablas de distribuciones.....	137
9.1 Tabla de la distribución poisson acumulada .....	137
9.2 Tabla de la distribución binomial acumulada .....	141
9.3 Tabla de la distribución normal .....	154
BIBLIOGRAFÍA .....	157

## PRESENTACIÓN

Con esta guía se pretende que el estudiante aprenda a manejar mediante el uso de procedimientos estadísticos, herramientas que podrá utilizar en el mejoramiento de su desempeño laboral y en muchos otros aspectos de la vida diaria.

En la guía desarrollaremos los conceptos básicos que permitan la comprensión de los métodos empleados en la estadística descriptiva, para la solución de casos relacionados con el análisis de información.

ESTADÍSTICA BÁSICA estudia la solución de problemas sobre probabilidades que requieren una cabal comprensión de algunas reglas fundamentales de probabilidad.

El texto introduce al estudiante en el mundo de la estadística, desarrollando competencias que lo ayudan a mejorar su desempeño laboral y asuntos de su profesión, para que tome decisiones y solucione problemas por medio de la aplicación de procedimientos estadísticos.



## INTRODUCCIÓN

La materia estadística e introducción a la probabilidad se adapta a las necesidades del curso semestral de las tecnologías e ingenierías, en las que se estudian los fundamentos de la estadística.

Con la guía ESTADÍSTICA BÁSICA se pretende llevar al alumno al mundo de la estadística, mostrándole el manejo de la información como recurso, y adquiere una nueva dimensión, dejando de ser una simple colección de hechos, para convertirse en una fuente de alimentación del proceso de toma de decisiones. Muestra cómo resolver problemas ante la incertidumbre, situación que todos enfrentamos tanto en el mundo de los negocios como en la vida cotidiana.

El orden que guarda este material en los capítulos siguientes es el enfoque pedagógico que tiene el Instituto Tecnológico Metropolitano en los microcurrículos de los diferentes programas.

Cada capítulo contiene una serie de ejercicios y aplicaciones que ayudan a los estudiantes a desarrollar habilidades para la aplicación de razonamientos estadísticos y resolver situaciones problemáticas e interpretación de resultados.

# CAPÍTULO 1

En este capítulo se presentan los principales conceptos básicos que permiten la comprensión de los métodos empleados en estadística descriptiva, para la solución de casos relacionados con el análisis de información.

## COMPETENCIA

Conceptualizar y reconocer la terminología básica en Estadística.

## EJE TEMÁTICO

Razones, proporciones y porcentajes, sumatoria y sus propiedades. Historia de la estadística.

## 1. CONCEPTOS PREVIOS

### 1.1 DEFINICIÓN DE SUMATORIA

Si  $x_1, x_2, x_3, \dots, x_n$  son  $n$  números, la suma de estos números  $x_1 + x_2 + x_3 + \dots + x_n$  se expresa simbólicamente mediante la escritura, esto es,

$$\sum_{i=1}^n x_i = x_1 + x_2 + x_3 + \dots + x_n$$

Así, por ejemplo, si  $x_i = i^3$  entonces,

$$\sum_{i=1}^5 x_i = \sum_{i=1}^5 i^3 = 1^3 + 2^3 + 3^3 + 4^3 + 5^3 = 225$$

### 1.2 PROPIEDADES

De la definición de  $\sum_{i=1}^n x_i$  se derivan las siguientes propiedades.

1. Si  $x_i = c$  (constante) entonces,

$$\sum_{i=1}^n x_i = \sum_{i=1}^n c = c + c + c + \dots + c = nc$$

Así, por ejemplo,  $\sum_{i=1}^5 4 = 4 + 4 + 4 + 4 + 4 = 5 * 4 = 20$

2. Si  $c$  es una constante que multiplica a cada una de las observaciones  $x_1, x_2, x_3, \dots, x_n$  entonces la suma de los  $n$  productos es igual a  $c$  multiplicada por la suma de las observaciones, esto es,

$$\sum_{i=1}^n cx_i = c \sum_{i=1}^n x_i$$

3. Si  $x_1, x_2, x_3, \dots, x_n$  e  $y_1, y_2, y_3, \dots, y_n$  son sucesiones de números, entonces

$$\sum_{i=1}^n (x_i + y_i) = \sum_{i=1}^n x_i + \sum_{i=1}^n y_i$$

4. Si  $x_1, x_2, x_3, \dots, x_n$  y  $y_1, y_2, y_3, \dots, y_n$  son dos sucesiones de números, entonces

$$\sum_{i=1}^n (x_i - y_i) = \sum_{i=1}^n x_i - \sum_{i=1}^n y_i$$

La *media aritmética* o simplemente *media* es la medida de posición más utilizada. La media aritmética representa el centro físico del conjunto de datos y se define como la suma de los valores observados, dividido por el total de observaciones. De una manera formal decimos que si  $x_1, x_2, x_3, \dots, x_n$  son  $n$  observaciones numéricas, entonces la media aritmética de estas  $n$  observaciones se denota y define de la siguiente manera:

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \dots + x_n}{n}$$

Entonces;  $\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$ , de donde se puede concluir que  $n\bar{x} = \sum_{i=1}^n x_i$

### 1.3 ACTIVIDAD

- Suponga que  $X = 1, 3, 5, 7, 9$  son las unidades producidas en una pequeña empresa en cada una de las primeras cinco horas de un día cualquiera de trabajo, suponga además que los costos de producción  $Y$  de las unidades producidas están dados por:  
 $Y = 3 + 2X$ .
  - Calcule los costos  $Y$  de producir 1, 3, 5, 7, 9 unidades.
  - Calcule  $\bar{X}$  y  $\bar{Y}$ . Compruebe que  $\bar{Y} = 3 + 2\bar{X}$ . ¿Qué puedes concluir?
- Si  $x_1, x_2, x_3, \dots, x_n$  y  $y_1, y_2, y_3, \dots, y_n$  son sucesiones de números y además  $y_i = a + bx_i$ , siendo  $a$  y  $b$  constantes entonces  $\bar{y}_i = a + b\bar{x}$ . Utilizando propiedades de sumatoria, justifica la propiedad anterior.

### 1.4 EJERCICIOS

- Suponga que se tienen dos sucesiones  $X$  y  $Y$ , apareadas de la manera siguiente:

X: 1 2 3 4 5

Y: 2 4 6 8 10

hallar el valor de cada una de las siguientes expresiones:

- $\left(\sum X\right)\left(\sum Y\right)$
  - $\left(\sum_{i=1}^4 x_i^2\right)\left(\sum_{i=1}^4 x_i\right)^2$
  - $\left(\sum XY\right) - \left(\sum X\right)\left(\sum Y\right)$
  - $\sum (X + Y)$
- Utilizar el símbolo  $\sum$  para representar las expresiones siguientes:
    - $x_1^2 + x_2^2 + x_3^2 + x_4^2$

- b.  $a x_1 + 2a x_2 + 3a x_3 + 4a x_4 + 5a x_5$
- c.  $c(x_1 + 1) + c(x_2 + 2) + c(x_3 + 3) + \dots + c(x_n + n)$
- d.  $(x_1 - 1)1^2 + (x_2 - 2)2^2 + (x_3 - 3)3^2 + (x_4 - 4)4^2 + (x_5 - 5)5^2$
3. Escriba con notación  $\sum$ .
- a.  $4 + 8 + 12 + 16 + 20$
- b.  $1 + r + r^2 + r^3 + \dots + r^n$
- c.  $3 + 9 + 27 + 81$
4. Hallar el valor de cada una de las expresiones siguientes:
- a.  $\sum_{i=1}^8 i$
- b.  $\sum_{i=1}^5 (i^2 + 3)$
- c.  $\sum_{k=1}^4 (k + 2)^2$
- d.  $\sum_{i=1}^7 (-1)^{i+2} (i^2 - 3)$
- e.  $\sum_{i=1}^4 \frac{4}{i(i+2)}$
5. Si  $\sum_{i=1}^4 x_i = 14$  y  $\sum_{i=1}^4 x_i^2 = 54$ ; calcular
- a.  $\sum_{i=1}^4 (x_i - 2)$
- b.  $\sum_{i=1}^4 (x_i + 2)$

c.  $\sum_{i=1}^4 (x_i - 2)^2$

d.  $\sum_{i=1}^4 3(2x_i^2 - 3x_i + 1)$

6. Calcular; utilizando propiedades

a.  $\sum_{i=1}^4 (i + 2)^2$

b.  $\sum_{i=1}^2 7i^2 - 3i$

c.  $\sum_{i=1}^3 \left( \frac{i^3 - i}{6} \right)$

7. Pruebe si las expresiones A y B son iguales:

$$A = \sum_{i=1}^n \sum_{j=1}^m (bx_{ij} - \bar{x}_{..})$$

$$B = (b-1)nm\bar{x}_{..} \quad \text{si, } \bar{X}_{..} = \frac{1}{nm} \sum_{i=1}^n \sum_{j=1}^m X_{ij}$$

# HISTORIA DE LA ESTADÍSTICA

## DESARROLLO HISTÓRICO DE LA ESTADÍSTICA

El término *estadística* se empleó para referirse a los datos del estado. En la Europa Medieval se utilizó la estadística para llevar datos sobre la población de un estado, los diezmos y los impuestos.

Hay algunos personajes fundamentales en el desarrollo de la estadística:

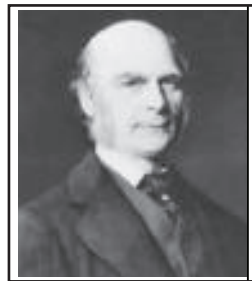
**ADOLPH QUETELET**  
(1796-1874)

Fue el primero en aplicar métodos modernos a un conjunto de datos y es considerado el padre de la estadística.



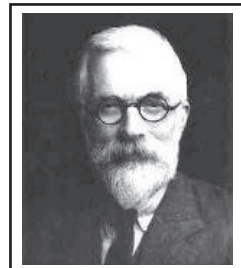
**FRANCIS GALTON**  
(1822-1911)

Creó la teoría de regresión que posteriormente refinó su discípulo Karl Pearson (1857-1936).



**RONALD FISHER**  
(1890-1962)

Es la figura más prominente hasta el presente y creó la teoría moderna de la estimación.



## CAPÍTULO 2

### COMPETENCIA

Describir y analizar gráficamente diferentes tipos de información.

#### EJE TEMÁTICO

Concepto de estadística como ciencia

Términos básicos

Métodos gráficos para describir información. Cualitativa

Gráfico de barras y circulares

Métodos gráficos para variables cuantitativas

Histogramas de frecuencia para datos no agrupados

Histograma de frecuencia para datos agrupados

### 2. ESTADÍSTICA DESCRIPTIVA

#### 2.1 DEFINICIONES DE TÉRMINOS ESTADÍSTICOS

*Estadística*: Es la ciencia de las matemáticas que se encarga de la selección, recolección, tabulación, presentación y análisis de la información que se utiliza en la toma de decisiones organizacionales.

La estadística se divide en dos grandes ramas, a saber:

*Estadística descriptiva*: Se dedica a la descripción, organización, síntesis y análisis de la información de interés; pero sin llegar a conclusiones fuertes o profundas, sobre la misma.

*Estadística inferencial*: Busca obtener conclusiones sólidas y más profundas, basado en el trabajo con muestras y su posterior generalización de resultados para la toma de decisiones y conclusiones sólidas.

La estadística inferencial nos permite inducir conclusiones de situaciones, sucesos o fenómenos previamente estudiados.

#### 2.2 CONCEPTOS ESTADÍSTICOS

Para comprender mejor la parte de estadística descriptiva, definiremos algunos términos muy utilizados, que nos servirán de guía para el desarrollo de los ejercicios.



**POBLACIÓN:** Se concibe como un conjunto total de elementos, datos, personas, atributos, medidas, acontecimientos u objetos, que poseen una o más características comunes y cuyas propiedades serán analizadas.

La población puede ser :

1. **POBLACIÓN FINITA:** Cuando es posible enumerar físicamente, TODOS los elementos que pertenecen a la población.
2. **POBLACIÓN INFINITA:** Cuando es imposible enumerar físicamente TODOS los elementos que pertenecen a la población.

**MUESTRA:** Es un subconjunto de la población, que se selecciona siguiendo ciertos procedimientos estadísticos, que se llama teoría de muestreo.

**PARÁMETRO:** Valor numérico que resume todos los datos de una población completa. Para determinar su valor es necesario utilizar la información poblacional completa.

**ESTADÍSTICO:** Es un valor numérico que resume todos los datos de una muestra y sirve como estimación del parámetro de la población.

**VARIABLE:** Es una característica, atributo o medida que se está analizando en un estudio estadístico.

La variable puede ser:

1. **VARIABLE CUALITATIVA:** Clasifica o describe un atributo o cualidad de los elementos de la población o muestra. (Atributos)
2. **VARIABLE CUANTITATIVA:** Los datos recolectados cuantifican un elemento de la población o muestra.

La variable cuantitativa puede ser:

- a) **VARIABLE CUANTITATIVA DISCRETA:** Cuando los valores que toma la variable son enteros que no se pueden partir.
- b) **VARIABLE CUANTITATIVA CONTINUA:** Cuando los valores que toma la variable se pueden partir.

**DATOS:** Conjunto de valores recolectados para la variable.

DATO: Valor de la variable asociado a un elemento de una población muestra.

EXPERIMENTO: Actividad planeada, cuyos resultados producen un conjunto de datos.

VARIABLE ALEATORIA: Una variable aleatoria (v.a) es una función que asigna a cada elemento de un espacio muestral un número real. O sea, una variable es aleatoria si toma diferentes valores como resultado de un experimento aleatorio.

## 2.3 EJERCICIOS: CONCEPTOS ESTADÍSTICOS

1. Utilizando la información suministrada en la situación siguiente, identifica los *conceptos básicos* de la estadística, tales como: población, muestra, variable, dato, datos, experimento, parámetro.

### SITUACIÓN 1

Un estudiante de estadística está interesado en determinar algo sobre el promedio del valor en pesos de los automóviles que pertenecen al cuerpo docente del Instituto Tecnológico Metropolitano.

La siguiente es la lista de información con la que el estudiante cuenta.

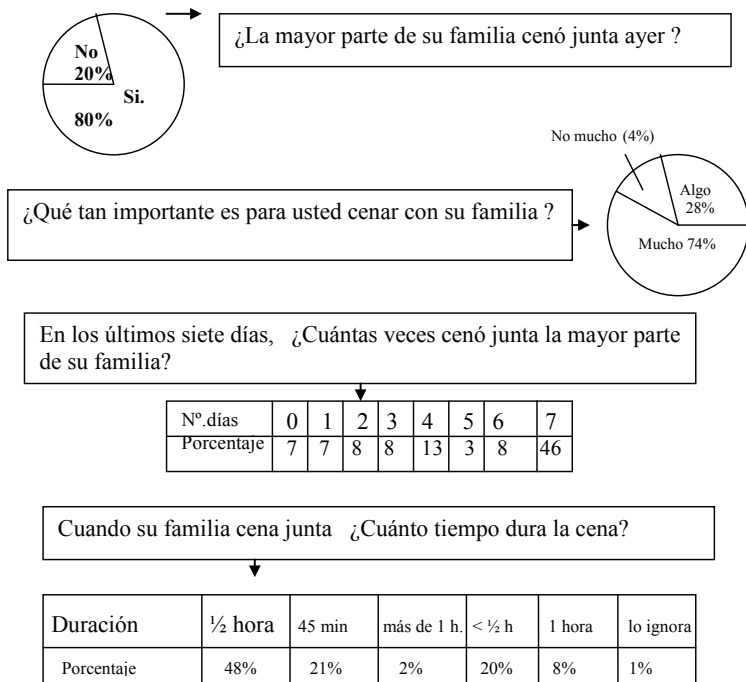
- Automóviles que pertenecen a los profesores del Departamento Idiomas: \_\_\_\_\_
- Colección de todos los automóviles que pertenecen a todos los miembros del cuerpo docente: \_\_\_\_\_
- Valor en pesos de cada automóvil individual: \_\_\_\_\_
- Los valores de los autos de cada uno de los docentes de idiomas: \_\_\_\_\_
- Le pregunto a cada docente de idiomas el valor de su auto: \_\_\_\_\_
- El valor promedio de todos los automóviles de los docentes del Instituto es de \$1.700.830: \_\_\_\_\_

- El automóvil del profesora Sánchez está valuado en \$14.594.000:

## SITUACIÓN 2

Comer juntos: sigue siendo importante.

A continuación se presentan los resultados obtenidos en una encuesta realizada respecto a la importancia de que una familia coma junta.



Con la información suministrada, responda las siguientes preguntas:

- Mencione las cuatro variables.
- ¿Qué tipo de variable es cada una?

### SITUACIÓN 3

Un fabricante de medicamentos está interesado en la proporción de personas que padecen hipertensión (presión arterial elevada) cuya condición pueda ser controlada por un nuevo producto desarrollado por la empresa. Se condujo un estudio en el que participaron 5000 personas que padecen de hipertensión, y se encontró que 80% de las personas pueden controlar su hipertensión con el medicamento. Suponiendo que las cinco mil personas son representativas del grupo con hipertensión, conteste las siguientes preguntas:

- a. ¿Cuál es la población?
- b. ¿Cuál es la muestra?
- c. Identifique el parámetro de interés
- d. Identifique la estadística y proporcione su valor
- e. ¿Se conoce el valor del parámetro?

### SITUACIÓN 4

La fábrica de gaseosas La Sed proyecta lanzar al mercado un nuevo sabor. Se realiza un test de aceptación de dicho sabor a 20 niños, utilizando una escala de 10 puntos, para medir el grado de aceptación. Uno de los niños (Pedro) aceptó el nuevo sabor con 7 puntos. Los puntos obtenidos en los 19 niños restantes son los siguientes: 2,6,7,4,5,5,9,8,7,5,1,8,4,7,7,7,6,5,4.

La muestra estuvo compuesta por igual número de niños de ambos sexos, de 6 a 12 años, pertenecientes a una escuela de Medellín, los cuales, en su mayoría, dieron una aceptación de 7 puntos.

1. Con la información, identifique:
  - a. ¿Cuál es la población?: \_\_\_\_\_
  - b. ¿Cuál es la muestra?: \_\_\_\_\_
  - c. ¿Cuál es la variable?: \_\_\_\_\_
  - d. ¿Es la variable cualitativa o cuantitativa?: \_\_\_\_\_
  - e. ¿Cuál es el experimento?: \_\_\_\_\_
  - f. ¿Cuál es un dato?: \_\_\_\_\_

- g. ¿Cuáles son los datos?: \_\_\_\_\_
- h. ¿Cuál es el estadístico?: \_\_\_\_\_
- i. ¿Cuál es el parámetro?: \_\_\_\_\_

2. Ejercicio

Identifique las siguientes expresiones como ejemplos de variables cualitativas o cuantitativas.

- a. Una encuesta de electores registrados según el candidato que apoyen.
  - b. El tipo de cuentas que se pueden registrar en una contabilidad.
  - c. Códigos utilizados para marcar productos en un supermercado
  - d. El tiempo mínimo necesario para una persona ejecute una tarea específica.
  - e. El número de páginas por trabajo que salen de la impresora de una computadora.
  - f. Registro por estratos de los habitantes de Medellín.
3. Resolver las siguientes preguntas:
- a. Describa en sus propias palabras cómo puede utilizarse la estadística para solucionar en varias disciplinas y ocupaciones.
  - b. Describa en sus propios términos la diferencia entre población y muestra, entre parámetro y un estadístico.
  - c. ¿Cuál es la diferencia entre una variable cuantitativa y una variable cualitativa? Dé ejemplos.
  - d. Diferencie entre una variable continua y una variable discreta. Dé ejemplo de cada una.
  - e. Una revista reciente reveló que los japoneses pronto controlarán hasta un 35% de las ventas de autos en los Estados Unidos, comparando con el 28% de finales de los años 80 está apenas un 8% por encima de lo ocurrido en 1970. ¿Esta información contiene estadística descriptiva, inferencias, o ambas? Explique.

- f. ¿Cuál es la diferencia entre la estadística descriptiva y la estadística inferencias? ¿Cuál cree usted constituye una forma más elevada de análisis estadístico y por qué?
  - g. Seleccione una población cualquiera que sea de su interés. Identifique variables cuantitativas y cualitativas de esa población que puedan seleccionarse para ser estudiadas.
  - h. Si los estadísticos están interesados realmente en poblaciones, ¿por qué realmente trabajan con muestras?
4. Determine si las siguientes variables son discretas o continuas:
- a. Número de cursos que los estudiantes de Costos y Presupuestos cursan en este semestre.
  - b. Número de pases atrapados por el beisbolista Tim Brown, receptor de los La Raiders.
  - c. Peso de los compañeros de equipo de Tim Brown.
  - d. Peso del contenido de las cajas de cereal.
  - e. Número de libros que usted leyó el año pasado.
5. El presidente de una fraternidad en el campus desea tomar una muestra de las opiniones de 112 miembros respecto a las actividades urgentes para el otoño:
- a. ¿Cuál es la población?
  - c. ¿Cuál es la muestra?

## 2.4 MÉTODOS GRÁFICOS PARA DESCRIBIR INFORMACIÓN

Un análisis estadístico comienza generalmente, con un estudio gráfico de los datos disponibles en una tabla.

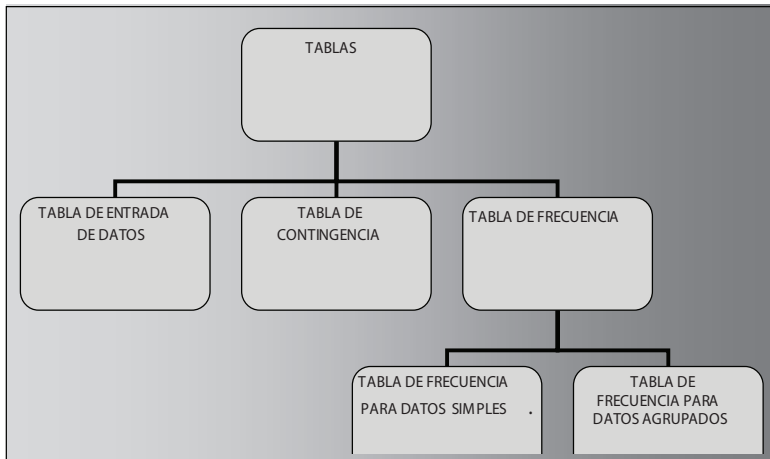
Cuando se tiene información considerable, acerca de una variable, es difícil hacer una interpretación rápida del comportamiento de ésta. Organizando estos datos en una tabla y a partir de elaborar algunas gráficas, se puede observar en forma clara e inmediata la variable.

## 2.5 TABLAS DE DATOS

**TABLA:** Es un cuadro que consiste en la disposición ordenada de los datos.

Las tablas sistematizan los resultados y ofrecen una visión numérica, sintética y global del fenómeno observado.

TIPO DE TABLAS



### 2.5.1 TABLA DE ENTRADA DE DATOS

Es aquella en la cual sólo aparecen los datos que se obtuvieron en la recolección de datos. Es la tabla más sencilla y se utiliza para registro de datos.

### 2.5.2 TABLA DE CONTINGENCIA

Llamada también tabla de doble entrada. Es aquella tabla que contiene los datos de dos variables y está formada de la siguiente manera: en la cabeza de las filas por los valores de una variable y las de las columnas por la otra, y en las casillas de la mitad, donde se cruzan las dos variables, van las frecuencias o número de elementos que reúne a la vez los valores de las dos variables. Para,  $i=1,2,3,\dots,m$  y  $j=1,2,3,\dots,n$

CONSTRUCCIÓN DE LA TABLA DE CONTINGENCIA

Variable $Y_j$ Variable $X_i$	$Y_1$	$Y_2$	...	$Y_n$	TOTAL
$X_1$	$X_1 Y_1$	$X_1 Y_2$	...	$X_1 Y_n$	$\sum_{j=1}^n X_1 Y_j$
$X_2$	$X_2 Y_1$	$X_2 Y_2$	...	$X_2 Y_n$	$\sum_{j=1}^n X_2 Y_j$
⋮	⋮	⋮	⋮	⋮	⋮
$X_m$	$X_m Y_1$	$X_m Y_2$	...	$X_m Y_n$	$\sum_{j=1}^n X_m Y_j$
TOTAL	$\sum_{i=1}^m X_i Y_1$	$\sum_{i=1}^m X_i Y_2$	...	$\sum_{i=1}^m X_i Y_n$	<b>n</b> tamaño muestra)

2.5.3 TABLA DE FRECUENCIAS

La tabla de frecuencias está formada por los valores de una variable cuantitativa y sus frecuencias correspondientes.

Definiremos qué es frecuencia:

FRECUENCIA: La frecuencia se refiere al número de repeticiones de una clase o de un valor de la variable.

2.5.3.1 TABLA DE FRECUENCIA SIMPLE

Está formada por los valores de una variable cuantitativa y sus frecuencias correspondiente. Generalmente se usa para un número de observaciones menores de 30 datos.



## PROCESO PARA LA CONSTRUCCIÓN DE LA TABLA DE FRECUENCIA SIMPLE

Utilizaremos el siguiente proceso para la elaboración de la tabla:

- Primera columna: Se organizan los datos de la variable en formas ascendente, sin repetir y a esta columna la llamamos  $x_i$ .
- Segunda columna: Se cuenta el número de veces que se repite el valor, y se coloca al frente de cada  $x_i$ . A esta frecuencia la llamamos frecuencia absoluta y la representamos por  $f_i$  o  $n_i$ .

$$\text{Entonces, } \sum_{i=1}^n f_i = n \quad (\text{n = tamaño de la muestra})$$

- Tercera columna: Se denomina Frecuencia Absoluta Acumulada. Está representada por  $F_i$  o  $N_i$ . Se calcula sumando la frecuencia absoluta  $f_i$  o  $n_i$  en cada valor, así:

$$\begin{aligned} F_1 &= f_1 \\ F_2 &= f_1 + f_2 \\ &\vdots \\ F_i &= \sum_{i=1}^n f_i \end{aligned}$$

- Cuarta columna: Frecuencia relativa, se representa por  $h_i$ . Se calcula dividiendo cada frecuencia absoluta  $f_i$  por el número total de observaciones  $n$ , así:

$$h_i = \frac{f_i}{n}, \text{ Entonces, } \sum_{i=1}^n h_i = 1$$

- Quinta columna: Frecuencia relativa acumulada, representada por  $H_i$ , correspondiente a cada valor de  $x_i$ . Se calcula sumando la frecuencia relativa en cada valor, así:

$$\begin{aligned} H_1 &= h_1 \\ H_2 &= h_1 + h_2 \\ H_3 &= h_1 + h_2 + h_3 \text{ y así sucesivamente} \\ &\vdots \end{aligned}$$

$$H_i = \sum_{i=1}^n h_i$$

La tabla para datos simples quedaría:

TABLA 1: TABLA DE DISTRIBUCIÓN DE FRECUENCIAS DATOS SIN AGRUPAR

$X_i$	$f_i$	$F_i$	$h_i$	$H_i$	$h_i^* \%$
$X_1$	$f_1$	$F_1$	$h_1 = f_1/n$	$h_1$	$h_1(100)$
$X_2$	$f_2$	$F_2$	$h_2 = f_2/n$	$h_1 + h_2$	$h_2(100)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$X_n$	$f_n$	$F_n$	$h_n = f_n/n$	$h_1 + h_2 + \dots + h_n$	$h_n(100)$

### 2.5.3.2 TABLA DE FRECUENCIA PARA DATOS AGRUPADOS

Se utiliza para las variables continuas y algunas discretas que presentan una gama de valores distintos considerables (con número de observaciones mayores de 30), se agrupan por intervalos o grupos.

#### PROCESO PARA LA CONSTRUCCIÓN DE LA TABLA DE FRECUENCIA PARA DATOS AGRUPADOS

1. Calculamos el número de intervalos de clase representado por **m**.  
**m**: Es el número de subgrupos en el cual se subdivide la variable.  
 Se puede calcular:

$$m = 1 + 3.3 \log n \quad (1)$$

$$m = \sqrt{n}$$

2. Cálculo el rango R: Que es la diferencia entre el valor mayor y el menor valor que toma la variable.

$$R = X_{\max} - X_{\min} \quad (2)$$

$X_{\max}$  : Valor máximo que toma la variable.

$X_{\min}$  : Valor mínimo que toma la variable.

3. Cálculo de la amplitud de los intervalos, se denota por la letra: **a**

$$a = \frac{R}{m} \quad (3)$$

4. Intervalos de clase: Esta es la primera columna de la tabla, comienza con el valor mínimo que toma la variable como el límite inferior de la primera clase, se le suma la amplitud como el cálculo del límite superior de la primera clase y este a su vez será el límite inferior ( $LI_1$ ), de la segunda clase; luego se le vuelve a sumar la amplitud para el límite superior ( $LS_1$ ), de la segunda clase, y así sucesivamente hasta llegar al último intervalo.

( $LI_1$ ): Límite inferior del intervalo o clase

( $LS_1$ ): Límite superior del intervalo o clase

i	Intervalos de clase	
1	$LI_1 = X_{\min}$	$LS_1 = LI_1 + a$
2	$LI_2 = LS_1$	$LS_2 = LI_2 + a$
⋮	⋮	⋮
n	$LI_n = LS_{n-1}$	$LS_n = LI_n + a$

$i = 1, 2, 3, \dots, n$

$$(LS_i) = (LI_i) + a \quad (4)$$

5. Se busca la marca de clase que denotamos por la nueva  $X_i$ , y es:

$$X_i = \frac{Ll_i + LS_i}{2} \quad (5)$$

6. La segunda columna de la tabla, para datos agrupados, es la frecuencia absoluta ( $f_i$ ): Es el número de observaciones que caen en el intervalo sin incluir el límite superior, es decir, número de datos mayores o iguales a ( $Ll_i$ ), pero menores que ( $LS_i$ ), o sea que el intervalo es cerrado hacia la izquierda y abierto a la derecha.  
[ )Intervalo.

Es importante anotar que el último intervalo de clase es cerrado [ ] a ambos lados, para no dejar información fuera del rango.

7. Se buscan las otras columnas, con las frecuencias  $F_i$ ,  $h_i$ ,  $H_i$ ,  $h_i^*$ , siguiendo el mismo procedimiento del de la tabla de frecuencia simple.

TABLA 2: TABLA DE DISTRIBUCIÓN DE FRECUENCIAS DATOS AGRUPADOS

Intervalos $Ll_i - LS_i$	$f_i$	$X_i$	$F_i$	$h_i$	$H_i$	$h_i^*\%$
$Ll_1 - LS_1$	$f_1$	$X_1$	$F_1$	$h_1 = \frac{f_1}{n}$	$h_1$	$h_1 (100)$
$Ll_2 - LS_2$	$f_2$	$X_2$	$F_2$	$h_2 = \frac{f_2}{n}$	$h_2 + h_2$	$h_2 (100)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
$Ll_n - LS_n$	$f_n$	$X_n$	$F_n$	$h_n = \frac{f_n}{n}$	$h_1 + h_2 + \dots + h_n$	$h_n (100)$

## 2.6 MÉTODOS GRÁFICOS

La presentación de la información mediante gráficos es algo que se analiza a diario y en forma casi natural por personas de diferentes profesiones. Los gráficos estadísticos nos permiten usar nuestra habilidad para visualmente procesar información de un gráfico. Luego, un gráfico es una de las mejores formas de conocer el material disponible pues facilita una comprensión global del problema en estudio. Los gráficos más usuales son:

### 2.6.1 GRÁFICO DE PASTEL O CIRCULAR

Llamado también el gráfico de sectores: Es un gráfico en forma de círculo, donde las categorías se basan en una proporcionalidad entre la frecuencia y el ángulo central de una circunferencia, de tal manera que a la frecuencia total le corresponde el ángulo central de  $360^\circ$ . Para construir se aplica la siguiente fórmula:

$$X = \text{frecuencia relativa} * 360^\circ / \text{frecuencia total}$$

También lo podemos representar en forma porcentual, se divide tantas partes como categorías tenga, de manera que el total del círculo sea el 100% y las cantidades se expresan en porcentaje.

Si la variable es cualitativa (rubio, moreno, alto bajo, etc.) se suele utilizar más este tipo de gráfico.

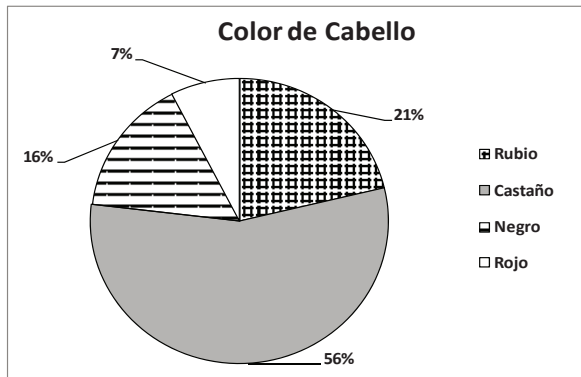
Ejemplo:

Supongamos que se tiene el color del cabello de un grupo de 160 estudiantes de la institución, el gráfico de pastel o circular, quedaría:

TABLA 1: COLOR DE CABELLO DEL UN GRUPO DE ESTUDIANTES DE LA INSTITUCIÓN

Categorías	No. de personas	% porcentaje	Angulo
Rubio	34	21	$77^\circ$
Castaño	89	56	$200^\circ$
Negro	25	16	$56^\circ$
Rojo	12	7	$27^\circ$
	Total = 160	100	$360^\circ$

## GRÁFICO DE SECTORES

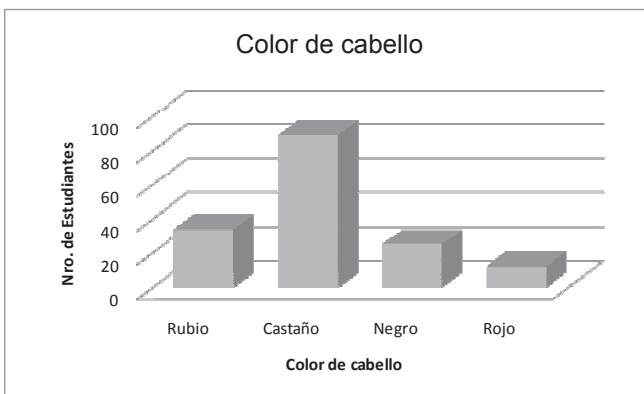


### 2.6.2 DIAGRAMA DE BARRAS

Se utiliza para graficar las frecuencias absolutas o relativas de una VARIABLE CUALITATIVA. En el eje de abscisas, situaremos las diferentes categorías de la variable. En el eje de ordenadas la frecuencia. Levantaremos barras o columnas SEPARADAS de altura correspondiente a la frecuencia adecuada.

#### Ejemplo

Utilizando el ejercicio anterior, obtenemos el siguiente diagrama de barras:



## BIBLIOGRAFÍA

- WALPOLE, Ronald, MYERS Raymon: *Probabilidad y estadística*, 4ª ed; México: Mc Graw Hill 1992, 797 p.
- CANAVOS, George *Probabilidad y estadística: aplicaciones y métodos*. México: McGraw-Hill, 1988, 651 p.
- BERENSON, Marck L, LEVINE David: *Estadística básica en administración conceptos y aplicaciones*, 6ª ed. México: Prentice-Hall, 1996, 943 p.
- BERENSON, Marck L, LEVINE David: *Estadística para administración y economía, conceptos y aplicaciones*. Bogotá: McGraw-Hill, 1993, 720 p.
- SOTO MAYOR Gabriel Velazco *Probabilidad y estadística para ingeniería y ciencias*. México: Thomson Learning, 2001, 326 p.
- MENDENHALL William: *Probabilidad y estadística para ingeniería y ciencias*, 4ª ed. México: Prentice Hall, 1997, 1182 p. + disquete.
- MONTGOMERY, Douglas *Probabilidad y estadística aplicadas a la ingeniería*: México: McGraw-Hill, 1994, 895 p.
- MILLER, Irwing: *Probabilidad y estadística para ingenieros*, 4ª ed, México: Prentice-Hall, 1992, 624.



*Estadística Básica*

se terminó de imprimir en febrero de 2010.

Para su elaboración se utilizó papel Bond de 75 g, en páginas interiores, y cartulina Propalcote 250 g para la carátula. Las fuentes tipográficas empleadas son Novarese Bk Bt 11 puntos, en texto corrido, y Myriad Pro 14 puntos en títulos.