

TEMA: ESTADÍSTICA

OBJETIVOS:

Interpretar y aplicar los conceptos y procedimientos básicos de la Estadística.

Estadística

En los primeros tiempos la estadística se limitaba simplemente al relevamiento de datos para luego ir incorporando paulatinamente nuevas técnicas para el análisis de la información recopilada.

Los complejos problemas de la sociedad actual y el incesante avance de la tecnología, generó la necesidad de una mayor formación estadística en nuestros alumnos, por ello esta materia intenta despertar en ustedes el interés que demanda la búsqueda de respuestas a diversas situaciones problemáticas y al desafío que representa poseer ciertos conocimientos para interpretarlos y comprenderlos.

La estadística se aplica a la física, agricultura, biología, medicina, economía, turismo, como estudio de mercado y control de calidad, censos de población, sondeos de opinión pública, etc.

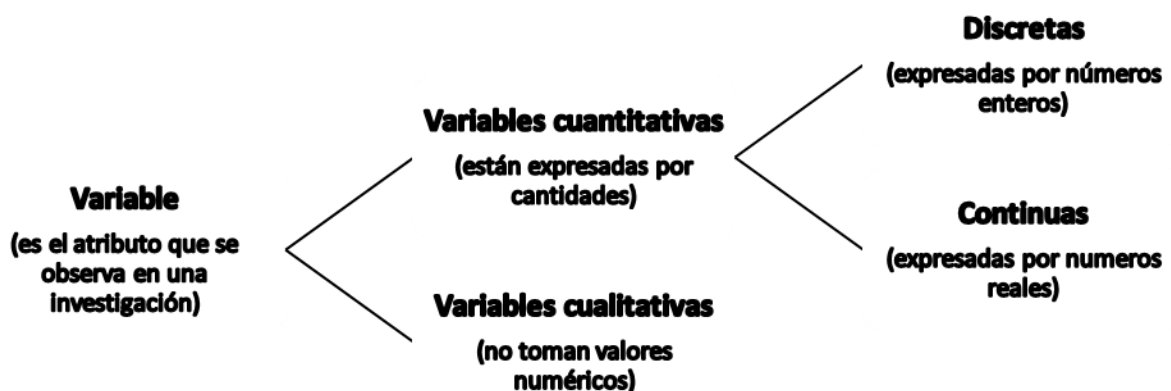
¿Qué es la Estadística?

La Estadística es la parte de las matemáticas que se ocupa de recoger, organizar y analizar grandes cantidades de datos para estudiar las características o el comportamiento de un colectivo.

Elementos de Estadísticas.

- **Población:** es el conjunto de individuos, que pueden ser personas o no, sobre los que se quiere obtener cierta información. Tamaño de la población se simboliza con (N) .
- **Individuo:** cada uno de los elementos de la población.
- **Muestra:** parte representativa de la población. Tamaño de la muestra la simbolizamos con (n) .
- **Variable estadística:** propiedad o característica de la población que estamos interesados en estudiar. (x_i)

Las variables se clasifican según el tipo de valor que pueden tomar.



- **Frecuencia absoluta:** es el número de veces que la variable toma un determinado valor se simboliza f_i .

La suma de las frecuencias absolutas es igual al número total de datos, n (es decir es el tamaño de la muestra):

$$f_1 + f_2 + f_3 + \dots + f_m = n$$

Para indicar resumidamente estas sumas se utiliza la letra griega Σ (sigma mayúscula) que se lee suma o sumatoria:

$$\sum_{i=1}^{i=m} f_i = n$$

- **Frecuencia relativa:** es la fracción del total de la muestra que le corresponde a ese dato, o sea, el cociente entre la frecuencia absoluta y el total de datos. Se simboliza f_r .

$$f_r = \frac{f_i}{n}$$

- **Frecuencia acumulada:** es la suma de las frecuencias absolutas de todos los valores inferiores o iguales al valor considerado. Se representa por F_i .

Ejemplo de lo visto hasta el momento.

Población: estudiantes de la orientación turismo.

Individuo: cada uno de los alumnos.

Variable cualitativa: color de ojos de los alumnos.

Variable cuantitativa discreta: edad del estudiante.

Variable cuantitativa continúa: la altura de los alumnos.

Actividades:

- 1) Indica, cual es la población, la variable y el tipo de variable.
 - a) Peso al nacer de los bebés que nacieron en Capital el año pasado.
 - b) Profesiones que quieren tener los estudiantes del último año escolar de la provincia de San Juan.
 - c) Número de mascotas hay en los hogares de La Toma.
 - d) Tiempo semanal que dedican a la lectura los estudiantes de Secundaria en el Colegio.
 - e) Número de tarjetas amarillas mostradas en los partidos de fútbol de la temporada pasada.
- 2) Indica que variables son cualitativas y cuales cuantitativas:
 - a) Comida Favorita.
 - b) Profesión que te gusta.
 - c) Número de goles marcados por tu equipo favorito en la última temporada.
 - d) Número de alumnos de tu colegio.
 - e) El color de los ojos de tus compañeros de clase.
 - f) Coeficiente intelectual de tus compañeros de clase.
- 3) De las siguientes variables indica cuáles son discretas y cuales continuas.
 - a) Número de acciones vendidas cada día en la Bolsa.
 - b) Temperaturas registradas cada hora en un observatorio.
 - c) Período de duración de un automóvil.
 - d) El diámetro de las ruedas de varios coches.
 - e) Número de hijos de 50 familias.
 - f) Censo anual de los españoles.

- 4) Determina en las siguientes variables cuales son cualitativa y cuantitativas (discretas o continuas).
- Gasto promedio por turista.
 - Las motivaciones de elección de un destino determinado.
 - Promedio de la temperatura durante la temporada de verano.
 - Nacionalidad de los turistas.
 - Hoteles de acuerdo al número de estrellas.
 - El impacto económico sobre el destino (pesos).
 - El grado de satisfacción del consumidor: Muy bueno-Bueno-Malo.
 - Número de hijos de los turistas que viajan en familia.

1.1. La recolección de datos y su interpretación

El instrumento más utilizado para recolectar información sobre cierto tipo de estudio es la **encuesta**; luego de realizada, se organiza en tablas, con la finalidad de realizar el tratamiento de la misma. Entre algunos conceptos integrantes, que figuran en las tablas tenemos:

Frecuencia absoluta: Es el número de veces en que se repite o aparece un valor, el total del número de veces, deberá coincidir con el tamaño de la muestra.

Frecuencia relativa: Es el cociente entre los valores de frecuencia absoluta y el tamaño de la muestra. Toma valores entre 0 y la unidad, debido a que son fracciones.

El valor total de todas las frecuencias relativas es 1.

De la frecuencia relativa se deriva su interpretación porcentual, donde se halla el producto entre la frecuencia relativa por 100%.

1.2. Tabla de frecuencia para datos no agrupados

Los **datos no agrupados** son valores obtenidos y recolectados a través de una encuesta, en una cantidad pequeña relativamente ($n < 30$), mismos que son analizados sin tipo de preclasificación.

Ejemplo 1

Se realiza un estudio para determinar la edad de veinte estudiantes del décimo año de EGB del Colegio «ABC», y se obtuvo los siguientes datos:

14, 15, 14, 14, 13, 16, 16, 15, 15, 15, 15, 15, 14, 14, 14, 14, 15, 16, 16, y 14.

Determinemos la tabla de frecuencias considerando la variable mencionada.

x_i	f_i	f_r	F
13	1	0,05	1
14	8	0,40	9
15	7	0,35	16
16	4	0,20	20
Total	20	1	

Actividades

- 1) Se realiza un estudio para determinar el número de horas que utilizan Internet; le preguntan a José; Juan; Mariana, Andrea, Luis; Alejandro; Silvia; Paulina; Marcos y Carlos, estudiantes del 5 año del secundario. ¿Cuántas horas al día usted utiliza el servicio de Internet? Identificar la población, muestra, variable y tipo de variable.
- 2) Los siguientes valores pertenecen a una encuesta realizada para conocer el número de veces que un estudiante de quinto año, revisa su celular en un día.

23 24 26 27 24 23 26 24 26 27

23 25 28 26 26 24 27 23 23 27

28 27 26 27 27 23 23 25 26 25

25 26 25 26 25 27 27 24 23 27

Determinar:

- a) Identificar la variable.
 - b) Determinar si es cualitativa o cuantitativa.
 - c) Identificar si es discreta o continua.
 - d) Realizar la tabla de frecuencias.
- 3) Se realizó un estudio acerca de los montos de ventas diarias, en pesos, de 25 comercios elegidos al azar. Los datos recopilados son los siguientes:

180 321 142 180 100

230 420 190 180 130

240 430 230 230 135

275 513 250 280 230

350 520 251 600 303

Determinar:

- a) Identificar la variable.
- b) Determinar si es cualitativa o cuantitativa.
- c) Identificar si es discreta o continua.
- d) Realizar la tabla de frecuencias.

Tabla de frecuencias para datos agrupados

Si la variable tiene muchos datos de distinto valor, antes de efectuar el recuento debemos agrupar los datos en intervalos.

Rango de datos: se determina mediante la diferencia entre el valor mayor de la variable ($x_{\text{máx}}$) y el valor menor ($x_{\text{mín}}$).

$$R = x_{\text{máx}} - x_{\text{mín}}$$

Los números extremos de cada intervalo se denominan “límite superior” y “límite inferior”

Tamaño o anchura de un intervalo: es la diferencia entre los límites que lo forman.

$$a = \frac{\text{límite superior} - \text{límite inferior}}{I}$$

Marca de clase: es el punto medio del intervalo

$$x_i = \frac{\text{límite superior} + \text{límite inferior}}{2}$$

Número de intervalos: cantidad de intervalos a formar

$$I = 1 + 3,3 \cdot \log n \quad \text{Fórmula de Sturges}$$

REGLAS GENERALES PARA FORMAR DISTRIBUCIONES DE FRECUENCIAS

- 1) Determinar el rango.
- 2) Todos los intervalos de clases deberán tener la misma amplitud.
- 3) Aplicar la fórmula de Sturges para hallar la cantidad de intervalos.
- 4) Determinar el ancho, a , dividiendo el rango en el número de intervalos.
- 5) Obtener la marca de clase.
- 6) Determinar el número de observaciones que caen dentro de cada intervalo de clase, es decir, hallar las frecuencias absolutas.
- 7) Obtener las frecuencias acumuladas de cada intervalo de clase.

Ejemplo: Se le pregunto a 90 estudiantes elegidos en forma aleatoria, sobre el gasto realizado el día anterior (transporte, fotocopias, artículos de librería). Las respuestas fueron las siguientes:

8 8 8 12 12 12 20 20 19 25 25 27 44 33 17 17 7 6 18 18 18
24 25 15 15 15 16 16 16 10 10 9 9 19 19 8 20 22 23 23 40 32
30 29 28 17 17 13 13 13 16 16 15 15 15 23 23 21 14 14 21 10 10
12 11 12 12 21 21 21 17 28 28 17 17 18 18 18 14 8 15 15 18 20
27 25 30 40 11 14

1) La tabulación comienza ordenando los datos en forma creciente, disponiéndolos en un diagrama de pilas.

Aclaración: Diagrama de pilas

Lo primero que se debe hacer es ordenar los datos de menor a mayor. Luego, se pueden reagrupar formando pilas verticales de datos repetidos dispuestos a la par y siguiendo un orden creciente de izquierda a derecha.

```

                                     15
                                     15  17 18
                                     15  17 18
                                     12
      8                               12  15 16 17 18
      8   10                         12  14 15 16 17 18
      8   10   12 13 14 15 16 17 18 19
      8  9 10 11 12 13 14 15 16 17 18 19
6  7  8  9 10 11 12 13 14 15 16 17 18 19
```


intervalos	x_i	f_i	F	f_r
[6 ; 12)	9	15	15	$= \frac{15}{90}=0,17$
[12 ; 18)	15	33	15+33=48	$= \frac{33}{90}=0,37$
[18 ; 24)	21	24	48+24=72	$= \frac{24}{90}=0,27$
[24 ; 30)	27	11	72+11=83	0,12
[30 ; 36)	33	4	87	0,04
[36 ; 42)	39	2	89	0,02
[42 , 48]	45	1	90	0,01
Total		90		1

Gráficos estadísticos

Un gráfico estadístico permite hacer una rápida lectura de los datos recolectados de una muestra.

Si la variable es **cualitativa** los gráficos más convenientes son el circular o el de barras.

↳ Gráfico circular:

Se utiliza con el objeto de visualizar la importancia que tiene cada categoría de la variable, en base a la cantidad de datos que contiene del conjunto total.

Se utiliza una circunferencia de un radio cualquiera, luego los ángulos que corresponden a cada categoría de la variable están asignados en correspondencia con la frecuencia relativa de c/u de ellos.

Es decir: ángulo de la categoría i-ésima

$$\alpha_i^\circ = \frac{f_i}{n} \cdot 360^\circ = f_r \cdot 360^\circ$$

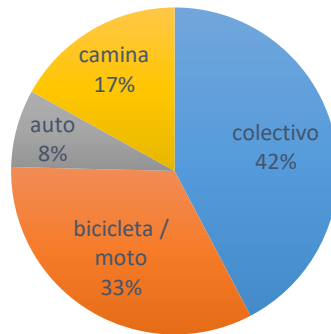
↳ Gráfico de barras

Cada clase se representa con una barra, se establece la altura de las barras teniendo en cuenta las frecuencias correspondientes.

Ejemplo para Variable cualitativa

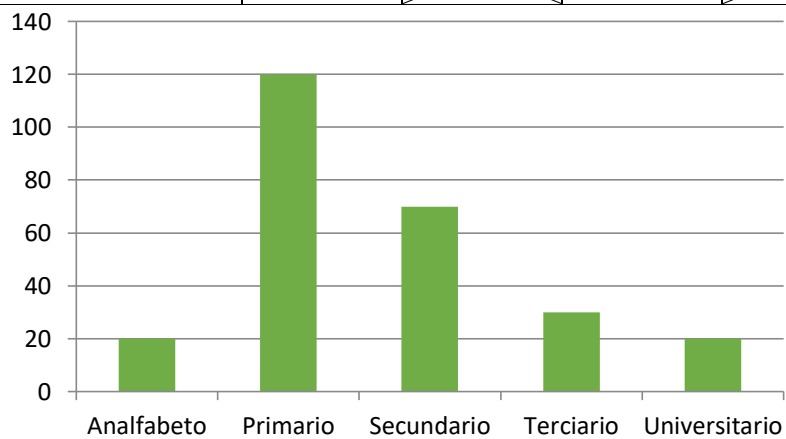
Se tomó una muestra de los alumnos del colegio ($n = 260$) y se les pregunto sobre el medio de transporte que utilizan para llegar al establecimiento. Las respuestas se presentan en la siguiente tabla:

Medio de transporte	f_i	f_r	α_i°	%
colectivo	110	0,42	$= 0,42 \cdot 360^\circ = 151^\circ$	$= 0,42 \cdot 100 \% = 42\%$
bicicleta / moto	86	0,33	$= 0,33 \cdot 360^\circ = 119^\circ$	$= 0,33 \cdot 100\% = 33\%$
auto	20	0,08	$= 0,08 \cdot 360^\circ = 29^\circ$	8%
camina	44	0,17	$= 0,17 \cdot 360^\circ = 61^\circ$	17%
Total	260	1	360°	100%



Nivel de instrucción alcanzado por el padre de los 260 alumnos entrevistados

Nivel de instrucción	f_i	F_i	f_r	F_r
Analfabeto	20	20	0,08	0,08
Primario	120	140	0,46	0,54
Secundario	70	210	0,27	0,81
Terciario	30	240	0,12	0,93
Universitario	20	260	0,08	1,01
Total	260		1,01	



Si la variable es **cuantitativa y dividida en intervalos de clases** los gráficos más convenientes son el histograma y el polígono de frecuencias.

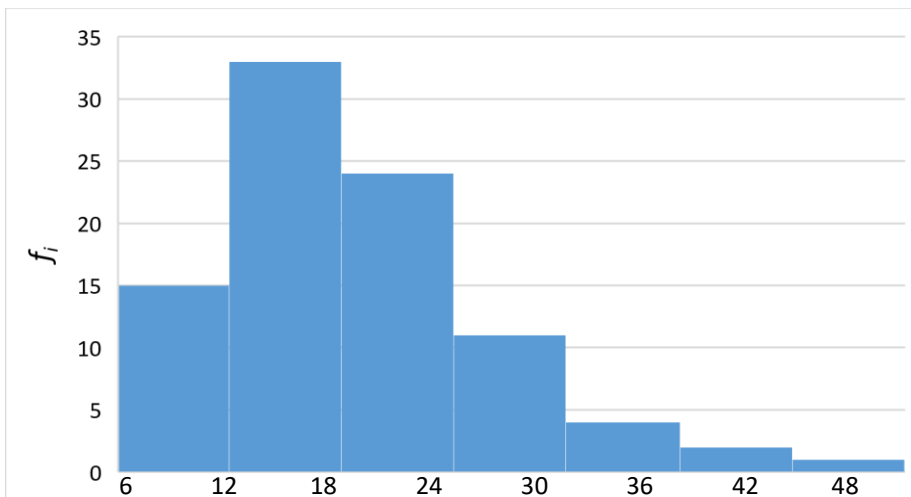
↳ Histograma:

En un sistema de coordenadas cartesianas se construyen rectángulos que tienen como base los intervalos de clases y como altura la frecuencia correspondiente a ese intervalo.

↳ Polígono de frecuencias

Es un gráfico de líneas trazado sobre las marcas de clase. Puede obtenerse uniendo los puntos medios de los techos de los rectángulos en el histograma.

Histograma de frecuencias absolutas



Polígono de frecuencias absolutas

