



La **estadística** es la parte de la matemática que se encarga de la recolección, organización y análisis de datos para cumplir con determinados objetivos.

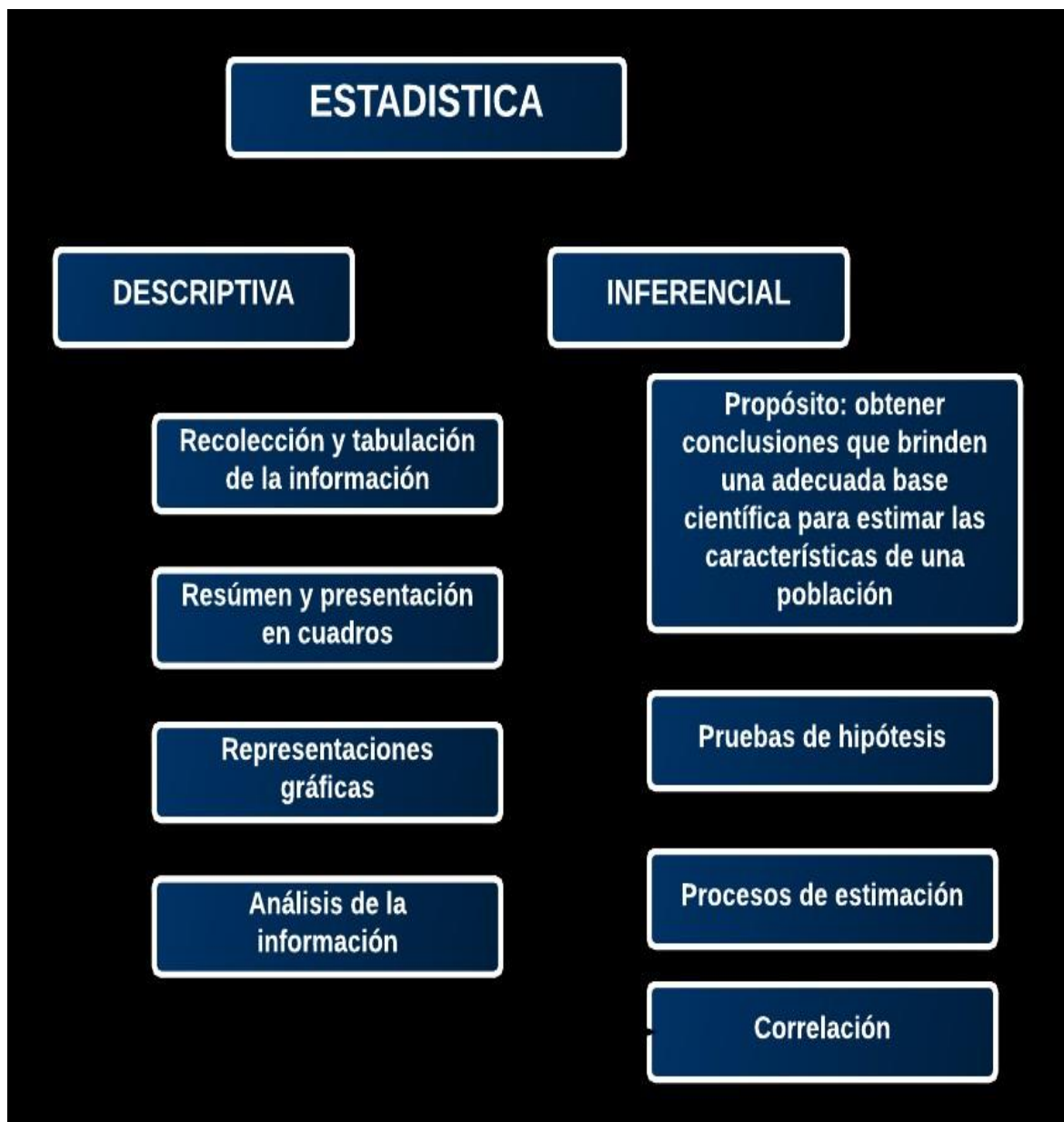
La etapa de recolección de datos se realiza a través de censos, encuestas e investigaciones. Gracias a los avances de la computación, los datos obtenidos pueden ser procesados rápidamente para recabar información.

En nuestro país periódicamente se llevan a cabo censos de población con el objetivo de averiguar la cantidad de habitantes y a partir de ello determinar, por ejemplo, las necesidades a nivel educativo y sanitario.

No hay un método estadístico más válido que otro, sino que la elección acerca de cuál usar dependerá de lo que se quiera estudiar o del tipo de aplicación que se esté investigando.

Si se piensa en un listado de población tipo censal en un tiempo concreto en el que figuren los datos personales de las personas que viven en cada domicilio de cada núcleo de población, lo que se está practicando aquí es **estadística descriptiva**.

Si posteriormente se toma una porción de esos mismos datos censales y se extrae con ellos ciertas conclusiones referidas a la totalidad de la población bajo análisis, haciendo uso de operaciones aritméticas, se trata de **estadística inferencial**.



Se llama **Población** al conjunto de individuos o elementos que se pretende estudiar estadísticamente mediante una encuesta, un censo o una investigación.

Cuando es difícil estudiar toda una población, se selecciona un subconjunto representativo llamado **Muestra**. A partir de los resultados obtenidos sobre una muestra se pueden conocer las características de la población. Cada uno de los temas sobre los que se estudia una población se denomina **variable estadística**.

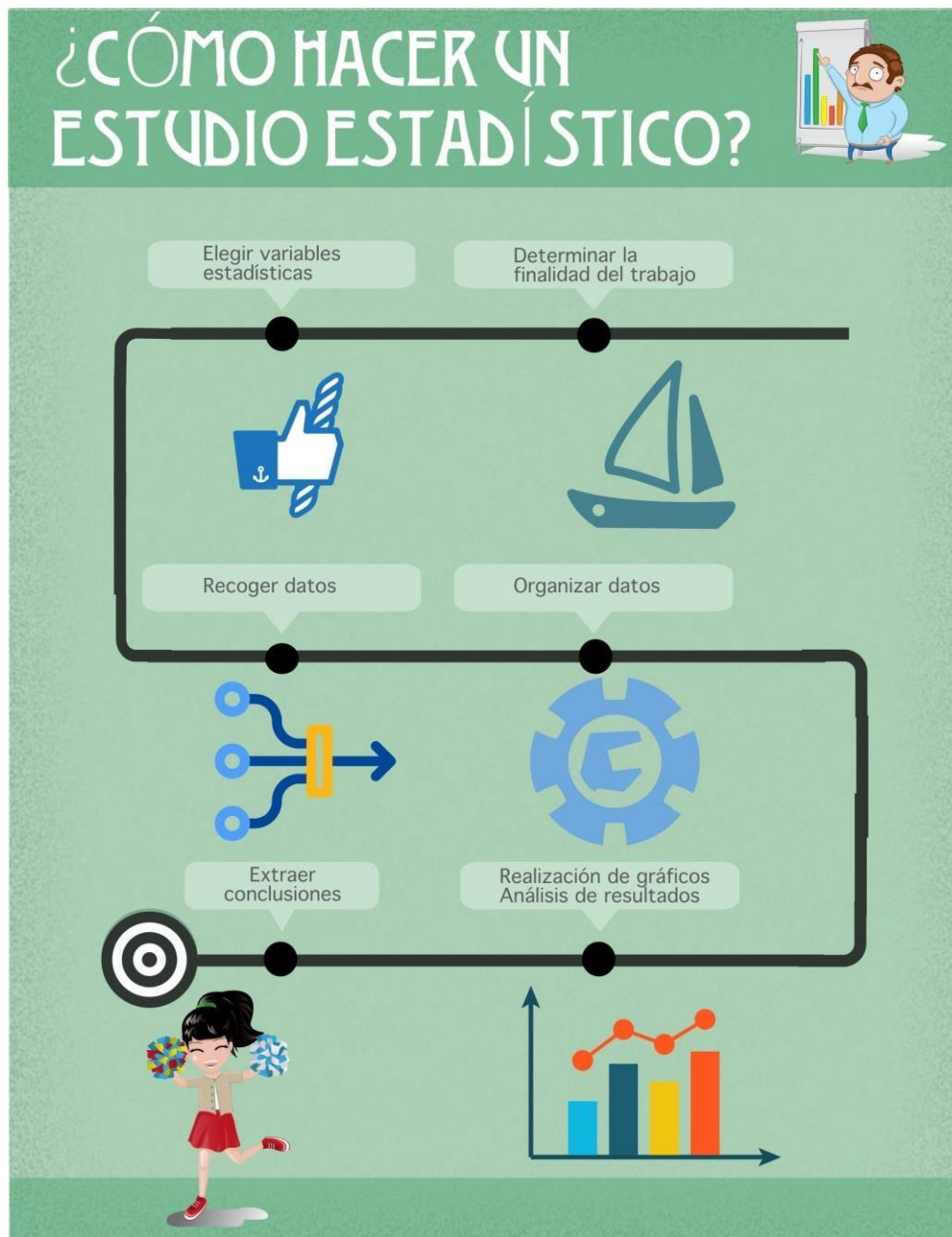


Tipos de variables estadísticas:

Variable cualitativa: Hace referencia a una cualidad. Se miden a partir de datos no numéricos. Ejemplos: el color de pelo de una persona; el desempeño en determinada actividad.

Variable cuantitativa: Hace referencia a una medida cuantitativa. Se miden a partir de datos numéricos. Ejemplos: la altura de una persona en centímetros o el peso de una persona en kilogramos.

ETAPAS DEL PROCESO ESTADÍSTICO:



1- Definición del Objeto de la investigación. 2- Recolección y Organización de datos. 3- Resumen (Gráficos y medidas descriptivas). 4- Análisis e Interpretación de resultados. 5- Toma Racional de Decisiones.

- 1- Definición del Objeto de la investigación: Debe plantearse un objeto de estudio que oriente todo el trabajo estadístico. - Población, muestra, muestreo, variables.
- 2- Recolección y Organización de datos: Van tomando forma los datos, se van tornando operables. Se calculan distribuciones de frecuencias, se ordenan en sentido creciente o decreciente los valores, se los separa por algún criterio válido.
- 3- Resumen (Gráficos y medidas descriptivas): En la etapa de resumen, se achica o reduce el gran cúmulo de datos recolectados, que es el gran

desafío de la estadística. Se calculan porcentajes, promedios y se construyen gráficos.

- 4- **Análisis de resultados:** Con los datos resumidos se puede analizar sus patrones de comportamiento, concentraciones, etc. El análisis se basa en aspectos objetivos, comprobables.
- 5- **Interpretación de resultados:** La interpretación es más subjetiva, se incorporan puntos de vista y opiniones propias del investigador que intentan dar respuestas o justificaciones a determinadas situaciones.
- 6- **Toma Racional de Decisiones:** Todo proceso estadístico, para que sea útil, debe concluir en decisiones que busquen llevar respuestas a las problemáticas planteadas o detectadas al momento de definir los objetivos, con la finalidad de superarlos.

MUESTREO

Primero hay que distinguir la palabra muestra de muestreo:

- **Muestra:** es un subconjunto representativo de la población.
- **Muestreo:** es el método que se utiliza para elegir a los elementos o individuos de la población que va a pasar a formar parte de la muestra.

Métodos de Muestreo probabilístico:

Es el procedimiento mediante el cual se selecciona al azar una determinada cantidad de elementos de la población para que conformen la muestra, teniendo todos de ellos la misma probabilidad de ser seleccionados.

*Muestreo aleatorio simple:

Primero se asigna un número a cada individuo de la población y luego a través de algún medio mecánico se eligen tantos sujetos como sea necesario para completar el tamaño de la muestra requerido.

*Muestreo aleatorio sistemático:

Este procedimiento exige numerar todos los elementos de la población para luego extraer 1 número aleatorio (i). Se parte de ese número aleatorio y los elementos que integran la muestra son los que ocupa los lugares $i, i+k, i+2k, i+3k, \dots, i+(n-1)k$, es decir se toman los individuos de k en k , siendo k el resultado de dividir el tamaño de la población entre el tamaño de la muestra: $k = N/n$.

*Muestreo aleatorio estratificado:

Consiste en considerar categorías típicas diferentes entre sí (estratos) que poseen gran homogeneidad respecto a alguna característica.

Lo que se pretende con este tipo de muestreo es asegurarse de que todos los estratos de interés estarán representados adecuadamente en la muestra. Cada estrato funciona independientemente, pudiendo aplicarse dentro de ellos el muestreo aleatorio simple o el estratificado para elegir los elementos concretos que formarán parte de la muestra. Tiende a asegurar que la muestra represente adecuadamente a la población en función de unas variables seleccionadas.

***Muestreo aleatorio por conglomerados:**

En el muestreo por conglomerados la unidad muestral es un grupo de elementos de la población que forman una unidad, a la que llamamos conglomerado. Las unidades hospitalarias, los departamentos universitarios, una caja de determinado producto, etc., son conglomerados naturales. En otras ocasiones se pueden utilizar conglomerados no naturales como, por ejemplo, las urnas electorales. Cuando los conglomerados son áreas geográficas suele hablarse de “muestreo por áreas”.

El muestreo por conglomerados consiste en seleccionar aleatoriamente un cierto número de conglomerados (el necesario para alcanzar el tamaño muestral establecido) y en investigar después todos los elementos pertenecientes a los conglomerados elegidos.

En el muestreo estratificado, se toma una muestra aleatoria de cada uno de los estratos, mientras que en el muestreo por conglomerados sólo se toman muestras de los grupos seleccionados. Una motivación común del muestreo por conglomerados es reducir los costos aumentando la eficiencia del muestreo.

Métodos de Muestreo No probabilístico:

Es el procedimiento de selección en el cual el investigador elige deliberadamente los objetivos a ser estudiados.

***Muestreo por cuotas o accidental:**

Se asienta generalmente sobre la base de un buen conocimiento de los estratos de la población y/o de los individuos más “representativos” o “adecuados” para los fines de la investigación. Mantiene, por tanto, semejanzas con el muestreo aleatorio estratificado, pero no tiene el carácter de aleatoriedad de aquél.

En este tipo de muestreo se fijan unas “cuotas” que consisten en un número de individuos que reúnen unas determinadas condiciones. Una vez determinada la cuota se eligen los primeros que se encuentren que cumplan esas características. Este método se utiliza mucho en las encuestas de opinión.

***Muestreo intencional o de conveniencia:**

Este tipo de muestreo se caracteriza por un esfuerzo deliberado de obtener muestras “representativas” mediante la inclusión en la muestra de grupos supuestamente típicos. Es muy frecuente su utilización en sondeos preelectorales de zonas que en anteriores votaciones han marcado tendencias de voto.

También puede ser que el investigador seleccione directa e intencionadamente los individuos de la población. El caso más frecuente de este procedimiento es el utilizar como muestra los individuos a los que se tiene fácil acceso (los profesores de universidad emplean con mucha frecuencia a sus propios alumnos).

*Muestreo bola de nieve:

Se localiza a algunos individuos, los cuales conducen a otros, y estos a otros, y así hasta conseguir una muestra suficiente. Este tipo se emplea muy frecuentemente cuando se hace estudios con poblaciones “marginales”, delincuentes, sectas, determinados tipos de enfermos, etc.

*Muestreo discrecional:

A criterio del investigador los elementos son elegidos sobre lo que él cree que pueden aportar al estudio.

DISTRIBUCIÓN DE FRECUENCIAS:

Los datos que se obtienen de una encuesta o investigación pueden organizarse en tablas que faciliten el análisis de la información.

En la puerta de una escuela de idiomas se realizó una encuesta para averiguar qué lengua extranjera estudian las personas que concurren a las clases.

Los datos que se obtuvieron fueron los siguientes:

IDIOMA	Cant. De personas que lo hablan	Parte que representan del total (Fracción)	Parte que representan del total (Expres. Decimal)	Porcentaje
Inglés	60			
Francés	20			
Alemán	10			
Portugués	35			
Italiano	13			
TOTALES				

Completen las tres últimas columnas de la tabla.

Para averiguar la parte del total que representa un valor de la variable, se divide el número de personas por el total de encuestados. El cociente se denomina frecuencia relativa y en la tabla anterior se expresan de tres formas distintas: como fracción, como expresión decimal y como porcentaje.

¿Qué es la frecuencia?

Se llama frecuencia a la cantidad de veces que se repite un determinado valor de la variable.

Es interesante ver cómo se distribuyen las frecuencias. Para ello, existen ciertos conceptos que debemos conocer:

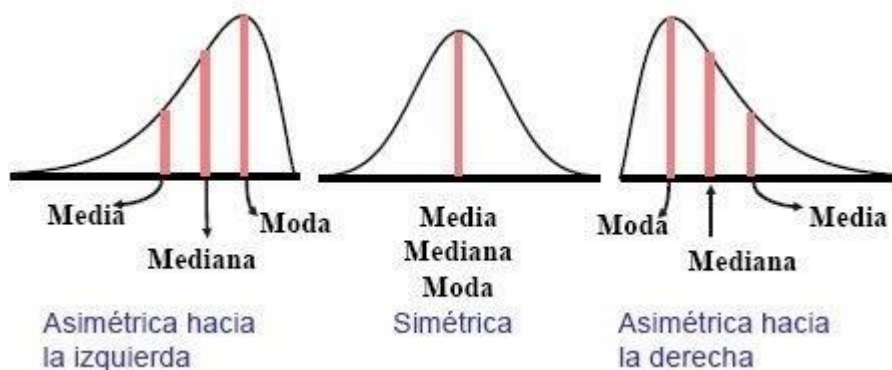
número de teléfonos móviles por hogar familiar	número de familias con esa cantidad	frecuencia absoluta acumulada	frecuencia relativa acumulada		porcentaje acumulado	
x_i	f_i	F_i	h_i	H_i	%	% acu.
2	12	12	0,30	0,30	30	30
3	16	28	0,40	0,70	40	70
4	8	36	0,20	0,90	20	90
5	3	39	0,08	0,98	7,5	97,5
6	1	40	0,03	1,00	2,5	100
N =		40				

- **Frecuencia absoluta** (f_i ó f_a): Es el número total de veces que se repite una observación. En ocasiones las observaciones se pueden presentar en intervalos.
- **Frecuencia relativa** (f_r ó h_i) : Es el número de la parte (respecto al total) que se repite una observación o un conjunto de ellas.
- **Frecuencia acumulada** (F_i ó F_a): Puede ser **relativa acumulada** o **absoluta acumulada**. Indica la cantidad acumulada hasta cierta observación.
- **Frecuencia porcentual** ($f_{\%}$ ó $\%$) : Es el número en porcentaje que se repite una observación o un conjunto de ellas.

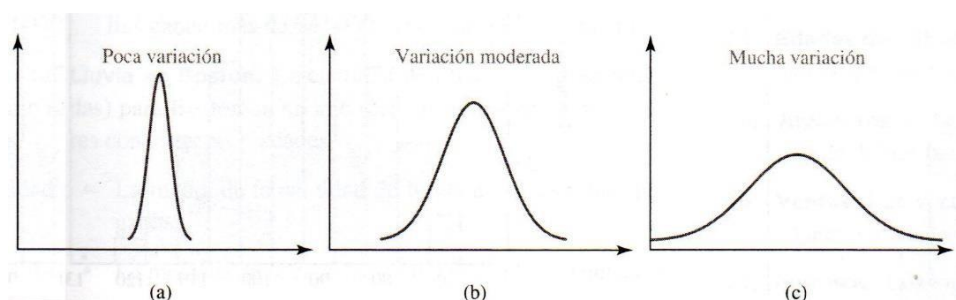
PARAMETROS ESTADISTICOS BASICOS:

Con el objetivo de resumir la información, se idearon diversas fórmulas que ofrecían medidas de un determinado tipo. Así, están aquellas que nos ofrecen información sobre el centro, otras sobre la dispersión o variabilidad y otras sobre la posición de un valor.

Medidas de tendencia central: Denominadas así porque ofrecen información sobre el centro de conjunto de datos. Por ejemplo, la media es una medida de tendencia o posición central ya que el promedio nos ofrece un valor centrado del conjunto de datos. ¿Dónde podríamos decir que se encuentra el punto medio? En el centro, en la mitad aproximadamente. Otro ejemplo de medida de tendencia central es la mediana.



Medidas de dispersión: También son conocidas como medidas de variabilidad. Por ejemplo, la desviación típica es una medida de variabilidad ya que nos dice si los valores de un conjunto de datos son muy dispares o no. Dos ejemplos más sobre medidas de dispersión podrían ser la varianza y el rango estadístico.



Medidas de posición: No son las más conocidas, pero se utilizan frecuentemente. Un ejemplo de ello, se encuentra en los percentiles o los deciles. Cuando un dato en concreto se encuentra en el percentil 90, quiere decir que por debajo de ese dato se encuentran el 90% de datos. Existen otras medidas de posición como los cuartiles o algunas variantes como el primer cuartil.

CUARTILES:



DECILES:



PERCENTILES:

