

TEMA 1. INTRODUCCIÓN Y ANÁLISIS GRÁFICO DE DATOS

¿QUÉ ES Y PARA QUÉ SIRVE LA ESTADÍSTICA?

1 La Estadística se ocupa de la recolección, agrupación, presentación, análisis e interpretación de datos.

La Estadística no son sólo resultados de encuestas, ni el cálculo de unos porcentajes, la estadística es un método científico que pretende sacar conclusiones a partir de unas observaciones hechas.

El estudio de la Estadística se divide habitualmente en tres partes:

- **Estadística Descriptiva** que centra la atención en los métodos gráficos y numéricos que se utilizan para resumir y procesar datos.
- **Probabilidad** que estudia las leyes del azar, las cuales permiten analizar las situaciones de incertidumbre utilizando probabilidades.
- **Inferencia Estadística** que centra la atención en la utilización de los datos para realizar predicciones, previsiones y estimaciones con el fin de tomar mejores decisiones.

ALGUNOS CONCEPTOS BÁSICOS.

- **Población:** conjunto de todos los elementos objeto de nuestro estudio. El tamaño de la población, N , puede ser muy grande o incluso infinito.
- **Muestra:** subconjunto, extraído de la población, (mediante técnicas de muestreo) cuyo estudio sirve para inferir características de toda la población. El *tamaño muestral* se denota por n .
- **Individuo:** cada uno de los elementos que forman la población o la muestra.
- **Parámetro:** medida numérica que describe una característica específica de una población.
- **Estadístico:** medida numérica que describe una característica específica de la muestra.

CLASIFICACIÓN DE VARIABLES.

Una variable estadística es una característica específica (edad, peso...), *objeto de estudio*, de un individuo o elemento de la población.

2

Una **variable** puede ser:

- **Categórica** (o **cualitativa**) si sus posibles valores son categorías o grupos.

Por ejemplo: las respuestas sí/no, el sexo, el estado civil...

A su vez las variables categóricas las dividiremos en:

- Variables **categóricas nominales**: aquellas en que no hay un orden natural en las categorías.

Por ejemplo, el estado civil: soltero, casado, viudo, separado.

- Variables **categóricas ordinales**: aquellas en que sí hay un orden natural en las categorías.

Por ejemplo, la nota de un examen: suspenso, aprobado, notable, sobresaliente.

- **Numérica** (o **cuantitativa**) si sus posibles valores son números.

Por ejemplo: la edad, el número de hijos...

Las variables numéricas las dividiremos en:

- Variables **numéricas discretas**: aquellas en el conjunto de posibles resultados es finito o infinito numerable.

Por ejemplo: número de estudiantes en una clase, número de personas que utilizan el transporte público...

- Variables **numéricas continuas**: aquellas en el conjunto de posibles resultados es infinito no numerable (*"pueden tener muchos decimales"*).

Por ejemplo: altura, tiempo de espera del autobús...

GRÁFICOS PARA DESCRIBIR VARIABLES CATEGÓRICAS.

Si tenemos un conjunto de observaciones de una variable categórica:

- Al número total de observaciones lo representaremos como n .
- Al número de categorías posibles lo representaremos como k .

3

Además:

- se llama **frecuencia** (o **frecuencia absoluta**) de una categoría al número de observaciones que hay en ella.

Las frecuencias las representaremos como n_1, n_2, \dots, n_k .

➤ Se cumple siempre que $n_1 + \dots + n_k = n$.

- se llama **frecuencia relativa** de una categoría a la proporción (o porcentaje) de todas las observaciones que hay en esa categoría.

Las frecuencias relativas las representaremos como f_1, f_2, \dots, f_k , y se obtienen dividiendo la frecuencia por el número total de observaciones, es decir:

$$f_1 = \frac{n_1}{n} \dots \dots \dots f_k = \frac{n_k}{n}$$

➤ Se cumple siempre que $f_1 + \dots + f_k = 1$.

El conjunto de observaciones de una variable categórica se describe utilizando una **tabla de frecuencias** y gráficos como un **diagrama de sectores**, **diagrama de barras** y **diagrama de Pareto**.