

TEMA 1. FUENTES DE INFORMACIÓN

Curso OCW de “Estadística Descriptiva con Excel para Grados de Ciencias Sociales”



1. INTRODUCCIÓN

La Estadística, como campo de estudio, es una ciencia que permite dar sentido a un conjunto de datos numéricos, extrayendo conclusiones sobre la realidad que nos rodea. Los datos son observaciones, medidas, sobre la cualidad o cantidad de algo o alguien en la naturaleza o en el mundo tales como por ejemplo la tasa de paro, los ingresos mensuales, las notas de los exámenes o los resultados electorales. A través de ellos llegamos a conocer un fenómeno o hecho que ayuda no solo a describirlo, sino también a extraer conclusiones y tomar decisiones.

Para ser útiles, los datos se deben recopilar y poner a nuestra disposición. Los datos nos llegan en diferentes formas o formatos y la mayoría son números o pueden ser traducidos a números para los análisis. La simple decisión de cómo medir y recopilar los datos relevantes para un estudio es el comienzo indispensable para utilizarlos en la solución de problemas.

2. FUENTES DE DATOS ESTADÍSTICOS

Las fuentes de información constituyen un recurso indispensable para adquirir conciencia de la realidad que se pretende estudiar o en la que se tienen que tomar decisiones. Los datos estadísticos necesarios para la comprensión de los hechos pueden obtenerse a través de diferentes tipos de fuentes.

Fuentes de datos primarias son las personas o instituciones que han recolectado directamente los datos y secundarias son las publicaciones y trabajos hechos por personas o entidades que no han recopilado directamente la información. Las más fiables son las oficinas gubernamentales o instituciones de reconocido prestigio que ofrecen plenas garantías sobre la información que proporcionan. A nivel práctico es aconsejable utilizar fuentes de datos primarias siempre que sea posible.

A pesar de que un sinnúmero de entidades recoge una información cada vez más completa y desagregada, siempre habrá ámbitos específicos en los que la información no se registre. De ahí que con cierta frecuencia el investigador deba obtener por sí mismo la información necesaria para un estudio concreto mediante la observación, experimentación o realización de encuestas.

En cualquiera de los casos es importante garantizar la calidad de la información, ya que las conclusiones estadísticas dependen en grado sumo de la exactitud de los datos que se utilicen.

Algunas Fuentes primarias de información importantes:

- Instituto Nacional de Estadística (INE).
- Instituto Aragonés de Estadística (IAEST).
- Servicios Municipales Estadísticos.
- Servicios Estadísticos de los Ministerios.
- Banco de España.
- Oficina Estadística de la Comunidad Europea (EUROSTAT).
- Banco Central Europeo (BCE).
- Census Bureau, World Bank, OCDE...

3. CONCEPTOS BÁSICOS

La Estadística es la herramienta indispensable para todas las disciplinas que pretenden obtener conclusiones a partir del análisis de conjuntos de datos. Más concretamente, la Estadística es la disciplina que proporciona una metodología para recoger, organizar, sintetizar, analizar datos y hacer inferencias a partir de ellos.

Por **población**, universo o colectivo se entiende el conjunto de elementos, individuos o entes, definidos sin ambigüedad, del que se pretende estudiar una serie de características o comportamientos. En Estadística el término **individuo** puede referirse a cosas tan diversas como personas, provincias, empresas, edificios, etc. Una vez delimitada con toda precisión cuál es la *población* objeto de estudio, quiénes son los *individuos* que la forman, deberemos indicar cuáles son los **caracteres**, denominadas **variables**, que de ellos nos interesan. Dado un conjunto de individuos, la observación de si se realiza o no un determinado carácter en los mismos o la medida de la intensidad con la que se verifica se llama **dato estadístico**.

Precisando un poco más llamaremos **datos** al conjunto de valores en bruto, sin orientación dada (la cifra 1,8); **información** son los datos dotados de una estructura, habitualmente concedida mediante la inclusión de las unidades correspondientes, que los hace apropiados para una determinada finalidad (la altura 1,8 m); finalmente,

conocimiento es la interpretación de la información de un dominio específico (la altura 1,8 m correspondiente a una mujer pigmea).

De cara a estudiar una determinada característica de los individuos de una población podemos, bien recoger la información de todos los individuos que la forman y que se denomina **censo**, o bien de sólo una parte significativa de ellos tomando una **muestra** de la población. Motivos tales como la economía, rapidez, calidad, imposibilidad o la observación destructiva entre otros hacen que se opte por la alternativa muestral.

4. TIPOS DE VARIABLES

Es obvio que las características a estudiar pueden ser muy distintas y todos los experimentos examinan algún tipo de variable. La razón por la cual es importante conocer qué tipo de variables estamos manejando es porque las técnicas estadísticas a utilizar dependerán de ello y por tanto la validez de las conclusiones que se extraigan de su análisis.

No es lo mismo la cotización en bolsa de una empresa que el nivel de instrucción de un parado. En tanto que la cotización refleja una característica que de manera natural se asocia a un número, el nivel de instrucción refleja una cualidad y su observación no necesita recurrir a un número, sino a una gama de categorías o modalidades. Este ejemplo permite diferenciar los dos grandes tipos de variables: cuantitativas y cualitativas.

Variabes cualitativas o atributos son aquellas que reflejan una cualidad del individuo cuya observación no va ligada necesariamente a un número. Dichas variables se miden mediante categorías o modalidades que deben de cumplir los tres requisitos siguientes: estar bien definidas, ser mutuamente excluyentes y todos y cada uno de los individuos deben estar incluidos en alguna de las categorías. Aunque en ocasiones pueda asociarse un número a cada una de las modalidades, estos números no tienen el significado de cantidades.

Variabes cuantitativas son aquellas que para indicar su valor en un individuo se acude de manera natural a un número, ya que corresponden a características que se representan mediante cantidades y con las que pueden llevarse a cabo operaciones aritméticas que tienen completo sentido. Algunos ejemplos pueden ser: la estatura, la

edad, el salario, la temperatura, un índice bursátil o concentración de sustancias tóxicas en el ambiente (medidas en partes por millón). Las variables cuantitativas pueden, a su vez, ser discretas o continuas.

Variable cuantitativa discreta es aquella que toma un número finito o infinito numerable de valores. Por ejemplo, el número de estudiantes en un curso, el número de miembros en la unidad familiar y los precios de las acciones en la Bolsa de Madrid.

Variable cuantitativa continua es aquella que toma un número infinito de valores en un intervalo finito. La observación de una variable conceptualmente continua suele realizarse en la práctica de forma discreta a causa de las limitaciones de los instrumentos de medida. En este sentido, por ejemplo, las estaturas de las personas pueden en principio tomar un número infinito de valores, pero en la práctica se miden con una precisión de centímetros. Por ejemplo, la estatura, el peso, la temperatura.