



Socioestadística I

Análisis estadístico en Sociología

Capítulo I. INTRODUCCIÓN A LA SOCIOESTADÍSTICA

I. INTRODUCCIÓN. Definición e historia.

I.1. ¿Que es la Sociestadística?

La estadística es la ciencias de las regularidades que se observan en conjuntos de fenómenos naturales. Es la colección de métodos científicos que permiten el análisis e interpretación de la información numérica.

Muchas veces se interpreta o se reconoce a la estadística como una herramienta que se utiliza para manipular números o para cuantificar todo aquello que no seamos capaces de interpretar simplemente analizando los fenómenos.

En otras ocasiones, la estadística no es popular porque se identifica con las matemáticas como un instrumento un tanto abstracto y de difícil comprensión.

I.2. Historia sobre los métodos estadísticos.

- En la antigüedad no se desarrolló porque era un término mal visto, o bien por la filosofía (óptica pagana) o bien por la religión (óptica cristiana). Resultaba impío o irreverente suponer que los fenómenos obedecían a las leyes de probabilidad (pagano) o a un mandato divino (religión)
- Hay una necesidad por la estadística en el s. XVII en diversos campos:

- Juegos de azar
- Reclutamiento
- Dinero
- Seguros
- Mortalidad
- Astronomía

- El papel de la demografía, será muy importante en la segunda mitad del s. XVIII con las necesidades que surgieron a la hora de administrar las poblaciones y la ciudadanía. La información censal y los movimiento de población.
- En 1850 aparece el término de estadística. Aparece ligado a la actividad gubernamental - estadista o política. Se empieza a ser necesaria la utilización de herramientas que permiten el control y seguimiento de ciertas facetas de la vida administrativa de un país.



- Finales del s. XIX, aparecen las teorías de las regularidades en diversas disciplinas, fue entonces cuando aparecieron las herramientas de la **DISTRIBUCIÓN DE FRECUENCIAS** y **DISTRIBUCIÓN DE PROBABILIDADES** que veremos durante el temario.
- Desde el final del s. XIX hasta el fin de la 2ª Guerra Mundial se consolida el término de estadística como una herramienta imprescindible a la hora de controlar y administrar una serie de elementos de la sociedad, el estado y en definitiva de lo que hoy conocemos por Naciones.
- En la actualidad pasa a ser una herramienta imprescindible para la investigación y para la administración tanto pública como privada

2. FUNCIONES DE LA ESTADÍSTICA. Estadística descriptiva e inferencial

- **Kruskal** definía la estadística como la que ofrece métodos generales para encontrar pautas recurrentes y construir modelos que faciliten el estudio.

FUNCIONES GENERALES que realizan los métodos estadísticos:

- a. Resumen de los datos y extracción de la información relevante de los mismos.
- b. Búsqueda y evaluación de los métodos y pautas que ofrecen los datos.
- c. Contribuir al diseño eficiente de experimentos y encuestas
- d. Facilitar la comunicación entre los científicos

Tiene DOS FUNCIONES MÁS, aunque estas son más una crítica:

- a. En muchas ocasiones se utiliza por utilizar. Parece que es más científico por que se utiliza la estadística.
- b. Utilización excesiva de datos estadísticos o que no son necesarios o que no se tendría que utilizar.

TIPOS DE ESTADÍSTICA:

1. **La Estadística Descriptiva:** es el conjunto de instrumentos y temas relacionados con la descripción de colecciones de observaciones estadísticas (se refiere tanto a la población como a una muestra de la misma).
2. **La Estadística Inferencial o inductiva:** Se ocupa de la lógica y de los procedimientos para la inferencia o inducción de propiedades de una población en base a los resultados obtenidos de una muestra conocida



La Estadística Inferencial: se basa en la teoría de las probabilidades, rama de las matemáticas, disciplina deductiva que suministra la base racional para un razonamiento inductivo que es lo que llamamos Estadística Inferencial.

EL PROCESO en la Estadística Inferencial es el siguiente: presentamos los instrumentos de análisis estadístico descriptivo para luego tratar los problemas que comporta la inducción de propiedades a partir de los resultados obtenidos sobre muestras representativas.

3. RELACIÓN ENTRE ESTADÍSTICA Y SOCIOLOGÍA

- La estadística se utiliza para operar con números que reflejan valores de mediciones que se suponen satisfacen determinados supuestos.
- En investigación siempre distinguimos dos fases:

- a. La elaboración de un marco teórico pertinente
- b. La fase analítica del proceso de investigación

Todo ello debe de entenderse como un proyecto unitario

- Por tanto, la estadística, si es una ayuda, pero nunca un sustituto para un buen razonamiento teórico y un buen que hacer metodológico.
- Según **Durkein** (1972), la estadística expresa cierto estado de alma colectiva. Tenemos los elemento necesarios para una precisa conceptualización en Sociología.
- Antes, teníamos pocos datos y se suplía con la creatividad del investigador y ahora, hay excesivos datos pero hay cierta dificultad para ordenarlos y teorizar con ellos.
- Cuando hablamos de Estadística hay que tener presente una serie de conceptos y que nos van a permitir trabajar mejor con esta herramienta dentro de la sociología. Estos son:

- a. **VARIABLES:** Preguntas que se formulan para medir una propiedad o característica de las personas entrevistadas.
- b. **OBSERVACIONES ESTADÍSTICAS (DATOS):** Son las puntuaciones o valores que tomas las variables. Las posibles respuestas o contestaciones a esas preguntas.
- c. **UNIDAD DE ANÁLISIS (CASOS):** Son los individuos o unidades a los que preguntamos o recabamos información para una investigación.



d. POBLACIÓN ESTADÍSTICA O UNIVERSO: Conjunto de todos los valores de las variables que desea medir el investigador en todas las unidades de análisis. Podemos encontrarnos con dos tipos de población:

FINITAS: se limita en el tiempo y en el espacio

INFINITAS: No se limita en el tiempo y en el espacio

e. MUESTRA: subconjunto de una población y que en muchas ocasiones utilizamos como base para el estudio o investigación y que posteriormente el resultado obtenido trataremos de inferir a la población a la que pertenece dicha muestra. De eso se ocupa la estadística inferencial o inductiva.

4. MEDICIÓN DE VARIABLES

4.1. También para el proceso de medición de las variables se utilizan una serie de conceptos

a. MEDICIÓN: es el procedimiento de asignación de numerales a objetos o acontecimientos de acuerdo con ciertas reglas

b. VARIABLE: cualquier característica o propiedad de un objeto o acontecimiento que contenga dos o más categorías que nos permiten clasificarlo potencialmente

c. INDICADORES: Elementos cuantitativos que sirven para medir un concepto (ejemplo: años de escolaridad, nivel de ingresos, tipo de ocupación...).

d. ÍNDICE: la combinación ponderada de los valores de varios indicadores forman un índice que toma valores numéricos concretos (ejemplo: estatus social que se obtiene mediante la combinación de varios indicadores)

e. VALIDEZ: hace referencia a que el procedimiento utilizado mida realmente lo que se pretende medir.

f. FIABILIDAD: un instrumento que al utilizarlo repetidas veces y en idénticas circunstancias produce los mismos resultados.

4.2. Características de las categorías o sistemas de medición:

a. Debe de ser exhaustivo (suficientes números de categorías para clasificar). Ejemplo: 0,1,2,3 y 4 o más (familias por número de hijos).

- La categoría "otros" se usa para crear un sistema clasificatorio más completo e inclusivo.



- “El no sabe, no contesta”. Siempre hay problemas para interpretar, pero en ocasiones resulta necesario.
- b. Buen sistema clasificatorio: categorías mutuamente exclusivas. Clasificar cada caso únicamente en una categoría (utilización por ejemplo del “menos de 25 años” o el “más de 25 años”. No deja nunca a ningún caso sin categoría.
- c. Que el procedimiento de medición sea siempre lo más preciso posible, que haya el mayor número de distinciones. Ejemplo: La división de los partidos políticos: La división: Partidos de Izquierda, Partidos de Derecha, Otros es menos precisa que: Izquierda, Centro, Derecha, Nacionalistas y Otros partidos.

4.3. Tipos de variables según el nivel de medición:

- a. **NOMINALES:** la propiedad o característica estudiada del objeto o acontecimiento solo se puede agrupar en categoría lógicamente exhaustivas y mutuamente excluyentes de forma que pueda establecerse claramente equivalencias o diferencias. Cuando decimos de algo o alguien una característica sin establecer ningún tipo de orden y sabiendo que cada característica es diferente de la otra. Las categorías pueden ser tanto nombres como números pero se interpreta como que son diferentes y no establecen ni un orden, ni superioridad.
 - Con las variables nominales se consigue realizar la **CLASIFICACIÓN** (la operación más simple y básica) de forma que los elementos se dividen en categorías y se deciden cuáles son más parecidas y cuáles son más diferentes
 - Las variables nominales tienen la propiedad de la equivalencia y engloba además las siguientes propiedades:
 1. Propiedad **REFLEXIVA:** Si $A = A$ para todo valor de A
 2. Propiedad **SIMÉTRICA:** Si $A = B$ también $B = A$
 3. Propiedad **TRANSITIVA:** Si $A = B$ y $B = C$ entonces $A = C$
 - Con escalas nominales no se pueden realizar operaciones aritméticas que se realizan usualmente con números (suma, resta, multiplicación y división), no se pueden calcular medias, medianas, etc. A veces, se hace necesario cuantificar las variables, en tal caso, se pueden convertir en variables ficticias, asignando a una categoría el valor 1 y al resto el valor 0, es lo que llamamos variables dicotómicas.
- b. **ORDINALES:** Son iguales que las nominales, con las mismas propiedades, pero además, incluyen la propiedad de que las categorías pueden ser ordenadas en el sentido de menor que o mayor que. No solo podemos clasificar, sino que además podemos ordenar. Asignamos una cualidad o característica de algo o de alguien estableciendo



un orden de las cualidades o características.

1. Propiedad **IRREFLEXIVA**: No es cierto que para toda A sea $A > A$
2. Propiedad **ASIMÉTRICA**: Si $A > B$ entonces $B < A$
3. Propiedad **TRANSITIVA**: Si $A > B$ y $B > C$ entonces $A > C$

- Tampoco podemos realizar operaciones aritméticas con este tipo de variables. Si podemos establecer un orden entre las categorías pero no la diferencia numérica entre esas categorías aunque le asignemos un valor numérico, sólo establece el orden.
- c. **INTERVALO**: Incluye las características de las nominales y ordinales pero además sus categorías se definen en términos de unidad de medición. Se asigna a un objeto o acontecimiento estudiado números que además de establecer un orden (**ORDINAL**) permite la interpretación de la diferencia entre dos medidas. Decimos de algo o de alguien una cualidad o característica estableciendo un orden y además en valor numérico dentro de ese orden o clasificación.
- Existencia de una unidad de medición común y constante, que permite asignar un número real a todos los pares de objetos del conjunto ordenado.
 - Hay problemas y dificultades en encontrar variables que realmente utilicen un nivel de medición de intervalo, y hay también problemas de errores y sesgos de utilizar medidas de intervalo con niveles nominales y ordinales, pero aquí si podemos realizar operaciones aritméticas, podemos especificar el cociente de cualquier par de intervalos.
- d. **COCIENTE, PROPORCIÓN O RAZÓN**: Tiene todas las características de una medida de intervalo y además se le puede asignar un punto de origen verdadero de valor 0. En este caso hablamos de una escala con cero absoluto o escala de cociente o proporción.

CONCLUSIONES:

- Para cada tipo de variable existen unos procedimientos apropiados para hacer uso mejor de la información. Si utilizamos procedimientos estadísticos apropiados para niveles bajos de medición con puntuaciones definidas a un nivel de medición alto, se producirá una pérdida de información, pues las propiedades de los niveles de medición son acumulativas. Ejemplo: utilizar un procedimiento estadístico apropiado para escalas ordinales con variables de intervalo. Al contrario, sería un error. Actuaríamos como si las variables contuvieran más información de la que realmente tienen.



- Los procedimientos estadísticos para niveles de medición más elevados (INTERVALO) permiten una descripción más concisa de los datos y con niveles de medición ordinal, un procedimiento para medidas de intervalo no produciría grandes errores en los resultados.

Existen otros tipos de clasificación de las variables:

VARIABLES CONTINUAS Y DISCRETAS:

- a. **CONTINUAS:** variables con infinitos valores fraccionados, valores en cualquier punto de la escala ininterrumpida, son las que llamamos **variables de intervalo**. Ejemplo: la edad en años (15,16,17,17 y 3 meses...).
- b. **DISCRETA:** En esta caso la escala de medición está interrumpida por espacios en la escala numérica. La mayor parte de las variables discretas son **variables nominales**. Ejemplo: el número de hijos de una familia. 1,2,3,4. No puede haber 2,5 hijos.

VARIABLES DEPENDIENTES E INDEPENDIENTES:

- a. **DEPENDIENTES:** los valores que toma la variable depende de los valores que presentan otras variables.
- b. **INDEPENDIENTES:** Al contrario, son aquellas variables cuyos valores explican a las variables dependientes.

BIBLIOGRAFÍA

- DURKHEIM, Emile: *Las reglas del método sociológico*, Buenos Aires, La Pléyade, 1972 (e.o., 1895).
- GARCÍA FERRANDO, M.: *Socioestadística. Introducción a la estadística en sociología*, Madrid, Alianza Editorial, 1989.
- GARCÍA FERRANDO, M.: *Sobre el método. Problemas de investigación empírica en Sociología*, Madrid, C.I.S., 1979.
- JIMÉNEZ BLANCO, José, et al.: *La conciencia regional en España*, Madrid, C.T.S., 1977.
- LINZ, Juan Jose: *Informe de la Encuesta sobre la Juventud 1977*, Madrid, Instituto de la Juventud, 1978.