

Introducción a la Estadística

Gabriel Cavada Ch.¹, Claudio Silva Z.¹

¹ División de Bioestadística y Demografía. Escuela de Salud Pública, Facultad de Medicina, Universidad de Chile.

¿Qué es Estadística? Etimológicamente, el vocablo deriva de Estado y significa “contar los bienes del Estado”. Los albores de esta disciplina se encuentran en la antigüedad; por ejemplo, de esta forma las autoridades del Egipto faraónico contaban sus bienes y registraban la profundidad del río Nilo en cada estación del año. Recuérdese también que Jesucristo nace en Belén porque un edicto del emperador romano ordena un censo para conocer el número y las características de los habitantes del Imperio, por lo que las personas debían concurrir a sus lugares de nacimiento y San José había nacido en Belén de Judá.

Sin embargo, la disciplina, desde aquellos lejanos tiempos hasta nuestros días ha ampliado su quehacer, perfeccionando tanto sus métodos que hoy es la disciplina más usada por el resto de las Ciencias y su desarrollo ha sido vertiginoso, sobretodo desde los inicios del siglo XX.

Así, hoy decimos que la Estadística es la disciplina que se ocupa de:

1. La recolección, organización y procesamiento de datos.
2. La obtención de inferencias al universo en estudio cuando se observa sólo una parte de éste.
3. Decidir bajo incertidumbre.

Método científico y estadístico

El hombre busca en forma ineludible una explicación racional a los fenómenos que lo rodean. El Método Científico le ayuda a organizar adecuadamente la observación de los hechos y a enunciar las leyes que los rigen. En la gran mayoría de los casos reales, el enunciado de las leyes se complica porque el fenómeno estudiado tiene multicausalidad; esta contingencia se trata de superar con el Método Estadístico.

Método Científico es un procedimiento que se aplica al ciclo completo de la investigación, desde el enunciado del problema hasta la evaluación de los resultados obtenidos. Al aplicar el Método Científico se distinguen las siguientes etapas:

1. Detección y enunciado del problema: es la descripción de una situación problema o es el planteamiento de una pregunta.
2. Formulación de la hipótesis: es una respuesta o explicación al problema enunciado que se hace en base al conocimiento científico existente.
3. Deducción de una consecuencia verificable: como la hipó-

tesis es una explicación general, a menudo ocurre que no se puede investigar en forma directa, luego se procede a deducir, lógicamente, consecuencias particulares de la hipótesis.

4. Verificación de la consecuencia: en ciencias exactas esto se realiza usando lógica pura; sin embargo, en ciencias no exactas la verificación se hace a través de la recolección de información o la observación de los fenómenos, lo que hace necesario la aplicación de Procedimientos Estadísticos.
5. Conclusión: consiste en la aceptación, modificación o rechazo total de la hipótesis planteada.

Método Estadístico es el que proporciona las estrategias necesarias para Recolectar y Analizar la información requerida. El Método Estadístico distingue dos etapas: Planificación y Ejecución.

En la Planificación se deben considerar los siguientes aspectos:

1. Definición de objetivos: es la descripción formal del problema que da origen a la investigación. Se debe señalar detalladamente lo que se va a investigar, el qué, cómo, dónde, cuándo y por qué.
2. Universo del estudio: es la definición del conjunto desde el cual se extraerá la información y hacia el que se generalizarán las conclusiones obtenidas.
3. Diseño de la muestra: la Teoría de Muestreo garantiza que la información que se genere permita proyectar válidamente las conclusiones al Universo de interés.
4. Definición de las unidades de observación (que objetos observaremos), las escalas de clasificación y las unidades de medida.
5. Preparación del Plan de Tabulación y análisis de la información: aquí se determinan las formas de presentar y analizar la información recolectada.

En la fase de Ejecución se pueden reconocer los siguientes aspectos:

1. Recolección de la información.
2. Elaboración de la información.
3. Análisis de los resultados.

Unidad de análisis, atributos, variables, escalas de medida, población y muestra

Unidad de análisis: Una vez definido el problema que se va a investigar, se definen naturalmente los objetos que serán obser-

Rincón de la Bioestadística

vados (seres humanos, piezas histológicas, etc.), los que llamaremos en lenguaje técnico unidades de análisis.

Atributos: Teniendo definidas las unidades de análisis, obviamente ellas presentan características que importan para nuestro estudio: Por ejemplo, si es de nuestro interés analizar antropométricamente un conjunto de seres humanos adultos, podemos consignar algunas características esenciales tales como edad, sexo, talla, peso, longitud de la circunferencia de la cintura, etc. Estas características las denominaremos atributos.

Variables: Cuando se han definido los atributos a estudiar, podemos ya observar unidades de análisis especificadas y los atributos quedan consignados como características únicas del objeto que estamos estudiando. Por ejemplo, si observamos una determinada persona podemos consignar:

- Edad: 50 años
- Sexo: Femenino
- Talla: 165 centímetros
- Peso: 72 kilogramos
- Longitud de la circunferencia de la cintura: 90 centímetros

Cuando los atributos ya han sido evaluados, reciben el nombre de variables del estudio.

Escalas de medida: Cuando procedemos a medir las variables del estudio, debemos tener presente que estamos consignando valores con unidades de medida y, por consiguiente, introduciendo unidades de medida. Estas escalas de medidas pueden ser: Nominales, Ordinales o Intervalares (o de Razón). Estas escalas tienen diferente Poder de Clasificación. Presentemos el siguiente esquema:

Variables	Escala de Medida	Poder de Clasificación
Cualitativas	Nominal	Sólo es capaz de nombrar o etiquetar la unidad de análisis. Por ej.: Sexo, Raza, Nacionalidad.
	Ordinal	Es capaz de nombrar, pero además introduce una jerarquía en las unidades observadas. Por ej.: Grado que se cursa en el sistema escolar básico, Nivel educacional.
Cuantitativas	Intervalar y de Razón	Es capaz de nombrar, jerarquizar, pero además permite hacer comparaciones matemáticas entre las unidades de análisis. Por ej.: Talla, Edad, Peso, Temperatura.

La distinción entre la Escala Intervalar y la de Razón es que en esta última el número cero indica ausencia de cantidad, mientras que en la primera, el cero es sólo un referente.

Estas escalas de medida son inherentes a la forma como se mide la variable. Las variables de escala intervalar, debido a su naturaleza numérica, se pueden clasificar en variables Discretas o Continuas. Las variables discretas, formalmente están relacionadas con los números Naturales (0, 1, 2, 3,...) es decir, su función es Contar; como ejemplo tenemos: cantidad de hijos, cantidad de caries, días trabajados, etc. Las variables continuas,

formalmente están relacionadas con los números Reales y su función es medir, en el sentido físico; como ejemplos tenemos: la Masa, la Longitud, el Tiempo, etc.

Una importante observación es que una variable inherentemente continua, por razones operativas se discretiza, esto es: si a usted le preguntan por su Edad, su respuesta estará en Años Cumplidos (30 años, 23 años,...); sin embargo, la Edad de una persona es una variable que indica tiempo de vida y, si se quisiera ser exacto, se tendría que contestar 30,213 años (30 años con 2 meses, 16 días, 16 horas y 19 minutos) lo que parece impracticable. La precisión con que se mide una variable va de acuerdo al interés de la investigación como se señaló en el párrafo de Método Estadístico.

Población: Llamamos Población al Conjunto Universo de las unidades de análisis; la población puede ser de tamaño finito o infinito. Por ejemplo, si se desea averiguar el diámetro de los eritrocitos en una persona determinada, la población en estudio son todos los eritrocitos que tiene esta persona, esta población en la práctica es infinita. Si se desea analizar el nivel de creatinina en pacientes en diálisis en Santiago de Chile, la población es finita.

Muestra: Como se puede apreciar, en muchos casos trabajar con una población completa puede resultar muy costoso o simplemente impracticable, de allí la necesidad de tomar una Muestra de la Población. Formalmente, una muestra es un Subconjunto Finito y factible de la Población que debe cumplir características ineludibles para lograr que las conclusiones estadísticas sean válidas; así, las características de una "buena muestra" son:

1. Aleatoria: garantiza que los elementos que componen la muestra fueron escogidos completamente al azar, es decir, sin predilección alguna por incluir o excluir determinada unidad de análisis.
2. El tamaño de la muestra, que es el número de unidades de análisis que se deben escoger; debe ser lo suficientemente grande como para garantizar la generalidad de los resultados. La determinación del tamaño de una muestra no es un problema trivial y constituye una especialización de la Estadística llamada Teoría de Muestreo.