

La evolución y desarrollo de la "Estadística" en el mundo actual, es tal que sería difícil dar una definición precisa de este concepto. A riesgo de agregar una más a las tantas existentes daremos la siguiente.

DEFINICIÓN 1.1 La Estadística es una ciencia que proporciona un conjunto de métodos que se utilizan para recolectar, resumir, clasificar, analizar e interpretar el comportamiento de los "datos" con respecto a una característica materia de estudio o investigación. En primera instancia se encarga de obtener información, describirla y luego usa esta información a fin de predecir "algo" respecto a la fuente de información.

La Estadística actual es el resultado de la unión de las disciplinas que evolucionaron independientemente hasta confluir en el siglo XIX: la primera es el "cálculo de probabilidades", que nace aproximadamente en el siglo XVII como teoría matemática de los juegos de azar; la segunda es la "Estadística" o ciencia del estado, del latín *status* (aunque sobre este significado etimológico de estadística no hay un criterio único de los autores, pues para unos se deriva del griego *statera* que significa balanza, para otros se deriva del alemán *staat* que significa estado) que estudia la descripción de los datos y tiene raíces más antiguas. La integración de ambas líneas de pensamiento da lugar a esta nueva ciencia.

1.2 DIVISIÓN DE LA ESTADÍSTICA

El campo de la Estadística generalmente está dividido en dos grandes áreas: Estadística Descriptiva y Estadística Inferencial.

1.2.1 ESTADÍSTICA DESCRIPTIVA

De la definición de Estadística, el lector notará dos aspectos bien remarcados, el primero: obtener información, procesarla y describirla, es lo que constituye la **Estadística Descriptiva**. Precisaremos con la definición siguiente.

DEFINICIÓN 1.2 Estadística Descriptiva es el conjunto de métodos que implican la recolección, presentación y caracterización de un conjunto de datos a fin de describir en forma apropiada las diversas características de estas. Es decir, un estudio estadístico se considera "**descriptivo**" cuando sólo se analiza y describe los datos.

EJEMPLO 1.1 Un gerente de personal desea conocer las aptitudes de cinco secretarías que trabajan en una dependencia particular de una compañía. Se aplica una prueba de aptitudes a las cinco secretarías y las calificaciones son 85, 90, 93, 82 y 95 puntos. Supongamos que la medida estadística que emplea el gerente de personal es la **aptitud**

promedio o media aritmética, la cual es la suma de los valores observados dividida entre el número de observaciones. Entonces, la calificación promedio es:

$$\frac{85 + 90 + 93 + 82 + 95}{5} = \frac{445}{5} = 89 \text{ puntos}$$

El resultado se limita a los datos obtenidos en este caso particular y no implica ninguna generalización acerca de las aptitudes de las secretarías de otras oficinas de la misma compañía. Es decir, el gerente está usando estadística para describir aptitudes de las secretarías de esa oficina. Este método es de naturaleza descriptiva, debido a que el promedio condensa y describe la información obtenida.

Los gráficos, tablas y mapas que muestra datos de tal forma que sean más fáciles de entender son todos ejemplos del uso de Estadística Descriptiva.

1.2.2 ESTADÍSTICA INFERENCIAL

El segundo aspecto de la definición de Estadística es: predecir "algo" con respecto a la fuente de información, es lo que constituye la **Estadística Inferencial o Inferencia Estadística**, la definición es la siguiente.

DEFINICIÓN 1.3 La Inferencia Estadística es el conjunto de métodos o técnicas que posibilitan la generalización o toma de las decisiones en base a una información parcial obtenida mediante técnicas descriptivas.

Es decir, un estudio estadístico, se considera inferencial cuando se pretende inferir o predecir conclusiones que añaden a toda la fuente de información de donde proviene los datos.

Ahora bien esta predicción se hace con un cierto grado de confianza; este grado de confianza se mide por la "probabilidad". Por tanto, el cálculo de probabilidades, piedra angular de la inferencia estadística está como puente entre las dos partes de la Estadística.

Aunque la Estadística descriptiva es importante para caracterizar y presentar información de los datos, sin embargo, el desarrollo de la Inferencia Estadística es lo que ha conducido a la gran expansión en la aplicación de los métodos estadísticos.

EJEMPLO 1.2 Suponga ahora en el ejemplo 1.1, que el gerente de personal desea conocer la aptitud promedio de todas las secretarías de la compañía, pero carece de tiempo e de los recursos para aplicar una prueba de aptitud a todas ellas. Entonces decide usar la aptitud promedio de las cinco secretarías para estimar la aptitud promedio de todas las secretarías de la compañía. El proceso de estimar esta aptitud promedio global será un problema de Inferencia Estadística.

2. Cinco neumáticos para automóvil de marca A y cuatro neumáticos de marca B se prueban para determinar su duración en servicio. La duración para los neumáticos marca A, son: 36 000, 29 000, 33 000, 37 000 y 40 000 km; para la marca B, son: 29 000, 31 000, 33 000 y 35 000 km. De las siguientes declaraciones hechas en base a estas cifras, identifique las que provienen de métodos descriptivos y las que provienen de inferencia estadística.
 - a. La duración promedio de los cinco neumáticos marca A es mayor que la de los cuatro neumáticos marca B.
 - b. Probablemente, la duración promedio de todos los neumáticos marca A sea casi 35 000 km, mientras que la de los neumáticos marca B sea aproximadamente de 32 000 km.
 - c. Si el precio de los neumáticos marca A es el mismo que de los neumáticos marca B, Ud. recomendaría los neumáticos marca A a todos sus amigos y parientes.
3. En cuatro pruebas de matemática, Juana recibió calificaciones de 17, 18, 15 y 14; mientras que Juan recibió calificaciones de 14, 13, 16 y 13. A partir de estos datos se llega a las siguientes conclusiones, ¿cuáles de éstas incluyen el método descriptivo y cuáles se obtienen mediante la inferencia estadística?
 - a. El promedio de las calificaciones de Juana es 16 y el promedio de las calificaciones de Juan es 14.
 - b. Juana es mejor estudiante que Juan.
 - c. Probablemente en la siguiente prueba, Juana obtenga calificaciones más elevadas que Juan.
 - d. La diferencia entre los dos promedios es 2 puntos.
4. ¿En qué contexto se emplea generalmente la palabra "estadística" en radio y televisión?
5. Clasifique cada una de las afirmaciones siguientes ya sea como inferencias o métodos descriptivos.
 - a. El año pasado, en la Universidad Nacional del Callao, el puntaje promedio del examen de admisión fue 85.
 - b. El Dr. García, un ecólogo, informó que en cierto río de la selva la carne de los peces contienen un promedio de 300 unidades de mercurio.
 - c. La compañía "RM" predijo quién sería el ganador en una elección presidencial después de conocer los resultados de las votaciones de 25 mesas de sufragio de las 2 800 mesas que hubo en total.

1.3 POBLACIÓN Y MUESTRA

La palabra "población" igual que sucede con el término "estadística" tiene varios significados. En el uso común, se refiere a todas las personas de una región, localidad o país. (La población de Lima es 6 millones; la población de China es 900 millones, etc). El concepto de "población", estadísticamente hablando, es un concepto más amplio de lo que se tiene al referirse a los habitantes de un Departamento o un País; población desde el punto de vista estadístico se define como sigue.

DEFINICIÓN 1.4 Población es la colección de todos los individuos, objetos u observaciones que poseen al menos una característica común.

Los términos población y universo, suelen usarse indistintamente.

EJEMPLOS 1.5

- a. Las edades de los estudiantes del Perú.
- b. Las edades de los alumnos del sistema universitario peruano.
- c. Los diámetros de la producción diaria de tuercas.
- d. Los pesos de los melones de una cosecha de la cooperativa los "meloneros".
- e. Las placas de los automóviles que circulan en un país.
- f. Los pacientes con poliomielitis sometidos a una determinada terapéutica de rehabilitación.
- g. Los enfermos de SIDA, tratado con uno de tres tratamientos diferentes.

En cada uno de los ejemplos anteriores, existe al menos una característica común para todos los elementos integrantes de la población. Y cada situación diferente implica una población diferente. En el ejemplo 1.5g, en lugar de una población, se tiene tres poblaciones diferentes, cada población está definida por cada uno de los tres tratamientos que se utilizan.

Es importante definir la población de acuerdo a la naturaleza y extensión del problema bajo estudio. Al hablar de naturaleza, entendemos la característica materia de estudio; es decir, si quisiéramos estudiar los pesos, alturas y edades de las personas, la población estará formada por los pesos, alturas y edades respectivamente. Y al hablar de extensión definimos la población tan extensa como sea necesario; es decir, si quisiéramos estudiar las características anteriores en los alumnos del sistema universitario peruano, la población estará referida al sistema universitario peruano; y si estudiáramos interesados en estudiar éstas características en los alumnos de las universidades de Lima, la población estará referida solamente a ellas. Al referirnos a la naturaleza del problema, optamos también por distinguir entre lo que podemos llamar población "objeto" y población "objetivo"; entendemos por población objeto, el conjunto de elementos materia de estudio y por población objetivo las diferentes

