

# Clase 1. Introducción a la Estadística

# Trabajo estadístico

- En la actividad diaria, permanentemente se toma conocimiento de informaciones de carácter **cuantitativo o cualitativo**, sobre diversos fenómenos, acontecimientos, manifestaciones o simplemente hechos ocurridos en nuestro ámbito. Algunas de estas informaciones son de carácter estadístico y otras no lo son.
- Por ejemplo, al concluir los partidos de fútbol programados para una particular fecha, los periodistas deportivos informan sobre los resultados obtenidos por los equipos de fútbol y la tabla de posición, del primero al último puesto. Esta información, que es expresada en forma cuantitativa y cualitativa (número de goles y número de orden que ocupa cada equipo), a nadie se le ocurriría llamarla **información estadística**. Su propósito es difundir resultados sobre acontecimientos que se suponen de interés para muchas personas. Sin embargo, esta misma información ofrecida como antecedentes históricos para pronosticar resultados de futuros partidos de fútbol es comúnmente llamada **estadísticas** del fútbol.

# ...¿Cuál es la diferencia?

- La diferencia consiste en el propósito para el cual se produce u ofrece la información. En el primer caso sólo da a conocer resultados; en el segundo caso estos mismos resultados son ofrecidos con el fin de que los interesados puedan **sacar conclusiones** (**realizar inferencias**) y eventualmente **tomar decisiones**.
- Igualmente, una empresa lleva su contabilidad en libros expresamente diseñados, donde se registra la información comercial y financiera de las operaciones de la empresa. Esta información es de **carácter contable y no estadística** dado que su propósito es llevar el control de las operaciones que realiza el establecimiento.
- Sin embargo esta misma información puede ser **estadística** si se presenta con fines de **analizar y proyectar** la marcha de dicha empresa. Una vez más, la diferencia obedece a que en el primer caso la información se registra con fines de control interno e impositivo mientras que en el segundo caso su propósito es **inferir o deducir** a base de esta información la situación presente y futura de la empresa.

## ...Más ejemplos

- El Registro Civil que se lleva en distintas dependencias del Estado recopilan los nacimientos, las defunciones o matrimonios ocurridos en un país. Dicha información tiene como propósito fundamental establecer derechos de los ciudadanos (nacionalidad; bienes conyugales; herencias etc.), y su información en tal sentido **no es estadística**. Sin embargo esta misma información de los Registros constituye la base de las llamadas **Estadísticas Vitales** las que pretenden ofrecer información básica para el estudio de la dinámica de la población de un país.
- Un ejemplo similar se encuentra en los registros de exportaciones e importaciones que lleva las Aduanas con fines impositivos y las llamadas **Estadísticas de Comercio Exterior** que son obtenidas de tales registros y tienen como propósito facilitar el estudio de políticas referentes a las exportaciones e importaciones.

## ...Por lo tanto:

- Para que una información sea considerada **estadística** debe estar orientada a facilitar la realización de estudios, de inferencias inductivas o la toma de decisiones en cualquier actividad o área del conocimiento.

### Áreas de Aplicación

- Los **ingenieros** de control de calidad recopilan datos sobre la fiabilidad de partes y productos fabricados, calidad de procesos, etc. para mejoramiento del producto.
- Las oficinas de estadística del **gobierno** publican cada mes nueva información numérica sobre la inflación y el desempleo, a través de índices de precios, tasa de desempleo, etc.
- Quienes se dedican a realizar previsiones, los **economistas**, los asesores financieros y los que determinan las políticas de una empresa, industria y del gobierno estudian estos datos para tomar decisiones basadas en la información obtenida.

## ...Más áreas de aplicación

- Con el fin de ofrecer un tratamiento adecuado a sus pacientes, los dentistas, los *médicos* y en general el personal de un centro de salud, deben entender la información estadística de las investigaciones que se publican en las revistas médicas sobre efectos de nuevas drogas, tratamientos de enfermedades, etc.
- En *política*, los funcionarios que ocupan cargos directivos consideran las estadísticas de la opinión pública para definir la legislación que quieren sus votantes.
- Las *empresas* basan sus decisiones en estudios de mercado sobre los patrones de compra de los consumidores, pruebas de nuevos productos, etc.

## ...Entonces

- De acuerdo con la experiencia, virtualmente toda persona involucrada en la toma de decisiones necesita conocimientos de **análisis estadístico**. Muy frecuentemente, en compañías grandes, se utiliza la **estadística** en forma habitual. Cuando se solicita personal para esos trabajos, se piden conocimientos sólidos de **análisis estadístico**.
- En cualquiera de estos u otros ejemplos se puede observar que tanto el **registro de los datos** que interesan, como su manejo o utilización, no siempre es simple y se necesitan procedimientos adecuados para llevarlos a cabo.

# Algunos conceptos importantes

- El **proceso de inferir** consiste en establecer relaciones de dependencia entre observaciones o postulados con conclusiones o deducciones.
- **Ejemplo 1:** Un automovilista transita por una autopista y **observa** que todos los automóviles vienen en sentido contrario. Entonces **concluye** que en esa autopista él es el único que transita correctamente y todos los demás lo hacen a contramano.
- **Ejemplo 2:** Por definición (o postulado), un número es par si es múltiplo de 2 y se **observa** que 8 es múltiplo de 2. Entonces se **concluye** que 8 es un número par.
- **Ejemplo 3:** Se requiere un promedio de 4.0 puntos para aprobar un Curso. Se **observa** que el promedio de un alumno antes de rendir el examen final es 5.8 puntos. Entonces se **concluye** que es muy posible que el alumno apruebe el Curso.

## ...Más ejemplos

- **Ejemplo 4:** Una persona que va a salir a la calle, debe decidir si lleva o no lleva paraguas. Como **dato** cuenta con la información del servicio meteorológico que pronostica lluvia con 70% de posibilidad. Entonces la persona **concluye** que va a llover y decide llevar el paraguas.
- **Ejemplo 5:** Por definición un animal es herbívoro si sólo se alimenta de vegetales y se conoce como **dato** que el oso come vegetales y carne. Luego se **concluye** que el oso es herbívoro.
- **Ejemplo 6:** Una encuesta a 400 personas de una ciudad muestra que el 60 % de las mujeres son “conservadoras” frente al 40% de los hombre. Entonces el investigador concluye que las mujeres de esa ciudad son más conservadoras que los hombres.
- Las inferencias pueden ser catalogadas en, al menos, dos grandes grupos a saber: **Inferencias Inductivas** e **Inferencias Deductivas**.

# Inferencia

- **Inferencias Inductivas** son las que, a partir de datos u observaciones en los cuales nos parece encontrar alguna semejanza o propiedad, concluimos en extender dichas propiedades o semejanzas a un universo mayor. Los ejemplos 1, 3 , 4 y 6 corresponden a inferencias inductivas. Se observa que la inferencia inductiva va de lo particular a lo general.
- **Inferencias Deductivas** son las que, a partir de enunciados o definiciones conocidos, concluye en una afirmación vinculada a los mismos. Los ejemplos 2 y 5 son ejemplo de inferencias deductivas. Se dice que estas inferencias van de lo general a lo particular.

# Objetivos de la Estadística

- La **Estadística** es la disciplina que trata con los métodos y procedimientos que facilitan la formulación de **inferencias inductivas**. En forma resumida diremos que el propósito de la **Estadística** es favorecer la realización de **inferencias inductivas** basadas en observaciones o datos, mediante:
  - 1.- El **resumen** de la información contenida en un conjunto de datos, sin que se pierdan aspectos relevantes para las **inferencias inductivas** que se puedan formular a base de tal resumen.
  - 2.- El **dimensionamiento** y, en lo posible, el control del inevitable error inherente a toda **inferencia inductiva** basadas en observaciones o datos.

# ...Conclusiones

- La primera conclusión que resulta de esta afirmación es que la **Estadística** es una disciplina de servicio.
- La segunda conclusión es que, dado que la **inferencia inductiva** es la base del desarrollo de la mayor parte de las ciencias así como de la toma de decisiones, la **Estadística** presta sus servicios tanto al desarrollo del conocimiento científico en las distintas ramas de las ciencias aplicadas, como a la toma de decisiones referidas a los más variados ámbitos y en particular a las que se enmarcan dentro del diseño, seguimiento y evaluación de políticas, programas y proyectos públicos y privados en las más diversas áreas.

# Métodos estadísticos

- Los métodos y procedimientos que utiliza la **Estadística** para cumplir su propósito de favorecer o apoyar la realización de **inferencias inductivas** basadas en **observaciones** pueden ser clasificados en dos grandes ramas de esta disciplina:
- **Estadística Descriptiva o Análisis de Datos** (Pre-Inferencial): Tiene como propósito resumir la información contenida en un conjunto de observaciones (con el propósito de favorecer o facilitar la formulación de **inferencias inductivas** por parte del usuario de ese resumen).
- **Inferencia Estadística**: El propósito de la Inferencia Estadística es realizar **inferencias inductivas** basadas en observaciones, controlando el riesgo de cometer el **error inferencial** (la posibilidad de **error** es inherente a toda **inferencia inductiva** puesto que nadie puede **inferir** con certeza sobre un todo observando apenas una parte).

# Observaciones y algunas definiciones

- Contar con un **conjunto de unidades o individuos (observaciones)**, es condición necesaria para llevar a cabo un trabajo estadístico. Estas observaciones pueden referirse a **características o cualidades** de las unidades (alto, medio, bajo; masculino, femenino; bueno; malo; etc.) o **cuantitativas** (ingreso monetario; estatura; peso; edad; número de alumnos; etc.)
- **Unidades de Estudio**: son los elementos cuyas características se desean estudiar.
- **Población o Universo de Estudio**: es el conjunto total de unidades o individuos de estudio que se desean investigar. Esta población puede ser finita o infinita. Si es finita, el número de unidades o elementos de la población se denota por  $N$ .

## ... más definiciones

- **Muestra:** es un subconjunto (una parte) de las unidades de la **población** cuyo estudio puede ser utilizado para sacar conclusiones sobre características de la **población**. Si es finito, el número de unidades o elementos de la **muestra** se denomina: tamaño de la muestra y denota por  $n$ .
- **Variables:** se refieren a las características o propiedades que se desean estudiar o investigar. Las variables se denotan con letras mayúsculas ( $X, Y, Z$  o bien  $X_1, X_2, X_3, \dots, X_n$ ). Los valores particulares que toma una variable se denota con la correspondiente letra en minúscula ( $x, y, z$  o bien  $x_1, x_2, x_3, \dots, x_n$ ).
- Las variables pueden ser **cuantitativas** o **cualitativas**.

# ...Variables

- **Variables cuantitativas**: si sus valores son números. Estas variables, a su vez, pueden ser clasificadas como:
- **Discretas** si el conjunto de sus valores posibles es finito o contable.
- **Continuas** en otro caso.
- Cuidado que podemos utilizar números simbólicamente. Debe tener sentido realizar **operaciones aritméticas** con esos símbolos. Por ejemplo, consideremos una variable que representa el resultados de un proceso de fabricación. Si “bueno” es representado por “1” y “malo” por “0” entonces no tiene sentido interpretar  $0 + 1 = 1$  como: malo + bueno=bueno. Por lo tanto en este caso 0 y 1 son sólo símbolos y la variable no es cuantitativa.

# ...Variables

- **Variables Cualitativas** si sus valores indican niveles (pequeño, mediano, grande) o categorías o clases (masculino, femenino; a favor, en contra; bueno, defectuoso: deficiente, regular, bueno, muy bueno, etc.). Las variables cualitativas se clasifican en:
- **Ordinales** cuando sus valores se pueden ordenar (pequeño, mediano, grande)
- **Nominales** cuando sus valores no son susceptibles de ordenación (mujer, hombre).
- **Unidad de Información:** son las unidades a las cuales se les mide o consulta sobre cada variable a investigar. La Unidad de Información puede coincidir con la Unidad de Estudio aunque no necesariamente debe ser así.

# ...Ejemplos

- **Ejemplo 7:** Se desea estudiar la edad, género, peso, estatura, e ingreso monetario del hogar de los alumnos que asisten a cierta universidad en un año determinado mediante las variables.
  - Unidades de estudio
  - Población
  - Muestra
  - Variables
  - Unidad de información
- **Ejemplo 8:** Se desea estudiar el tiempo empleado por un alumno en resolver un examen.
  - Unidad de estudio
  - Unidad de información
  - Población
  - Muestra