

# INTRODUCCIÓN A LA ESTADÍSTICA

Facultad de Ciencias Básicas  
Universidad Católica del Maule

Talca, Chile

- La **estadística** desempeña un papel fundamental, y cada vez más relevante, en prácticamente todas las áreas de la ciencia.
- **Estadística** proviene de la palabra **estado** porque originalmente estaba relacionada estrictamente al estado (**gobierno**).
- La **estadística** se utiliza como herramienta al servicio de las ciencias donde la variabilidad y la incertidumbre forman parte de su naturaleza, entre estas se encuentran: ciencias actuariales, ciencias ambientales, ciencias agropecuarias, ciencias básicas, ciencias económicas y administrativas, ciencias de la ingeniería, ciencias naturales, ciencias de la salud, ciencias sociales, políticas y humanas, entre otros campos del conocimiento científico y del comportamiento humano.

“LA ESTADÍSTICA ES  
EL ARTE DE TRANSFORMAR  
DATOS EN INFORMACIÓN”

# INTRODUCCIÓN

La **estadística** es también conocida como la **ciencia de los datos** o **ciencia estadística** y comprende un conjunto de métodos que permiten recolectar, organizar y procesar datos, para después analizar los resultados obtenidos transformándolos en información, como un apoyo eficiente a la toma de decisiones.



- **Unidad de análisis (UA):** entidad que proporciona un dato.
- **Variable estadística:** Es cualquier característica o atributo que se desea estudiar acerca de las UAs y que se espera varíe de una UA a otra. Comúnmente, éstas se denotan con letras mayúsculas (X, Y, Z, W, etc).
- **Recorrido de una variable:** Es el conjunto de todos los posibles valores de una variable. Comúnmente, se identifican con el símbolo  $R_X$ .
- **Medición u observación:** es el proceso que asigna un valor de la variable a una UA.
- **Dato estadístico:** Es la observación, conteo o medición de una variable estadística sobre una UA.

# CONCEPTOS ESTADÍSTICOS

- **Población:** es el universo o conjunto total de UAs. Es sobre este conjunto para quien se obtendrán las conclusiones finales.
- **Muestra:** es cualquier subconjunto de la población.
- **Muestra aleatoria:** es un subconjunto de la población que posee un grado de representatividad adecuado.
- **Error de medición:** es la diferencia entre el valor exacto que posee la UA y el valor obtenido en la medición.





- **Censo:** estudio que involucra a toda la población. No contiene margen de error (o error de precisión), pero es costoso.
- **Muestreo:** estudio que involucra a una muestra. Contiene error de muestreo, pero es menos costoso.
- **Sesgo:** componente de un estudio que hace que la muestra no sea representativa, pues distorsiona las conclusiones del estudio e impide que se lleve a cabo un proceso inferencial.
- **Parámetros  $\theta$ :** información relacionada a la población y que es de interés para el investigador. Los más usados son la media poblacional  $\mu$ , la varianza poblacional  $\sigma^2$ , la proporción poblacional  $p$  y el total poblacional  $\tau$ .
- **Estimadores  $\hat{\theta}$ :** aproximaciones de los parámetros  $\theta$  basadas en una muestra aleatoria.
- **Estadísticas:** medidas que describen la muestra y sirven para estimar los parámetros poblacionales. Las más conocidas son la media y la desviación estándar.

Los métodos estadísticos pueden clasificarse en dos:

- ① Descriptivos.
- ② Inferenciales.

## **Métodos descriptivos**

Se preocupan de describir el conjuntos de datos. Generalmente están relacionados al análisis preliminar o exploratorio de los datos. En el caso de un censo son los únicos métodos que se pueden aplicar.

## **Métodos inferenciales**

Son los encargados de llevar adecuadamente los resultados de una muestra aleatoria hacia la población objetivo. Estos se dividen en métodos inferenciales paramétricos y no-paramétricos.



## **Variable estadística**

Es cualquier característica o atributo que se desea estudiar acerca de las UAs y que se espera varíe de una UA a otra. Comúnmente, éstas se denotan con la letra  $X$ .

## **Recorrido de una variable**

Es el conjunto de todos los posibles valores de una variable. Comúnmente, se identifican con el símbolo  $R_X$ .

## **Dato estadístico**

Es la observación, conteo o medición de una variable estadística sobre una UA.

Ejemplos de variables estadísticas:

- X: Marca de un auto;  $R_X = \{\text{Chevrolet, Toyota, Suzuki, Mazda}\}$
- Y: Grado de lealtad a una tienda;  $R_Y = \{\text{bajo, medio, alto}\}$
- Z: Número de hijos;  $R_Z = \{0, 1, 2, 3, \dots\}$
- W: Precio de un producto;  $R_W = ]0, \infty[$

Las variables estadísticas pueden clasificarse en dos:

- ① **Cualitativas:** Representan un atributo o una cualidad y los datos generados para esta variables son producto de observaciones.
- ② **Cuantitativas:** Representan una cantidad y los datos generados para esta variables son producto de mediciones.

$$\text{Variable} \left\{ \begin{array}{l} \text{Cualitativa o categórica (cualidad o atributo)} \\ \text{Cuantitativa o numérica (cantidad)} \left\{ \begin{array}{l} \text{Discreta (contable –conteo–)} \\ \text{Continua (no contable –medición–)} \end{array} \right. \end{array} \right.$$

Según el esquema anterior, si el dato de una variable cuantitativa se generó a través de un proceso de conteo entonces se dice que la variable cuantitativa es discreta, en caso contrario la variable cuantitativa es continua y admite valores decimales.

# ESCALAS DE MEDICIÓN I

- La **escala nominal** se usa para variables cualitativas en donde no es posible establecer un orden entre ellos. Ejemplo: marca de un auto,  $R_X = \{\text{Chevrolet, Toyota, Suzuki, Mazda}\}$ .
- La **escala ordinal** es usada en el mismo sentido que la nominal pero ahora sí es posible establecer un orden, pero no distancias. Ejemplo: grado de lealtad a una tienda  $R_Y = \{\text{bajo (1), medio (2), alto (3)}\}$ .

# VARIABLES Y ESCALAS

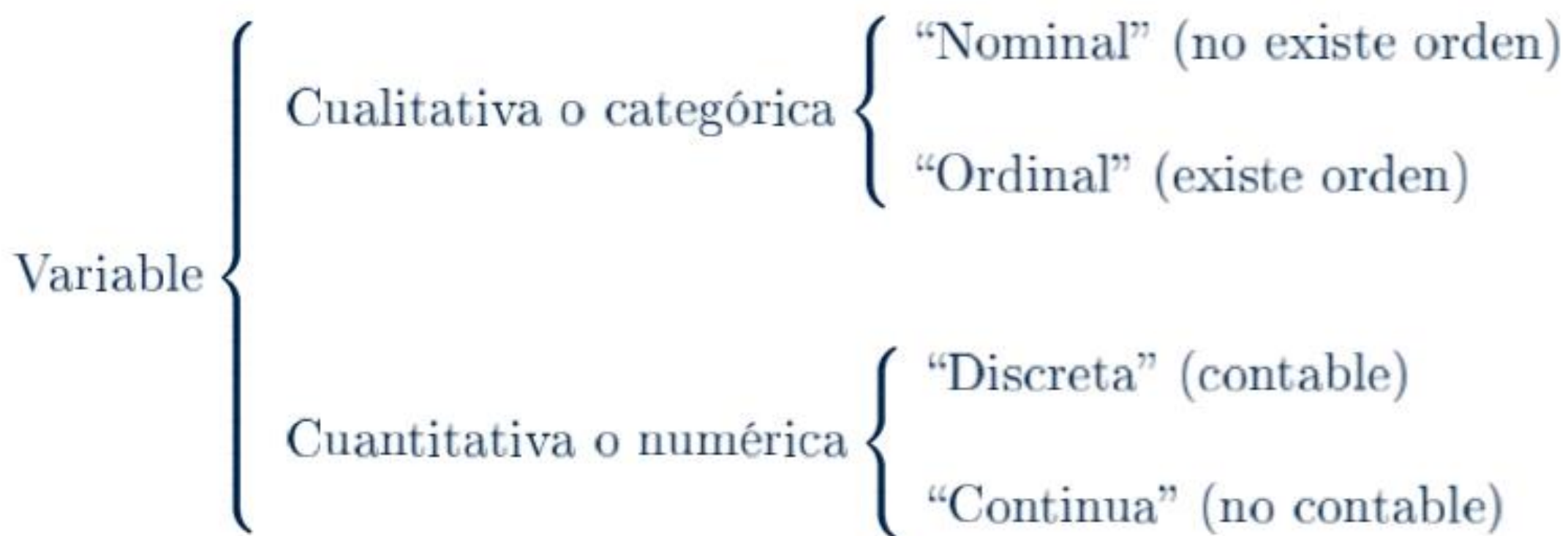
Ejemplos de clasificación de variables estadísticas y escalas de medición:

Variable	Recorrido	Tipo
Marca de un auto	{Chevrolet, Toyota, Suzuki, Mazda}	Cualitativa/Nominal
Grado de lealtad	{bajo, medio, alto}	Cualitativa/Ordinal
Número de hijos	{0, 1, 2, 3, ...}	Cuantitativa/Discreta
Precio de un producto	$[0, \infty[$	Cuantitativa/Continua



## RESUMEN: VARIABLES Y ESCALAS

Una vez reconocidas las escalas de medición, es posible, haciendo uso de éstas, reclasificar los tipos de variables, quedando de la siguiente manera:





## EJERCICIOS:

En las siguientes situaciones identifique la población, UA, variable y su recorrido. Además clasifique la variable según su tipo (cualitativa o cuantitativa discreta o continua) y escala (nominal u ordinal).

- En la carrera *ICI* interesa analizar el rendimiento académico de sus estudiantes en insuficiente, suficiente, bueno y muy bueno.
- Una entidad regional del gobierno de Chile está interesado en cuantificar el número de niños Haitianos de la región del Maule.
- En la Facultad de Ciencias Básicas interesa conocer el puntaje promedio en la prueba de CI de sus estudiantes.
- En la UCM interesa conocer la proporción de estudiantes a favor de la donación de órganos.
- En la carrera *ICI* interesa analizar el peso de sus estudiantes en kilogramos.