

# Probabilidad y Estadística Fundamental

Estadística descriptiva - Clasificación de variables y escalas de medición

Profesor: Nicolás López

Universidad Nacional de Colombia



# Contenido

Repaso definiciones iniciales

Clasificación de las variables por escala

Clasificación de las variables por dimensión

Clasificación de las variables por clase

Conclusión

Ejemplo de aplicación



# Contenido

Repaso definiciones iniciales

Clasificación de las variables por escala

Clasificación de las variables por dimensión

Clasificación de las variables por clase

Conclusión

Ejemplo de aplicación



# Repaso definiciones iniciales

## Recordando

- ▶ Población y muestra.
- ▶ Unidad experimental.
- ▶ Diferencia entre **variable** y **dato** → ???



# Repaso definiciones iniciales

## Recordando

- ▶ Población y muestra.
- ▶ Unidad experimental.
- ▶ Diferencia entre **variable** y **dato** → **medición**



# Repaso definiciones iniciales

## Medición

Es un proceso sistemático de asignación de un valor a las características de interés de las UE. Esta medición puede ser bastante concreta (como la estatura en cms de una persona) ó bastante abstracta (como inteligencia, ó algún otro *constructo*).



# Contenido

Repaso definiciones iniciales

Clasificación de las variables por escala

Clasificación de las variables por dimensión

Clasificación de las variables por clase

Conclusión

Ejemplo de aplicación



# Clasificación de las variables

## Tipos de datos según su escala de medida

Al observar una variable, podemos encontrar una diferenciación por la "forma" en la cual la medimos. Esta clasificación induce un mayor nivel de **sofisticación** en la medición:

- 1 Nominal. Ej: Sexo al nacer  
Relación de igualdad - desigualdad
- 2 Ordinal. Ej: Escolaridad  
Relación de igualdad - desigualdad + Orden
- 3 Intervalo. Ej: Puntaje IQ, temperatura ( $^{\circ}\text{C}$ ,  $^{\circ}\text{F}$ )  
Relación de igualdad - desigualdad + Orden + Cero relativo
- 4 Razón. Ej: Ingresos mensuales, temperatura ( $^{\circ}\text{Kelvin}$ )  
Relación de igualdad - desigualdad + Orden + Cero absoluto





# Clasificación de las variables

Tipos de datos según su **escala de medida**. Ejemplo detallado 1

Para la temperatura en grados centígrados:

- + Se puede determinar si dos temperaturas son iguales ó no:  $10^{\circ}\text{C}$  es igual a  $10^{\circ}\text{C}$ , pero diferente a  $20^{\circ}\text{C}$ .
- + Se puede determinar si dos temperaturas son mayores o menores entre si:  $10^{\circ}\text{C}$  mayor a  $0^{\circ}\text{C}$ , pero menor a  $20^{\circ}\text{C}$  .



# Clasificación de las variables

Tipos de datos según su **escala de medida**. Ejemplo detallado 1

Adicionalmente, iguales intervalos implican iguales diferencias, por ejemplo, una diferencia de  $10^{\circ}\text{C}$  tiene el mismo significado en cualquier punto de la escala:

- + La diferencia entre  $0^{\circ}\text{C}$  y  $10^{\circ}\text{C}$  es igual a la diferencia entre  $10^{\circ}\text{C}$  y  $20^{\circ}\text{C}$ .
- Sin embargo,  $0^{\circ}\text{C}$ , no implica ausencia de temperatura.

Con lo cual, la temperatura en  $^{\circ}\text{C}$  cumple con las características de una escala de intervalo, pero no de razón.



# Clasificación de las variables

Tipos de datos según su **escala de medida**. Ejemplo detallado 2

Figura 1: Tomado de *The Fundamental Process of Measurement*

Athlete	True Time	Watch V (Nominal)	Watch W (Ordinal)	Watch X (Ordinal)	Watch Y (Interval)	Watch Z (Ratio)
A	10	23	11	2	21	20
B	11	12	14	3	23	22
C	13	20	15	4	27	26
D	20	19	18	5	41	40
E	13	20	15	4	27	26
S	0	26	9	1	1	0

*A group of athletes are trying out for the track team at the university, and they are all being timed in the 100 meter dash. Several coaches record times for each athlete. However, they use some very unusual stopwatches. The stopwatches are at several different levels of measurement*



# Clasificación de las variables

Tipos de datos según su **escala de medida**. Ejemplo detallado 2

## Notas

- ▶ Las escalas de medida tienen una estructura jerárquica, es decir, están ordenadas: cada una posee sus propiedades junto a las propiedades de las anteriores escalas.
- ▶ ¿Qué pasará si hago el mismo análisis de la tabla, intercambiando la primera columna por la última?



# Clasificación de las variables

Tipos de datos según su **escala de medida**. Ejemplo detallado 2

Da lo mismo!!!... es decir que cualquiera de estas dos columnas puede considerarse como aquella denotada como "true times".

- ▶ Puede que la primera columna tenga más sentido que la última, pues da el **tiempo en segundos** para cada atleta, una **unidad de medida** que hemos usado toda la vida. Sin embargo, ambas columnas dan la misma información.
- ▶ Si los inventores del reloj hubieran definido un segundo como la mitad de tiempo que dura actualmente, la última columna sería **más familiar** para nosotros.
- ▶ Parece arbitrario porque lo es. De la misma manera que medir nuestra estatura, ya sea en centímetros o en pulgadas, ambas son **equivalentes**.



# Contenido

Repaso definiciones iniciales

Clasificación de las variables por escala

Clasificación de las variables por dimensión

Clasificación de las variables por clase

Conclusión

Ejemplo de aplicación



# Clasificación de las variables

## Tipos de datos según su **dimensión**

Resultan **datos univariados** cuando se mide una sola variable en una sola unidad experimental. Resultan **datos bivariados** cuando se miden dos variables en una sola unidad experimental. Resultan **datos multivariados** cuando se miden más de dos variables.



# Contenido

Repaso definiciones iniciales

Clasificación de las variables por escala

Clasificación de las variables por dimensión

Clasificación de las variables por clase

Conclusión

Ejemplo de aplicación

Nominal



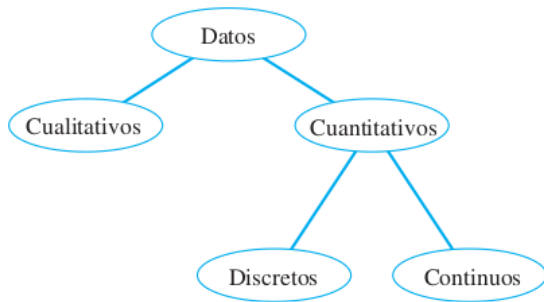


## Clasificación de las variables

## Tipos de datos según su clase

Al observar una variable, podemos clasificar esta observación en las siguientes clases:

Figura 2: Tomado de Introduccion a la Probabilidad y Estadística



# Contenido

Repaso definiciones iniciales

Clasificación de las variables por escala

Clasificación de las variables por dimensión

Clasificación de las variables por clase

Conclusión

Ejemplo de aplicación



# Conclusión

- ▶ Hay tres principales clasificaciones de las variables en estadística: dimensión, clase y escala.
- ▶ Es importante saber la tipología de las variables de estudio para determinar las técnicas estadísticas apropiadas a ser utilizadas.
- ▶ Tener clara la clasificación de las variables de interés determina la manera más clara y óptima de presentar los resultados a la audiencia.



# Ejemplo de aplicación

## Diccionario de variables

Se busca recolectar la información de una muestra de trabajadores. Las características de interés son listadas en un **diccionario de variables**, el cual denota las características de escala, dimensión y clase para cada variable del conjunto de datos. Las variables de interés son:

- ▶ Máximo nivel de escolaridad alcanzado.
- ▶ Edad en años cumplidos.

¿Cuál es la clasificación de cada una de las variables por escala, dimensión y clase? ¿cómo registraría esta información de manera organizada?



# Contenido

Repaso definiciones iniciales

Clasificación de las variables por escala

Clasificación de las variables por dimensión

Clasificación de las variables por clase

Conclusión

Ejemplo de aplicación

