



# **Unidad 1: ESTADISTICA DESCRIPTIVA**

## **Distribuciones de frecuencias**

**Probabilidad y Estadística**  
**Año 2020**

# Manejo de Datos

**Lote de datos:** Es una serie de mediciones de una o más características de interés que llamaremos variables, estas variables se dividen en dos tipos:

- Variables Cualitativas o Categóricas
- Variables Cuantitativas

# Variables Cualitativas o Categóricas

- Son las que describen cualidades o atributos, pero no toman valores numéricos.  
Ejemplos:

- a) “Tamaño de empresas” . Categorías: {pequeñas, medianas, grandes}
- b) “Causas de falla de una máquina”. Categorías: {Fluctuaciones de corriente, Error del operador, Engranajes desgastados, etc.}
- c) Nacionalidad
- d) “Tipos de medios de transporte según distancia que recorren”. Categorías: {corta, media, larga}

- Se puede observar que en el ejemplo a) y d) las categorías tienen un orden natural, en este caso se dice que la variable tiene **Escala Ordinal**
- Si las categorías no tienen ningún orden, como en el ejemplo b) y c), se dice que la variable tiene **Escala Nominal**

# Variables Cuantitativas

Son aquellas que toman valores numéricos.

Ejemplos:

a) “Cantidad de artículos fabricados en un día por una empresa”

Valores que puede tomar:  $\{0, 1, 2, \dots, n\}$

b) “Longitud de un tornillo en mm.”

Valores que puede tomar:  $(0, \infty)$

- Una variable que puede tomar un número finito de valores o infinito numerable, como en el ejemplo a), se llama variable **Discreta**. Por lo general las variables discretas toman valores enteros.
- Una variable que puede tomar valores en todo un intervalo se llama variable **Continua**, como es el caso del ejemplo b)

# Distribuciones de Frecuencia

- La distribución de frecuencias es una tabla donde se presenta la información resumida de la siguiente forma: en una columna están las clases (o categorías) y en otra columna la cantidad de datos en cada clase, es decir que en lugar de conocer el valor exacto de cada dato sólo conocemos a qué clase pertenece.
- La tabla sacrifica parte de la información a cambio de obtener mayor legibilidad.

# Distribuciones de Frecuencia

- **Primer caso:** Supongamos que se cuenta con datos de una *variable Cualitativa* ( *nominal u ordinal* )
- En este caso los datos están clasificados en categorías

## Ejemplo : Variable Cualitativa

Los accidentes en una planta fabril se clasifican de acuerdo a la zona del daño. Los datos son: manos, ojos, manos, brazos, manos, ojos, ojos, ojos, piernas, pie, pie, brazos, piernas, brazos, manos, ojos.

$n = 16$

# Distribuciones de Frecuencia

## Ejemplo : Variable Cualitativa

| Zona del daño | Cantidad de Accidentes |
|---------------|------------------------|
| Manos         | = 4                    |
| Ojos          | = 5                    |
| Brazos        | = 3                    |
| Piernas       | = 2                    |
| Pie           | = 2                    |

***Accidentes en el mes de Noviembre de 2000. Planta I.***

| Zona del daño | Cantidad de Accidentes |
|---------------|------------------------|
| Manos         | 4 (25%)                |
| Ojos          | 5 (31%)                |
| Brazos        | 3 (19%)                |
| Piernas       | 2 (13%)                |
| Pie           | 2 (13%)                |
| Total         | 16                     |

# Distribuciones de Frecuencia

- **Segundo caso:** Datos provenientes de una *variable Cuantitativa discreta con pocos valores*

- **Ejemplo: Variable Cuantitativa Discreta**

Número de accidentes automovilísticos en 10 intersecciones céntricas de una ciudad durante un fin de semana de Diciembre.

|   |   |   |   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 2 | 3 | 0 | 1 | 4 | 1 | 0 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|---|



# Distribuciones de Frecuencia

- **Ejemplo: Variable Cuantitativa Discreta**

| Nº de Accidentes | Cantidad de intersecciones |
|------------------|----------------------------|
| 0                | = 3                        |
| 1                | = 2                        |
| 2                | = 2                        |
| 3                | = 2                        |
| 4                | = 1                        |

***Nº de Accidentes automovilísticos en 10 intersecciones céntricas de una ciudad. Diciembre de 2000.***

| Nº de Accidentes | Cantidad de Intersecciones |
|------------------|----------------------------|
| 0                | 3 ( 30%)                   |
| 1                | 2 ( 20%)                   |
| 2                | 2 ( 20%)                   |
| 3                | 2 ( 20%)                   |
| 4                | 1 ( 10%)                   |
| Total            | 10 (100%)                  |

# Distribuciones de Frecuencia

- **Tercer Caso:** Observaciones de una *variable Cuantitativa Continua o Discreta con muchos valores*
- En este caso existe la necesidad de construir las clases
- Para construir las clases, se subdivide el rango de datos en subintervalos, según la siguiente regla

# Distribuciones de Frecuencia

## Regla

- Selección del número de clases  $n^{\circ} \text{ de clases} = 1.75\sqrt[3]{n}$   
donde n es la cantidad de datos (tamaño muestral)
- Selección de la amplitud de cada clase  $Amp \text{ de clase} = \frac{\text{max} - \text{min}}{n^{\circ} \text{ de clases}}$
- Elegir intervalos semiabiertos y disjuntos que cubran **todo el rango de datos**
- Elegir los limites de los intervalos, en lo posible, enteros y múltiplos de 5.
- Elegir los limites de los intervalos, de manera tal que no haya clases vacías, en lo posible

# Distribuciones de Frecuencia

- Tercer Caso:
- ***Ejemplo : Variable Cuantitativa Continua***

Los siguientes datos son los tiempos de procesamiento por computadora, en segundos, de ciertas solicitudes:

|       |       |       |       |
|-------|-------|-------|-------|
| 10.53 | 12.24 | 15.82 | 25.35 |
| 28.21 | 20.55 | 22.83 | 18.51 |
| 19.88 | 14.35 | 15.41 | 28.56 |
| 24.20 | 23.32 | 19.53 | 18.96 |
| 25.62 | 27.20 | 30.43 | 26.31 |

Tiempos en seg. Redondeados:

|    |    |    |    |
|----|----|----|----|
| 11 | 12 | 16 | 25 |
| 28 | 21 | 23 | 19 |
| 20 | 14 | 15 | 29 |
| 24 | 23 | 20 | 19 |
| 26 | 27 | 30 | 26 |

# Distribuciones de Frecuencia

- Ejemplo : Variable Cuantitativa Continua***

\*  $n = 20$

\* nº de clases =  $1.75 \sqrt[3]{n} = 4.75 \cong 5$

\* amplitud de clase =  $\frac{30 - 11}{5} = 3.8 \cong 4$

| Clases  | Marcas |
|---------|--------|
| (10,14] | = 3    |
| (14,18] | = 2    |
| (18,22] | = 5    |
| (22,26] | = 4    |
| (26,30] | = 6    |
| Total   | 20     |

| Clases  | Marcas |
|---------|--------|
| [10,15) | = 3    |
| [15,20) | = 4    |
| [20,25) | = 6    |
| [25,30) | = 6    |
| [30,35) | = 1    |
| Total   | 20     |

# Distribuciones de Frecuencia

- ***Ejemplo : Variable Cuantitativa Continua***

*Distribución de frecuencias:*

***Tiempos de Procesamiento por PC de solicitudes. Febrero 2001.***

| Tiempos en seg. | Cantidad de Solicitudes | %   |
|-----------------|-------------------------|-----|
| [10,15)         | 3                       | 15  |
| [15,20)         | 4                       | 20  |
| [20,25)         | 6                       | 30  |
| [25,30)         | 6                       | 30  |
| [30,35)         | 1                       | 5   |
| Total           | 20                      | 100 |