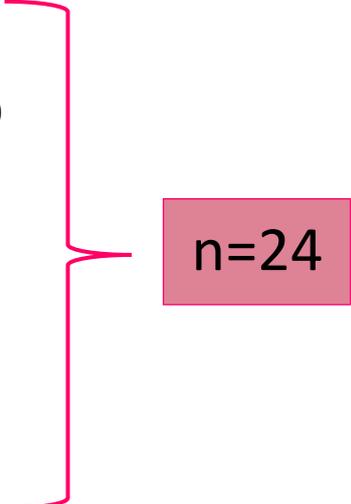


## DISTRIBUCIONES DE FRECUENCIA

Cuando el tamaño muestral ( $n$ ) o el tamaño poblacional ( $N$ ) es muy grande, se dificulta extraer información de la serie de datos no agrupados, denominados datos crudos o raw data.

Los datos crudos suelen tener habitualmente la unidad experimental y el valor que toma esa unidad experimental, por ejemplo:

$x_1 = 10$	$x_8 = 13$	$x_{15} = 7$	$x_{22} = 11$	
$x_2 = 8$	$x_9 = 8$	$x_{16} = 7$	$x_{23} = 10$	
$x_3 = 14$	$x_{10} = 12$	$x_{17} = 10$	$x_{24} = 12$	
$x_4 = 13$	$x_{11} = 13$	$x_{18} = 11$		
$x_5 = 7$	$x_{12} = 10$	$x_{19} = 11$		
$x_6 = 9$	$x_{13} = 9$	$x_{20} = 9$		
$x_7 = 8$	$x_{14} = 9$	$x_{21} = 7$		

Para que los datos sean útiles, necesitamos organizar nuestras observaciones de modo que podamos distinguir patrones y llegar a conclusiones lógicas, es por eso que, en esos casos (por convención  $n > 20$ , pero puede variar según la bibliografía), lo conveniente es realizar un **Agrupamiento en distribuciones de frecuencia**.

### **¿Qué hace una distribución de frecuencia?**

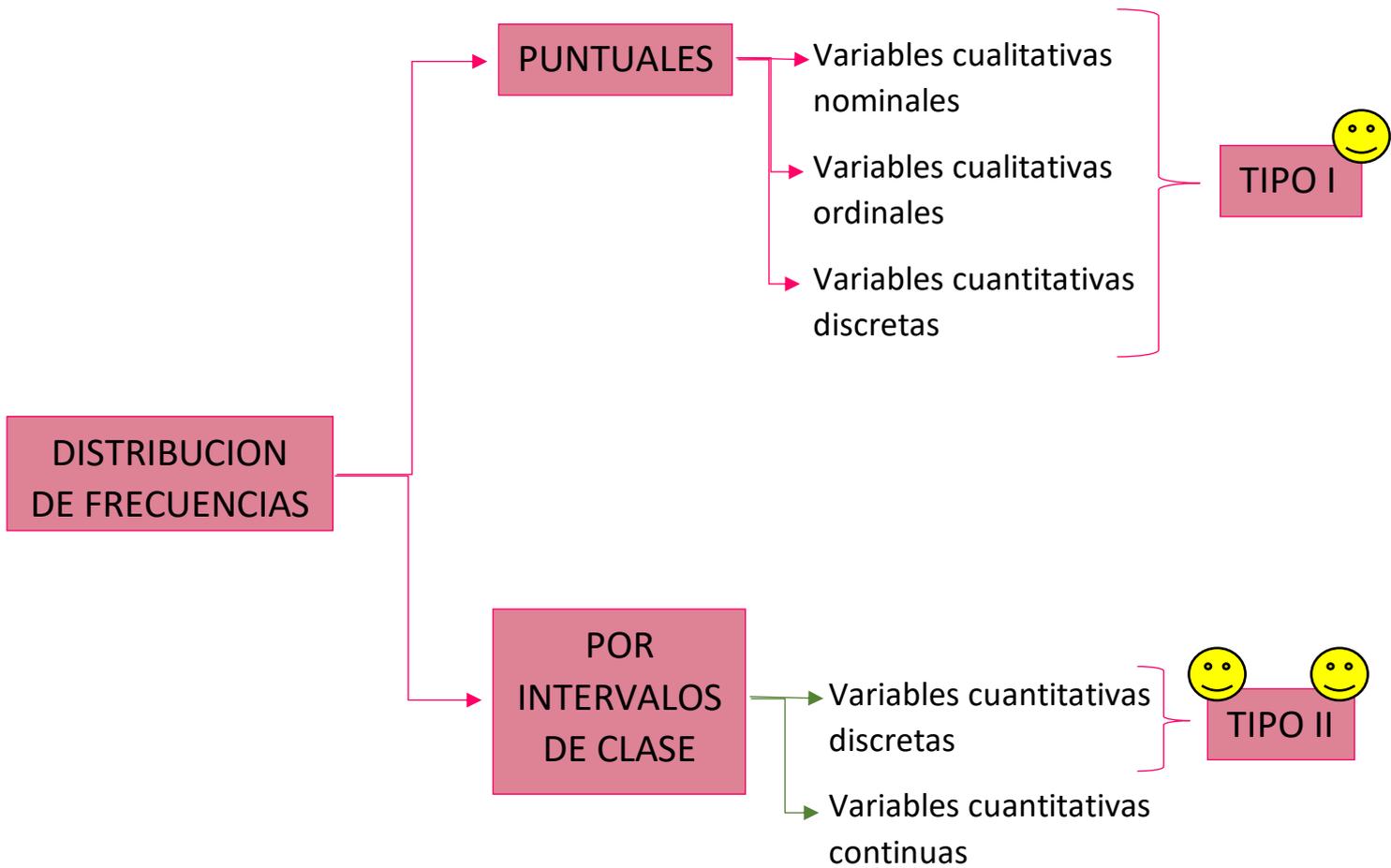
Una distribución de frecuencias *resume valores numéricos y no numéricos organizándolos en un conjunto de **clases***.

Dicho de otra forma, una distribución de frecuencias *muestra el número de observaciones del conjunto de datos que caen en cada una de las clases*. Si se puede determinar la frecuencia con la que se presentan los valores en cada clase de un conjunto de datos, se puede construir una distribución de frecuencia.

### **¿Qué debemos tener en cuenta a la hora de realizar distribución de frecuencia?**

Se debe tener en cuenta:

- **Tamaño de la muestra o de la población:** Debido a que si sus tamaños son pequeños no tiene sentido estadístico realizar una distribución de frecuencias.
- **Qué tipo de variable se está estudiando:** El tipo de frecuencia que se pueda llegar a realizar dependerá del tipo de datos con los que estamos trabajando, es decir, no será lo mismo si estamos agrupando por ejemplo datos cuantitativos continuos que datos cualitativos.
- **El cardinal del espacio muestral:** Su tamaño nos indicara, también, que tipo de distribución podemos utilizar y si, en algunos casos, debemos reconfigurar el experimento.



 TIPO I →  $\#\Omega \leq 15$   
 TIPO II →  $\#\Omega > 15$

ACLARACION: En el caso de que se esté trabajando con **variables cualitativas** (ya sean ordinales o nominales) cuyo cardinal del espacio muestral sea **mayor a 15**, se deberá **reconfigurar el experimento**, a fines de disminuir dicho  $\#\Omega$  y, de esa forma, poder realizar una distribución de frecuencia puntual.

## **TIPOS DE DISTRIBUCIONES DE FRECUENCIA**

### **Frecuencia Absoluta:**

Se trata de contar la cantidad de veces que se repite un valor en una variable. Al sumar las frecuencias absolutas se obtiene un resultado idéntico a la cantidad total de datos (n).

$$\sum_{i=1}^{K=10} f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_7 + f_8 + f_9 + f_{10} = n$$

Ejemplo:

Un alumno secundario, a lo largo de su ciclo lectivo, se saca en las evaluaciones diferentes notas: 5, 8, 6, 10, 7, 9, 10, 8, 7.50, 8, 7. Frente a estos valores se puede decir que la frecuencia absoluta de 8 es 3, debido a que se repite en 3 oportunidades.

El número de veces que aparece un valor se representa con  $f_i$ , donde el subíndice  $i$  representa cada uno de los valores.

### **Frecuencia Relativa:**

Te da la proporción de la clase respecto al total observado.

La proporción o frecuencia relativa es igual al número de datos en cada clase dividido entre el número total de datos:

$$fr_i = \frac{f_i}{n}$$

Ejemplo: Si existen 80 datos y la frecuencia de cierta clase es de 20, la proporción de datos en esa clase es:  $20/80 = 0,25$

Cuando se comparan dos o más grupos, es más útil conocer la proporción o el porcentaje del total de cada grupo que conocer el conteo de frecuencia de cada uno. En estos casos se crea una *distribución de frecuencia relativa* o una *distribución porcentual*.

Si los dos o más grupos tienen muestras de tamaños diferentes se debe utilizar una distribución de frecuencia relativa o una distribución porcentual.

Frecuencia Porcentual:

Se obtiene multiplicando la proporción (frecuencia relativa) en cada clase por 100%. Siguiendo el ejemplo anterior, el porcentaje sería  $0.25 \times 100\% = 25\%$

Frecuencia Absoluta o Porcentual Acumulada:

Es la cantidad de veces que se repite una frecuencia absoluta, o porcentual, simple hasta cierto tipo de clase. Es decir, es la suma de las frecuencias simples y se simboliza  $F_i$ .

Solo podemos realizar frecuencias acumuladas con datos numéricos. Con datos cualitativos nominales no tiene sentido realizar una frecuencia acumulada, mientras que, con datos cualitativos ordinales existe una discusión al respecto, pero en el curso no los acumularemos. Dicho de otra forma, para cualquier tipo de datos categóricos solo se pueden realizar distribuciones de frecuencia simples.

Ejemplo aplicando todas las frecuencias:

Consideremos las notas de los exámenes finales de la Materia Estadística de un curso hipotético de la universidad de Buenos Aires

# $\Omega_x = 10$

N = 50

1	4	5	6	8
2	4	5	6	8
2	4	5	7	8
3	4	5	7	8
3	4	5	7	8
3	5	5	7	8
4	5	6	7	9
4	5	6	7	9
4	5	6	8	9
4	5	6	8	10

Con la información de los “datos de notas” podemos armar la distribución de frecuencias.

$X_i$	$f_i$	$f_{ri}$	$f_{ri} \times 100\%$	$F_i$
1	1	0,02	2%	1
2	2	0,04	4%	3
3	3	0,06	6%	6
4	9	0,18	18%	15
5	11	0,22	22%	26
6	6	0,12	12%	32
7	6	0,12	12%	38
8	8	0,16	16%	46
9	3	0,06	6%	49
10	1	0,02	2%	50