	<b>INSTITUCIÓN EDUCATIVA JULIA RESTREPO – TULUÁ</b> <b>AÑO 2024</b>		
	<b>Departamento de matemáticas</b>	<b>PERIODO: Uno</b>	
	<b>Guía No 1: Distribución de frecuencias para datos agrupados</b>		<b>Grado: 8º</b>
	<b>Asignatura: Estadística</b>		<b>Docente: Emerson Restrepo P</b>

### Indicadores de desempeño:

- Utilizar distribuciones de frecuencias y algunos conceptos básicos de estadística para ordenar, interpretar y analizar datos.
- Interpreta los datos representados en diferentes tablas de frecuencias.

# DISTRIBUCIÓN DE FRECUENCIAS PARA DATOS AGRUPADOS

## MOMENTO UNO: EXPLORACION

Para presentar el reporte, Carlos puede realizar una tabla como la 1 en la cual agrupe los libros de acuerdo con su peso en intervalos disjuntos de igual amplitud.

### Explora

En una biblioteca se van a reubicar los libros en cinco estantes, de manera que los más pesados se ubiquen en los estantes inferiores y los más livianos en la parte superior.

Carlos, el encargado, debe presentar un reporte que permita conocer la cantidad de libros que hay en la biblioteca de acuerdo con su peso.

- ¿Cómo puede presentar el reporte?

Peso de los libros (en gramos)	Número de libros
[100, 200)	40
[200, 300)	65
[300, 400)	45
[400, 500)	35
[500, 600)	35
[600, 700)	38

Tabla 1



De la tabla se puede inferir que hay 258 libros en la biblioteca, que los más livianos pesan entre 100 g y 200 g, y los más pesados entre 600 g y 700 g.

## Distribución de clases o datos agrupados

La distribución de frecuencias agrupadas o tabla con datos agrupados se emplea si las variables toman un número grande de valores o la variable es continua.

Se agrupan los valores en intervalos que tengan la misma amplitud denominados clases. A cada clase se le asigna su frecuencia correspondiente.

La tabla de distribución de frecuencias está formada por los siguientes elementos:

- ⌘ **Intervalos de clase:** se ubican en la primera columna de la tabla. En cada intervalo se incluye el grupo de datos que está entre el valor donde inicia (límite inferior) y el valor en el que termina (límite superior). Tienen la particularidad de ser disjuntos y ordenados, lo cual implica que cada dato queda incluido en uno y solo un intervalo.
- ⌘ **Frecuencia del intervalo:** número de individuos o datos que están clasificados en cada intervalo. Se representa con la letra  $f$ .
- ⌘ **Frecuencia relativa:** relación entre la frecuencia y el total de la muestra o población; se representa como  $f_r$ . Esta relación se expresa mediante el cociente  $\frac{f}{n}$  en donde  $n$  es el número de datos. Al multiplicar  $f_r$  por 100 se obtiene el porcentaje de los datos que están en dicha clase.
- ⌘ **Frecuencia acumulada:** es la sumatoria del número de individuos que están en los intervalos anteriores y la frecuencia del intervalo. Se representa por  $F$ .
- ⌘ **Frecuencia relativa acumulada:** se representa por  $F_r$  y se plantea como la relación entre la frecuencia acumulada y el total de la muestra, es decir,  $F_r = \frac{F}{n}$ .
- ⌘ **Marca de clase:** es el punto medio de cada intervalo y se considera como el dato más representativo de dicho intervalo. Se representa por  $M_c$ .

Con el siguiente ejemplo vamos a explicar algunos conceptos que deben tener en cuenta

## Ejemplo

En una biblioteca se van a reubicar los libros en cinco estantes, de manera que los más pesados se ubiquen en los estantes inferiores y los más livianos en la parte superior. Carlos, el encargado, debe presentar un reporte que permita conocer la cantidad de libros que hay en la biblioteca de acuerdo con su peso. ¿Cómo puede presentar el reporte?



## Conceptos

**Intervalo de clase:** son divisiones o categorías en las cuales se agrupan un conjunto de datos ordenados con características comunes. En cada intervalo se incluye el grupo de datos que está entre el valor donde inicia (límite inferior -  $L_i$ ) y el valor en el que termina (límite superior -  $L_s$ ).

Peso de los libros (en gramos)	Número de libros
[100, 200)	40
[200, 300)	65
[300, 400)	45
[400, 500)	35
[500, 600)	35
[600, 700)	38



### Ten en cuenta

Cada uno de los intervalos en los que se agrupan los datos incluye el extremo izquierdo pero no el derecho. Por ejemplo, el intervalo  $[9, 11)$  indica que este incluye el valor 9 pero no el valor 11.

**Numero de intervalos (NI):** representa la cantidad de intervalos o clases en la cuales se va resumir la información. En la tabla representa el número de filas que se deben generar

¿Número de intervalos?

Peso de los libros (en gramos)	Número de libros
[100, 200)	40
[200, 300)	65
[300, 400)	45
[400, 500)	35
[500, 600)	35
[600, 700)	38

$$NI = \sqrt{n}$$

Regla de Sturge

$$NI = 1 + 3,322 \log n$$

*n*: numero de datos

**Marca de clase (M<sub>i</sub> o X<sub>i</sub>):** es el punto medio de cada intervalo y se considera como el dato más representativo de dicho intervalo. Se suman los extremos del intervalo y se divide el resultado entre dos.

Peso de los libros (en gramos)	Número de libros
[100, 200)	40
[200, 300)	65
[300, 400)	45
[400, 500)	35
[500, 600)	35
[600, 700)	38

¿Cuál es el punto medio o la mitad?  
R//150

¿Cómo se calcula la marca de clase?

$$M_i = \frac{L_i + L_s}{2}$$

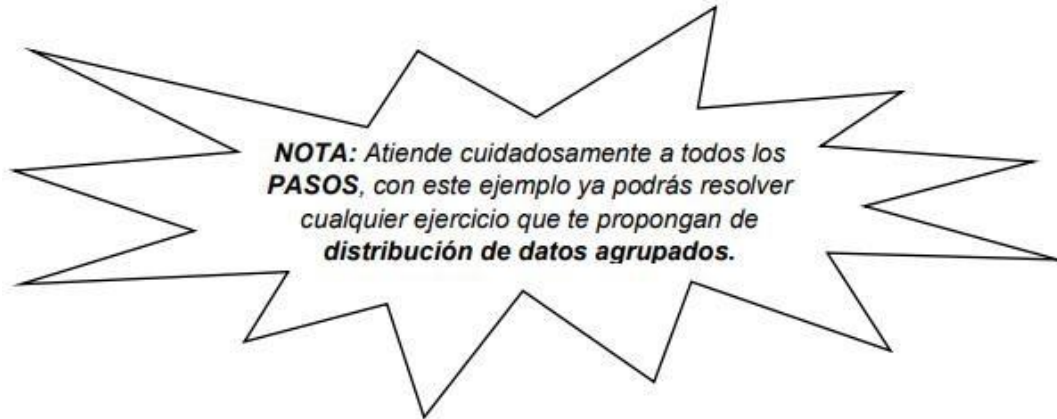
**Rango de la distribución (R):** es la diferencia entre el dato mayor (D<sub>M</sub>) y el dato menor (D<sub>m</sub>) de la base o conjunto de datos.

$$R = D_M - D_m$$

**Amplitud (A):** representa el tamaño de cada intervalo de clase, es decir, el número de valores que concurren en una clase determinada.

$$A = \frac{\text{Rango}}{\text{Numero de intervalos}} = \frac{D_M - D_m}{\sqrt{n}}$$

Para elaborar una tabla de distribución de frecuencias de datos agrupados se realiza el siguiente procedimiento que se explicara con el siguiente ejemplo:



**Ejemplo:**

Para determinar los factores óptimos de crecimiento de un hongo comestible, un equipo de ingenieros de alimentos realizó un experimento que consistió en cultivar dicho hongo en 50 muestras diferentes y observar, después de 60 días, los cuerpos fructíferos que generó cada cultivo.

123	116	167	198	165	148	169	110	121	100
145	132	145	126	176	189	163	101	120	109
135	127	178	187	180	166	134	129	118	102
167	185	183	177	156	145	167	143	132	121
145	128	119	117	140	121	164	129	132	140

**Solución:**

**Paso 1:** Se calcula el número de intervalos:

¿Cuál es el intervalo de clase? R// La variable: Número de cuerpos fructíferos



¿Número de datos? R// 50 datos

¿Cuántos intervalos de clase? R// usamos la fórmula para calcularlos

$$NI = \sqrt{n}$$
$$NI = \sqrt{50} = 7,07 \cong 7$$

**Paso 2:** Se calcula el rango de la distribución, que es la diferencia entre el dato mayor y el dato menor:

$R = D_M - D_m$

123	116	167	198	165	148	169	110	121	100
145	132	145	126	176	189	163	101	120	109
135	127	178	187	180	166	134	129	118	102
167	185	183	177	156	145	167	143	132	121
145	128	119	117	140	121	164	129	132	140

Dato menor ( $D_m$ )  
Primer dato del primer intervalo

$$R = 198 - 100 = 98$$

**Paso 3:** Se calcula el tamaño o amplitud de cada intervalo:

¿Cuál es la amplitud del intervalo? R// usamos la fórmula para calcularlo

$$A = \frac{R}{NI} = \frac{D_M - D_m}{\sqrt{n}}$$
$$A = \frac{98}{7} = 14$$

**PASO 5: Procedemos a construir finalmente la tabla de frecuencia.**

**Construcción de las clases o intervalos**

Se construyen los intervalos. Para ello, se toma el dato menor como límite inferior del primer intervalo y a este valor se le suma el tamaño o amplitud del intervalo para encontrar el límite superior.

Columna - Intervalos de clase (Números de cuerpos fructíferos)

- **Primer intervalo**

Límite inferior= 100

Límite superior=  $Li + A$

Límite superior=  $100 + 14 = 114$

[100,114) →

Número de cuerpos fructíferos
[100,114)
[114,128)

- **Segundo intervalo**

Límite inferior= 114 (donde termina el primer intervalo)

Límite superior=  $Li + A$

Límite superior=  $114 + 14 = 128$

[114,128)

Columna - Marca de clase ( $M_i$ )

- Primera marca de clase

$$M_i = \frac{L_i + L_s}{2}$$

$$M_i = \frac{100 + 114}{2} = 107$$

Número de cuerpos fructíferos	Marca de clase ( $M_i$ )
[100,114)	107
[114,128)	121

- Segunda marca de clase

$$M_i = \frac{L_i + L_s}{2}$$

$$M_i = \frac{114 + 128}{2} = 121$$

Para entender mejor: El intervalo cerrado [,] indica que el número que lo acompaña está incluido en los datos de ese grupo y el intervalo abierto (,) indica que el número que lo acompaña NO está incluido en este intervalo, y hará parte del intervalo que le siga, así:

En el intervalo [2 - 7), los días de ausencia laborales que entran en este intervalo son 2,3 ,4, 5 y 6, el número 7 es el que acompaña al intervalo abierto y este **NO** entra en la agrupación de este intervalo.

Si mi intervalo es [12 - 17), solo contaré cuantos números 12, 13, 14, 15 y 16, hay en la tabla (Son los señalados con amarillo en la tabla, [12 - 17) = 20), el número 17 no lo contaré y no lo incluiré, este hará parte del siguiente intervalo que es [17 - 22).

En el intervalo [27 - 32], es un intervalo cerrado en [27 y es cerrado en 32], por la tanto, todos los números estarán incluidos en el intervalo, es decir los números que contaré son 27, 28, 29, 30, 31 y 32.

**Esto pasa cuando es el último intervalo.**



Número de cuerpos fructíferos	Marca de clase (Mi)	Frecuencia absoluta (f)	Frecuencia relativa (fr)			Frecuencia absoluta acumulada (F)	Frecuencia relativa acumulada (Fr)	
			Fracción	Decimal	Porcentaje		Decimal	Porcentaje
[100,114)	107	5	$\frac{5}{50}$	0,10	10	5	0,1	10
[114,128)	121	11	$\frac{11}{50}$	0,22	22	16	0,32	32
[128,142)	135	10	$\frac{10}{50}$	0,20	20	26	0,52	52
[142,156)	149	6	$\frac{6}{50}$	0,12	12	32	0,64	64
[156,170)	163	9	$\frac{9}{50}$	0,18	18	41	0,82	82
[170,184)	177	5	$\frac{5}{50}$	0,10	10	46	0,92	92
[184,198)	191	4	$\frac{4}{50}$	0,08	8	50	1,00	100

**Frecuencia absoluta (fi):** Corresponde al número de veces que se repite cada dato.

Para calcular la **Frecuencia absoluta acumulada (Fi):**

**Paso 1:** Escribimos la primera frecuencia absoluta (5) en la casilla de frecuencia absoluta acumulada (5).

**Paso 2:** Luego sumamos la primera frecuencia absoluta acumulada con la segunda frecuencia absoluta (5+11 = 16), el resultado lo escribimos al frente en la Frecuencia absoluta acumulada.

**Paso 3:** Seguimos haciendo ese proceso (16+10 = 26) luego (26+6=32) y así hasta llegar a 50, que corresponde a la última frecuencia absoluta acumulada, la acula debe ser igual a la cantidad de datos observados.

Para hallar la **Frecuencia relativa (fr o hi):**

**Paso 1:** Sumamos todas las frecuencias absolutas para saber cuál es el total.

**Paso 2:** Luego de tener el total, tomamos cada frecuencia absoluta y la dividimos entre este total.

**Paso 3:** Esto lo hacemos con cada frecuencia es decir ( $5 / 50 = 0,1$ ) en la siguiente casilla ( $11 / 50 = 0,22$ ), luego la siguiente ( $10 / 50 = 0,2$ ) y así con todas en orden hasta completar la tabla. La frecuencia relativa se puede ver de manera porcentual, solo debes multiplicar el valor que te dio de la división por 100

Para calcular la **Frecuencia relativa acumulada ( $F_r$  o  $H_i$ )**: se obtiene de ir sumando las frecuencias relativas de una clase o grupo de la muestra con la anterior.

**Paso 1:** Escribimos la primera frecuencia relativa (0.1) en la casilla de frecuencia relativa acumulada (0.1).

**Paso 2:** Luego sumamos la primera frecuencia relativa acumulada con la segunda frecuencia relativa ( $0.1+0.22 = 0.32$ ), el resultado lo escribimos al frente en la Frecuencia relativa acumulada.

**Paso 3:** Seguimos haciendo ese proceso ( $0.32+0.2 = 0.52$ ) luego ( $0.52+0.12=0.64$ ) y así hasta llegar a 1.0 aproximadamente, corresponde a la última frecuencia absoluta acumulada. La última casilla de la columna de la frecuencia absoluta acumulado debe ser siempre aproximadamente 1.0.

### Conclusiones de ejemplo

- 11 de los hongos dieron entre 114 y 128 cuerpos fructíferos. Este valor fue el más alto registrado en el experimento y equivale a 22%
- 10 de los hongos dieron entre 28 y 142 cuerpos fructíferos. Este valor equivale a 20%
- El 42% de los hongos produce entre 114 y 142 cuerpos fructíferos.
- Tan solo el 8% de los hongos produce entre 184 y 198 cuerpos fructíferos

### MOMENTO TRES: PRACTICA

1. A 40 estudiantes se les pidió que estimen el número de horas que habrían dedicado a estudiar la semana pasada (tanto en clase como fuera de ella), obteniéndose los siguientes resultados:

36	30	47	60	32	35	40	50
54	35	45	52	48	58	60	38
32	35	56	48	30	55	49	39
58	50	65	35	56	47	37	56
58	50	47	58	55	39	58	45

Elaborar la tabla de distribución de frecuencias y presentar algunas conclusiones relacionadas con la información.

2. En un centro comercial, se consultó la edad a todas las personas que entraban entre las 12:00 h y 12:30 h. Los resultados obtenidos fueron los siguientes:

15	73	1	65	16	3	42
36	42	3	61	19	36	47
30	45	29	73	69	34	23
22	21	33	27	55	58	17
4	17	48	25	36	11	4
54	70	51	3	34	26	10

### MOMENTO CUATRO: TRANSFERENCIA Y VALORACION

3. Observa la siguiente tabla.

Recorrido (km)	f (BUSES)	fr	%	F	Fr	Mi
150-156			12			
157-163	8		32			
164-170						
171-177			28			
178-184			4			

- a) Completa la tabla del recorrido (en kilómetros) de los buses de una empresa  
 b) Responde, ¿Cuántos buses tiene la empresa?
4. Complete los siguientes datos, de acuerdo a las notas obtenidas por un curso en la prueba de matemáticas de 40 estudiantes.

8	8	6	5	3	<u>2</u>	2	5	10	4
9	3	6	6	2	4	7	5	9	3
7	4	7	4	6	3	9	10	6	8
6	5	9	9	8	5	<u>10</u>	9	4	7

Valor máximo \_\_\_\_\_ Valor mínimo \_\_\_\_\_

Rango \_\_\_\_\_ Amplitud para  $k = 4$  \_\_\_\_\_

Completa la tabla de frecuencia para las notas registradas en la tabla agrupadas en 4 intervalos de la tabla anterior.

Intervalos	Frecuencia absoluta (f)	Frecuencia absoluta acumulada (F)	Frecuencia relativa (ni)	Frecuencia relativa porcentual (%)
[2 - 4)	7	7	$7 / 40 = 0,175$	17,5%
[4 - )		17	$10 / 40 = 0,25$	
[ - )				25 %
[8 - ]				

