

1 Datos agrupados.

Cuando hablamos de datos agrupados, la media y la mediana son iguales.

Si contamos con valores muy dispersos, dentro de un conjunto de datos, en vez de unas cuantas repetidos, los podemos redistribuir en una tabla de distribución de frecuencias para datos agrupados.

El procedimiento para redistribuir los valores consiste en crear intervalos numéricos, a los que llamamos intervalo de clase, y contamos cuántos datos se encuentran dentro de cada intervalo de clase.

Cuando llevamos a cabo este proceso, estamos convirtiendo valores discretos en valores continuos. Cabe destacar que el ancho de cada intervalo, marca de clase, respecto al promedio de datos o valores contenidos puede no ser constante, pero para fines de este curso se espera que cada intervalo sea de igual magnitud.

ej. Los siguientes datos corresponden a los aciertos obtenidos por 50 alumnos que realizaron su examen para el ingreso al nivel medio superior.

26	48	67	72	78	83	90	93	106	119
42	49	70	73	79	83	90	95	107	121
42	53	70	74	80	85	92	96	107	122
43	51	70	75	81	87	93	102	110	127
46	64	72	76	81	87	93	105	119	128

Haciendo uso de estos datos, construir una tabla de distribución de frecuencias para datos agrupados, comenzando por las clases 15-29, 30-44, 45-59, ...

		Límites reales		Intervalos de clase*			
# Clase	Intervalos de clase	f	χ_i	LRI - LAS	\bar{f}	% Tro	f · x
1	15-29	1	22	14.5 - 29.5	1	$\frac{100}{50} \cdot 1 = 2\%$	22
2	30-44	3	37	29.5 - 44.5	4	$2 \cdot 4 = 8\%$	111
3	45-59	5	52	44.5 - 59.5	9	$2 \cdot 9 = 18\%$	260
4	60-74	9	67	59.5 - 74.5	18	$2 \cdot 18 = 36\%$	603
5	75-89	12	82	74.5 - 89.5	30	$2 \cdot 30 = 60\%$	984
6	90-104	9	97	89.5 - 104.5	39	$2 \cdot 39 = 78\%$	873
7	105-119	7	112	104.5 - 119.5	46	$2 \cdot 46 = 92\%$	784
8	120-134	4	127	119.5 - 134.5	50	$2 \cdot 50 = 100\%$	508

$$\begin{aligned} \% Tro &= \frac{100 \cdot f}{n} \\ \chi_i &= \frac{LRI_i + LAS_i}{2} \\ LRI_i &= \frac{LIC_i + LSC_{i-1}}{2} \\ LSC_i &= \frac{LIC_i + LIC_{i+1}}{2} \end{aligned}$$

LIC_i : frecuencia acumulada
 LSC_i : frecuencia acumulada
 LRI_i : límite inferior de clase i-ésima
 LSC_i : límite superior de clase i-ésima

LIC_i = límite inferior de clase i-ésima
 LSC_i = límite superior de clase i-ésima

f = frecuencia de clase.
 Los intervalos son criterios de cada persona.
 *Notas: por acuerdo, no deben de traslaparse.

Como se podrá haber notado, Los intervalos de clase utilizados fueron indicados por el propio problema, y están construidos de tal forma que no se traslapen, es decir, que no compartan límites. [15-29, 30-44] Las clases.

Los límites reales representan los valores "reales" entre los cuales se considera la clase, se calcula haciendo uso de los límites superiores e inferiores de la clase previa y posterior a la clase en cuestión, respectivamente.

# Clase	Intervalos de clase	LRI - LRS	
2	30 - 44	29.5 - 44.5	$LSC_2 = 44$ $LIC_2 = 30$
3	45 - 59	44.5 - 59.5	$LSC_3 = 59$ $LIC_3 = 45$
4	60 - 74	59.5 - 74.5	$LSC_4 = 74$ $LIC_4 = 60$

Para calcular el LRI de la clase #3, tomamos el límite superior de clase de la clase anterior [3-1=2], y sumamos el límite inferior de clase de nuestra clase [3], finalmente dividimos entre dos para promediar.

$$LRI_3 = \frac{LSC_2 + LIC_3}{2} ; \begin{matrix} LSC_2 = 44 \\ LIC_3 = 45 \end{matrix}$$

$$LRI_3 = \frac{44 + 45}{2} = 44.5$$

Para calcular el LRS de la clase #3 utilizamos la misma lógica, pero utilizando el LIC de nuestra clase (3) y el LIC de la clase posterior (3+1=4).

$$LRS_3 = \frac{LSC_3 + LIC_4}{2} ; \begin{matrix} LSC_3 = 59 \\ LIC_4 = 60 \end{matrix}$$

$$LRS_3 = \frac{59 + 60}{2} = 59.5$$

Para encontrar el límite real inf. de clase de la clase #1, tomamos el LSC_0 como uno imaginario (no matemáticamente), como si existiera otro intervalo previo al nuestro, lo mismo con el LIC para la última clase pero uno posterior al mismo.

# Clase	Intervalos de clase	LRI - LRS	
0	0 - 14	-	*
1	15 - 29	14.5 - 29.5	
⋮			
8	120 - 134	119.5 - 134.5	
9	135 - 149	-	*

* No se colocan en la tabla, solo están aquí para ilustrar el concepto.