

CAPÍTULO 3

MEDIDAS DE TENDENCIA CENTRAL

A veces, de los datos recolectados ya organizados en alguna de las formas vistas en capítulos anteriores, se desea encontrar una especie de punto central en función de sus frecuencias. En Estadística se conocen tres diferentes, llamadas *medidas de tendencia central*, cuya utilización varía de acuerdo con lo que se desee del conjunto de datos recolectados. Esas tres medidas de tendencia central son *la media*, *la mediana* y *la moda*.

Cada una de ellas se estudiará en dos partes: primero, cuando los datos están organizados en tablas de distribución de frecuencias simple y, segundo, cuando están organizados en intervalos. Además, a veces difieren las fórmulas para calcular alguna de ellas si se trata de poblaciones o de muestras. En caso de que no se diga nada, deberá entenderse que la fórmula es la misma para ambas.

3.1 LA MEDIA

La media, llamada también *media aritmética*, es la medida de tendencia central conocida popularmente como “promedio”.

3.1.1 LA MEDIA PARA FRECUENCIAS SIMPLES

Cuando los datos recolectados han sido organizados en una tabla de distribución de frecuencias simples, la media, para poblaciones como para muestras, se puede calcular por medio de la fórmula

$$\bar{x} = \frac{\sum fx}{n}$$

en donde:

\bar{x} = media o promedio.

$\sum fx$ = suma de las frecuencias por su correspondiente dato nominal.

n = suma de todas las frecuencias (número de datos recolectados).

Para calcular la media, debe añadirse una columna fx a la tabla original en la que se registren los resultados correspondientes al producto de la frecuencia por su valor nominal (fx).

Ejemplo 1: Las calificaciones de Matemáticas de los grupos A y B se muestran en la tabla de la derecha. Calcular el promedio (la media) obtenido por esos grupos.

Solución: Debe añadirse a la tabla original una columna encabezada por fx en donde se anotarán los resultados correspondientes a las multiplicaciones de cada valor nominal x por su frecuencia f respectiva.

Por ejemplo, para la primera fila de la tabla:

$$fx = 2 \times 0 = 0$$

para la segunda fila:

$$fx = 3 \times 1 = 3$$

para la tercera fila:

$$fx = 3 \times 2 = 6$$

La tabla completa con las tres columnas queda como se muestra a la derecha. La suma de los valores de la columna fx es 544, de manera que utilizando la fórmula para el promedio, recordando que n es la suma de todas las f , se obtiene:

calificaciones x	f
0	2
1	3
2	3
3	6
4	8
5	9
6	17
7	22
8	10
9	6
10	5

calificaciones x	f	fx
0	2	0
1	3	3
2	3	6
3	6	18
4	8	32
5	9	45
6	17	102
7	22	154
8	10	80
9	6	54
10	5	50

91 544

$$\bar{x} = \frac{544}{91}$$

$$\bar{x} = 5.97$$

Ejemplo 2: Los precios de los 97 artículos que se venden en una tienda están señalados en la tabla de la derecha. Calcular el precio promedio (la media) que existe en dicha tienda

Solución: Debe añadirse, como en el ejemplo anterior, a la tabla original una columna encabezada por fx en donde se anotarán los resultados correspondientes a las multiplicaciones de cada valor nominal x por su frecuencia f respectiva.

Por ejemplo, para la primera fila:

$$fx = 5 \times 250.00 = 1250.00$$

para la segunda fila:

$$fx = 3 \times 305.50 = 916.5$$

para la tercera fila:

$$fx = 4 \times 330.00 = 1320.00$$

La tabla completa con las tres columnas queda como se muestra a la derecha. La suma de los valores de la columna fx es 45 260.2, de manera que utilizando la fórmula para el promedio, se obtiene:

$$\bar{x} = \frac{45260.2}{97}$$

$$\bar{x} = 466.6$$

precios x	f
\$ 250.00	5
\$ 305.50	3
\$ 330.00	4
\$ 395.75	6
\$ 400.00	8
\$ 465.80	8
\$ 500.00	18
\$ 512.35	20
\$525.00	13
\$ 530.00	9
\$ 540.40	2
\$ 550.00	1

97

precios x	f	fx
\$ 250.00	5	1250.00
\$ 305.50	3	916.5
\$ 330.00	4	1320.00
\$ 395.75	6	2374.5
\$ 400.00	8	3200.00
\$ 465.80	8	3726.40
\$ 500.00	18	9000
\$ 512.35	20	10247.00
\$525.00	13	6825.00
\$ 530.00	9	4770.00
\$ 540.40	2	1080.80
\$ 550.00	1	550.00

97

45260.2

CUESTIONARIO 3.1

- 1) Calcular el promedio de los datos organizados en las tablas obtenidas en el cuestionario 2.3, página 16

3.1.2 LA MEDIA PARA FRECUENCIAS POR INTERVALOS

Cuando los datos recolectados han sido organizados en una tabla de frecuencias por intervalos, la media para poblaciones como para muestras se puede calcular por medio de la fórmula

$$\bar{x} = \frac{\sum fx}{n}$$

en donde:

\bar{x} = media

x = punto medio del intervalo.

$\sum fx$ = suma de las frecuencias por su correspondiente dato nominal.

n = suma de todas las frecuencias (número de datos recolectados).

Obsérvese que es la misma fórmula que la correspondiente a los datos organizados en tablas de distribución de frecuencias simples, en donde la única diferencia es la interpretación de la x . En una representa el valor nominal, en ésta el punto medio del intervalo. De hecho, esta situación se va a repetir en las otras dos medidas de tendencia central que faltan de estudiar aún, la mediana y la moda, ya que también se estudiarán en dos casos: cuando los datos estén organizados en tablas con frecuencias simples o cuando estén por intervalos.

Para calcular la media de datos organizados por intervalos, deben añadirse ahora dos columnas a la tabla original: la primera columna añadida es para anotar el valor del punto medio x del intervalo y la otra en la que se registren los resultados correspondientes al producto de la frecuencia por el correspondiente valor del punto medio del intervalo (fx).

Ejemplo 1: Calcular el promedio (la media) de los valores agrupados en intervalos de la tabla de la derecha.

Solución: Deben añadirse a la tabla original dos columnas encabezadas por x y por fx , en donde se anotarán los resultados correspondientes a los puntos medios de cada intervalo y al producto de la frecuencia por ese punto medio.

<i>intervalos</i>	<i>f</i>
0 - 2	12
3 - 5	13
6 - 8	23
9 - 11	16
12 - 14	18

82

La tabla completa con las cuatro columnas queda como se muestra abajo a continuación:

<i>intervalos</i>	<i>f</i>	<i>punto medio x</i>	<i>fx</i>
0 - 2	12	1	12
3 - 5	13	4	52
6 - 8	23	7	161
9 - 11	16	10	160
12 - 14	18	13	234

82

619

La suma de los valores de la columna fx es 619, de manera que utilizando la fórmula para el promedio, donde hay que recordar que $n = \sum f$, se obtiene:

$$\bar{x} = \frac{\sum fx}{n}$$

$$\bar{x} = \frac{619}{82} = 7.548$$

CUESTIONARIO 3.2

- 1) Calcular el promedio de los datos organizados en intervalos en las tablas del cuestionario 2.3, página 16.

3.2 LA MODA

La **moda** es la medida de tendencia central que se define como aquel valor nominal que tiene la frecuencia mayor. Por lo tanto, una distribución de frecuencias puede tener más de una moda o, inclusive, no tener moda cuando todos los datos tienen frecuencia 1.

La moda se simboliza con sus dos primeras iniciales: *Mo*

3.2.1 LA MODA PARA DISTRIBUCIÓN DE FRECUENCIAS SIMPLES

Cuando los datos recolectados han sido organizados en una tabla de frecuencias simples, la moda se obtiene buscando en la columna de frecuencias el o los valores que tengan mayor frecuencia.

Ejemplos: De las dos tablas siguientes, localizar la moda de cada una de ellas.

TABLA A		TABLA B	
x	f	x	f
35	12	100	25
40	8	200	29
45	13	300	27
50	11	400	29
55	16	500	22
60	10	600	24
65	11	700	28
70	15	800	25
75	14	900	28
80	5	1000	19

Solución: Para la tabla **A** basta recorrer la columna de las frecuencias y localizar que la mayor frecuencia es $f = 16$, que corresponde al dato nominal 55. Por lo tanto la moda es $Mo = 55$.

Para la tabla **B** igualmente basta localizar que la mayor frecuencia es $f = 29$, que corresponde a los datos nominales 200 y 400. Por lo tanto la moda es $Mo = 200$ y también $Mo = 400$, es decir, la tabla **B** tiene dos modas. Se dice que es **bimodal**.

3.2.2 MODA PARA DISTRIBUCIÓN DE FRECUENCIAS POR INTERVALOS

Cuando los datos recolectados han sido organizados en una tabla de frecuencias por intervalos, la moda se obtiene buscando en la columna de frecuencias el o los valores que tengan mayor frecuencia. Entonces la moda es el punto medio del intervalo que tiene la mayor frecuencia.

Ejemplos: De las dos tablas siguientes, localizar la moda de cada una de ellas.

TABLA A		TABLA B	
<i>intervalo</i>	<i>f</i>	<i>intervalo</i>	<i>f</i>
3 - 5	52	11 - 20	75
6 - 8	68	21 - 30	62
9 - 11	53	31 - 40	77
12 - 14	41	41 - 50	55
15 - 17	70	51 - 60	77
18 - 20	64	61 - 70	70
21 - 23	69	71 - 80	58
24 - 26	68	81 - 90	63
27 - 29	55	91 - 100	69
30 - 32	52	101 - 110	72
33 - 35	66	111 - 120	77
36 - 38	61	121 - 130	76
39 - 41	45	131 - 140	75

Solución: Para la tabla **A** basta recorrer la columna de las frecuencias y localizar que la mayor frecuencia es $f = 70$, que corresponde al intervalo 15-17. Por lo tanto, la moda es el punto medio de ese intervalo, es decir, $Mo = 16$.

Para la tabla **B** recorriendo la columna de las frecuencias se localiza que la mayor frecuencia es $f = 77$, que corresponde a los intervalos 31 - 40, 51 - 60 y 111 - 120. Es **trimodal**. Por lo tanto las modas son $Mo = 35.5$, $Mo = 55.5$ y $Mo = 115.5$ que son los puntos medios de los intervalos correspondientes.

CUESTIONARIO 3.3

- 1) Localizar **la moda** de los datos organizados en las tablas obtenidas en el cuestionario 2.3
- 2) Localizar **la moda** de los datos organizados en intervalos en las tablas obtenidas en el cuestionario número 2.3.

3.3 LA MEDIANA

La **mediana** es la medida de tendencia central que se define como aquel valor nominal que tiene, dentro de un conjunto de datos ordenados, arriba y abajo de él, el mismo número de datos nominales. En otras palabras, es el dato que está a la mitad, es el dato que divide en dos partes iguales a un conjunto de datos.

Por ejemplo, del conjunto 3, 3, 3, 4, 5, 5, 5, 5, 5, 6, 6, 7, 8, 9, 9 el cinco remarcado en color rojo y subrayado es el que está a la mitad del conjunto ordenado, ya que antes de él existen 7 datos y después de él también.

Para facilitar la localización de la mediana en una tabla, conviene agregarle una columna en la que se anoten las frecuencias acumuladas *fa*. Entonces, el número total de datos recolectados más uno, dividido entre dos da el dato central d_c :

$$d_c = \frac{n + 1}{2}$$

Ese resultado se busca en la columna de las frecuencias acumuladas y al dato nominal que le corresponda, es la mediana.

La mediana se simboliza con las letras: *Mdn*.

3.3.1 LA MEDIANA PARA DISTRIBUCIÓN EN FRECUENCIAS SIMPLES

Cuando los datos recolectados han sido organizados en una tabla de frecuencias simples, la mediana se obtiene buscando en la columna de frecuencias acumuladas el valor que esté situado exactamente a la mitad, conforme a la fórmula anterior. Los casos de conflicto se explican en los siguientes ejemplos.

Ejemplo 1: Localizar la mediana del conjunto de calificaciones mostrado en la siguiente tabla.

Solución: A la tabla original ya se le añadió la columna de frecuencias acumuladas.

La máxima frecuencia acumulada, que es lo mismo que el número total de datos nominales, es de 117. Es decir, la tabla corresponde a las calificaciones de 117 alumnos.

El dato central respecto de los valores nominales es el valor nominal 5 (ver tabla de la derecha). Ése no es la mediana.

El dato central respecto de las frecuencias, no de los datos nominales x , se obtiene sumando 1 al 117 y dividiéndolo entre dos, es decir,

$$d_c = \frac{117 + 1}{2} = 59$$

x	f	fa
0	1	1
1	1	2
2	3	5
3	5	10
4	6	16
5	7	23
6	11	34
7	15	49
8	25	74
9	20	94
10	23	117

117

Diagrama de anotaciones: Una flecha roja apunta hacia arriba desde la fila $x=5$ con el texto "5 datos nominales". Una flecha roja apunta hacia abajo desde la fila $x=5$ con el texto "5 datos nominales". Una flecha roja apunta hacia la izquierda desde la celda $(x=8, fa=74)$ con el texto "éste no". Una flecha roja apunta hacia la izquierda desde la celda $(x=8, fa=74)$ con el texto "éste no".

que significa que el dato ordinal 59, el quincuagésimo noveno, es el que está situado a la mitad de todos. Observando la columna de las frecuencias acumuladas se ve que hasta la calificación 7, contadas desde el principio, van apenas 49 alumnos, mientras que hasta la calificación 8 ya van 74. Esto significa que dentro de la calificación 8 está el alumno número 59, que es el central. Por lo tanto, la mediana es $Mdn = 8$.

Dicho de otra forma: cuando se fueron ordenando una por una las calificaciones, al contar el último 7 se llevaban en ese momento 49 calificaciones ordenadas (ver tabla). Al continuar, el 50° (quincuagésimo) dato o calificación fue de valor 8; el 51° (quincuagésimo primer) dato fue también de valor 8; el 52° (quincuagésimo segundo) dato o calificación fue también de valor 8, y así sucesivamente hasta el 74° (septuagésimo cuarto). Eso significa que el 59° dato correspondió al valor nominal $x = 8$.

El error más común que se comete a la hora de intentar localizar la mediana es buscar el dato nominal x central en vez del dato ordinal, o sea, el error consiste en buscar en la columna de los datos nominales x el que está a la mitad y eso no es. En el ejemplo anterior, el dato nominal x central es el 5 y ése no es la mediana.

Ejemplo 2: Localizar la mediana del conjunto de datos mostrado en la siguiente tabla de la derecha.

Solución: Dicha tabla es la tabla original a la que ya se le añadió la columna de frecuencias acumuladas.

La máxima frecuencia acumulada, que es lo mismo que el número total de datos nominales, es de 89. Es decir, la tabla corresponde a 89 datos recolectados.

Los datos centrales respecto de la tabla son los valores nominales 800 y 900 (ver tabla de la derecha). Ésos no son la moda.

El valor central respecto de las frecuencias, no de los datos nominales x , se obtiene sumando 1 al 89 y dividiéndolo entre dos, es decir,

x	f	fa
500	10	10
525	10	20
550	13	33
600	12	45
700	11	56
750	7	63
800	7	70
900	5	75
950	5	80
1000	3	83
1200	2	85
1250	2	87
1350	1	88
1400	1	89

$$d_c = \frac{1 + 89}{2} = 45$$

que significa que el dato ordinal 45 (el que ocupa el lugar 45) es el que está situado a la mitad de todos, o sea, hay 44 antes y 44 después de él. Observando la columna de las frecuencias acumuladas se ve que hasta el dato nominal $x = 600$, contados desde el principio, van 45 datos recolectados. Esto significa que de los 12 datos nominales, $x = 600$, justamente el último de ellos ocupa el orden 45, que es el central. Por lo tanto, la mediana es $Mdn = 600$.

Dicho de otra forma: cuando se fueron ordenando uno por uno los datos nominales, al contar el último correspondiente al valor $x = 550$, se llevaban hasta allí 33 datos recolectados (ver tabla). Al continuar, el 34° (trigésimo cuarto) dato fue de valor $x = 600$, el 35° (trigésimo quinto) dato fue de valor $x = 600$, el 36° (trigésimo sexto) dato fue de valor $x = 600$, y así sucesivamente hasta el 45° (cuadragésimo quinto), o sea 12 más (que es la frecuencia del dato nominal $x = 600$) Eso significa que el 45° (cuadragésimo quinto) dato correspondió al valor nominal $x = 600$.

Obsérvese que la mediana $Mdn = 600$ no es el que está situado a la mitad de la columna de los datos nominales x .

3.3.2 CASOS DE CONFLICTO

Ejemplo 3: Localizar la mediana del conjunto de datos mostrado en la siguiente tabla.

Solución: Debe entenderse que la tabla original consta solamente de las dos primeras columnas x y f . La que aparece a la izquierda es dicha tabla original a la que ya se le añadió la columna de frecuencias acumuladas.

La máxima frecuencia acumulada, que es lo mismo que el número total de datos nominales, es de 574. Es decir, la tabla corresponde a 574 datos recolectados.

El valor central respecto de las frecuencias, no de los datos nominales x , se obtiene sumando 1 al 574 y dividiéndolo entre dos, es decir,

x	f	fa
45	100	100
46	106	206
47	63	269
48	112	381
49	81	462
50	70	532
51	27	559
52	5	564
53	10	574

574

$$d_c = \frac{1 + 574}{2} = 287.5$$

que significa que tanto el dato ordinal 287 como el 288 (el que ocupa en orden el lugar 287 y el 288) son los que están situados a la mitad de todos. Observando la columna de las frecuencias acumuladas se ve que dentro del conjunto de 112 datos recolectados con valor nominal $x = 48$, están los que ocupan el orden 287 y 288, de manera que en este caso no hay conflicto para determinar la mediana y ésta es $Mdn = 48$.

Dicho de otra forma: cuando se fueron ordenando uno por uno los datos recolectados, al contar el último dato nominal con valor $x = 47$, se llevaban en ese momento 269 datos recolectados ordenados (ver tabla). Al continuar, el 270° (bicentésimo septuagésimo) dato fue de valor 48; el 271° (bicentésimo septuagésimo primer) dato fue también de valor 48; el 272° (bicentésimo septuagésimo segundo) dato fue también de valor 48, y así sucesivamente hasta 112 más (que es la frecuencia del dato nominal $x = 48$) Eso significa que el 287° (bicentésimo octogésimo séptimo) dato, lo mismo que el 288° (bicentésimo octogésimo octavo), correspondieron al valor nominal $x = 48$.

Obsérvese que la mediana $Mdn = 48$ no es el que está situado a la mitad de la columna de los datos nominales x .

Ejemplo 4: Las edades de un grupo de 28 personas van de los 45 a los 58 años, mostrado ya organizado en la siguiente tabla. Localizar la mediana de dicho conjunto de datos.

Solución: Debe entenderse que la tabla original consta solamente de las dos primeras columnas x (edades) y f . La que aparece a la izquierda es dicha tabla original a la que ya se le añadió la columna de frecuencias acumuladas. Este es un ejemplo para ilustrar el caso de conflicto para localizar la mediana.

La máxima frecuencia acumulada, que es lo mismo que el número total de datos nominales, es de 28. Es decir, la tabla corresponde a 28 datos recolectados.

<i>edad</i> x	f	fa
45	2	2
46	1	3
47	3	6
48	3	9
49	5	14
50	6	20
51	2	22
52	4	26
53	2	28

28

El valor central respecto de las frecuencias, no de los datos nominales x , se obtiene sumando 1 al número total de datos 28 y dividiéndolo entre dos, es decir,

$$f_c = \frac{28 + 1}{2} = 14.5$$

que significa que tanto el dato ordinal 14 (décimo cuarto) como el 15 (décimo quinto) son los que están situados a la mitad de todos, ya que debe comprenderse que no existen datos ordinales fraccionarios, sino solamente enteros, o sea, existe el 1^{er} dato, el 2^o dato, el 3^{er} dato, el 8^o dato, el 19^o dato, etc, pero no puede existir el 14.5^o (el décimo cuarto punto cinco) dato, si acaso el 14^o (décimo cuarto) exactamente.

Observando la columna de las frecuencias acumuladas se ve que el 14^o dato fue el último correspondiente al dato nominal $x = 49$, mientras que el 15^o dato es el primero que corresponde al dato nominal $x = 50$, por lo que se entra en el conflicto de cuál de los dos ha de ser la mediana. En casos como éste, la media de ambos se toma como la mediana, es decir, la mediana es $Mdn = 49.5$.

CUESTIONARIO 3.4

- 1) Localizar **la mediana** de los datos organizados en las tablas obtenidas en el cuestionario 2.3.

3.3.3 LA MEDIANA PARA FRECUENCIAS POR INTERVALOS O AGRUPADAS

Cuando los datos recolectados han sido organizados en una tabla de frecuencias por intervalos, la mediana podría considerarse el punto medio del intervalo en donde se localiza el valor central de las todas frecuencias; sin embargo, se acostumbra más bien localizar con exactitud un punto dentro de ese intervalo que sea el más representativo, por lo que la mediana se obtiene por medio de una fórmula.

Para utilizar la fórmula mencionada debe añadirse primero a la tabla original una columna de frecuencias acumuladas.

Se le llama **clase de la mediana** al intervalo en donde se encuentra la mediana. **Clase** o **intervalo** es lo mismo. El resultado que se obtiene con la fórmula es un valor que se encuentra dentro de la clase de la mediana.

La fórmula correspondiente es

$$M_{dn} = L + \left(\frac{\frac{n}{2} - f_a}{f} \right) i$$

en donde:

M_{dn} = mediana

L = límite inferior de la clase (o intervalo) de la mediana.

n = número total de datos.

f_a = frecuencia acumulada en la clase (intervalo) inmediata anterior a la clase (intervalo) de la mediana.

f = frecuencia de la clase de la mediana.

i = diferencia entre los límites de la clase (intervalo) de la mediana.

Para localizar el intervalo de la mediana se procede igual que para datos no agrupados.

Ejemplo 1: Localizar **la mediana** del conjunto de datos organizados en intervalos, mostrado en la siguiente tabla.

Solución: A la tabla original ya se le añadió la columna de frecuencias acumuladas.

La máxima frecuencia acumulada, que es lo mismo que el número total de datos nominales, es de 49.

El dato central respecto de las frecuencias, no de los datos nominales x , se obtiene sumando 1 al 49 y dividiéndolo entre dos, es decir,

$$d_c = \frac{49 + 1}{2} = 25$$

que significa que el dato ordinal 25 (vigésimo quinto) es el que está situado a la mitad de todos. Observando la columna de las frecuencias acumuladas se ve que hasta el intervalo **151 - 180**, contadas desde el principio, van apenas 23 datos ordenados, mientras que hasta el intervalo **181 - 210** ya van 34. Esto significa que dentro del intervalo **181 - 210** está el dato ordinal número 25, que es el central. Por lo tanto, la clase de la mediana es **181 - 210**.

Dicho de otra forma: cuando se fueron ordenando uno por uno los datos recolectados, al contar el último del intervalo **151 - 180** se llevaban en ese momento 23 datos ordenados (ver tabla). Al continuar, el 24° (vigésimo cuarto) dato fue ya del intervalo **181 - 210**; el 25° (vigésimo quinto) dato fue también de la clase **181 - 210**, y así sucesivamente hasta el 34° (trigésimo cuarto), o sea 11 más (que es la frecuencia del intervalo) Significa que el 25° (vigésimo quinto) dato correspondió al intervalo **181 - 210**.

Así que se tienen los siguientes datos para ser sustituidos en la fórmula:

$$\begin{aligned} L &= 181 \\ n &= 49 \\ fa &= 23 \\ f &= 11 \\ i &= 210 - 181 = 29 \end{aligned}$$

<i>intervalo</i>	<i>f</i>	<i>fa</i>
1 - 30	1	1
31 - 60	1	2
61 - 90	3	5
91 - 120	5	10
121 - 150	6	16
151 - 180	7	23
181 - 210	11	34
211 - 240	15	49

49

de modo que

$$Mdn = 181 + \left(\frac{\frac{49}{2} - 23}{11} \right) 29$$

$$Mdn = 181 + (0.13636) 30$$

$$Mdn = 185.09$$

Obsérvese que, como se dijo antes, el valor de la mediana está adentro del intervalo de la clase de la mediana, es decir, adentro de 181 - 210. Lo contrario sería una contradicción, pues si se afirma que el intervalo (la clase) de la mediana es 181 - 210, precisamente dentro de ese intervalo debe estar la mediana.

Ejemplo 2: Localizar la mediana del conjunto de datos organizados en intervalos mostrado en la siguiente tabla.

Solución: A la tabla original ya se le añadió la columna de frecuencias acumuladas.

La máxima frecuencia acumulada, que es lo mismo que el número total de datos nominales, es de 75.

El dato central respecto de las frecuencias, no de los datos nominales x se obtiene sumando 1 al 75 y dividiéndolo entre dos, es decir,

$$d_c = \frac{75 + 1}{2} = 38$$

que significa que el dato ordinal 38 (trigésimo octavo) es el que está situado a la mitad de todos. Observando la columna de las frecuencias acumuladas se ve que hasta el intervalo 14 - 22, contadas desde el principio, van apenas 27 datos ordenados, mientras que hasta el intervalo 23 - 31 ya van 57. Esto significa que dentro del intervalo 23 - 31 está el dato ordinal número 38, que es el central. Por lo tanto, la clase de la mediana es 23 - 31.

<i>intervalo</i>	<i>f</i>	<i>fa</i>
5 - 13	16	16
14 - 22	11	27
23 - 31	30	57
32 - 40	5	62
41 - 49	2	64
50 - 58	7	71
59 - 67	1	72
68 - 76	3	75

75

Dicho de otra forma: cuando se fueron ordenando uno por uno los datos recolectados, al contar el último del intervalo 14 - 22 se llevaban en ese momento 27 datos ordenados (ver tabla). Al continuar, el 28° (vigésimo octavo) dato fue ya del intervalo 23 - 31; el 29° (vigésimo noveno) dato fue también del mismo intervalo 23 - 31, y así sucesivamente hasta el 57° (quincuagésimo séptimo), o sea 30 más (que es la frecuencia del intervalo) Eso significa que el 38° (trigésimo octavo) dato correspondió al intervalo 23 - 31.

Así que para calcular la mediana, se tienen los siguientes datos:

$$\begin{aligned} L &= 23 \\ n &= 75 \\ fa &= 27 \\ f &= 30 \\ i &= 31 - 23 = 8 \end{aligned}$$

y sustituyendo:

$$Mdn = 23 + \left(\frac{\frac{75}{2} - 27}{30} \right) 8$$

$$Mdn = 23 + (0.35) 8$$

$$Mdn = 25.8$$

Obsérvese que, como se dijo antes, el valor de la mediana está adentro del intervalo de la clase de la mediana, es decir, adentro de 23 - 31. Lo contrario sería una contradicción, pues si se afirma que el intervalo (la clase) de la mediana es 23 - 31, precisamente dentro de ese intervalo debe estar la mediana.

En el caso de que la clase de la mediana sea el primer intervalo, como antes de él no hay ninguno dato, tampoco hay ninguna frecuencia acumulada. Así que en la fórmula el valor de fa será igual a cero. Ver el siguiente ejemplo.

Ejemplo 3: Localizar la mediana del conjunto de datos organizados en intervalos, mostrado en la siguiente tabla.

Solución: A la tabla original ya se le añadió la columna de frecuencias acumuladas.

La máxima frecuencia acumulada, que es lo mismo que el número total de datos nominales, es de 45.

El dato central respecto de las frecuencias, no de los datos nominales x se obtiene sumando 1 al 45 y dividiéndolo entre dos, es decir,

$$d_c = \frac{45 + 1}{2} = 23$$

que significa que el dato ordinal 23 (vigésimo tercero) es el que está situado a la mitad de todos. Observando la columna de las frecuencias acumuladas se ve que el primer intervalo acaba en el dato 35 (trigésimo quinto), por lo tanto allí está el 23 (vigésimo tercero). El primer intervalo es el de la clase de la mediana.

Así que para calcular la mediana, se tienen los siguientes datos:

$$\begin{aligned} L &= 1 \\ n &= 45 \\ fa &= 0 \\ f &= 35 \\ i &= 9 - 1 = 8 \end{aligned}$$

sustituyendo en la fórmula:

$$Mdn = 1 + \left(\frac{\frac{45}{2} - 0}{35} \right) 8$$

$$Mdn = 1 + \left(\frac{22.5}{35} \right) 8$$

<i>intervalo</i>	<i>f</i>	<i>fa</i>
1 - 9	35	35
10 - 18	5	40
19 - 27	2	42
28 - 36	3	45

45

$$Mdn = 1 + 5.142$$

$$Mdn = 6.14$$

LUIS CASTRO PÉREZ
www.luiscastrop.com

LUIS CASTRO PÉREZ
www.luiscastrop.com

LUIS CASTRO PÉREZ
www.luiscastrop.com

LUIS CASTRO PÉREZ
www.luiscastrop.com

CUESTIONARIO 3.5

- 1) Localizar **la mediana** de los datos organizados en las tablas obtenidas en el cuestionario 2.3, problemas 1 a 5, página 16.