Estadígrafos de Posición

Cuartiles, deciles y percentiles

Lectura e interpretación de datos

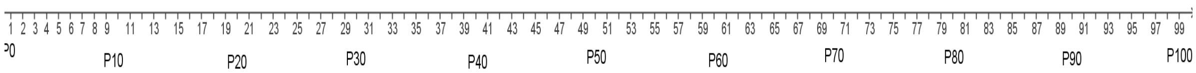
Los cuartiles forman uno de los métodos más útiles, eficaces y utilizados para describir grupos de observaciones (GENE V. GLAS, 1996).

Los cuartiles son valores que dividen a la distribución en partes iguales. Específicamente son intervalos comprendiendo un mismo número de valores. "Un cuartil es un punto en una escala numérica que se supone abarca una serie de observaciones dividiéndolas en dos grupos, cuyas respectivas proporciones se conocen" (GENE V. GLAS, 1996).

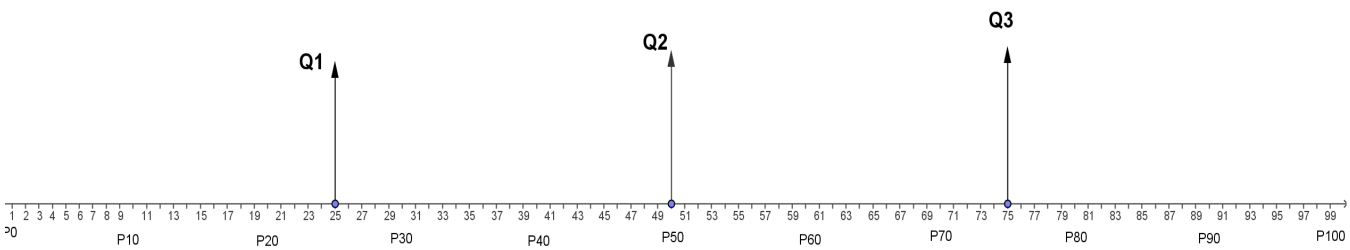
Los cuartiles, los deciles y los percentiles son los más utilizados. Veamos en qué consiste cada uno de ellos.

Percentiles: son 99 valores que dividen en cien partes iguales a un grupo de datos ordenados.

El percentil P es el punto bajo el cual se halla el P por ciento de las puntuaciones (GENE V. GLAS, 1996).



Cuartiles: son los tres valores que dividen a una distribución en cuatro partes iguales. Son un caso particular de los percentiles.

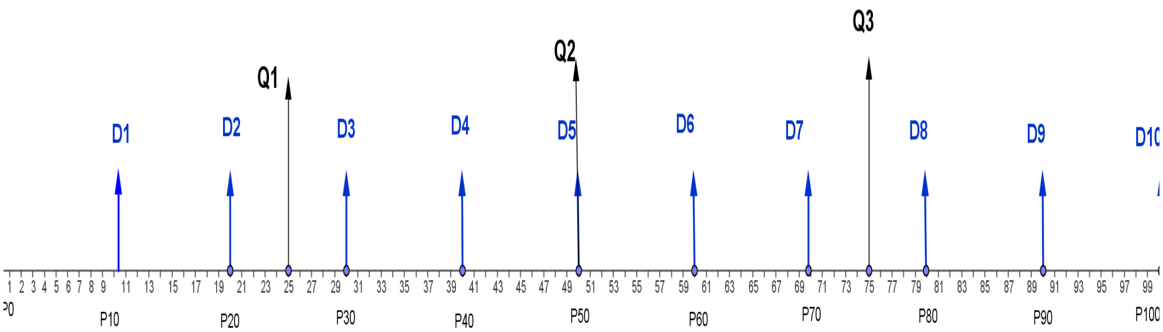


- El primer cuartil Q_1 es el menor valor que es mayor a una cuarta parte de los datos.
- El segundo cuartil Q_2 (la mediana) es el menor valor que es mayor a la mitad de los datos.
- El tercer cuartil Q_3 es el menor valor que es mayor a tres cuartas partes de los datos.

"Tres puntos adicionales, además de la mediana, en una distribución, son los cuartiles: puntos que dividen a una distribución en cuatro partes o cuartos. Esos cuartiles (Q_1, Q_2, Q_3) pueden estimarse a partir de la curva de ojiva, buscando los valores que correspondan a los valores del porcentaje acumulado de $25, 50$ y 75"

El punto que divide al cuarto inferior (25%) de los tres cuartos superiores de la distribución, es el primer cuartil (Q_1); el segundo cuartil (Q_2) es idéntico a la mediana y es el percentil 50; el tercer cuartil (Q_3) divide el cuarto superior de los tres cuartos inferiores de la distribución" (HOPKINS & HOPKIS, 1997, p. 18).

DECILES: son los nueve valores que dividen al conjunto de datos ordenados en diez partes iguales. Son también un caso particular de los percentiles.



"Análogamente, los valores que dividen los datos en diez partes iguales se llaman deciles y se representan por (D_1, D_2, D_3,..., D_9), mientras los valores divisores de los datos en cien partes iguales se llaman percentiles y se representan por $(P_1, P_2, P_3, ..., P_{99})$. El quinto decil y el quincuagésimo percentil, corresponden con la mediana. Los percentiles P_{25} y P_{75} se corresponden con el primer y tercer cuartil, respectivamente. En conjunto, cuartiles, deciles, percentiles y otros valores obtenidos por subdivisiones análogas de los datos, se llaman cuantiles" (Murray R. Spiegel, 1995, pp. 49-50).

Fórmula para calcular cualquier percentil de una distribución de frecuencias, cuyo intervalo de clase es 1:

$$P_p = L_i + \frac{p_n - f_a}{f}$$

De donde:

$L_i =$ límite inferior real de la puntuación del intervalo de longitud 1, que contiene la cuarta parte de las frecuencias. Partiendo del extremo inferior de la distribución.

$f_a =$ frecuencia acumulada hasta L_i.

$f =$ es la frecuencia del intervalo en el que se halla la cuarta parte de n, el total de frecuencias.

En una distribución de frecuencias agrupadas, para calcular un percentil, es muy parecido al caso anterior (no agrupados).

$$P_p = L_i + \frac{p_n - f_a}{f}(A)$$

De donde:

$L_i =$ límite inferior real de la puntuación del intervalo que contiene la frecuencia pn partiendo del extremo inferior.

$f_a =$ frecuencia acumulada hasta L_i.

$f =$ es la frecuencia en el intervalo en el cual se halla la cuarta parte de p_n y A es la amplitud de cualquier intervalo de puntuaciones.

$A =$ es la amplitud del intervalo.

Estadígrafos de Posición

EJEMPLO 1:

PUNTUACIÓN DEL EXAMEN	FRECUENCIA (F)	FRECUENCIA ACUMULADA (F_a)
20	1	150
19	2	149
18	5	147
17	3	142
16	12	139
15	10	127
14	19	117
13	20	98
12	33	78
11	12	45
10	6	33
9	10	27
8	5	17
7	4	12
6	2	8
5	2	6
4	3	4
3	1	1
2	0	0
1	0	0
	$n = 150$	

Determina el P_{25}, el vigésimo quinto percentil, en una distribución de frecuencias acumuladas cuando la clase de marca del intervalo es uno (DATOS NO AGRUPADOS).

“En una escuela primaria, el maestro aplicó una prueba de 20 ítems a 150 estudiantes; la puntuación obtenida corresponde al número de respuestas correctas. Se registraron los siguientes datos:

La construcción de una distribución de frecuencias acumuladas nos facilita el trabajo al realizar el cálculo de cualquier percentil, aunque el intervalo de clase sea igual a 1, como es este caso, por lo que la tabla quedaría:

Paso 1: hallar $(0.25)n$ dividiendo n entre cuatro: $\frac{150}{4} = 37.5$

Paso 2: determinar cuál es el límite inferior real L de la clase de puntuaciones, que incluya a la persona cuya puntuación es de 37.5, iniciando desde el más bajo.

Como puedes observar en la tabla:

PUNTUACIÓN DEL EXAMEN	FRECUENCIA (F)	FRECUENCIA ACUMULADA (F_a)
11	12	45
10	6	33

Como 33 personas tuvieron una puntuación inferior a 10 y 45 inferior a 11, se observa que la frecuencia 37.5 se encuentra en la clase: $10.5 – 11.5$. De donde L_i es 10.5

Paso 3: Para calcular la parte del intervalo que corresponde a la porción dividida por 37.5, utilizaremos el siguiente procedimiento:

Resta $(0.25)n$ de la frecuencia acumulada (f_a) hasta llegar a L_i, que es igual a 33, punto donde se observa se acumularon frecuencias, por lo tanto tendremos que:

$$(0.25)n – (fa) = 37.5 - 33 = 4.5$$

Paso 4: divide el resultado anterior por la frecuencia f del intervalo que contiene la puntuación 37.5, es decir:

$$\frac{4.5}{12} = 0.375$$

Con esto se ha calculado qué fracción del intervalo de clase se encuentra por debajo de la frecuencia $(0.25)n$. En el intervalo $10.5 - 11.5$ existen 12 frecuencias y como $\frac{4.5}{12} = 0.375$, se observa que las primeras $4\ ^1/_2$ frecuencias ocupan este intervalo.

Paso 5: ahora suma el resultado anterior a $L_i = 0.375 + 10.5 = 10.875$.

Aquí se aprecia que el 25% de las puntuaciones se encuentran por debajo de 10.875, así como el 75% de las puntuaciones se encuentran por encima de 10.875.

Todo esto lo puedes simplificar con la fórmula expresada con anterioridad en esta lección:

Fórmula para calcular cualquier percentil de una distribución de frecuencias, cuyo intervalo de clase es 1:

$$P_p = L_i + \frac{p_n - f_a}{f}$$

$$P_{25} = 10.5 + \frac{37.5 - 33}{12} = 10.5 + 0.375 = 10.875$$

Como puedes observar, resulta aplicable para cualquier percentil que se pida. Se describieron los pasos uno a uno para que puedas aplicar la fórmula. Conforme adquieras la destreza, puedes eliminar los pasos indicados y aplicar únicamente la fórmula como se mostró con anterioridad.

INTERPRETACIÓN DEL RESULTADO:

El lugar número 25 del cien por ciento de la población de ítems de esta escuela se encuentra en el ítem 10.87.

Estadígrafos de Posición

EJEMPLO 2:

En una prueba de aptitud matemática aplicada a un grupo de 50 estudiantes en una escuela preparatoria, se obtuvieron las siguientes calificaciones: $n = 50$

88	**74**	**77**	**69**	**79**
33	86	78	66	69
38	65	65	49	75
44	39	63	78	70
77	79	84	75	97
90	64	89	82	71
99	68	74	73	54
56	62	78	91	63
78	85	81	81	82
72	86	66	90	76

Estadígrafos de Posición

La construcción de una distribución de frecuencias acumuladas nos facilita el trabajo al realizar el cálculo de cualquier percentil, aunque el intervalo de clase (A) no sea igual a 1, como es este caso, por lo que la tabla quedaría:

i	$L_s - L_i$	f_i	x_i	F_i
1	98 – 89	6	93.5	6
2	88 – 79	11	83.5	17
3	78 – 69	16	73.5	33
4	68 – 59	10	63.5	43
5	58 – 49	3	53.5	46
6	48 – 39	2	43.5	48
7	38 – 29	2	33.5	50
		$\sum = 50$		$\sum = 50$

NOTA: esta tabla ya la habíamos realizado en la lección 1.

Paso 1: hallar $(0.25)n$ dividiendo n entre cuatro: $\frac{50}{4} = 12.5$.

Paso 2: determinar cuál es el límite inferior real L de la clase de puntuaciones, que incluya a la persona cuya puntuación es de 12.5, iniciando desde el más bajo.

i	$\square_s - L_i$	f_i
3	78 – 69	16
4	68 – 59	10

$L = 68.5 - 78.5$ De donde Li es 68.5

Paso 3: para calcular la parte del intervalo que corresponde a la porción dividida por 12.5, utilizaremos el siguiente procedimiento:

$$(0.25)n - (f_a) = 12.5 - 43 = -30.5$$

Paso 4: divide el resultado anterior por la frecuencia f del intervalo que contiene la puntuación 12.5, es decir:

$$\frac{-30.5}{16} = -1.90$$

Paso 5: ahora suma el resultado anterior a $L_i = -1.90 + 68.5 = 66.59$.

Todo esto lo puedes simplificar con la fórmula expresada con anterioridad en esta lección:

$$P_p = L_i + \frac{p_n - f_a}{f}(A)$$

$$P_{25} = 68.5 + \frac{12.5 - 43}{16}(10) = 68.5 - 19.06 = 49.43$$

Recuerda que A es el tamaño del intervalo.

Interpretación: el percentil 25 de estas calificaciones corresponde a la nota 49.43.

Se puede decir que el 25% de las calificaciones son menores a 49.43, aunque en realidad no puede tomarse al pie de la letra la afirmación, ya que pueden existir errores de apreciación al momento de determinar percentiles.

Este procedimiento lo puedes aplicar para calcular cualquier percentil que se te pida en datos agrupados.

En resumen:

Para calcular cuartiles $(Q_1,\ Q_2,\ Q_3 \text{ ó } Q_4)$

$$Q_1 = L_i + \frac{\frac{\Sigma f}{4} - F_i}{f}(A)$$

Q_1 = Cuartil uno.

L_i = Límite inferior de la clase que contiene a la cuartila uno (correspondiente a la clase donde está la primera frecuencia acumulada que contiene a la cuarta parte del conjunto $\frac{\Sigma f}{4}$).

F_i = Frecuencia absoluta acumulada anterior a la que contiene al cuartil 1.

f = Frecuencia absoluta de la clase que contiene al cuartil 1.

i = Intervalo de la clase que contiene a Q_1.

$$Q_3 = L_i + \frac{\frac{3\Sigma f}{4} - F_i}{f}(A)$$

Q_3 = Cuartil tres.

L_i = Límite inferior de la clase que contiene al cuartil tres (correspondiente a la clase donde la primera frecuencia acumulada que contiene a las tres cuartas parte de la distribución$\frac{3\Sigma f}{4}$).

F_i = Frecuencia absoluta acumulada anterior a la que contiene al cuartil 3.

f = Frecuencia absoluta de la clase que contiene al cuartil 3.

A = Intervalo de la clase que contiene a Q_3.

Este razonamiento se puede utilizar para el cálculo de otras cuartilas:

Quintilas:

$$5_1 = L_i + \frac{\frac{\Sigma f}{5} - F_i}{f}(A) \quad \text{(quintil uno)}$$

$$5_2 = L_i + \frac{\frac{2\,\Sigma f}{5} - F_i}{f}(A) \quad \text{(quintil dos)}$$

$$5_3 = L_i + \frac{\frac{3\,\Sigma f}{5} - F_i}{f}(A) \quad \text{(quintil 3)}$$

Estadígrafos de Posición

$$5_4 = L_i + \frac{\frac{4\,\Sigma f}{5} - F_i}{f}(A) \quad \text{(quintil cuatro)}$$

Decilas:

$$D_1 = L_i + \frac{\frac{\Sigma f}{10} - F_i}{f}(A) \quad \text{(decil uno)}$$

$$D_2 = L_i + \frac{\frac{2\,\Sigma f}{10} - F_i}{f}(A) \quad \text{(decil dos)}$$

$$D_3 = L_i + \frac{\frac{3\Sigma f}{10} - F_i}{f}(A) \quad \text{(decil tres)}$$

$$D_9 = L_i + \frac{\frac{9\,\Sigma f}{10} - F_i}{f}(A) \quad \text{(decil nueve)}$$

Percentiles:

$$P_1 = L_i + \frac{\frac{\Sigma f}{100} - F_i}{f}(A) \quad \text{(percentil uno)}$$

$$P_2 = L_i + \frac{\frac{2\,\Sigma f}{100} - F_i}{f}(A) \quad \text{(percentil dos)}$$

$$P_{15} = L_i + \frac{\frac{15\,\Sigma f}{100} - F_i}{f}(A) \quad \text{(percentil quince)}$$