

Análisis estadístico para datos agrupados

Grado 7

Introducción: cuando hay demasiada información...

- **Análisis de una situación común.**

Se necesita realizar un estudio sobre el número de bebidas (refrescos, lácteos, gaseosas) que consume la comunidad educativa del colegio en un mes. De la población del colegio se elige una muestra de 100 personas. Conviene realizar el análisis con las medidas estadísticas básicas?

31	18	10	27	14	31	18	10	27	14
12	24	16	28	20	12	24	16	28	20
13	30	22	9	26	13	30	22	9	26
19	11	23	15	32	19	11	23	15	32
25	17	29	21	8	25	17	29	21	8
31	18	10	27	14	31	18	10	27	14
12	24	16	28	20	12	24	16	28	20
13	30	22	9	26	13	30	22	9	26
19	11	23	15	32	19	11	23	15	32
25	17	29	21	8	25	17	29	21	8



Datos agrupados y su objetivo

Cuando el estudio de una situación contienen una gran cantidad de elementos (más de 20 datos, números, personas encuestadas, etc.), para facilitar el análisis y los cálculos es necesario agruparlos para luego realizar tratamientos similares de moda, media, mediana, desviación entre otros.

Por esa razón, al manejo de grandes cantidades de números y su interpretación en estadística se denomina análisis de datos agrupados.

Procedimientos para datos agrupados

Los elementos son clasificados por grupos, los cuales tienen una medida representativa. De acuerdo a la ocurrencia de aparición o frecuencia de esa medida se calculan moda, media, mediana, desviación, ...

- Rango

Es la resta del elemento máximo con el mínimo. Muestra el ancho numérico en que varían los elementos:

$$R = \text{Valor máx} - \text{Valor mín}$$

- Agrupación

Los elementos son clasificados por grupos o intervalos, contando la cantidad de elementos cuyo valor está entre los extremos del intervalo.

Estos intervalos se denominan intervalos de clase; pueden ser conocidos (la situación ya los plantea) o no-conocidos (hay que calcularlos de la situación).

- Frecuencia

Es el número de elementos que ocurre para cada intervalo de clase. Es un número absoluto.

- Marca de clase

Es la medida representativa de cada intervalo de clase. Es el valor promedio de los límites del intervalo:

$$Mc = (V. \text{ superior} + V. \text{ inferior})/2$$

Ejemplo de aplicación

Estudio sobre el número de bebidas que consume la comunidad educativa del colegio en un mes

- **Rango**

$$R = 32 - 8 = 24$$

- **Tabla de frecuencia**

Tabla de frecuencia		
Intervalo	Frecuencia	Marca de clase
1 - 10	12	5,5
11 - 20	40	15,5
21 - 30	40	25,5
31 - 40	8	35,5
Total	100	

- Interpretación

- 8 miembros de la comunidad consumen 5,5 bebidas al mes (medida estadística del intervalo 1-10).
- 40 miembros consumen 15,5 y 25,5 bebidas al mes respectivamente.
- Se afirma que 80 de cada 100 miembros consumen entre 15,5 y 35,5 bebidas al mes!

Medidas estadísticas en datos agrupados

Estudio sobre el número de bebidas que consume la comunidad educativa del colegio en un mes

- **Moda (M_o)**

Es la marca de clase con mayor frecuencia. Del ejemplo en cuestión, hay dos modas:

$M_o=15,5$ bebidas y $M_o=25,5$ bebidas

- **Media ($\langle X \rangle$)**

También llamado promedio ponderado, es un número que se encuentra en la mitad de las medidas estadísticas (marcas de clase).

Es una medida apropiada donde se tiene en cuenta el peso o la frecuencia de cada marca de clase. Para su cálculo, se multiplica cada marca de clase por su frecuencia y luego sumar dichos valores. Finalmente dividir la suma por el total de elementos. Del ejemplo,

Frecuencia (F)	Marca Clase (MC)	F x MC
12	5,5	66
40	15,5	620
40	25,5	1020
8	35,5	284
	Total	1990

$\langle X \rangle = 1990/100 = 19,90$ bebidas por mes

Medidas estadísticas en datos agrupados

- Mediana (Me)

Similar a datos NO agrupados, es el valor que ocupa el valor central de todos los elementos cuando están ordenados de mayor a menor o viceversa. Está medida sólo es posible en datos cuantitativos.

Según el total de elementos (aquí simbolizado con N), para obtener la mediana se diferencian dos casos. Los elementos se asume que están ordenados de la forma:

Dato1, Dato2, Dato3, Dato4,...,
Dato(i),...

- Total de elementos impar

La mediana tiene por resultado el dato central de los N elementos, tal número es

$$Me = \text{Dato } (N+1)/2$$

- Total de elementos par

La mediada tiene por resultado el promedio de los dos datos centrales de los N elementos, tales datos son

$$\text{Dato } N/2, \text{ Dato } N/2+1$$

Luego,

$$Me = \text{promedio datos centrales}$$

Aunque existen otros métodos para hallar la mediana, el mostrado aquí es muy básico.

Representación de datos agrupados

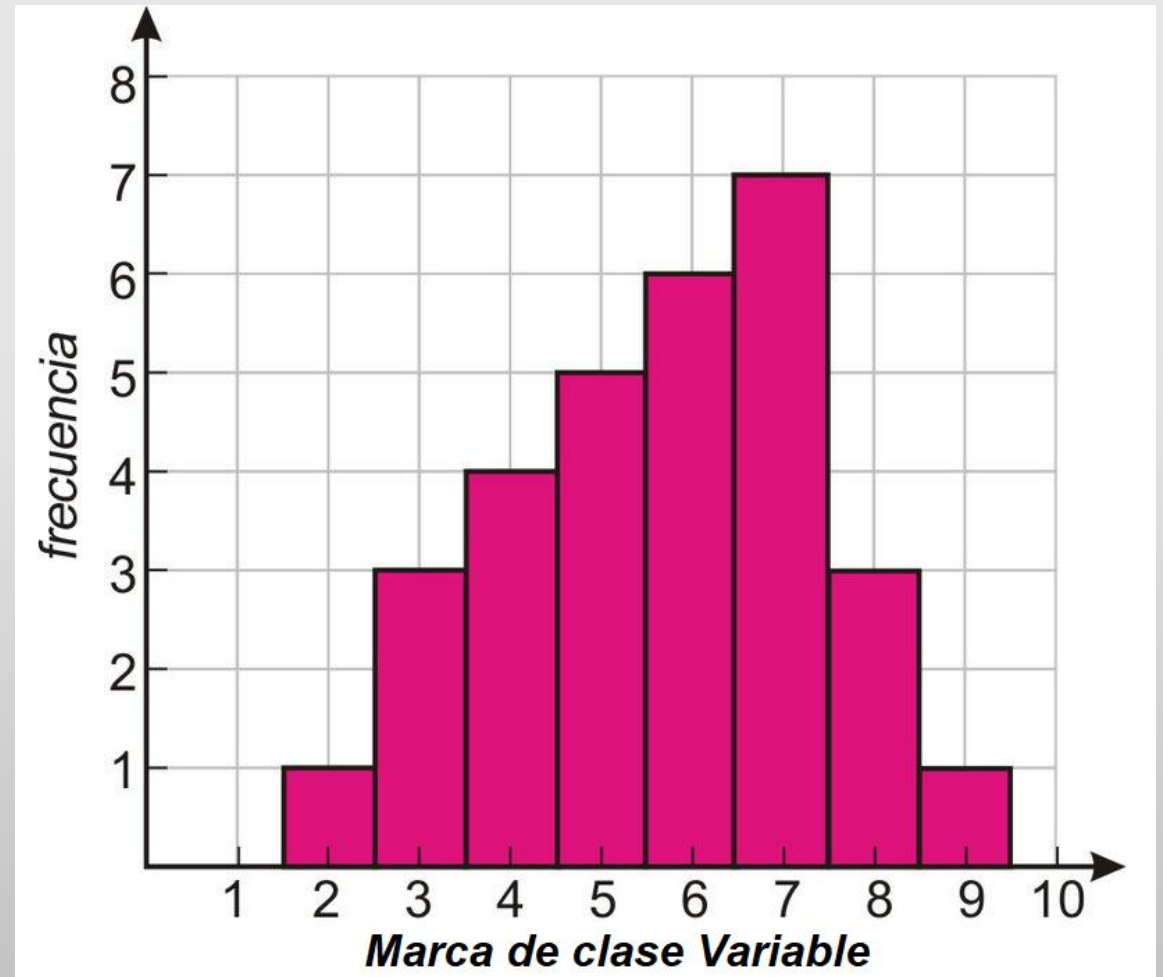
- Histograma

Es un gráfico de barras verticales en el cual el eje vertical muestra las frecuencias y en el eje horizontal las marcas de clase de la variable.

Su fin es apreciar la forma y dispersión de los elementos, logrando mostrar fácil y rápidamente la comprensión de la variable a estudiar.

Su construcción es muy simple y necesita de una tabla de frecuencias y marcas de clase.

- Ejemplo plantilla histograma



Actividad 1

Realizar un estudio de datos agrupados para la siguiente situación:

En los últimos meses los Bogotanos han expresado su inconformismo con el servicio de transporte SITP, por la demora de las rutas. Consultando a 30 usuarios, se les pregunto “¿cuánto demora su ruta de transporte?”. Ellos respondieron, dando su tiempo en minutos:

33 18 4 45 13 27 48 60 23 33
39 45 37 23 30 19 37 42 55 22
45 43 40 19 28 32 35 53 23 7

- 1. Encuentre el rango.
- 2. Complete la tabla de frecuencia.

Tiempo (min)	Frecuencia	Marca Clase
1 – 10		
11 – 20		
21 – 30		
31 – 40		
41 – 50		
51 – 60		

- 3. Halle la marca de cada clase.
- 4. Encuentre la moda del estudio.
- 5. Encuentre el promedio del estudio para la demora del transporte.

Actividad 2

Estudiar la situación siguiente usando análisis de datos agrupados; realizar:

1. Rango
2. Completar tabla de frecuencia.
3. Hallar moda, mediana y promedio.
4. Escriba interpretación(es) de su(s) resultado(s).
5. Representación gráfica (histograma).

Las notas del Qüiz de estadística de 30 estudiantes de algún grado 7 de otro lugar del universo, para la tercera clase del periodo 4 fueron (valoración 0 – 100, mínima nota aprobación 60):

45, 58, 10, 37, 49, 49, 55, 40, 55, 59, 50, 52, 52, 59, 60, 60, 67, 69, 62, 61, 37, 47, 46, 30, 50, 66, 15, 60, 49, 30

Nota	Frecuencia	Marca Clase
1 – 20		
21 – 40		
41 – 60		
61 – 80		
81 – 100		

