

# Distribución de frecuencias con datos agrupados

---

Cuando los valores de la variable son muchos, conviene agrupar los datos en intervalos o clases para así realizar un mejor análisis e interpretación de ellos.

- Para construir una tabla de frecuencias con datos agrupados, conociendo los intervalos, se debe determinar la frecuencia absoluta ( $f_i$ ) correspondiente a cada intervalo, contando la cantidad de datos cuyo valor está entre los extremos del intervalo. Luego se calculan las frecuencias relativas y acumuladas, si es pertinente.
- Si no se conocen los intervalos, se pueden determinar de la siguiente manera: (recuerda que los intervalos de clase se emplean si las variables toman un número grande de valores o la variable es continua).
  - Se busca el valor máximo de la variable y el valor mínimo. Con estos datos se determina el rango.
  - Se divide el rango en la cantidad de intervalos que se desea tener, (por lo general se determinan 5 intervalos de lo contrario es ideal que sea un número impar por ejemplo 5, 7, 9) obteniéndose así la amplitud o tamaño de cada intervalo.
  - Comenzando por el mínimo valor de la variable, que será el extremo inferior del primer intervalo, se suma a este valor la amplitud para obtener el extremo superior y así sucesivamente.

**Método Empírico:** este método depende del criterio del evaluador de los datos, por lo tanto es arbitrario. Dice lo siguiente.

$$5 \geq k \geq 20$$

Veamos como se resuelve el siguiente ejercicio:

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

- Construye una tabla de frecuencias cuyos datos estén agrupados en ocho intervalos.

1. Para poder construir la tabla de frecuencias lo primero que debemos hacer es calcular el rango.

El rango da la idea de proximidad de los datos a la media. Se calcula restando el dato menor al dato mayor.

El dato mayor y el menor lo hemos destacado con color rojo:

$$\text{Dato mayor} - \text{dato menor} = 73 - 1 = 72$$

Por lo tanto; Rango = 72

2. Se nos sugiere que debemos agrupar en 8 intervalos o clases, con este dato podemos calcular la amplitud o tamaño de cada intervalo, dividiendo el valor

$$\text{Amplitud} = \frac{\text{Rango}}{\text{N}^\circ \text{ de intervalos}}$$

del rango por la cantidad de intervalos que se desean obtener (en este caso son 8).

$$72 / 8 = 9$$

Por lo tanto la amplitud de cada intervalo será de 9.

3. *Ahora podemos comenzar a construir la tabla de frecuencias:*

Hay distintas formas de construir los intervalos dependiendo del tipo de variable que estemos trabajando.

- a. **Variables cuantitativas discretas:** solo pueden tomar un número finito de valores. Siendo por lo general estos valores los números naturales 1, 2, 3...Un ejemplo son el número de hijos, el número de habitaciones de una vivienda, el número de matrimonios de una persona. Cuando categorizamos variables discretas los límites de clase son idénticos a los límites reales. Por ejemplo, el número de personas que viven en una familia podemos agruparlo, De 1 hasta 2 (0 es imposible no hay ninguna familia sin ningún miembro) De 3 hasta 4, De 5 hasta 7.
- b. **Variables cuantitativas continuas:** Las variables continuas, por el contrario, pueden, tomar un número infinito de valores en cualquier intervalo dado. En este caso los valores se agrupan en intervalos cuyos límites inferior y superior serían los siguientes:

Inferior :  $L_i$

Superior:  $L_{i-1}$

Habitualmente, los intervalos se consideran cerrados a la izquierda y abiertos a la derecha, es decir que el extremo inferior está incluido en el intervalo, pero el extremos superior no.

Es importante mencionar que las clases o intervalos para las variables continuas pueden ser de tres tipos:

**Abiertas:** clases abiertas tienen límites determinados (a,b), pero los valores que la contienen comprenden valores muy cercanos a estos límites sin comprenderlos a ellos mismos, esto se representa con un intervalo definido entre paréntesis (). Esto quiere decir que esta clase contiene valores desde a hasta b pero no contiene exactamente a ni b solo valores muy cercanos.

**Cerradas:** las clases cerradas, además de los valores que están entre a y b, los contiene a ellos, y se representa con corchetes [a,b].

**Semi-abiertas:** pueden contener a o b más los valores que están entre ellos, y se puede representar con un corchete y un paréntesis, por ejemplo, (a,b], en este caso no contiene el valor a y si los valores de b, además de los valores que están entre estos.

c. **Registro discreto de variables continuas:** Cuando la variable considerada es continua pero ocurre que la precisión del instrumento de medida se limita a un número finito de datos, existe la opción de construir los intervalos de tal forma que ambos extremos estén incluidos en él. Ejemplo:

50 a 52,

53 a 55,

56 a 58,

59 al 61,

62 al 64

Estos serían los límites aparentes de los intervalos.

Con esta información construiremos la tabla en esta ocasión con el último método explicado.

Intervalo (i)	Edades	Frecuencia Absoluta (fi)	Frecuencia Acumulada (Fi)	Frecuencia Relativa (hi)	Frecuencia Relativa Acumulada (Hi)
1	1 - 10	7	7	$7 / 42 = 0,17$	$7 / 42 = 0,17$
2	11 - 20	6	$7+6 = 13$	$6 / 42 = 0,14$	$13 / 42 = 0,31$
3	21 - 30	8	$13+8 = 21$	$8 / 42 = 0,19$	$21 / 42 = 0,5$
4	31 - 40	6	$21+6 = 27$	$6 / 42 = 0,14$	$27 / 42 = 0,64$
5	41 - 50	5	$27+5 = 32$	$5 / 42 = 0,12$	$32 / 42 = 0,76$
6	51 - 60	4	$32+4 = 36$	$4 / 42 = 0,1$	$36 / 42 = 0,86$
7	61 - 70	4	$36+4 = 40$	$4 / 42 = 0,1$	$40 / 42 = 0,95$
8	71 - 80	2	$40+2 = 42$	$2 / 42 = 0,05$	$42 / 42 = 1$
	Amplitud 9	N: 42		$hi = fi / N$	$Hi = Fi / N$

**Responder las siguientes preguntas:**

- a. Del total de personas encuestadas, ¿cuántas personas tienen entre 31 y 40 años?

*Respuesta: Observamos los datos obtenidos en la tabla y tenemos que:*

4	31 - 40	6	$21+6 = 27$	$6 / 42 = 0,14$	$27 / 42 = 0,64$
---	---------	---	-------------	-----------------	------------------

El dato lo obtenemos de la columna de la frecuencia absoluta.

Recuerda que:

Frecuencia absoluta Corresponde a la cantidad de veces que se repite un dato.

Denotamos este valor por  $f_i$ .

Por lo tanto la respuesta es 6 personas.

- b. Del total de personas encuestadas, ¿cuántas personas tienen 60 o menos años?

*Respuesta: Observamos los datos obtenidos en la tabla y tenemos que:*

6	51 - 60	4	$32+4=36$	$4 / 42= 0,1$	$36 / 42= 0,86$
---	---------	---	-----------	---------------	-----------------

El dato lo obtenemos de la columna de frecuencia absoluta acumulada.

Recuerda que:

Frecuencia absoluta acumulada es la suma de las frecuencias absolutas observadas hasta el intervalo  $i$ .

En este caso es el intervalo 6. Por lo tanto la respuesta es 36 personas tienen 60 o menos años.

- c. ¿Cuál es la probabilidad de, que al elegir al azar a un persona consultada, esta tenga entre 11 y 20 años?

*Respuesta: Observamos los datos obtenidos en la tabla y tenemos que:*

2	11 - 20	6	$7+6= 13$	$6 / 42= 0,14$	14%
---	---------	---	-----------	----------------	-----

El dato lo obtenemos de la columna de frecuencia relativa.

Recuerda que:

Frecuencia relativa Corresponde a la probabilidad de pertenecer a cierta categoría. Se puede expresar en tantos por ciento.

En este caso es el intervalo 2, ya que es ahí donde se encuentran las edades entre 11 y 20 años.

Entonces la respuesta es: La probabilidad es 14%.

Por último vamos a repasar el concepto de:

**Frecuencia relativa acumulada (Hi)**, Es la probabilidad de observar un valor menor o igual al valor que toma la variable en estudio en ese intervalo.

Se calcula dividiendo  $F_i$  por el número total de datos. También puedes calcularlo Sumando la frecuencia relativa de cada grupo con la frecuencia relativa acumulada del grupo anterior.

Si haces correctamente estos cálculos, el último grupo tendrá una frecuencia acumulada de 1, o muy cerca de 1, permitiendo redondear el error.

Recuerda que este valor se puede expresar como porcentaje, para esto solo debes multiplicar el valor obtenido por 100 y listo!!!

**Este cálculo te sirve en el caso de que te pregunten:**

- d. Si le preguntas a una persona cualquiera ¿Cuál es la probabilidad de que tenga 50 años o menos?

5	41 - 50	5	$27+5= 32$	$5 / 42= 0,12$	$32 / 42= 0,76$	76%
---	---------	---	------------	----------------	-----------------	-----

*Respuesta: La probabilidad es de un 76%*

Fuente: <http://www.portaleducativo.net/octavo-basico/791/Tablas-de-frecuencias-con-datos-agrupados>

Editado: Edufuturo

Palabras: 1215