

Estadística descriptiva

Jessica Nathaly Pulzara Mora
jessica.pulzara@udea.edu.co

Departamento de ingeniería de sistemas



**UNIVERSIDAD
DE ANTIOQUIA**

Tabla de frecuencias variable cuantitativa

Intervalos I_i	Marca de clase	Frecuencias		
		Absolutas f_i	Relativas	Porcentajes
I_1		f_1		
I_2		f_2		
...		...		
I_k		f_k		
Total		n	1.00	100.00

Tabla de frecuencias – variable cuantitativa

Marca de clase: Punto medio de los extremos del intervalo.

$$[a, b) \rightarrow m_i = \frac{a + b}{2}$$

Tabla de frecuencias – variable cuantitativa

Proceso de Agrupamiento en Intervalos (Clases):

1. Identificar el valor mínimo (x_{min}) y máximo (x_{Max}) de los datos.
2. Calcular el rango de los datos:

$$R = x_{Max} - x_{min}$$

3. Determinar la cantidad de intervalos $[a; b)$:

$$k = 1 + 3.32 \cdot \log_{10}(n) \quad (\text{Regla de Sturges})$$

Tabla de frecuencias – variable cuantitativa

4. Calcular la amplitud del intervalo. Redondear al entero superior:

$$A = \frac{R}{k}$$

5. Establecer intervalos.
6. Realizar el conteo de datos para cada intervalo (frecuencia absoluta).
7. Calcular las otras frecuencias y la marca de clase.

Verificar que se cumple la condición $(A)(k) \geq R$; lo cual asegura que ninguna observación quede fuera de la distribución de frecuencias.

Ejemplo

Datos correspondientes al número de bicicletas prestadas durante varias semanas por una estación del sistema de bicicletas públicas:

Lun.	Mar.	Mier.	Jue.	Vier.	Sáb.
68	65	12	22	79	31
63	43	32	43	27	28
42	25	49	27	22	25
27	74	38	49	23	45
30	51	42	28	24	12
36	36	27	23	25	57
28	42	31	19	44	51
32	28	50	46	30	43
12	38	21	16	24	69
	47	23	49		

Tabla de frecuencias – variable cuantitativa

Vamos a construir la tabla de frecuencias y a realizar algunas interpretaciones.

Solución:

- Rango:

$$R = x_{Max} - x_{min} = 79 - 12 = 67$$

- Número de intervalos:

$$k = 1 + 3.32 \log_{10}(57) = 6.83 \approx 7$$

- Amplitud

$$A = \frac{R}{k} = \frac{67}{7} = 9.57 \approx 10$$

Tabla de frecuencias correspondiente:

Intervalo de Clase	m_i	f_i	F_i	f_{Ri}	F_{Ri}	$f_{Ri}\%$	$F_{Ri}\%$
[12, 22)	17	6	6	0,11	0,11	11%	11%
[22, 32)	27	22	28	0,38	0,49	38%	49%
[32, 42)	37	6	34	0,11	0,60	11%	60%
[42, 52)	47	16	50	0,28	0,88	28%	88%
[52, 62)	57	1	51	0,02	0,90	2%	90%
[62, 72)	67	4	55	0,07	0,97	7%	97%
[72, 82)	77	2	57	0,03	1	3%	100%

Respondamos,

- ¿Cuántas veces se prestaron de 22 a 32 ciclas?
- ¿Qué porcentaje de las veces se prestaron entre 72 y 82 ciclas?

Análisis de 2 variables cualitativas

Tablas de contingencia

Tablas de contingencia

- Exclusiva para variables **cualitativas**
- Son tablas cruzadas: las filas son las clases de una variable, y las columnas las clases de otra variable.

Ejemplo

Con el fin de determinar que materias opcionales, en deportes, se podrían proponer a los estudiantes de una universidad, se hizo necesario hacer un estudio acerca de la actividad física que ellos realizan en su tiempo libre, para ello, hizo una encuesta entre 156 estudiantes de varias carreras y se les preguntó acerca de su deporte preferido para practicar en tiempo libre. Los resultados se muestran en la siguiente tabla de contingencia:

DEPORTE PREFERIDO							
G		fútbol	Baloncesto	Voleibol	Atletismo	Otros	Total
	Hombre	35	14	10	6	7	72
	Mujer	3	38	19	13	11	84
	Total	38	52	29	19	18	156

DEPORTE PREFERIDO							
G		fútbol	Baloncesto	Voleibol	Atletismo	Otros	Total
	Hombre	35/156	14/156	10/156	6/156	7/156	72/156
	Mujer	3/156	38/156	19/156	13/156	11/156	84/156
	Total	38/156	52/156	29/156	19/156	18/156	156/156

DEPORTE PREFERIDO							
G		fútbol	Baloncesto	Voleibol	Atletismo	Otros	Totales
	Hombre	22.44%	8.97%	6.41%	3.85%	4.49%	46.15%
	Mujer	1.92%	24.36%	12.18%	8.33%	7.05%	53.85%
	Totales	24.36%	33.33%	18.59%	12.18%	11.54%	100%

Interpretemos,

- El 24.36 % de la personas que fueron encuestadas fueron mujeres que prefieren practicar el baloncesto.
- El 3.85 % de los hombres se inclinan por el atletismo y el 1.92 % de las mujeres prefieren el futbol, el cual es el porcentaje más bajo de la muestra.
- El deporte que más practican independientemente del género, es el baloncesto con un 33.33 %.

Tabla de frecuencia bidireccional

Gender\ Hobby	Dance	Sports	TV	Total
Male	2	10	8	20
Female	16	6	8	30
Total	18	16	16	50



Tabla de frecuencia relativa bidireccional

Gender\ Hobby	Dance	Sports	TV	Total
Male	0,04	0,20	0,16	0,4
Female	0,32	0,12	0,16	0,6
Total	0,36	0,32	0,32	1

Tabla de frecuencia relativa bidireccional por fila

Gender\ Hobby	Dance	Sports	TV	Total
Male	0,10	0,50	0,40	1
Female	0,53	0,20	0,27	1

Tabla de frecuencia relativa bidireccional por columna

Gender\ Hobby	Dance	Sports	TV
Male	0,11	0,63	0,50
Female	0,89	0,38	0,50
Total	1	1	1

Ejercicio

En una central de procesamiento de datos se revisan los tiempos de respuesta a diversos usuarios (39), los datos se muestran a continuación (tiempo en segundos):

25 31 42 22 36 31 45 50 52 35 33 32 37
28 20 28 35 28 39 43 48 51 40 37 36 29
31 25 39 23 26 40 47 44 55 30 28 38 28

Construir la tabla de frecuencia y escribir algunas conclusiones.