

ESTADÍSTICA UNIDIMENSIONAL

PARTE 1: Análisis estadístico de los datos de pluviometría en la región de Almuñécar recogidos el lunes 11 de febrero en la finca "El Zahorí"

1. CLASIFICACIÓN VARIABLES ESTADÍSTICA

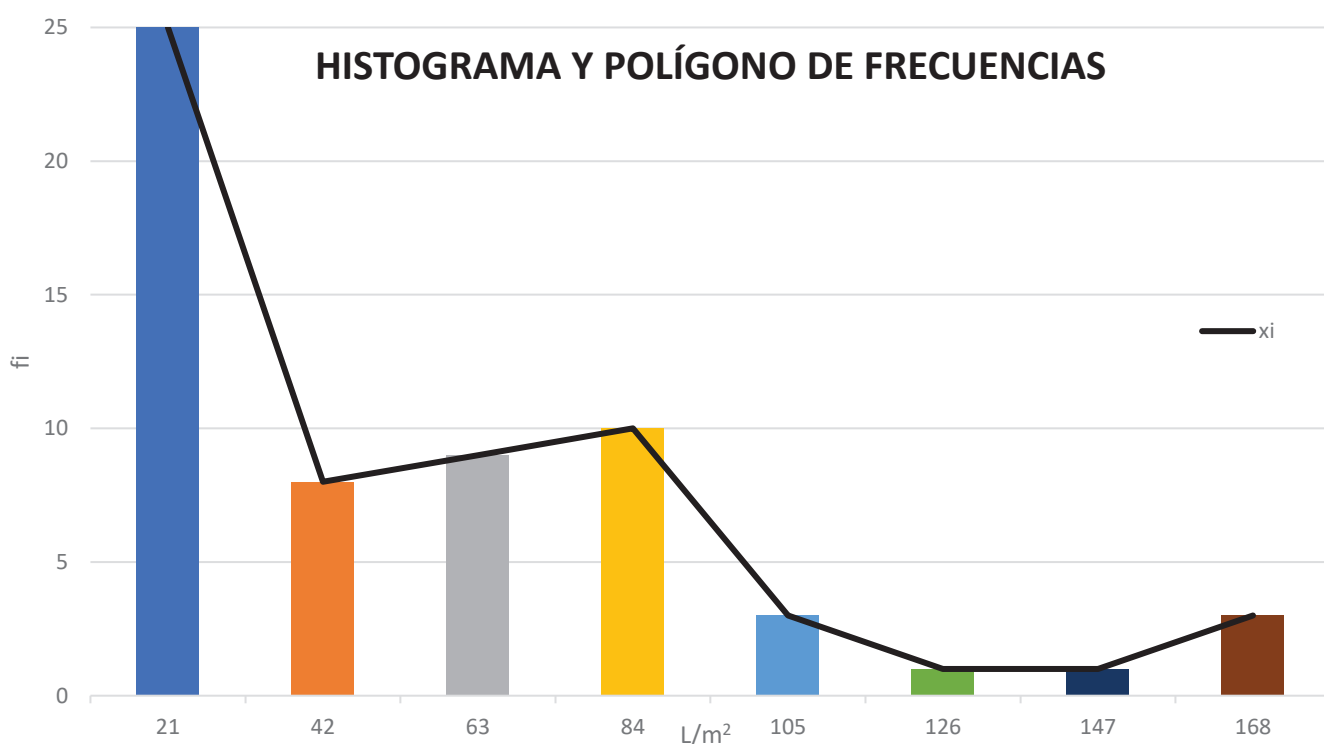
Variable estadística cuantitativa, continua.

Agrupamos los datos \rightarrow N.º de intervalos = $\sqrt{N} = \sqrt{60} = 7,7 \approx 8$ intervalos

$$\text{Amplitud de los intervalos} = \frac{\text{Max}-\text{Min}}{\sqrt{60}} = \frac{165,5-0}{\sqrt{60}} = 21,37 \approx 21$$

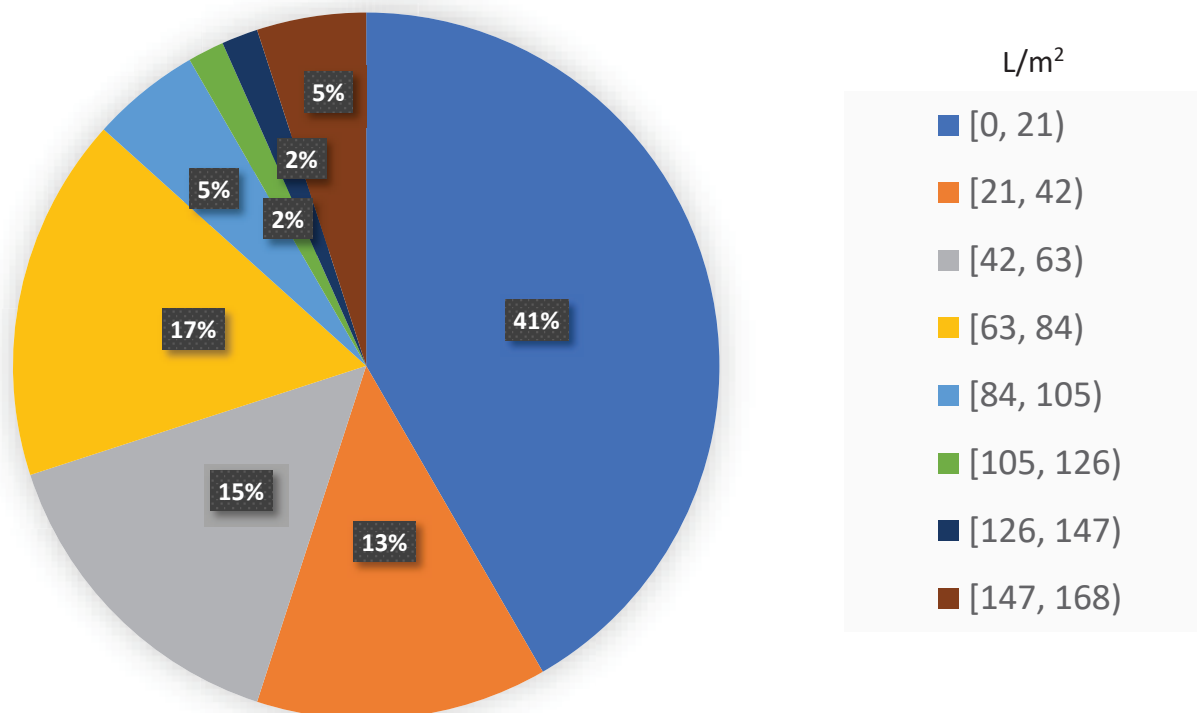
2. TABLA DE FRECUENCIAS

L/m ²	xi	fi	Fi	hi = fi/N	Hi	Pi = hi · 100	αi Ángulo	xi · fi	xi ² · fi
[0, 21)	10,5	25	25	0,42	0,42	42%	150º	262,5	2756,25
[21, 42)	31,5	8	33	0,13	0,55	13%	48º	252	7938
[42, 63)	52,5	9	42	0,15	0,7	15%	54º	472,5	24806,25
[63, 84)	73,5	10	52	0,17	0,87	17%	60º	735	54022,5
[84, 105)	94,5	3	55	0,05	0,92	5%	18º	283,5	26790,75
[105, 126)	115,5	1	56	0,02	0,94	2%	6º	115,5	13340,25
[126, 147)	136,5	1	57	0,02	0,96	2%	6º	136,5	18632,25
[147, 168)	157,5	3	60	0,05	1	5%	18º	472,5	74418,75
		N = 60		1		100%	360º	2730	222705



3. GRÁFICOS ESTADÍSTICOS

DIAGRAMA DE SECTORES



Los porcentajes corresponden a el dato P_i en la tabla de frecuencias

4. MEDIDAS DE CENTRALIZACIÓN

La **media** se encuentra al sumar todos los números de los datos y luego al dividir entre el número de valores en el conjunto. La **mediana** es el valor medio cuando un conjunto de datos se ordena de menor a mayor. La **moda** es el número que se presenta con más frecuencia en un conjunto de datos.

$$\text{Media} \rightarrow \bar{X} = \frac{\sum f_i \cdot x_i}{N} = \frac{2730}{60} = 45,5$$

Mediana \rightarrow El primer valor que cumple que $F_i > \frac{60}{2} = 30$ es $F_2 = 33$

Por tanto, el intervalo mediano = [21, 42)

$$\text{Mediana} = \text{Me} = L_i + c \cdot \frac{N/2 - F_{i-1}}{f_i} = 21 + 21 \cdot \frac{60/2 - 25}{8} = 34,13$$

Moda \rightarrow La frecuencia mayor es 25 que corresponde al intervalo [0, 21), que es el intervalo modal.

5. MEDIDAS DE POSICIÓN

Las medidas de posición son cuartiles. Los cuartiles son los tres valores de la variable que dividen a un conjunto de datos ordenados en cuatro partes iguales. Q_1 , Q_2 y Q_3 determinan los valores correspondientes al 25%, al 50% y al 75% de los datos. Q_2 coincide con la mediana.

El primer valor que cumple que $F_i > \frac{60}{4} = 15$ es $F_1 = 25$

$Q_1 \rightarrow 25\% \text{ de } 60 = 15 \rightarrow F_1 = 25 \rightarrow \text{Corresponde al intervalo } [0, 21)$

Por interpolación $\rightarrow Q_1 = Li + c \cdot \frac{N/4 - F_{i-1}}{f_i} = 0 + 21 \cdot \frac{60/4 - 0}{25} = 12,6 \rightarrow [0, 21)$

$Q_2 = Me = 34,13$

$Q_3 \rightarrow 75\% \text{ de } 60 = 45 \rightarrow \text{Corresponde al intervalo } [63, 84)$

El primer intervalo que cumple que $F_i > \frac{3 \times 60}{4} = 45$ es $F_3 = 52$

Por interpolación $\rightarrow Q_3 = Li + c \cdot \frac{3N/4 - F_{i-1}}{f_i} = 63 + 21 \cdot \frac{3 \cdot 60/4 - 42}{10} = 69,3 \rightarrow [63, 84)$

6. MEDIDAS DE DISPERSIÓN

Las medidas de dispersión son parámetros estadísticos que indican como se alejan los datos respecto de la media aritmética. Sirven como indicador de la variabilidad de los datos.

La varianza es una medida de dispersión que representa la variabilidad de una serie de datos respecto a su media. La desviación típica es la medida de dispersión más común, que indica qué tan dispersos están los datos con respecto a la media. El coeficiente de variación es la relación entre la desviación típica de una muestra y su media.

Varianza $\rightarrow \sigma^2 = \overline{X^2} - \bar{X}^2$

$$\sigma^2 = \frac{\sum x_i^2 \cdot f_i}{N} - \bar{X}^2 = \frac{222705}{60} - 45,5^2 = 1641,5$$

La desviación típica (σ) mide cuánto se separan los datos.

Desviación típica $\rightarrow \sigma = \sqrt{\sigma^2} = \sqrt{1641,5} = 40,5$

El coeficiente de variación es la relación entre la desviación típica de una muestra y su media.

Coeficiente de variación $\rightarrow CV = \frac{\sigma}{\bar{X}} = \frac{40,5}{45,5} = 0,89$

7. CONCLUSIÓN Y ANÁLISIS

En la mayoría de los meses que se han producido precipitaciones estas no han superado los 21 L/m^2 , la cual corresponde al 41% del total. Las mínimas precipitaciones se encuentran entre 126 L/m^2 y 147 L/m^2 , cada una corresponde al 2% del total de litros recogidos. La media de todos los años es de $45,5 \text{ L/m}^2$

Podemos concluir que es un zona poco lluviosa, con un clima seco.

PARTE 2: Realiza una encuesta

Encuesta realizada a los alumnos de 1º BACH (científico y humanidades) sobre el pasado examen de inglés.

1. CLASIFICACIÓN VARIABLES ESTADÍSTICA

Variable estadística cuantitativa, discreta.

DATOS:

Notas	0	1	2	3	4	5	6	7	8	9	10
Número de alumnos	1	2	0	6	4	11	2	8	7	5	3

2. TABLA DE FRECUENCIAS

xi	fi	Fi	hi = fi/N	Hi	Pi = Hi · 100	αi Ángulo	xi · fi	xi ² · fi
0	1	1	0,02	0,02	2%	7,2º	0	0
1	3	4	0,06	0,08	6%	21,6º	3	3
2	0	4	0	0,08	0%	0º	0	0
3	6	10	0,12	0,2	12%	43,2º	18	54
4	4	14	0,08	0,28	8%	28,8º	16	64
5	11	25	0,22	0,5	22%	79,2º	55	275
6	2	27	0,04	0,54	4%	14,4º	12	72
7	8	35	0,16	0,7	16%	57,6º	56	392
8	7	42	0,14	0,84	14%	50,4º	56	448
9	5	47	0,1	0,94	10%	36º	45	405
10	3	50	0,06	1	6%	21,6º	30	300
	50		1		100%	360º	291	2013

3. GRÁFICOS ESTADÍSTICOS

DIAGRAMA DE BARRAS Y POLÍGONO DE FRECUENCIAS

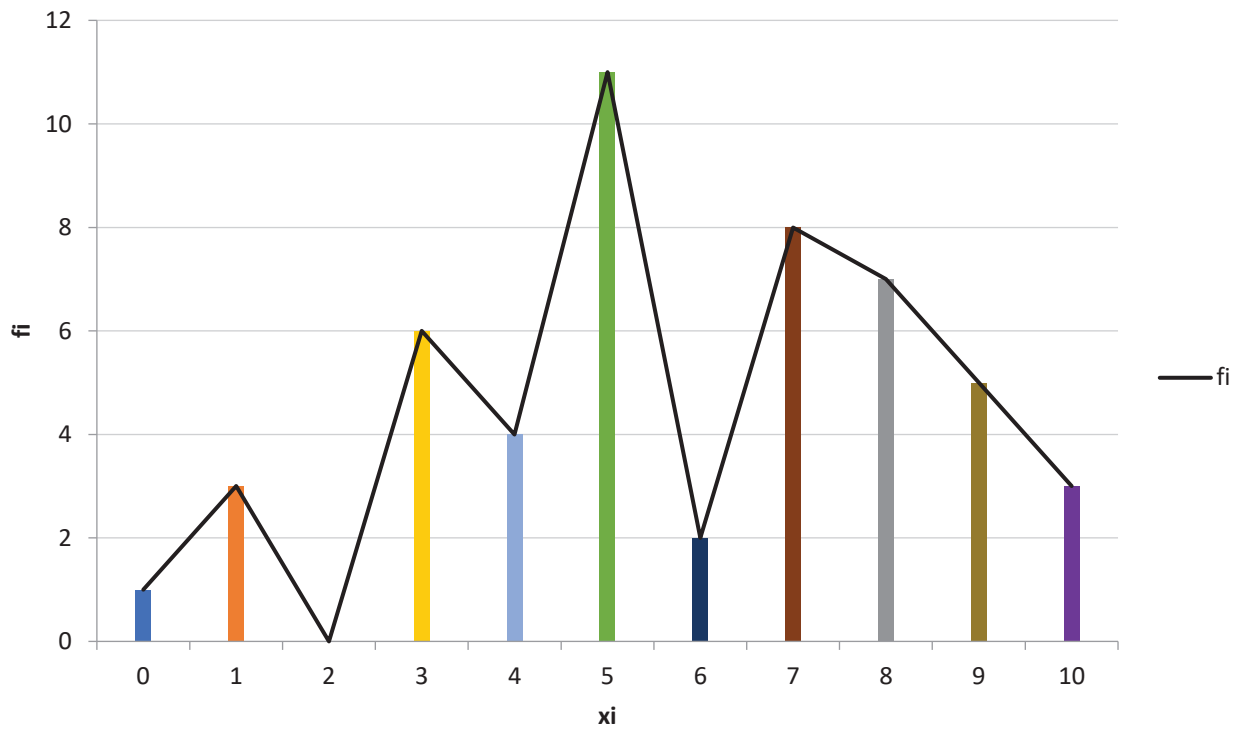
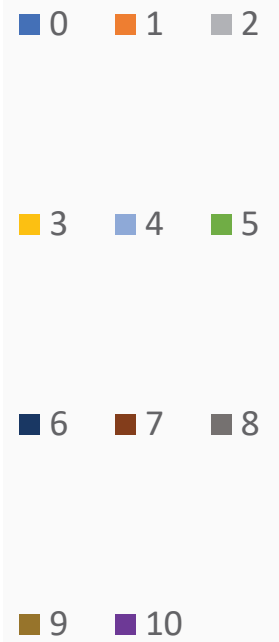
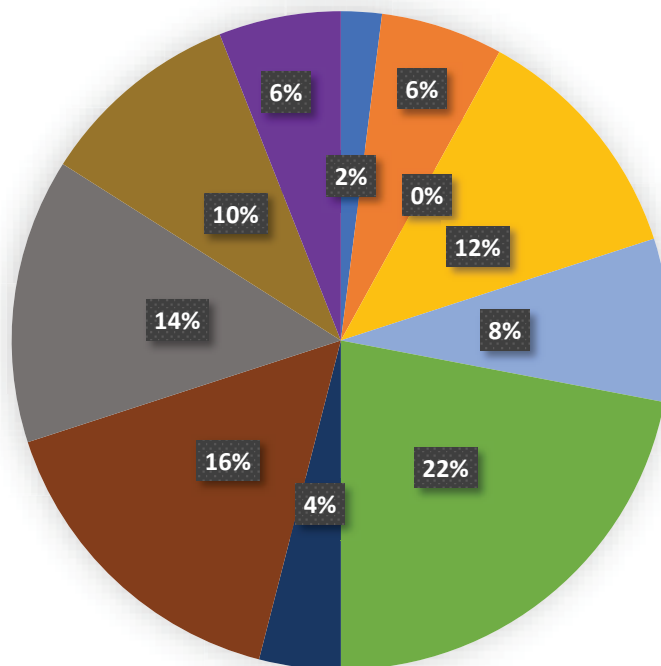


DIAGRAMA DE SECTORES



4. MEDIDAS DE CENTRALIZACIÓN

La **media** se encuentra al sumar todos los números de los datos y luego al dividir entre el número de valores en el conjunto. La **mediana** es el valor medio cuando un conjunto de datos se ordena de menor a mayor. La **moda** es el número que se presenta con más frecuencia en un conjunto de datos.

$$\text{Media} \rightarrow \bar{X} = \frac{\sum fi \cdot xi}{N} = \frac{291}{50} = 5,82$$

Moda \rightarrow Dato con mayor frecuencia $M_o = 5$

Mediana \rightarrow Valor que ocupa la posición central de los datos después de ordenarlos

Al haber 50 datos hay que fijarse en el primer valor que cumple $F_i > 25$. Por tanto :
 $Me = 6$

5. MEDIDAS DE POSICIÓN

Las medidas de posición son cuartiles. Los cuartiles son los tres valores de la variable que dividen a un conjunto de datos ordenados en cuatro partes iguales. Q_1 , Q_2 y Q_3 determinan los valores correspondientes al 25%, al 50% y al 75% de los datos. Q_2 coincide con la mediana.

$Q_1 \rightarrow 25\%$ de 50 = 12,5 \rightarrow El primer valor que cumple $F_i > 12,5$ es 4 $\rightarrow Q_1 = 4$

$Q_2 \rightarrow 50\%$ de 50 = 25 $\rightarrow Q_2 = 6 = Me$

$Q_3 \rightarrow 75\%$ de 50 = 37,5 $\rightarrow Q_3 = 8$

6. MEDIDAS DE DISPERSIÓN

Las medidas de dispersión son parámetros estadísticos que indican como se alejan los datos respecto de la media aritmética. Sirven como indicador de la variabilidad de los datos.

La varianza es una medida de dispersión que representa la variabilidad de una serie de datos respecto a su media. La desviación típica es la medida de dispersión más común, que indica qué tan dispersos están los datos con respecto a la media. El coeficiente de variación es la relación entre la desviación típica de una muestra y su media.

$$\text{Varianza} \rightarrow \sigma^2 = \overline{X^2} - \bar{X}^2$$

$$\sigma^2 = \frac{\sum xi^2 \cdot fi}{N} - \bar{X}^2 = \frac{2013}{50} - 5,82^2 = 6,3876$$

$$\text{Desviación típica} \rightarrow \sigma = \sqrt{\sigma^2} = \sqrt{6,3876} = 2,5274$$

$$\text{Coeficiente de variación} \rightarrow CV = \frac{\sigma}{\bar{X}} = \frac{2,5274}{5,82} = 0,4343$$

7. CONCLUSIÓN Y ANÁLISIS

La mayoría de los estudiantes se encuentran han aprobado, un 72%. 11 estudiantes han obtenido un 5, siendo la nota más común. Ningún estudiante ha sacado un 2 y solo 1 ha sacado un 0. La media de los estudiantes es de un 5,82.

A pesar de que la mayoría están aprobados, las notas que mayor porcentaje tienen no son muy altas.