



**Facultad de Ciencias Económicas**

Universidad Nacional de Jujuy

**ESTADÍSTICA**  
**Año: 2022**

# ESTADÍSTICA DESCRIPTIVA

## CLASE PRACTICA





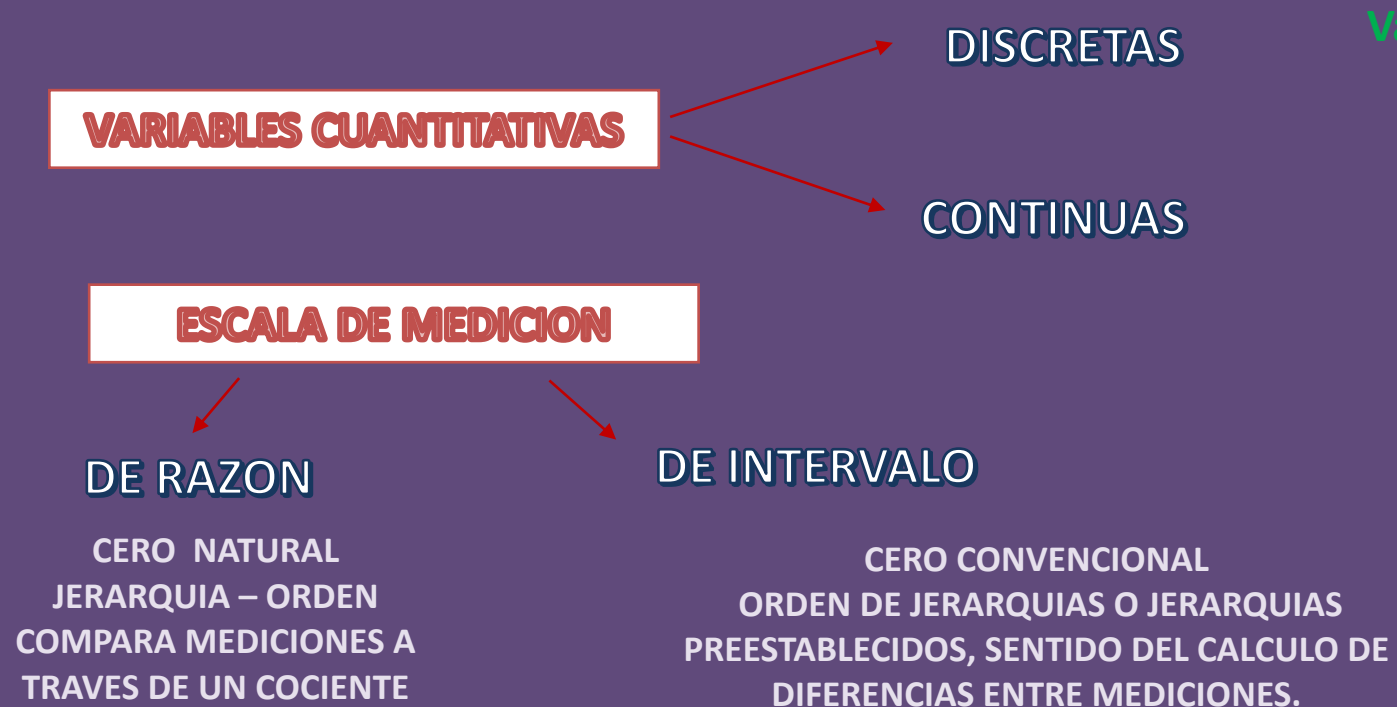
## Problema 5

La cantidad diaria de fallecimientos por COVID-19 en marzo de 2022 en una provincia del norte argentino fue:

0 3 2 6 1 4 0 3 2 4 2 0 3 1 1 5 4 0 1 5 0 1 2 3 2 1 4 0 3 2 2

## a) Identificar la variable, clasificarla e indicar su escala de medición

**X: Cantidad diaria de fallecimientos por COVID-19 en marzo 2022 en una provincia del norte argentino.**



**Variable Cuantitativa Discreta. Escala de Medición: De Razón.**

### Aclaración

La Población en estudio son las personas fallecidas por día en una provincia del norte argentino.  
¿Que queremos saber de esta población?

Para ello, **DEFINIMOS la VARIABLE**



b) Elaborar una tabla de frecuencias en las que se incluyan: frecuencia absoluta, absoluta acumulada, relativa, relativa porcentual y relativa acumulada. Interpretar dos frecuencias de cada tipo.

## TABULACIÓN DE DATOS



Donde **n** es la cantidad de observaciones

**SERIE SIMPLE**  
( $n < 30$ )

$x_i : x_1, x_2, x_3, \dots, x_n$

**DIAGRAMA DE TALLO Y HOJAS**

**SERIE O DISTRIBUCIÓN DE FRECUENCIAS**  
( $n \geq 30$ )

$x_i : x_1, x_2, x_3, \dots, x_k$

$f_i : f_1, f_2, f_3, \dots, f_k$

Donde **K** es la cantidad de valores que asume la variable

\* Sin Intervalos (K son los valores de la variable)

\* Con Intervalos (donde K es la cantidad de Intervalos)



## OTRAS CLASES DE FRECUENCIAS...

RELATIVAS:  $r_i = f_i / n$        $\sum_i r_i = 1$        $i = 1, \dots k$

RELATIVAS PORCENTUALES:  $r_i \cdot 100^k$        $\sum_{i=1} r_i = 100 \%$

## Frecuencias Acumuladas para una serie de frecuencias sin intervalos

$\leq$  que

$f_i$     $F_i$

$x_1$

$f_1$     $F_1 = f_1$

$x_2$

$f_2$     $F_2 = f_1 + f_2 = F_1 + f_2$

$x_3$

$f_3$     $F_3 = f_1 + f_2 + f_3 = F_2 + f_3$

$x_4$

$f_4$     $F_4 = f_1 + f_2 + f_3 + f_4 = F_3 + f_4$

.

.

$x_k$

$f_k$     $F_k = f_1 + \dots + f_{k-1} = F_{k-1} + f_k$     $= n$



Gráfico escalonado creciente

## Para una serie de frecuencias sin intervalos

$\geq$  que

$f_i$      $F_i$

$x_1$

$f_1$      $F_1 = n$

$x_2$

$f_2$      $F_2 = n - f_1 = F_1 - f_1$

$x_3$

$f_3$      $F_3 = n - f_1 - f_2 = F_2 - f_2$

$x_4$

$f_4$      $F_4 = n - f_1 - f_2 - f_3 = F_3 - f_3$

•

•    •

$x_k$

$f_k$      $F_k = f_k$

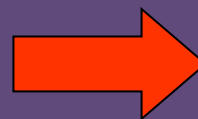


Gráfico escalonado decreciente



## Tabla de Frecuencias

| $x_i$ | $f_i$ | $r_i$  | $r_i\%$ | $F_i \downarrow$ | $F_i \uparrow$ | $Fi\% \downarrow$ | $Fi\% \uparrow$ | $x_i \cdot f_i$ | $(x_i - \bar{x})$ | $(x_i - \bar{x})^2$ | $(x_i - \bar{x})^2 \cdot f_i$ |
|-------|-------|--------|---------|------------------|----------------|-------------------|-----------------|-----------------|-------------------|---------------------|-------------------------------|
| 0     | 6     | 0,1935 | 19,35%  | 6                | 31             | 19,35%            | 100,00%         | 0               | -2,16             | 4,67                | 27,99                         |
| 1     | 6     | 0,1935 | 19,35%  | 12               | 25             | 38,71%            | 80,65%          | 6               | -1,16             | 1,35                | 8,07                          |
| 2     | 7     | 0,2258 | 22,58%  | 19               | 19             | 61,29%            | 61,29%          | 14              | -0,16             | 0,03                | 0,18                          |
| 3     | 5     | 0,1613 | 16,13%  | 24               | 12             | 77,42%            | 38,71%          | 15              | 0,84              | 0,71                | 3,53                          |
| 4     | 4     | 0,1290 | 12,90%  | 28               | 7              | 90,32%            | 22,58%          | 16              | 1,84              | 3,39                | 13,54                         |
| 5     | 2     | 0,0645 | 6,45%   | 30               | 4              | 96,77%            | 9,68%           | 10              | 2,84              | 8,07                | 16,13                         |
| 6     | 1     | 0,0323 | 3,23%   | 31               | 1              | 100,00%           | 3,23%           | 6               | 3,84              | 14,75               | 14,75                         |
|       | 31    | 1,0000 | 100,00% |                  |                |                   |                 | 67              |                   | 32,94               | 84,19                         |

## Interpretación:

### Frecuencias Absolutas:

- En 6 días del mes de marzo no se produjeron fallecimientos por COVID-19 en una provincia de norte argentino.
- En 5 días del mes de marzo se produjeron tres fallecimientos por COVID -19 en una provincia del norte argentino.

### Frecuencias Relativas:

- 0,19 es la proporción de días del mes de marzo en donde se produjo un fallecimiento por COVID-19 en un día en una provincia del norte argentino.
- La proporción de días en las que se produjeron seis (6) Fallecimientos por COVID-19 es del 0,023 durante el mes de marzo en una provincia del norte argentino.

### Frecuencia Relativa Porcentual:

- En el 22,58% de los días del mes de marzo se produjeron dos fallecimientos por COVID-19, en una provincia del norte argentino .
- En un 6,45% de los días se produjeron cinco fallecimientos por COVID-19 durante el mes de marzo en una provincia del norte argentino.



### Frecuencias acumuladas: (menor que)

- En 12 días del mes de marzo en una provincia del norte argentino se produjo un fallecimiento o ninguno por COVID-19.
- En el 77,42% de los días del mes de marzo en una provincia del norte argentino se produjeron de tres o menos fallecimientos por COVID-19.

### Frecuencias acumuladas: (mayor que)

- En 19 días del mes de marzo se produjeron dos fallecimientos o mas por COVID-19, en una provincia del norte argentino.
- El 22,58% de los días del mes de marzo hubo cuatro fallecimientos o mas por COVID-19, en una provincia del norte argentino.



c) Representar gráficamente las distribuciones de frecuencias.



Graficos Variables  
Cuantitativas

Serie simple o datos no agrupados:  
**NO** tiene representación grafica

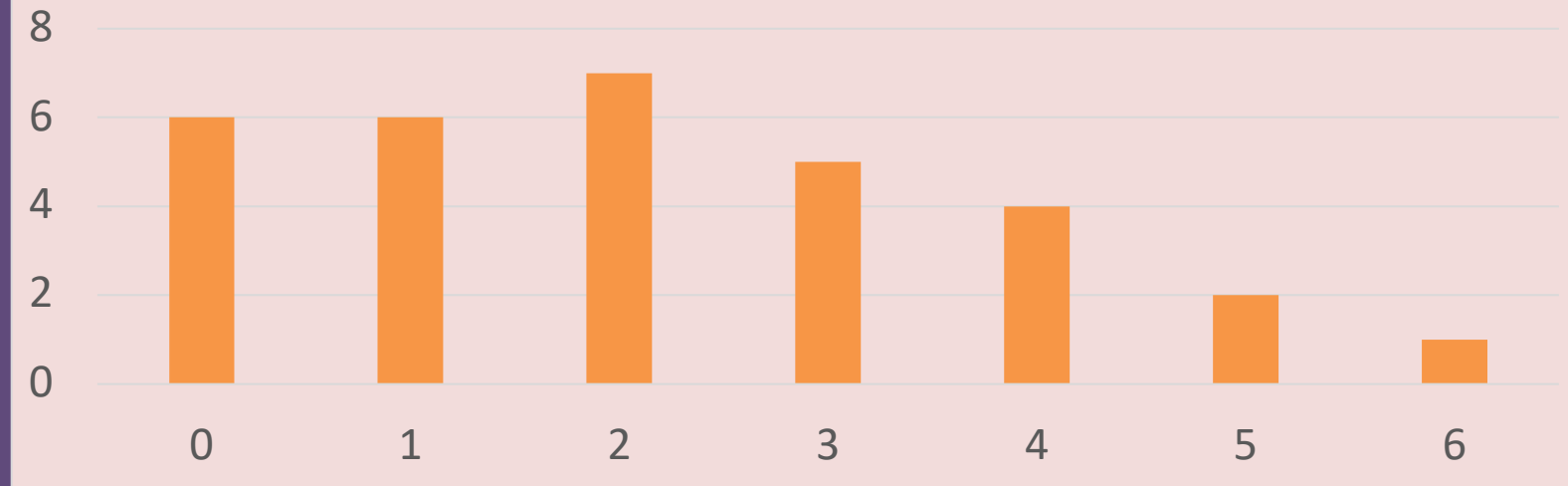
Serie de frecuencias o datos  
agrupados

Variable discreta  
**GRAFICO DE BASTONES**

Variable Continua  
**HISTOGRAMA O POLIGONO DE  
FRECUENCIAS**



Cantidad diaria de fallecimientos por COVID-19  
durante el mes Marzo 2022 en una provincia del  
norte argentino.



Fuente: Aquí colocamos de donde se obtubieron los  
datos y fecha de los mismos

Cantidad de fallecimientos por día en COVID-19 producidos en el mes de Marzo en una provincia del norte argentino

GRAFICO ESCALONADO  
 FRECUENCIAS ACUMULADAS  $F_i$  ↓

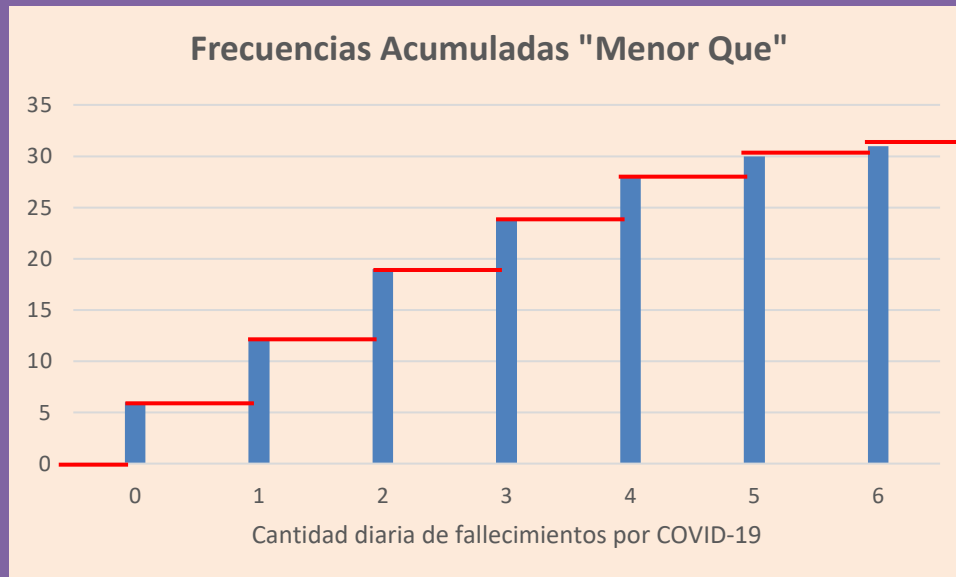
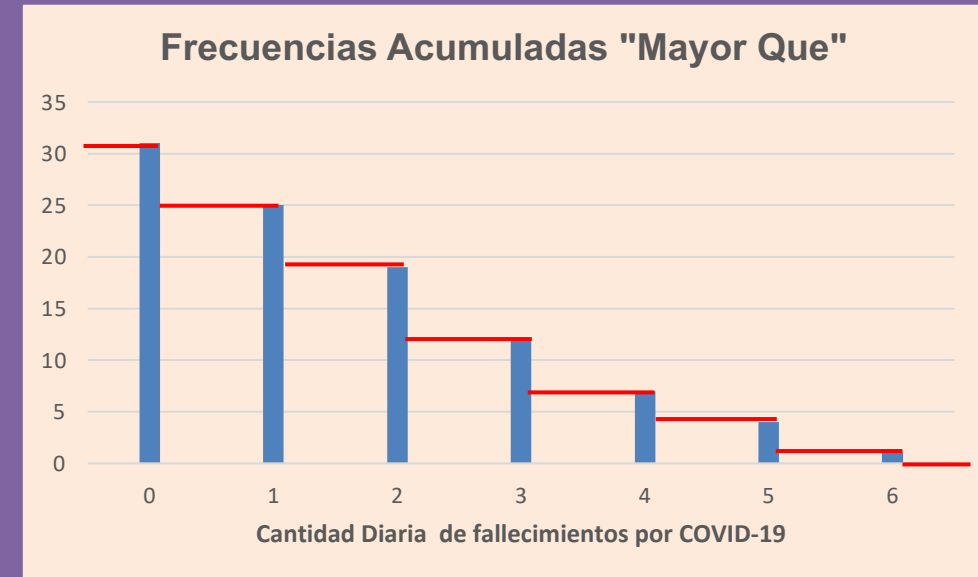


GRAFICO ESCALONADO  
 FRECUENCIAS ACUMULADAS  $F_i$  ↑





**d) ¿Qué porcentaje de días no hubo fallecimientos por COVID-19?**

**En el 19,36% de los días no hubo fallecimientos por COVID-19.**

**e) ¿En cuantos días hubieron mas de tres decesos por COVID-19?**

**En doce días se produjeron tres decesos por COVID-19 o mas.**



f) Calcular las medidas de tendencia central y de dispersión. Interpretar cada una de ellas.

### MEDIDAS DE TENDENCIA CENTRAL

#### Media aritmética: $\bar{x}$ ( $\mu$ )

En símbolos :  $\bar{x} = (\sum_{i=1}^k x_i f_i) / (\sum_i f_i)$   $i = 1, 2, 3, \dots, k$

$$\bar{x} = 67/31 = 2,16$$

#### Interpretación:

En promedio la cantidad de fallecimientos por COVID – 19 que se produjo en el mes de Marzo, en una provincia del norte argentino, es de 2 fallecidos.

## Mediana (de una serie de frecuencias sin intervalos)

Me: 1)  $F_i \downarrow$

2)  $n/2 = 31/2 = 15,5$

3) Buscar la  $F_i \downarrow > 15,5 = 19 \rightarrow x = 2$

### Interpretación:

El 50% de los días del mes de marzo de 2022, se produjeron dos (2) o menos fallecidos por Covid-19 y el restante 50% de los días dos (2) o más fallecidos, en una provincia del norte argentino.

## Moda

$M_o = 2$

Interpretación: El número de fallecidos por COVID -19 en el mes de marzo, en una provincia del norte argentino, que mas se repite es dos (2).

## Medidas de dispersión

Rango o recorrido:  $R = X_{\text{máx.}} - X_{\text{min.}}$

$$R = 6 - 0 = 6$$

La variabilidad máxima en la cantidad de fallecidos por COVID-19 en una provincia del norte argentino, durante el mes de marzo, es de 6 (seis).

## Medidas de dispersión

Rango Intercuartílico:  $RI = Q3 - Q1 = 3 - 1 = 2$

**Q1:** 1)  $F_i$

2)  $1n/4 = (31/4) * 1 = 7,75$   $F_i = 12$

3)  $Q1 = 1$

**Q3:** 1)  $F_i$

2)  $3n/4 = (31/4) * 3 = 23,25$   $F_i = 24$

3)  $Q3 = 3$

En el 50% central de la cantidad de fallecidos por Covid-19 en el mes de marzo en una provincia del norte argentino, tiene una variabilidad de 2 (dos) fallecidos.

Varianza:

$$S^2 = \frac{\sum_i (x_i - \bar{x})^2 f_i}{n-1} = 84,19/30 = 2,8063$$



La dispersión en la cantidad de fallecidos es 3 (tres), en terminos de unidades cuadráticas, con respecto a la media.

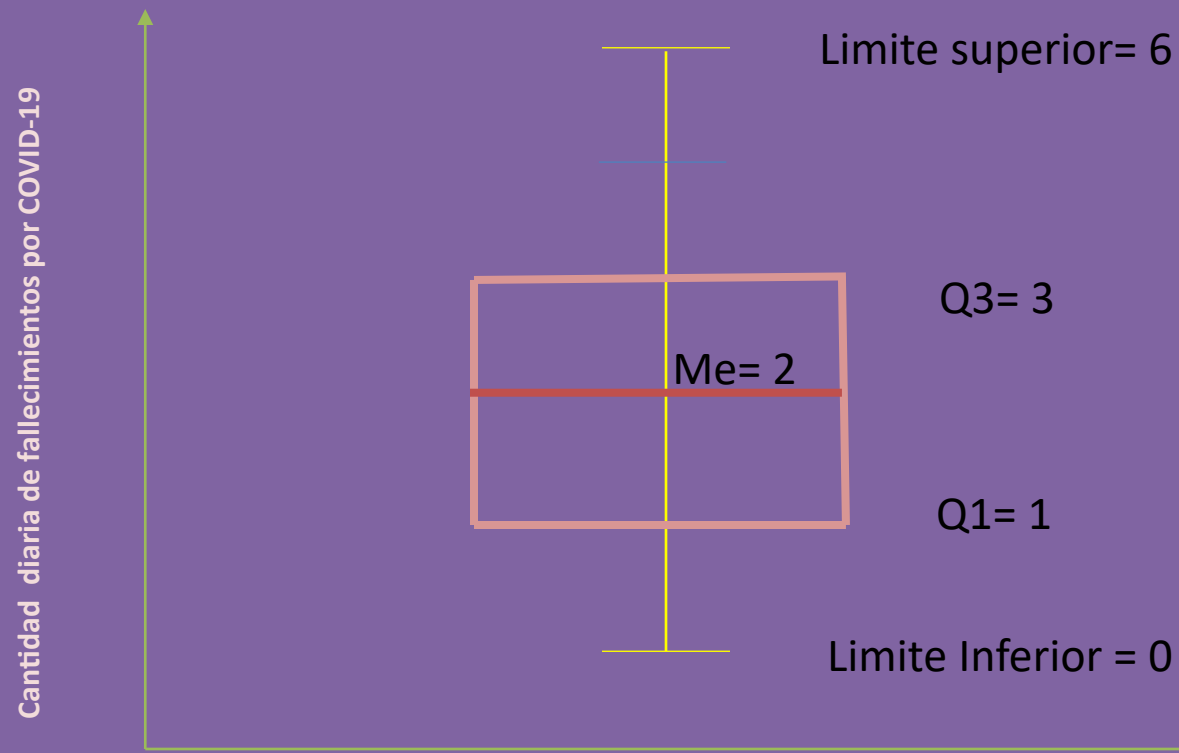
Desviación estándar: 
$$S = \sqrt{\frac{\sum_i (xi - \bar{x})^2 f_i}{n-1}} = \sqrt{2,8063} = 1,6752$$

La dispersion promedio en la cantidad de fallecimientos que se producen por dia, en el mes de marzo en una provincia del norte argentino, es de dos (2) veces por encima y por debajo de la media.

Aclaracion: al tratarse de una variable discreta es correcto interpretar la misma en numeros enteros.



## g) Graficar e interpretar el diagrama de caja (Boxplot).



### INTERPRETACION

En el diagrama de Boxplot o Caja podemos observar que el 50% central de los datos (la caja) no se encuentra en el centro de la distribución, sino levemente corrida hacia los valores mas bajos de la variable, con lo cual tiene una leve distribución asimétrica positiva.

h) El 50% central de la cantidad de fallecimientos diarios por COVID.19 ¿qué variabilidad tiene?

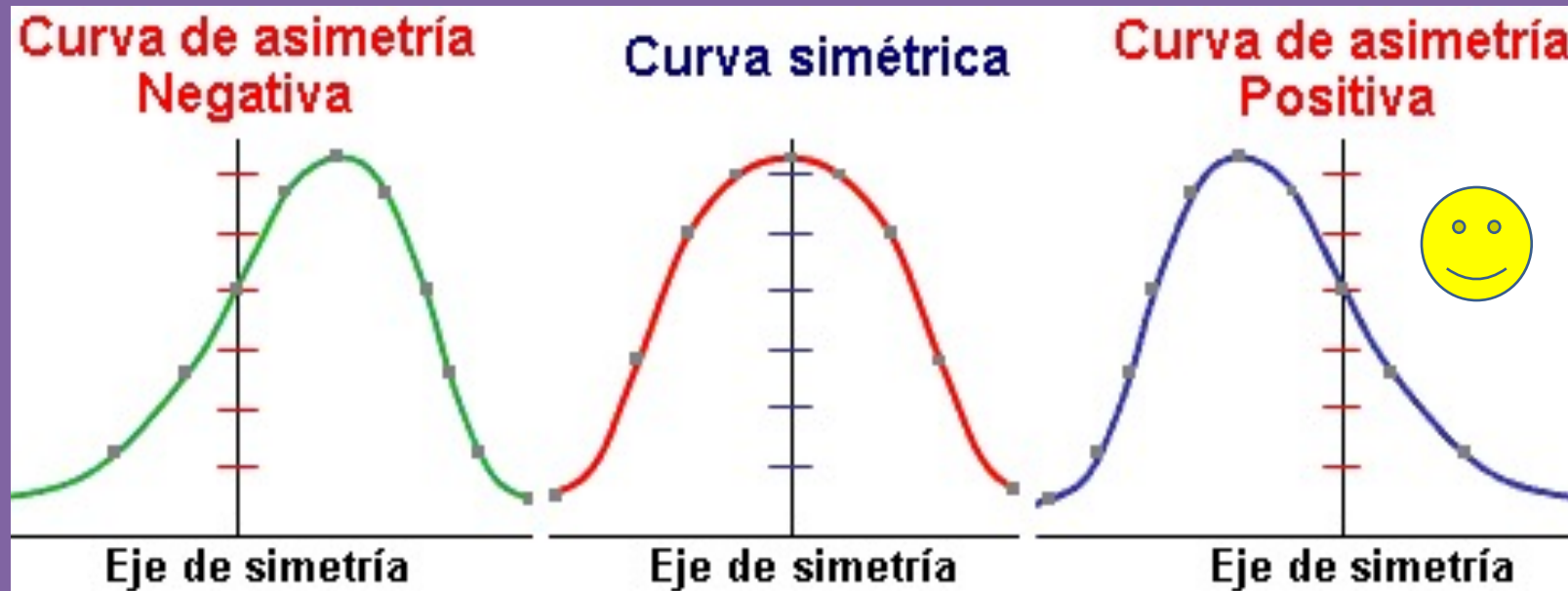
El 50 % central de la cantidad de fallecimientos diarios por COVID-19 tiene una variabilidad de dos (2) fallecimientos.

j) Calcular y comentar los coeficientes de Asimetría y Curtosis.

Coeficiente de asimetria:  $A_s = \frac{\bar{x} - M_o}{s} = (2,16 - 2) / 1,6752 = 0,0952$

\*Campo de variación:  $-1 \leq A_s \leq 1$ \*

La curva de nuestros datos es prácticamente simétrica, con una leve tendencia a una asimetría positiva, influenciada por los días que se produjeron mas de tres muertes diarias por COVID-19. Lo que significa que en la mayoría de los días la cantidad de fallecimientos es de 2 (dos).



Coeficiente de kurtosis:

Kurtosis → Agudeza

$$K = \frac{\frac{1}{2} (Q_3 - Q_1)}{(P_{90} - P_{10})}$$

- $(K \Rightarrow 0)$  la distribución es *Platicúrtica*
- $(K \Rightarrow 0,5)$  la distribución es *Leptocúrtica*
- $(K \Rightarrow 0,25)$  la distribución es *Mesocúrtica*

Calculos auxilires:

P10: 1)  $F_i \downarrow$

$$2) (n/100)*10 = (31/100)*10 = 3,1$$

$$3) F_i \downarrow > 3,1 = 6 \quad \rightarrow \quad P10 = 0$$

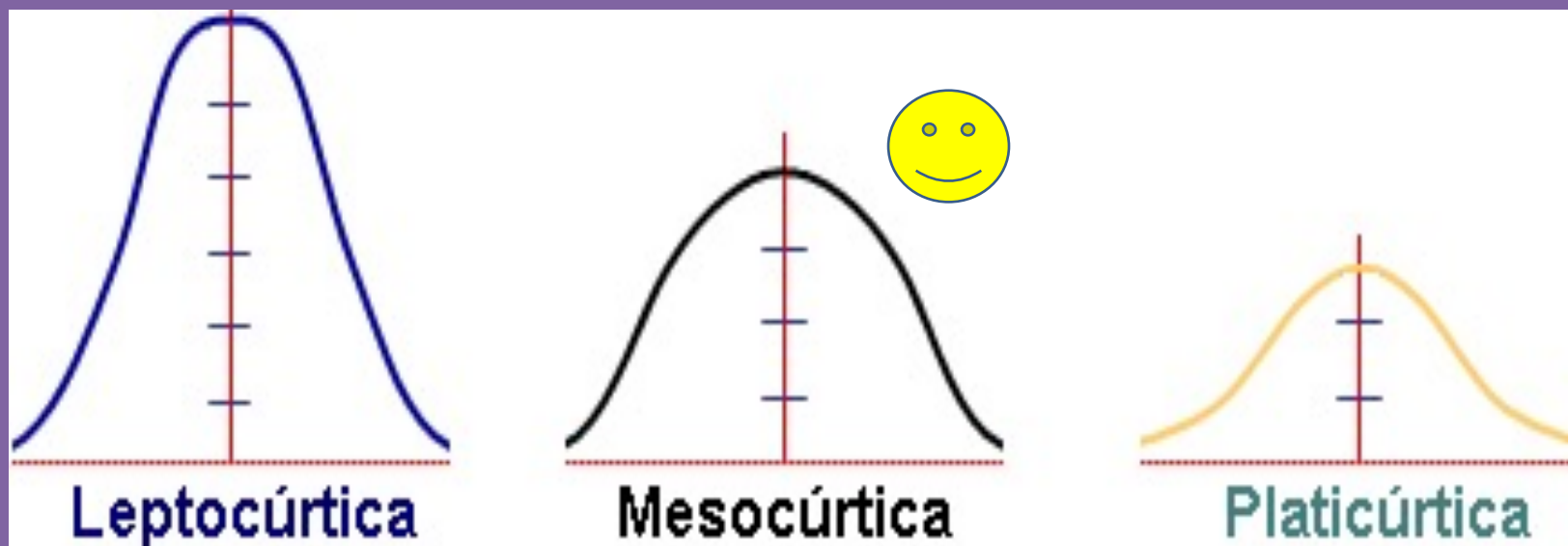
P90: 1)  $F_i \downarrow$

$$2) (n/100)*90 = (31/100)*90 = 27,9$$

$$3) F_i \downarrow > 27,9 = 28 \quad \rightarrow \quad P90 = 4$$

$$K = \frac{\frac{1}{2} (3-1)}{(4-0)} = 0,25$$

( $K \Rightarrow 0,25$ ) la distribución es *Mesocúrtica*



Se determina con el coeficiente de Kurtosis, que presentan un grado de concentración normal los datos obtenidos para el estudio de la cantidad de fallecimientos diarios por Covid-19, que se produjeron durante el mes de marzo en una provincia del norte argentino.

*Muchas Gracias*