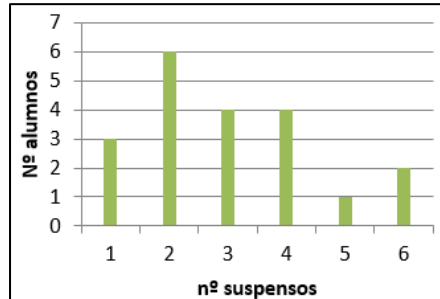


PARTE 1: ESTADÍSTICA UNIDIMENSIONAL Y BIDIMENSIONAL

1. El siguiente diagrama de barras representa el número de suspensos que han tenido los alumnos de 1º bachillerato en la asignatura de matemáticas.



- a) Calcula la tabla de frecuencias
b) Calcula las medidas de centralización.
c) Calcula las medidas de dispersión.
d) ¿Qué porcentaje de alumnos tienen menos de 5 suspensos?
2. Se ha realizado una encuesta a 6 propietarios, preguntando el número de personas que habitan el hogar familiar y el número de habitaciones que tiene la casa, obteniendo la tabla dada:

Nº de personas	3	5	4	6	5	4
Nº de habitaciones	2	4	3	5	5	4

- a) Calcula el coeficiente de correlación, indicando todos los cálculos necesarios e interpretando el resultado.
b) Estima el número esperado de personas que viven en el hogar, si éste tiene 6 habitaciones. ¿Es buena la estimación?
3. De una variable bidimensional (X, Y) sabemos que:
- La recta de regresión de Y sobre X es $y = 16,26 - 2,88x$
 - La covarianza es $\sigma_{xy} = -3,98$.
- Calcula la desviación típica de la variable X.

PARTE 2: PROBABILIDAD

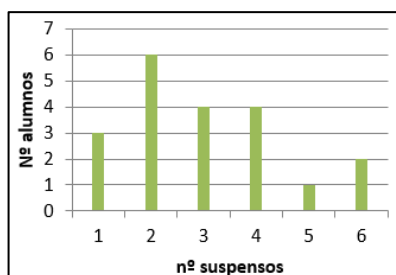
1. En un edificio inteligente dotado de energía solar y eólica, se sabe que la energía suministrada cada día proviene de placas solares con probabilidad 0,4, de molinos eólicos con probabilidad 0,26 y de ambos tipos de instalaciones con probabilidad 0,12. Elegido un día al azar, calcúlese la probabilidad de que la energía sea suministrada al edificio:
 - a) Por alguna de las dos instalaciones.
 - b) Solamente con alguna de las dos.

2. En un cierto país donde la enfermedad VIRIS es endémica, se sabe que un 12% de la población padece dicha enfermedad. Se dispone de una prueba para detectar la enfermedad, pero no es totalmente fiable ya que da positiva en el 90% de los casos de personas realmente enfermas y en el 5% de personas sanas.
 - a) Calcula la probabilidad de que, elegida al azar una persona del país, la prueba de positiva.
 - b) Calcula la probabilidad de que esté sana una persona a la que la prueba le ha dado positivo.

3. La probabilidad de que un tenista obtenga un punto de saque directo es de 0,02.
 - a) Si durante un partido realiza 8 servicios, ¿cuál es la probabilidad de que consiga al menos 2 saques directos?
 - b) Si durante un torneo realiza 3000 servicios, ¿cuál es la probabilidad de que consiga más de 80 saques directos?

PARTE 1: ESTADÍSTICA UNIDIMENSIONAL Y BIDIMENSIONAL

1. El siguiente diagrama de barras representa el número de suspensos que han tenido los alumnos de 1º bachillerato en la asignatura de matemáticas.



- a) Calcula la tabla de frecuencias:

x_i	f_i	F_i	h_i	H_i	%	$x_i \cdot f_i$	$x_i^2 \cdot f_i$
1	3	3	0,15	0,15	15%	3	3
2	6	9	0,3	0,45	30%	12	24
3	4	13	0,2	0,65	20%	12	36
4	4	17	0,2	0,85	20%	16	64
5	1	18	0,05	0,9	5%	5	25
6	2	20	0,1	1	10%	12	72
$N = 20$			1		100%	60	224

- b) Calcula las medidas de centralización.

Calculamos la media: $\bar{x} = \frac{\sum_{i=1}^k x_i f_i}{N} = \frac{60}{20} = 3$ suspensos

La moda es: $Mo = 2$ porque es el valor de mayor frecuencia absoluta, el que más se repite.

Al ser un número par de datos ($N = 20$), la mediana es la media de los dos datos centrales después de ordenarlos de menor a mayor, que corresponden a los situados en las posiciones 10 y 11, que son 3 y 3 respectivamente. Por tanto, la mediana es: $Me = 3$ suspensos.

- c) Calcula las medidas de dispersión.

El rango es la diferencia entre el mayor y el menor dato, es decir: $R = 6 - 1 = 5$

La varianza se calcula de la siguiente forma:

$$\sigma^2 = \frac{\sum_{i=1}^k x_i^2 f_i}{N} - \bar{x}^2 = \frac{224}{20} - 3^2 = 2,2$$

La desviación típica se define como la raíz cuadrada de la varianza: $\sigma = \sqrt{2,2} = 1,48$

El coeficiente de variación es $CV = \frac{\sigma}{\bar{x}} = 0,493 \approx 49,3\%$

- d) ¿Qué porcentaje de alumnos tienen menos de 5 suspensos?

Observando la tabla de frecuencias, podemos concluir que el porcentaje de alumnos con menos de 5 suspensos es 85%.

2. Se ha realizado una encuesta a 6 propietarios, preguntando el número de personas que habitan el hogar familiar y el número de habitaciones que tiene la casa, obteniendo la tabla dada:

Nº de personas	3	5	4	6	5	4
Nº de habitaciones	2	4	3	5	5	4

- a) Calcula el coeficiente de correlación, indicando todos los cálculos necesarios e interpretando el resultado.
Se trata de datos de una variable bidimensional, que relaciona el número de habitaciones (X) y el número de personas que habitan en el hogar (Y) de 6 inquilinos.

El coeficiente de correlación es una medida que indica el grado de dependencia lineal que existe entre ambas variables y toma siempre un valor comprendido entre -1 y 1.

La forma de calcularlo es a partir de la covarianza y de las desviaciones típicas de las variables:

$$r = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$$

Comenzamos calculando la media de habitaciones:

$$\bar{x} = \frac{2 + 4 + 3 + 5 + 5 + 4}{6} = \frac{23}{6} = 3,83 \simeq 4 \text{ habitaciones}$$

La media de personas en el hogar es:

$$\bar{y} = \frac{3 + 5 + 4 + 6 + 5 + 4}{6} = \frac{27}{6} = 4,5$$

Calculamos la covarianza de X y de Y:

$$\begin{aligned} \sigma_{xy} &= \frac{\sum x_i \cdot y_i}{6} - \bar{x} \cdot \bar{y} = \\ &= \frac{2 \cdot 3 + 4 \cdot 5 + 3 \cdot 4 + 5 \cdot 6 + 5 \cdot 5 + 4 \cdot 4}{6} - 3,83 \cdot 4,5 = \frac{109}{6} - 3,83 \cdot 4,5 = 0,9317 \end{aligned}$$

Calculamos las desviaciones típicas. En primer lugar, calculamos la varianza de X y la varianza de Y:

$$\sigma_x^2 = \frac{\sum x_i^2 \cdot f_i}{6} - \bar{x}^2 = \frac{95}{6} - 3,83^2 = 1,1644 ; \text{ Luego, } \sigma_x = \sqrt{1,1644} = 1,07$$

$$\sigma_y^2 = \frac{\sum y_i^2 \cdot f_i}{6} - \bar{y}^2 = \frac{127}{6} - 4,5^2 = 0,917 ; \text{ Luego, } \sigma_y = \sqrt{0,917} = 0,9576$$

Por tanto, el coeficiente de correlación es: $r = \frac{0,9317}{1,07 \cdot 0,9576} = 0,9093$ Hay una dependencia lineal fuerte y positiva entre ambas variables.

- b) Estima el número esperado de personas que viven en el hogar, si éste tiene 6 habitaciones. ¿Es buena la estimación?

La ecuación de la recta de regresión de Y sobre X es la siguiente:

$$y = \frac{\sigma_{xy}}{\sigma_x^2} \cdot x + \bar{y} - \frac{\sigma_{xy}}{\sigma_x^2} \cdot \bar{x}$$

Sustituyendo:

$$y = \frac{0,9317}{1,1644} \cdot x + 4,5 - \frac{0,9317}{1,1644} \cdot 3,83 = 0,8 \cdot x + 1,42$$

Así, para una casa con $x=6$ habitaciones, se esperan: $y = 0,8 \cdot 6 + 1,42 = 6,22 \simeq 6 \text{ personas}$

Como el coeficiente de correlación es un valor cercano a 1, podemos concluir que la estimación es fiable.

3. De una variable bidimensional (X, Y) sabemos que:

- La recta de regresión de Y sobre X es $y = 16,26 - 2,88 x$
- La covarianza es $\sigma_{xy} = -3,98$.

Calcula la desviación típica de la variable X.

La ecuación de la recta de regresión de Y sobre X es la siguiente:

$$y = \frac{\sigma_{xy}}{\sigma_x^2} \cdot x + \bar{y} - \frac{\sigma_{xy}}{\sigma_x^2} \cdot \bar{x}$$

Y sabemos que, la ecuación es: $y = -2,88 \cdot x + 16,26$

Por tanto, identificando los coeficientes:

$$\begin{cases} (1) & \frac{\sigma_{xy}}{\sigma_x^2} = -2,88 \\ (2) & \bar{y} - \frac{\sigma_{xy}}{\sigma_x^2} \cdot \bar{x} = 16,26 \end{cases}$$

Como además $\sigma_{xy} = -3,98$, de la igualdad (1) obtenemos el valor de la desviación típica de X:

$$\frac{-3,98}{\sigma_x^2} = -2,88 \rightarrow \frac{-3,98}{-2,88} = \sigma_x^2 \rightarrow \sigma_x^2 = 1,382 \rightarrow \sigma_x = 1,176$$

PARTE 2: PROBABILIDAD

1. En un edificio inteligente dotado de energía solar y eólica, se sabe que la energía suministrada cada día proviene de placas solares con probabilidad 0,4, de molinos eólicos con probabilidad 0,26 y de ambos tipos de instalaciones con probabilidad 0,12. Elegido un día al azar, calcúlese la probabilidad de que la energía sea suministrada al edificio:

a) Por alguna de las dos instalaciones.

b) Solamente con alguna de las dos.

Definimos A="energía solar"; B="energía eólica".

Sabemos que $p(A) = 0,4$ $p(B) = 0,26$ $p(A \cap B) = 0,12$

a) $p(A \cup B) = p(A) + p(B) - p(A \cap B) = 0,4 + 0,26 - 0,12 = 0,54$

b) $p((A \cap \bar{B}) \cup (B \cap \bar{A})) = p(A \cup B) - p(A \cap B) = 0,54 - 0,12 = 0,42$

2. En un cierto país donde la enfermedad VIRIS es endémica, se sabe que un 12% de la población padece dicha enfermedad. Se dispone de una prueba para detectar la enfermedad, pero no es totalmente fiable ya que da positiva en el 90% de los casos de personas realmente enfermas y en el 5% de personas sanas.

a) Calcula la probabilidad de que, elegida al azar una persona del país, la prueba de positiva.

b) Calcula la probabilidad de que esté sana una persona a la que la prueba le ha dado positivo.

Definimos E="padecer la enfermedad"; P="prueba positiva"

Del enunciado sabemos que: $p(E) = 0,12 \rightarrow p(\bar{E}) = 1 - 0,12 = 0,88$; $p(P/E) = 0,9$; $p(P/\bar{E}) = 0,05$

a) Utilizamos el Teorema de la probabilidad total:

$$p(P) = p(E) \cdot p(P/E) + p(\bar{E}) \cdot p(P/\bar{E}) = 0,12 \cdot 0,9 + 0,88 \cdot 0,05 = 0,152$$

b) Utilizamos el Teorema de Bayes:

$$p(\bar{E}/P) = \frac{p(\bar{E}) \cdot p(P/\bar{E})}{p(P)} = \frac{0,88 \cdot 0,05}{0,152} \cong 0,29$$

3. La probabilidad de que un tenista obtenga un punto de saque directo es de 0,02.

a) Si durante un partido realiza 8 servicios, ¿cuál es la probabilidad de que consiga al menos 2 saques directos?

b) Si durante un torneo realiza 3000 servicios, ¿cuál es la probabilidad de que consiga más de 80 saques directos?

a) Llamamos X a la variable "número de saques directos", se trata de una binomial $X \hookrightarrow B(8, 0'02)$.

La función de probabilidad asociada a X es:

$$p(X = k) = \binom{8}{k} \cdot 0,02^k \cdot 0,98^{8-k} \quad k = 0, 1, \dots, 8$$

Nos piden:

$$p(X \geq 2) = 1 - p(X < 2) = 1 - [p(X = 0) + p(X = 1)] = 1 - \left[\binom{8}{0} \cdot 0'02^0 \cdot 0'98^8 + \binom{8}{1} 0,02 \cdot 0'98^7 \right] \cong 0,01034$$

b) En este caso $X \hookrightarrow B(3000, 0'02)$. Como $np(1-p) > 10$ se puede aproximar por una normal $N(3000 \cdot 0,02, \sqrt{3000 \cdot 0'02 \cdot 0'98})$ y operando tenemos $N(60, 7,67)$.

Por tanto:

$$\begin{aligned} p(X > 80) &= 1 - p(X \leq 80) = 1 - p(X \leq 80'5) = 1 - p\left(Z \leq \frac{80'5 - 60}{7'67}\right) = 1 - p(Z \leq 2'67) \\ &= 1 - 0'9962 = 0'0038 \end{aligned}$$