

2. CORRELACION Y REGRESIÓN

2.1. CORRELACIÓN

2.1.1. Introducción

En los tratados anteriores hemos estudiado diferentes temas de estadística atendiendo sólo a una variable, y hemos dado procedimientos para el cálculo de estadísticos de muestras que permiten describir e interpretar la distribución de los valores de esa variable. Pero los requerimientos de las ciencias van más allá del comportamiento de una variable, y en muchos de sus problemas necesitan investigar la relación entre dos o más variables. Por ejemplo, ¿existe alguna relación entre la altura de los árboles y el diámetro de sus troncos a cierto nivel del suelo? ¿Existe alguna relación entre el grado de oscuridad nocturna, la temperatura ambiente y el apetito de los ratones? ¿Hay alguna relación entre las calificaciones obtenidas por los estudiantes en sus exámenes de admisión a la universidad y su rendimiento académico en ella? ¿Existe alguna relación entre la alimentación de las personas y las afecciones cardíacas? Si hay relaciones de asociación entre las variables, entonces ¿será posible hacer estimaciones del valor de una de ellas conociendo el valor de la otra variable?. Precisamente éste es el tema central de la correlación.

Para investigar la relación entre dos variables necesitamos estudiar distribuciones bidimensionales, esto es por lo tanto el nivel de relación que existe entre las variables.

2.1.2. CLASES DE CORRELACIÓN.

2.1.2.1. Por la relación entre las variables.

2.1.2.1.1. Lineal:

Relación que une a las variables puede ser representado por una línea recta.

2.1.2.1.2. No lineal:

Relación que une a las variables pero no puede ser representada por una línea recta.

2.1.2.2. Fenómenos que relacionan:

2.1.2.2.1. Simple:

Cuando a cada variable dependiente corresponde una independiente.

2.1.2.2.2. Múltiple:

Cuando a cada variable dependiente corresponde varias independientes.

Ejemplo.

El rendimiento en cada uno de las materias en el paro del año

2.1.2.3. Por la cantidad existente entre los fenómenos que relacionan

2.1.2.3.1. Perfecta:

Cuando la variable dependiente e independiente tienen variaciones equivalentes.

2.1.2.3.2. Imperfecta:

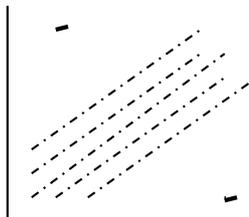
Cuando las variaciones de las dos variables no son equivalentes .

2.1.2.3.3. Nula:

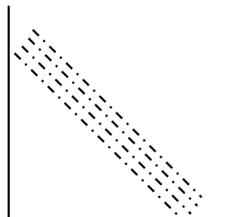
Cuando no existe ningún tipo de relación en las variaciones de las variables .

2.1.3. Interpretación de la correlación.

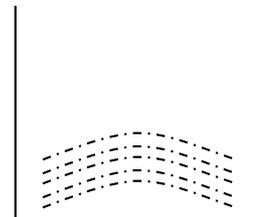
Para interponer el diagrama de dispersión nos basamos en los siguientes gráficos.



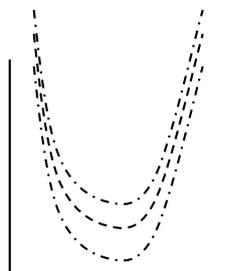
SIMPLE
LINEAL
POSITIVA



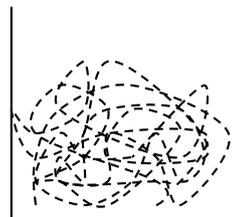
SIMPLE
LINEAL
NEGATIVA



SIMPLE NO
LINEAL



SIMPLE NO
LINEAL

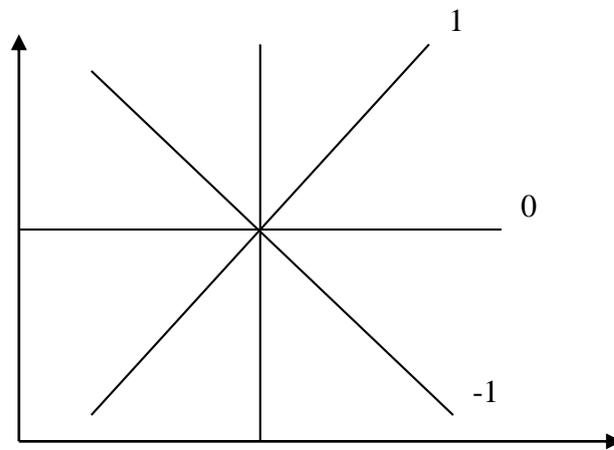
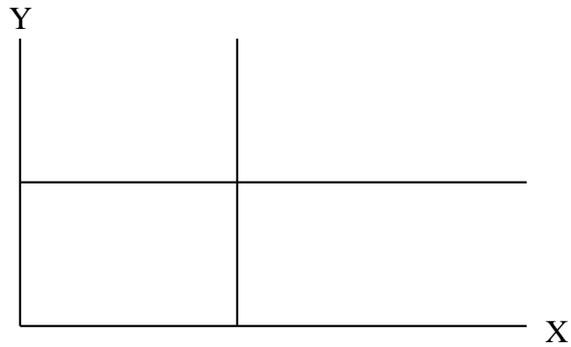


SIMPLE NO
LINEAL Y
NULA

2.1.4. Pasos para la interpretación de una correlación.

2.1.4.1. Graficar la variable en un diagrama de dispersión.

- 2.1.4.2. Obtener los promedios o media aritmética de cada variable
- 2.1.4.3. Graficar los promedios en el diagrama de dispersión de manera que se genere un nuevo plano cartesiano.
- 2.1.4.4. Sacar los porcentajes de puntos ubicados en cada cuadrante de estenuevo plano.
- 2.1.4.5. Luego de realizar los cálculos del coeficiente e interpretarlo.



Como base para poder interpretar el r , nos apoyaremos en el siguiente cuadro de referencia

r =Coeficiente de Correlación

Va de + 1 a - 1

	1		Perfecta positiva
0,90	a	0,99	Muy alta positiva
0,70	a	0,89	Alta positiva
0,40	a	0,69	Moderada positiva
0,20	a	0,39	Baja positiva
0,01	a	0,19	Muy baja positiva
	0		
-0,19	a	-0,01	Muy baja negativa
-0,39	a	-0,20	Baja negativa
-0,69	a	-0,40	Moderada negativa
-0,89	a	-0,70	Baja negativa
-0,99	a	-0,90	Muy baja negativa
	-1		Perfecta negativa

EJERCICIO

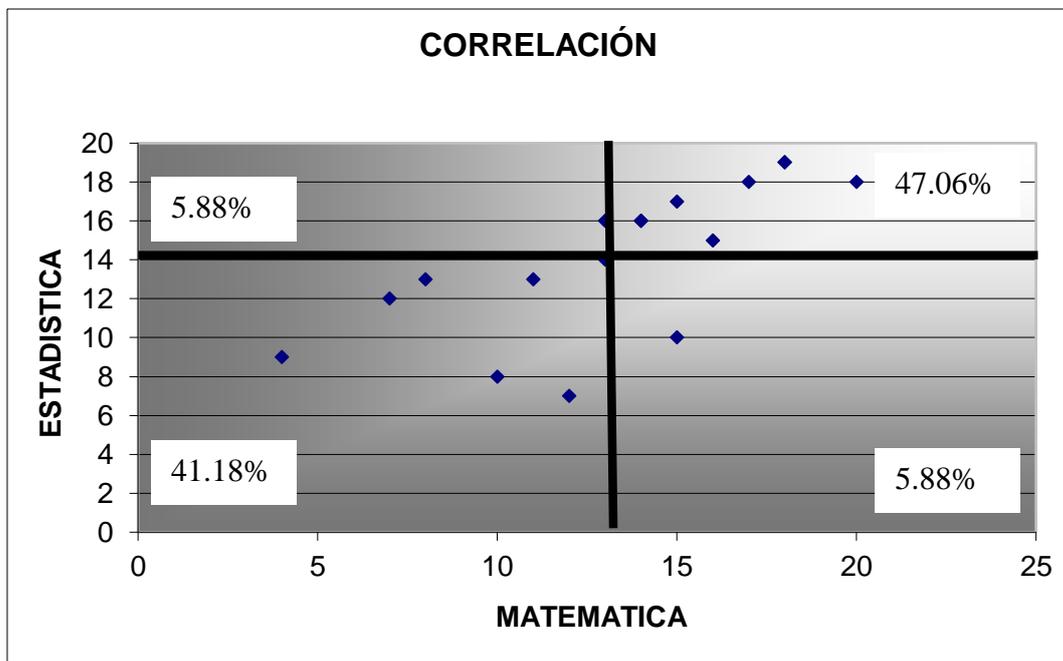
Calcular la correlación existente entre las calificaciones obtenidas por los estudiantes de un curso en dos diferentes materias.

Se recomienda ordenar los datos de acuerdo a la primer variable

Est.	MAT X	EST Y	Est.	MAT X	EST Y
A	15	10	C	20	18
B	18	19	B	18	19
C	20	18	O	18	19
D	13	16	M	17	18
E	10	8	I	16	15
F	7	12	N	15	17
G	4	9	A	15	10
H	8	13	K	14	16
I	16	15	L	14	16
J	13	14	D	13	16
K	14	16	J	13	14
L	14	16	Q	12	7
M	17	18	P	11	13
N	15	17	E	10	8
O	18	19	H	8	13
P	11	13	F	7	12
Q	12	7	G	4	9
				225	240
				ΣX	ΣY

$$\bar{X} = \frac{\sum X}{n} = \frac{225}{17} = 13.24$$

$$\bar{Y} = \frac{\sum Y}{n} = \frac{240}{17} = 14.12$$



INTERPRETACIÓN.

La correlación existente entre la nota de mate. y estadística, los estudios el 17 es simple lineal positivo donde se demuestra que 47.06% de estudiantes no manifiesta ningún tipo de novedad d5, tiene problema en matemáticas y no en estadística, el 41.18% en atención especial porque se encuentran por debajo de los promedios en ambas asignaturas, el otro 5.88% restante, manifiesta dificultades en Estadística y no en Matemáticas.