UNIVERSIDAD NACIONAL DE JAÉN

UNIVERSIDAD NACIONAL DE JAÉN

DEPARTAMENTO ACADÉMICO DE CIENCIAS BÁSICAS Y

APLICADAS

INGENERÍA MECÁNICA Y ELÉCTRICA

Manual de correlación y regresión Lineal

Autores:

Dra. Rosario Yaqueliny Llauce Santamaria

Dra. Marcela Yvone Saldaña Miranda

Mg. Mario Félix Olivera Aldana

JAÉN – PERÚ, FEBRERO 2024

Índice

Índice de Tablas

Introducción

El coeficiente de correlación y la regresión lineal examinan la relación entre dos variables continuas X e Y. El coeficiente de correlación mide el grado de asociación lineal entre X e Y sin asumir una dirección específica en la relación entre ambas variables. Por otro lado, la regresión lineal simple cuantifica cómo cambia el valor promedio de la variable Y(dependiente o respuesta) en función de los cambios en la variable X(independiente o explicativa)

En este manual se incluyen ejemplos prácticos y aplicaciones en diversos campos, como la física y la ingeniería, para resolver situaciones problemáticas la vida real. Se discutirán en detalle las aplicaciones de ingeniería, particularmente en ingeniería mecánica y eléctrica. Estos campos hacen un uso extensivo de la correlación para analizar el rendimiento de los componentes y evaluar la eficiencia energética, así como la regresión para modelar el comportamiento de los materiales y optimizar los procesos de fabricación. Por ejemplo, la correlación se puede utilizar para comprender cómo los cambios en la corriente eléctrica afectan la temperatura de un motor, mientras que la regresión puede ayudar a predecir cómo se comportará un material en determinadas condiciones de tensión y temperatura. Finalmente, incluiremos una sección de ejercicios aplicados que le permitirá practicar y reforzar sus conocimientos, resolver problemas del mundo real, reforzar conceptos teóricos y demostrar la versatilidad y el poder predictivo de estos métodos estadísticos.

Los Autores

1. Correlación

La correlación tiene como objetivo evaluar la dirección y la fuerza de la asociación entre dos variables cuantitativas. Esto nos permite entender la intensidad de su relación y determinar si el aumento en el valor de una variable está asociado con un incremento o una disminución en el valor de la otra variable.

Para evaluar la asociación entre dos variables, la aproximación inicial generalmente se realiza mediante un diagrama de dispersión

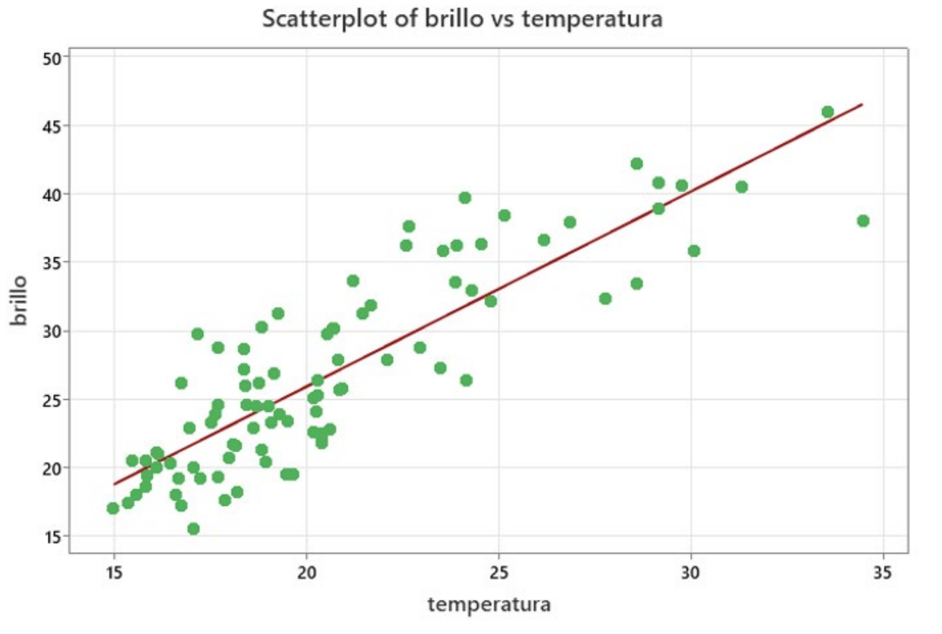

Figura 1. Scatterplot of brillo vs temperatura

En el diagrama de dispersión se observa una posible relación lineal entre la temperatura y el Scatterplot.

La nube de puntos proporciona una visualización inicial de la relación entre dos variables. Sin embargo, para medir la fuerza y dirección de esta asociación, es necesario calcular un coeficiente de correlación que cuantifique la relación entre las variables.

2. Coeficiente de Correlación

Existen dos coeficientes de correlación: el coeficiente de Pearson (paramétrico) se emplea cuando las variables cumplen con los criterios de normalidad y evalúa específicamente la adecuación de la relación lineal entre dos variables cuantitativas y el coeficiente de Spearman (no paramétrico) se usa cuando las variables no cumplen con los criterios de normalidad o son ordinales, y mide cualquier tipo de asociación, no necesariamente lineal.

3. Coeficiente de Correlación de Pearson

También conocido como coeficiente de correlación lineal, evalúa la fuerza y la dirección de la relación lineal entre dos variables cuantitativas. Este coeficiente, representado por la letra “r”, oscila entre -1 y 1, donde:

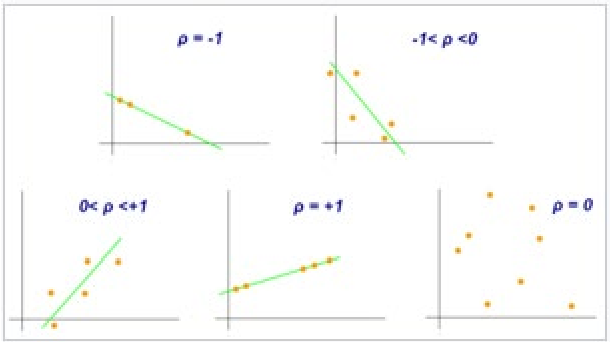

Figura 2. Diagramas de dispersión

Valores y su Interpretación:

- r =1 correlación positiva perfecta, a medida que una variable aumenta, la otra también lo hace en una relación lineal exacta.
- r = −1 correlación negativa perfecta, Cuando una variable aumenta, la otra disminuye de manera perfectamente lineal.
- r = 0 no hay correlación lineal, no existe una relación lineal entre las variables, aunque podría haber otro tipo de relación (no lineal).
- **Valores positivos (0 < r< 1):** Correlación lineal positiva. A medida que una variable aumenta, la otra también tiende a aumentar, pero no de manera perfecta. El valor de r indica la fuerza de la asociación: valores cercanos a 1 sugieren una relación lineal fuerte y positiva, mientras que valores cercanos a 0 indican una relación lineal más débil

- **Valores negativos (-1 < r < 0):** Correlación lineal negativa. A medida que una variable aumenta, la otra tiende a disminuir. Valores cercanos a -1 indican una relación lineal fuerte y negativa, mientras que valores cercanos a 0 indican una relación lineal más débil

Y se calcula usando:

$$r = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\sqrt{[n\sum x_i^2 - (\sum x_i)^2][n\sum y_i^2 - (\sum y_i)^2]}}$$

Donde:

- n = es el número de pares de datos.
- $\sum xy$ suma del producto de los pares de datos.
- $\sum x$ suma de los valores de la variable x
- $\sum y$ suma de los valores de la variable y
- $\sum x^2$ suma de los cuadrados de los valores de x
- $\sum y^2$ suma de los cuadrados de los valores de y

4. Covarianza

Es una medida estadística que indica la dirección de la relación lineal entre dos variables cuantitativas. Mide cómo varían juntas dos variables respecto a sus medias. A diferencia del coeficiente de correlación, que estandariza la medida, la covarianza proporciona una medida en las unidades de los variables originales.

La fórmula de covarianza entre dos variables X e Y es:

$$Cov(X,Y) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

Donde:

- X_i e Y_i son valores individuales de las variables X e Y
- $\bar{X}$ y $\bar{Y}$ corresponde a las medias de las variables X e Y, respectivamente.

- n es el número de pares de datos

Interpretación de la covarianza:

- **Covarianza positiva**: Indica que las dos variables tienden a aumentar o disminuir juntas.
- **Covarianza negativa**: una variable tiende a aumentar mientras que la otra disminuye.
- **Covarianza cercana a cero**: no hay una relación lineal clara entre las dos variables.

Asimismo, tener en cuenta que la covarianza no proporciona una medida de la fuerza de la relación, ya que su valor depende de las unidades de las variables. Por esta razón, se suele usar el coeficiente de correlación, que estandariza la covarianza para ofrecer una medida de la fuerza y dirección de la relación.

5. Regresión Lineal simple

Es una técnica estadística utilizada para modelar y analizar la relación entre dos variables cuantitativas: independiente (X) y dependiente (Y). Su objetivo es encontrar la mejor línea recta que prediga el valor de Y en función de X

Modelo de Regresión Lineal Simple

La fórmula general es:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Donde:

y: V. dependiente

x: V. independiente

β_0 intersección o el término constante

β_1 pendiente de la línea de regresión, indica el cambio en Y que no se explica por X.

ε término de error, que captura la variabilidad en Y que no se explica por X

6. Ejercicios Desarrollados

Ejemplo 1:

Se tienen los siguientes datos sobre las horas de estudio y las calificaciones en un examen para 5 estudiantes:

Tabla 1

Datos de Horas de Estudio (X) y Calificación (Y)

Estudiante	Horas de Estudio (X)	Calificación (Y)
1	2	65
2	3	70
3	5	85
4	1	55
5	4	80

Determinar el coeficiente de correlación de Pearson entre las horas de estudio y las calificaciones:

Calculamos:

- Media de X = $\frac{2+3+5+1+4}{5}$= 3

- Media de Y= $\frac{65+70+85+55+80}{5}$= 71

1. **Desviaciones estándar:**

- Para X:

$$\boldsymbol{\sigma_X} = \sqrt{\frac{(\mathbf{2-3})^2 + (\mathbf{3-3})^2 + (\mathbf{5-3})^2 + (\mathbf{1-3})^2 + (\mathbf{4-3})^2}{\mathbf{5}}}$$

$$= \mathbf{1.414}$$

- Para Y:

$$\boldsymbol{\sigma_y} = \sqrt{\frac{\mathbf{(65-71)^2+(70-71)^2+(85-71)^2+(55-71)^2+(80-71)^2}}{\mathbf{5}}} = \mathbf{10.68}$$

2. **Covarianza:**

$$\boldsymbol{Cov(X,Y)}$$

$$= \frac{\mathbf{(2-3)(65-71)+(3-3)(70-71)+(5-3)(85-71)+(1-3)(55-71)+(4-3)(80-7}}{\mathbf{5}}$$

$$= \mathbf{12.6}$$

3. **Coeficiente:**

$$\boldsymbol{r} = \frac{\boldsymbol{Cov(X,Y)}}{\boldsymbol{\sigma X \sigma Y}} = \mathbf{0.83}$$

Conclusión:

El coeficiente es aproximadamente 0.83, lo que muestra una fuerte correlación positiva entre las horas de estudio y las calificaciones.

Ejemplo 2 (Teoría):

Se ha encontrado un coeficiente de correlación de Pearson de -0.65 entre el número de horas que los empleados pasan en redes sociales durante el trabajo y su rendimiento en una evaluación de productividad.

¿qué implica este valor del coeficiente?

Solución: El coeficiente -0.65 expresa una relación negativa moderada del número de horas que los empleados pasan en redes sociales y su rendimiento en la evaluación de productividad. En otras palabras, más tiempo pasa en las redes sociales, es probable que tu rendimiento en tareas relacionadas con la productividad baje.

Ejemplo 3:

Imaginemos que queremos estimar cuánto costará una casa según su tamaño en metros cuadrados. Contamos con la siguiente información:

Tabla 2

Datos de Tamaño (m^2) y Precio (USD)

Tamaño (m^2	*Precio (USD)*
50	150,000
60	180,000
70	210,000
80	240,000
90	270,000

Queremos ajustar una regresión lineal simple para determinar el precio y basado en el tamaño x.

Solución:

1. **Definir el modelo:**

   Fórmula , modelo de regresión lineal simple:

$$y = \beta_0 + \beta_1 x + \epsilon$$

   Donde:

   β_0 intersección

   β_1 pendiente, y ϵ es el error.

2. **Calcular los parámetros:**

   Primero, necesitamos calcular la media de X y Y:

$$x = \frac{50 + 60 + 70 + 80 + 90}{5} = 70$$

$$\mathrm{y} = \frac{150{,}000 + 180{,}000 + 210{,}000 + 240{,}000 + 270{,}000}{5} = 210{,}000$$

Luego, calculamos la pendiente β_1 y la intersección β_0:

$$\beta_1 = \frac{\sum_{i=1}^{n}(x_i - x)(y_i - y)}{\sum_{i=1}^{n}(x_i - x)^2}$$

$$\beta_1 = \frac{(50-70)(150{,}000-210{,}000)+(60-70)(180{,}000-210{,}000)+(70-70)(210{,}000-210{,}000)}{(50-70)^2+(60-70)^2+(70-70)^2+(80-70)^2+(90-70)^2} \ldots$$

$$\ldots\frac{+(80-70)(240{,}000-210{,}000)+(90-70)(270{,}000-210{,}000)}{(50-70)^2+(60-70)^2+(70-70)^2+(80-70)^2+(90-70)^2} = 3000$$

Para β_0:

$$\beta_0 = \bar{y} - \beta_1\bar{x} = 210{,}000 - 3{,}000 \cdot 70 = 210{,}000 - 210{,}000 = 0$$

Así que la ecuación de la regresión es:

$$\mathrm{y} = 0 + 3{,}000\mathrm{x}$$

Interpretación:

La pendiente $\beta_1 = 3000$ indica que, por cada metro cuadrado adicional, el precio de la casa aumenta en 3,000 USD. La intersección $\beta_0 = 0$ no tiene sentido práctico en este caso, ya que un tamaño de 0 m² no corresponde a una casa real, pero es parte del modelo matemático.

Ejemplo 4: Regresión Lineal Múltiple

Se desea averiguar el precio de una casa teniendo en cuenta dos factores: su tamaño en metros cuadrados x_1 y número de habitaciones x_2. Los datos son los siguientes:

Tabla 3

Datos de Tamaño (m²), Habitaciones y Precios (USD)

Tamaño (m²)	Habitaciones	Precios (USD)
50	2	150,000
60	3	180,000
70	3	210,000
80	4	240,000
90	4	270,000

Queremos ajustar un modelo de regresión lineal múltiple.

Solución:

Definir el modelo:

El modelo de regresión lineal múltiple es:

$$y = \beta_0 + \beta_1 x + \epsilon$$

Calcular los parámetros:

Los cálculos exactos son más complejos y generalmente se realizan usando software estadístico. El proceso incluye resolver el sistema de ecuaciones de los parámetros.

Aquí presentamos una forma simplificada de hacerlo con software o herramientas estadísticas (como R, Python, o incluso Excel). Sin embargo, el método manual involucra matrices y álgebra lineal, lo cual puede ser extenso para detallar aquí.

Interpretación:

Supongamos que obtienes los siguientes parámetros:

$$\beta_0 = 500{,}000$$

$$\beta_1 = 2{,}000 (pormetrocuadrado)$$

$$\beta_2 = 10{,}000 (por\ habitacion)$$

Entonces la ecuación sería:

$$y = 50{,}000 + 2{,}000x_1 + 10{,}000x_2$$

Esto indica que el precio base de la casa es 50,000 USD. Cada metro cuadrado adicional aumenta el precio en 2,000 USD, y cada habitación adicional aumenta el precio en 10,000 USD.

5. Aplicaciones de la Correlación y Regresión en la Ingeniería Mecánica y Eléctrica

5.1. Correlación

Análisis de Desempeño de Componentes

El análisis de desempeño de componentes en sistemas mecánicos y eléctricos frecuentemente utiliza la correlación para identificar relaciones entre variables críticas. Por ejemplo, en un motor eléctrico, la correlación entre la corriente de entrada y la temperatura del motor puede revelar cómo varía el rendimiento con respecto al calor generado (Montgomery & Runger, 2018)

Evaluación de la Eficiencia Energética

La eficiencia energética en sistemas eléctricos puede ser evaluada mediante técnicas de correlación. Por ejemplo, la relación entre el voltaje y el consumo de energía en distintos dispositivos puede ser analizada para optimizar su uso y reducir pérdidas energéticas ((Chen & Chen, 2020).

5.2. Regresión

Modelado de Comportamiento de Materiales

En ingeniería mecánica, la regresión se utiliza para modelar el comportamiento de materiales bajo diferentes condiciones de estrés y temperatura. Los modelos de regresión pueden predecir cómo un material se deformará bajo ciertas cargas, lo cual es crucial para el diseño de estructuras seguras y eficientes (Callister & Rethwisch, 2018)

Optimización de Procesos de Manufactura

La regresión es un método clave usado en optimizar procesos de manufactura. Permite desarrollar modelos predictivos para ajustar parámetros de proceso, como la velocidad de corte y la alimentación, para minimizar defectos y maximizar la eficiencia (Groover, 2019)

Análisis de Fallas

Este análisis en sistemas eléctricos y mecánicos puede beneficiarse del uso de la regresión. Por ejemplo, la regresión logística puede emplearse para modelar la probabilidad de falla de componentes en función: la edad del componente, las condiciones de operación y el historial de mantenimiento (Jardine, Lin,, & Banjevic, 2006)

5.2.1. Regresión Múltiple y Sistemas Complejos

Predicción de Vida Útil de Componentes

La regresión múltiple se utiliza para determinar la vida útil de componentes en sistemas complejos, considerando múltiples variables influyentes. Este tipo de análisis es importante para el mantenimiento predictivo y la gestión de activos (Meeker & Escobar, 1998).

Modelos de Automatización y Control

En este modelo la regresión múltiple puede ser utilizada para crear modelos que explican la relación entre diversas variables de entrada y salida. Estos modelos son fundamentales para diseñar controladores que aseguren el rendimiento óptimo del sistema (Ogata, 2010)

6. Ejercicios aplicados a la Carrera
En una fábrica, se desea analizar el vínculo que tiene el esfuerzo con la deformación de cierto material que desea trabajar en factoría

Ejemplo 5
Tabla 4
Datos de Material, Esfuerzo (psi) y Deformación (mm)

Material	1	2	3	4	5	6	7	8	9	10	11
Esfuerzo (psi)	35	34	37	36	35	36	35	33	37	36	36
Deformación (mm)	2	0	3	4	3	4	3	1	4	3	5

Utilizaremos a continuación el diagrama de dispersión:

Diagrama de dispersión

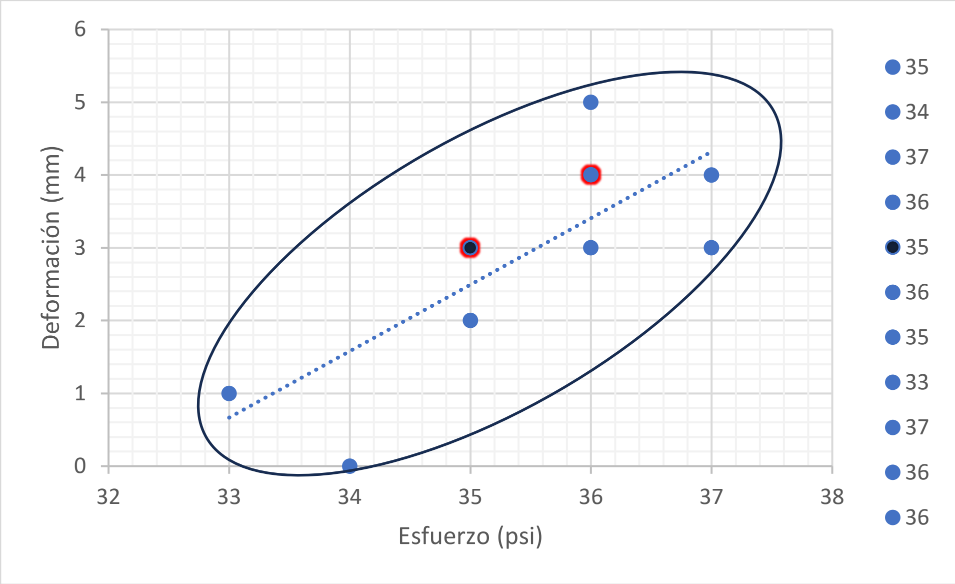

Figura 3

Tenemos una tendencia positiva. Por ende, la relación es positiva.

Calculamos y/o Trazamos la Línea de Regresión, junto con ello el Coeficiente de Regresión.

Formula a emplear: $Y_c = a + bX$, Donde:

- ✓ Y_c: Y dado un valor de X.
- ✓ a : Punto de intersección con el eje y cuando x=0
- ✓ b : Dependencia de la línea de regresión (coeficiente de regresión)

$$a = \bar{Y} - b\bar{X} \qquad \bar{Y} = \frac{\sum Y}{n} \qquad \bar{X} = \frac{\sum X}{n} \qquad b = \frac{n(\sum xy) - (\sum x)(\sum y)}{n \sum x^2 - (\sum x)^2}$$

Ejemplo 6

Tabla 5

Cálculo de $(X_i)^2$

Material	X	Y	$X_i * Y_i$	$(X_i)^2$
1	35	2	70	1225
2	34	0	0	1156
3	37	3	111	1369
4	36	4	114	1296
5	35	3	105	1225

6	36	4	144	1296
7	35	3	105	1225
8	33	1	33	1089
9	37	4	148	1369
10	36	3	108	1296
11	36	5	180	1296
Suma	390	32	1118	13842

Ahora reemplazamos en $b = \frac{n(\sum xy)-(\sum x)(\sum y)}{n\sum x^2-(\sum x)^2}$

$$b = \frac{11(1118) - (390)(32)}{11(13842) - (390)^2}$$

$$b = -\frac{91}{81} = -1.12$$

Remplazamos en $a = \bar{Y} - b\bar{X}$

$$a = \frac{\sum Y}{n} - b\frac{\sum X}{n}$$

$$a = \frac{32}{11} - b\frac{390}{11}$$

$$a = 2.909 - (-1.12)(35.454)$$

$$a = 42.617$$

Remplazamos en $Y_c = a + bX$

$$Y_c = 42.617 - 1.12X$$

Ahora graficamos la línea de regresión:

Calculamos Y_c , para X=33 y X=37

Ejemplo 7

Tabla 6

Datos de X y Y_c

X	Y_c
33	5.657
37	1.177

Diagrama de dispersión

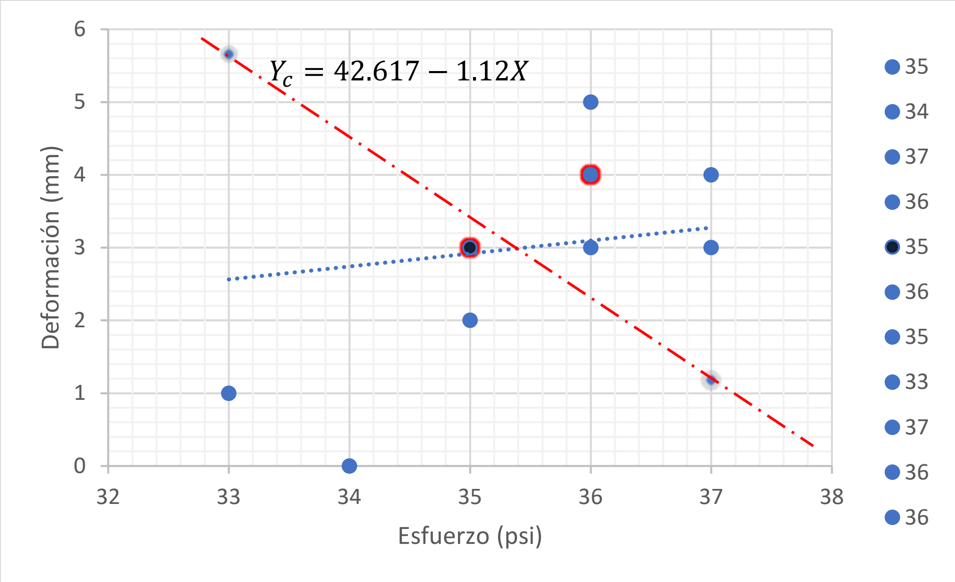

Figura 4

Caculo de coeficiente de correlación de Pearson

Tenemos la siguiente formula: $r = \frac{n(\sum xy)-(\sum x)(\sum y)}{\sqrt{([n\sum x^2-(\sum x)^2][n\sum y^2-(\sum y)^2])}}$

Tabla 7

Cálculo de $(Y_i)^2$

Materiales	X	Y	$X_i * Y_i$	$(X_i)^2$	$(Y_i)^2$
1	35	2	70	1225	4
2	34	0	0	1156	0
3	37	3	111	1369	9
4	36	4	114	1296	16
5	35	3	105	1225	9
6	36	4	144	1296	16
7	35	3	105	1225	9

8	33	1	33	1089	1
9	37	4	148	1369	16
10	36	3	108	1296	9
11	36	5	180	1296	25
Suma	390	32	1118	13842	114

Remplazamos $r = \frac{11(1118)-(390)(32)}{\sqrt{([11(13842)-(390)^2][11(114)-(32)^2])}}$

$$r = -0.942$$

$$|r| = 0.942$$

El valor de r se asemeja a uno; la relación lineal entre las variables es bastante fuerte.

Ejemplo 8

Determinar el vínculo entre la temperatura de operación de un motor eléctrico y su eficiencia energética

Temperatura de operación (°C)

Eficiencia energética (%)

Haremos la recopilación de datos:

Supongamos que se han recogido los siguientes datos de un motor eléctrico en diferentes condiciones de operación:

Tabla 8

Datos de Temperatura de operación (°C) y Eficiencia energética (%)

Temperatura de operación (°C)	Eficiencia energética (%)
20	92
25	91
30	89
35	88
40	85

45	83
50	80
55	78
60	75

Una vez recopilada lo datos, vamos a hacer el análisis correspondiente

Análisis de Correlación

Cálculo del coeficiente de correlación.

$$[r = \frac{\sum(X-\overline{X})(Y-\overline{Y})}{\sqrt{\sum(X-\overline{X})^2 \sum(Y-\overline{Y})^2}}]$$

Análisis:

Cálculo de la regresión lineal:

La fórmula de la recta de regresión es:

$$[Y = a + bX]$$

Donde:

- (Y) Eficiencia energética.
- (X) Temperatura de operación.
- (a) Ordenada al origen (intercepto).
- (b) Pendiente de la recta (coeficiente de regresión).

Resultados del Ejercicio de Correlación y Regresión

Coeficiente de Correlación:

$$[r = -0.992]$$

Este valor indica una fuerte correlación negativa entre la temperatura de operación y la eficiencia energética. A medida que aumenta la temperatura, la eficiencia energética disminuye.

Regresión Lineal:

$$[\text{Eficiencia} = 101.89 - 0.433 \times \text{Temperatura}]$$

Donde:

El intercepto (a) es 101.89.

La pendiente (b) es -0.433.

Interpretación

Intercepto (101.89): Cuando la temperatura de operación es 0°C, la eficiencia energética esperada del motor es 101.89%. Aunque este valor no es realista en términos físicos, es útil para la ecuación de regresión.

Pendiente (-0.433): Por cada aumento de 1°C en la temperatura de operación, la eficiencia energética disminuye en 0.433%.

Análisis de Resultados

R-cuadrado (0.984): el valor de (R^2) expresa que el 98.4% de la variación en la eficiencia energética se puede explicar por la variación en la temperatura de operación.

Probabilidad del F-Statistic (1.49e-07Este valor muestra que el modelo de regresión es muy relevante y confiable.

Este tipo de análisis es útil en ingeniería mecánica y eléctrica para optimizar las condiciones operativas y mejorar el rendimiento de los equipos.

Ejemplo 9

Una muestra aleatoria de 9 trabajadores eléctricos aparentemente buenos en sus trabajos entre 10 a 15 trabajadores produjo los siguientes datos respecto a los trabajadores y la eficiencia en porcentaje:

Trabajadores	Eficiencia %
10	86
10.5	88
11	92
12.5	93
13	95
13.5	96
14	97
14.5	98
15	99

- Encontrar la ecuación de la regresión lineal.
- En el diagrama de dispersión, trazar la recta de la ecuación de regresión lineal.
- Si el trabajador 13, ¿cuál es la eficiencia?

Solución:

Para determinar los coeficientes en una regresión lineal, se emplean las ecuaciones para la pendiente y la intersección,

X es trabajadores.

Y es la eficiencia.

Sumatorias:

$$\sum X = 10 + 10.5 + 11 + 12.5 + 13 + 13.5 + 14 + 14.5 + 15 = 114$$

$$\sum Y = 86 + 88 + 92 + 93 + 95 + 96 + 97 + 98 + 99 = 844$$

$$\sum XY = 10 * 86 + 10.5 * 88 + 11 * 92 + 12.5 * 93 + 13 * 95 + 13.5 * 96 + 14 * 97 + 14.5 * 98 + 15 * 99 = 10753.5$$

$\sum X2 = 100 + 110.25 + 121 + 156.25 + 169 + 182.25 + 196 + 210.25 + 225$
$= 1470$

Pendiente:

$$m = \frac{n(\sum xy) - ((\sum x) \ \ (\sum y)}{n(\sum x^2) - (\sum x)^2}$$

$$m = \frac{9 \times (10753.5) - (114)(844)}{9 \times (1470) - (114)^2} = 2.42$$

Intersección:

$$b = \frac{844 - (2.42)(114)}{9} = 63.12$$

La ecuación de regresión es:

$\mathrm{y} = 63.12 + 2.42(x)$

Si el trabajador 13, ¿cuál es la eficiencia?

$\mathrm{y} = 63.12 + 2.42 \ \ x\ 13 = \ 63.12 + 31.46 = 94.58\%$

La eficiencia del trabajador 13 es de 94.58%

Ejemplo 10

En una entidad eléctrica trabajan cuatro técnicos operarios. Los años que han pasado de sus títulos y la cantidad de meses que han trabajado en esta empresa, aquí se muestran cada uno:

años(x)	3	4	5	6
meses(y)	4	3	2	1

- calcular el valor del coeficiente de correlación lineal e interpretarlo.

solución:

Xi	Yi	Xi*Yi	X_i^2	Y_i^2
3	4	12	9	16

4	3	12	16	9
5	2	10	25	4
6	1	6	36	1
18	10	40	86	30

$$X = \frac{18}{4} = 4{,}5$$

$$Y = \frac{10}{4} = 2{,}5$$

$$\sigma_x^2 = \frac{86}{4} - 4.5^2 = 1.25$$

$$\sigma_y^2 = \frac{30}{4} - 2.5^2 = 1.25$$

$$\sigma_{xy} = \frac{40}{4} - 4.5 * 2.25 = -1.25$$

$$r = \frac{-1.25}{\sqrt{1.25} * \sqrt{1.25}} = \frac{-1.25}{1.25} = -1$$

interpretación:

Por lo tanto , la correlación es perfecta e inversa.

7. Conclusiones

- La correlación y la regresión son métodos estadísticos clave para entender y modelar las relaciones entre variables cuantitativas Estas técnicas nos ayudan a identificar patrones y hacer predicciones, asimismo facilita la toma de decisiones que se basan en datos en diferentes áreas, especialmente en ingeniería.
- En este manual, hemos explorado cómo la correlación evalúa el vínculo entre variables, mientras que la regresión facilita la predicción de resultados. Estas técnicas son fundamentales para que los procesos de ingeniería optimicen y mejoren el rendimiento de los sistemas, demostrando su utilidad práctica en aplicaciones del mundo real.
- Dominar la correlación y la regresión permite a los profesionales convertir los datos en información valiosa, impulsando la innovación y la mejora continua. Estas

herramientas no sólo ayudan a comprender las condiciones actuales, sino que también permiten a los profesionales predecir y dar forma a escenarios futuros, tomando decisiones informadas.

8. Referencias

Callister, W. D., & Rethwisch, D. G. (2018). *Materials science and engineering: An introduction (10th ed.).* Wiley.

Chen, X., & Chen, Y. (2020). *Energy efficiency and management in industrial applications. Springer.*

Groover, M. P. (2019). *Fundamentals of modern manufacturing: Materials, processes, and systems (7th ed.).* Wiley.

Jardine, A. K., L. D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing, 20*(7), 1483-1510.

Meeker, W. Q., & Escobar, L. A. (1998). *Statistical methods for reliability data.* Wiley.

Montgomery, D. C., & Runger, G. C. (2018). *Applied statistics and probability for engineers (7th ed.).* Wiley.

Ogata, K. (2010). *Modern control engineering (5th ed.).* Prentice Hall.

UG, C. D. (2018). *Unidad didáctica 5: Correlación y regresión.* Obtenido de https://blogs.ugto.mx/enfermeriaenlinea/unidad-didactica-5-correlacion-y-regresion/