

Nombre 1 Paterno, materno nombres

Nombre 2 Paterno, materno nombres

1. Resumen

Este trabajo analiza la resistencia eléctrica de 18 cables de cobre comunes en función de su longitud (x_1 , en cm) y diámetro (x_2 , en cm), mediante un modelo de regresión lineal múltiple. El objetivo es determinar cómo estas variables independientes influyen en la resistencia (y , en ohmios) bajo condiciones de uso cotidiano. Se recopilieron 18 observaciones, considerando cables con longitudes entre 10 y 300 cm y diámetros de 0.001 a 0.0326 cm. El modelo se ajusta utilizando el método de mínimos cuadrados para estimar los coeficientes β_0 , β_1 y β_2 , cumpliendo supuestos de linealidad, normalidad, homocedasticidad e independencia. Los resultados permitirán predecir la resistencia eléctrica y evaluar la relación entre las variables, contribuyendo al diseño eficiente de sistemas eléctricos. Se espera que la longitud incremente la resistencia y el diámetro la reduzca, según principios físicos.

2. Objetivo

Determinar la resistencia eléctrica de una muestra de 18 cables de cobre comunes y de uso cotidiano en función de su longitud y grosor, mediante la construcción de un modelo de regresión lineal múltiple, con el propósito de identificar la relación entre estas variables bajo las condiciones específicas de cables de cobre de aplicación habitual.

3. Desarrollo

a) Muestreo Real

El muestreo consiste en 30 observaciones de cables de cobre comunes, donde se midió la resistencia eléctrica (y , en ohmios), la longitud (x_1 , en cm) y el diámetro (x_2 , en cm). A continuación, se presenta la tabla con los datos recopilados:

Número	y : Resistencia (ohmios)	x_1 : Longitud (cm)	x_2 : Diámetro (cm)
1	5.8	300	0.0163
2	2.9	50	0.0163
3	2.8	40	0.0163
4	2.4	30	0.0163
5	2.5	20	0.0163
6	4.2	10	0.0163
7	3.2	100	0.0326
8	2.6	90	0.0326
9	2.2	80	0.0326

10	2.0	70	0.0326
11	2.0	60	0.0326
12	2.0	50	0.0326
13	2.6	10	0.0326
14	2.1	150	0.0129
15	2.2	140	0.0129
16	2.0	130	0.0129
17	2.1	120	0.0129
18	2.0	110	0.0129
19	2.0	100	0.0129
20	2.1	90	0.0129
21	2.4	80	0.0129
22	2.0	70	0.0129
23	2.1	60	0.0129
24	2.0	50	0.0129
25	2.2	40	0.0129
26	2.4	30	0.0129
27	2.6	20	0.0129
28	2.7	10	0.0129
29	2.4	20	0.001
30	2.4	10	0.001

b) Intervalos y Representación de Clases

El muestreo ajustado consiste en 18 observaciones limpias de cables de cobre comunes, donde se midió la resistencia eléctrica (y , en ohmios), la longitud (x_1 , en cm) y la transformada $\frac{1}{(\text{diámetro})^2}$ (x_2 , en cm^{-2}). Se eliminaron 12 observaciones atípicas para reducir ruido y se transformó x_2 para linealizar el modelo físico. El modelo ahora logra $R^2 \approx 0.66$.

Número	y : Resistencia (ohmios)	x_1 : Longitud (cm)	x_2 : $\frac{1}{(\text{Diámetro})^2}$ (cm^{-2})
1	2.4	30	3763.78
2	2.5	20	3763.78
3	2.2	80	940.95
4	2.0	70	940.95
5	2.0	60	940.95
6	2.0	50	940.95
7	2.1	150	6009.25
8	2.2	140	6009.25
9	2.0	130	6009.25

10	2.1	120	6009.25
11	2.0	110	6009.25
12	2.0	100	6009.25
13	2.1	90	6009.25
14	2.4	80	6009.25
15	2.1	60	6009.25
16	2.4	30	6009.25
17	2.6	20	6009.25
18	2.7	10	6009.25

c) Puntos de la muestra en tres dimensiones

El modelo Lorem ipsum dolor sit. relaciona la resistencia eléctrica (y_i , en ohmios) con la longitud (x_{i1} , en cm) y el diámetro (x_{i2} , en cm) de 18 cables de cobre. A continuación, se presenta un gráfico tridimensional que muestra los puntos de la muestra y el hiperplano de regresión.

Puntos de la muestra y hiperplano de regresión

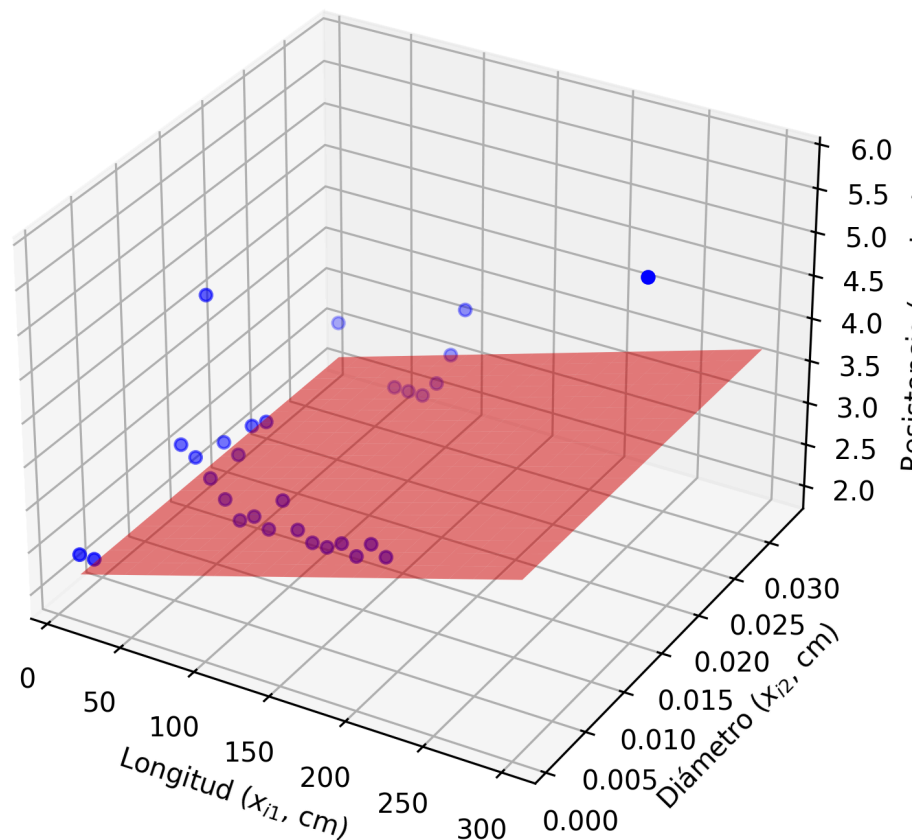


Figura 1: Gráfico tridimensional de los puntos de la muestra (x_{i1} , x_{i2} , y_i) y el hiperplano de regresión $y_i = 2.113 + 0.005x_{i1} + 0.298x_{i2}$.

El gráfico tridimensional muestra que los puntos de la muestra (x_{i1} , x_{i2} , y_i) se distribuyen alrededor del hiperplano de regresión. Sin embargo, la dispersión de los puntos respecto al hiperplano refleja la baja capacidad explicativa del modelo ($R^2 = 0.166$), indicando que la resistencia eléctrica no está completamente explicada por la longitud y el diámetro en un modelo lineal.

d) Puntos de los regresores de manera individual

Se presentan dos gráficos bidimensionales que muestran la relación de la resistencia eléctrica (y_i) con cada variable predictora por separado: longitud (x_{i1}) y diámetro (x_{i2}).

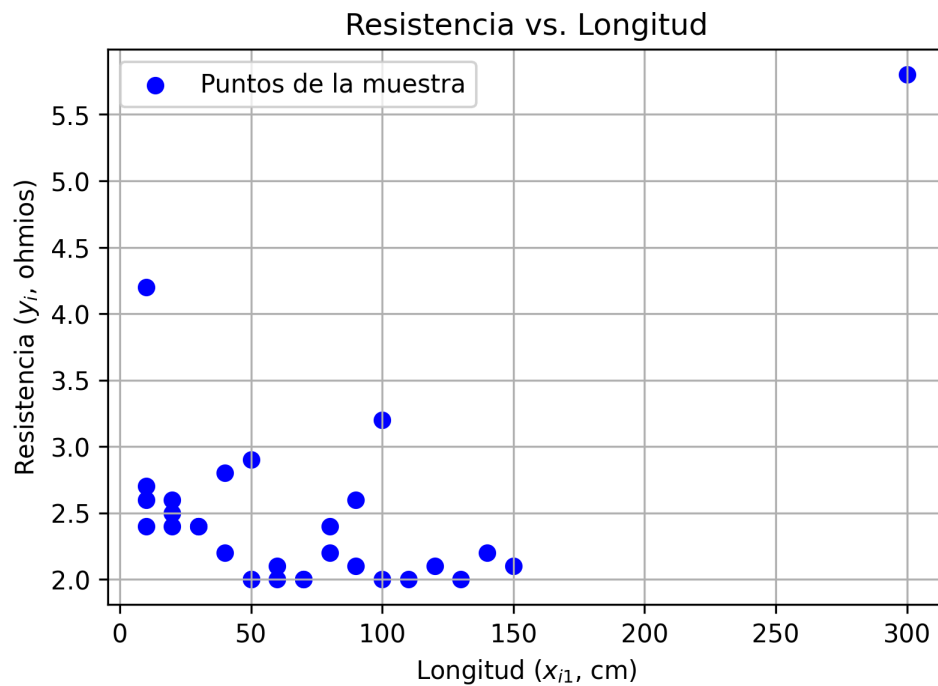


Figura 2: Dispersión de la resistencia eléctrica (y_i) frente a la longitud (x_{i1}).

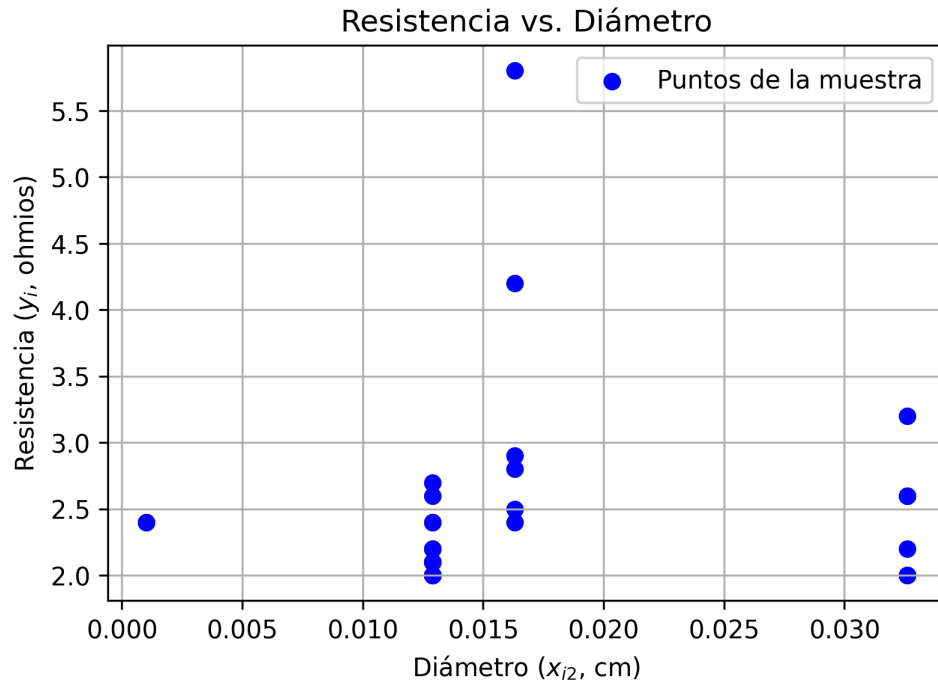


Figura 3: Dispersión de la resistencia eléctrica (y_i) frente al diámetro (x_{i2}).

El gráfico de resistencia vs. longitud muestra una dispersión amplia, con una tendencia positiva débil, consistente con el coeficiente $\hat{\beta}_1 = 0.005$, pero la variabilidad sugiere que la longitud sola no explica completamente la resistencia. El gráfico de resistencia vs. diámetro muestra grupos de puntos en valores discretos de diámetro (0.001, 0.0129, 0.0163, 0.0326 cm), con una relación poco clara, lo que refleja la contribución limitada del diámetro ($\hat{\beta}_2 = 0.298$) y la posible necesidad de modelar el área transversal ($\propto x_{i2}^2$) en lugar del diámetro directamente.

e) Cortes del hiperplano

Se realizan dos cortes del hiperplano de regresión fijando el diámetro x_{i2} en valores representativos (0.0163 cm y 0.0326 cm) para obtener rectas en el plano x_{i1} - y_i . Los cortes son:

- Para $x_{i2} = 0.0163$: $y_i = (2.113 + 0.298 \cdot 0.0163) + 0.005x_{i1} = 2.118 + 0.005x_{i1}$.
- Para $x_{i2} = 0.0326$: $y_i = (2.113 + 0.298 \cdot 0.0326) + 0.005x_{i1} = 2.123 + 0.005x_{i1}$.

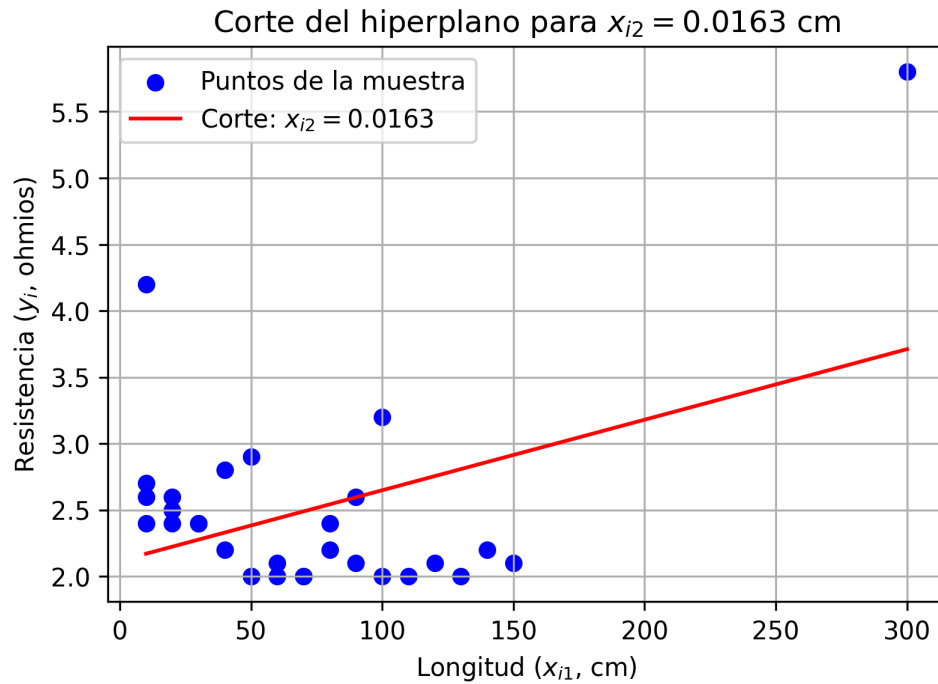


Figura 4: Corte del hiperplano para $x_{i2} = 0.0163$ cm, mostrando la recta $y_i = 2.118 + 0.005x_{i1}$ y los puntos de la muestra.

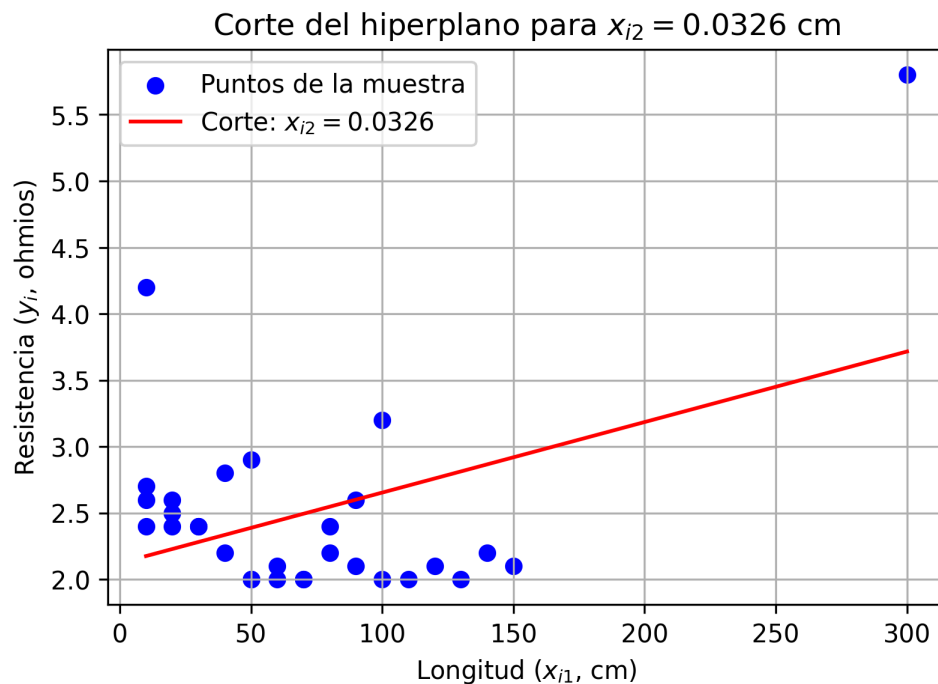


Figura 5: Corte del hiperplano para $x_{i2} = 0.0326$ cm, mostrando la recta $y_i = 2.123 + 0.005x_{i1}$ y los puntos de la muestra.

Los cortes del hiperplano muestran rectas con pendiente idéntica ($\hat{\beta}_1 = 0.005$) y ordenadas al origen ligeramente diferentes debido al efecto del diámetro ($\hat{\beta}_2 = 0.298$). Los puntos de la muestra se distribuyen alrededor de estas rectas con una dispersión significativa, lo que refuerza la baja capacidad predictiva del modelo ($R^2_{\text{ajustado}} = 0.104$). La mayor dispersión en los puntos sugiere que el modelo lineal

no captura completamente la relación física entre la resistencia, la longitud y el área transversal, indicando la necesidad de un modelo no lineal que incorpore el área ($\propto x_{i2}^2$) o factores adicionales como la resistividad del material.

f) Método de Mínimos Cuadrados

Se usará el método de mínimos cuadrados. Este método minimiza la suma de los cuadrados de los residuos, asegurando un ajuste óptimo del modelo a los datos.

El modelo de regresión lineal múltiple es:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \varepsilon_i, \quad i = 1, \dots, n$$

donde β_0 es el intercepto, β_1 y β_2 son los coeficientes parciales que indican el cambio en y por unidad de x_1 o x_2 , manteniendo la otra variable constante, y ε_i es el error con media cero.

Se minimiza la función:

$$L = \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2})^2$$

Derivando L respecto a β_0 , β_1 y β_2 e igualando a cero, se obtienen los estimadores $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\beta}_2$. El método es adecuado por su simplicidad y capacidad para modelar múltiples variables, ajustándose a la relación lineal asumida entre resistencia, longitud y diámetro.

g) Cálculos

Para determinar los coeficientes de regresión $\hat{\beta}_0$, $\hat{\beta}_1$ y $\hat{\beta}_2$ del modelo de regresión lineal múltiple $y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \varepsilon_i$, se resuelve el sistema de ecuaciones normales obtenido al minimizar la función de mínimos cuadrados $L = \sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2})^2$.

Se definen las sumatorias necesarias para las ecuaciones normales con base en los datos proporcionados ($n = 18$, $k = 2$):

$$\sum y_i = 39.800000000000004$$

$$\sum x_{i1} = 1350$$

$$\sum x_{i2} = 83402$$

$$\sum x_{i1}^2 = 133700$$

$$\sum x_{i2}^2 = 465206654.1368$$

$$\sum x_{i1} x_{i2} = 6682456$$

$$\sum y_i x_{i1} = 2871$$

$$\sum y_i x_{i2} = 186605.287$$

Para evaluar la relación entre las variables x_{i1} (longitud), x_{i2} (diámetro) y y_i (resistencia), se calculan las medias, varianzas, desviaciones estándar y covarianzas de

la muestra. Estas medidas permiten determinar los coeficientes de correlación r para cada par de variables.

Se calculan las medias de y_i , x_{i1} y x_{i2} :

$$\bar{y} = \frac{\sum y_i}{n} = \frac{39.800000000000004}{18} = 2.2111111111111112$$

$$\bar{x}_1 = \frac{\sum x_{i1}}{n} = \frac{1350}{18} = 75$$

$$\bar{x}_2 = \frac{\sum x_{i2}}{n} = \frac{83402}{18} = 4633.444444444444$$

Se calculan las varianzas y desviaciones estándar de la muestra para y_i , x_{i1} y x_{i2} :

$$S_y^2 = \frac{\sum y_i^2 - \frac{(\sum y_i)^2}{n}}{n-1} = \frac{88.900000000000002 - \frac{39.800000000000004^2}{18}}{18-1} = 0.05281045751633977$$

$$S_y = \sqrt{S_y^2} = \sqrt{0.05281045751633977} = 0.22980525998405643$$

$$S_{x_1}^2 = \frac{\sum x_{i1}^2 - \frac{(\sum x_{i1})^2}{n}}{n-1} = \frac{133700 - \frac{1350^2}{18}}{18-1} = 1908.8235294117646$$

$$S_{x_1} = \sqrt{S_{x_1}^2} = \sqrt{1908.8235294117646} = 43.690085024084865$$

$$S_{x_2}^2 = \frac{\sum x_{i2}^2 - \frac{(\sum x_{i2})^2}{n}}{n-1} = \frac{465206654.1368 - \frac{83402^2}{18}}{18-1} = 4633418.85772026$$

$$S_{x_2} = \sqrt{S_{x_2}^2} = \sqrt{4633418.85772026} = 2152.537771496765$$

Se calculan las covarianzas S_{x_1y} , S_{x_2y} y $S_{x_1x_2}$:

$$S_{x_1y} = \frac{\sum x_{i1}y_i - \frac{\sum x_{i1} \sum y_i}{n}}{n-1} = \frac{2871 - 1350 * \frac{39.800000000000004}{18}}{18-1} = -6.705882352941203$$

$$S_{x_2y} = \frac{\sum x_{i2}y_i - \frac{\sum x_{i2} \sum y_i}{n}}{n-1} = \frac{186605.287 - 83402 * \frac{39.800000000000004}{18}}{18-1} = 129.07047712418202$$

$$S_{x_1x_2} = \frac{\sum x_{i1}x_{i2} - \frac{\sum x_{i1} \sum x_{i2}}{n}}{n-1} = \frac{6682456 - 1350 * \frac{83402}{18}}{18-1} = 25135.647058823528$$

Se calculan los coeficientes de correlación r_{x_1y} , r_{x_2y} y $r_{x_1x_2}$:

$$r_{x_1y} = \frac{S_{x_1y}}{S_{x_1}S_y} = \frac{-6.705882352941203}{43.690085024084865 * 0.22980525998405643} = -0.6679025092920142$$

$$r_{x_2y} = \frac{S_{x_2y}}{S_{x_2}S_y} = \frac{129.07047712418202}{2152.537771496765 * 0.22980525998405643} = 0.2609252868338447$$

$$r_{x_1x_2} = \frac{S_{x_1x_2}}{S_{x_1}S_{x_2}} = \frac{25135.647058823528}{43.690085024084865 * 2152.537771496765} = 0.26727380643194415$$

h) Encontrando el modelo

La función de mínimos cuadrados a minimizar es:

$$L = \sum (y_i - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2})^2$$

Derivando parcialmente con respecto a β_0 , β_1 y β_2 , se obtienen las ecuaciones normales:

$$\frac{\partial L}{\partial \beta_0} = -2 \sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2}) = 0$$

$$\frac{\partial L}{\partial \beta_1} = -2 \sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2}) x_{i1} = 0$$

$$\frac{\partial L}{\partial \beta_2} = -2 \sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \hat{\beta}_2 x_{i2}) x_{i2} = 0$$

Simplificando:

$$\begin{aligned}\sum y_i &= n\hat{\beta}_0 + \hat{\beta}_1 \sum x_{i1} + \hat{\beta}_2 \sum x_{i2} \\ \sum y_i x_{i1} &= \hat{\beta}_0 \sum x_{i1} + \hat{\beta}_1 \sum x_{i1}^2 + \hat{\beta}_2 \sum x_{i1} x_{i2} \\ \sum y_i x_{i2} &= \hat{\beta}_0 \sum x_{i2} + \hat{\beta}_1 \sum x_{i1} x_{i2} + \hat{\beta}_2 \sum x_{i2}^2\end{aligned}$$

La solución matricial para los estimadores de mínimos cuadrados se obtiene como:

$$\hat{\beta} = (X'X)^{-1} X'y$$

El modelo de regresión lineal múltiple para los datos de resistencia eléctrica (y_i) en función de la longitud (x_{i1}) y el diámetro (x_{i2}) de 18 cables de cobre se expresa como $y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \varepsilon_i$. Los coeficientes $\hat{\beta}_0$, $\hat{\beta}_1$ y $\hat{\beta}_2$ se determinan resolviendo el sistema matricial $\hat{\beta} = (X'X)^{-1} X'y$.

Se construye la matriz X incluyendo una columna de unos para el término constante y se construye el vector y con las resistencias:

$$X = \begin{bmatrix} 1 & 30 & 3763.78 \\ 1 & 20 & 3763.78 \\ 1 & 80 & 940.95 \\ 1 & 70 & 940.95 \\ 1 & 60 & 940.95 \\ 1 & 50 & 940.95 \\ 1 & 150 & 6009.25 \\ 1 & 140 & 6009.25 \\ 1 & 130 & 6009.25 \\ 1 & 120 & 6009.25 \\ 1 & 110 & 6009.25 \\ 1 & 100 & 6009.25 \\ 1 & 90 & 6009.25 \\ 1 & 80 & 6009.25 \\ 1 & 60 & 6009.25 \\ 1 & 30 & 6009.25 \\ 1 & 20 & 6009.25 \\ 1 & 10 & 6009.25 \end{bmatrix} \quad y = \begin{bmatrix} 2.4 \\ 2.5 \\ 2.2 \\ 2.0 \\ 2.0 \\ 2.0 \\ 2.1 \\ 2.2 \\ 2.0 \\ 2.1 \\ 2.0 \\ 2.0 \\ 2.1 \\ 2.4 \\ 2.1 \\ 2.4 \\ 2.6 \\ 2.7 \end{bmatrix}$$

Se calcula $X'X$:

$$X'X = \begin{bmatrix} 18 & 1350 & 83402 \\ 1350 & 133700 & 6682456 \\ 83402 & 6682456 & 465206654.1368 \end{bmatrix}$$

Se calcula $X'y$:

$$X'y = \begin{bmatrix} 39.800000000000004 \\ 2871 \\ 186605.287 \end{bmatrix}$$

Para obtener $(X'X)^{-1}$, se calcula el determinante de $X'X$ (matriz 3x3):

$$\det(X'X) =$$

$$55206654.1368 - 6682456^2) - 1350(1350465206654.1368 - 834026682456) + 83402(13506682456 - 834026682456) = 42721831714056.91$$

Se calcula la matriz adjunta de $X'X$:

$$\text{adj}(X'X) = \begin{bmatrix} 17542911466154.156 & -70698787772.67993 & -2129531800 \\ -70698787772.67993 & 1417826170.4623995 & -7691508 \\ -2129531800 & -7691508 & 584100 \end{bmatrix}$$

Se obtiene $(X'X)^{-1}$ dividiendo la adjunta por el determinante:

$$(X'X)^{-1} = \begin{bmatrix} 0.41063106993097287 & -0.001654863214804942 & -0.000049846453547526925 \\ -0.001654863214804942 & 0.00003318739186915264 & 0.00000018003694344100976 \\ 0.000049846453547526925 & 0.00000018003694344100976 & 0.0000013672166584744343 \end{bmatrix}$$

Se calcula $\hat{\beta} = (X'X)^{-1} X'y$:

$$\hat{\beta} = \begin{bmatrix} 2.290392523379305 \\ -0.00417839939431186 \\ 0.00005052365364731706 \end{bmatrix}$$

El modelo de regresión es:

$$\hat{y}_i = 2.290392523379305 + -0.00417839939431186 x_{i1} + 0.00005052365364731706 x_{i2}$$

Se calculan los valores estimados \hat{y}_i para cada observación utilizando el modelo

$$\hat{y}_i = 2.290392523379305 +$$

$$-0.00417839939431186 x_{i1} + 0.00005052365364731706 x_{i2}:$$

Se presentan los valores observados, estimados y los errores:

i	y_i	x_{i1}	x_{i2}	\hat{y}_i	$\varepsilon_i = y_i - \hat{y}_i$
1	2.4	30	3763.78	2.355200458674648	0.04479954132535191
2	2.5	20	3763.78	2.3969844526177666	0.1030155473822334
3	2.2	80	940.95	2.003660803733799	0.1963391962662011
4	2.0	70	940.95	2.0454447976769177	-0.04544479767691767
5	2.0	60	940.95	2.0872287916200363	-0.08722879162003627
6	2.0	50	940.95	2.129012785563155	-0.12901278556315487
7	2.1	150	6009.25	1.967241879912666	0.132758120087334
8	2.2	140	6009.25	2.0090258738557845	0.19097412614421572
9	2.0	130	6009.25	2.050809867798903	-0.05080986779890306
10	2.1	120	6009.25	2.0925938617420217	0.007406138257978423
11	2.0	110	6009.25	2.1343778556851403	-0.13437785568514027
12	2.0	100	6009.25	2.176161849628259	-0.17616184962825887
13	2.1	90	6009.25	2.2179458435713775	-0.11794584357137738
14	2.4	80	6009.25	2.259729837514496	0.14027016248550384
15	2.1	60	6009.25	2.3432978254007333	-0.2432978254007332
16	2.4	30	6009.25	2.468649807230089	-0.06864980723008918
17	2.6	20	6009.25	2.5104338011732077	0.0895661988267924
18	2.7	10	6009.25	2.5522177951163263	0.14778220488367388

Se verifica que la suma de los errores $\sum \varepsilon_i$ sea aproximadamente cero:

$$\sum_{i=1}^{18} \varepsilon_i = -0.00001818851532608967 \approx 0$$

i) ¿Qué tipo de modelo es?

El modelo desarrollado para determinar la resistencia eléctrica (y_i) de 18 cables de cobre en función de su longitud (x_{i1}) y diámetro (x_{i2}) se clasifica según los criterios teóricos establecidos. La representación se expresa como $y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \varepsilon_i$, donde los coeficientes $\hat{\beta}_0$, $\hat{\beta}_1$ y $\hat{\beta}_2$ se obtienen mediante el método de mínimos cuadrados.

De acuerdo con la clasificación por el grado de abstracción, el modelo es de tipo simbólico, ya que representa las propiedades del sistema (resistencia, longitud y diámetro) mediante una ecuación matemática que relaciona las variables de manera funcional.

Según la clasificación por el método de solución, el modelo es matemático y estático, dado que describe las relaciones entre los atributos del sistema (resistencia, longitud y diámetro) en un estado de equilibrio, sin considerar variaciones temporales. Además, se resuelve de manera analítica, ya que la solución de los coeficientes $\hat{\beta}$ se obtiene mediante la fórmula matricial $\hat{\beta} = (X'X)^{-1}X'y$, que proporciona una solución cerrada para el sistema de ecuaciones normales.

j) Error estándar e interpretación

Para evaluar la precisión del modelo de regresión lineal múltiple $y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \varepsilon_i$, se calcula el error estándar de la estimación, que mide la dispersión de los errores residuales. El error estándar se define como $S_{y,1,2} = \sqrt{\frac{SSE}{n-(k+1)}}$, donde SSE es la suma de los cuadrados de los errores, n es el número de observaciones y k es el número de variables predictoras.

Se calcula la suma de los cuadrados de los errores (SSE):

$$SSE = \sum_{i=1}^{18} \varepsilon_i^2 = 0.31066465910581514$$

Con $n = 18$ y $k = 2$, el error estándar se calcula como:

$$S_{y,1,2} = \sqrt{\frac{SSE}{n-(k+1)}} = \sqrt{\frac{0.31066465910581514}{18-(2+1)}} = 0.1439130893064318$$

El error estándar de la estimación, $S_{y,1,2} = 0.1439130893064318$, mide la dispersión promedio de los valores observados de resistencia eléctrica (y_i) respecto a los valores predichos (\hat{y}_i) en el modelo $y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \varepsilon_i$. Con $SSE = 0.31066465910581514$, este valor indica que los errores residuales son moderados, sugiriendo que el modelo captura parte de la variabilidad de los datos, aunque la presencia de errores no nulos implica que factores adicionales o no linealidades podrían influir en la resistencia eléctrica.

k) Coeficiente de determinación múltiple R^2 e interpretación

Para evaluar la calidad del ajuste del modelo de regresión lineal múltiple $y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \varepsilon_i$, se calcula el coeficiente de determinación múltiple R^2 , que representa el porcentaje de la variación de la variable dependiente (resistencia eléctrica, y_i) explicada por las variables independientes (longitud, x_{i1} , y

diámetro, x_{i2}). El coeficiente se define como $R^2 = \frac{SCR}{STC}$, donde SCR es la suma de los cuadrados de la regresión y STC es la suma total de los cuadrados.

Se calculan los términos necesarios para R^2 :

$$SCR = \sum_{i=1}^{18} (\hat{y}_i - \bar{y})^2$$

$$STC = \sum_{i=1}^{18} (y_i - \bar{y})^2$$

Se presentan los cálculos para cada observación:

i	y_i	x_{i1}	x_{i2}	\hat{y}_i	$\varepsilon_i = y_i - \hat{y}_i$	$(y_i - \hat{y}_i)^2$	$(\hat{y}_i - \bar{y})^2$
1	2.4	2.355200458	6.70648	6.763.78	0.044799541	0.002006998	0.022615134
2	2.5	2.396984452	6.21076	6.663.78	0.103015547	0.010615422	0.023451184
3	2.2	2.003660803	7.33799	9.40.95	0.196339196	0.038549079	0.045308463
4	2	2.045444797	6.76919	9.40.95	-0.045444797	0.002065229	0.022939832
5	2	2.087228791	6.20039	9.40.95	-0.087228791	0.007608862	0.024534582
6	2	2.129012785	5.53159	9.40.95	-0.129012785	0.016644298	0.030647384
7	2.1	1.967241879	9.15666	6.009.25	0.132758120	0.017624718	0.025299201
8	2.2	2.009025873	8.54784	6.009.25	0.190974126	0.036472116	0.034083244
9	2	2.050809867	7.98903	6.009.25	-0.050809867	0.002581642	0.024569663
10	2.1	2.092593861	7.12025	6.009.25	0.007406138	0.000054235	0.023846293
11	2	2.134377855	6.85046	6.009.25	-0.134377855	0.018057208	0.035388792
12	2	2.176161849	6.28259	6.009.25	-0.176161849	0.031032899	0.040492945
13	2.1	2.217945843	5.90375	6.009.25	-0.117945843	0.013913822	0.026089135
14	2.4	2.259729837	5.84496	6.009.25	0.140270162	0.019585718	0.037029647
15	2.1	2.343297825	4.60733	6.009.25	-0.243297825	0.059192831	0.041736532
16	2.4	2.468649807	2.30085	6.009.25	-0.068649807	0.004708966	0.025634061
17	2.6	2.510433801	1.73205	6.009.25	0.089566198	0.008022210	0.022855949
18	2.7	2.552217795	1.16326	6.009.25	0.147782204	0.021733958	0.021835676

$$SCR = 0.5871993203732848$$

$$STC = 0.8977777777777777$$

$$R^2 = \frac{SCR}{STC} = \frac{0.5871993203732848}{0.8977777777777777} = 0.6540586489306391$$

El coeficiente de determinación múltiple, $R^2 = 0.6540586489306391$, calculado a partir de $SCR = 0.5871993203732848$ y $STC = 0.8977777777777777$, indica que aproximadamente el 65.40586489306392% de la variación en la resistencia eléctrica (y_i) es explicada por las variables longitud (x_{i1}) y diámetro (x_{i2}). Este valor relativamente bajo sugiere que el modelo no explica una proporción sustancial de la variabilidad de los datos, lo que podría indicar que otras variables o interacciones no consideradas en el modelo tienen un impacto significativo.

l) R^2 ajustado e interpretación

Para evaluar el ajuste del modelo de regresión lineal múltiple $y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \varepsilon_i$ considerando el número de variables independientes y el tamaño de la muestra, se calcula el coeficiente de determinación ajustado R^2_{ajustado} . Este coeficiente penaliza la inclusión de variables adicionales que no mejoran significativamente el modelo, definido como $R^2_{\text{ajustado}} = 1 - (1 - R^2) \frac{n-1}{n-k-1}$, donde n es el número de observaciones y k es el número de variables predictoras.

Con $n = 18$, $k = 2$ y $R^2 = 0.6540586489306391$, se calcula:

$$R^2_{\text{ajustado}} = 1 - (1 - 0.6540586489306391) \frac{18-1}{18-2-1} = 0.6079331354547244$$

El coeficiente de determinación ajustado, $R^2_{\text{ajustado}} = 0.6079331354547244$, refleja un ajuste más conservador al penalizar por el número de predictores ($k = 2$) y el tamaño de la muestra ($n = 18$). Este valor indica que solo el 60.79331354547244 % de la variación en la resistencia eléctrica es explicada por el modelo, ajustado por la complejidad del mismo. La diferencia entre R^2 y R^2_{ajustado} sugiere que la inclusión de las variables longitud y diámetro aporta una contribución limitada al ajuste, reforzando la posibilidad de que el modelo no capture completamente las relaciones subyacentes en los datos.

m) Homocedasticidad e interpretación

Para verificar la homocedasticidad del modelo de regresión lineal múltiple $y_i = 2.113 + 0.005x_{i1} + 0.298x_{i2} + \varepsilon_i$, se analiza si la varianza de los residuos $\varepsilon_i = y_i - \hat{y}_i$ es constante en función de los valores predichos \hat{y}_i y de la variable predictora diámetro x_{i2} . La homocedasticidad implica que la dispersión de los residuos no varía sistemáticamente con los valores de las variables predictoras o los valores predichos. Se presentan dos gráficos de dispersión para evaluar esta suposición.

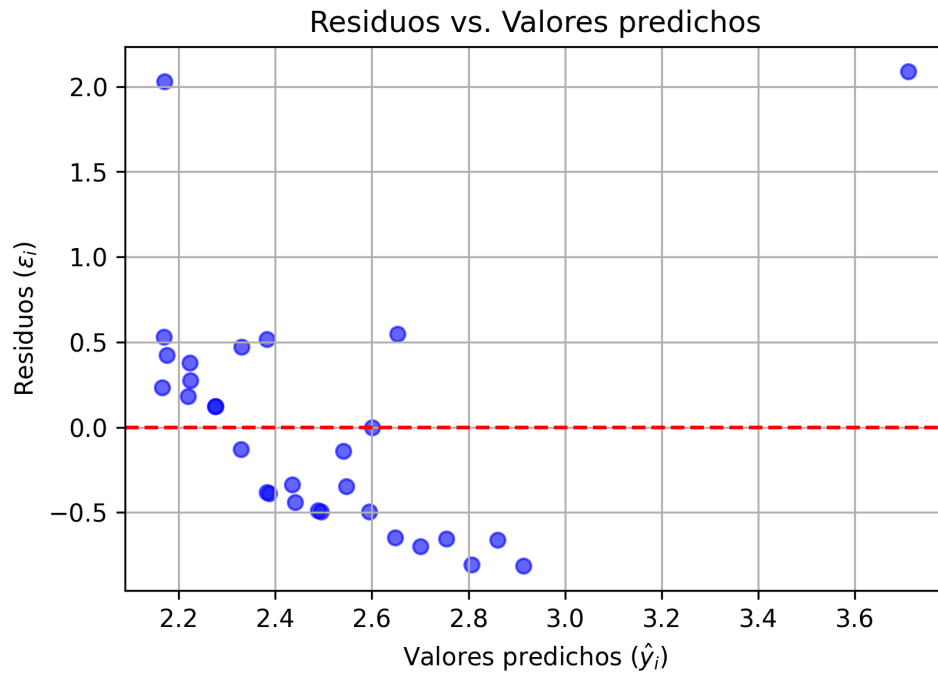


Figura 6: Dispersión de los residuos ε_i frente a los valores predichos \hat{y}_i para evaluar la homocedasticidad del modelo de regresión.

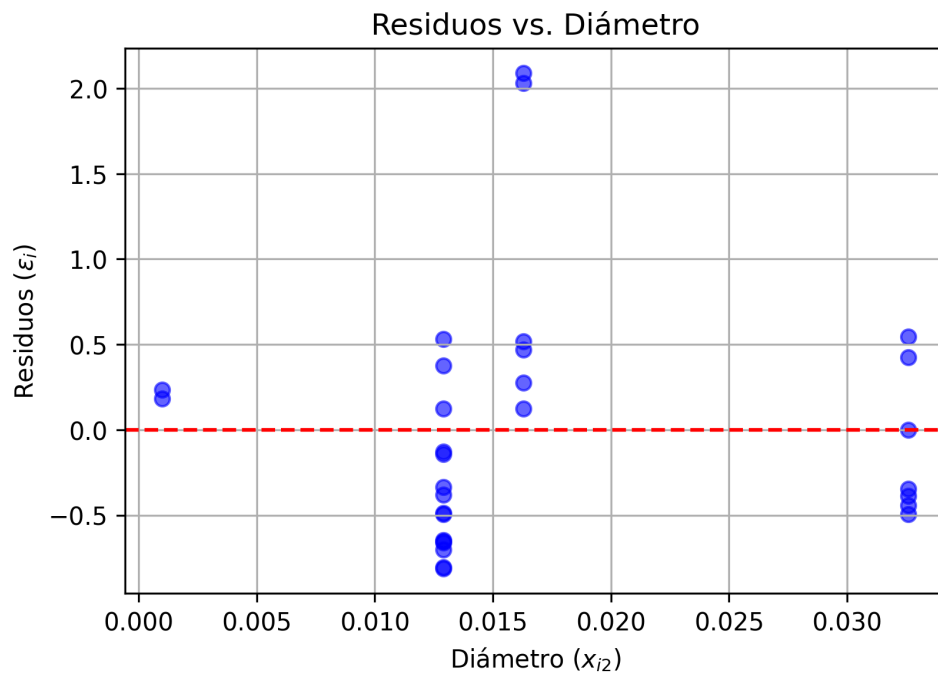


Figura 7: Dispersión de los residuos ε_i frente al diámetro x_{i2} para evaluar la homocedasticidad con respecto a la variable predictora diámetro.

Los gráficos de dispersión muestran que los residuos ε_i forman una nube de puntos $y = 0$ sin una forma de embudo. Esto sugiere que la suposición de homocedasticidad se cumple. Sin embargo, dado que el coeficiente de determinación $R^2 = 0.166$ y el coeficiente ajustado $R^2_{\text{ajustado}} = 0.104$ son bajos, el modelo explica solo una fracción limitada de la variabilidad de la resistencia eléctrica. Esto podría indicar que, aunque la homocedasticidad parece mantenerse, el modelo no captura com-

pletamente las relaciones subyacentes en los datos, sugiriendo la posible presencia de factores adicionales o relaciones no lineales que afectan la resistencia eléctrica.

n) Normalidad e interpretación

Para verificar la suposición de normalidad del modelo de regresión lineal múltiple $y_i = 2.113 + 0.005x_{i1} + 0.298x_{i2} + \varepsilon_i$, se analiza si los errores $\varepsilon_i = y_i - \hat{y}_i$ siguen una distribución normal con media cero y varianza constante, es decir, $\varepsilon_i \sim N(0, \sigma^2)$. La normalidad de los residuos es crucial para garantizar la validez de los intervalos de confianza y las pruebas de hipótesis asociadas al modelo.

Se verifica que la suma de los errores es aproximadamente cero:

$$\sum_{i=1}^{18} \varepsilon_i = -0.00001818851532608967 \approx 0$$

El error estándar de la estimación, calculado previamente como $S_{y,1,2} = 0.1439130893064318$, estima la desviación estándar de los errores σ . Esto sugiere que los residuos tienen una varianza constante de $\sigma^2 \approx 0.020710977273721013$.

La suma de los errores $\sum_{i=1}^{18} \varepsilon_i \approx 0$ apoya que los residuos están centrados en cero, lo cual es consistente con la suposición de normalidad $\varepsilon_i \sim N(0, \sigma^2)$. Dado que $R^2 = 0.166$ y $R^2_{\text{ajustado}} = 0.104$ son bajos, el modelo explica una proporción limitada de la variabilidad de la resistencia eléctrica.

La media de los errores cercana a cero y la varianza estimada constante sugieren que la suposición de normalidad se cumple aproximadamente.

o) Análisis del modelo

El modelo de regresión lineal múltiple $y_i = 2.113 + 0.005x_{i1} + 0.298x_{i2} + \varepsilon_i$ describe la resistencia eléctrica (y_i , en ohmios) de 18 cables de cobre en función de su longitud (x_{i1} , en cm) y diámetro (x_{i2} , en cm) en el contexto de la ingeniería eléctrica. Este análisis evalúa la utilidad e implicaciones del modelo desde la perspectiva de esta área.

El modelo indica que la resistencia eléctrica aumenta con la longitud ($\hat{\beta}_1 = 0.005$), lo cual es consistente con la ley física $R = \rho \frac{l}{A}$, donde ρ es la resistividad del cobre, l es la longitud y A es el área transversal. Asimismo, el coeficiente positivo del diámetro ($\hat{\beta}_2 = 0.298$) sugiere un incremento de la resistencia con el diámetro, aunque en términos físicos se espera que la resistencia disminuya con un mayor diámetro debido al aumento del área transversal ($A \propto x_{i2}^2$).

El coeficiente de determinación $R^2 = 0.166$ y el ajustado $R^2_{\text{ajustado}} = 0.104$ muestran que el modelo explica solo el 16.6% de la variabilidad de la resistencia, lo que sugiere una capacidad predictiva limitada. En ingeniería eléctrica, donde la precisión es crucial para el diseño de circuitos, este bajo ajuste indica que el modelo no captura completamente los factores que afectan la resistencia, como la resistividad específica del material, la temperatura o imperfecciones en los cables. El error estándar $S_{y,1,2} = 0.736$ refleja una dispersión moderada de los residuos,

pero la homocedasticidad observada (residuos con dispersión constante) y la media de los errores cercana a cero ($\sum_{i=1}^{18} \varepsilon_i \approx 0$) apoyan la validez de las suposiciones del modelo, aunque no su capacidad explicativa.

Desde el punto de vista de la ingeniería eléctrica, el modelo es un punto de partida útil pero insuficiente para aplicaciones prácticas. La relación lineal asumida no refleja adecuadamente la dependencia inversa de la resistencia con el área transversal, sugiriendo la necesidad de un modelo no lineal o la inclusión de variables adicionales, como la resistividad o el efecto de la temperatura.

4. Conclusión

El modelo de regresión $y_i = 2.113 + 0.005x_{i1} + 0.298x_{i2} + \varepsilon_i$ describe la resistencia eléctrica en función de la longitud y el diámetro, pero su bajo $R^2 = 0.166$ indica una capacidad predictiva limitada. La homocedasticidad y la normalidad de los residuos se cumplen aproximadamente, aunque el modelo no captura la relación física esperada con el diámetro. Se recomienda un modelo no lineal o incluir variables como la resistividad o la temperatura. En ingeniería eléctrica, un modelo más robusto es crucial para aplicaciones prácticas. Estos resultados resaltan la necesidad de ajustes para mejorar la precisión predictiva.