

**UNIVERSIDAD NACIONAL
JOSÉ FAUSTINO SÁNCHEZ CARRIÓN**



**FACULTAD DE CIENCIAS
ESCUELA PROFESIONAL DE ESTADÍSTICA E INFORMÁTICA**

MODELOS LINEALES

**AUTORA:
BLANCO NAPURI CANDY EMPERATRIZ**

HUACHO – PERÚ

2021

TRABAJO DE INVESTIGACION

Se cree que la energía eléctrica consumida cada mes por una planta química está relacionada con la temperatura ambiental promedio x_1 , el número de días del mes x_2 , la pureza promedio del producto x_3 y las toneladas fabricadas del producto x_4 . Se dispone de datos históricos que se presentan en la siguiente tabla:

Y	x1	x2	x3	x4
240	25	24	91	100
236	31	21	90	95
290	45	24	88	110
274	60	25	87	88
301	65	25	91	94
316	72	26	94	99
300	80	25	87	97
296	84	25	86	96
267	75	24	88	110
276	60	25	91	105
288	50	25	90	100
261	38	23	89	98

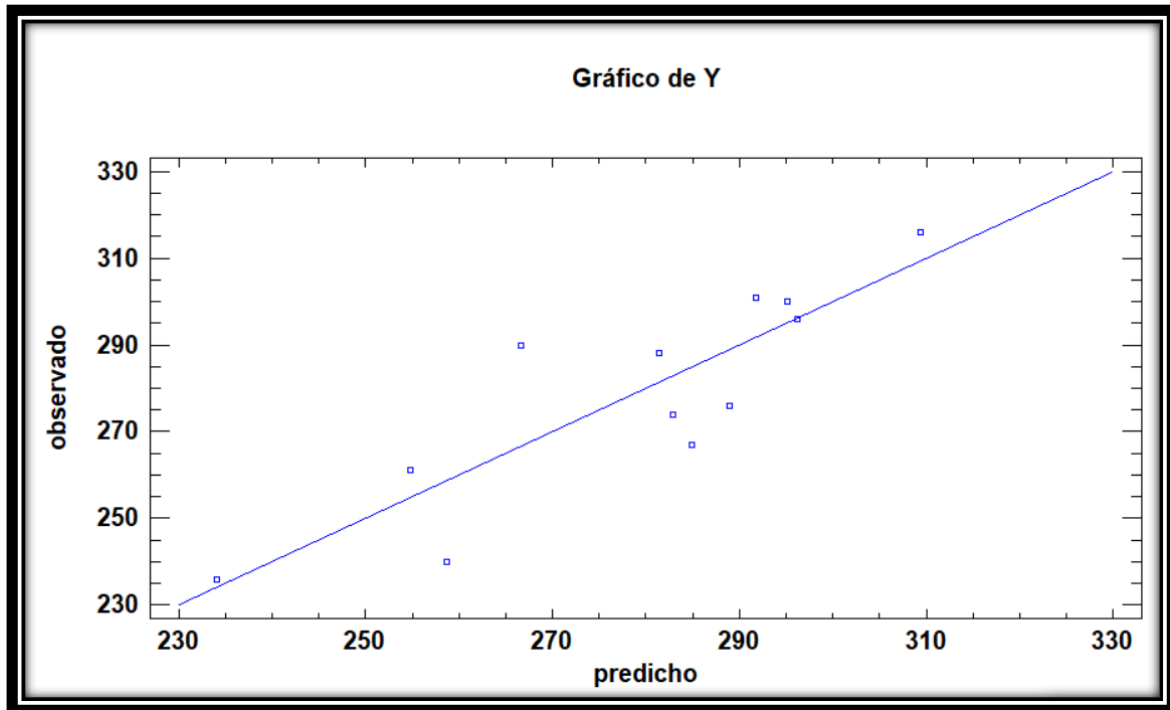
- a) Determinar el modelo de regresión.

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_p X_p + \varepsilon$$

$$Y = -102.731 + 0.65371X_1 + 8.92364X_2 + 1.43746X_3 + 0.0136093X_4$$

- b) Analizar los supuestos del modelo. Interprete.

➤ Linealidad:



Como podemos visualizar en la gráfica anterior que si hay linealidad positiva ya que los datos se ajustan a la línea de la tendencia.

➤ No autocorrelación y homocedasticidad:

H_0 = Los errores no están autocorrelacionados (los errores son independientes).

H_1 = Los errores están autocorrelacionados (los errores no son independientes).

Estadístico de Durbin-Watson

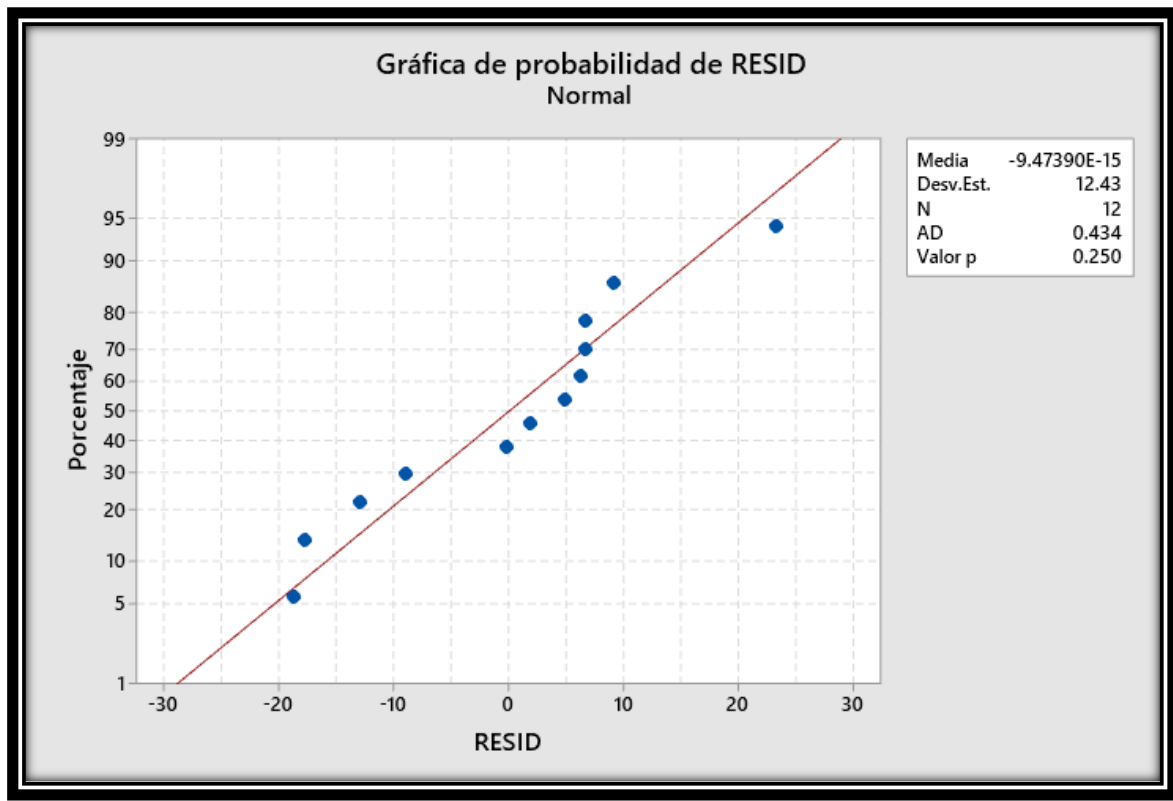
Estadístico de Durbin-Watson = 1.77159

Como el D.W = 1.77159 está contenido dentro de los intervalos de 1.5 a 2.5, entonces se asume que los errores son independientes o no existen de los errores. Además, se asume que los errores tienen una varianza constante (homocedasticidad).

➤ Normalidad:

H_0 = Los errores siguen una distribución normal.

H_1 = Los errores no siguen una distribución normal.



A.D. = 0.434, con un valor $p = 0.250$

Decisión: Como valor $p > \alpha$; no se rechaza H_0 .

Conclusión: Al 5% de nivel de significancia, se asume que los errores siguen una distribución normal. Por lo tanto, se cumple el supuesto de normalidad de los errores.

c) Hacer el análisis respectivo del planeamiento de hipótesis. Interprete.

✓ **Planteamiento de hipótesis:**

$$H_0: \beta_1; \beta_2; \beta_3; \beta_4 = 0$$

$$H_1: \beta_1; \beta_2; \beta_3; \beta_4 \neq 0$$

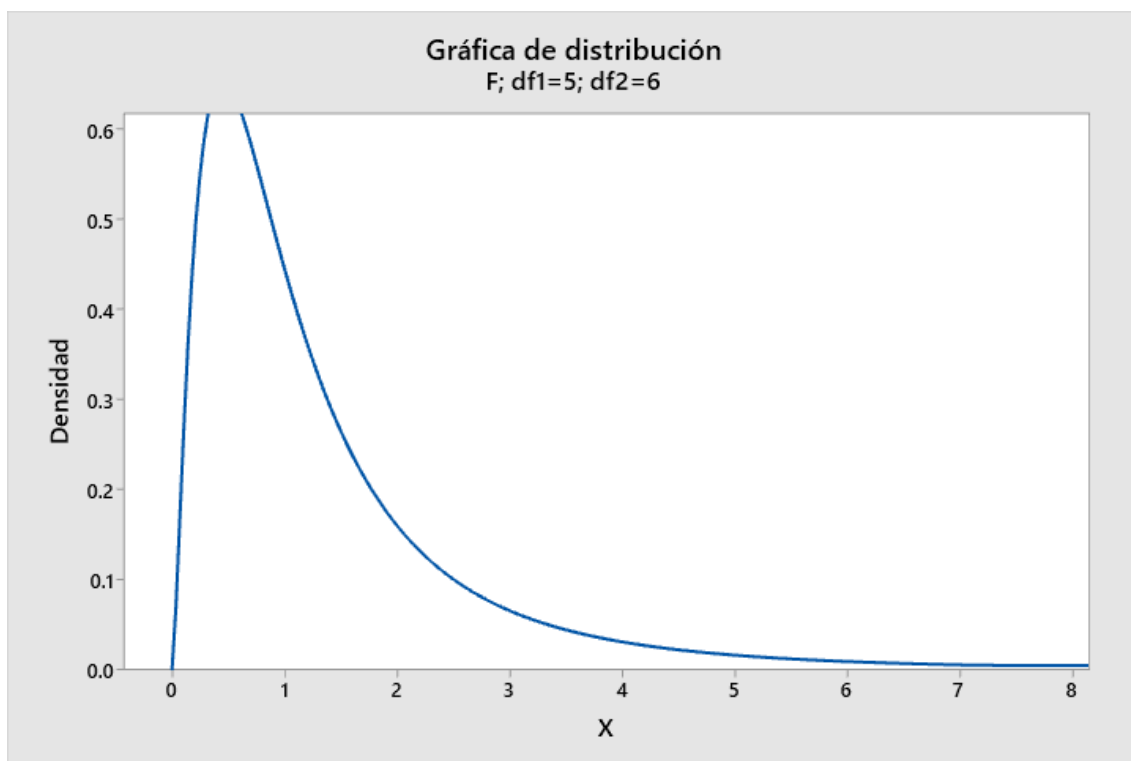
✓ **Nivel de significancia:**

$$\alpha = 0.05$$

✓ **Estadístico de prueba:**

Prueba F

✓ **Región de Rechazo:**



✓ **Decisión:**

Acepto H_0 , asumiendo que todas las variables independientes son iguales a cero.

d) Determine e interprete el coeficiente de determinación.

Resumen del modelo

S	R-cuad.	R-cuad. (ajustado)	R-cuad. (pred)
15.5793	74.47%	59.89%	7.75%

- El 74.47% de la variabilidad de energía eléctrica consumida es explicada por temperatura ambiental promedio, el número de días del mes, pureza promedio del producto y las toneladas fabricadas del producto.

A. Terry and Associates.

El tiempo para entregar equipos médicos Terry and Associates es un centro especializado en análisis médicos en Denver, Colorado. Una de sus principales fuentes de ingreso es un equipo que se utiliza para detectar grandes cantidades de plomo en la sangre. Los empleados de los talleres automotrices, de la industria jardinera y los pintores comerciales de casas están expuestos a grandes cantidades de plomo y, por tanto, se deben someter a la prueba en forma aleatoria. Es costoso realizar la prueba, por tanto, los equipos se entregan bajo pedido en diversos lugares en el área de Denver. Kathleen Terry, la propietaria, está preocupada por el costo adecuado de cada entrega. Para investigar esto, la señorita Terry reunió información sobre una muestra aleatoria de 50 entregas recientes. (Véase la Tabla de la columna siguiente.) Los factores relacionados con el costo de entrega del equipo son:

- Prep. El tiempo que transcurre entre el pedido hecho por teléfono a la compañía y el momento en que el equipo está listo para entregarse.*
- Entrega El tiempo de recorrido real de la planta de Terry al cliente*
- . • Millaje La distancia en millas de la planta de Terry al cliente.*

<i>Muestra</i>	<i>Costo</i>	<i>Prep</i>	<i>Entrega</i>	<i>Millaje</i>
1	32.6	10	51	20
2	23.37	11	33	12
3	31.49	6	47	19
4	19.31	9	18	8
5	28.35	8	88	17
6	22.63	9	20	11
7	22.63	9	39	11
8	21.53	10	23	10
9	21.16	13	20	8
10	21.53	10	32	10
11	28.17	5	35	16
12	20.42	7	23	9
13	21.53	9	21	10
14	27.55	7	37	16
15	23.37	9	25	12
16	21.18	8	23	10
17	21.54	11	21	9
18	15.99	8	13	4
19	17.96	12	12	4
20	25.22	6	41	4
21	24.29	3	28	13
22	22.76	4	26	10
23	28.17	9	54	16
24	19.68	7	18	8
25	25.15	6	50	13
26	20.36	9	19	7
27	21.16	3	19	8
28	25.95	10	45	14
29	18.76	12	12	5
30	18.76	8	16	5
31	24.29	7	35	13
32	19.56	2	12	6
33	22.63	8	30	11
34	21.16	5	13	8
35	21.16	11	20	8
36	19.68	5	19	8
37	18.76	5	14	7
38	17.96	5	11	4
39	23.37	10	25	12
40	25.22	6	32	14
41	27.06	8	44	16
42	21.96	9	28	9
43	22.63	8	31	11
44	19.68	7	19	8
45	22.76	8	28	10
46	21.96	13	18	9
47	25.95	10	32	14
48	26.14	8	44	15
49	24.29	8	34	13
50	24.35	3	33	12

1. *Analice los supuestos del modelo para luego establecer el modelo que explique la relación entre estas variables.*

❖ No autocorrelación y homocedasticidad:

H_0 = *Los errores no están autocorrelacionados (los errores son independientes).*

H_1 = *Los errores están autocorrelacionados (los errores no son independientes).*

Estadístico de Durbin-Watson

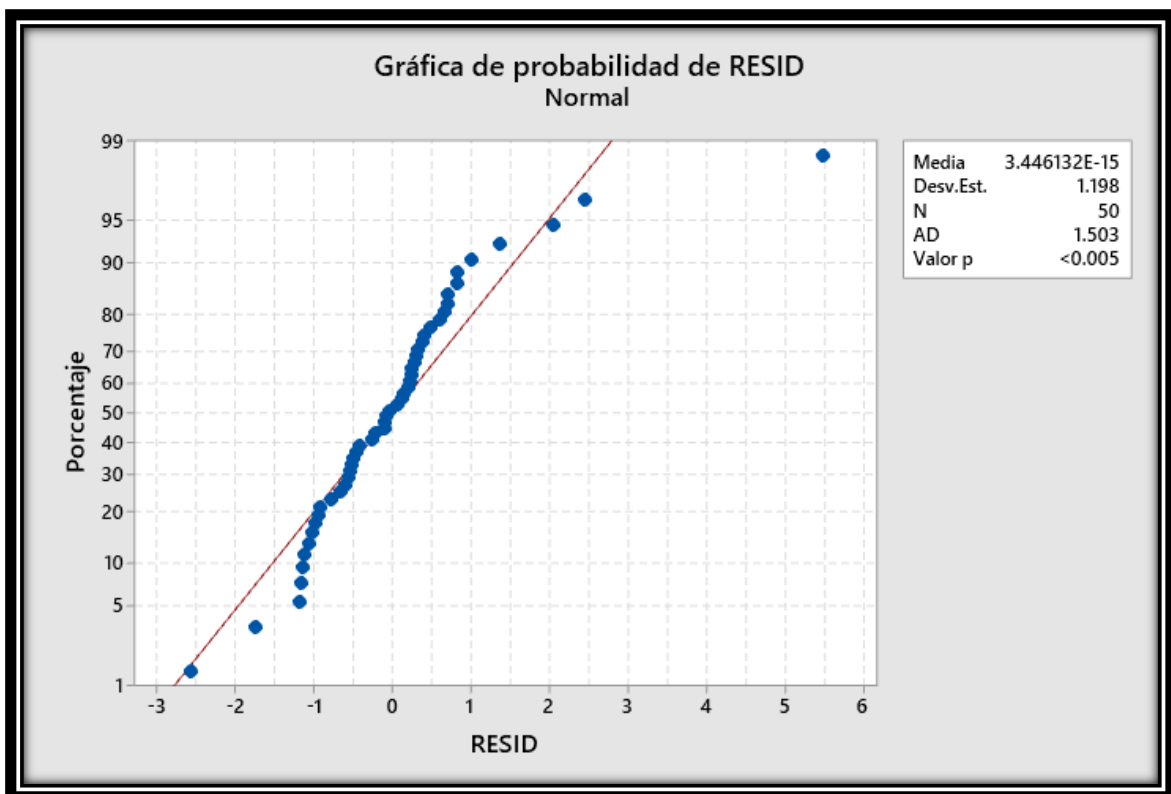
Estadístico de Durbin-Watson = 2,05963

Observamos que el valor del estadístico Durbin – Watson es 2,06 y se encuentra entre los valores de 1 a 3 entonces aceptamos la hipótesis nula, llegando a concluir que los errores son independientes y si hay homocedasticidad.

❖ Normalidad:

H_0 : *Los errores se distribuyen de manera normal.*

H_1 : *Los errores no se distribuyen de manera normal.*



A.D. = 1.503, con un valor $p < 0.005$

Decisión: Como valor $p < \alpha$; se rechaza H_0 .

Conclusión: Al 5% de nivel de significancia, se asume que los errores no siguen una distribución normal. Por lo tanto, no se cumple el supuesto de normalidad de los errores.

2. Desarrolle una ecuación de regresión lineal múltiple que describa la relación entre el costo de entrega y las otras variables.

$$Y = 14,7301 - 0.0449X_1 + 0.0685X_2 + 0.6191X_3$$

- *¿Estas tres variables explican una cantidad razonable de la variación de la variable dependiente?*

Resumen del modelo

S	R-cuad.	R-cuad. (ajustado)	R-cuad. (pred)
1.23668	87.99%	87.21%	77.07%

Dado que el coeficiente de determinación nos da el valor aproximado del 87.99% podemos concluir que las variables independientes si inciden con respecto a la variable dependiente (Costo de entrega del equipo).

- *Calcule el costo de entrega para un equipo que incluye 10 minutos de preparación, 30 minutos de entrega y que debe cubrir una distancia de 14 millas.*

$$Y = 14,7301 - 0.0449(10) + 0.0685(30) + 0.6191(14)$$

$$Y = 25.0035$$