

NOMBRE: Jennifer Priscila de León Flores y Valeria García Salazar

FECHA: 2/Julio/2021

Considere los siguientes datos que corresponden al cloro residual en una alberca en diversos momentos después de haberse tratado con químicos:

número de horas	cloro residual (partes por millón)
2	1.8
4	1.5
6	1.4
8	1.1
10	1.1
12	0.9

Nota: a continuación, se enlistan los incisos con los elementos a revisar y seguido se muestra con letras cursivas la nota u observación buscando queden claras las indicaciones

- a) **Identifique la variable dependiente y la independiente; grafique los datos x vs y , describa la gráfica mencionando si es posible la existencia de relación entre las variables. (Aquí la descripción es previa al análisis, sin trampa, a ojo de buen cubero)**

Variable dependiente: y = cloro residual

Variable independiente: x = número de horas



En la gráfica podemos observar la relación entre X (número de horas) y Y (cloro residual), mientras el número de horas aumenta, el cloro residual disminuye. Por lo que decimos que tiene una relación inversa.

- b) Realice un análisis de regresión a un nivel 0.05 de significancia y complete la tabla de modelos linealizables correspondiente a los datos. Identifique el mejor modelo justificando dicha elección. (Pegar y llenar la tabla que te mostré en el archivo “modelos linealizables”)

TABLA DE MODELOS LINEALIZABLES						
MODELO	Ecuación estimada	Modelo lineal asociado	Prueba de significancia			R ²
			Hipótesis	p valor	Conclusión	
Lineal	y estimada=-0.085751+1.9	*****	H ₀ : $\beta_1=0$ H _a : β_1 dif de 0	0.000864	Regresión significativa, β_1 dif de 0	95.23%
Potencia	y estimada=2.43x ^{-0.3663}	y=-0.3663x+0.8913	H ₀ : $\beta_1=0$ H _a : β_1 dif de 0	0.002209	Regresión significativa, β_1 dif de 0	92.42%
Exponencial	y estimada=2.012 ^{-0.066x}	y=0.699-0.066x	H ₀ : $\beta_1=0$ H _a : β_1 dif de 0	0.000491	Regresión significativa, β_1 dif de 0	96.40%
Logaritmo	y estimada=2.17-0.4862ln(x)	y=-0.4862x+2.17	H ₀ : $\beta_1=0$ H _a : β_1 dif de 0	0.00058	Regresión significativa, β_1 dif de 0	96.09%
Recíproco	y estimada=x/(1.031x+1.082)	y=-1.082x+1.031	H ₀ : $\beta_1=0$ H _a : β_1 dif de 0	0.037361	Regresión significativa, β_1 dif de 0	70.18%

El modelo que más se ajusta a estos datos es el exponencial debido a que tiene un mayor porcentaje de R² con 96.40%.

- c) Calcule los intervalos de confianza del 95% para los parámetros de la regresión elegida y comente al respecto. (De la tabla anterior deberás elegir el mejor y después calcular sólo para ese que elegiste los intervalos de confianza)

Los intervalos de confianza para el modelo Exponencial con una significancia del 0.05 son:

	Intervalo	
	Li	Ls
B₀	0.56118824	0.8379456
B₁	-0.08401422	-0.04848186

El intervalo de confianza de β_1 representa la potencia multiplicada por x de la base del modelo exponencial.

El intervalo de confianza de β_0 representa la base del modelo exponencial.

- d) En caso de existir evidencia de regresión al origen recalcule la ecuación de regresión bajo el modelo elegido y realice la prueba de significancia. (si existe evidencia de regresión el modelo generado será mi nuevo modelo elegido, aquí incluyes la ecuación del modelo y la prueba de significancia)**

Como se puede observar en la tabla del inciso c), el intervalo de confianza de β_0 no contiene el origen, es decir no contiene al 0, por lo tanto, NO existe evidencia de regresión al origen.

- e) Calcule los residuales y residuales estandarizados del modelo elegido e indique la presencia de datos atípicos. (Muestras los residuales y luego haces la comparación para datos atípicos y de existir mencionas cuales son)**

Análisis de residuos

Observación	Pronóstico Y*	Residuos	Residuos estándares
1	0.567070838	0.020715827	0.432616644
2	0.43457476	-0.02910965	-0.607908151
3	0.302078681	0.034393555	0.718253955
4	0.169582603	-0.07427242	-1.551059829
5	0.037086525	0.058223655	1.215907179
6	-0.095409554	-0.00995096	-0.207809798

Según los resultados estándares, no hay datos atípicos pues ninguno es mayor a 2 o menor a -2.

- f) Analice el cumplimiento de supuestos de los residuales incluyendo gráficas y comentarios al respecto (se menciona qué revisas, cómo se revisa y si se cumple o no)**

MEDIA CERO Y DISTRIBUCION NORMAL

# CLASES	2.45	3
MIN	-	
Max	0.07427242	
Rango	0.05822366	
Ancho	0.13249608	
	0.04416536	0.0442

MEDIA	0
DESV EST	0.04788495

Pruebas de hipótesis:

H0: Los residuales provienen de población normal con media cero

H1: Los residuales provienen de alguna otra distribución

Clases	Li	Ls	Frecuencia	Probabilidad	Esperado	Cociente
1	-	-	1	0.264997358	1.58998415	0.21892123
2	-	-	2	0.351017797	2.10610678	0.00534572
3	-	-	3	0.383984845	2.30390907	0.21031324
			6	1	6	0.43458019 EP

EP	0.43458019
X2 de tabla	3.84145882

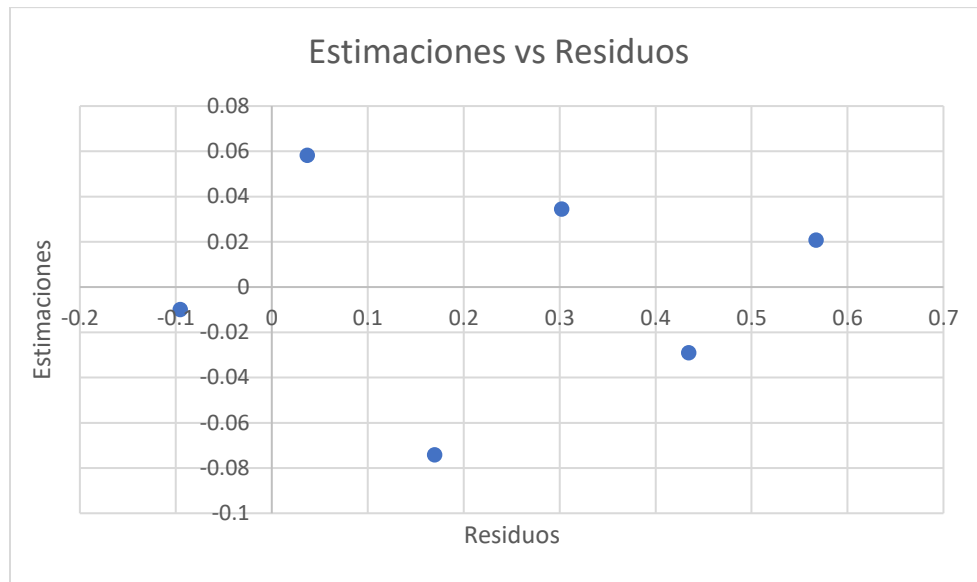
Región de rechazo:

Rechazo H0 si $EP=0.4345 > X2 \text{ de tabla}=3.8414$

NO rechazo H0

Los residuales provienen de distribución normal con media cero con una confianza del 95%.

VARIANZA CONTANTE



De la gráfica no se aprecian patrones visibles, por lo que se cumple varianza constante para residuales.

INCORRELACION

Pruebas de hipótesis:

H0: Los residuales están incorrelacionados, independientes entre si

H1: Los residuales están correlacionados

Residuos	c/u cuadrado	Restas	Cuad restas
0.02071583	0.000429145		
-0.02910965	0.000847372	-0.04982548	0.002482578
0.03439356	0.001182917	0.06350321	0.004032657
-0.07427242	0.005516393	-0.10866598	0.011808295
0.05822366	0.003389994	0.13249608	0.017555211
-0.00995096	9.90216E-05	-0.06817462	0.004647778
	0.011464842	<--denominador	0.04052652 <--numerador

EP=	3.53485186
-----	------------

alfa =
0.05

	dL	dU
6	?	?
15	1.08	1.36
20	1.2	1.41

Podemos notar que $EP > dL$ y además $EP < dU$, esto es, $dL < dU < EP$. Entonces:

NO Rechazo H0

Con 95% de confianza los residuales están incorrelacionados, son independientes.

Conclusión de supuestos:

DADO LO ANTERIOR SE TIENE QUE LOS RESIDUALES CUMPLEN CON LOS SUPUESTOS DEL MODELO EXPONENCIAL.

- g) Grafique los datos x vs y sobreponiendo la ecuación del modelo elegido. (La misma grafica del inciso a pero con la línea de ajuste del modelo que elegiste)

X	Y
Número de horas	Cloro residual (partes por millón)
2	1.8
4	1.5
6	1.4
8	1.1
10	1.1
12	0.9

