

Universidad del Valle de Guatemala  
5 de febrero de 2021  
Rudik R. Rompich - Carné: 19857

Estadística 2 - Eugenio Aristondo

## Hoja de trabajo 1

La resistencia a la tensión de cierto elemento estructural es una característica importante. La tabla siguiente da información de la resistencia a la tensión ( $y$ ), el espesor de la soldadura ( $x_1$ ), la altura del poste ( $x_2$ ), el grosor del elemento ( $x_3$ ), la longitud del elemento ( $x_4$ ), el ancho de la unión ( $x_5$ ) y el ancho del poste ( $x_6$ ).

### 1. I

Elabore la matriz de correlación para este conjunto de datos y elija como regresores preliminares aquellos cuya correlación con los demás es menor o igual a  $|0.30|$ .

	$y$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
$y$	1						
$x_1$	0.323060711	1					
$x_2$	-0.2496878	-0.06432696	1				
$x_3$	0.690828596	0.235210518	0.00307297	1			
$x_4$	-0.55743632	-0.0689409	0.446279737	-0.22477219	1		
$x_5$	-0.33477435	-0.21904291	0.469937292	-0.0404318	0.400760914	1	
$x_6$	-0.24809027	-0.38948249	0.1927617	-0.21524481	0.384335731	0.429373254	1

### 2. II. RLS

a) Encuentre el mejor modelo de regresión lineal simple. Y verifique los supuestos.

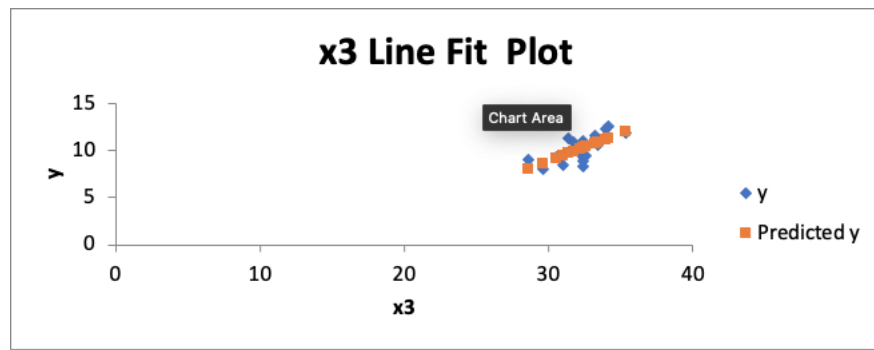
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.6908286							
R Square	0.47724415							
Adjusted R Square	0.44649381							
Standard Error	1.01072192							
Observations	19							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	1	15.854553	15.854553	15.5199612	0.00105705			
Residual	17	17.3664996	1.0215588					
Total	18	33.2210526						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-8.9710443	4.85449081	-1.8479887	0.08207828	-19.213125	1.27103606	-19.213125	1.27103606
$x_3$	0.5930553	0.15053929	3.9395382	0.00105705	0.27544515	0.91066544	0.27544515	0.91066544

Lo que implica que el mejor modelo de regresión lineal simple es:

$$y = 0,593055296x_3 - 8,971044274 \quad (1)$$

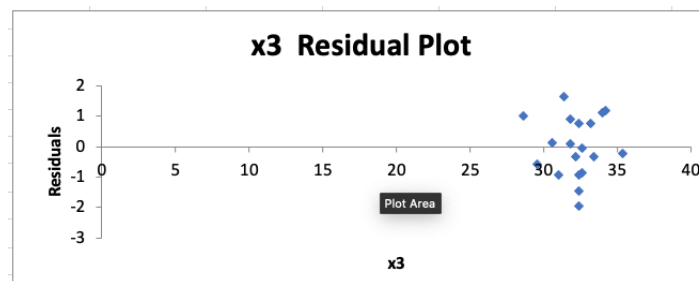
Por otra parte, los supuestos:

- Linealidad: Según el  $R^2$  es de 0.47724415, lo que implica que es un modelo con deficiencias y no tan preciso. El supuesto, sin embargo, se cumple.



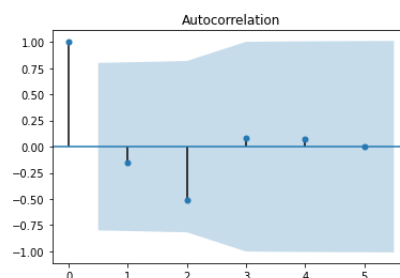
- Homocedasticidad: La variación de residuos es el mismo para cada valor de X.

Como se observa en el análisis de residuos, parece indicar que casi todos los puntos cumplen el supuesto, aunque el residuo de 10.5 indica una leve diferencia en este supuesto.



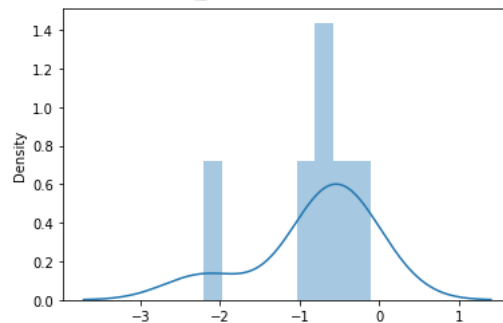
- Independencia: Las observaciones son independientes de cada una.

Como se observa en el análisis de independencia, el supuesto parece cumplirse.

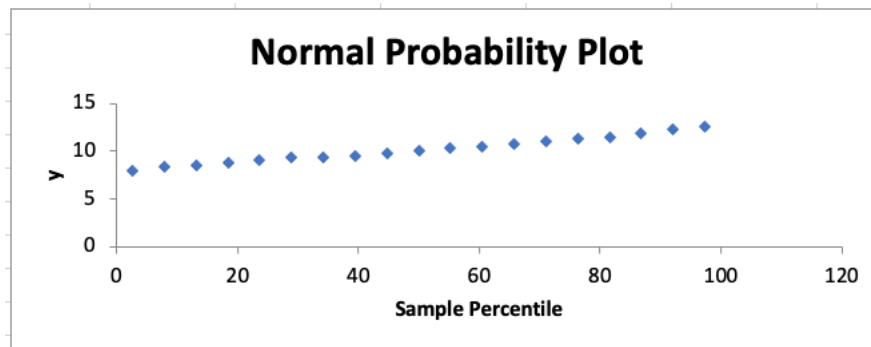


- Normalidad: Por cada valor arreglado de X, Y es normalmente distribuida.

Es una distribución normal con una leve variación.



El análisis de promedios parece mostrar que el supuesto es correcto.



### 3. III. RLM

b) Haga el análisis completo de regresión lineal múltiple mediante la eliminación en reversa tomando en cuenta todos los regresores y las interacciones dobles solamente de aquellos regresores cuya correlación excede el  $|0.30|$

	y	x1	x3	x4	x5	x1x3	x1x4	x1x5	x3x4	x3x5	x4x5
y	1										
x1	0.32306071	1									
x3	0.6908286	0.23521052	1								
x4	-0.55743632	-0.0689409	-0.22477219	1							
x5	-0.33477435	-0.21904291	-0.0404318	0.40076091	1						
x1x3	0.59596274	0.86017706	0.69742834	-0.16336122	-0.18154909	1					
x1x4	-0.07774459	0.78274873	0.0608108	0.56549097	0.0683037	0.6118557	1				
x1x5	0.07911223	0.78011131	0.1907737	0.19296213	0.43803983	0.67561144	0.76503966	1			
x3x4	0.07256364	0.13521376	0.58692769	0.65624951	0.31280649	0.41127377	0.52567634	0.32446855	1		
x3x5	0.24955297	0.01619498	0.6836967	0.13419881	0.7011913	0.37081884	0.10135064	0.46322553	0.64902673	1	
x4x5	-0.53099338	-0.16968812	-0.15616582	0.85052933	0.82187992	-0.20311817	0.38914725	0.36908577	0.59186229	0.48973772	1

c) Plantee el modelo final.

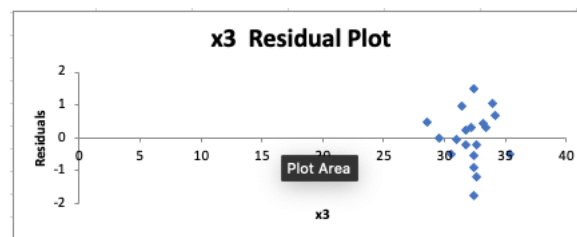
SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.804723826							
R Square	0.647580437							
Adjusted R Sq	0.603527991							
Standard Error	0.855414697							
Observations	19							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	2	21.51330377	10.75665188	14.7002154	0.000237947			
Residual	16	11.70774886	0.731734304					
Total	18	33.22105263						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	4.656305295	6.394818452	0.728137214	0.477051654	-8.90010423	18.21271482	-8.90010423	18.21271482
x3	0.511325821	0.130753272	3.910615879	0.001245605	0.234141267	0.788510374	0.234141267	0.788510374
x4	-0.12418333	0.044655992	-2.7808885	0.013357398	-0.21884981	-0.02951686	-0.21884981	-0.02951686

d) Plantee la ecuación ajustada.

$$y = 0,511325821x_3 - 0,1241833x_2 + 4,656305295$$

e) Haga análisis completo de residuos.

Análisis de residuos de la variable  $x_3$



Análisis de residuos de la variable  $x_4$

