

REGRESION LINEAL SIMPLE

REGRESION LINEAL

- *En la búsqueda de mejoras o en la solución de problemas es necesario, frecuentemente, investigar la relación entre factores (o variables). Para lo cual existen varias herramientas estadísticas, entre los que se encuentran el diagrama de dispersión, el análisis de correlación y el análisis de regresión.*
- *El análisis de regresión puede usarse para explicar la relación de un factor con otro(s). Para ello, son necesarios los datos, y estos pueden obtenerse de experimentos planeados, de observaciones de fenómenos no controlados o de registros históricos.*

Regresión lineal simple

- *Sean dos variables X y Y . Supongamos que se quiere explicar el comportamiento de Y con el de X . Para esto, se mide el valor de Y sobre un conjunto de n valores de X , con lo que se obtienen n parejas de puntos $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$.*
- *A Y se le llama la variable dependiente o la variable de respuesta y a X se le conoce como variable independiente o variable regresora.*
- *Supongamos que las variables X y Y están relacionadas linealmente y que para cada valor de X , Y es una variable aleatoria. Es decir, supongamos que cada observación de Y puede ser descrita por el modelo.*

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

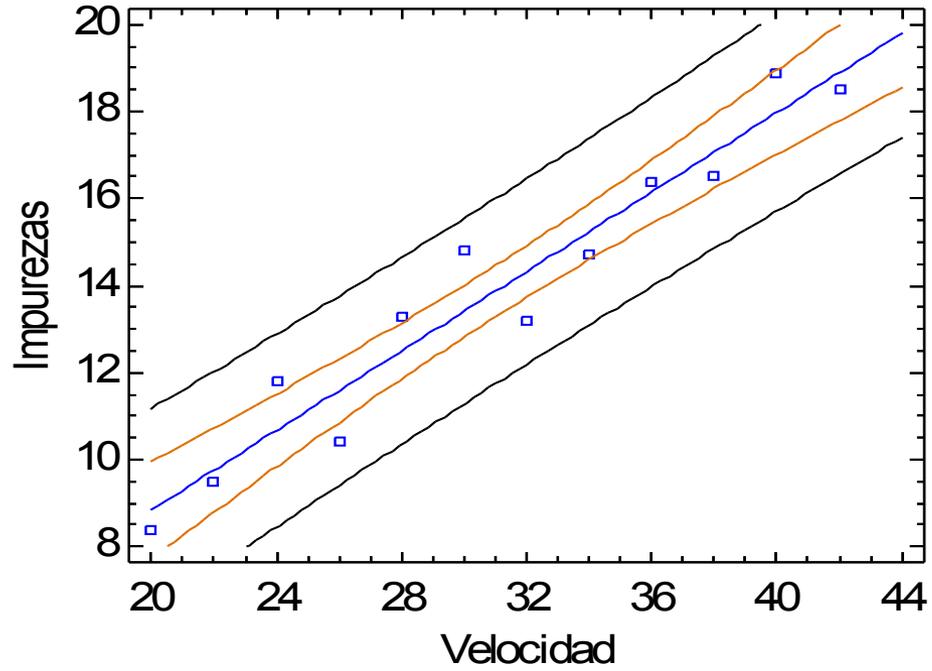
Ejemplo 5

En una fábrica de pintura se desea investigar la relación entre la velocidad de agitación X y el porcentaje de impurezas en la pintura Y . Mediante un diseño experimental se obtienen los siguientes datos.

<i>Velocidad</i>	<i>Impurezas</i>
20	8.4
22	9.5
24	11.8
26	10.4
28	13.3
30	14.8
32	13.2
34	14.7
36	16.4
38	16.5
40	18.9
42	18.5

Gráfico del Modelo Ajustado

$$\text{Impurezas} = -0.289277 + 0.456643 \cdot \text{Velocidad}$$



	<i>Mínimos Cuadrados</i>	<i>Estándar</i>	<i>Estadístico</i>	
<i>Parámetro</i>	<i>Estimado</i>	<i>Error</i>	<i>T</i>	<i>Valor-P</i>
<i>Intercepto</i>	-0.289277	1.22079	-0.236959	0.8175
<i>Pendiente</i>	0.456643	0.0384385	11.8798	0.0000

<i>Fuente</i>	<i>Suma de Cuadrados</i>	<i>Gl</i>	<i>Cuadrado Medio</i>	<i>Razón-F</i>	<i>Valor-P</i>
<i>Modelo</i>	<i>119.275</i>	<i>1</i>	<i>119.275</i>	<i>141.13</i>	<i>0.0000</i>
<i>Residuo</i>	<i>8.45142</i>	<i>10</i>	<i>0.845142</i>		
<i>Total (Corr.)</i>	<i>127.727</i>	<i>11</i>			

Coeficiente de Correlación = 0.96635

R-cuadrada = 93.3832 por ciento

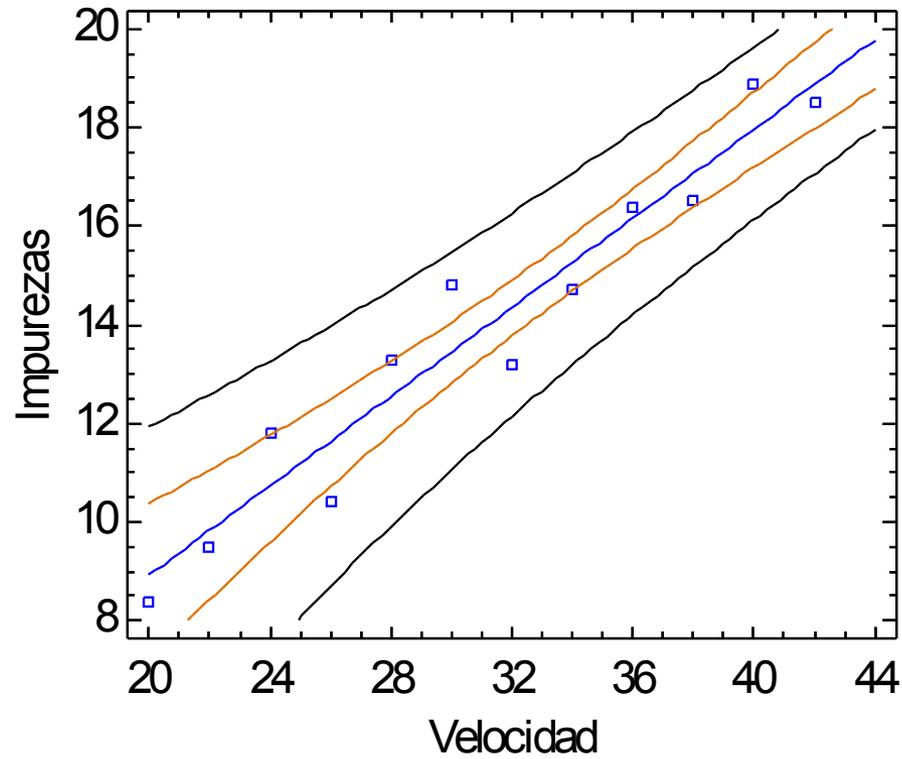
		95.00%		95.00%	
	Predicciones	Límite	Predicción	Límite	Confianza
X	Y	Inferior	Superior	Inferior	Superior
20	8.84359	6.5127	11.1745	7.73128	9.9559
28	12.4967	10.3493	14.6442	11.852	13.1415
35	15.6932	13.5339	17.8526	15.0099	16.3766
40	17.9765	15.7094	20.2435	17.005	18.948

Modelo	Correlación	R-Cuadrada
Cuadrado Doble	0.9668	93.48%
Raíz Cuadrada de X	0.9666	93.43%
Raíz Cuadrada-Y Log-X	0.9665	93.40%
Lineal	0.9663	93.38%
Raíz Cuadrada Doble	0.9658	93.27%
Multiplicativa	0.9651	93.14%
Curva S	-0.965	93.13%
Logaritmo de X	0.9644	93.02%
Doble Inverso	0.9639	92.91%
Cuadrado de Y	0.9637	92.87%
Raíz Cuadrada de Y	0.9628	92.69%
Logarítmico-Y Raíz Cuadrada-X	0.9615	92.46%
Raíz Cuadrada-Y Inversa de X	-0.9606	92.28%
Cuadrado de X	0.9593	92.03%
Cuadrado-Y Raíz Cuadrada-X	0.9586	91.89%
Exponencial	0.9557	91.34%
Inversa de X	-0.953	90.82%
Inversa-Y Log-X	-0.9524	90.71%
Cuadrado-Y Log-X	0.9511	90.46%
Raíz Cuadrada-X Cuadrado-X	0.9505	90.34%
Inversa-Y Raíz Cuadrada-X	-0.9433	88.98%
Log-Y Cuadrado-X	0.9382	88.03%
Inversa de Y	-0.9321	86.87%
Cuadrado-Y Inversa de X	-0.9292	86.34%
Inversa-Y Cuadrado-X	-0.9046	81.82%
Logístico	<sin ajuste>	
Log probit	<sin ajuste>	

MODELO CUADRADO DOBLE

Gráfico del Modelo Ajustado

$$\text{Impurezas} = \sqrt{-1.51282 + 0.202685 \cdot \text{Velocidad}^2}$$



MODELO CUADRADO DOBLE

	<i>Mínimos Cuadrados</i>	<i>Estándar</i>	<i>Estadístico</i>	
<i>Parámetro</i>	<i>Estimado</i>	<i>Error</i>	<i>T</i>	<i>Valor-P</i>
<i>Intercepto</i>	-1.51282	18.5639	-0.0814926	0.9367
<i>Pendiente</i>	0.202685	0.0169294	11.9723	0.0000

<i>Fuente</i>	<i>Suma de Cuadrados</i>	<i>Gl</i>	<i>Cuadrado Medio</i>	<i>Razón-F</i>	<i>Valor-P</i>
<i>Modelo</i>	91204.9	1	91204.9	143.34	0.0000
<i>Residuo</i>	6362.97	10	636.297		
<i>Total (Corr.)</i>	97567.9	11			

Coeficiente de Correlación = 0.966842

R-cuadrada = 93.4784 por ciento