# APLICACIÓN DE LA REGRESIÓN LINEAL Y NO LINEAL A PROBLEMAS DE INGENIERÍA

Teresa Pérez Sosa[a], Dirma Yanes Quintero[b]

*a. Centro de Estudio de Fabricación Avanzada y Sostenible (CEFAS). Universidad de Matanzas.*

*b. Departamento de Matemática. Universidad de Matanzas.*

**Resumen:** Los modelos lineales tienen la ventaja de ser fácilmente interpretables, sin embargo, pueden tener limitaciones importantes en capacidad predictiva, mientras que la regresión no lineal genera una ecuación para describir la relación no lineal entre una variable de respuesta continua y una o más variables predictoras y predice nuevas observaciones. Esta última será utilizada cuando no se pueda modelar adecuadamente la relación con parámetros lineales. En el presente trabajo se describe la aplicación de modelos lineales y no lineales en problemas de ingeniería, pues la correcta elección de un modelo adecuado, que describa los datos, proporciona elementos de juicio suficientes para la toma de decisiones en condiciones de incertidumbre. Determinándose para ello, qué modelo utilizar para representar los datos y ajustarlos al modelo más adecuado teniendo en cuenta la bondad del ajuste dentro del rango de datos medidos experimentalmente y/o fuera del dicho rango (debido al carácter predictivo de las funciones).

***Palabras claves***: regresión lineal, regresión no lineal, problemas de ingeniería.

**Abstract:** The lineal models have the advantage of being easily interpretable, however, they can have important limitations in capacity to predict, while the not lineal regression generates an equation to describe the not lineal relationship between a variable of continuous answer and an or more variables that predict and it predicts new observations. This last one will be used when you cannot model the relationship appropriately with lineal parameters. Present work describes the application of lineal and not lineal models in engineering problems, because the correct election of an appropriate model that describes the data, provides enough trial elements for the taking of decisions under conditions of uncertainty. Being determined for it, how model to use to represent the data and to adjust them to the most appropriate pattern keeping in mind the kindness of the adjustment inside the range of data measured experimentally and/or outside of the this range (due to the capacity to predict of the functions).

***Keywords***: lineal regression, not lineal regression, engineering problems.

## 1. Introducción.

Las diferentes ciencias por su naturaleza experimental se han propuesto en las diferentes épocas que marcan nuestra historia, describir el mundo que nos rodea. Desde los griegos hasta la época actual, se ha pretendido encontrar diversos modelos matemáticos, que nos ayuden a comprender toda la naturaleza del universo, desde la química, física, astronomía, botánica, etc., se han buscado ecuaciones que permitan predecir el comportamiento de un determinado sistema.

En la mayoría de las investigaciones –específicamente en la rama de la ingeniería– en las cuales se realizan n mediciones, observaciones o experimentos de donde se obtengan datos de diferentes variables; es fundamental determinar algún tipo de relación de dependencia entre las variables con el fin de hacer predicciones o pronósticos de eventos futuros de acuerdo con el comportamiento de ellas. De ahí que el presente trabajo se proponga como objetivo analizar un ejemplo donde se aplique la regresión lineal y no lineal en un problema de ingeniería. [1]

En el análisis de regresión una de las dos variables, que llamamos X, puede considerarse como variable ordinaria, es decir se puede medir sin error apreciable. La otra variable Y, es una variable aleatoria. A X se la llama variable independiente (algunas veces variable controlada) y nuestro interés es la dependencia de Y en términos de X.

Supongamos que en cierto experimento aleatorio tratamos de manera simultánea dos variables, una variable ordinaria X y una variable aleatoria Y. Efectuamos el experimento de tal manera que seleccionamos primero n valores $x_1$,$x_2$ ,...,xn de X y luego para cada $X_j$ obtenemos un valor observado $Y_j$ de Y. Entonces, tenemos una muestra de n parejas de valores:

$$(x_1, y_1),(x_2, y_2)....(x_n, y_n)\ , \tag{1}$$

Podemos graficar las n parejas como puntos del plano. Con el objetivo es hallar alguna función que describa aproximadamente el diagrama de puntos anterior, en el rango considerado de la variable X. En primer lugar es importante elegir la función más adecuada. [2]

Las clases de funciones más utilizadas son las siguientes:

- Polinomiales

a) Lineales

$$f(x.a) = a_0 + a_1 x\ , \tag{2}$$

b) Cuadráticas

$$f(x.a) = a_0 + a_1 x + a_2 x^2\ , \tag{3}$$

c) En general de grado menor o igual a m

$$f(x.a) = a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m\ , \tag{4}$$

- Potenciales

$$f(x.a) = a_0 \cdot x^{a1}\ , \tag{5}$$

- Exponenciales

$$f(x.a) = a_0 \cdot (a_1)^x\ , \tag{6}$$

- Logarítmicas

$$f(x.a) = a_0 \cdot a_1 \ln x\ , \tag{7}$$

En la actualidad y dado el gran avance tecnológico, se han desarrollado poderosas herramientas computacionales que permiten desarrollar modelos con gran facilidad. SPSS es un poderoso y completo software estadístico que permite buscar simultáneamente entre diferentes modelos de carácter tanto lineal como no lineal, cual es el más adecuado para un problema en específico, el mismo será utilizado en el procesamiento de los datos del ejemplo que se propone a continuación.

## 2. Materiales y métodos.

Ejemplo:

Los siguientes datos se corresponden a las mediciones de la velocidad del aire y del coeficiente de evaporación de las gotas de combustible en una turbina de propulsión.

**Tabla 1.** Datos se corresponden a las mediciones de la velocidad del aire y del coeficiente de evaporación de las gotas de combustible en una turbina de propulsión.

| Velocidad del aire (cm/s) X | Coeficiente de evaporación ($mm^2$/s) Y |
|---|---|
| 20 | 0,18 |
| 60 | 0,37 |
| 100 | 0,35 |
| 140 | 0,78 |
| 180 | 0,56 |
| 220 | 0,75 |
| 260 | 1,18 |
| 300 | 1,36 |
| 340 | 1,17 |
| 380 | 1,65 |

En la aplicación de la regresión se tiene que $\hat{y} = a + bx$ es la recta de regresión lineal, el parámetro b se llama coeficiente de regresión.

Su valor expresa el incremento de $\hat{y}$ cuando x aumenta una unidad. Si b toma un valor positivo, la variable $\hat{y}$ crece al crecer x, y en consecuencia la recta es creciente, lo que indica que la dependencia entre las variables es directa.

Si b toma un valor negativo, la variable $\hat{y}$ decrece al crecer x, y en consecuencia la recta es decreciente, lo que indica que la dependencia entre las variables es inversa. Si b = 0, la recta es horizontal y no hay dependencia entre las variables, ya que las variaciones de x no provocan variación de $\hat{y}$.

Además se considera importante obtener el valor del Coeficiente de Determinación $R^2$ como la parte relativa de la variación total que viene explicada por el modelo.

$$R^2 = \frac{SC\,\mathrm{Re}\,g}{SCT}, \qquad (8)$$

- Toma valores entre 0 y 1 ($0 \leq R^2 \leq 1$).
- Todo ajuste mínimo cuadrático debe venir acompañado de su respectivo coeficiente de determinación para poder conocer el poder representativo de la función de ajuste, es decir el valor explicativo del modelo.
- Generalmente si $R^2 \geq 0.90$ se acepta el ajuste, en caso contrario se debe buscar otro modelo.

**3. Análisis de los resultados.**

- Regresión Lineal

Diagrama de dispersión: en el cual se indica la recta de regresión.

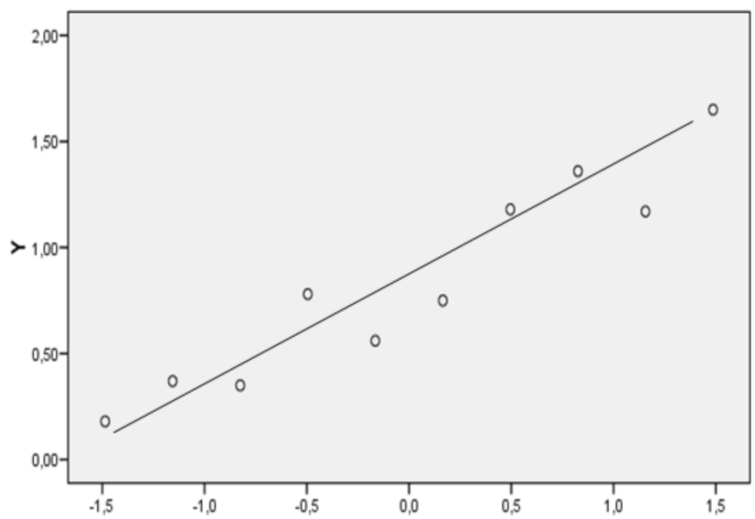


**Figura 1**. Diagrama de dispersión. Regresión Lineal.

Análisis de Varianza:

**Tabla 2.** Análisis de la varianza aplicado al modelo.

| Modelo | | Suma de cuadrados | gl | Media cuadrática | F | Sig. |
|---|---|---|---|---|---|---|
| 1 | Regresión | 1,935 | 1 | 1,935 | 76,492 | ,000(a) |
| | Residual | ,202 | 8 | ,025 | | |
| | Total | 2,137 | 9 | | | |

Proponiendo un nivel de significación $\alpha=0,05$

Como el valor de Sig de $F<0,05$ se rechaza la hipótesis nula, por lo tanto $\beta \neq 0$

Parámetros del Modelo:

**Tabla 3.** Parámetros del modelo lineal e intervalos de confianza del 95% para el coeficiente de regresión y la ordenada al origen

| Modelo | | Coeficientes no estandarizados | | Coeficientes estandarizados | t | Sig. |
|---|---|---|---|---|---|---|
| | | B | Error típ. | Beta | | |
| 1 | (Constante) | ,069 | ,101 | | ,686 | ,512 |
| | X | ,004 | ,000 | ,951 | 8,746 | ,000 |

Como el valor de Sig de t<0,05 se rechaza la hipótesis nula, por lo la velocidad del aire influye significativamente sobre el coeficiente de evaporación.

Coeficiente de Determinación:

**Tabla 4.** Coeficiente de determinación de la regresión lineal.

| Modelo | R | R cuadrado | R cuadrado corregida | Error típ. de la estimación |
|---|---|---|---|---|
| 1 | ,951(a) | ,905 | ,893 | ,15905 |

$R^2$=0,905≥0,90 por lo tanto la regresión lineal es un muy buen ajuste.

- Regresión Cuadrática

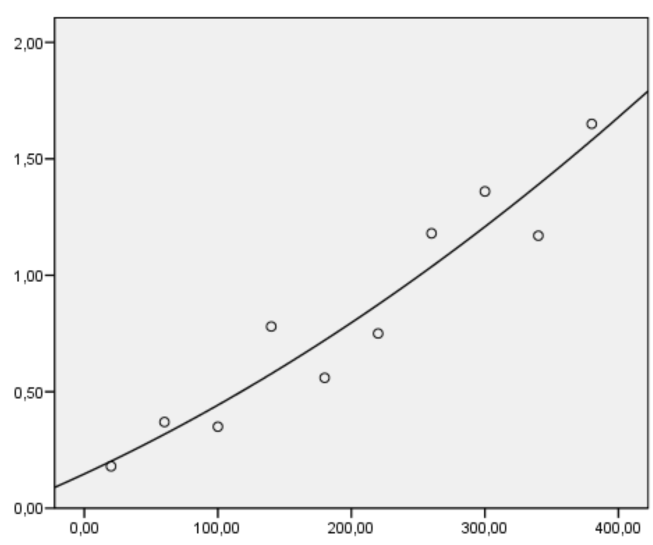


**Figura 2.** Diagrama de dispersión. Regresión Cuadrática.

**Tabla 5.** Resumen del modelo y estimaciones de los parámetros de la Regresión Cuadrática.

| Ecuación | Resumen del modelo | | | | | Estimaciones de los parámetros | | |
|---|---|---|---|---|---|---|---|---|
| | R cuadrado | F | gl1 | gl2 | Sig. | Constante | b1 | b2 |
| Cuadrático | ,911 | 35,684 | 2 | 7 | ,000 | ,147 | ,003 | 2,91E-006 |

- Regresión Potencial

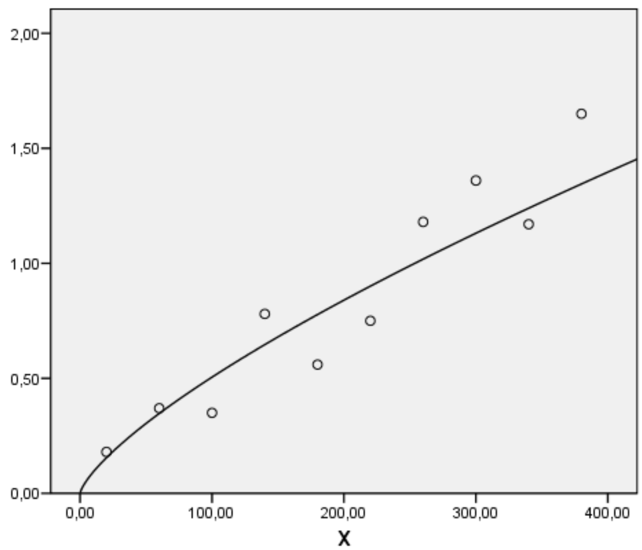


**Figura 3**. Diagrama de dispersión. Regresión Potencial.

**Tabla 6.** Resumen del modelo y estimaciones de los parámetros de la Regresión Potencial.

| Ecuación | Resumen del modelo | | | | | Estimaciones de los parámetros | |
|---|---|---|---|---|---|---|---|
| | R cuadrado | F | gl1 | gl2 | Sig. | Constante | b1 |
| Potencia | ,903 | 74,317 | 1 | 8 | ,000 | ,017 | ,734 |

- Regresión Exponencial

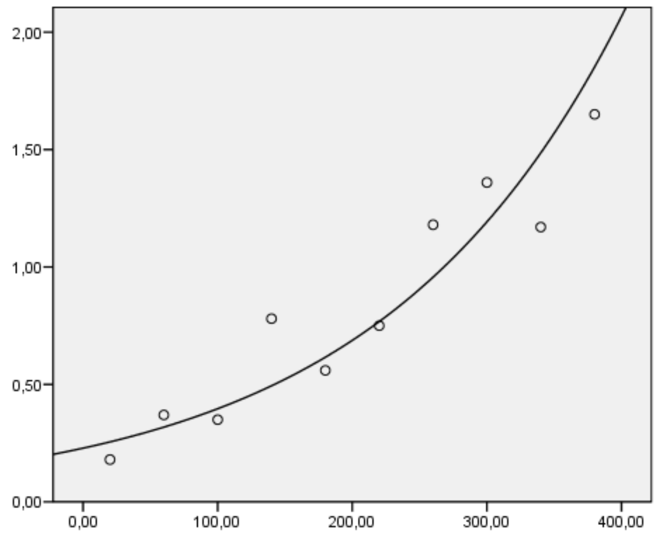


**Figura 4.** Diagrama de dispersión. Regresión Exponencial.

**Tabla 7.** Resumen del modelo y estimaciones de los parámetros de la Regresión Exponencial.

| Ecuación | Resumen del modelo | | | | | Estimaciones de los parámetros | |
|---|---|---|---|---|---|---|---|
| | R cuadrado | F | gl1 | gl2 | Sig. | Constante | b1 |
| Exponencial | ,887 | 62,684 | 1 | 8 | ,000 | ,229 | ,006 |

- Regresión Logarítmica

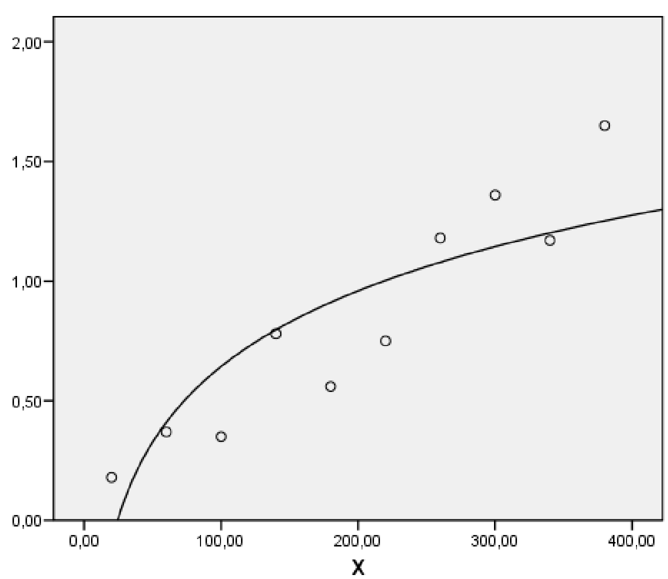


**Figura 5.** Diagrama de dispersión. Regresión Logarítmica.

**Tabla 8.** Resumen del modelo y estimaciones de los parámetros de la Regresión Logarítmica.

| Ecuación | Resumen del modelo | | | | | Estimaciones de los parámetros | |
|---|---|---|---|---|---|---|---|
| | R cuadrado | F | gl1 | gl2 | Sig. | Constante | b1 |
| Logarítmica | ,736 | 22,286 | 1 | 8 | ,002 | -1,457 | ,456 |

Comparación e las regresiones estimadas:

**Tabla 9.** Regresiones estimadas.

| Tipo de Regresión | Coeficiente de Determinación |
|---|---|
| Lineal | 0,905 |
| Cuadrática | 0,911 |
| Potencial | 0,903 |
| Exponencial | 0,887 |
| Logarítmica | 0,736 |

Es posible observar (Tabla 9) que las regresiones Lineal, Cuadrática y Potencial son muy buenas pues en los tres casos $R^2 \geq 0{,}90$. Sin embargo en los casos de la Exponencial y la Logarítmica las regresiones no se consideran buenas pues $R^2 < 0{,}90$.

El mejor ajuste es a través de una función cuadrática $R^2=0{,}911$

## 4. Conclusiones.

La aplicación de distintas regresiones sobre un mismo problema relacionado con la ingeniería nos permite realizar comparaciones, sin limitarse solamente al caso lineal.

La facilidad de que nos brindan las nuevas tecnologías permite en poco tiempo efectuar comparaciones para la correcta elección de un modelo adecuado, que describa los datos en problemas de ingeniería, así como nos proporciona elementos de juicio suficientes para la toma de decisiones en condiciones de incertidumbre.

## Referencias.


[1] J. Reyes, "Una aplicación de la Regresión no lineal a los Hormigones Reciclados", Tesis de Maestría en Ingeniería de los Materiales, 2016. Universidad de Alicante. España.

[2] C. Minnaard, "Modelos de regresión lineales y no lineales: su aplicación en problemas de ingeniería", Segundo Congreso Argentino de Ingeniería Mecánica, 2010.