# TEMA 8. MODELOS DE REGRESIÓN NO LINEAL.

Curso OCW de "Estadística Descriptiva con Excel para Grados de Ciencias Sociales"

Grupo Decisión Multicriterio Zaragoza (GDMZ)

## 1. REGRESIÓN NO LINEAL

Aunque la regresión lineal tiene aplicación en numerosos problemas, en ocasiones, bien la naturaleza de la relación que liga las variables, bien las características concretas de los datos, exigen la utilización de ajustes de funciones no lineales.

Para establecer otras funciones para la forma de dependencia de tipo no lineal, el procedimiento de ajuste en base al criterio de ajuste de mínimos cuadrados requiere adaptar la expresión del Error Cuadrático Medio a la correspondiente ecuación y su correspondiente optimización:

$$\boldsymbol{ECM} = \sum_{i=1}^{N}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{N}(y_i - f(x_i)\,)^2$$

No obstante, existen funciones matemáticas no lineales con la particularidad de que sus ecuaciones se pueden "reformular", en forma de una ecuación lineal, eso sí, entre otras variables. Este proceso se conoce como linealización y conlleva realizar un cambio de variable.

Por ejemplo, si la función a ajustar es una exponencial, es decir, $Y = ae^{bX}$, tomando logaritmos en ambos miembros obtenemos lo siguiente:

$$\ln Y = \ln a + b\,X$$

de manera que, llamando Y' = ln Y y a' = ln a, se tiene:

$$Y' = a' + b\,X$$

que vuelve a ser la expresión de un modelo de regresión lineal simple.

Análogamente, si la función a ajustar es potencial del tipo $Y = aX^b$, de nuevo tomando logaritmos en ambos miembros se tiene:

$$\ln Y = \ln a + b \ln X$$

de manera que llamando Y' = ln Y, a' = ln a y X' = ln X, se tiene:

$$Y' = a' + b\,X'$$

que vuelve a ser la expresión de un modelo de regresión lineal simple.

A continuación, se dan algunos modelos de regresión no lineales y las transformaciones que habría que hacer para aplicar las técnicas de regresión lineal para estimarlos:

- Regresión hiperbólica:

$$Y = a + \frac{b}{X} \Leftrightarrow Y = a + bX'$$
$$\text{con} \quad X' = \frac{1}{X}$$

- Regresión logarítmica:

$$Y = a + b \ln X \Leftrightarrow Y = a + bX'$$
$$\text{con} \quad X' = \ln X$$

- Regresión exponencial:

$$Y = ae^{bX} \Leftrightarrow \ln Y = \ln a + bX \Leftrightarrow Y' = a' + bX$$
$$\text{con} \quad Y' = \ln Y$$
$$a' = \ln a$$

- Regresión potencial:

$$Y = aX^b \Leftrightarrow ln\, Y = ln\, a + b\, ln\, X \Leftrightarrow Y' = a' + bX'$$
$$con \quad X' = ln\, X\,; \quad Y' = ln\, Y$$
$$a' = ln\, a$$

## 2. INTERPRETACIÓN DE LOS RESULTADOS DE UNA REGRESIÓN

De poco valen los resultados obtenidos en una regresión si luego no se saben interpretar en el contexto en el que el modelo se ha planteado. Dicha interpretación debe tener en cuenta, por un lado, la ecuación del modelo y, por el otro, las variables utilizadas, X e Y, en el estudio. Lo habitual es que sea un modelo que busque analizar el impacto ejercido por la variable independiente X sobre la variable independiente Y así como su poder explicativo.

El impacto se suele estudiar analizando el valor y el signo del coeficiente de regresión b así como la forma de la curva estimada. El parámetro a es una estimación del valor de Y para un valor concreto de X (usualmente 0 o 1) y su sentido económico depende de si el valor predicho corresponde a una interpolación o a una extrapolación. En este último caso su validez suele ser más dudosa cuanto mayor sea el grado de extrapolación asociado a la predicción hecha. El poder explicativo se analiza, en términos relativos, con el coeficiente de determinación y, en términos absolutos, utilizando la desviación típica del residual.

En la siguiente tabla (Tabla 1) damos algunas pautas sobre cómo interpretar algunos de los coeficientes de los modelos anteriores:

**Tabla 1.** Interpretación de los coeficientes de regresión

| Nombre | Ecuación | Parámetro a | Parámetro b |
|---|---|---|---|
| **Lineal** | $Y = a+bX$ | Valor de Y predicho para X=0<br><br>Tiene sentido si es una interpolación | Incremento esperado sobre Y cuando X se incrementa en 1 |
| **Exponencial** | $Y = ab^X$ | Valor de Y predicho para X=0<br><br>Tiene sentido si es una interpolación | Incremento esperado sobre Y en términos relativos cuando X se incrementa en 1. El modelo suele estar asociado a fenómenos en los que se dan leyes de rendimientos marginales crecientes |
| **Logarítmico** | $Y = a+b \ln X$ | Valor de Y predicho para X=1<br><br>Tiene sentido si es una interpolación | Incremento esperado sobre Y cuando X se incrementa en 1 unidad porcentual. El modelo suele estar asociado a fenómenos en los que se dan leyes de rendimientos marginales decrecientes |
| **Potencial** | $Y = aX^b$ | Valor de Y predicho para X=1<br><br>Tiene sentido si es una interpolación | Elasticidad de Y respecto a X |
| **Hiperbólico** | $Y = a + \frac{b}{X}$ | Valor asintótico de Y cuando $X \to \infty$ | El modelo suele estar asociado a fenómenos en los que se dan leyes de rendimientos marginales decrecientes con valores asintóticos y determina el ritmo al que se alcanza dicho límite |

Respecto al poder explicativo del modelo suelen darse como buenos, en términos relativos, modelos con coeficientes de determinación superiores al 75% y con ajustes medios aquéllos con coeficientes superiores al 50%. Sin embargo, sólo el coeficiente de determinación no vale. Hay que analizar la desviación típica residual del modelo cuyo valor suele determinar la validez práctica en el contexto del problema. Dicha validez requiere, por parte del analista, un conocimiento específico del mismo así como de los objetivos del estudio y no se pueden, por tanto, pautas generales.