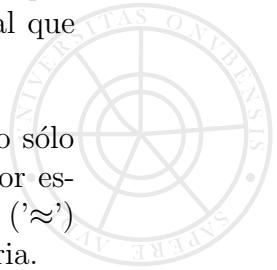
Tema 6

Modelos de regresión

6.1 Introducción

El término **Análisis de Regresión** describe una colección de técnicas estadísticas que sirven como base para realizar inferencias sobre la relación existente entre dos o más variables en estudio. Por ejemplo, podríamos pensar en estudiar la relación que existe entre la lluvia caída por metro cuadrado, (x), y la cosecha obtenida, (y). Está claro que la relación entre estas variables no es exacta (en el sentido de que no existe una función f tal que $y = f(x)$), aunque no se puede negar que existe cierta relación entre ambas variables por lo que se podría estudiar la posibilidad de encontrar una función matemática f tal que $y \approx f(x)$.



En lo que sigue estudiaremos el modelo más sencillo posible que resulta cuando sólo hay una variable independiente x y la función f es lineal. Con la notación anterior escribiríamos la relación entre x e y como $y \approx \beta_0 + \beta_1 x$. La condición de similaridad ('$\approx$') se traduce estadísticamente en la adición de un error ε en forma de variable aleatoria.

Así, el **modelo de regresión lineal simple** es $y = \beta_0 + \beta_1 x + \varepsilon$, donde y es la variable respuesta medida, β_0 y β_1 son los parámetros de la regresión, y ε es el error del modelo.

En la práctica dispondremos de un conjunto de n pares de observaciones experimentales, $(x_1, y_1), (x_2, y_2), \ldots (x_n, y_n)$, y el objetivo será encontrar estimaciones de β_0 y β_1 a partir de estos pares. De este modo, para cada (x_i, y_i), tenemos que

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad i = 1, 2, \ldots, n$$

Durante todo el tema supondremos que los x_i son valores no aleatorios mientras que los ε_i son variables aleatorias con media 0, varianza constante σ^2 e incorreladas entre sí.

La **interpretación del modelo** [1] es: para un valor fijo de x, la variable aleatoria y tiene una media igual a $E[y] = \beta_0 + \beta_1 x$ y una varianza igual a σ^2 independiente de x.

El **objetivo**, como ya se ha indicado, es estimar los coeficientes regresión, β_0 y β_1, labor que comenzamos a tratar en la siguiente sección.

[1] Ocasionalmente utilizaremos la notación $E[y|_x] = \beta_0 + \beta_1 x$, interpretando a $E[y|_x]$ como el valor esperado de y para un valor específico de x.

6.2 Formulación de mínimos cuadrados

Llamaremos b_0 y b_1 a los **estimadores** de β_0 y β_1 respectivamente, y denotaremos por $\widehat{y_i} = b_0 + b_1 x_i$ al **valor estimado** de la respuesta cuando $x = x_i$, esto es, el valor de y que teóricamente correspondería a $x = x_i$.

Se trata entonces de encontrar b_0 y b_1 de modo que cada valor real de la variable respuesta, y_i, y el correspondiente valor estimado de la misma, $\widehat{y_i}$, difieran lo menos posible. Para ello usaremos el método de mínimos cuadrados, que consiste en determinar los valores de b_0 y b_1 que hacen mínima la **suma de cuadrados residuales**, esto es, la suma de los cuadrados de las diferencias entre cada valor y_i y el correspondiente valor estimado $\widehat{y_i}$, es decir:

$$\min_{b_0,b_1} \sum_{i=1}^{n} (y_i - \widehat{y_i})^2$$

o lo que es igual

$$\min_{b_0,b_1} \sum_{i=1}^{n} (y_i - b_0 - b_1 x_i)^2$$

Para ello, b_0 y b_1 deben ser solución del sistema de ecuaciones lineales:

$$\begin{cases} \frac{\partial}{\partial b_0}\Big[\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)^2\Big] = 0 \\ \\ \frac{\partial}{\partial b_1}\Big[\sum_{i=1}^{n}(y_i - b_0 - b_1 x_i)^2\Big] = 0 \end{cases}$$

Resolviendo el sistema, se llega a que la solución viene dada por:

$$\boxed{b_1 = \frac{S_{xy}}{S_x^2}} \qquad \boxed{b_0 = \overline{y} - \frac{S_{xy}}{S_x^2} \cdot \overline{x}}$$

donde S_{xy} es la *covarianza* entre x e y, cuya expresión es: $S_{xy} = \frac{\sum_{i=1}^{n} x_i \cdot y_i}{n} - \overline{x} \cdot \overline{y}$

Por lo tanto la recta estimada es:

$$\boxed{y - \overline{y} = \frac{S_{xy}}{S_x^2}(x - \overline{x})}$$

Además, bajo la hipótesis establecida de que los x_i son no aleatorios, los errores ε_i son incorrelados y $E[\varepsilon_i] = 0$ y $Var[\varepsilon_i] = \sigma^2$, se verifica que $E[b_0] = \beta_0$ y $E[b_1] = \beta_1$, esto es, b_0 y b_1 son estimadores insesgados de β_0 y β_1.

6.3 Estimación de la varianza del error

El estimador de máxima verosimilitud de σ^2 viene dado por:

$$\boxed{\widehat{\sigma^2} = \sum_{i=1}^{n} \frac{(y_i - \widehat{y_i})^2}{n}}$$

Por otra parte, se puede comprobar que el estimador anterior es un estimador sesgado de σ^2 y que su esperanza es $E(\widehat{\sigma^2}) = \frac{n-2}{n}\sigma^2$. Se deduce entonces que un estimador insesgado de σ^2 viene dado por

$$\boxed{S^2 = \frac{\sum_{i=1}^{n}(y_i - \widehat{y}_i)^2}{n-2}}$$

.

A S^2 se le denomina **cuadrados medios del error** y $n-2$ son los **grados de libertad del error o del residuo**[2].

6.4 Descomposición de la variabilidad total

Supongamos dado un conjunto de datos experimentales y desarrollado sobre ellos el modelo de regresión lineal simple tal como hemos descrito en este tema. Se puede interpretar el modelo ajustado como una explicación de la variación observada en la variable y (respuesta), que se medirá con respecto a $\overline{y}$. Por supuesto, es importante que los valores ajustados $\widehat{y}_i$ estén próximos a los reales y_i. Si esto ocurre, la variación de los $\widehat{y}_i$ alrededor de $\overline{y}$ estará próxima a la variación de los y_i en torno a $\overline{y}$ (puede comprobarse que $\overline{\widehat{y}} = \overline{y}$).

En consecuencia, se consideran dos fuentes de variación:

1. La suma total de cuadrados: $\mathrm{SS}_{\mathrm{Tot}} = \sum_{i=1}^{n}(y_i - \overline{y})^2$.

2. La suma de cuadrados debida a la regresión: $\mathrm{SS}_{\mathrm{Reg}} = \sum_{i=1}^{n}(\widehat{y}_i - \overline{y})^2$.

Se puede comprobar que:

$$\sum_{i=1}^{n}(y_i - \overline{y})^2 = \sum_{i=1}^{n}(\widehat{y}_i - \overline{y})^2 + \sum_{i=1}^{n}(y_i - \widehat{y}_i)^2$$

esto es, $SS_{\mathrm{Tot}} = SS_{\mathrm{Reg}} + SS_{\mathrm{Res}}$. Esta ecuación representa la siguiente relación entre las sumas de cuadrados:

$$\begin{pmatrix} \text{variabilidad total} \\ \text{de la respuesta} \end{pmatrix} = \begin{pmatrix} \text{variabilidad explicada} \\ \text{por el modelo} \end{pmatrix} + \begin{pmatrix} \text{variabilidad} \\ \text{no explicada} \end{pmatrix}$$

La situación deseable es que $\mathrm{SS}_{\mathrm{Reg}}$ sea grande en comparación con $\mathrm{SS}_{\mathrm{Res}}$, pues $\mathrm{SS}_{\mathrm{Reg}}$ se puede interpretar como la variación de y producida por los cambios en x, mientras que $\mathrm{SS}_{\mathrm{Res}}$ es la variación debida a los errores ε_i del modelo.

[2]En este tema, S^2 denota los cuadrados medios del error y no varianza de una variable aleatoria o un conjunto de datos. La varianza de los datos correspondientes a la variable x o la variable y se denotará por S_x^2 y S_y^2, respectivamente.

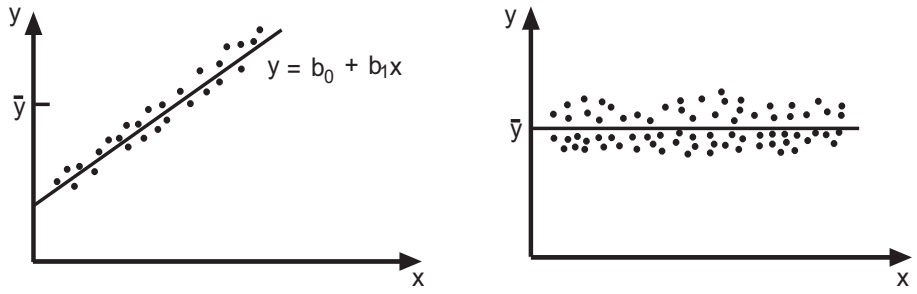


Figure 6.1: Dos ajustes de regresion

La descomposición de $\mathrm{SS}_{\mathrm{Tot}}$ es una herramienta muy sencilla para determinar si la variación explicada por el modelo de regresión es **real** o **casual**.

Si es **real**, significa que hay una tendencia lineal clara en la relación entre x e y, esto es, $\beta_1 \neq 0$. En tal caso $\mathrm{SS}_{\mathrm{Reg}}$ será una parte apreciable de la $\mathrm{SS}_{\mathrm{Tot}}$.

La figura 6.4 muestra dos conjuntos de datos ficticios. En el primero observamos que existe una pendiente no nula (concretamente, positiva) en la regresión y como consecuencia $E[y|_x]$ aumenta cuando aumenta x. En este caso se aprecia que, efectivamente, existe relación lineal entre las variables.

En el segundo conjunto de datos la recta de ajuste es horizontal y no se aprecia dependencia lineal entre las variables ya que el comportamiento de la variable dependiente (y) no se ve afectado por el valor que toma la variable independiente (x). En este caso $\mathrm{SS}_{\mathrm{Reg}} = 0$ y la variación completa en y se debe, únicamente, a la variación del error alrededor de la linea ajustada $\widehat{y} = \overline{y}$.

Si bien en estos dos ejemplos es evidente el carácter significativo o no de la regresión, en general será necesario el uso de un contraste de hipótesis para determinar si la regresión es significativa.

6.5 Contrastes de hipótesis e intervalos de confianza

A partir de la recta de regresión podemos obtener información sobre las siguientes cuestiones:

1. La variable x, ¿realmente influye de manera lineal en la variable respuesta y?
2. ¿Existe un ajuste adecuado de los datos y el modelo?

La primera cuestión se puede responder mediante un test de hipótesis sobre la pendiente de la recta, β_1. Para ello se resuelve el contraste

$$\begin{cases} H_0: & \beta_1 = 0 \\ H_1: & \beta_1 \neq 0 \end{cases}$$

Rechazar H_0 en favor de H_1 lleva a la conclusión de que x influye significativamente en la respuesta de manera lineal. Sin embargo, es necesario tener en cuenta que el rechazo de H_0 simplemente indica que se detectó una tendencia, no implica nada sobre la calidad

del ajuste ni sobre la capacidad del modelo para hacer predicciones sobre la variable dependiente. La resolución de este tipo de contrastes se detallará en la Sección 6.5.1.

La respuesta a la segunda cuestión la obtendremos a partir del *coeficiente de determinación* y el *coeficiente de determinación corregido*. En la Sección 6.6 veremos cómo calcular e interpretar estos coeficientes.

6.5.1 El T-test y la estimación mediante intervalos de confianza

Se pueden realizar tests de hipótesis separadamente sobre β_0 y β_1. Para ello es necesaria la hipótesis de normalidad sobre los errores. Con esta hipótesis, se puede demostrar que

$$\frac{b_1 - \beta_1}{S}\sqrt{nS_x^2} \sim t_{n-2}$$

Si estamos interesados en realizar el contraste:

$$\begin{cases} H_0: & \beta_1 = \beta_{1,0} \\ H_1: & \beta_1 \neq \beta_{1,0} \end{cases}$$

donde $\beta_{1,0}$ es una constante determinada, un estadístico adecuado es

$$t = \frac{b_1 - \beta_{1,0}}{S}\sqrt{n}S_x$$

Fijado un nivel de significación α para el contraste, se rechará H_0 si:

$$\boxed{\begin{array}{c} b_1 \geq \beta_{1,0} + \frac{S}{S_x\sqrt{n}} t_{n-2,1-\alpha/2} \\ \\ \text{o bien} \\ \\ b_1 \leq \beta_{1,0} - \frac{S}{S_x\sqrt{n}} t_{n-2,1-\alpha/2} \end{array}}$$

En lo relativo a β_0, para resolver el contraste

$$\begin{cases} H_0: & \beta_0 = \beta_{0,0} \\ H_1: & \beta_0 \neq \beta_{0,0} \end{cases}$$

un estadístico apropiado será

$$t = \frac{b_0 - \beta_{0,0}}{S\sqrt{\frac{1}{n} + \frac{\overline{x}^2}{nS_x^2}}}$$

que, bajo H_0, sigue una distribución t_{n-2}.

Como consecuencia, rechazaremos la hipótesis nula cuando

$$\boxed{\begin{array}{c} b_0 \geq \beta_{0,0} + S\sqrt{\frac{1}{n} + \frac{\overline{x}^2}{nS_x^2}} \cdot t_{n-2,1-\alpha/2} \\ \\ \text{o bien} \\ \\ b_0 \leq \beta_{0,0} - S\sqrt{\frac{1}{n} + \frac{\overline{x}^2}{nS_x^2}} \cdot t_{n-2,1-\alpha/2} \end{array}}$$

Si se desea calcular un intervalo de confianza bien para la pendiente, β_1, o para la intersección, β_0, del ajuste,

Un intervalo de confianza para β_1 con nivel de confianza $100(1-\alpha)\%$ viene dado por:

$$\boxed{\left(b_1 - \sqrt{\frac{S^2}{nS_x^2}} \cdot t_{n-2,1-\alpha/2}, b_1 + \sqrt{\frac{S^2}{nS_x^2}} \cdot t_{n-2,1-\alpha/2}\right)}$$

y un intervalo de confianza para β_0 al $(1-\alpha)100\%$ será:

$$\boxed{\left(b_0 - S\sqrt{\frac{1}{n} + \frac{\overline{x}^2}{nS_x^2}} \cdot t_{n-2,1-\alpha/2}, b_0 + S\sqrt{\frac{1}{n} + \frac{\overline{x}^2}{nS_x^2}} \cdot t_{n-2,1-\alpha/2}\right)}$$

Por otra parte, dado un valor x_0 de la variable x, $\hat{y}(x_0) = b_0 + b_1 x_0$ representa la respuesta estimada y se puede considerar como un estimador del valor medio de y cuando $x = x_0$, esto es: $E(y|_{x_0}) = \beta_0 + \beta_1 x_0$. Bajo la condición de errores "normales", $\hat{y}(x_0)$ sigue una distribución normal, y un intervalo de confianza para $E(y|_{x_0})$ viene dado por

$$\boxed{\left(\hat{y}(x_0) - S \cdot \sqrt{\frac{1}{n} + \frac{(x_0-\overline{x})^2}{nS_x^2}} \cdot t_{n-2,1-\alpha/2}, \hat{y}(x_0) + S \cdot \sqrt{\frac{1}{n} + \frac{(x_0-\overline{x})^2}{nS_x^2}} \cdot t_{n-2,1-\alpha/2}\right)}$$

6.6 Calidad del modelo ajustado

Responderemos ahora a la siguiente cuestión: ¿se ajustan los datos al modelo de forma adecuada? Para ello usaremos el coeficiente de determinación, que se define como:

$$R^2 = 1 - \frac{SS_{\text{Res}}}{SS_{\text{Tot}}}.$$

Alternativamente, en el caso de la regresión lineal, debido a la relación que existe entre la suma de cuadrados total y la suma de cuadrados debidos a la regresión y los residuos, puede escribirse como:

$$R^2 = \frac{SS_{\text{Reg}}}{SS_{\text{Tot}}} = \frac{\sum_{i=1}^{n}(\widehat{y}_i - \overline{y})^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}$$

El coeficiente R^2 representa la proporción de la variación en la respuesta que es explicada por el modelo y verifica que $0 \leq R^2 \leq 1$. Cuando el ajuste entre las variables es perfecto se verifica que $R^2 = 1$. Cuando no existe relación lineal entre las variables, $R^2 = 0$.

En el caso de la regresión lineal simple se verifica además que

$$R^2 = \frac{S_{xy}^2}{S_x^2 \cdot S_y^2}$$

es decir, $R^2 = r^2$, donde r es el coeficiente de correlación lineal de Pearson.

¿Qué valor es aceptable para R^2? Depende del tipo de estudio y del campo científico en el que se trabaje. Un químico interesado en calibrar una composición de alta precisión necesitará un valor de R^2 muy alto, quizá superior a 0.99, mientras que un científico que estudia el comportamiento humano puede considerarse afortunado si observa un valor de R^2 ligeramente superior a 0.7. Además, determinados fenómenos pueden, por su propia naturaleza, modelarse con más precisión que otros.

Aunque el coeficiente de determinación puede interpretarse fácilmente, existen algunos peligros en su uso. Por ejemplo, es un criterio peligroso para la comparación de modelos candidatos, porque cualquier término adicional del modelo (como un término cuadrático) hará decrecer SS_{Res} (o al menos no lo aumentará) y por tanto hará crecer R^2 (o al menos no disminuirá). Así, R^2 puede ser artificialmente alto a causa de un "sobreajuste", es decir, de la inclusión de demasiados términos en el modelo. Un incremento en R^2 no implica que el elemento adicional sea necesario. Por ello, para evitar este problema, en la comparación de modelos candidatos se utiliza el **coeficiente de determinación corregido** definido como

$$\overline{R}^2 = 1 - \frac{SS_{\text{Res}}/(n-p)}{SS_{\text{Tot}}/(n-1)}$$

donde p es el número de parámetros del modelo.

6.7 El modelo de regresión lineal múltiple

El modelo desarrollado anteriormente se puede generalizar al caso en que sea necesario el uso de más de una variable regresora.

Consideremos un experimento en el que los datos generados son del tipo

y	x_1	x_2	$\ldots$	x_k
y_1	x_{11}	x_{21}	$\ldots$	x_{k1}
y_2	x_{12}	x_{22}	$\ldots$	x_{k2}
$\vdots$	$\vdots$	$\vdots$		$\vdots$
y_n	x_{1n}	x_{2n}	$\ldots$	x_{kn}

Donde la variable y es la "respuesta" observada y $x_1, \ldots, x_k$ representan magnitudes que se pueden medir de forma exacta o con un error despreciable. En la fila i-ésima aparecen representados los datos observados en el experimento i-ésimo para la variable respuesta y para la variable regresora.

Si suponemos que cada y_i está relacionada de forma lineal con las variables $x_1, x_2, \ldots, x_k$, obtenemos el modelo:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_k x_{ki} + \varepsilon_i \quad i = 1, 2, \ldots, n; \qquad n \geq k+1$$

donde ε_i es un error incorrelado de observación en observación, con media 0 y varianza σ^2 y los x_{ij} son no aleatorios.

Un modelo es lineal cuando lo es en los parámetros β_i. Por ejemplo, son lineales los modelos

$$
\begin{aligned}
y &= \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon \\
y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon \\
y &= \beta_0 + \beta_1 lnx_1 + \beta_2 lnx_2 + \beta_3 lnx_3 + \varepsilon \\
lny &= \beta_0 + \beta_1 (\tfrac{1}{x_1}) + \beta_2 (\tfrac{1}{x_2}) + \varepsilon
\end{aligned}
$$

y no son lineales

$$
\begin{aligned}
y &= \beta_0 + \beta_1 {x_1}^{\beta_2} + \beta_3 x_2^{\beta_4} + \varepsilon \\
y &= \tfrac{\beta_0}{1+e^{-(\beta_1 x_1 + \beta_2 x_2)}} + \varepsilon
\end{aligned}
$$

6.7.1 Modelo lineal general

El modelo descrito anteriormente se puede escribir de la forma

$$Y = X\beta + \varepsilon$$

donde

$$
Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad
X = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix}, \quad
\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad
\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}
$$

ε e Y son vectores aleatorios, X es una matriz $n \times p$ de datos, con $p = k + 1$, que es el número total de parámetros del modelo, y recibe el nombre de matriz de diseño.

6.7.2 Método de mínimos cuadrados

Se trata de buscar el estimador b del vector β que satisfaga:

$$\frac{\partial}{\partial b}[(Y - Xb)'(Y - Xb)] = 0$$

Si X es una matriz de rango total, $b = (X'X)^{-1}X'Y$. Un estimador insesgado de σ^2 viene dado por

$$S^2 = \frac{(Y - Xb)'(Y - Xb)}{n - p} = \sum_{i=1}^{n} \frac{(y_i - \widehat{y_i})^2}{n - p}$$

donde p es el número de parámetros que deben ser estimados e $\widehat{y_i}$ es la respuesta estimada en el i-ésimo dato.

Bajo la condición de que $E[\varepsilon] = \Theta$ (vector nulo), se puede comprobar que b es un estimador insesgado de β. Si además suponemos que los errores ε_i son incorrelados y $Var[\varepsilon_i] = \sigma^2$, entonces $Var[b] = \sigma^2(X'X)^{-1}$.

Al igual que ocurría en el caso de la regresión lineal simple, se tiene la siguiente relación entre SS_{Tot}, SS_{Reg} y SS_{Res}:

$$\sum_{i=1}^{n}(y_i-\overline{y})^2=\sum_{i=1}^{n}(\widehat{y}_i-\overline{y})^2+\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2$$

siendo la descomposición de los grados de libertad (bajo la hipótesis de normalidad de los errores) como sigue:

$$n-1=(p-1)+(n-p)$$

El análisis de la varianza se utiliza para resolver el contraste:

$$\begin{cases} H_0: & \beta_1=\beta_2=\ldots=\beta_k=0 \\ H_1: & \beta_1\neq 0 \ \text{ó} \ \beta_2\neq 0 \ \ldots \ \text{ó} \ \beta_k\neq 0 \end{cases}$$

La tabla para el cálculo del estadístico F del contraste queda así:

Fuente	Suma de cuadrados	Grados de libertad	Cuadrados medios	F
Regresión	$\mathrm{SS}_{\mathrm{Reg}}$	$p-1$	$\frac{\mathrm{SS}_{\mathrm{Reg}}}{p-1}=\mathrm{MS}_{\mathrm{Reg}}$	$F=\frac{\mathrm{MS}_{\mathrm{Reg}}}{S^2}$
Residuos	$\mathrm{SS}_{\mathrm{Res}}$	$n-p$	$\frac{\mathrm{SS}_{\mathrm{Res}}}{n-p}=S^2$	
Total	SS_{Tot}	$n-1$		

Dado un nivel de significación α, rechazaremos H_0 cuando $F\geq f_{p-1,n-p,1-\alpha}$. Rechazar H_0 implica que las variables regresoras influyen en la respuesta de manera lineal.

El coeficiente de determinación se define e interpreta del mismo modo que en el caso de la regresión lineal simple, y nuevamente puede calcularse como:

$$R^2=\frac{SS_{\mathrm{Reg}}}{SS_{\mathrm{Tot}}}=\frac{\sum_{i=1}^{n}(\widehat{y}_i-\overline{y})^2}{\sum_{i=1}^{n}(y_i-\overline{y})^2}$$

Recordamos que en la comparación de modelos se utiliza el coeficiente de determinación corregido.

Por otra parte, y aunque no entraremos en ese terreno, al igual que ocurría en el caso de la regresión lineal simple, es posible calcular intervalos de confianza así como realizar contrastes de hipótesis acerca de los parámetros que intervienen en el modelo (siempre bajo la hipótesis de normalidad de los errores).

6.8 Regresión no lineal

Hasta ahora habíamos considerado modelos con estructura lineal en los parámetros. En numerosas ocasiones aparecen situaciones experimentales que requieren el uso de modelos no lineales.

Aplicar el método de mínimos cuadrados directamente sobre estos modelos puede conducir a ecuaciones complicadas que no se pueden resolver con herramientas algebraicas y requieren el uso de métodos numéricos.

Por este motivo, no vamos a profundizar en la resolución de estos problemas y veremos, únicamente, cómo un cambio de variable adecuado puede ayudar a obtener una solución aproximada cuando se trata de ajustar determinados modelos. Esta solución no deja de ser una aproximación pues con el cambio de variable es posible que no respetemos ciertas hipótesis como los errores aditivos o la hipótesis de normalidad.