

# Regressión lineal múltiple

Se tiene un sistema de ecuaciones

$y = xb$  o  $xb=y$ ,  $x$  una matriz de coeficientes  
y se toman n muestras

Número de Muestra i	Muestra
1	$(x_{11}, x_{12}, x_{13}, \dots, x_{1k}, y_1)$
2	$(x_{21}, x_{22}, x_{23}, \dots, x_{2k}, y_2)$
.	.
n	$(x_{n1}, x_{n2}, x_{n3}, \dots, x_{nk}, y_n)$

Modelo de RLM:

$$\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k = y$$

Objetivo: estimar los diferentes parámetros  $\beta_i$ , con  $i=0, \dots, k$ , a partir de  $b_i$ , con  $i=0, \dots, k$ .

$$b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_k x_k = \hat{y}$$

Evaluamos cada muestra en el modelo:

Sistema de ecuaciones lineales con "n" ecuaciones y "k" incógnitas

$$\begin{cases} b_0 + x_{11}b_1 + x_{12}b_2 + \dots + x_{1k}b_k = y_1 \\ b_0 + x_{21}b_1 + x_{22}b_2 + \dots + x_{2k}b_k = y_2 \\ \vdots \\ b_0 + x_{n1}b_1 + x_{n2}b_2 + \dots + x_{nk}b_k = y_n \end{cases}$$

Tamaño:  $n \times k$

Forma aumentada del sistema:

$$\left( \begin{array}{cccc} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{array} \right) \left( \begin{array}{c} b_0 \\ b_1 \\ \vdots \\ b_k \end{array} \right) = \left( \begin{array}{c} y_1 \\ y_2 \\ \vdots \\ y_n \end{array} \right)$$

Matriz asociada al sistema

$$\Downarrow \\ n \times (k+1)$$

Vector de incógnitas

$$\Downarrow \\ (k+1) \times 1$$

Vector solución

$$\Downarrow \\ n \times 1$$

$$X \quad b = Y$$

Sistema  $Xb = Y$

## Modelos de regresión lineal múltiple

Se calculan las derivadas parciales de  $SS_E$  con respecto a cada  $b_i$ , se igualan a cero y se resuelve el sistema. El siguiente teorema muestra una forma de calcular estos estimadores:

**TEOREMA:** los estimadores por el método de mínimos cuadrados de los parámetros  $\beta_0, \beta_1, \beta_2, \beta_3, \dots, \beta_k$ , a partir de la muestra  $(x_i, y_i), i = 1, 2, \dots, n$ , son los siguientes:

$$b = (X^t \cdot X)^{-1} \cdot (X^t \cdot y)$$

Las operaciones se pueden resolver en la calculadora

Ejemplo 100 Se tienen las siguientes observaciones

$x_1$	0	1	1	0	-1
$x_2$	1	0	-1	0	1
$Y$	1	4	6	3	0

1. Estime la ecuación de regresión de  $Y$  como función lineal de  $x_1, x_2$ .

$X_1$	$X_2$	$Y$
0	1	1
1	0	4
1	-1	6
0	0	3
-1	1	0

Modelo

$$Y = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 \text{ a partir de}$$

$$\hat{Y} = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2$$

Se debe agregar una columna de 1's al principio y luego los demás e igualan a  $\hat{Y}$

$$\hat{Y} = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2$$

$$X = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix} \quad Y = \begin{pmatrix} 1 \\ 4 \\ 6 \\ 3 \\ 0 \end{pmatrix} \quad b = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}$$

$$b = (X^T \cdot X)^{-1} \cdot (X^T \cdot Y) = \begin{pmatrix} 3 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}$$

R/ El modelo de RLM es  $\hat{Y} = 3 + 1 \cdot X_1 - 2 \cdot X_2$

La calcu solo permite 8x9 max.

# Regresion NO lineal multiple

Ejemplo 1: Se tienen las siguientes observaciones en la tabla adjunta:

- Estime los coeficientes en la ecuación

$$z = \beta_0 + \beta_1 x + \beta_2 y + \beta_3 y^2$$

- Determine  $z$  cuando  $x = 2$  y  $y = 7$

x	y	z
3	5	-46
8	9	-110
13	4	10
16	6	-12
20	2	64

Se deben estimar los coeficientes de :

$$z = \beta_0 + \beta_1 x + \beta_2 y + \beta_3 y^2$$

Sea  $x_1 = x$ ,  $x_2 = y$ ,  $x_3 = y^2$

Reemplazamos en  $z$  usando  $b$  y las  $x_i$

$$z = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3$$

(construyendo nueva tabla)

x	y	$y^2$	$z$
$x_1$	$x_2$	$x_3$	
3	5	$5^2$	-46
8	9	$9^2$	-110
13	4	$4^2$	10
16	6	$6^2$	-12
20	2	$2^2$	64

$$X = \begin{pmatrix} 1 & 3 & 5 & 5^2 \\ 1 & 8 & 9 & 9^2 \\ 1 & 13 & 4 & 4^2 \\ 1 & 16 & 6 & 6^2 \\ 1 & 20 & 2 & 2^2 \end{pmatrix}$$

Recordar los  $z$

$$\begin{aligned} b &= (X^T \cdot X)^{-1} \cdot (X^T \cdot y) \\ &= \begin{pmatrix} 2 \\ 9 \\ -7 \\ -1 \end{pmatrix} \end{aligned}$$

$$y = \begin{pmatrix} -46 \\ -110 \\ 10 \\ -12 \\ 64 \end{pmatrix} \quad b = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

El modelo de RNL es  $z = 2 + 9x - 7y - y^2$

Determine  $z$  cuando  $x = 2$  y  $y = 7$

$$z = 2 + 9x - 7y - y^2$$

$$z = 2 + 9 \cdot 2 - 7 \cdot 7 - 7^2$$

$$\boxed{z = -88}$$

Ejemplo 2: Dadas las siguientes observaciones  
estime los coeficientes de la ecuación

$$z = \beta_0 x^{\beta_1} y^{\beta_2}$$

- Determine  $z$  cuando  $x = 2$  y  $y = 7$

x	y	z
1	0,5	0,06
1,5	1	6,7
2	2	512
1	2	64

$$z = \beta_0 x^{\beta_1} - y^{\beta_2}$$

$$\ln(z) = \ln(\beta_0 x^{\beta_1} - y^{\beta_2})$$

$$\ln(z) = \ln(\beta_0 \cdot x^{\beta_1}) + \ln(y)^{\beta_2}$$

$$\ln(z) = \ln(\beta_0) + \underbrace{\beta_1 \ln(x)}_{b_1} + \underbrace{\beta_2 \ln(y)}_{b_2}$$

Modelo de estimación para  $z = \beta_0 \cdot x^{\beta_1} y^{\beta_2}$

$$z_j = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2$$

## Construyendo nueva tabla

$$\ln(z) = \ln(\beta_0) + \beta_1 \ln(x) + \beta_2 \ln(y)$$

$\downarrow$        $\downarrow$        $\downarrow$        $\downarrow$        $\downarrow$        $\downarrow$   
 $z_1$        $b_0$        $b_1$        $x_1$        $b_2$        $x_2$

X	Y	Z
1	0,5	0,06
1,5	1	6,7
2	2	512
1	2	64

$$\begin{array}{lll}
 x_1 = \ln(x) & x_2 = \ln(y) & z_1 = \ln(z) \\
 \ln(1) & \ln(0,5) & \ln(0,06) \\
 \ln(1,5) & \ln(1) & \ln(6,7) \\
 \ln(2) & \ln(2) & \ln(512) \\
 \ln(1) & \ln(2) & \ln(64)
 \end{array}$$

$$X = \begin{pmatrix} 1 & \ln(1) & \ln(0,5) \\ 1 & \ln(1,5) & \ln(1) \\ 1 & \ln(2) & \ln(2) \\ 1 & \ln(1) & \ln(2) \end{pmatrix} \quad Y = \begin{pmatrix} \ln(0,06) \\ \ln(6,7) \\ \ln(512) \\ \ln(64) \end{pmatrix}$$

$$b = (X^T \cdot X)^{-1} \cdot (X^T \cdot Y)$$

$$b = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}$$

$$= \begin{pmatrix} 0,6748 \\ 3,0078 \\ 5,0260 \end{pmatrix} \begin{matrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{matrix}$$

$$\ln(z) = 0,6748 + 3,0078 \ln(x) + 5,0260 \ln(y)$$

Volviendo al modelo original

$$\ln(z) = \ln(\beta_0) + \beta_1 \ln(x) + \beta_2 \ln(y)$$

$\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow$   
 $z_1 \quad b_0 \quad b_1 \quad x_1 \quad b_2 \quad x_2$

$$\ln(z) = 0,6748 + 3,0078 \ln(x) + 5,0260 \ln(y)$$

$\downarrow \quad \downarrow \quad \downarrow \quad \downarrow \quad \downarrow$   
 $z_1 \quad b_0 \quad b_1 \quad x_1 \quad b_2 \quad x_2$

$$z = \beta_0 x^{\beta_1} \cdot y^{\beta_2}$$

El  $\beta_0$  es el resultado de  $\ln(\beta_0)$ , entonces  
hay que volver hacia atrás

$$\ln(\beta) = 0,6748$$
$$e^\beta = 0,6748$$
$$e \approx 1,9628$$

Finalmente

$$z = \beta_0 x^{\beta_1} \cdot y^{\beta_2}$$

$$z = 1,9628 \cdot x^{3,0078} \cdot y^{5,0260}$$

Determine  $z$  cuando  $x = 2$  y  $y = 7$

$$z = 1,9628 \cdot x^{3,0078} \cdot y^{5,0260}$$

$$z = 1,9628 \cdot 2^{3,0078} \cdot 7^{5,0260} = 279130,9293$$