Ejercicios de Estadística

Temas: Regresión no lineal
Titulaciones: Química

Alfredo Sánchez Alberca (asalber@ceu.es)

CEU Universidad San Pablo

En un estudio se ha medido el calor liberado en una reacción química en distintos instantes desde el comienzo de la reacción, obteniendo los siguientes datos:

Tiempo en minutos	2,5	3,7	4,1	5,3	6,2
Calor en calorías	15,9	44,5	65,6	206,5	498,7

Se pide:

1. Calcular el modelo de regresión lineal del calor sobre el tiempo. Según este modelo ¿Cuándo cambiarán las calorías por cada minuto que pase?
2. Calcular el modelo de regresión exponencial del calor sobre el tiempo.
3. Utilizando el mejor de los dos modelos anteriores, predecir el calor generado a los 5 minutos de la reacción. ¿Es fiable la predicción? Justificar la respuesta.

1. Calcular el modelo de regresión lineal del calor sobre el tiempo. Según este modelo ¿Cuándo cambiarán las calorías por cada minuto que pase?

Datos
X = Tiempo en minutos
Y = Calor liberado en calorías

X	2,5	3,7	4,1	5,3	6,2
Y	15,9	44,5	65,6	206,5	498,7

$y = 125'8161x - 382'3181$

Recta de regresión y sobre x : $y = \bar{y} + \frac{S_{xy}}{S_x^2}(x - \bar{x}) = 166'24 + \frac{207'1436}{1'6464}(x - 4'36) =$

$\bar{x} = \frac{\sum x_i}{n} = \frac{2'5 + 3'7 + 4'1 + 5'3 + 6'2}{5} = \frac{21'8}{5} = 4'36$ min

$S_x^2 = \frac{\sum x_i^2}{n} - \bar{x}^2 = \frac{2'5^2 + 3'7^2 + 4'1^2 + 5'3^2 + 6'2^2}{5} - 4'36^2 = \frac{103'28}{5} - 4'36^2 = 1'6464$ min^2

$\bar{y} = \frac{\sum y_j}{n} = \frac{15'9 + 44'5 + 65'6 + 206'5 + 498'7}{5} = 166'24$ cal

$S_y^2 = \frac{\sum y_j^2}{n} - \bar{y}^2 = \frac{15'9^2 + 44'5^2 + 65'6^2 + 206'5^2 + 498'7^2}{5} - 166'24^2 = \frac{297867'25}{5} - 166'24^2 =$
$= 31940'3344$ cal^2

$S_{xy} = \frac{\sum x_i y_i}{n} - \bar{x}\bar{y} = \frac{2'5 \cdot 15'9 + \dots + 6'2 \cdot 498'7}{5} - 4'36 \cdot 166'24 =$
$= \frac{4659'34}{5} - 4'36 \cdot 166'24 = 207'1436$ min.cal

2. Calcular el modelo de regresión exponencial del calor sobre el tiempo.

X	2,5	3,7	4,1	5,3	6,2
Y	15,9	44,5	65,6	206,5	498,7
z	2'7663	3'7955	4'1836	5'3303	6'2120

Datos

$X =$ Tiempo en minutos

$Y =$ Calor liberado en calorías

$\bar{x} = 4{,}36$ min

$s_x^2 = 1{,}6464$ min^2

Modelo de regresión exponencial $y = e^{a+bx} = e^{0'9357x+0'378}$

$z = \ln y = \ln(e^{a+bx}) = a+bx$

Recta de regresión de z sobre x: $z = \bar{z} + \frac{S_{xz}}{S_x^2}(x-\bar{x}) = 4'4575 + \frac{1'5405}{1'6464}(x - 4'36) =$

$= 0'9357x + 0'378$

$\bar{z} = \frac{\sum z_i}{n} = \frac{2'7663 + \cdots + 6'2120}{5} = 4'4575$ log cal

$S_z^2 = \frac{\sum z_i^2}{n} - \bar{z}^2 = \frac{2'7663^2 + \cdots + 6'2120^2}{5} - 4'4575^2 = \frac{106'5616}{5} - 4'4575^2 = 1'4427$ log^{2}cal

$S_{xz} = \frac{\sum x_i z_i}{n} - \bar{x}\cdot\bar{z} = \frac{2'5\cdot 2'7663 + \cdots + 6'2\cdot 6'2120}{5} - 4'36\cdot 4'4575 =$

$= \frac{104'8768}{5} - 4'36\cdot 4'4575 = 1'5405$ min.(log(cal))

3. Utilizando el mejor de los dos modelos anteriores, predecir el calor generado a los 5 minutos de la reacción. ¿Es fiable la predicción? Justificar la respuesta.

$$r^2_{lineal} = \frac{s_{xy}^2}{s_x^2 s_y^2} = \frac{207'1436^2}{1'6464 \cdot 31940'3344} = \underline{0'816}$$

$$r^2_{exponencial} = \frac{s_{xz}^2}{s_x^2 \cdot s_z^2} = \frac{1'5405^2}{1'6464 \cdot 1'4427} = \underline{0'9991}$$

$$y = e^{0'9357x + 0'378}$$

$$y(5) = e^{0'9357 \cdot 5 + 0'378} = \underline{157'0193}\ \text{cal}$$

Datos

$X =$ Tiempo en minutos

$Y =$ Calor liberado en calorías

$Z = \log(Y)$

$s_x^2 = 1{,}6464$ min^2

$s_y^2 = 31940{,}3344$ cal^2

$s_z^2 = 1{,}4427 \log^2(\text{cal})$

$s_{xy} = 207{,}1436$ min·cal

$s_{xz} = 1{,}5405$ min· log(cal)

Aunque el ajuste es muy bueno, no podemos decir que la predicción sea fiable porque el tamaño muestral es muy pequeño