

Ejercicios de Estadística

Temas: Regresión lineal y no lineal

Titulaciones: Medicina, Farmacia

Alfredo Sánchez Alberca

asalber@ceu.es

<http://aprendeconalf.es>



CEU

*Universidad
San Pablo*



La siguiente tabla muestra las tasas de incidencia de gripe por cada 100.000 habitantes registradas al cabo de un número de días desde el comienzo de el estudio.

Días	1	5	8	12	20	26	38	44
Tasa gripe	60	66	71	80	106	132	194	235

Se pide:

1. Calcular la tasa de incidencia de gripe a los 50 días desde el comienzo del estudio mediante un modelo de regresión lineal.
2. ¿Cuánto varía la tasa de incidencia de gripe cada día según el modelo lineal?
3. Calcular la tasa de incidencia de gripe a los 50 días desde el comienzo del estudio mediante un modelo de regresión exponencial.
4. ¿Cuál de las predicciones anteriores es más fiable? Razonar la respuesta.

Utilizar las siguientes sumas para los cálculos ($X = \text{Días}$ e $Y = \text{Tasa de gripe}$):

$$\begin{aligned} \sum x_i &= 154, \sum \log(x_i) = 19.8494, \sum y_j = 944, \sum \log(y_j) = 37.2024, \\ \sum x_i^2 &= 4690, \sum \log(x_i)^2 = 60.2309, \sum y_j^2 = 140918, \sum \log(y_j)^2 = 174.8363, \\ \sum x_i y_j &= 25182, \sum \log(x_i) y_j = 2795.2484, \sum x_i \log(y_j) = 772.3504, \\ \sum \log(x_i) \log(y_j) &= 96.1974. \end{aligned}$$

1. Calcular la tasa de incidencia de gripe a los 50 días desde el comienzo del estudio mediante un modelo de regresión lineal.

Datos

$X \equiv$ Días

$Y \equiv$ Tasa de gripe

$$\sum x_i = 154$$

$$\sum \log(x_i) = 19.8494$$

$$\sum y_j = 944$$

$$\sum \log(y_j) = 37.2024$$

$$\sum x_i^2 = 4690$$

$$\sum \log(x_i)^2 = 60.2309$$

$$\sum y_j^2 = 140918$$

$$\sum \log(y_j)^2 = 174.8363$$

$$\sum x_i y_j = 25182$$

$$\sum \log(x_i) y_j = 2795.2484$$

$$\sum x_i \log(y_j) = 772.3504$$

$$\sum \log(x_i) \log(y_j) = 96.1974$$

2. ¿Cuánto varía la tasa de incidencia de gripe cada día según el modelo lineal?

Datos

$X \equiv$ Días

$Y \equiv$ Tasa de gripe

Recta de regresión de Y sobre X :

$$y = 39.7951 + 4.0626x$$

3. Calcular la tasa de incidencia de gripe a los 50 días desde el comienzo del estudio mediante un modelo de regresión exponencial.

Datos

$X \equiv$ Días

$Y \equiv$ Tasa de gripe

$$\bar{x} = 19.25$$

$$s_x^2 = 215.6875$$

$$\bar{y} = 118$$

$$s_y^2 = 3690.75$$

$$s_{xy} = 876.25$$

$$\sum \log(x_i) = 19.8494$$

$$\sum \log(y_j) = 37.2024$$

$$\sum \log(x_i)^2 = 60.2309$$

$$\sum \log(y_j)^2 = 174.8363$$

$$\sum \log(x_i)y_j = 2795.2484$$

$$\sum x_i \log(y_j) = 772.3504$$

$$\sum \log(x_i) \log(y_j) = 96.1974$$

4. ¿Cuál de las predicciones anteriores es más fiable?
Razonar la respuesta.

Datos

$$X \equiv \text{Días}$$

$$Y \equiv \text{Tasa de gripe}$$

$$\bar{x} = 19.25$$

$$s_x^2 = 215.6875$$

$$\bar{y} = 118$$

$$s_y^2 = 3690.75$$

$$s_{xy} = 876.25$$

$$\overline{\log(y)} = 4.6503$$

$$s_{\log(y)}^2 = 0.2293$$

$$s_{x \log(y)} = 7.0255$$