

Práctica VI: Relación entre variables. Regresión

Estadística I

Curso 2006/2007

La siguiente tabla muestra la superficie (en m^2 .) y el precio (en euros) del alquiler de cada uno de los 10 pisos incluidos en una muestra seleccionada en una ciudad:

Superficie	100	90	80	110	85	95	85	80	70	95
Alquiler	470	510	420	550	440	450	470	430	430	510

1. ¿Son linealmente independientes el precio del alquiler y la superficie de un piso?
2. Si queremos predecir el precio del alquiler de un piso en base a su superficie, ¿qué modelo de regresión se ajusta mejor a los datos, el modelo lineal o el modelo exponencial?
3. Utilizando el modelo (de los dos propuestos) que mejor se ajusta a los datos, ¿cuáles serían las predicciones para los alquileres de pisos de 98 y 150 m^2 ? Comenta la confianza que te merecen dichas predicciones.
4. Obtener los residuos del modelo de regresión anterior y realizar un estudio descriptivo básico de los mismos.
5. Generar una muestra $\{(x_i, Y_i)\}_{i=1}^{100}$, con $x_i = i/100$, $i = 1, \dots, 100$ e $Y_i = (2 + 3x_i) + \varepsilon_i$, donde ε_i es una muestra aleatoria de una variable $N(0, 1)$.
 - a) Obtener el diagrama de dispersión de (x, Y) .
 - b) Ajustar un modelo de regresión lineal a la nube de puntos y observar las estimaciones de los parámetros.
 - c) Valorar la bondad del ajuste y compararlo con modelos alternativos.
 - d) Calcular el coeficiente de correlación lineal entre Y y x .
 - e) Repetir el proceso para los valores de $\sigma = 0, 0.5, 2$ y 5 .