

EJERCICIOS PROPUESTOS ANÁLISIS DE REGRESIÓN

Ejercicio 1:

Volver, de nuevo, al ajuste lineal realizado en el ejercicio 4, tras eliminar todos los datos que daban residuos atípicos, y resolver las siguientes cuestiones:

- Indicar que modelo es el más adecuado y el valor de R^2 para dicho modelo.
(Solución: Raíz cuadrada-Y, $R^2 = 0.9188$)
- Ajustar el modelo anterior e indicar su estimación.
(Solución: $Edad = (3.96728 + 0.1684TiempEmp)^2$)
- Indicar si hay otros modelos que sean más o menos igual de adecuados.
(Solución: el modelo lineal y exponencial también parecen igualmente adecuados)
- Estudiar si hay residuos atípicos en el nuevo modelo y eliminarlos del estudio. Indicar si hay un nuevo modelo que se ajuste mejor a esos datos, su estimación y su R^2 .
(Solución: Tras eliminar los residuos atípicos el mejor modelo es el exponencial,
 $Edad = \text{Exp}(2.8697 + 0.05664TiempEmp)$, $R^2 = 0.95$)
- Indicar la predicción de $Edad$ que da el modelo indicado en el apartado anterior para un empleado que lleva 14 años en la empresa.
(Solución: La predicción es de 38 años)

NOTAS Y RESPUESTAS:

EJERCICIOS PROPUESTOS ANÁLISIS DE REGRESIÓN

Ejercicio 2:

- Para los datos del fichero **miempresa.sf3** añade una nueva variable denominada *PagaEx* e introduce sus datos:

Id	1	2	3	4	5	6	7	8	9	10
PagaEx	1889	1851	1828	1783	1843	1813	1864	1860	1768	1838
Id	11	12	13	14	15	16	17	18	19	20
PagaEx	1827	1796	1850	1867	1839	1776	1833	1857	1909	1880
Id	21	22	23	24	25	26	27	28	29	30
PagaEx	1900	1853	1864	1826	1798	1779	1877	1841	1866	1866
Id	31	32	33	34	35	36	37	38	39	40
PagaEx	1828	1888	1828	1793	1844	1856	1899	1871	1824	1781

- Realiza un diagrama de dispersión para las variables $X = \text{Salari07}$ e $Y = \text{PagaEx}$ y determina los siguientes valores:
 - a) Valor de la variable *PagaEx* del empleado (hombre) que en la nube de puntos está algo alejado del conjunto.
 - b) Coeficiente de Correlación Lineal.
 - c) Covarianza de las variable bajo estudio.
 - d) Varianza de la variable *PagaEx*.
- Realiza un estudio de regresión lineal simple para la variable *PagaEx* en función de la variable *Salari07* y determina los siguientes valores:
 - a) Valor de la ordenada del modelo estimado.
 - b) Valor de la pendiente del modelo estimado.
 - c) Modelo ajustado.
 - d) Coeficiente de Determinación.
 - e) Varianza Residual del modelo.

NOTAS Y RESPUESTAS:

EJERCICIOS PROPUESTOS ANÁLISIS DE REGRESIÓN

Ejercicio 3:

- Para los datos del fichero **miempresa.sf3**, con la nueva variable *PagaEx* introducida en la práctica anterior, vuelve a realizar un estudio de regresión lineal simple para las variables $X = \text{Salari07}$ e $Y = \text{PagaEx}$ y determina los siguientes valores:
 - a) El valor de la variable *Id* de los residuos atípicos y eliminarlos del estudio.
 - b) Nuevo modelo ajustado y el Coeficiente de Determinación.
 - c) El mejor modelo que se puede ajustar, su estimación y R^2 .
 - d) Valor de la variable *PagaEx* que predice el modelo del apartado anterior para el empleado con $Id=27$ y error de dicha estimación.
 - e) Valor de la variable *PagaEx* que predice el modelo del apartado (c) para el empleado con *Salari07* 1950 euros.

NOTAS Y RESPUESTAS:

EJERCICIOS PROPUESTOS ANÁLISIS DE REGRESIÓN

Ejercicio 4:

Realiza un estudio de regresión lineal simple para la variable *Seguro07* en función de la variable *PagaEx* y determina los siguientes valores:

- a) El modelo ajustado y el Coeficiente de Determinación.
- b) El mejor modelo que se puede ajustar, después del lineal, su estimación y R^2 .
- c) Valor de la variable *Seguro07* que predice el modelo del apartado anterior para los empleados viudos y error de dicha estimación.
- d) Mirar la columna de residuos del ajuste del apartado (b) e indicar si se puede sacar alguna conclusión sobre las estimaciones y los errores indicados en el apartado anterior.
- e) Valor de la variable *Seguro07* que predice el modelo del apartado (b) para el empleado con *PagaEx* 1850 euros.
- f) Valor de la variable *Seguro07* que predice el modelo del apartado (b) para el empleado con *PagaEx* 1500 euros. Explicar este resultado.

NOTAS Y RESPUESTAS:

EJERCICIOS PROPUESTOS ANÁLISIS DE REGRESIÓN

Ejercicio 5:

En el fichero `preciopisos.sf3` se recogen datos sobre el precio de algunas viviendas seleccionadas aleatoriamente de Granada. Se pretende explicar la variación del precio de los pisos a partir de variables como la superficie en metros cuadrados, el número de cuartos de baños y la edad del inmueble. Realiza un ajuste de regresión lineal múltiple y contesta razonadamente las siguientes preguntas.

- a) Representa el gráfico de dispersión matricial
- b) Obtén la ecuación ajustada de regresión. Interpreta los coeficientes de regresión estimados
- c) ¿Es significativo el modelo de regresión considerado?
- d) ¿Puede eliminarse alguna variable del modelo? Razona la respuesta
- e) ¿Qué precio estimado tendrá un piso de 120 metros cuadrados, con 2 cuartos de baño y una edad de 15 años?
- f) ¿Son todas las variables válidas para explicar el precio de las viviendas?
- g) Obtén el valor de la suma de cuadrados debido a la regresión. Interpretación
- h) ¿Qué porcentaje del aumento del precio de los pisos no está explicado por el modelo propuesto? Razona la respuesta
- i) ¿Hay indicios de alteraciones en la normalidad de los residuos en el gráfico probabilístico normal?
- j) ¿Existe patrón o tendencia en el gráfico de los residuos frente a las predicciones? ¿A qué conclusiones puedes llegar?
- k) ¿Son independientes los residuos?

NOTAS Y RESPUESTAS: