

Actividad. Regresión Lineal Simple

1. Si pretendiésemos explicar un suceso y/o fenómeno acontecido en el pasado ¿Podemos inferir la respuesta asociada a dichos eventos en base a los restos materiales presentes?
2. Haciendo referencia al análisis de correlación lineal de Pearson, ¿establece este algún tipo de relación causa-efecto de una variable sobre otra(s)?
3. Define causalidad. Exponga algún ejemplo.
4. ¿Podrías mencionar los parámetros involucrados en la ecuación de regresión lineal?
5. En un plano cartesiano, si afirmo que el eje 'x' también se denomina eje de ordenadas, ¿estoy en lo cierto?
6. ¿Sabrías diferenciar entre recta de regresión y plano de regresión?
7. ¿Cuáles son los supuestos (o hipótesis) del análisis de regresión lineal?
8. Dados los siguientes datos, calcula la recta de regresión que mejor se adapte a nuestra nube de puntos siendo "cuentas" la variable dependiente o de respuesta y "distancia" la variable independiente o explicativa.

cuentas	distancia
110	1.1
2	100.2
6	90.3
98	5.4
40	57.5
94	6.6
31	34.7
5	65.8
8	57.9
10	86.1

Figura 1: Tabla de datos referidos a número de cuentas por yacimiento y distancia (km) del yacimiento a la mina. Fuente: Elaboración propia.

9. ¿Serías capaz de interpretar el significado de los parámetros de la ecuación de regresión?
10. ¿Qué implicaciones conlleva obtener un intercepto con valor '0'?
11. ¿Qué ponderación lleva a cabo el análisis de regresión lineal para calcular los valores de los parámetros que configuran la recta de regresión?
12. ¿Cuál sería el error asociado a mi modelo en la estimación del número de cuentas para un yacimiento que se encuentra a 1.1 km de la mina?
13. ¿Cómo calcularías los residuos del modelo dado los siguientes datos?

cuentas	predicciones
6	-6.682842
98	85.520196
40	28.938591
94	84.216973
31	53.69983
5	19.924631
8	28.504183
10	-2.121561

Figura 2: Tabla de datos referidos a número de cuentas por yacimiento y su estimación atendiendo al modelo lineal. Fuente: Elaboración propia.

14. Con los datos residuales, verifica si se cumple (o no) el supuesto de normalidad.
15. ¿Que 2 de conjuntos (de datos) se han de emplear en la modelización lineal? ¿Cómo llevarías a cabo la preparación de estos?
16. Evalúa la capacidad predictiva del modelo implementando una validación cruzada simple.

17. Si mis coeficientes de regresión se han calculado con un intervalo de confianza del 95% ¿cuál será la probabilidad de que la correlación lineal entre los coeficientes de regresión y la variable de respuesta o explicada se deba al azar? ¿Y si tengo un nivel de significación (Alpha) de 0.01, con que Intervalo de Confianza he obtenido mis coeficientes de regresión?
18. Si las estimaciones arrojadas por mi modelo lineal resultan menos precisas (mayor error) en un determinado rango de valores con respecto a otro, decimos ¿qué hay indicios de homocedasticidad o heterocedasticidad?
19. ¿Qué medida de precisión estadística nos indica el % de variabilidad explicada de la variable dependiente por nuestro modelo lineal?
20. Explica la diferencia entre una observación atípica y una observación que produzca lo que se conoce en estadística como “apalancamiento” del modelo?