



Práctica 5 – Regresión lineal y correlación

OBJETIVOS

Los objetivos de esta práctica son los siguientes:

- Continuar el estudio estadístico con los datos que provienen de la encuesta.
- Estudio de la correlación entre dos variables.
- Realizar ajustes lineales entre dos variables. Medir la bondad de dichos ajustes. Hacer predicciones.
- Usar funciones estadísticas para el estudio de la correlación y la regresión con la hoja de cálculo Excel.

EJERCICIO 1. Regresión lineal y correlación con datos de encuesta (Hoja 'Estatura-Peso')

Se han seleccionado los datos de las variables Estatura y Peso para aquellos estudiantes que respondieron a ambas cuestiones. Los datos son los que se recogen en la hoja 'Estatura-Peso'.

- a) Realiza un diagrama de dispersión para estudiar la relación entre las variables. ¿Qué conclusiones puedes extraer de dicho diagrama?
- b) ¿Qué variable, la estatura o el peso, es más homogénea?
- c) Calcula la covarianza e interpreta el resultado. Calcula una medida que te permita cuantificar la intensidad de la relación lineal entre las variables.
- d) Hallar la expresión de la recta de regresión del peso sobre la estatura utilizando funciones de Excel. Interpreta los parámetros de la recta obtenida. Aprovechando el gráfico realizado en el apartado a), muestra en el mismo la recta de regresión. Obtener también la expresión de la recta de regresión de la estatura sobre el peso.
- e) Calcula los valores teóricos y los residuos. Obtén la varianza residual y la varianza explicada. Comprueba que la varianza residual coincide con la varianza de los residuos (¿por qué?) y que la varianza explicada coincide con la varianza de los valores teóricos.



- f) Estudiar la bondad del ajuste anterior numéricamente. Interpreta la medida que utilices para medir dicha bondad. Estudiar la bondad del ajuste anterior a partir del gráfico de los residuos.
- g) Predice el peso que tendría un estudiante con una altura de 180 cm. ¿Es fiable la predicción? ¿Y con una altura de 210 cm? ¿Y con una altura de 194 cm?

EJERCICIO 2. Regresión lineal y correlación (Hoja 'Sandías')

Se sabe que existe una relación lineal entre la cantidad de lluvia caída en el mes anterior a la recolección (X) en Dl. por m² y la producción de sandías (Y) en kg por m². En un estudio llevado a cabo en diferentes lugares del país, se han obtenido los siguientes resultados:

X	6	10	12	14	16	18	22	24	26	32
Y	40	44	46	48	52	58	60	68	74	80

- a) Representar gráficamente las variables mediante el diagrama de dispersión.
- b) ¿Qué variable, la cantidad de lluvia o la producción de sandías, es más homogénea?
- c) Calcular la covarianza y el coeficiente de correlación lineal. ¿Son las variables incorreladas? ¿Existe una relación lineal directa, inversa o no la hay?
- d) Obtener la recta de regresión mínimo cuadrática de la producción de sandías sobre la cantidad de lluvia caída en el mes anterior a la recolección. Interpretar los coeficientes de dicha recta.
- e) Evaluar la bondad del modelo ajustado en el apartado anterior numérica y gráficamente. Interpretar los resultados.
- f) Predecir la producción de sandías si la lluvia caída ha sido de 23,5 Dl por m² ¿Es una predicción fiable?

PROBLEMAS PROPUESTOS

En la hoja de Excel aparecen, además, tres hojas nuevas (Publicidad, Producción Agrícola y Población) para que los alumnos puedan realizar más ejercicios, similares a los anteriores.