



SERIE TEMA 5 REGRESIÓN LINEAL.

AutoraS: Védaur Allié Sarmiento Torres y Sandra Alvarado López.

1.- De datos sobre la resistencia a la compresión (psi) “X” y la permeabilidad intrínseca “Y” de varias mezclas de concreto, se obtuvieron los siguientes resultados:

$$n = 14, \sum y_i = 572, \sum y_i^2 = 23,53, \sum x_i = 43, \sum x_i^2 = 157.42, \sum x_i y_i = 1697.80$$

Suponga que las dos variables están relacionadas según el modelo de regresión lineal simple.

- Obtener la ecuación, utilizando el método de mínimos cuadrados.
- Usa la ecuación de la línea ajustada para predecir qué permeabilidad se observaría cuando la resistencia a la compresión es 4.3 psi.

Respuestas: a) $\beta_0 = 48.01, \beta_1 = -2.33$ b) 37.991

2.- La fábrica Plastics AB, fabrica tensores plásticos de cinturones de seguridad para General Motors. Después del ensamble y la pintura final, se colocan las partes en una banda transportadora que las lleva a la estación de inspección final.

La velocidad con la que las piezas pasan a dicha inspección depende de la velocidad de la banda (pies por minuto). Aunque es deseable contar con líneas más rápidas, a la gerencia le preocupa que un incremento de la velocidad impida que los inspectores identifiquen las partes defectuosas.

Para probar esta teoría, la empresa lleva a cabo un experimento en el que el mismo lote de partes, con un número conocido de piezas defectuosas, se inspecciona a diferentes velocidades.

Se recaban los siguientes datos:

| Velocidad de la línea | Número de partes defectuosas encontradas |
|-----------------------|--|
| 20 | 23 |
| 20 | 21 |
| 30 | 19 |
| 30 | 16 |
| 40 | 15 |
| 40 | 17 |
| 50 | 14 |
| 50 | 11 |

- Trace un diagrama de dispersión con la velocidad de la línea como variable independiente.
- ¿Qué indica el diagrama elaborado en el inciso anterior sobre la relación entre las dos variables?

- c) Encuentre la ecuación de regresión estimada.
- d) Prediga la cantidad de partes defectuosas que se encontrarán en la línea a una velocidad de 25 pies por minuto.
- e) Diga si el modelo obtenido es recomendable para estimar el número de partes defectuosas encontradas cuando la velocidad de la línea es de 15 (pies por minuto)

Respuestas: c) $\beta_0 = 27.5$, $\beta_1 = -0.3$ d) 20

3.- En cierta región, se han registrado los niveles anuales de precipitación pluvial en pulgadas (X) y la producción de trigo (Y) durante los últimos 8 años. Los datos se muestran a continuación:

| Precipitación pluvial (ln) (X) | Producción de trigo (toneladas por acre) (Y) |
|--------------------------------|--|
| 12.9 | 62.5 |
| 7.2 | 28.7 |
| 11.3 | 52.2 |
| 18.6 | 80.6 |
| 8.8 | 41.6 |
| 10.3 | 44.5 |
| 15.9 | 71.3 |
| 13.1 | 54.4 |

- a) Obtener la ecuación del mejor ajuste utilizando el método de mínimos cuadrados.
- b) Calcular el coeficiente de correlación.
- c) ¿Cómo interpretas el valor que obtuviste para el coeficiente de correlación?
- d) Para un año en particular suponga que se produjeron 66.53 toneladas de trigo por acre, ¿Cuál fue la cantidad anual de precipitación pluvial?

Respuestas: a) $\beta_0 = 0.2292$, $\beta_1 = 4.42$ b) 0.982 d) 14.81 ln.

4.- Una revista de ciclismo está interesada en la revisión de bicicletas, principalmente de carreras. Uno de los factores más importantes en la selección de una bicicleta de carreras es el peso de esta.

Los siguientes datos contienen el peso en libras y el precio en dólares de 10 marcas de las bicicletas revisadas.

| Peso | Precio |
|------|--------|
| 17.8 | 2100 |
| 16.1 | 6250 |
| 14.9 | 8370 |
| 15.9 | 6200 |
| 17.2 | 4000 |
| 13.1 | 8600 |
| 16.2 | 6000 |
| 17.1 | 2580 |
| 17.6 | 3400 |

14.1 | 8000

- Determine la ecuación de regresión estimada para saber el precio de una bicicleta cuando se conoce su peso.
- ¿Cuál es la bondad de ajuste de la ecuación de regresión estimada? Explique el valor obtenido.
- ¿Cuál es el valor del coeficiente de correlación? Indique qué tan fuerte es la relación entre el peso y el precio de las bicicletas.

Respuestas: a) $\beta_0 = 28574$, $\beta_1 = -1439$ b) 0.8637 c) -0.9293

5.- La fuerza de impulso de un motor (Y) es una función de la temperatura de escape (X) en °F cuando otras variables de importancia se mantienen constantes. Considere los siguientes datos.

| Temperatura de escape (X) | Fuerza de impulso (Y) |
|---------------------------|-----------------------|
| 1760 | 4300 |
| 1652 | 4650 |
| 1485 | 3200 |
| 1390 | 3150 |
| 1820 | 4950 |
| 1665 | 4010 |
| 1550 | 3810 |
| 1700 | 4500 |
| 1270 | 3008 |

- Obtener la ecuación del mejor ajuste utilizando el método de mínimos cuadrados.
- Obtener el coeficiente de determinación e interpreta el valor que se obtiene.

Respuestas: a) $\beta_0 = -1847.63$, $\beta_1 = 3.653$ b) 0.856

6.- A partir de los siguientes datos sobre el volumen de producción y costos de una operación de manufactura de rines de acero, responda las siguientes preguntas.

| Volumen de producción (unidades) | Costo total |
|----------------------------------|-------------|
| 400 | 4000 |
| 450 | 5000 |
| 550 | 5400 |
| 600 | 5900 |
| 700 | 6400 |
| 750 | 7000 |

- Determine la ecuación de regresión estimada a efecto de pronosticar el costo total dado un volumen de producción determinado.
- ¿Cuál es el costo variable por unidad producida?

- c) ¿Qué porcentaje de la variación en el costo total puede explicarse por el volumen de producción?
- d) Use la prueba t para la pendiente, con $\alpha=0.05$, para probar si el volumen de producción se relaciona de forma significativa con el costo total e indique su conclusión.

Compruebe su conclusión del inciso anterior con la construcción del intervalo de confianza para la pendiente, con un nivel de confianza del 95%. ¿Cuál es su conclusión?

Respuestas: a) $\beta_0 = 1246.7$, $\beta_1 = 7.6$ b) 7.6 c) -0.9587

7.- La siguiente tabla muestra la carga pico de potencia para una planta generadora de energía eléctrica y la temperatura alta diaria para una muestra aleatoria de 10 días. Pruebe la hipótesis de que el coeficiente de correlación poblacional r , entre la carga pico de potencia y la temperatura alta, es cero contra la hipótesis alternativa de que sea positivo. Use nivel de significancia de 0.05.

| Temperatura alta (°F) | Carga de pico |
|-----------------------|---------------|
| 95 | 214 |
| 82 | 152 |
| 90 | 156 |
| 81 | 129 |
| 99 | 254 |
| 100 | 266 |
| 93 | 210 |
| 95 | 204 |
| 93 | 213 |
| 97 | 150 |

8.- En la producción de herramientas, el método para deformar acero a temperatura normal es de gran importancia.

Podemos esperar que este proceso afecte la dureza del acero. Para investigar esta relación se ha tomado la muestra de la siguiente tabla.

| Deformación X (mm) | Dureza Y (kg/mm ²) |
|--------------------|--------------------------------|
| 6 | 68 |
| 9 | 67 |
| 11 | 65 |
| 13 | 53 |
| 22 | 44 |
| 26 | 40 |
| 28 | 37 |

- a) Trace el diagrama de dispersión.
- b) Encuentre la recta de regresión estimada.

- c) Calcule la varianza residual.
- d) Construya el intervalo de confianza del 95% para la pendiente e interprételo.
- e) Calcule el coeficiente de correlación e indica que tan fuerte es la relación lineal de las variables.
- f) Realice una prueba de hipótesis no direccional del coeficiente de correlación, para probar la ausencia de correlación, con un nivel de significancia del 10%.

Respuestas: a) $\beta_0 = 75.72$, $\beta_1 = -1.3196$ c) $S_x^2 = 11.022$ e) -0.9773

9.- En la siguiente tabla aparece información acerca de ocho automóviles de cuatro cilindros considerados entre los más eficientes en consumo de combustible en 2006. Los tamaños de los motores se dan en volumen total de cilindros, medido en litros (L).

| Automóvil | Volumen de cilindros (X) | Caballos de Fuerza (Y) |
|----------------|--------------------------|------------------------|
| Honda Civic | 1.8 | 51 |
| Toyota Prius | 1.5 | 51 |
| VW Golf | 2.0 | 115 |
| VW Beetle | 2.5 | 150 |
| Toyota Corolla | 1.8 | 126 |
| VW Jetta | 2.5 | 150 |
| Mini Cooper | 1.6 | 118 |
| Toyota Yaris | 1.5 | 106 |

- a) Realiza un diagrama de dispersión y obtén el coeficiente de correlación y de determinación e interprétalos.

Respuesta: 0.685 0.4692

10.- Se está estudiando el nivel de ruido en decibeles producido por las máquinas excavadoras de una construcción, evaluando diferentes niveles de potencia en Watts.

| Potencia (watts) X | Ruido (decibeles) Y |
|-----------------------|------------------------|
| 100 | 55 |
| 500 | 80 |
| 800 | 86 |
| 900 | 89 |
| 1000 | 90 |

- a) Estimar el modelo de regresión lineal.
- b) Estimar el modelo logarítmico mediante una transformación lineal.
- c) Estimar el modelo hiperbólico mediante una transformación lineal.

Por el criterio del coeficiente de determinación, indicar cuál de los tres modelos anteriores proporciona un mejor ajuste.