

Tema 6

Modelos de regresión

- 6.1** Ajustar una recta a través de los cinco puntos siguientes. Obtener las estimaciones para β_0 y β_1 . Representar gráficamente los puntos así como la recta obtenida.

x	-2	-1	0	1	2
y	3	2	1	1	0.5

¿Presentan los datos suficiente evidencia para indicar que la pendiente β_1 difiere de cero? (Utilizar un nivel de significación del 5%). Obtener un intervalo de confianza al 95% para β_0 y para β_1 . Calcular e interpretar una medida de la bondad del ajuste realizado.

- 6.2** Los experimentos diseñados para medir valores de CL50 en la investigación de los efectos de cierto producto tóxico en peces se efectúan con dos métodos diferentes. En un método el agua fluye continuamente a través de los tanques del laboratorio dinámico y el otro método tiene condiciones de agua en reposo. A fin de establecer los criterios para sustancias tóxicas, la Agencia para la protección ambiental pretende ajustar todos los resultados a la condición dinámica, por lo que se requiere un modelo para relacionar los dos tipos de observaciones. Las observaciones acerca de ciertos productos tóxicos en ambas condiciones, estática y dinámica, dieron los siguientes resultados (las mediciones están dadas en partes por millón)

Producto tóxico	CL50 dinámico, y	CL50 estático, x
1	23.00	39.00
2	22.30	37.50
3	9.40	22.20
4	9.70	17.50
5	0.15	0.64
6	0.28	0.45
7	0.75	2.62
8	0.51	2.36
9	28.00	32.00
10	0.39	0.77

Ajustar el modelo $y = \beta_0 + \beta_1 x + \varepsilon$. Interpretar los resultados. Estimar el valor dinámico para un producto tóxico con un valor estático de CL50 de $x = 12$ partes por millón.

- 6.3** Se realizó un experimento para observar el efecto de un incremento de la temperatura en la efectividad de un antibiótico. Se almacenaron tres porciones de una onza del antibiótico durante el mismo periodo de tiempo a cada una de las siguientes temperaturas: 30^0 , 50^0 , 70^0 y 90^0 . Las lecturas de la efectividad observadas a la temperatura del periodo experimental fueron:

Lecturas de la efectividad (y)	38, 43, 29	32, 26, 33	19, 27, 23	14, 19, 21
Temperatura (x)	30^0	50^0	70^0	90^0

- (a) Mediante el método de mínimos cuadrados, ajusta una recta a los datos observados.
- (b) Representar gráficamente los puntos y la recta obtenida.
- 6.4** Con objeto de determinar la relación entre la razón agua/cemento y la resistencia de las vigas fabricadas con esa mezcla de materiales, se realizó una serie de pruebas en las que se recogieron los siguientes resultados:

Razón agua/cemento	1.21	1.29	1.37	1.46	1.62	1.79
Resistencia	1.302	1.231	1.061	1.040	0.803	0.711

- (a) Ajustar el modelo $y = \beta_0 + \beta_1 x + \varepsilon$.
- (b) Probar $H_0 : \beta_1 = 0$ frente a $H_1 : \beta_1 < 0$, con $\alpha = 0.05$. ¿Que se podría decir cuando $x = 0$?
- (c) Encontrar un intervalo de confianza al 90% para la resistencia esperada de la viga cuando se utiliza una relación agua/cemento de 1.5. ¿Qué ocurre con el intervalo de confianza si tratamos de estimar promedios de resistencia para razones de agua/cemento de 0.3 o 2.7?
- 6.5** En la siguiente tabla la variable “ x ” representa la fuerza de tensión aplicada a una barra de acero en miles de libras e “ y ” es el estiramiento resultante en milésimas de pulgada:

x	1	2	3	4	5	6
y	14	33	40	63	76	85

- (a) Representar gráficamente los datos para comprobar si es razonable suponer que la regresión de y sobre x es lineal.
- (b) Encontrar la recta de mínimos cuadrados y predecir el estiramiento cuando la tensión es de 3.5 miles de libras. Dar un intervalo de confianza al 95% para el estiramiento medio de la barra cuando se le aplica esa tensión.
- (c) Construir un intervalo de confianza para la pendiente con un nivel de confianza del 95%.

- 6.6** Las materias primas utilizadas en la producción de una fibra sintética son almacenadas en un lugar donde no se tiene control sobre la humedad. Las mediciones de la humedad relativa en el lugar de almacenamiento y la humedad en una muestra de las materias primas (ambas en porcentajes) en 12 días dieron los siguientes resultados:

Humedad (x)	Contenido de humedad (y)
42	12
35	8
50	14
43	9
48	11
62	16
31	7
36	9
44	12
39	10
55	13
48	11

- Determinar la recta de regresión de y sobre x . Dar una medida de la bondad del ajuste realizado.
- ¿Es razonable suponer que la humedad del almacén influye de forma lineal en la humedad de la fibra? (Tomar $\alpha = 0.05$)
- Encontrar un intervalo de confianza del 99% para el contenido medio de humedad de las materias primas cuando la humedad del lugar de almacenamiento es del 40%.

- 6.7** Una compañía de seguros desea relacionar la cantidad y (en miles de euros) de daños producidos por el fuego, en incendios residenciales, con la distancia x (en Km.), entre la residencia y el parque de bomberos más próximo. Una muestra de 10 incendios arroja los siguientes datos:

x	3.4	1.8	4.6	2.3	3.1	5.5	0.7	3	2.6	4.3
y	26.2	17.8	31.3	23.1	27.5	36	14.1	22.3	19.6	31.3

- Construir un modelo de regresión lineal del tipo $y = \beta_0 + \beta_1 x + \epsilon$.
- Obtener un intervalo de confianza al 95% para la pendiente.
- Construir un intervalo de confianza al 99% para el gasto medio producido por un incendio ocasionado a 3.5 Km. de la estación de bomberos más cercana.

- 6.8** Ajustar un modelo del tipo $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \epsilon$ a los datos de la siguiente tabla. Representar gráficamente los puntos así como la parábola ajustada.

x	-3	-2	-1	0	1	2	3
y	1	0	0	-1	-1	0	0

- 6.9** Las estadísticas de enfermedades en Florida para la década que terminó en 1976 demuestran que la hepatitis infecciosa tenía las tasas de incidencia que se presentan en la siguiente tabla (expresadas en casos por población de 100.000).

x	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976
y	10.5	18.5	22.6	27.2	31.2	33.0	44.9	49.4	35.0	27.6

- (a) Sea y la tasa de incidencia y x el año “codificado” (0 para 1967, 1 para 1968, hasta 9 para 1976) . Ajustar un modelo del tipo $y = \beta_0 + \beta_1 x + \varepsilon$.
- (b) A los mismos datos, ajustar el modelo $y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon$.

¿Qué modelo es más adecuado?

- 6.10** Los datos muestrales siguientes fueron registrados por una compañía de mudanzas sobre los pesos de seis equipos, las distancias a que fueron trasladados y los daños sufridos:

Peso (1000 Kg.)	Distancia (1000 Km.)	Daño (100 E.)
4.0	1.5	16
3.0	2.2	11.2
1.6	1.0	6.9
1.2	2.0	9
3.4	0.8	12.3
4.8	1.6	18.6

- (a) Ajustese una ecuación de la forma $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$.
- (b) Estimar los daños que sufrirá un equipo que pesa 2400 Kg. cuando se traslada 1200 Km.

- 6.11** Se realiza un experimento para determinar la duración de vida (y) de ciertos circuitos electrónicos en función de dos variables de agrupación x_1 y x_2 , con los siguientes resultados:

y	11	8	73	21	46	30
x_1	-10	0	10	-10	0	10
x_2	0	-5	5	0	5	-5

- (a) Ajustar a los datos un modelo del tipo $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$.
- (b) Calcular e interpretar una medida de la bondad del ajuste realizado.

- 6.12** Mediante el método de los mínimos cuadrados, ajustar una función de la forma $y = kx^b$, donde $k > 0$ y $b > 0$, a la nube de puntos:

X	1.1	1.5	2	2.5	2.9
Y	1.4	2.3	4.2	6.4	9

Determinar una medida de la bondad del ajuste realizado.

- 6.13** La siguiente tabla muestra los valores experimentales de la presión P , de una masa de gas, correspondientes a diferentes valores de su volumen V .

Volumen en pulgadas cúbicas	54.3	61.8	72.4	88.7	118.6	194
Presión en libras/pulgada cuadrada	61.2	49.5	37.6	28.4	19.2	10.1

Según los principios de la termodinámica, existe una relación entre las variables que tiene la forma $P \cdot V^\lambda = C$, donde λ y C son constantes.

- (a) Hallar los valores de λ y C .
- (b) Hallar la ecuación que liga a P y V .
- (c) Estimar P cuando $V = 100$ pulgadas cúbicas.

6.14 El número de cm. y que una estructura recién construida se hunde en el suelo, viene dado por

$$y = 3 - \beta e^{-\alpha x}$$

donde x es el número de meses que lleva construido.

Del estudio de seis estructuras se recogen los siguientes datos

x	2	4	6	12	18	24
y	1.07	1.88	2.26	2.78	2.97	2.99

Estimar α y β y calcular una medida de la bondad del ajuste realizado. (Sugerencia: obsérvese que la relación entre $\ln(3 - y)$ y x es lineal)

