

HOJA 3 : MODELOS DE REGRESIÓN LINEALIZABLES

1. En un experimento controlado, se observa el tamaño Y (en cm) alcanzado por seis plantas durante una semana, tras suministrárseles diferentes cantidades X de hormona del crecimiento. Los resultados son los siguientes:

$X(mgr)$	2	5	10	15	20	25
$Y(cm)$	8	16	20	23	25	26

- a) Ajustar los datos a un modelo de regresión $Y = a + b \ln X$, dibujando los datos y la curva correspondiente, y evaluando la bondad del ajuste.
 b) Si por regulación sanitaria X debe ser menor que 14 mgr, ¿qué tamaño máximo alcanzarán las plantas?
 c) A una planta de 24 cm, ¿qué cantidad probable de hormona se le habrá suministrado?
2. En Botánica, el *muestreo de áreas contiguas* permite estimar el número de especies distintas que habitan un ecosistema. El recuento se realiza comenzando con un cuadrado de área 1 m^2 , y duplicando en cada paso el área con cuadrados contiguos. El modelo que relaciona $N =$ “número de especies diferentes” con $S =$ “superficie en m^2 ” suele cumplir la relación

$$N = a \log_2 S + b,$$

donde los parámetros miden $a =$ índice de diversidad, $b =$ número medio de especies por unidad de área.

En una zona protegida se observa que

$S :$	1	2	4	8	16	32	64
$N :$	3	4	7	11	16	19	21

- a) Ajustar dicho modelo a los datos, estimando los parámetros a y b , y valorando la calidad del ajuste.
 b) Estimar el número de especies distintas en una parcela de 28 m^2 , y el tamaño aproximado de una parcela con 9 especies distintas.
3. En un laboratorio se estudia la descomposición del nitrometano CH_3NO_2 en presencia de exceso de base. Para ello se mide la concentración $C(t)$ de nitrometano en varios instantes de tiempo, obteniéndose los datos

t (en seg)	1/4	1/2	3/4	1	1'5
C (en $[mM]$)	3'86	2'59	1'84	1'21	0'51

- (a) Representar los datos y ajustarlos a una curva $C(t) = ae^{-bt}$, evaluando la calidad del ajuste.
 (b) Da una estimación de la concentración inicial, y del tiempo que tarda en bajar la concentración hasta 1 mM.
4. En un estudio sobre proliferación de bacilos de la fiebre tifoidea, se expusieron cultivos del bacilo en aguas estancadas durante diferentes períodos de tiempo. Los siguientes datos representan los tiempos de exposición t (en días), y las correspondientes concentraciones de bacilos q medidas en las muestras (en miles/ cm^3).

$t :$	0	1	2	3	5	7	10	15
$q :$	1	1.2	1.6	2.1	3.5	5.6	12.8	43.4

- a) Ajustar una recta y una exponencial a los datos, dibujando datos y curvas. Interpretar cuál de los dos ajustes es más razonable.
 b) Si continuara esta evolución, ¿cuándo se llegaría al millón de bacilos/ cm^3 ?
5. Los siguientes datos corresponden a la evolución del peso celular (en mgr/ml) y al % de nitrato en un cultivo de algas durante 3 días (mediciones cada 24 horas).

Tiempo (T)	Peso (X)	% nitrato (Y)
Inicio	0,07	12,5
1 día	0,19	10,4
2 días	0,52	7,8
3 días	1,07	4,5

- a) Ajustar una curva a la evolución temporal del peso.
 b) Ajustar primero una recta y luego una exponencial a los datos “peso” (X) y “% de nitrato” (Y). ¿Cuál de los dos ajustes parece más razonable?
 c) Mediante lo obtenido en a) y b) estimar el % de nitrato que había en el cultivo al cabo de 36 horas.

6. En biología animal es habitual que partes distintas del cuerpo crezcan con diferentes escalas. Esto se denomina *crecimiento alométrico* (en lugar de isométrico), y corresponde a una relación potencial entre las medidas X e Y de ambas partes

$$Y = aX^b.$$

Se quiere estudiar experimentalmente la relación entre el perímetro craneal X y la altura Y de bebés de 0 a 24 meses, obteniéndose los datos

Edad (meses)	0	3	6	12	18	24
Perímetro craneal (X)	35	41	44	47'2	48'8	49'7
Altura (Y)	50	61'5	67'5	76	82	85

- a) Ajustar los datos a un modelo alométrico $Y = aX^b$, evaluando el ajuste obtenido.
 b) Estimar el perímetro craneal de un bebé de altura 70 cm.
 c) ¿En qué edad es mayor el ratio X/Y entre tamaño de la cabeza y longitud corporal?
7. En Ecología de aves marinas, la regla de Darlington predice que cuando la superficie de nidificación se duplica, el número de especies que la habita aumenta en una proporción constante. Esto se traduce en un modelo potencial

$$N = aS^b,$$

donde N =número de especies y S =superficie de la isla. En una zona insular protegida se recogen los datos

supeficie de la isla (kms ²)	820	387	320	30	12	8	1.8
número de especies	51	40	37	19	14	14	8

- (i) Ajustar los datos a un modelo del tipo anterior.
 (ii) Estimar el número de especies en una isla cercana de superficie 140 Km².
 (iii) ¿Por qué factor se debe multiplicar el área para que se duplique el número de especies?
8. En las reacciones bioquímicas en presencia de una enzima, el modelo teórico de Michaelis-Menten predice que la velocidad de la reacción V aumenta con la concentración inicial de sustrato S según la fórmula

$$V = \frac{aS}{b + S},$$

donde a = velocidad máxima de la reacción y b =concentración de semi-saturación.

En el laboratorio se obtienen experimentalmente los siguientes datos

S (mol/ml)	0'10	0'25	0'50	1	2	4	8
V (mol/min)	21	34'6	44'1	61'1	73'8	73'9	76'3

- (i) Calcula la recta de regresión de $(\frac{1}{S}, \frac{1}{V})$, y estima el valor de a, b usando que $\frac{1}{V} = \frac{b}{aS} + \frac{1}{a}$
 (ii) Calcula la recta de regresión de $(S, \frac{S}{V})$, y estima el valor de a, b usando que $\frac{S}{V} = \frac{b}{a} + \frac{S}{a}$.
 (iii) ¿Cuál de los ajustes anteriores parece más razonable?
9. Un modelo de crecimiento propuesto por J. Monod predice que, para un cultivo de bacterias, la tasa de crecimiento per cápita, T , depende de la concentración de glucosa en el cultivo, C , según la regla

$$T = \frac{aC}{k + C}, \quad \text{donde } a, k > 0.$$

En un laboratorio se miden los datos siguientes

C (% glucosa)	0,005	0,01	0,02	0,04	0,1	0,15	0,2	0,4
T	0,04	0,14	0,23	0,38	0,72	0,81	0,97	1,06

- (i) Estimar a, k linealizando el modelo de la forma $\frac{1}{T} = \frac{k}{a} \frac{1}{C} + \frac{1}{a}$.
 (ii) Estimar a, k linealizando el modelo de la forma $T = a - k \frac{T}{C}$.
 (iii) ¿Cuál de los dos ajustes anteriores te parece más razonable?

Ejercicios complementarios para practicar con ordenador

10. En una piscifactoría se investiga si la concentración \mathbf{C} de clorofila de fitoplancton (en $\mu\text{gr/l}$) durante los meses de verano influye en la producción anual de peces \mathbf{P} (en Kgs por m^3 y año). Se recogen los siguientes datos, correspondientes a mediciones en diversas piscinas

\mathbf{C}	0'3	0'9	2'4	4'7	5'2	7'3
\mathbf{P}	4	6	11	18	19	21

- (i) Dibuja los puntos en el plano \mathbf{C}, \mathbf{P} , y decide (según su aspecto y su coeficiente de correlación) qué tipo de ajuste es mejor de entre $\mathbf{P} = a + b\mathbf{C}$, $\mathbf{P} = a + b \log \mathbf{C}$ y $\mathbf{P} = a\mathbf{C}^b$.
- (ii) Una vez decidido el tipo de ajuste, determina los parámetros a y b , y dibuja en una misma gráfica los puntos y la curva que mejor se ajusta a ellos. ¿Es bueno el ajuste obtenido?
- (iii) Una nueva variedad de fitoplancton tiene una concentración de clorofila de $3 \mu\text{gr/l}$. Estima la producción de peces en este caso. Si esta producción parece poca para el empresario, ¿cuál debería ser la concentración de clorofila para que la producción de peces aumente en un 20%?
11. Considera las siguientes medidas experimentales de la presión parcial de O_2 (en mmHg) a distintas alturas h (en Kms sobre el nivel del mar):

h	0	5	10	15	20	50
P	159	88	50	26	14	0.5

- (i) Dibuja la nube de puntos en el plano h, P , y decide según su aspecto qué tipo de ajuste es mejor de entre $P = a + bh$, $P = ae^{bh}$, $P = ah^b$.
- (ii) Una vez decidido el tipo de ajuste, determina los parámetros a y b , y dibuja en una misma gráfica los puntos y la curva que mejor se ajusta a ellos. ¿Es bueno el ajuste obtenido?
- (iii) Estima las presiones parciales de O_2 a la que se verían sometidos
- los tripulantes de un globo aerostático que alcance los 18 kms de altitud.
 - los mineros que trabajen a 500 metros de profundidad.
12. En un estudio del efecto de la puromicina en una reacción enzimática que sigue el modelo de Michaelis-Menten, $V = aS/(b + S)$ se han obtenido en el laboratorio los siguientes datos (en experimentos con réplica)

S (ppm)	0'02	0'02	0'06	0'06	0'11	0'11	0'22	0'22	0'56	0'56	1'10	1'10
V (dpm)	76	47	97	107	123	139	159	152	191	201	207	200

- (i) Usando regresión en $(\frac{1}{S}, \frac{1}{V})$, ajusta los datos a la curva, y estima los parámetros a, b .
- (ii) Usando regresión en $(S, \frac{S}{V})$, ajusta los datos a la curva, y estima los parámetros a, b .
- (iii) Usa *regresión no-lineal* para ajustar los datos a la curva, y estima los parámetros a, b .
- (iv) Representa en un mismo gráfico las curvas obtenidas y decide cuál de los tres procedimientos parece más adecuado.