MATEMÁTICAS. 1º **grado Bioquímica**

HOJA 2 : REGRESIÓN LINEAL

1. Se han tomado cinco muestras de glucógeno, a las que se ha aplicado una cantidad de glucogenasa (en mmol/l) anotando en cada caso la velocidad de reacción (en μmol/min):

concentr glucogenasa	1	2	3	0′25	0′5
veloc reacción	18	35	60	8	10

   (a) Justificar si la velocidad de reacción aumenta con la concentración de glucogenasa.

   (b) Pronosticar la velocidad de reacción en una muestra con 2'5 mmol/l de glucogenasa.

2. El método de Lowry para determinar la concentración de proteínas en un cultivo consiste en añadir un cierto reactivo y posteriormente medir la intensidad de azul. La ley experimental predice que dicha intensidad Y depende linealmente de la concentración de proteína X. Para calibrar un *colorímetro* en el laboratorio, se miden las intensidades de azul de 8 concentraciones conocidas (en μg/ml), obteniéndose los datos

Concentr	0	25	50	100	150	200	250	300
Intensidad	0	0.070	0.146	0.240	0.359	0.461	0.562	0.683

   (a) Hallar la recta de regresión $Y = a + bX$, la covarianza y el coeficiente de correlación.

   (b) Extraemos una muestra de suero, y medimos su intensidad de azul $y = 0.160$. Da un valor estimado para la concentración de proteína en dicha muestra.

3. El número de manatíes muertos en las costas de Florida aumenta cada año. Se cree puede tener relación con la presencia de lanchas motoras, que hieren o matan a muchos de ellos. Se presenta una tabla que contiene, para cada año, el número de licencias para motoras (en miles) expedidas en Florida y el número de manatíes muertos.

Año	Licencias	Manatíes	Año	Licencias	Manatíes
1997	447	13	2004	559	34
1998	460	21	2005	585	33
1999	481	24	2006	614	33
2000	498	16	2007	645	39
2001	513	24	2008	675	43
2002	512	20	2009	711	50
2003	526	15	2010	719	47

   a) Dibujar un diagrama de dispersión del número de manatíes muertos sobre el número de licencias anuales. ¿Qué nos dice el diagrama sobre la relación entre esas dos variables?

   b) Hallar la recta de regresión del número de manatíes muertos sobre el número de licencias anuales. ¿Hay evidencia de una relación lineal entre ambos factores?

   c) El estado de Florida quisiera mantener la cantidad anual de manatíes fallecidos por debajo de 40; dar una estimación del número de licencias que debería permitir.

4. La tabla siguiente presenta tres conjuntos de datos preparados por el estadístico Frank Anscombe para ilustrar los peligros de hacer cálculos sin antes representar los datos. *Los tres conjuntos de datos tienen la misma correlación y la misma recta de regresión.*

   Conjunto de datos A:

X	10	8	13	9	11	14	6	4	12	7	5
Y	8,04	6,95	7,58	8,81	8,33	9,96	7,24	4,26	10,84	4,82	5,68

   Conjunto de datos B:

X	10	8	13	9	11	14	6	4	12	7	5
Y	9,14	8,14	8,74	8,77	9,26	8,10	6,13	3,10	9,13	7,26	4,74

   Conjunto de datos C:

X	8	8	8	8	8	8	8	8	8	8	19
Y	6,58	5,76	7,71	8,84	8,47	7,04	5,25	5,56	7,91	6,89	12,50

a) Calcula la correlación y la recta de regresión para los tres datos y comprueba que son iguales.

b) Dibuja los diagramas de dispersión con sus rectas de regresión correspondientes.

c) ¿En cuál de los tres casos utilizarías la recta de regresión para predecir Y dado $X = 14$.

5. Se tienen dos variables x e y con medias 7 y 10 y desviaciones típicas 3 y 2. Se conoce además que el coeficiente de correlación es $0'7$. Escribir las ecuaciones de las rectas de regresión para (x, y) y para (y, x), y calcular el error cuadrático medio en cada caso.

6. El tiempo de vida de ciertos insectos parece que se ve influido por el porcentaje de humedad del hábitat en que se encuentran. Para estudiar esto, se toman muestras en 10 hábitats diferentes de las variables X="Porcentaje medio de humedad" e Y="Tiempo medio de vida", obteniéndose un porcentaje medio de humedad del 59% y un tiempo medio de vida 28'7 días. El resto de datos se resumen como:

$$\sum x_i^2 = 35140; \qquad \sum y_i^2 = 8573; \qquad \sum x_i y_i = 17260$$

a) Expresar el tiempo medio de vida en función de la humedad mediante la recta de regresión. Evaluar el ajuste.

b) ¿Cuál sería el tiempo medio de vida estimado si la humedad es del 65%?

7. En poblaciones con capacidad limitada, el modelo logístico establece que la tasa de crecimiento per cápita T, depende linealmente del número de individuos N, según la regla

$$T(N) = r\Big(1 - \frac{N}{K}\Big),$$

donde r =tasa máxima de crecimiento, y K =capacidad máxima de la población. Queremos aplicar este modelo a la población mundial, utilizando los datos:

año	1965	1970	1975	1980	1985	1990	1995	2000	2005
N ($\times 10^9$)	3,335	3,692	4,068	4,435	4,831	5,264	5,674	6,071	6,454
T ($\times 10^{-2}$)	2,648	2,421	2,308	2,067	2,051	2,055	1,810	1,632	1,484

(i) Calcula una recta de regresión para $T = a + bN$, y usa esta información para estimar los valores de r y K.

(ii) Dibuja los datos anteriores y su recta de regresión, y evalúa la bondad del ajuste.

8. En reacciones bioquímicas en presencia de una enzima, la velocidad de la reacción V depende de la concentración de sustrato $[S]$, según la ecuación de Michaelis-Menten

$$V = \frac{v_{\max}\,[S]}{k + [S]}, \tag{1}$$

donde $k, v_{\max}$ son parámetros positivos. Para estimar estos dos parámetros, se puede buscar la recta de regresión entre $1/[S]$ y $1/V$, ya que (1) implica

$$\frac{1}{V} = \frac{k}{v_{\max}}\frac{1}{[S]} + \frac{1}{v_{\max}}.$$

Para una cierta reacción, en el laboratorio se obtienen los datos

$[S]$	0	1	5	15	40	65	80	90	100
V	0	37	75	91	96	98	98.5	99	100

(i) Escribe en una tabla el valor de los datos $1/[S]$ y $1/V$.

(ii) Calcula la recta de regresión $\frac{1}{V} = a + b\frac{1}{[S]}$, y evalua la bondad del ajuste.

(iii) Utiliza lo anterior para obtener una estimación de los parámetros k y $v_{\max}$.

(iv) ¿Qué velocidad de reacción pronosticarías para $[S] = 50$?