

Estadística

Ejercicios TEMA 2. Curso 2011/2012

Grado de Ingeniería Química

2.1 Indicar si las siguientes afirmaciones son verdaderas o falsas y justificarlo brevemente:

- En una muestra se obtiene un valor de la correlación entre X e Y de 0.24. En este caso X e Y muestran una dependencia inversa.
- El valor de la correlación entre dos variables X e Y es el mismo tanto si se pide la recta de regresión de X en función de Y ($X=a+bY$) como si se pide la recta de Y en función de X ($Y=a+bX$).
- El coeficiente de correlación tiene las unidades de la variable.
- Un coeficiente de correlación de -0.87 entre dos variables X e Y, muestra una relación lineal fuerte entre ambas variables.

2.2 Conociendo los siguientes datos:

$$\begin{aligned} \sum_{i=1}^6 x_i &= 464 & \sum_{i=1}^6 x_i^2 &= 36354 \\ \sum_{i=1}^6 y_i &= 482 & \sum_{i=1}^6 y_i^2 &= 39022 & \sum_{i=1}^6 x_i y_i &= 37613 \end{aligned}$$

a) Rellene los huecos de esta tabla:

\bar{x}	$S_n^2(X)$	S(X)	\bar{y}	$S^2(Y)$	Cov(X,Y)

- Estudiar la relación lineal entre esas dos variables.
- Obtener el modelo de regresión lineal de X en función de Y.
- Según el modelo obtenido determinar el valor de y para un valor de x=80

2.3 (8 puntos) Para un conjunto de piezas de mármol que están siendo sometidas a un proceso de pulido, se han recogido datos sobre el peso (Y) que han disminuido desde el inicio del mismo (en gramos) y el tiempo (X) que llevan sometidas a tal proceso (en minutos), de los cuáles se tiene la información:

$$\begin{aligned} \sum_{i=1}^8 x_i &= 77 & \sum_{i=1}^8 y_i &= 75.8 \\ \sum_{i=1}^8 x_i^2 &= 905 & \sum_{i=1}^8 y_i^2 &= 879.34 & \sum_{i=1}^8 x_i \cdot y_i &= 891 \end{aligned}$$

A partir de esta información se pide:

- (3 puntos) Estime el modelo lineal que explica el peso perdido por las piezas en función del tiempo que lleva el proceso de pulido.
- (2.5 puntos) Según el modelo considerado, qué peso se espera que va a perder una pieza que está un cuarto de hora sometida al proceso.
- (2.5 puntos) Calcule el coeficiente de correlación entre ambas variables e indique, razonadamente, el tipo de relación que tienen.
- (2.5 puntos) Calcule la fracción de varianza explicada por el modelo.

2.4 En la siguiente tabla se recogen datos de la presión sanguínea y la edad de una muestra de 12 individuos:

Edad (años)	56	42	70	35	64	47	53	49	38	42	68	60
Presión sanguínea (mmHg)	147	122	160	118	150	130	146	145	113	145	152	155

- Estudiar la relación entre ambas variables.
- Calcular la recta de regresión de la presión sanguínea en función de la edad.
- Estimar la presión sanguínea para una edad de 65 años.
- Determinar el error estándar de la estimación.

- 2.5** Un ciclista se desplaza en línea recta con un movimiento uniforme para el cual según las leyes de la mecánica su posición x en un instante t vendrá dada por la ecuación $x = x_0 + vt$ donde x_0 es la posición inicial y v la velocidad.

Se han tomado los siguientes valores de su posición x en metros y el tiempo t en segundos:

x (metros)	14	26.2	37.7	51	61.8	76	84.2
t (segundos)	2	4	6	8	10	12	14

A partir de estos datos estimar:

- el coeficiente de correlación
 - los valores de la posición inicial y la velocidad del ciclista por medio de una regresión lineal.
 - el espacio recorrido por el ciclista transcurridos 9 segundos.
 - el error estandar de la estimación y la fracción de varianza explicada por el modelo.
- 2.6** En un experimento se han recogido los siguientes datos de presión de un gas para varios valores de su volumen, manteniendo constante la temperatura.

V (mm^3)	54.3	61.8	72.4	88.7	118.6	194.0
P (mm de Hg)	61.2	49.5	37.6	28.4	19.2	10.1

Los principios de la termodinámica relacionan los valores de volumen (V) y presión (P) de un proceso adiabático mediante la expresión $PV^\gamma = C$ siendo γ y C dos constantes. Calcular:

- Los valores de γ y C expresando la presión en función del volumen.
 - Estimar el valor de la presión para un volumen de $90 mm^3$
- 2.7** A partir de los datos de presión de vapor del propano y de la temperatura que se muestran en la siguiente tabla,

T	10	20	40	60	70	100
P	47.5	56.5	79.5	108.6	125.5	188.5

- Linealice la ecuación $P = e^{(A + \frac{B}{T})}$ y determine mediante regresión lineal los parámetros A y B para los datos de la tabla.
 - Con el modelo resultante estime la presión del propano para una temperatura de 65 grados.
- 2.8** Para determinar el efecto de los residuos de una empresa química en un río cercano se toman medidas de la concentración de nitrato en el agua del río en distintos puntos. Con el fin de establecer un sistema de medida automático en el río se toman muestras de forma manual y con el nuevo sistema de medida automático para validar este nuevo sistema. La concentración de nitrato obtenido con cada sistema de medida se muestran en la siguiente tabla:

Sist. Manual	50	35	130	160	85	310	280	460	410	585
Sist. Automático	90	40	160	210	90	360	250	480	330	593

- A partir del valor de la correlación determinar la validez del sistema automático que se quiere establecer.
 - Si el modelo lineal es apropiado, calcular la recta de regresión del sistema automático en función de las medidas obtenidas con el manual.
 - Calcular las predicciones que se obtienen del modelo obtenido para cada una de las medidas resultantes del sistema de medida manual.
 - Determinar los errores que se comenten al hacer esas predicciones.
- 2.9** En la siguiente tabla se recogen datos de un experimento en el que se ha analizado el tiempo de secado de un tipo de pintura dependiendo de la cantidad añadida de un compuesto que reduce el tiempo de secado.

Tiempo de secado (horas)	12	10.5	10	9	8.5	8	7.5	7.2	7
Aditivo (gramos)	1	2	3	4	5	6	7	8	9

Sabiendo que la relación entre la variable tiempo (T) y la cantidad de aditivo (C) se expresa mediante un modelo del tipo $T = \beta C^\alpha$ siendo β y α dos constantes, responder a las siguientes cuestiones:

- dibuje un diagrama de dispersión para verificar que es razonable suponer que la relación es de tipo potencial.
- los valores de β y α .
- el tiempo de secado de la pintura cuando se mezclan 9.5 gramos de aditivo.

2.10 En un estudio se han recogido los siguientes valores de los ingresos totales de 40 familias frente a los gastos fijos por mes en euros:

Gastos \ Ingresos	Ingresos				
	300-600	600-840	840-1100	1100-1350	1350-1700
35-70	2				
70-110	1	3	5		
110-150			8	10	
150-180				6	2
180-300					3

Calcular:

- la recta de regresión mínimo cuadrática de los gastos fijos sobre los ingresos.
- el coeficiente de correlación lineal.
- con la recta calculada estime, los gastos fijos de una familia cuyos ingresos son de 200 euros.