

Ejercicios

TEMA 1: ESTADÍSTICA DESCRIPTIVA

- 1.1** Al validar un método analítico rutinario para detectar residuos de tianfenicol en alimentos se obtienen los siguientes resultados:

Día 1	54.72	56.78	53.33	48.59	53.80	52.06
Día 2	44.31	45.48	49.69	52.17	54.69	66.58
Día 3	52.55	47.34	64.89	45.76	63.52	59.13

Calcula los estadísticos media, mediana, desviación típica muestral, coeficiente de variación, mínimo y máximo para cada uno de los días. Utiliza estas medidas para comparar los días entre sí. [FUENTE: Ellison, Barwick y Duguid Farrant, *Practical Statistics for the Analytical Scientist*, 2009.]

- 1.2** Para investigar la reproducibilidad de un método de determinación de selenio en alimentos, se realizaron nueve medidas sobre un lote de arroz tostado y se obtuvieron (en $\mu\text{g/g}$) los valores:

0.07 0.07 0.08 0.07 0.07 0.08 0.08 0.09 0.08

Calcula la media, la cuasivarianza, la desviación típica muestral, el coeficiente de variación, la mediana y el rango de estos datos. Si se sustituye la primera observación por 0.04, ¿qué efecto tiene sobre estas medidas estadísticas? [FUENTE: Moreno, García y Mariné, *Analyst*, 1983.]

- 1.3** En un experimento se midió la temperatura de sublimación del Ir y del Rh repetidas veces; se obtuvieron los valores siguientes:

IRIDIO								
136.6	145.2	151.5	162.7	159.1	159.8	160.8	173.9	160.1
160.4	161.1	160.6	160.2	159.5	160.3	159.2	159.3	159.6
160.0	160.2	160.1	160.0	159.7	159.5	159.5	159.6	159.5

RODIO							
126.4	135.7	132.9	131.5	131.1	131.1	131.9	132.7
133.3	132.5	133.0	133.0	132.4	131.6	132.6	132.2
131.3	131.2	132.1	131.1	131.4	131.2	131.1	131.1

- a) Calcula media, mediana, cuasivarianza, desviación típica muestral, coeficiente de variación y distancia intercuartílica en cada conjunto de datos.
- b) Calcula los percentiles 90, 75, 50, 25 y 10.
- c) Representa los datos mediante diagramas de cajas (*boxplots*) paralelos.
- d) Utiliza la información obtenida en los apartados anteriores para comparar las temperaturas de sublimación del Ir y el Rh.
- 1.4** Al comercializar un cargamento de carbón deben comunicarse algunas de sus propiedades con precisión, pues éstas determinan la calidad del cargamento. Una propiedad importante es el llamado *poder calorífico superior* (PCS), que se utiliza para describir el contenido energético del combustible. Para determinar el PCS (en MJ/kg) de un cargamento de carbón de Osterfeld se ha utilizado el protocolo ISO 1928. Éstos son los datos resultantes (ordenados de menor a mayor):

23.640 23.678 23.690 23.700 23.712 23.727 23.730 23.740
23.760 23.771 23.778 23.780 23.796 23.800 23.830 23.840
23.850 23.860 23.860 23.870 23.877 23.890 23.940

- a) Se denotan los datos por x_i , $i = 1, \dots, 23$. Si $\sum_{i=1}^{23} x_i = 547.119$ y $\sum_{i=1}^{23} x_i^2 = 13014.882627$, calcula la media, la desviación típica y el coeficiente de variación muestrales.
- b) Representa gráficamente los datos mediante un histograma definido sobre los intervalos $[23.60, 23.68]$, $(23.68, 23.76]$, $(23.76, 23.84]$, $(23.84, 23.92]$, $(23.92, 24.00]$.
- c) Determina los valores de la mediana y de los cuartiles primero y tercero. Representa gráficamente el diagrama de caja de los datos.

1.5 En la competición escolar de valoración organizada por el Royal Australian Chemical Institute (RACI) en 1997, cada equipo, compuesto por tres estudiantes de secundaria, tenía que determinar la concentración de una misma solución de ácido acético, utilizando sendas soluciones de sosa cáustica y de ácido clorhídrico de concentraciones conocidas. La concentración real de ácido acético era 0.1147 M. A continuación se reproducen, ordenados de menor a mayor, los resultados (en M) obtenidos por un estudiante de cada uno de 21 equipos que participaron en la competición. [FUENTE: Hibbert, D.B. y Gooding, J.J., *Data Analysis for Chemistry*, Oxford University Press, 2006]

0.1134 0.1138 0.1139 0.1141 0.1142 0.1143 0.1143
 0.1144 0.1144 0.1145 0.1146 0.1148 0.1150 0.1150
 0.1152 0.1153 0.1155 0.1158 0.1177 0.1219 0.1222

- a) Calcula la mediana y los cuartiles primero y tercero. Dibuja e interpreta el diagrama de caja de los datos.
- b) Dibuja e interpreta el histograma de los datos, agrupando las observaciones en los seis intervalos con extremos en

0.112, 0.114, 0.116, 0.118, 0.120, 0.122, 0.124.

1.6 Con el objetivo de incrementar las precipitaciones en zonas desérticas, se evalúa un método que consiste en bombardear («sembrar») una nube con yoduro de plata (AgI). El experimento se diseñó de la siguiente manera: de cada nube que se podía sembrar, se decidió al azar si se la bombardeaba o no. Una nube no bombardeada se llama nube control. Presentamos la cantidad de agua caída de 25 nubes sembradas y 26 nubes control:

NUBES SEMBRADAS								
129.6	31.4	2745.6	489.1	430.0	302.8	119.0	4.1	92.4
17.5	200.7	274.7	7.7	1656.0	978.0	198.6	703.4	1697.8
	334.1	118.3	255.0	115.3	242.5	32.7	40.6	
NUBES CONTROL								
26.1	26.3	87.0	95.0	372.4	0.01	17.3	24.4	11.5
321.2	68.5	81.5	47.3	28.6	830.1	345.5	1202.6	36.6
4.9	4.9	41.1	29.9	163.0	244.3	147.8	21.7	

Realiza un análisis descriptivo de los datos para estudiar si la siembra de las nubes provoca un aumento de precipitaciones.

1.7 Se realizan cinco determinaciones independientes de la fluorescencia de las soluciones ácidas de quinina a seis concentraciones distintas. Los resultados obtenidos son:

Concentración en ng/ml (X)	0	10	20	30	40	50
Intensidad de fluorescencia (Y)	4	22	44	60	75	104
	3	20	46	63	81	109
	4	21	45	60	79	107
	5	22	44	63	78	101
	4	21	44	63	77	105

Si $\sum_{i=1}^{30} y_i = 1574$; $\sum_{i=1}^{30} y_i^2 = 117200$; y $\sum_{i=1}^{30} x_i y_i = 56690$; determina la ecuación de la recta de regresión de la fluorescencia sobre la concentración, y el coeficiente de correlación. Representa los datos y la recta de regresión en un diagrama de dispersión. ¿Es bueno el ajuste?

1.8 En un estudio del complejo formado entre los iones del europio (Eu) y el ácido piridin-2.6-dicarboxílico (APDC), se determinaron los valores de absorbancia de disoluciones con distintas concentraciones de APDC-Eu. Los resultados fueron

APDC-Eu	0.2	0.4	0.6	0.8	1.0	1.2	1.4	1.6
Absorbancia	0.008	0.014	0.024	0.034	0.042	0.050	0.055	0.065
APDC-Eu	1.8	2.0	2.4	2.8	3.2	3.6	4.0	
Absorbancia	0.068	0.076	0.077	0.073	0.066	0.063	0.058	

Realiza un análisis de regresión lineal con estos datos y extrae conclusiones.

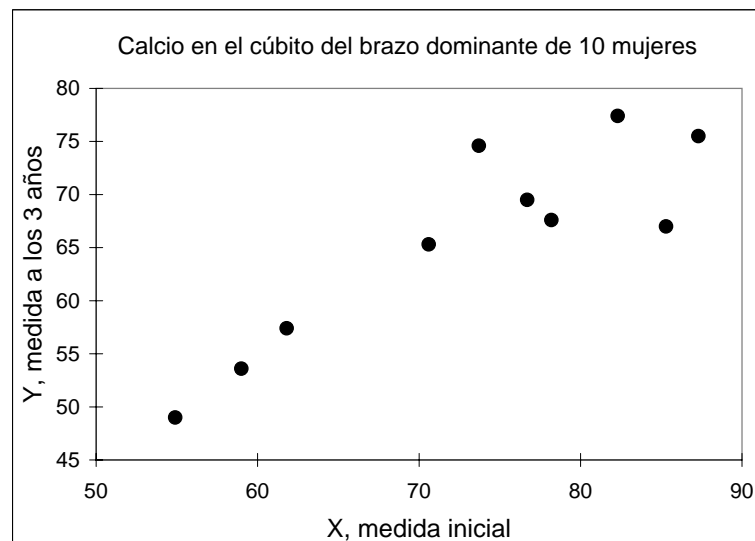
1.9 A fin de determinar la pérdida de calcio óseo con la edad, se realizan dos medidas, separadas tres años, del nivel de calcio en el cúbito del brazo dominante en 10 mujeres con edad avanzada.

Mujer	Inicial (X)	tras 3 años (Y)	Mujer	Inicial (X)	tras 3 años (Y)
1	87.3	75.5	6	78.2	67.6
2	59.0	53.6	7	73.7	74.6
3	76.7	69.5	8	61.8	57.4
4	70.6	65.3	9	85.3	67.0
5	54.9	49.0	10	82.3	77.4

Datos resumidos:

$$\sum_{i=1}^{10} x_i = 729.8 \quad \sum_{i=1}^{10} y_i = 656.9 \quad \sum_{i=1}^{10} x_i^2 = 54399.10 \quad \sum_{i=1}^{10} y_i^2 = 43977.99 \quad \sum_{i=1}^{10} x_i y_i = 48801.26$$

a) Calcula la recta de regresión de Y sobre X. Dibújala (aproximadamente) sobre el diagrama de dispersión de los datos que aparece a continuación.



b) Calcula e interpreta el coeficiente de correlación entre X e Y.

c) Para las observaciones de X, halla la mediana, los cuartiles, el rango total y el rango intercuartílico. Dibuja el diagrama de caja.