ESTADÍSTICA

ANALISIS DE DATOS HOJA 1

1. El número de accidentes mortales registrados en una población durante 10 días fueron: $3, 6, 4, 8, 1, 5, 3, 3, 7, 2$. Se pide:

   (a) Diagrama de barras.

   (b) Media, mediana y moda.

   (c) Varianza, desviación típica y coeficiente de variación.

   (d) Cuartiles primero y tercero: Rango intercuartílico.

2. Se ha obtenido el peso de 40 niños recién nacidos:

Clases (Kg.)	Fr. Absoluta	Clases (Kg.)	Fr. Absoluta
2.15 – 2.45	2	3.35 – 3.65	7
2.45 – 2.75	3	3.65 – 3.95	3
2.75 – 3.05	9	3.95 – 4.25	3
3.05 – 3.35	13		

   Se pide:

   (a) Histograma y poligono de frecuencias.

   (b) Completar la tabla de frecuencias.

   (c) Media, mediana y moda.

   (d) Varianza y desviación típica.

   (e) Rango intercuartílico.

   (f) Porcentaje de bebes cuyo peso está comprendido entre 3 y 4 Kg.

3. Para obtener información sobre el porcentaje de albumina en suero proteico de las personas, se analizaron 40 personas entre 2 y 40 años, con los siguientes resultados:

70.2	70.7	72.8	68.7	71.9	64.4	62.4	60.4	62.8	67
62.9	65.9	67.5	66.6	67.8	73.5	68.4	63.1	72.3	73.4
70.1	70	62.3	65	65.2	72.1	66.1	65.5	62	63.3
72.4	69.4	66.3	71.3	70.2	68.5	69.1	66.4	65.2	70.1

   Se pide:

   (a) Suponiendo que un error del 0.5% en el porcentaje de albumina medido no es importante, agrupar los datos en clases de amplitud 2 y tabular los resultados obtenidos.

   (b) Hallar el porcentaje de personas cuyo porcentaje de albumina está comprendido entre 62 y 72.

4. Sean $x_1, x_2, ...x_n$ e $y_1, y_2, ...y_n$ los datos de las dos características X e Y tomados sobre n unidades. Demostrar que el valor medio de la variable Z obtenida como suma de X e Y, es la suma de los valores medios de éstas.

5. Demostrar que si construimos una característica Z mezclando n_1 valores observados de X y n_2 de Y, el valor medio de Z es :

$$\overline{z} = \frac{n_1}{n_1 + n_2}\overline{x} + \frac{n_2}{n_1 + n_2}\overline{y}$$

6. Demostrar que si multiplicamos todos los valores observados de una variable por $k > 0$, la desviación típica y la media quedarán también multiplicados por k mientras el coeficiente de variación (cociente entre la desviación típica y la media en valor absoluto) quedara invariante.

7. Demostrar que al multiplicar los valores observados de una característica X por k_1 y los valores de Y por k_2, el coeficiente de correlación entre ambos no varía si k_1 y k_2 son del mismo signo.

8. Los siguientes datos corresponden al porcentaje de impurezas en el agua de una muestra de pozos de una cierta zona:

22	8	15	19	13	23	23	9	20	17	21	13	11
19	26	17	23	14	24	21	17	15	14	11	10	26
11	5	21	13	15	13	7	16	15	11	20	17	13
10												

   (a) Construir una tabla de frecuencias agrupando los datos en 5 clases de amplitud 5 y tomando como limite inferior de la primera clase 4.5. Dibujar el correspondiente histograma y poligono de frecuencias.
   (b) Obtener la media, mediana y cuartiles para los datos iniciales. Comparar estos valores con los obtenidos mediante los datos agrupados.

9. Dada la tabla adjunta de distribución de frecuencias, se pide:

x_i	n_i	N_i	F_i
3		6	
6		11	0.275
9	9		
12		27	
13	10		0.925
15			1.0

   (a) Completar la tabla
   (b) Hallar el tercer cuartil
   (c) Desviación media respecto a la mediana

(d) Coeficiente de variación de Pearson

10. Se han medido los pesos y alturas de seis personas, obteniéndose los datos siguientes:

Pesos	65	60	65	63	68	68
Alturas	1.70	1.50	1.68	1.70	1.75	1.80

Se quiere saber:

(a) ¿Qué medidas están más dispersas los pesos o las alturas?

(b) ¿Cuál es el coeficiente de variación de Pearson en cada caso?

(c) ¿Cuál es el coeficiente de variación media (respecto a la mediana)?.

11. Dada la siguiente distribución, ¿Qué centil corresponde a 222?. ¿Qué centil corresponde a 230?.

x_i	210-215	215-220	220-225	225-230	230-235
n_i	2	10	11	5	2