



1. Clasificar las siguientes variables como discretas o continuas. También clasificarlas según los niveles de medida (categórico nominal/ordinal, cuantitativo de intervalo o de razón).
 - a) Parte del día en términos de AM o PM.
 - b) Brillo medido con un sensor de luminosidad.
 - c) Brillo medido según el juicio de las personas.
 - d) Ángulos medidos en grados entre 0 y 360.
 - e) Medallas de bronce, plata y oro de las Olimpiadas.
 - f) Altura sobre el nivel del mar.
 - g) Número de pacientes de un hospital.
 - h) Número ISBN de los libros.
 - i) Grado de transparencia de un cuerpo según la clasificación: opaco, translúcido, transparente.
 - j) Rango militar.
 - k) Distancia al centro del campus.
 - l) Densidad de una sustancia en gramos por centímetro cúbico.
 - m) Número de abrigo en el guardarropa.

2. En una muestra aleatoria de familias, se contó el número de hijos (variable X). Los resultados están en la siguiente tabla:

2	1	1	3	1	2	5	1	2	3
4	2	3	2	1	4	2	3	2	1

Se pide:

- a) Tipo de variable X en cuanto a los niveles de medida y en cuanto al tipo de valores que toma.
- b) Tamaño de la muestra.
- c) Número de valores diferentes de la variable X .
- d) Recorrido.
- e) Tabla de frecuencias absolutas y relativas.
- f) Tabla de frecuencias acumuladas absolutas y relativas.
- g) Diagrama de barras con las frecuencias absolutas.
- h) Diagrama de tartas.

3. Las alturas de 20 alumnos del grado de Ingeniería Informática en la UFV son las recogidas en la siguiente tabla (valores expresados en cm):

180	177	178	181	175	184	180	181	179	180
184	186	167	183	177	186	181	177	173	181

Se pide:

- Estableciendo una serie de intervalos con los siguientes los límites: [165, 169], [170, 174], [175, 179], [180, 184], [185, 189], indicar cuáles son las marcas de clase y las fronteras.
 - Tabla de frecuencias absolutas y relativas con dichos intervalos.
 - Tabla de frecuencias acumuladas absolutas y relativas con dichos intervalos.
 - Dibujar el histograma de frecuencias relativas.
 - Dibujar el diagrama de barras de frecuencias relativas acumuladas.
4. A la pregunta de cuál es su comida favorita, las frecuencias relativas en % de las respuestas obtenidas son de la siguiente manera:

Huevos fritos	40%
Filete de ternera	25%
Macarrones	15%
Jamón serrano	10%
Gulas con gambas	5%
Otra	5%

Se pide:

- Tipo de variable que representa la comida elegida.
 - Si se entrevistó a 640 personas, calcular las frecuencias absolutas de cada respuesta.
 - Diagrama de barras con las frecuencias absolutas.
 - Diagrama de tartas con las frecuencias relativas.
5. Los salarios semanales de 65 empleados de la empresa X vienen expresados en la siguiente distribución de frecuencias. Se pide:
- El límite inferior de la 6ª clase.
 - El límite superior de la 4ª clase.
 - La marca de la 3ª clase.
 - Las fronteras de la 5ª clase.
 - Las anchuras del 5º intervalo.
 - La frecuencia absoluta de la 3ª clase.
 - La frecuencia relativa de la 3ª clase
 - El intervalo de clase con máxima frecuencia (intervalo de clase modal).

- i) El porcentaje de empleados que cobran menos de 280,00 € a la semana.
- j) El porcentaje de empleados que cobran menos de 300,00 € pero al menos 260,00 € a la semana.

Salario semanal (€)	Nº de empleados
250,00 – 259,99	8
260,00 – 269,99	10
270,00 – 279,99	16
280,00 – 289,99	14
290,00 – 299,99	10
300,00 – 309,99	5
310,00 – 319,99	2

6. Con los datos del ejercicio anterior, se pide:
- a) Calcular la tabla de frecuencias relativas.
 - b) Dibujar el histograma de frecuencias relativas.
 - c) Calcular las frecuencias relativas acumuladas.
 - d) Dibujar el polígono de frecuencias relativas acumuladas.
7. Tras la contratación de 5 nuevos empleados, repetir el histograma de salarios, donde ahora la tabla queda modificada con intervalos de amplitud no constante.

Salario semanal (€)	Nº de empleados
250,00 – 259,99	8
260,00 – 269,99	10
270,00 – 279,99	16
280,00 – 289,99	15
290,00 – 299,99	10
300,00 – 319,99	8
320,00 – 379,99	3

8. Se lanzan 5 monedas 1000 veces. El número de lanzamientos en los que han salido 0, 1, 2, 3, 4 y 5 caras se muestra en la tabla. Se pide:
- a) Representar los datos mediante un diagrama de barras.
 - b) Construir una tabla con los % de tiradas que han dado un nº de caras menor que 0, 1, 2, 3, 4, 5 o 6. ¿Cómo se llama esto que acabamos de calcular?
 - c) Representar los datos del apartado anterior mediante una curva formada por segmentos (función plot).

Nº de caras	Nº de tiradas (frecuencia absoluta)
0	38
1	144
2	342
3	287
4	164
5	25

9. Las notas de un estudiante en seis exámenes fueron las siguientes:

8,4	9,1	7,2	6,8	8,7	7,8
-----	-----	-----	-----	-----	-----

Se pide calcular:

- Media aritmética.
- Media aritmética ponderada, teniendo en cuenta que las dos primeras notas corresponden cada una a un 30% de la nota final, las dos siguientes al 15% cada una, y las dos últimas al 5% cada una.
- Mediana.

10. Dados los siguientes números:

5	3	6	5	4	5	2	8	6	5
4	8	3	4	5	4	8	2	5	4

Se pide calcular de dos maneras, con la fórmula habitual o con la de las frecuencias absolutas:

- Media aritmética.
- Media geométrica.
- Media armónica.
- Media cuadrática.

11. Con los datos del ejercicio anterior, calcular también la mediana (hacerlo de dos maneras: con los datos en bruto y con los datos agrupados) y la moda.

12. Los salarios anuales de 8 empleados de una empresa son: 15000 €, 16000 €, 16500 €, 17000 €, 15600 €, 16300 €, 16000 € y 40000 €. Se pide:

- Calcular la media aritmética.
- Calcular la media aritmética recortando de manera simétrica un 20% de los datos extremos.
- Calcula mediana.

Comentar todos los resultados.

13. Sean otra vez las alturas de los 20 alumnos del ejercicio 3. Utilizando la distribución de frecuencias de aquel ejercicio (es decir, considerando la altura como una variable continua y agrupando los datos), se pide calcular:

- La media aritmética.
- La mediana.
- La moda.
- Repetir los cálculos anteriores sin agrupar los datos. Comparar y discutir los resultados obtenidos.

14. En una fábrica de tuercas se toma una muestra de 40 unidades para un estudio de calidad. Los pesos de las tuercas (con precisión de 0,1 g) vienen indicados en la siguiente tabla:

Peso (gramos)	Frecuencia absoluta
11,8 – 12,6	3
12,7 – 13,5	5
13,6 – 14,4	9
14,5 – 15,3	12
15,4 – 16,2	5
16,3 – 17,1	4
17,2 – 18,0	2

Se pide calcular:

- Identificar las marca de clase, los límites de los intervalos y las fronteras.
 - La media aritmética.
 - La mediana.
 - La moda.
15. Con los datos del ejercicio anterior, dibujar un histograma de frecuencias absolutas. Para ello crear un vector f con el valor de las frecuencias absolutas. Crear a su vez un vector m con las marcas de clase de cada intervalo. Establecer con la función `names` que los nombres de las componentes de f son precisamente los valores del vector m . Esto se hace así: `names(f) <- m`.

Nótese que para datos ya agrupados, como es el caso, el histograma se dibuja en R con `barplot`, con el parámetro `space=FALSE` para que las barras salgan sin espacios intermedios. Si los datos no estuvieran agrupados, el histograma lo dibujaríamos con la función `hist`.

Identificar el valor de la abcisa que divide al histograma en dos mitades exactamente iguales.

16. Repetir el ejercicio anterior, pero esta vez dibujando un polígono de frecuencias relativas acumuladas (función `plot`). Calcular el valor de la abcisa que corresponde con la ordenada del 50%.
17. Utilizando la función `rchisq`, generar un total de 1000 números aleatorios según la distribución χ^2 con 10 grados de libertad. Consultar la ayuda de R para los detalles. Almacenar los valores obtenidos en un vector para su uso posterior. Se pide:
- Dibujar un histograma de valores. Probar con intervalos de anchura 2.
 - Calcular la media de los datos (sin agrupar).
 - Calcular la mediana de los datos (sin agrupar).

- d) Calcular la moda de los datos (agrupados según los intervalos del histograma).
- e) Comprobar la validez de la aproximación: $(\text{media} - \text{moda}) \approx 3 * (\text{media} - \text{mediana})$.
- f) Dibujar los números obtenidos mediante un diagrama de cajas (`boxplot`). Analizar el gráfico.

18. Con los datos del ejercicio 3, calcular:

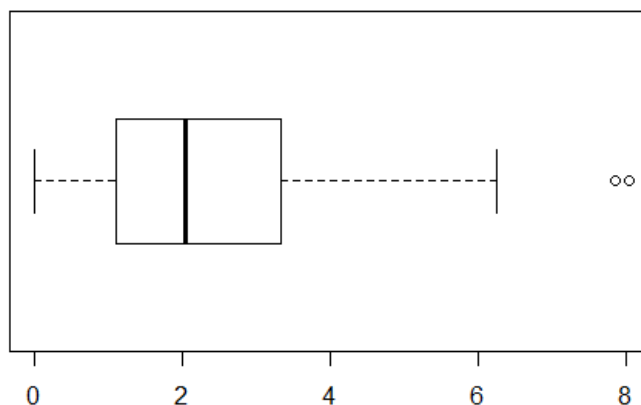
- a) El primer cuartil.
- b) El tercer cuartil.
- c) El primer decil.
- d) El octavo decil.
- e) El percentil del 56%.

19. Con los datos del ejercicio 3, se pide:

- a) Dibujar un diagrama de cajas (`boxplot`). Analizar el gráfico.
- b) El recorrido.
- c) El recorrido intercuartil.
- d) El recorrido semiintercuartil.
- e) La desviación típica sesgada y no sesgada.
- f) Coeficiente de variación de Pearson.

20. Repetir el cálculo de la desviación típica sesgada de los datos del ejercicio 3 utilizando la fórmula de la diferencia de la media de los cuadrados y el cuadrado de la media.

21. Representamos gráficamente una variable de la siguiente manera:



Razonar si es esperable que la media aritmética de la variable sea mayor, menor o igual que la mediana. ¿Y la moda respecto de la mediana?

22. Un grupo de niños de Primaria tienen el siguiente Coeficiente Intelectual.

Marca	70	74	78	82	86	90	94	98	102	106	110	114	118	122	126
Frec. Abs.	4	9	16	28	45	66	85	72	54	38	27	18	11	5	2

Se pide calcular:

- a) La media aritmética.
- b) La desviación típica.

23. Con los datos del ejercicio anterior, se pide:

- a) Convertir la variable a unidades estándar. Añadir dos marcas de clase adicionales, 66 y 130, con frecuencia 0.
- b) Calcular las frecuencias relativas de cada intervalo.
- c) Representar gráficamente como un polígono de frecuencias (función p_{lot}) el valor de la frecuencia relativa respecto de la variable en unidades estándar.

24. En un examen de Estadística la nota media de 30 alumnos fue de 7,8, con una desviación típica de 0,8. Sin embargo en un examen de Complejidad Computacional, la media fue de 7,3, con una desviación típica de 7,6. ¿En qué asignatura fue mayor la dispersión absoluta? ¿Y la dispersión relativa? ¿En cuál de las dos asignaturas se obtuvieron mejores resultados? Justificar la respuesta.

25. Con los datos de las tuercas del ejercicio 14, calcular el 1er y 2º coeficiente de asimetría de Pearson y compararlo con el coeficiente de Fisher. ¿Qué tipo de asimetría tiene?

26. Recordemos los datos de salarios del ejercicio 5. Se pide calcular:

- a) La media aritmética.
- b) La mediana.
- c) La moda.
- d) Los coeficientes primero y segundo de asimetría de Pearson.

27. Para las alturas del ejercicio 3, se pide calcular:

- a) El coeficiente de asimetría de Fisher.
- b) La curtosis.

28. Para el ejercicio 21 de los coeficientes intelectuales, se pide calcular:

- a) El coeficiente de asimetría de Fisher.
- b) La curtosis.

29. En 1965 A. J. Lea recogió datos sobre la temperatura anual media en varias ciudades de Gran Bretaña, Noruega y Suecia y la tasa de mortalidad en un tipo de cáncer de pecho de mujeres. Los datos que obtuvo fueron los siguientes:

Temperatura anual media (°F)	Índice de mortalidad
51.3	102.5
49.9	104.5
50.0	100.4
49.2	95.9
48.5	87.0
47.8	95.0
47.3	88.6
45.1	89.2
46.3	78.9
42.1	84.6
44.2	81.7
43.5	72.2
42.3	65.1
40.2	68.1
31.8	67.3
34.0	52.5

Se pide:

- Calcular el índice de correlación lineal entre ambas variables.
- Calcular la covarianza entre ambas variables.
- Calcular la varianza de la temperatura.
- El valor medio de la temperatura y el valor medio del índice de mortalidad.
- Representar gráficamente los datos en un diagrama de dispersión, colocando la temperatura en el eje X y la mortalidad en el eje Y. Usar el color rojo.
- Calcular la recta de mínimos cuadrados que mejor se ajusta a los datos.
- Superponer dicha recta a la gráfica del apartado a. Usar el color azul.

30. Las notas de 14 alumnos de Algoritmos y de Interacción Persona-Ordenador son las siguientes:

Alg.	3,3	1	6	5	8	2	5,4	7	7,5	6,5	9	10	5,1	5,8
IPO	2,4	3,5	5,5	5	9	1	4	7,5	6	7	10	8	4,3	3

Se pide representar los datos en un diagrama de dispersión y calcular el mejor ajuste lineal a los datos.