



Dpto. Matemática Aplicada y Estadística

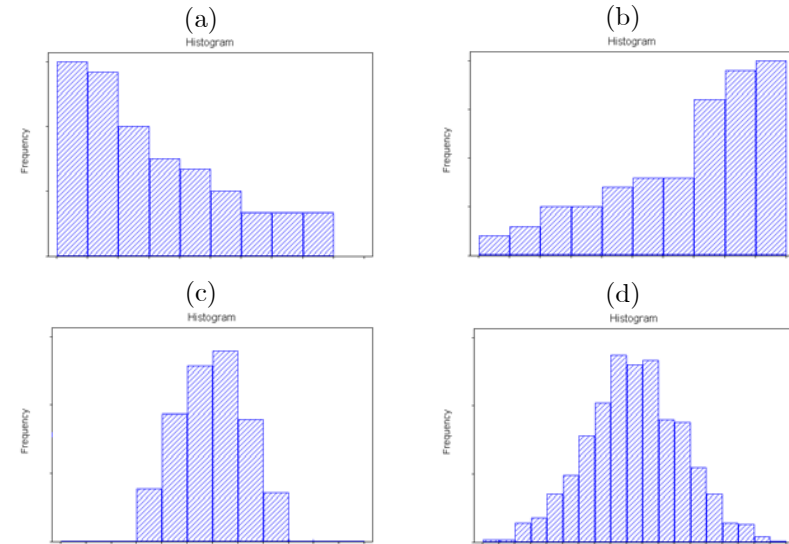
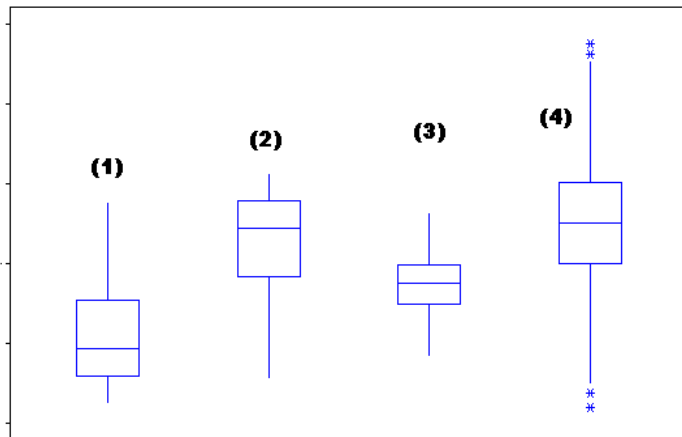
**Grado en IIAA y Grado en IHJ.**  
**Asignatura: Estadística Aplicada. Curso 2011-2012**  
**HOJA 1A: Estadística Descriptiva Unidimensional.**

1. Un determinado estudio intenta medir la corrosión de un determinado material a partir de ultrasonidos. Tras analizar 20 datos se obtuvieron las siguientes medidas descriptivas:

$$\bar{x}=90 \quad \bar{x^2}=10123 \quad \min(x_i)=40 \quad \max(x_i)=125 \quad Q_1=72.5 \quad Me=90 \quad Q_3=96.5$$

A partir de estos datos, se pide:

- (a) Obtener el diagrama de cajas y bigotes asociado e indicar si podemos admitir que no existen datos atípicos. ¿Entre qué valores se encontrarían las observaciones no atípicas?
  - (b) ¿Qué medidas de centralización y dispersión estimas más adecuadas para resumir la información contenida en el conjunto de datos? Razona tu respuesta.
2. Identificar qué diagrama de caja y bigotes corresponde a cada uno de los histogramas siguientes y comentar las características más relevantes de cada conjunto de datos. Razonar cuál es la medida de centralización y de dispersión más adecuada para la distribución correspondiente.



3. Para analizar *el tiempo de respuesta de una base de datos de consulta* se tomaron 20 datos correspondientes a 4 semanas consecutivas obteniéndose los siguientes resultados:

	Lunes	Martes	Miércoles	Jueves	Viernes
Semana 1	4.32	7.14	9.21	10.46	12.1
Semana 2	5.2	7.5	9.34	10.53	16.6
Semana 3	6.39	8.4	9.51	11.3	17.5
Semana 4	7.06	8.97	9.71	11.5	21.25

- (a) Calcular la media, mediana, cuartiles y desviación típica correspondientes a estos datos.
- (b) Construir de manera detallada el diagrama de caja y bigotes del *tiempo de respuesta de la base de datos* indicando en cada línea del diagrama su valor numérico ¿Existen datos atípicos? ¿Qué representa un dato atípico en una muestra? Comentar las características más relevantes del anterior diagrama.
- (c) Agrupar los datos en intervalos de la misma amplitud y representar gráficamente la distribución de frecuencias. Describir las características más relevantes de dicha gráfica. ¿Qué medidas de posición central y de dispersión son más adecuadas para resumir los datos? Razona tu respuesta. Este gráfico corrobora los comentarios dados en el apartado anterior.

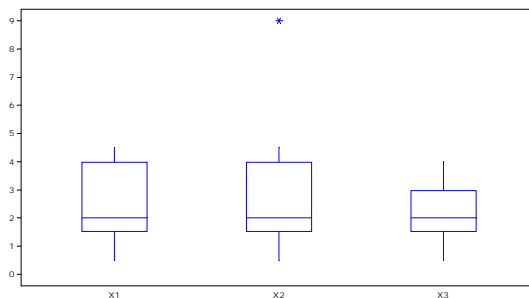
4. En la tabla adjunta se presentan los datos relativos al consumo de agua de 45 granjas dedicadas a la producción de trigo de una determinada región durante un año hidrológico:

Consumo (en Dam <sup>3</sup> )	Nº de granjas
$1.2 \leq X < 1.8$	5
$1.8 \leq X < 2.4$	10
$2.4 \leq X < 3$	8
$3 \leq X < 3.6$	10
$3.6 \leq X < 4.2$	7
$4.2 \leq X < 4.8$	3
$4.8 \leq X < 5.4$	2

- (a) Representar gráficamente la variable en un histograma. Describir las características más relevantes de dicha gráfica. ¿Qué medidas de posición central y de dispersión son más adecuadas para resumir los datos?
- (b) Calcular el consumo medio de agua en las granjas de la muestra. ¿En qué intervalo se sitúa la mediana? ¿Cuál es el intervalo modal?
- (c) ¿Qué porcentaje de granjas tienen un consumo de agua inferior a 2.4 Dam<sup>3</sup>? ¿Qué porcentaje de granjas tienen un consumo mayor o igual a 3 Dam<sup>3</sup>? ¿y a 3.6 Dam<sup>3</sup>/100 Km? ¿Qué porcentaje de granjas tienen un consumo de agua comprendido en el intervalo  $1.8 \leq X < 4.2$  Dam<sup>3</sup>?, ¿y en el intervalo  $1.8 \leq X < 2.4$  Dam<sup>3</sup>?
- (d) ¿En cuántas granjas el consumo de agua superó los 3.2 Dam<sup>3</sup> de agua?
5. La siguiente tabla muestra los resúmenes descriptivos que proporciona un programa de ordenador para el crecimiento durante 10 años de una determinada especie:

Media	2.5	Mínimo	0.5
Mediana	2	Máximo	4.5
Desviación típica	1.35	Rango Intercuartílico	2.5

Indicar cuál de estos diagramas de caja y bigotes puede representar a este conjunto de datos. Dar razones de por qué los otros no pueden representar a los datos resumidos en la tabla.



6. Se han tomado muestras de dos empresas con un número total de empleados similares y se han observado los salarios percibidos. La información obtenida se recoge en la siguiente tabla:

Empresa 1		Empresa 2	
Salarios	Nº empleados	Salarios	Nº empleados
10-20	24	5-15	19
20-30	36	15-25	42
30-40	20	25-55	35
40-50	20	55-75	20
50-100	50	75-95	13

- (a) Calcular el salario medio y la desviación típica de los salarios percibidos por los empleados en cada empresa. ¿En cuál de las dos empresas los salarios presenta menor dispersión relativa?
- (b) ¿Cuál es el salario percibido por el mayor número de empleados de la primera empresa? Justifica tu respuesta en función de la medida de centralización utilizada.
- (c) Si en la segunda empresa clasificamos a un empleado en el grupo en que se encuentra el 50% de los menos remunerados, ¿cuál será el máximo salario que podrá percibir?
7. Los valores observados de una variable Y son:

19.6 7.9 8.1 4.2 2.9 3.7 1.3 2.2

- (a) ¿Entre qué valores podemos decir que se encuentran los datos no atípicos? ¿Existen datos atípicos?
- (b) Calcular la media y la mediana. Compara sus valores, ¿por qué crees que son tan distintas?