



Examen de prácticas con Statistix

Nombre:

Problema 1

En el fichero **azufre.txt** se encuentran los datos correspondientes a un estudio sobre la *contaminación de un determinado suelo en función de la concentración de azufre del mismo*. Después de importar los datos desde el fichero **azufre.txt** (la primera línea del fichero contiene los nombres de las variables: CONTAMINA y AZUFRE), que se encuentra en la dirección habitual, responder a las cuestiones siguientes:

1. Realizar un diagrama de caja y bigotes de la variable CONTAMINACIÓN. ¿Qué representa cada línea de este diagrama? Comenta las características más relevantes del diagrama. **(2 ptos)**
2. Omitir los datos atípicos correspondientes a la contaminación, si los hubiera, y trabajar en adelante sin dichos datos. Calcular las medias y desviaciones típicas de las dos variables. ¿Afecta a estas medidas la presencia de valores atípicos? Razona tu respuesta. **(2 ptos)**
3. Queremos buscar un modelo que explique la *evolución de la contaminación en función del nivel de azufre*. Realiza una nube de puntos de ambas características. ¿Qué tipo de modelo parece más adecuado? Proponer un modelo concreto a partir de los datos y comentad la bondad del ajuste. **(2 ptos)**
4. A partir del modelo estimado, ¿cuál sería la contaminación de un suelo que presenta un nivel de azufre igual a 1.75? ¿y si el nivel de azufre es de 8.5? **(1 pto)**

Problema 2

En el fichero **espesor.txt** se encuentran los datos que se obtuvieron de una muestra de tamaño 1386 de la distribución del espesor de la madera en los postes telefónicos (W.A. Shewhart. *Economic Control of Quality of Manufactured Products*. Macmillan and Co, N.Y. 1932, pag.66). Importar los datos del fichero **espesor.txt**, que se encuentra en la dirección habitual.

1. ¿Cuál es el porcentaje de postes en la muestra cuyo espesor está comprendido entre 1.7 y 4.7 pulgadas? (**1 pto**)

¿Y entre 1.5 pulg. y 5.1 pulg.?(**1 pto**)

2. ¿Qué modelo de distribución continua te parece más adecuado para describir el comportamiento de la variable aleatoria X = "Espesor de la madera de un poste telefónico seleccionado al azar"? Razona tu respuesta. (**1 pto**)

3. Suponiendo que la media de X es 2.9 y su desviación típica es 0.8 (pulgadas), calcular las siguientes probabilidades: (**2 ptos**)

(a) $Pr(1.7 \leq X < 4.7) =$

(b) $Pr(1.5 \leq X < 5.1) =$

4. Proporcionar un intervalo de confianza al 98% para la media de la variable aleatoria X . ¿Qué interpretación tiene el intervalo obtenido? (**2 ptos**)

5. ¿Cuánto vale el error de estimación para el intervalo obtenido?. Si la desviación típica de la variable X fuese más pequeña, ¿aumentaría dicho error de estimación? Razona tu respuesta. (**1 pto**)