



Dpto. Matemática Aplicada y Estadística

Ingeniero Técnico de Telecomunicación (Telemática)
Estadística (septiembre 2005)

Prácticas con Statistix

Nombre:

1. Un determinado estudio medioambiental informa sobre los resultados obtenidos sobre la aparición de cloro y el porcentaje de suelo urbano cercano a lo largo de un acuífero. Los resultados obtenidos fueron almacenados en el fichero **contamina.txt**
 - (a) Dibuja un diagrama de cajas-bigotes para ambas características y determina si existe algún valor atípico. (si es así indica su valor). Indica los valores exactos de cada una de las líneas que componen cada diagrama. **(1 punto)**

 - (b) Realizar el diagrama de dispersión entre ambas características y determina un modelo que permita obtener el contenido de cloro en el agua en función del % de suelo urbano cercano. Calcular el valor de R^2 y comentar la bondad del ajuste. **(1 puntos)**

 - (c) A partir del estudio realizado, ¿qué influencia tiene el suelo urbano en la concentración de cloro obtenida? **(0.5 punto)**

 - (d) ¿Podrías determinar cuál sería la cantidad de cloro si el % de suelo urbano fuese de 1.5?, ¿y si fuese de 2.5? **(0.5 puntos)**

 - (e) En una determinada zona del acuífero se ha detectado una concentración de 21.7. ¿podemos determinar el % de suelo urbano cercano?. Razona tu respuesta. **(0.5 puntos)**

2. Se ha realizado un estudio de salinidad de ciertas disoluciones, midiéndose en ellas la concentración de ClNa en mg/l. Los resultados obtenidos se almacenaron en el fichero **clna.dat**. Una vez recuperado el fichero, que se encuentra en su ruta habitual, responder a las siguientes cuestiones

(a) Realiza el histograma correspondiente tomando 7 clases y límite inferior 28 y límite superior 224 (**copiarlo en la parte de atrás**). Indicar la frecuencia de cada clase: **(1 punto)**

Clase	Frecuencia

(b) ¿Qué modelo de distribución podemos asociar a la variable **contenido en ClNa** en función de los datos observados?. Razona tu respuesta. **(1 punto)**

(c) Si suponemos que la variable " **$X = \ln(\text{ClNa})$** " sigue una distribución normal, determinar sus parámetros, obtener $\Pr(X < 5)$ así como la proporción de datos en la muestra cuyo valor es inferior a 5. **(0.5 puntos)**

(d) Obtener un intervalo de confianza al 97% para el promedio de la variable **X**. ¿Cómo interpretas el resultado obtenido? **(1 punto)**

(e) ¿Podemos afirmar que el promedio de la variable **$X = \ln(\text{ClNa})$** es superior a 4.75?. Responder a la pregunta anterior planteando la prueba estadística correspondiente y comentando el p-valor obtenido. **(1 punto)**

3. Una determinada empresa está interesada en estudiar las fluctuaciones que se producen en una determinada serie temporal con el fin de intentar predecir el valor futuro en función de valores anteriores (los datos se encuentran en el fichero **serie.sx**). Si suponemos que el proceso tiene varianza constante, determinar cuántas observaciones anteriores deberían tenerse en consideración en un instante dado, para poder predecir el siguiente valor. Explica el procedimiento que has utilizado **(1 pto)** y propón un modelo explicativo **(1 pto)**