
La recta de regresión

Práctica

1. El proceso de lectura desde una unidad de almacenamiento a la memoria 'buffer' del ordenador puede verse como un proceso en dos etapas. En una primera etapa, la unidad de lectura busca el sector en el que se encuentra la información. El tiempo empleado en esta etapa se denominará 'retraso latente' (*latency delay*).

Una vez encontrada la información, ésta se transmite desde la unidad de almacenamiento a la memoria invirtiendo un tiempo proporcional al tamaño de la información. Llamaremos al tiempo empleado en esta etapa 'tiempo de transmisión'. Cada vez que repetimos la operación de lectura de un fichero de M kilobytes, los tiempos totales invertidos no serán los mismos, pues dependerán de si dicha información se encuentra en sectores contiguos (sólo cometemos un retraso latente), o por el contrario ha de ir saltando de un lugar a otro de la unidad de almacenamiento (consumimos varios retrasos latentes).

Para comprobar el estado de una unidad de almacenamiento se ejecuta un protocolo en el que se leen ficheros de varios tamaños (en kilobytes) y se cronometra el tiempo empleado en la lectura (milisegundos). Periódicamente se realiza este protocolo y se analizan los datos obtenidos. De esta manera se hace un seguimiento del estado de la unidad de almacenamiento y se decide si es conveniente realizar alguna actividad de mantenimiento. Los datos del fichero **Retrasolatente.sf3** contienen los resultados de uno de estos análisis (extraídos del libro *Probability and Statistics for Computer Science*, Johnson, ed. Wiley). Se pide:

- a. Analiza gráficamente la relación entre el tamaño del fichero que se lee (variable **Kilobytes**) y el tiempo empleado en su lectura (variable **Milisegundos**)
 - b. Calcula la recta de regresión que sirva para predecir el tiempo empleado en leer un fichero en función de sus tamaño en kilobytes
 - c. ¿Cuánto se tardará, aproximadamente, en leer un fichero de 350 kilobytes?
 - d. Utilizando la regresión que prediga el tamaño del fichero en función del tiempo empleado, ¿qué tamaño de fichero prevés que podremos leer en 200 milisegundos?
2. El fichero **AlumnosIndustriales.sf3** tiene varias variables de un conjunto de 95 estudiantes de primer curso de Ingeniería Industrial. Se quiere saber
 - a. ¿Es razonable intentar predecir la altura de una persona de esta muestra a partir de su peso?
 - b. Utilizando la recta de regresión adecuada: ¿qué altura se prevé para una persona que pese 70 kilos?
 - c. Utilizando la recta de regresión adecuada: ¿qué peso se prevé para una persona que mida 175 cm?
 - d. Una persona que mida 185 cm y pese 72 kilos ¿puede considerarse delgada, gruesa, normal?
 - e. Distingamos ahora por sexo (Select:sexo=1 ó sexo=0). ¿Qué altura se prevé para un chico que pese 65 kilos? ¿Y para una chica de ese mismo peso?

Práctica a entregar: Elige **uno** de los dos ejercicios propuestos.

Extensión máxima: 2 hojas por una cara.