



Universidad Politécnica de Cartagena
Dpto. Matemática Aplicada y Estadística
Acceso a Ingeniería en Edificación
Estadística

Examen Febrero 2010 (II).

IMPORTANTE : Duración máxima: 1:30 hora. Se valorará el razonamiento de los argumentos utilizados en la resolución de los problemas.

Problema 1

En el fichero de datos obrasnuevas.txt, se puede encontrar los datos correspondientes a las viviendas a construir, según visados de arquitectos técnicos, desde enero de 2001 hasta octubre de 2009. La columna “Superficie” corresponde a la superficie media de las viviendas nuevas del tipo considerado para el mes y el año asociados.

- a). Importar los datos en un conjunto de datos (data frame) que llamarán `obrasnuevas`. ¿Cuántos filas y cuántas columnas contiene el conjunto de datos.? **(0.2 pts)**
- b). Realizar un diagrama de caja-bigotes de la variable superficie en función del tipo de viviendas. ¿Para qué tipo de superficie, presenta mayor dispersión los datos? **(0.3 pts)**
- c). Explicar qué representan los distintos componentes (líneas y/o puntos) de un diagrama caja-bigotes. **(0.4 pts)**
- d). Identificar dos atípicos de la variable superficie para el tipo ”otros”. ¿A qué mes y año corresponden? **(0.4 pts)**
- e). Introducir una nueva variable en el conjunto de datos que se llame `superficie_total` que contenga la superficie total de las viviendas visadas para un tipo, un año y un mes determinado. Se indicará el procedimiento seguido. **(0.3 pts)**
- f). Obtener el promedio de la superficie total de las vivienda visadas, desglosando el indicador por año y también según el tipo. ¿En qué tipo se ha sufrido el mayor descenso de la superficie construida? **(0.4 pts)**

Problema 2

Con el fin de determinar si existe relación entre la cantidad de polímeros de látex incluida durante el proceso de mezclado de cemento Portland y su resistencia adhesiva a tensión, una empresa encargada de realizar certificaciones de obras toma una muestra de tamaño 10, obteniendo los siguientes resultados:

<i>Polímeros/Cemento(mg/kg)</i>	13.5	11.0	13.0	11.2	12.0	13.2	12.0	13.5	11.2	13.0
<i>Resistencia(kgf/cm²)</i>	17.5	16.6	17.2	16.6	17.0	17.3	16.9	17.3	16.8	17.1

- a). Calcular la media y varianza asociada a cada una de las variables. **(0.3 pts)**
- b). Realizar un ajuste lineal por mínimos cuadrados de la resistencia respecto a la cantidad de polímeros añadida en la mezcla. **(0.4 pts)**
- c). Comentar sobre la bondad del ajuste. **(0.3 pts)**
- d). Calcular el coeficiente de correlación entre ambas variables. **(0.3 pts)**
- e). Deducir, suponiendo el modelo lineal válido, el incremento promedio que se espera en la resistencia promedio si se aumenta la cantidad de polímeros en 5 unidades. **(0.4 pts)**
- f). Deducir, suponiendo que la relación proporcionada por las rectas de regresión es válida, el valor estimado para la resistencia si la cantidad de polímero agregado es de 11.5 mg/kg. **(0.3 pts)**

Problema 3

Se realizan mediciones de la resistencia de unos bloques de hormigón. Denotamos por X la variable "Valor obtenido en una medición realizada al azar".

Suponemos a partir de ahora que se puede modelizar la distribución de X por una distribución Normal.

Se realizan 8 mediciones de la resistencia de unos determinados bloques, obteniéndose los siguientes datos: 5.3, 4.2, 7.2, 6.3, 5.5, 6.5, 4.8, 5.1

- a). Construir un intervalo de confianza al 98% para el centro de la distribución de X . **(0.3 pts)**
- b). ¿Qué interpretación tiene un intervalo de confianza? En particular, explique que quiere decir "un nivel de confianza de 95%" **(0.4 pts)**
- c). Si en la estimación anterior queremos cometer un error inferior a una unidad, determinar el tamaño de la muestra mínimo necesario para garantizar este objetivo. (Para este cálculo, se podrá aproximar el valor de la varianza poblacional σ , por 4). **(0.3 pts)**