Estadística I
Grado en Matemáticas, UAM, 2017-2018
Trabajo computacional

Instrucciones

- Para realizar este trabajo debéis organizaros en grupos (como máximo de tres personas). Se comunicará a `pablo.fernandez@uam.es` la composición de los grupos no más tarde del **23 de diciembre de 2017**.
- **Entregables**. Se deberán enviar, electrónicamente,
  - las hojas de cálculo creadas en el trabajo;
  - y una (breve) memoria explicativa (en pdf).
- La **fecha límite de entrega** de los trabajos es el **28 de diciembre de 2018**. Aunque, por supuesto, se puede enviar en cualquier momento anterior.
- Está colgado en la red un pequeño manual de excel, por si fuera de utilidad.
- Se sugiere crear una hoja de cálculo para cada ejercicio. Parte del examen es que estas hojas de cálculo estén bien organizadas y sea sencillo seguir la información contenida en ellas.
- En la memoria se recogerán los *resultados* obtenidos en cada ejercicio, gráficas ilustrativas, y los comentarios y conclusiones que consideréis oportunos. Se valorará la organización y la (buena) presentación y redacción de la memoria.

Ejercicio 1. En la hoja de cálculo adjunta encontrarás las series históricas de 10 años de cotizaciones diarias de las acciones de Iberdrola y del índice IBEX 35. Junto a ellas se han calculado los correspondientes rendimientos (variaciones porcentuales) diarias.

En el ejercicio solo trabajaremos con estas series de rendimientos diarios.

1a. Cuestiones de estadística descriptiva y regresión lineal

- Toma la serie de datos del IBEX y
  - calcula la media y la cuasidesviación típica muestral, además de los rendimientos mínimo y máximo en el periodo de estudio;
  - construye un histograma tomando, por ejemplo, 60 clases (equiespaciadas) entre el mínimo y el máximo rendimiento. (Como alternativa, y para una mejor visualización, puedes decidir que la primera y la última clases sean algo más grandes que las demás).

- Toma ahora la pareja de series de rendimientos de Iberdrola e IBEX, calcula su coeficiente de correlación muestral y determina la recta de regresión, indicando la bondad del ajuste. Acompáñalo con un gráfico de la nube de puntos (gráfico de dispersión en excel).

1b. Intervalos de confianza y contraste de hipótesis

Suponemos que tanto los rendimientos de Iberdrola como los del IBEX siguen un modelo normal.

- Calcula un intervalo de confianza con $\alpha = 5\,\%$ para la media de los rendimientos de Iberdrola.
- Contrasta la hipótesis $H_0 : \mu_{\text{IBEX}} = 0\,\%$ con nivel de significación $\alpha = 1\,\%$.
- Creemos que la variabilidad de los rendimientos del IBEX ha de ser significativamente menor que la de los rendimientos de Iberdrola. Plantea la hipótesis adecuada y contrástala con nivel de significación $\alpha = 5\,\%$.

Ejercicio 2.

La variable X sigue una $\mathcal{N}(\mu, \sigma^2)$. (Recuerda que entonces $X = \mu + \sigma\, Z$, donde $Z \sim \mathcal{N}(0, 1)$).

Sabemos (teorema de Fisher-Cochran) que, para muestras $(X_1, \ldots, X_n)$, las variables $\overline{X}$ y S^2 son independientes. Además, $\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n)$ y $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$.

Este ejercicio plantea el análisis numérico (vía simulación) de estas cuestiones.

2a. Generación de muestras del par $(\overline{X}, S^2)$

Tomamos $\mu = 1$, $\sigma^2 = 2$ y $n = 20$. Seguimos el siguiente esquema:

- sorteamos una muestra $(x_1, \ldots, x_{20})$ de tamaño 20 de normales con esos parámetros;
- hallamos el valor de la media $\overline{x}$ y la cuasivarianza s^2 muestrales;
- repetimos el experimento un buen número de veces (por ejemplo, 3000) y vamos anotando los sucesivos valores de $\overline{x}$ y s^2;
- finalmente, copiamos en valores estas 3000 parejas de datos.

2b. Análisis de las muestras obtenidas

Usando la muestra obtenida en el apartado anterior,

- estima la media y la varianza de $\overline{X}$;
- calcula la proporción de muestras en las que $1 \leq \overline{X} \leq 1.2$ y $1.4 \leq S^2 \leq 2.0$ (simultáneamente), y compáralo con el *producto* de las proporciones de cada suceso por separado. (Puedes repetir este ejercicio para otros intervalos para $\overline{X}$ y S^2).

(Como ejercicio adicional y extra, puedes estimar por máxima verosimilitud el número de grados de libertad de la χ^2 con la que se distribuye S^2. La instrucción `= distr.chicuad(x;n;falso)` devuelve el valor de la función de densidad de una χ^2 con n grados de libertad en el punto x).