Estadística I
Grado en Matemáticas, UAM, 2018-2019
Trabajo computacional 1. Simulación

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 **9 de noviembre de 2018**.
- **Entregables**. Se deberán enviar, electrónicamente,
  - las hojas de cálculo creadas en el trabajo;
  - y una (breve) memoria explicativa (en pdf).
  - Convendrá que tanto las hojas de cálculo como el pdf tengan nombres identificativos, del tipo `ej1-apellidos` o similares.
- La **fecha límite de entrega** de los trabajos es el **20 de noviembre 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.
- Parte del examen es que estas hojas de cálculo estén bien organizadas y sea sencillo seguir la información contenida en ellas.
- En la (breve) 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.

La variable X viene definida por la siguiente función de densidad:

$$f(x) = \begin{cases} 4x & \text{si } x \in [0, 1/2], \\ 4(1-x) & \text{si } x \in (1/2, 1], \\ 0 & \text{en el resto de los casos.} \end{cases}$$

Observa que X es una variable triangular (simétrica) en el intervalo $[0, 1]$.

El objetivo del ejercicio es estimar, por simulación, el valor de

$$\mathbf{E}(Z), \qquad \text{donde } Z = e^X - X + \cos(X).$$

Para ello, se sugiere el siguiente procedimiento:

- escribe una fórmula explícita para $F(x)$, la función de distribución de X;
- escribe una fórmula explícita para $F^{-1}(u)$, la inversa de la función de distribución de X;
- usa la expresión de F^{-1} y la instrucción `aleatorio()` de excel para generar 3000 muestras de X;
- y transforma esas 3000 muestras de X en 3000 muestras de Z siguiendo la especificación de arriba;
- finalmente, estima $\mathbf{E}(Z)$ a través de la media aritmética de las muestras de Z.

Por cierto, por si sirve de referencia, el valor exacto de $\mathbf{E}(Z)$ es

$$\mathbf{E}(Z) = -\frac{1}{2} - 8\sqrt{e} + 8\cos(1/2) + 4e - 4\cos(1)\,,$$

que es aproximadamente 2.043.

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 $(X_1, \dots, X_n)$, 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 $(x_1, \dots, x_{20})$ 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 la muestra obtenida

Usando la muestra obtenida en el apartado anterior,

- estima la media y la varianza de $\overline{X}$;
- estima la media y la varianza de S^2;
- construye un histograma de las muestras de $\overline{X}$ y otro de las muestras de S^2;
- calcula la proporción de muestras en las que $1 \leq \overline{X} \leq 1.2$ y $1.6 \leq S^2 \leq 2.0$ (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).