PROBABILIDADES Y ESTADÍSTICA (C)

ADICIONAL PRÁCTICA 5 (para resolver usando R u otro software)

1. El objetivo de este ejercicio es utilizar la ley de los grandes números para aproximar integrales de funciones continuas en intervalos finitos, tal como se propone en el ejercicio 12 de la Práctica 5. Esta vez procuraremos aproximar la probabilidad de que una variable $Z \sim \mathcal{N}(0,1)$ tome valores en un intervalo $[a,b]$. Es decir, queremos aproximar numéricamente la siguiente integral

$$\int_a^b \frac{1}{\sqrt{2\pi}}\, exp^{-\frac{x^2}{2}}\, dx\ . \tag{1}$$

Para ello, consideremos la siguiente propuesta:

   (a) Sea U una variable con distribución uniforme en el intervalo $[0,1]$: $U \sim \mathcal{U}[0,1]$. Dados $a, b \in \mathbb{R}$, encontrar α y β de forma tal que $Y = \alpha U + \beta$ tenga distribución uniforme en el intervalo $[a,b]$: $Y \sim \mathcal{U}[a,b]$.

   (b) A partir de $U_1, U_2, \ldots$ variables i.i.d., $U_i \sim \mathcal{U}[0,1]$ construir variables aleatorias $Y_1, Y_2, \ldots$ i.i.d. tales que $Y_i \sim \mathcal{U}[a,b]$. ¿A qué tiende en probabilidad

$$\frac{\sum_{i=1}^{n} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{Y_i^2}{2}}}{n} ?$$

   (c) Aproximar la integral (1) mediante

$$(b-a)\,\frac{\sum_{i=1}^{n} \frac{1}{\sqrt{2\pi}}\, e^{-\frac{Y_i^2}{2}}}{n}\ ,$$

   para los siguientes valores de a y b, y para $n = 100$, 1000 y 50000.

      i. $a = -1.96$ y $b = 1.96$.
      ii. $a = -2$ y $b = 1$.
      iii. $a = 0$ y $b = 2.34$.

   (d) Comparar estos resultados con aquellos que obtendría a partir de la tabla de la función Φ o utilizando un comando provisto por el software.

2. En este ejercicio estudiaremos la distribución del promedio de variables independientes e idénticamente distribuidas y a través de los histogramas correspondientes analizaremos el comportamiento de estas distribuciones a medida que promediamos un número creciente de variables aleatorias.

   Para ello generaremos una muestra de variables aleatorias con una distribución dada y luego calcularemos el promedio de cada muestra. Replicaremos ésto mil veces, es decir, generaremos una muestra aleatoria de la variable $\overline{X}$ de tamaño 1000. Observe que, en principio, desconocemos la distribución de $\overline{X}$. A partir de todas las replicaciones realizaremos un histograma para los promedios obtenidos para obtener una aproximación de la densidad o la función de probabilidad de $\overline{X}$.

(a) Considerar dos variables aleatorias X_1 y X_2 independientes con distribución $U(0,1)$ y el promedio de ambas, es decir,

$$\overline{X} = \frac{X_1 + X_2}{2}\,.$$

Generando una muestra de dos variables aleatorias con distribución $U(0,1)$ computar la variable promedio. Replicar 1000 veces y a partir de los valores replicados realizar un histograma. ¿ Qué características tiene este histograma?

(b) Aumentemos a cinco las variables promediadas. Considerar ahora 5 variables aleatorias uniformes independientes, es decir $X_1, X_2, \ldots, X_5$ i.i.d. con $X_i \sim U(0,1)$ y definir

$$\overline{X} = \frac{1}{5}\sum_{i=1}^{5} X_i\,.$$

Generando muestras de cinco variables aleatorias con distribución $U(0,1)$ computar la variable promedio. Repetir 1000 veces y realizar un histograma para los valores obtenidos. Comparar con el histograma anterior. ¿Qué se observa?

(c) Aumentemos aún más la cantidad de variables promediadas. Generando muestras de 30 variables aleatorias con distribución $U(0,1)$ repetir el ítem anterior. ¿Qué se observa?

(d) Idem anterior generando muestras de 200 variables aleatorias. ¿Qué pasa si se aumenta el tamaño de la muestra?

(e) Repetir los ítems a) a d) generando ahora variables con distribución $\mathcal{E}(\lambda)$ para distintos valores de λ. Comparar los resultados obtenidos.

(f) Repetir los ítems a) a d) generando ahora variables con distribución $Bi(1,p)$ para $p = 0.5, 0.01$ y 0.0001.

<u>Conclusiones</u>: