PRÁCTICA 6

ANÁLISIS ESTADÍSTICO DE DATOS SIMULADOS.

Ejercicio 1. Genere n valores de una variable aleatoria normal estándar de manera tal que se cumplan las condiciones: $n \geq 30$ y $S/\sqrt{n} < 0.1$, siendo S el estimador de la desviación estándar de los n datos generados.

a) ¿Cuál es el número esperado de datos que deben generarse para cumplir las condiciones?

b) ¿Cuál es el número de datos generados efectivamente?

c) ¿Cuál es la media muestral de los datos generados?

d) ¿Cuál es la varianza muestral de los datos generados?

Ejercicio 2. Estime mediante el método de Monte Carlo la integral

$$\textbf{i)} \int_0^1 \frac{e^x}{\sqrt{2x}}\,dx, \qquad \textbf{ii)} \int_{-\infty}^{\infty} x^2 \exp(-x^2)\,dx.$$

a) Indique cómo se obtiene mediante simulación el valor de la integral.

b) Genere al menos 100 valores y detenéngase cuando la desviación estándar muestral S del estimador sea menor que 0.01.

Ejercicio 3. Calcule mediante un método de Monte Carlo las siguientes integrales:

$$\textbf{i)} \quad \int_{\pi}^{2\pi} \left(\frac{\sin(x)}{x}\right) dx \qquad \textbf{ii)} \int_0^{\infty} \frac{3}{3+x^4}$$

a) Obtenga mediante simulación en computadora el valor de la integral deteniendo la simulación cuando el semi-ancho del intervalo de confianza del 95% sea justo inferior a 0.001.

b) Indique cuál es el número de simulaciones N_s necesarias en la simulación realizada para lograr la condición pedida y complete con los valores obtenidos la siguiente tabla (usando **4** decimales):

N^o de sim.	$\bar{I}$	S
1 000		
5 000		
7 000		
$N_s =$		

Ejercicio 4. Para $U_1, U_2, \ldots$ variables aleatorias uniformemente distribuidas en el intervalo $(0,1)$, se define:

$$N = \text{Mínimo}\left\{ n : \sum_{i=1}^{n} U_i > 1 \right\}$$

Esto es, N es igual a la cantidad de números aleatorios que deben sumarse para exceder a 1.

a) Observe que $E[N] = e$ por lo cual puede aproximar e con la media muestral $\bar{N}$.

b) Calcule la varianza de $\bar{N}$.

c) Estime la varianza del estimador $\bar{N}$ correspondiente a 1000 ejecuciones de la simulación y utilice este resultado para dar una estimación de e mediante un intervalo de confianza de 95%.

Ejercicio 5. Considere una sucesión de números aleatorios $\{U_i\}_i$ y sea M el primer n tal que la variable U_n es menor que su variable predecesora. Es decir,

$$M = n \quad \text{tal que} \quad U_1 \leq U_2 \leq \cdots \leq U_{n-1} \text{ y } U_n < U_{n-1}$$

a) Justifique que $P(M > n) = 1/n!$, $n \geq 0$.

b) Utilize la identidad

$$E[M] = \sum_{n=0}^{\infty} P(M > n)$$

para mostrar que $E[M] = e$.

c) Utilize el resultado del item anterior para estimar e mediante 1000 ejecuciones de una simulación.

d) Calcule el valor de la varianza muestral del estimador del item (c) y dé una estimación de e mediante un intervalo de confianza de 95%.

Ejercicio 6. Estime π sorteando puntos uniformemente distribuidos en el cuadrado cuyos vértices son: $(1,1)$, $(-1,1)$, $(-1,-1)$, $(1,-1)$, y contabilizando la fracción que cae dentro del círculo inscrito de radio 1. Obtenga un intervalo de ancho menor que 0.1, el cual contenga a π con el 95% de confianza. ¿Cuántas ejecuciones son necesarias?

Ejercicio 7. Sean $X_1, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas con media μ desconocida. Para a y b constantes dadas, $a < b$, nos interesa estimar

$$p = P\left(a < \overline{X} - \mu < b\right)$$

sabiendo que $\overline{X} = \sum_{i=1}^{n} X_i/n$.

a) Explique como utilizar el método “bootstrap” para estimar p.

b) Estime p asumiendo que para $n = 10$, los valores de las variables X_i resultan 56, 101, 78, 67, 93, 87, 64, 72, 80 y 69 y $a = -5$, $b = 5$.

Ejercicio 8. Sean $X_1, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuidas con varianza σ^2 desconocida. Se planea estimar σ^2 mediante la varianza muestral

$$S^2 = \sum_{i=1}^{n} (X_i - \overline{X})^2/(n-1)$$

Si $n = 2$, $X_1 = 1$ y $X_2 = 3$, ¿cuál es la estimación “bootstrap” de $\mathrm{Var}(S^2)$?

Ejercicio 9. Sean $X_1, \ldots, X_n$ variables aleatorias normales, independientes e idénticamente distribuidas con media μ y varianza σ^2 desconocidas. Considere el estimador T dado por:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}},$$

donde S es la desviación estándar muestral y $\bar{X}$ es la media muestral.

a) Dada una muestra en particular, explique detalladamente cómo puede utilizarse el método “bootstrap” para estimar la probabilidad $p = P(T > a)$, donde a es un número real conocido.

i) ¿Qué valor de μ se emplea en la implementación de este método?

ii) ¿Cuántas muestras "bootstrap" distintas pueden generarse a partir de una dada muestra empírica de tamaño n?

b) La siguiente muestra de 10 valores fue obtenida mediante un generador de variables aleatorias normales:

3.0592,	2.3304,	2.8548,	1.2546,	2.1628,	4.9828,	5.4259,	0.9078,	4.5811,	3.2749

Calcule los valores *muestrales* de $\bar{X}$ y S^2.

c) Dado $a = 2.50$, estimar p con 100000 simulaciones a partir de la muestra dada, utilizando el método de "bootstrap". Escriba el resultado con **5** decimales.

Ejercicio 10. Fundamentos teóricos hacen pensar que existe un nuevo tipo de estrella variable explosiva, cuya curva de luz decae en 9 dias. Dada la dificultad de observar este tipo de fenómenos, solo se conocen 10 candidatos. Las duraciones de las explosiones en los mismos son las siguientes (en dias):

$$7.5, \quad 12.3, \quad 8.8, \quad 7.9, \quad 9.3, \quad 10.4, \quad 10.9, \quad 9.6, \quad 9.1, \quad 11.2.$$

(a) Asumiendo que las explosiones observadas corresponden a un mismo fenómeno, calcule el valor esperado empírico de su duración.

(b) Obtenga la aproximación bootstrap del error cuadrático medio $ECM(\bar{X}, \mu)$.

(c) Determine, usando la técnica de bootstrap, la probabilidad de que la media muestral $\bar{X}$ se encuentre en el rango 8 a 10.

Ejercicio 11. Los siguientes datos son mediciones de dureza (o grado de temple) obtenido en muestras de metal templado con agua salada

144.98 145.04 145.02 145.04 145.03 145.03 145.04 144.97 145.05 145.03 145.02 145 145.02

Realize una aproximación Bootstrap de

$$P\left[|S^2 - \sigma^2| > 0.02\right]$$

Ejercicio 12. Considerar un sistema con un único servidor en el cual los clientes potenciales llegan de acuerdo con un proceso de Poisson de razón 4.0. Un cliente potencial entrará al sistema sólo si hay tres o menos clientes en el sistema al momento de su llegada. El tiempo de servicio de cada cliente está distribuido según una exponencial de parámetro 4.2. Después del instante $T = 8$ no entran más clientes al sistema (los tiempos están dados en horas).

a) Realize un estudio de simulación para estimar el tiempo promedio que un cliente pasa en el sistema.

b) Aplique el método "bootstrap" para estudiar el error cuadrático medio de su estimador.