Modelos y Simulación – Licenciatura en Computación

Guía N°5: Análisis Estadístico de Datos Simulados

Problema 1: Generar 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 la desviación estándar muestral 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?

e) Comente los resultados de los items (c) y (d). ¿Son sorprendentes?

Problema 2: Estimar mediante el método de Monte Carlo la integral

$$\int_0^1 \exp(x^2)\, dx \; .$$

Generar al menos 100 valores y detenerse cuando la desviación estándar del estimador sea menor que 0.01.

Problema 3: Para $U_1, U_2, \ldots$ variables aleatorias uniformemente distribuídas en el intervalo $(0,1)$, se define:

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

Es decir, N es igual a la cantidad de números aleatorios que deben sumarse para exeder a 1. Como se mostró en el Problema 3 de la Guía N°2, $E[N] = e$.

Calcular la varianza del estimador $E[N]$ correspondiente a 1000 ejecuciones de la simulación y dar una estimación de e mediante un intervalo de confianza de 95%.

Problema 4: Considere una sucesión de números aleatorios y sea M el primero que es menor que su predecesor. Es decir,

$$M = \{n : U_1 \leq U_2 \leq \ldots \leq U_{n-1} > U_n\}$$

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

b) Utilizar la identidad

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

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

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

d) Calcular la varianza del estimador del item (c) y dar una estimación de e mediante un intervalo de confianza de 95%.

Problema 5: Estimar π sorteando puntos uniformemente distribuídos en el cuadrado cuyos vértices son: $(1,1)$, $(-1,1)$, $(-1,-1)$, $(1,-1)$, y contabilizando la fracción que cae dentro del círculo inscripto de radio 1. Obtener un intervalo de ancho menor que 0.1, el cual contenga a π con el 95% de confianza. ¿Cuántas ejecuciones son necesarias?

Problema 6: Sean $X_1, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuídas con media μ desconocida. Para a y b constantes dadas, $a < b$, nos interesa estimar $p = P\left(a < \sum_{i=1}^{n} X_i/n - \mu < b\right)$.

a) Explicar como utilizar el método "bootstrap" para estimar p.

b) Estimar 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. Sean $a = -5$ y $b = 5$.

Problema 7: Sean $X_1, \ldots, X_n$ variables aleatorias independientes e idénticamente distribuídas 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 $\mathrm{Var}(S^2)$?

Problema 8: 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 distribuído según una exponencial de parámetro 4.2. Despues del instante $T = 8$ no entran más clientes al sistema (los tiempos están dados en horas). Realizar un estudio de simulación para estimar el tiempo promedio que un cliente pasa en el sistema. Aplicar el método "bootstrap" para estudiar el error cuadrático medio de su estimador.