INTEGRACIÓN MONTECARLO

MAURICIO BARRERA R.
Ingeniero Industrial
Profesional Universitario
Universidad Tecnológica de Pereira
mauriciobarrera@utp.edu.co

ALVARO TREJOS CARPINTERO
Estadístico
Profesor Asistente
Master en Investigación de Operaciones y Estadística
Universidad Tecnológica de Pereira
alvarot@utp.edu.co

PATRICIA CARVAJAL OLAYA
Estadística
Profesor Asistente
Master en Investigación de Operaciones y Estadística
Universidad Tecnológica de Pereira
pacarva@utp.edu.co

Grupo de investigación: "Estudio y aplicación de herramientas estadísticas modernas en la solución de problemas del entorno" – Multivariado -

RESUMEN
Este documento presenta un método para evaluar integrales definidas utilizando números aleatorios, método que es muy poderoso sobre todo cuando se trata de resolver integrales en múltiples dimensiones.

PALABRAS CLAVES: Integral**,** Simulación, Monte Carlo.

ABSTRACT
This paper presents a method to evaluate integrals using random numbers, method that is very powerful especially when it is a question of solving integrals in multiple dimensions.

KEYWORDS: Integral, Simulation, Monte Carlo.

1. INTRODUCCION

Los modelos de simulación son útiles para describir algunos fenómenos de la realidad. Así la simulación trata fundamentalmente de construir modelos abstractos en una computadora de tal forma que describan la parte esencial del comportamiento de un sistema de interés, pudiéndose entonces diseñar y realizar experimentos con el modelo y posteriormente extraer conclusiones de sus resultados. Los modelos de simulación tienen su origen en los trabajos de John Von Neumann, pero solo es apreciable su poder cuando se introduce el computador moderno.

Para la implementación de un modelo de simulación se hace necesario entonces, tener en primera instancia una fuente de números aleatorios que representen las observaciones del fenómeno, unas transformaciones de estas observaciones para alimentar el modelo, un procesamiento de las salidas del modelo y un análisis de las respectivas estadísticas resultantes.

Fecha de Recepción: 31 Agosto de 2006
Fecha de Aceptación: Dejar en blanco

En este documento se va a mostrar como los números aleatorios, además de ser la base fundamental de cualquier modelo de simulación, también permite, entre otras, hacer evaluaciones de integrales definidas, herramienta útil en las ciencias y en la ingeniería.

2. INTEGRACIÓN

Sea $f(x)$ una función, deseamos resolver entonces el problema de calcular la siguiente integral:

$$\theta = \int_0^1 f(x)dx \tag{1}$$

Para calcular el valor de la integral observemos que si **U** es una variable aleatoria distribuida uniformemente en el intervalo (0,1), entonces es posible expresar a θ como:

$$\theta = E(f(u)) \qquad (2)$$

Si Ui,...,Up son Variables aleatorias independientes y uniformes, entonces g(Ui), ….,g(Up) son variables aleatorias independientes e idénticamente distribuidas con valor esperado θ. Utilizando la ley fuerte de los grandes números que enuncia que con probabilidad 1.

$$\sum_{i=1}^{P} \frac{f(u_i)}{P} \rightarrow E(f(u)) = \theta \qquad (3)$$

Por consiguiente, es posible aproximar el valor de la integral, generando una gran cantidad de números aleatorios Ui, evaluar $f(x)$ en cada Ui, sumar estas cantidades y dividir entre el numero de números aleatorios.; obteniendo entonces una aproximación al verdadero valor de la integral cuando la cantidad de números aleatorios tiende al infinito.

3. LEY DE LOS GRANDES NUMEROS

Aunque el tema central de este documento es la integración a través del uso de números aleatorios, es conveniente hacer algunas claridades y justificaciones sobre el teorema que sustenta todo el método.

Primero que todo se debe mostrar un teorema de Concentración de distribuciones llamado “Desigualdad de Chebyshev” que va a ser la base de todo nuestro análisis**.**

Si **X** es una variable aleatoria con media μ y varianza σ^2**,** entonces para cualquier valor de $k > 0$ se tiene que

$$P[|X - \mu| \geq k\sigma] \leq \frac{1}{k^2} \qquad (4)$$

Esta desigualdad muestra que la cantidad de área de la distribución, situada fuera del intervalo $\mu - k\sigma \leq x \leq \mu + k\sigma$ es a lo sumo igual a $\frac{1}{k^2}$, dando así una idea de la concentración de una variable aleatoria sin importar como se distribuya.

Ahora si se tiene una sucesión de variables aleatorias $X1, X2,..., Xn$, Independientes con media μ y desviación estándar σ^2 se tiene para cada ε > 0**,** aplicando la desigualdad de Chebyshev [4]

$$P\left[\left|\frac{\sum x_i}{n} - \mu\right| \geq \frac{k\sigma}{\sqrt{n}}\right] \leq \frac{1}{k^2} \qquad (5)$$

Si se reemplaza $\frac{\kappa\sigma}{\sqrt{n}}$ por ε en (5) se tiene

$$P\left\{\left|\overline{x} - \mu\right| \geq \varepsilon\right\} = \frac{\sigma^2}{n\varepsilon^2} \qquad (6)$$

Si $n \rightarrow \infty$ infinito en (6), se convierte en

$$P\left\{\left|\overline{x} - \mu\right| \geq \varepsilon\right\} = 0 \qquad (7)$$

De este modo, la probabilidad de que $\bar{x}$ difiera de la media μ en mas de ε **,** con epsilon pequeño, tiende a ser igual a cero a medida que aumente el tamaño de la muestra.

Este teorema llamado “Ley Débil de los Grandes Números” seria suficiente para nosotros si lo que interesara fuera la convergencia en probabilidad de

$$\frac{\sum f(Ui)}{n} \quad \text{a} \quad E\big(f(u)\big) \qquad (8)$$

En realidad, el interés radica en la convergencia casi segura del promedio de variables aleatorias

$$\frac{\sum f(Ui)}{n} \quad \text{a} \quad E\big(f(u)\big) \qquad (9)$$

O en otras palabras el problema radica en garantizar la convergencia

$$\lim_{n\to\infty} \frac{\sum f(Ui)}{n} = E\big(f(U)\big) \qquad (10)$$

Afortunadamente el problema fue solucionado ya por los matemáticos a través de un teorema llamado “Ley Fuerte de los Grandes Números”**,** que enuncia que bajo algunas consideraciones muy generales el promedio de variables aleatorias de una muestra convergen con probabilidad 1 a la esperanza matemática de la variable en cuestión.

Así es la “Ley Fuerte de los Grandes Números” el teorema que justifica el método de integración Monte Carlo como llaman algunos autores [4]. Aunque la demostración del teorema no es el objetivo del presente trabajo se recomienda la consulta de textos avanzados de probabilidad como el indicado en la referencia [5] en el cual se explica con detalle la validez y las condiciones bajo las cuales funciona.

4. METODOLOGIA

Habiendo abordado ya las cuestiones técnicas referentes a la justificación de los teoremas de convergencia se procede a ilustrar con un ejemplo el proceso de cálculo de integrales a través de números aleatorios.

Supongamos que se requiere calcular el valor de $\int_0^1 x^3 dx$

Es evidente que esta integral da como resultado ¼ pero sirve como ejemplo.

El proceso entonces es el siguiente:

- Generar una gran cantidad de números aleatorios entre 0 y 1 (para nuestro ejemplo n =1000).
- Evaluar $f(Ui)$ (en el ejemplo Ui^3)
- Sumar y dividir entre n ,

$$\frac{\sum f(Ui)}{n} = \frac{\sum_{i=1}^{1000} Ui^3}{1000}$$

- Aproximar la integral al resultado

$$\hat{\theta} = \frac{\sum f(Ui)}{n} \rightarrow \int_0^1 f(x)dx \quad (11)$$

En el ejemplo de integrar x^3 entre 0 y 1 se utilizó una serie de números aleatorios uniformes generados en Excel con una aproximación de 0.254 con un error del 1.6% con respecto al verdadero valor de la Integral.

La cuestión radica ahora en encontrar alguna transformación apropiada para calcular cualquier integral definida.

$$\theta = \int_a^b f(x)dx \quad (12)$$

Si se hace la sustitución:

$$z = \frac{x-a}{b-a} \qquad dz = \frac{dx}{b-a}$$

Se tiene que

$$\theta = \int_0^1 f(a + (b-a)z)(b-a)dz \quad (13)$$

O en otros términos

$$\theta = \int_0^1 g(z)dz \quad (14)$$

Con $g(z) = (b-a) f(a + (b-a)z)$

Y queda solucionando el problema de la integral con límites [a, b] de integración.

Así sustituciones análogas permiten resolver las integrales impropias tipo

$$\theta = \int_0^\infty f(x)dx \quad (15)$$

Haciendo simplemente la sustitución

$$z = \frac{1}{x+1}$$

Con $dz = -\frac{1}{(x+1)^2}dx = -z^2 dx$

Y obteniendo

$$\theta = \int_0^1 \frac{1}{z^2} f(1/z - 1)dz \quad (16)$$

Este método de evaluar integrales puede ser mejorado con algunas técnicas de reducción de varianza, como lo son el uso de las variables antitéticas, variables de control, entre otros.

5. CONCLUSIONES

Aunque es cierto que un método numérico tradicional como la regla de Simpson es mas eficiente para resolver integrales unidimensionales, resulta ineficaz para una integral multidimensional y es aquí donde es poderoso el uso de números aleatorios, así para evaluar integrales de orden mayor o igual a cinco 5 es superior el método de números aleatorios que la regla del trapecio y para dimensiones superiores a ocho 8 los números aleatorios se vuelven superiores a cualquier método numérico de integración. [4]

En trabajos adicionales se mostraran métodos para resolver ecuaciones integrales y sistemas de ecuaciones lineales con números aleatorios y más exactamente con simulación Monte Carlo.

6. BIBLIOGRAFIA

[1] LAW-KELTON. Simulation Modelling and Analysis Third edition. Mc Graw Hill 2000

[2] THOMPSON James R. Wiley John and Sons. A Modellers Approach. 2000

[3] ROSS. Sheldon. Simulación. Prentice Hall. 1999

[4] INSUA - INSUA - MARTIN. Simulación Métodos y Aplicaciones Alfa omega. Limusa 1991

[5] CRAMER Harald. Métodos Matemáticos de Estadística. Aguilar.1968