

**Confiabilidad estructural de componentes
mecánicos con daño
edición 2020**

**Sesión 6: Introducción al Método de Monte
Carlo**

docentes del curso: R. Mussini, H. Cancela

dictado semestre 1 - 2020

Contenido:

1. Introducción.
2. Esquema básico de Monte Carlo.
3. Estimación de integrales de Lebesgue.
4. Integrales de Lebesgue-Stieltjes.

Introducción

El método Monte Carlo permite encontrar soluciones aproximadas a una variedad de problemas matemáticos a través de la realización de experimentos de muestreo estadístico en una computadora.

De manera un tanto sorprendente a primera vista, el método es aplicable tanto a problemas que no tienen absolutamente ningún contenido probabilístico, como a otros cuya estructura es inherentemente estocástica.

En general, los métodos numéricos que se basan en el empleo de evaluar n puntos en el seno de un espacio m -dimensional para producir una solución aproximada tienen un error que decrece con orden $n^{-1/m}$ en el mejor caso, lo que los convierte en extremadamente ineficientes cuando m es alto (fenómeno conocido como "maldición de la dimensionalidad").

En cambio, los métodos de Monte Carlo obtienen estimaciones con un error absoluto del orden $n^{-1/2}$, independientemente de m . Este punto se convierte en la ventaja fundamental del método, que en muchos casos es el único aplicable.

Bases históricas

El desarrollo de los métodos de Monte Carlo está basado en tres desarrollos matemáticos distintos, pero relacionados:

1. El desarrollo de las probabilidades a partir del siglo XVII. A partir del estudio de los juegos de azar, el desarrollo de la noción de variable aleatoria y de secuencia de resultados sucesivos del juego como una secuencia de eventos aleatorios. Posteriormente (S. XIX y XX) el reconocimiento de que la esperanza de una función de variables aleatorias continuas se expresa como una integral llevó a entender que es posible sortear de manera aleatoria números, transformarlos de acuerdo a las reglas prescritas, y de esta forma obtener una solución aproximada al cálculo de una integral en un problema que no tiene ningún contenido probabilístico (como por ejemplo el cálculo aproximado de π).
2. A fines del S. XIX, el estudio de caminatas aleatorias (random walks) permitió conectar las mismas con la solución de ecuaciones diferenciales

y ecuaciones integro-diferenciales. En ese momento, la aplicación discutida fue la solución analítica de dichas ecuaciones como forma de obtener soluciones (exactas o aproximadas) a los problemas planteados en términos de las caminatas aleatorias.

3. En el S. XX, y particularmente vinculado al desarrollo de la energía atómica en la segunda mitad del siglo, apareció la necesidad de resolver ecuaciones diferenciales e integrales parciales en espacios de dimensiones muy altas, que escapaba a las herramientas numéricas existentes. La propuesta que surgió fue el revertir el razonamiento previo, y en lugar de emplear la solución de las ecuaciones diferenciales como aproximación de un sistema estocástico, desarrollar modelos de caminatas aleatorias de las ecuaciones a resolver, y emplear una computadora para realizar muestras repetidas del modelo aleatorio y de esta forma obtener soluciones aproximadas al problema original. Esta propuesta fue la que recibió el nombre de Método de Monte Carlo, y está en la base del desarrollo posterior de la temática.

4. Un cuarto elemento, que apareció posteriormente, fue el desarrollo de la teoría de complejidad (computacional) a partir de los años 1970. Esta teoría dio una base sólida para identificar clases de problemas donde el cálculo exacto (o aproximado) insumía tiempos que crecían de manera exponencial; en algunos de estos, la propia teoría ha servido para demostrar que empleando Monte Carlo es posible en tiempos polinomiales encontrar resultados dentro de los errores de aproximación deseados.

Lecturas opcionales adicionales

Comentario: En caso de que el vínculo no esté vigente, buscar otro equivalente para suplirlo. Compartirlo en el foro.

- Elementos sobre la historia de MMC: (lecturas opcionales)
<http://people.sc.fsu.edu/~pbeerli/mcmc/metropolis1.pdf>
(último acceso 2020-06-20).
<http://web.archive.org/web/20160921205607/http://www.landorsimulation.com:80/formacion-con-simulacion/el-mundo-en-movimiento/historia-de-la-simulacion> (último acceso 2020-06-20). Comentario: se trata de una versión archivada en Internet Archive; el original,
<http://www.landorsimulation.com/formacion-con-simulacion/el-mundo-en-movimiento/historia-de-la-simulacion/> ya no está accesible.
<http://web.archive.org/web/20070519170205/http://stud2.tuwien.ac.at/~e9527412/history.html> (último acceso

2020-06-20). Comentario: se trata de una versión archivada en Internet Archive; el original, <http://stud2.tuwien.ac.at/e9527412/history.html> ya no está accesible.

- Artículo sobre “Monte Carlo method” en Wolfram MathWorld: <http://mathworld.wolfram.com/MonteCarloMethod.html> (último acceso 2020-06-20).
- Artículo sobre “Monte Carlo integration” en Wolfram MathWorld: <http://mathworld.wolfram.com/MonteCarloIntegration.html> (último acceso 2020-06-20).
- Discusión sobre impacto: <http://www.csm.ornl.gov/ssi-expo/MChist.html> (último acceso 2020-06-20).
- Artículo sobre “Monte Carlo method” en Encyclopedia of Mathematics: <https://mathworld.wolfram.com/MonteCarloMethod.html>

[//www.encyclopediaofmath.org/index.php/Monte-Carlo_method](http://www.encyclopediaofmath.org/index.php/Monte-Carlo_method)
(último acceso 2020-06-20).

- Presentación “Monte Carlo Methods: Early History and The Basics”
por Michael Mascagni,
http://www.cs.fsu.edu/~mascagni/MC_Basics.pdf (último acceso
2020-06-20).

Notación y repaso de elementos básicos de probabilidad

- X variable aleatoria (discreta o continua).
- F_X distribución de probabilidad de X , $F_X(x) = \text{Prob}(X \leq x)$.
- Si X es continua, f_X función de densidad de probabilidad de X (tal que $F_X(x) = \int_{-\infty}^x f_X(t)dt$).
- $E(X)$ esperanza de X ; muchas veces denotamos $\phi = E(X)$ el valor que se desea calcular a través del muestreo de X .
 - Si X es continua, $E(X) = \int_{-\infty}^{\infty} t f_X(t)dt$.
 - Si X es discreta y toma valores en un conjunto C ,
 $E(X) = \sum_{x \in C} x \text{Prob}(X = x)$.
- $\text{Var}(X) = E((X - E(X))^2) = E(X^2) - E(X)^2$ varianza de X ; muchas veces denotada σ_X^2 .

- $DE(X) = \sqrt{\text{Var}(X)}$ desviación estándar de X , muchas veces denotada σ_X .
- $CV(X) = DE(X)/E(X)$ coeficiente de variación de X , es una medida de la desviación o dispersión de una distribución de probabilidad, normalizada teniendo en cuenta el valor esperado.
- $\mathbf{X} = (X_1, X_2, \dots, X_m)$ vector aleatorio de dimensión m (compuesto por m variables aleatorias distintas, dependientes o independientes).
- $X^{(1)}, X^{(2)}, \dots, X^{(n)}$ - muestra de n variables aleatorias independientes con la misma distribución de X .
- Notar la diferencia entre un vector aleatorio de variables distintas, y una muestra de n variables equidistribuidas.

Distribuciones básicas

Se recuerda las siguientes distribuciones, que serán empleadas en la discusión subsiguiente y en algunos ejemplos y ejercicios.

- Distribución uniforme entre a y b , $U(a, b)$:
 - p.d.f $f_U(x) = 0$ si $x < a$ o $x > b$; $f_U(x) = 1/(b - a)$ si $a \leq x \leq b$;
 - $F_U(x) = 0$ si $x < a$; $F_U(x) = (x - a)/(b - a)$ si $a \leq x < b$,
 $F_U(x) = 1$ si $x \geq b$
- Distribución exponencial de parámetro $\lambda > 0$, $E(\lambda)$:
 - p.d.f $f_E(x) = 0$ si $x < 0$; $f_E(x) = \lambda e^{-\lambda x}$ si $x \geq 0$;
 - $F_E(x) = 0$ si $x < 0$; $F_E(x) = 1 - e^{-\lambda x}$ si $x \geq 0$.
- Distribución normal de parámetros μ y $\sigma > 0$, $N(\mu, \sigma)$:
 - p.d.f $f_N(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$.

Esquema básico de un Método Monte Carlo

Supongamos que deseo calcular un cierto valor ϕ , y conozco una variable aleatoria X con distribución F_X tal que $\phi = E(X)$.

El método de Monte Carlo en su versión más simple consiste en

1. *sortear* valores para un conjunto $X^{(1)}, X^{(2)}, \dots, X^{(n)}$, de variables aleatorias i.i.d. (independientes e idénticamente distribuidas) a X .
2. Calcular $S_n = X^{(1)} + \dots + X^{(n)}$, la suma de los n valores sorteados.
3. Calcular $\hat{X} = S_n/n$.
4. Calcular $\hat{V} = \sum_{i=1}^n (X^{(i)})^2 / (n(n-1)) - \hat{X}^2 / (n-1)$.

Se dice que \hat{X} es un *estimador* de ϕ ; discutiremos en las próximas transparencias los argumentos que llevan a pensar que con alta probabilidad sus valores son cercanos.

Comentarios

- Por sortear entendemos generar aleatoriamente, siguiendo la distribución de probabilidad F_X .
- El conjunto de valores sorteados para $X^{(1)}, X^{(2)}, \dots, X^{(n)}$ se llama *muestra*, o equivalentemente *conjunto de replicaciones de X* .
- n es el *tamaño de la muestra*, también llamado *número de replicaciones*.
- \hat{X} es en sí misma una variable aleatoria, de esperanza igual a ϕ . Resulta interesante estimar (si existe) la varianza de \hat{X} . Formalmente, si la varianza de X existe (y la denotamos $\sigma_X^2 = \text{Var}(X)$), aplicando las hipótesis de equidistribución e independencia de las observaciones $X^{(i)}$,

sabemos que

$$\begin{aligned}\text{Var}(\hat{X}) &= \text{Var}(S_n/n) = \text{Var}\left(\frac{\sum_{i=1}^n X^{(i)}}{n}\right) = \\ &= \sum_{i=1}^n \frac{\text{Var}(X^{(i)})}{n^2} = \frac{n\text{Var}(X)}{n^2} = \frac{\text{Var}(X)}{n}.\end{aligned}$$

Sin embargo, en general $\text{Var}(X)$ no se conoce. Como alternativa, podemos emplear la propia muestra para obtener un estimador de $\text{Var}(X)$, el estimador insesgado más habitual es

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (X^{(i)} - \hat{X})^2.$$

Realizando manipulaciones llegamos al siguiente estimador de la

varianza de \hat{X} ,

$$\hat{V} = \sum_{i=1}^n (X^{(i)})^2 / (n(n-1)) - \hat{X}^2 / (n-1).$$

Motivación del método

Es claro que el método de Monte Carlo no provee el valor exacto deseado, sino una aproximación, con un cierto error.

La justificación inicial del uso de Monte Carlo proviene de dos teoremas centrales de la probabilidad y la estadística, la Ley Débil de los Grandes Números y el Teorema del Límite Central (o Teorema Central del Límite).

Sea $X^{(1)}, X^{(2)}, \dots, X^{(n)}$ un conjunto de variables aleatorias i.i.d. (independientes e idénticamente distribuidas).

Sea $X^{(1)} + X^{(2)} + \dots + X^{(n)}$. Si existe la esperanza $\mu = E(X^{(i)})$, entonces la Ley Débil de los Grandes Números indica que, para todo $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \text{Prob} \left(\left| \frac{S_n}{n} - \mu \right| > \epsilon \right) = 0.$$

La interpretación es que si se suma n muestras independientes $X^{(i)}$, la

probabilidad que la suma (normalizada por n) esté lejos del valor exacto a estimar μ tiende a 0 con n .

Si, adicionalmente, existe la varianza $\sigma^2 = E((X^{(i)} - \mu)^2)$, el Teorema del Límite Central implica que

$$\lim_{n \rightarrow \infty} \text{Prob} \left(\frac{S_n - n\mu}{\sigma\sqrt{n}} < a \right) = (2\pi)^{-1/2} \int_{-\infty}^a e^{-x^2/2} dx.$$

Dado que el término de la derecha es la distribución de probabilidad de una variable aleatoria normal de media 0 y varianza 1, este teorema indica cuál es el comportamiento asintótico de la distribución del error cometido al emplear S_n como estimador de μ .

Ambos resultados proveen la motivación para aplicar Monte Carlo, ya que indican que con un número suficientemente alto de experimentos, es posible estimar el parámetro deseado incurriendo en pequeño error con alta probabilidad, y permiten cuantificar asintóticamente la relación entre estos dos valores (error y probabilidad) a través de la distribución normal.

Sin embargo, es preciso ser cauteloso en la aplicación práctica del método, ya que las implementaciones reales no verifican las hipótesis de estos dos teoremas. Por un lado, las limitaciones computacionales imponen una cota superior a los valores de n que se pueden emplear (y cuando se emplean números pseudo-aleatorios, la naturaleza cíclica de estos hace que no sea posible obtener una cantidad arbitraria de muestras independientes). Por otra parte, las estimaciones de errores a partir del Teorema del Límite Central sólo son válidas asintóticamente, pero su calidad para un valor de n dado depende de la velocidad de convergencia de la distribución de $S_n - n\mu$ a la distribución normal, lo que introduce una nueva fuente de error.

Es posible usar como alternativa otras fórmulas para derivar información sobre el error cometido por el método, tal como se discute en próximas sesiones.

Ejemplo 1

Supongamos que tenemos un satélite, que para su funcionamiento depende de que al menos 2 paneles solares de los 5 que tiene disponibles estén en funcionamiento, y queremos calcular ϕ la vida útil esperada del satélite (el tiempo promedio de funcionamiento hasta que falla, usualmente conocido en la literatura como MTTF - Mean Time To Failure).

Supongamos que cada panel solar tiene una vida útil que es aleatoria, y está uniformemente distribuída en el rango [1000 hs, 5000 hs] (valor promedio: 3000 hs). [NOTA: en la práctica, no es común usar distribuciones uniformes, que no reflejan bien el comportamiento real - en el ejemplo se postulan estas distribuciones por sencillez conceptual].

Para estimar por Monte Carlo el valor de ϕ , haremos n experimentos, cada uno de los cuales consistirá en sortear el tiempo de falla de cada uno de los paneles solares del satélite, y observar cual es el momento en el cuál han fallado 4 de los mismos, esta es la variable aleatoria cuya esperanza es el tiempo promedio de funcionamiento del satélite.

El valor promedio de las n observaciones nos proporciona una estimación de ϕ .

Exper. nro.	Tiempo hasta falla de					satélite, $X^{(i)}$
	Panel 1	Panel 2	Panel 3	Panel 4	Panel 5	
1	3027	1738	2376	4685	4546	4546
2	4162	4029	4615	3455	3372	4162
3	3655	2896	1378	4010	4144	4010
4	2573	2649	2117	3956	1281	2649
5	2977	2724	1355	2268	3262	2977
6	3756	4190	1749	3398	2581	3756
Prom.	-	-	-	-	-	$S_n/n = 3683$

Table 1: Una simulación detallada con $n = 6$ experimentos.

De esta simulación, tenemos un valor estimado para la vida útil esperada del satélite de 3683. Un indicador del error que podemos estar cometiendo es la varianza o equivalentemente la desviación estándar de S_n , que en este caso es (haciendo los cálculos) 297.

Seudocódigo básico de un Método Monte Carlo

Supongamos que deseo calcular un cierto valor ϕ , y conozco una variable aleatoria X con distribución F_X tal que $\phi = E(X)$.

Procedimiento EstimaciónMonteCarlo (integer n , real \hat{X} , real \hat{V})

Parámetro de entrada: n , *tamaño de la muestra*

Parámetros de salida: \hat{X} , *estimador de ϕ* ; \hat{V} , *estimador de $\text{Var}(\hat{X})$*

1. $\hat{X} = 0$. /* Inicialización */
2. $\hat{V} = 0$.
3. For $i = 1, \dots, n$ do
 - 3.1 Sortear un valor de la variable $X^{(i)}$ con distribución F_X
 - 3.2 $\hat{X} = \hat{X} + X^{(i)}$ /* Acumular*/
 - 3.3 $\hat{V} = \hat{V} + (X^{(i)})^2$ /* Acumular*/
4. $\hat{X} = \hat{X}/n$
5. $\hat{V} = \hat{V}/(n * (n - 1)) - \hat{X}^2/(n - 1)$

Para la implementación computacional de Monte Carlo, se supone siempre posible el conseguir muestras de variables aleatorias uniformes entre 0 y 1 ($U(0, 1)$), y el generar muestras de otras distribuciones a partir de la transformación de las variables uniformes. En la próxima sesión discutiré más a fondo este tema, esencial en la práctica.

Para dar los elementos necesarios para poder programar implementaciones, se adelantan los siguientes conceptos:

- Bibliotecas para generar números pseudo-aleatorios: conjunto de funciones que permiten generar secuencias de números que se comportan de forma razonablemente similar a una secuencia de variables aleatorias independientes con distribución uniforme entre 0 y 1.
 - Semilla: valor dado para inicializar la secuencia, semillas distintas resultan en secuencias distintas.
 - Función de inicialización: inicializa la secuencia con una semilla.
 - Función de sorteo: proporciona el próximo número aleatorio dentro de la secuencia.

- Generación de una v.a. $X = U(a, b)$ a partir de una v.a. $U = U(0, 1)$: se sortea el valor de U , y se calcula $X = a + (b - a)U$.
- Generación de una v.a. $X = E(\lambda)$ a partir de una v.a. $U = U(0, 1)$: se sortea el valor de U , y se calcula $X = -\ln(U)/\lambda$.

Ejemplo 2 - implementación

El siguiente ejemplo muestra una implementación en lenguaje C de Monte Carlo aplicado en un caso muy sencillo.

Problema: se desea calcular la esperanza de la vida útil de un satélite, cuyo equipamiento principal (sujeto a fallos) consiste en dos computadoras y un equipamiento de transmisión. El satélite funciona mientras al menos una de las dos computadoras y el equipo de transmisión funcionen. En el momento del lanzamiento, las computadoras poseen una vida útil aleatoria de distribución uniforme entre 0 y 500 hs; y el equipamiento de transmisión posee una vida útil aleatoria de distribución uniforme entre 0 y 1500 hs.

Solución: Definir un experimento consistente en sortear la vida útil (el tiempo hasta falla) de las dos computadoras y el equipo de transmisión, y en base a estos calcular el tiempo hasta falla del satélite. Para sortear estos valores, se emplea una rutina que genera números pseudoaleatorios.

Repetir este experimento un número fijo de veces (por ejemplo 1000), y

calcular la estimación de la esperanza de vida útil y de la desviación de esta estimación utilizando las fórmulas vistas previamente.

El código correspondiente (utilizando la función estándar `drand48` para generación de números aleatorios) está disponible en <http://www.fing.edu.uy/inco/cursos/mmc/codigoC/satelite.c>

También está disponible el código (utilizando el generador “Mersenne Twister”, disponible en <http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/emt.html>) (último acceso: 2020-06-20): <http://www.fing.edu.uy/inco/cursos/mmc/codigoC/satelite-version2.c>

Ambos códigos fueron compilados con `gcc`, usando la opción `-lm` (biblioteca matemática).

Definiciones / nomenclatura del método de Monte Carlo

- *Estimador puntual*: denota una expresión matemática que transforma un conjunto de datos (derivados de una muestra obtenida experimentalmente) en un único número, conocido como *estimación puntual*, que sirve como aproximación para el valor exacto (pero desconocido) del volumen (o más generalmente de la integral) de interés.
- *Estimador de intervalo o Intervalo de confianza*: denota una expresión matemática que transforma un conjunto de datos (derivados de una muestra obtenida experimentalmente) para calcular un intervalo numérico que, con una probabilidad especificada, contiene el valor exacto del volumen.
- En la práctica, en general se emplea un intervalo de confianza para complementar y servir como indicación del error potencial cometido por un estimador puntual.

- Un punto central en todo método de Monte Carlo es el número de sorteos independientes o replicaciones que deben realizarse para garantizar una cierta cota del error. Este número es llamado *tamaño de muestra*, y depende del error aceptable en cada estudio, de las características del problema (y del tiempo computacional disponible).
- Muchas veces es posible (eventualmente empleando algo de información previamente conocida) calcular antes del muestreo, un *tamaño de muestra de peor caso* que garantiza que, si se emplea al menos este número de experimentos, el estimador puntual resultante tendrá un error de acuerdo a las especificaciones deseadas.

En general el tamaño de muestra de peor caso resulta en un esfuerzo computacional mayor que el estrictamente necesario, pero conocerlo con anticipación puede ser útil en varios contextos. Por un lado, si no hay limitaciones en tiempo de cálculo, el usuario puede emplear este tamaño de muestra para asegurarse la calidad del resultado. Si por el contrario el tiempo de cálculo es una limitación importante, es posible emplear el conocimiento previo del tamaño de muestra de peor caso para varias

alternativas como forma de elegir aquella más conveniente, y también para estudiar como mejorar la eficiencia computacional del método.

- En algunos casos también es posible calcular a priori *tamaño de muestra de mejor caso*, que dan una cota inferior por debajo de la cuál es seguro que no puede alcanzarse las especificaciones de error requeridas. La relación entre el tamaño de muestra de mejor y peor caso sirven como indicadores de que tan pesimista es esta segunda medida.

Integración en múltiples variables

Discutiremos como emplear el método de Monte Carlo para calcular la integral de una función φ sobre una región \mathcal{R} acotada (que supondremos ha sido escalada de manera que $\mathcal{R} \subseteq \mathcal{J}^m$, donde \mathcal{J}^m es el hipercubo unitario m -dimensional $[0, 1]^m = [0, 1] \times \cdots \times [0, 1]$).

Formalmente, consideremos la integral en el sentido de Lebesgue

$$\zeta(\mathcal{R}) = \int_{\mathcal{R}} \varphi(\mathbf{x}) d\mathbf{x},$$

donde φ es una función Lebesgue integrable en múltiples variables, definida en una región \mathcal{R} .

El concepto de integral de Lebesgue es una de las generalizaciones clásicas fundamentales en la evolución del concepto de integral, y las funciones Lebesgue integrables incluyen funciones acotadas y no acotadas que cumplen ciertas condiciones de teoría de la medida (ver [https:](https://)

[//www.encyclopediaofmath.org/index.php/Lebesgue_integral](http://www.encyclopediaofmath.org/index.php/Lebesgue_integral) y http://en.wikipedia.org/wiki/Lebesgue_integral por más información).

En algunos casos (muy particulares), es posible evaluar $\zeta(\mathcal{R}) = \int_{\mathcal{R}} \varphi(\mathbf{x}) d\mathbf{x}$ por métodos analíticos. En la enorme mayoría de las ocasiones, esto no es posible, y es necesario aplicar métodos numéricos. La evaluación de estas integrales es uno de los problemas clásicos del análisis numérico, que si bien posee ya una amplia literatura, sigue siendo objeto de investigación y de propuesta de nuevos métodos para clases particulares de funciones.

Entre los métodos determinísticos aplicables, se encuentran las fórmulas de cuadratura y el uso de secuencias equi-distribuidas. Como alternativa, tenemos la opción de emplear Monte Carlo. Cada alternativa tiene sus ventajas y sus limitaciones.

Por comodidad, en la discusión subsecuente supondremos que $\varphi(\mathbf{x}) = 0$ para todo $\mathbf{x} \in \mathcal{J}^m \setminus \mathcal{R}$, esto hace que los valores de las integrales sobre \mathcal{R} y sobre \mathcal{J}^m coincidan.

Estimación por Monte Carlo

Sea \mathcal{R} una región (eventualmente de volumen desconocido $\lambda(\mathcal{R})$), en el hipercubo unitario m -dimensional $\mathcal{J}^m = [0, 1]^m = [0, 1] \times \cdots \times [0, 1]$.

Si \mathcal{R} fuera de tamaño arbitrario, pero acotada, se puede transformar (mediante traslación y homotecia) para que quede incluida en \mathcal{J}^m . Suponemos que \mathcal{R} está definida por un conjunto de desigualdades y de relaciones implícitas entre las variables espaciales $0 \leq x_i \leq 1$, $i = 1, \dots, m$.

Sea $\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)}$ una secuencia de vectores aleatorios uniformemente distribuidos en \mathcal{J}^m . Entonces

$$\bar{\zeta}_n(\mathcal{R}) = \frac{1}{n} \sum_{i=1}^n \varphi(\mathbf{X}^{(i)})$$

es un estimador insesgado de $\zeta(\mathcal{R})$, con error estándar

$\sqrt{(\int_{\mathcal{R}} \varphi^2(\mathbf{x}) d\mathbf{x} - \zeta^2(\mathcal{R}))/n}$. Siempre que $\int_{\mathcal{R}} \varphi^2(\mathbf{x}) d\mathbf{x} < \infty$, el error es de $O(n^{-1/2})$.

El siguiente pseudocódigo presenta el esquema más básico del método Monte Carlo aplicado a la integración de una función.

Procedimiento IntegraciónMonteCarlo (función φ , entero n , real $1 - \delta$)

Entrada: función a integrar φ , n tamaño de la muestra, $1 - \delta$ nivel de confianza

Salida: $\bar{\zeta}_n$ estimación de la integral, $V[\bar{\zeta}_n]$ estimación de la varianza

1. $S = 0; T = 0; /*$ Inicialización $*/$
2. For $j = 1, \dots, n$ do
 - 2.1 Sortear $\mathbf{X}^{(j)}$ con distribución uniforme en \mathcal{J}^m ;
 - 2.2 If $j > 1$ then $T = T + (1 - 1/j) (\varphi(\mathbf{X}^{(j)}) - S/(j - 1))^2$;
 - 2.3 $S = S + \varphi(\mathbf{X}^{(j)})$. $/*$ Acumular en S y T $*/$
3. $\bar{\zeta}_n = S/n$; $/*$ Estimador puntual de $\zeta(\mathcal{R})$ $*/$
4. $\hat{\sigma}_n^2 = T/(n - 1)$; $/*$ Estimador puntual de la varianza de $\varphi(\mathbf{X}^{(j)})$ $*/$
5. $V[\bar{\zeta}_n] = \hat{\sigma}_n^2/n$; $/*$ Estimador puntual de la varianza de $\bar{\zeta}_n$ $*/$
5. Calcular $[I_1(S, n, \delta), I_2(S, n, \delta)]$ $/*$ un intervalo de confianza de nivel $(1 - \delta)$ para $\zeta(\mathcal{R})$ $*/$

Hacemos notar por un lado que se acumula el valor de la función φ en cada punto sorteado; recordamos que si queremos integrar sobre $\mathcal{R} \subset \mathcal{J}^m$, definimos $\varphi(\mathbf{x}) = 0$ para todo \mathbf{x} que pertenezca al hipercubo pero no a \mathcal{R} .

Los puntos 2.3, 3 y 4 corresponden a una manera alternativa para calcular un estimador de la varianza. Las dos fórmulas clásicas son $\hat{\sigma}_n^2 = \frac{1}{n-1} \sum_{j=1}^n (\varphi(\mathbf{X}^{(j)}) - \bar{\zeta}_n)^2$, y $\hat{\sigma}_n^2 = \frac{1}{n-1} \left(\sum_{j=1}^n \varphi^2(\mathbf{X}^{(j)}) - n\bar{\zeta}_n^2 \right)$. La primera tiene el inconveniente de que es necesario acumular todos los valores de $\varphi(\mathbf{X}^{(j)})$ hasta el final del programa para hacer el cálculo, la segunda tiene problemas numéricos, dado que acumula valores elevados al cuadrado y puede producirse pérdida de precisión. La fórmula recursiva mostrada en el pseudocódigo es equivalente a estas dos, y es más robusta del punto de vista numérico (existe también una fórmula recursiva similar para la esperanza).

Lecturas adicionales:

- Demostración de la fórmula recursiva en el libro del curso, pag. 68; o en la enciclopedia MathWorld, entrada “Sample Variance Computation”,

página `http:`

`//mathworld.wolfram.com/SampleVarianceComputation.html`

(último acceso: 2020-03-26).

- Fórmula de Welford (una manera alternativa de calcular la varianza a medida que se realiza el muestreo, sin almacenar todos los valores generados), discutida en el blog de John D.Cook, página `http://www.johndcook.com/blog/standard_deviation/` (último acceso: 2020-03-26).

Cálculo de intervalo de confianza empleando la aproximación normal

Bajo un conjunto de condiciones bastante generales, que incluyen $\int_{\mathcal{R}} \varphi^4(\mathbf{x}) d\mathbf{x} < \infty$, es posible demostrar que $(\bar{\zeta}_n - \zeta)/(\hat{\sigma}_n^2/n)^{1/2}$ converge en distribución a una v.a. normal $N(0, 1)$ cuando $n \rightarrow \infty$. Esto es lo mismo que afirmar que

$$\lim_{n \rightarrow \infty} \text{Prob} \left(\frac{\bar{\zeta}_n - \zeta}{(\hat{\sigma}_n^2/n)^{1/2}} \leq \beta \right) = F_{N(0,1)}(\beta) = \Phi(\beta),$$

y por lo tanto si queremos un nivel de confianza de $1 - \delta$, podemos construir el intervalo

$$\left(\bar{\zeta}_n - \Phi^{-1}(1 - \delta/2)(\hat{\sigma}_n^2/n)^{1/2}, \bar{\zeta}_n + \Phi^{-1}(1 - \delta/2)(\hat{\sigma}_n^2/n)^{1/2} \right),$$

que es un intervalo de confianza asintóticamente válido (cuando $n \rightarrow \infty$).

Es muy común usar este intervalo en la práctica, aunque se mantienen los comentarios realizados en el caso de la estimación de volúmenes: dificultad para conocer la tasa de convergencia a la distribución normal, que puede no ser uniforme en ζ ; error introducido al emplear $\hat{\sigma}_n^2$ en lugar de σ^2 (desconocido); existencia de correlaciones entre las estimaciones de la media y de la varianza.

Para salvar estos problemas se requiere información o condiciones adicionales sobre la función a integrar; cuando estas no se conocen o están disponibles, la aproximación normal sigue siendo la herramienta básica para evaluar el error, aunque es necesario tener en cuenta que suele ser optimista (resultando en un intervalo de confianza que no es siempre suficientemente ancho para que la cobertura efectiva alcance la nominal).

Determinación del número de replicaciones empleando la aproximación normal

Cuando se tiene una especificación de error (ϵ, δ) predeterminada, y se desea calcular el número de replicaciones para alcanzar este error, se debe proceder en estas etapas:

1. realizar un conjunto de n' pruebas preliminares, para estimar la varianza (estimada por $\hat{\sigma}_{n'}^2$);
2. calcular el tamaño de muestra requerido de acuerdo a la aproximación normal, que es $n_N(\epsilon, \delta) = \lceil (\Phi^{-1}(1 - \delta/2))^2 \hat{\sigma}_{n'}^2 / \epsilon^2 \rceil$,
3. realizar un conjunto de $N \geq n_N(\epsilon, \delta)$ nuevas pruebas, con semillas diferentes (para asegurar la independencia entre la estimación del valor de N y los experimentos en sí).

Estimación de integrales de Lebesgue-Stieltjes

Si bien toda la presentación la hemos hecho para integrales de Lebesgue, definidas en un dominio acotado (supuesto por simplicidad incluido en \mathcal{J}^m), el método de Monte Carlo puede emplearse en un cuadro más general, a través de la observación que cuando existe la integral de Lebesgue $\zeta(\mathcal{R}) = \int_{\mathcal{R}} \varphi(\mathbf{x}) d\mathbf{x}$, su valor coincide con el de una integral de Lebesgue-Stieltjes

$$\zeta = \int_{\mathcal{Z}} \kappa(\mathbf{z}) dF(\mathbf{z}),$$

donde \mathcal{Z} es una región incluida en R^m (eventualmente no acotada), $\kappa()$ es una función medible (eventualmente no acotada) en \mathcal{Z} , y $F()$ es una distribución en los conjuntos medibles de \mathcal{Z} (es decir, $0 \leq F(\mathcal{A}) \leq F(\mathcal{B}) \leq 1$ para todo \mathcal{A}, \mathcal{B} conjuntos medibles tales que $\mathcal{A} \subseteq \mathcal{B} \subseteq \mathcal{Z}$), cuyas formas están determinadas por \mathcal{R} y φ . (ver https://www.encyclopediaofmath.org/index.php/Lebesgue-Stieltjes_integral, https://www.encyclopediaofmath.org/index.php/Stieltjes_integral, <http://www.math.utah.edu/~li/L-S%20integral.pdf>)

Al ser F una función de distribución, ζ puede interpretarse como la esperanza de $\kappa(\mathbf{Z})$, donde \mathbf{Z} es un vector aleatorio de dimensión m y distribución F .

A través de esta propiedad, vemos que es posible emplear el esquema de Monte Carlo realizando n muestras independientes $\mathbf{Z}^{(i)}$ con distribución F , y tomando $\hat{\zeta}_n = n^{-1} \sum_{i=1}^n \kappa(\mathbf{Z}^{(i)})$ como estimador de ζ .

En este caso también se puede construir un intervalo de confianza

$$\left(\hat{\zeta}_n - \Phi^{-1}(1 - \delta/2)(\hat{\sigma}_n^2/n)^{1/2}, \hat{\zeta}_n + \Phi^{-1}(1 - \delta/2)(\hat{\sigma}_n^2/n)^{1/2} \right)$$

asintóticamente válido ($n \rightarrow \infty$).

Dependiendo de las varianzas de $\bar{\zeta}_n$ y $\hat{\zeta}_n$ (y tiempos computacionales de los dos esquemas), puede ser más ventajoso emplear uno u otro de los dos esquemas.

El siguiente pseudocódigo corresponde al método Monte Carlo aplicado para calcular una integral de Lebesgue-Stieltjes.

Procedimiento MonteCarlo-Lebesgue-Stieljes ($\kappa, dF, n, 1 - \delta$)

Entrada: función a integrar κ , función densidad de probabilidad dF , n tamaño de la muestra, $1 - \delta$ nivel de confianza

Salida: $\hat{\zeta}_n$ estimación de la integral, $V[\hat{\zeta}_n]$ estimación de la varianza

1. $S = 0; T = 0;$ /* Inicialización */
2. For $j = 1, \dots, n$ do
 - 2.1 Sortear $\mathbf{Z}^{(j)}$ con densidad de probabilidad dF ;
 - 2.2 If $j > 1$ then $T = T + (1 - 1/j) (\kappa(\mathbf{Z}^{(j)}) - S/(j - 1))^2$;
 - 2.3 $S = S + \kappa(\mathbf{Z}^{(j)})$. /* Acumular en S y T */
3. $\hat{\zeta}_n = S/n$; /* Estimador puntual de $\zeta(\mathcal{R})$ */
4. $\hat{\sigma}_n^2 = T/(n - 1)$; /* Estimador puntual de la varianza de $\kappa(\mathbf{Z}^{(j)})$ /
5. $V[\hat{\zeta}_n] = \hat{\sigma}_n^2/n$; /* Estimador puntual de la varianza de $\hat{\zeta}_n$ */
5. Calcular $[I_1(S, n, \delta), I_2(S, n, \delta)]$ /* un intervalo de confianza de nivel $(1 - \delta)$ para $\zeta(\mathcal{R})$ */

Preguntas/ayuda para el estudio y la autoevaluación de lo aprendido; no es parte de la evaluación formal del curso ni es necesario entregar las respuestas.

- ¿Cuáles fueron los principales desarrollos matemáticos que permitieron el surgimiento de los Métodos de Monte Carlo?
- ¿Quiénes fueron las personas que desarrollaron el método en su forma moderna y le dieron su nombre?
- ¿En qué tipos de aplicaciones se suele usar métodos de Monte Carlo?
- ¿Cuál es el esquema básico del Método de Monte Carlo?
- ¿Qué es el tamaño de la muestra del método?
- ¿Qué es el estimador calculado por el método?

Ejercicio

Supongamos que un componente mecánico tiene una cierta resistencia aleatoria R con distribución uniforme $[200, 500]$, y está sujeta a una carga S con distribución uniforme $[0, 400]$ [NOTA - estas distribuciones no son realistas, se toman sólo a modo de ejemplo inicial sencillo].

- Escribir la probabilidad de falla como una integral.
- Calcular analíticamente la probabilidad de falla.
- Escribir un programa (en el lenguaje de su preferencia) que permita estimar utilizando Monte Carlo esa probabilidad de falla (estimador puntual e intervalo de confianza).
- Ejecute el programa con los siguientes tamaños de muestra: $m = 10; 10^3; 10^6$ y reporte los resultados (incluyendo también los tiempos de cálculo necesarios en cada caso).

- Como ayuda y para validar el programa, en el EVA del curso está disponible una planilla electrónica (formatos LibreOffice y Excel), que implementan la estimación por Monte Carlo de este ejemplo. La planilla está preparada con $m = 100$. Modificarla para utilizarla con $m = 10$ y 10^3 , y comparar con los resultados del programa.