UNIVERSIDAD DE SONORA

DIVISIÓN DE CIENCIAS EXACTAS Y NATURALES
DEPARTAMENTO DE MATEMÁTICAS

INTEGRACIÓN NUMÉRICA POR EL MÉTODO DE MONTE CARLO

tesis

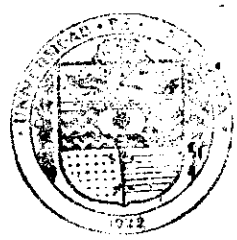
que presenta

Sonia Guadalupe Sosa León

para obtener el grado de

Licenciado en Matemáticas

Mayo 1997



A mis padres,

a mis hermanos,

a la raza.

Contenido

Introducción

El problema de asignar un valor a

$$I = \int_a^b f(x)\,dx$$

en el caso de un integrando f suficientemente regular (por ejemplo continua por tramos), es de gran importancia pues aparece en diversas ramas de la ciencia.

Encontrar tal número viene a ser un problema en cuya solución generalmente recurrimos al método que se apoya en el segundo Teorema Fundamental del Cálculo Integral, mediante el cual

$$\int_a^b f(x)\,dx = F(b) - F(a)$$

donde $F(x)$ es la antiderivada de $f(x)$.

En el caso en que $F(x)$ pueda obtenerse por métodos analíticos y resulte ser una función sencilla de evaluar, este método puede ofrecernos el mejor cálculo, sin embargo, algunas veces nos encontramos con que la integral de una función sencilla nos lleva a una función muy complicada para evaluar o donde las evaluaciones sólo son aproximaciones, por ejemplo:

$$\int_0^x \frac{dt}{1+t^4} = \frac{1}{4\sqrt{2}} \log \frac{x^2+x\sqrt{2+1}}{x^2-x\sqrt{2+1}} + \frac{1}{2\sqrt{2}} \left\{ \arctan \frac{x}{\sqrt{2}-x} + \arctan \frac{x}{\sqrt{2}+x} \right\},$$

donde el problema no sólo es el número de cálculos que se deben realizar sino que, debido a las funciones logaritmo y arcotangentes que involucra, el problema puede ser resuelto solo con cierto grado de aproximación.Un caso diferente es el de la integral $\int e^{-x^2} dx$, para la cual ni siquiera podemos obtener $F(x)$ en términos de una combinación finita de funciones usuales (polinomios, trigonométricas, trigonométricas inversas, exponencial, logaritmo, etc.).

Estos son algunos de los inconvenientes que se presentan al utilizar uno de los métodos de solución analítica para resolver un problema específico como es el de la integral definida.

El hecho de que frecuentemente se presenten problemas como los anteriores nos lleva al estudio de métodos de *Integración Numérica*, donde una integral es aproximada por una expresión del tipo

$$\int_a^b f(x)\,dx \cong w_1 f(x_1) + w_2 f(x_2) + ... + w_n f(x_n), \quad -\infty \leq a \leq b \leq \infty,$$

llamada fórmula de cuadratura, donde $x_1, x_2, ..., x_n$ son n puntos (o abcisas) tomadas del intervalo de integración y los números $w_1, w_2, ..., w_n$ son n "pesos" asignados a los puntos. Ejemplos clásicos de estas fórmulas son los métodos de Newton-Cotes y cuadratura gaussiana de Chebyshev y Hermite.

Conforme se ha ido generalizando el uso de las computadoras, poderosos programas de integración simbólica se han ido complementando con el análisis numérico para resolver problemas que involucran el cálculo de miles de integrales, problemas de integración múltiple, diferenciación, ecuaciones diferenciales ordinarias, análisis de errores de redondeo y muchos otros.

En el caso de la aproximación a la solución de integrales múltiples (que es el de mayor importancia) generalmente se utilizan los llamados métodos de cuadratura, sin embargo éstos, además de volverse sumamente difíciles de

manipular, tienen el problema de que conforme la dimensión del espacio de integración crece, el número de nodos requeridos para mantener una cota de error crece drásticamente, como es el caso de la regla del trapecio [1].

El método de Monte Carlo, desarrollado en los 40's, nos ofrece una manera menos complicada de obtener la aproximación a la integral y la posibilidad de acabar con el inconveniente de la dimensionalidad en la cota para el error.

El método consiste básicamente en lo siguiente: representar la solución analítica del problema como un parámetro de cierta población hipotética, el cual en la mayoría de los casos es el valor esperado de la distribución que describe a la población. Posteriormente a partir de una muestra de la población estimamos dicho parámetro, la cual será representada por una sucesión de números aleatorios.

El principal motivo que nos lleva al desarrollo del presente trabajo se basa en que el método de Monte Carlo presenta facilidad y eficiencia en la solución de integrales múltiples, y el hecho de que en la literatura existente [3,11,12,13], por lo regular sólo se desarrolla para integrales sencillas, que es el caso (al menos en términos de la expresión del error), en que los métodos tradicionales lo superan.

El trabajo está estructurado de la siguiente forma: en el Capítulo 1 desarrollamos la teoría referente a los números aleatorios y simulación de variables aleatorias; en el Capítulo 2 formulamos el método de Monte Carlo y lo aplicamos a la solución de integrales en regiones acotadas, integrales impropias y series; y el Capítulo 3 consiste en el análisis de error. Finalmente hacemos algunos comentarios sobre el trabajo, a manera de conclusión.

A lo largo del desarrollo del trabajo nos referimos reiteradamente a ciertos resultados de la teoría de probabilidad, éstos al igual que los programas utilizados se encuentran en dos apéndices al final del trabajo.

Capítulo 1

Simulación de Variables Aleatorias

1.1 Introducción

La simulación surge como una técnica alternativa para tratar problemas que son demasiado costosos para resolverse experimentalmente o demasiado complicados para ser tratados analíticamente.

Una variante de la simulación es la que se conoce como método de Monte Carlo el cual tiene su base en la teoría de probabilidad. Esta técnica se aplica a dos tipos de problemas. Primero en aquellos que llevan inmerso componentes estocásticas o de incertidumbre y segundo en los problemas matemáticos determinísticos que no pueden resolverse (si existe solución) por métodos determinísticos efectivos. En este caso mediante el método de Monte Carlo se pueden obtener soluciones aproximadas simulando variables aleatorias, como lo veremos más adelante, en la solución de integrales.

En este capítulo (Sección 1.3) se estudiarán las técnicas más importantes para simular variables aleatorias. Para esto en la Sección 1.2 se da una breve explicación de la manera en que obtenemos los números aleatorios, los cuales constituyen la base fundamental de la simulación mediante el método de Monte Carlo.

1.2 Números Aleatorios

Los números "elegidos aleatoriamente" pueden tener distintas aplicaciones como son: simular problemas físicos, seleccionar una muestra aleatoria, resolver problemas por métodos numéricos, recreación, entre otras.

El problema de generar números aleatorios no es trivial ya que no se tiene un concepto claro de aleatoriedad. De manera general, lo que se pide para que una sucesión de números sea aleatoria es que cada uno de ellos se obtenga por mera "casualidad" no teniendo relación con ningún otro y que tenga cierta probabilidad de pertenecer a un rango dado. Es decir, los números deben ser independientes y comportarse con cierta distribución.

Formas intuitivas de conseguir números aleatorios involucran lanzamiento de dados o de dardos, sacar bolas marcadas de una urna, hacer rodar una canica en una ruleta o tirar cartas de una baraja, pero físicamente resulta casi imposible que todos los artefactos usados (dado, dardo, ruleta, etc.) sean imparciales o *justos,* y poco práctico ya que se necesita repetir el experimento un gran número de veces. Existen también procedimientos electrónicos simulando a los artefactos anteriores con los que podemos conseguir sucesiones binarias de números aleatorios.

Alrededor de 1930 L.H.C.Tippett publicó una tabla con 40000 números aleatorios y desde entonces se han construido máquinas para generarlos mecánicamente, sin embargo, debido a su poca utilidad práctica, las tablas fueron abandonadas.

En este sentido, se buscó la manera de obtener sucesiones de números de una forma determinística y recursiva para generarlos por medio de una computadora, las cuales de manera general deben satisfacer las siguientes propiedades: (i) que se comporten aproximadamente con cierta distribución; (ii) la propiedad anterior será invariante sobre ciertas reglas de selección de una subsucesión. En base a estas características podemos comparar diferentes generadores de números. A este tipo de sucesiones se les llama pseudo-aleatorias. En el nombre pseudo-aleatorio se oculta de cierta forma la incredulidad de que la computadora, una máquina tan precisa y determinística, pueda producir números aleatorios. En nuestro caso, a pesar de ser obtenidas de esta manera las seguiremos llamando sucesiones aleatorias.

El primero en proponer sucesiones de este tipo fue J. Von Newman a mediados de siglo. Mediante su propuesta cada número era obtenido al elevar al cuadrado el anterior y extraer del resultado los dígitos que se encuentran al centro.

El método más usado para generar números aleatorios es el método de congruencias. Este puede tomar varias formas, sin embargo la más común está definida por

$$x_{n+1} = (\alpha x_n + c) \ (mod\, m), \qquad n \geq 0 \tag{1.1}$$

donde x_i, α, c y m son números enteros y $0 < x_i < m$. A esta sucesión se le llama lineal congruencial. En el caso homogeneo ($c = 0$) se le llama método multiplicativo congruencial.

Para obtener una sucesión de números en $(0, 1)$, obtenemos primero una sucesión de enteros $x_1, x_2, x_3, ...$ como la definida anteriormente y posteriormente generamos la sucesión

$$u_n = \frac{x_n}{m}, \qquad n = 1, 2, 3, ...$$

En la implementación de los métodos de Monte Carlo se requiere de grandes sucesiones de números aleatorios, para obtenerlas es necesario poner cuidado en la selección de α y m en (1.1), ya que principalmente de éstos depende el ciclo de la sucesión.

Los generadores de números aleatorios que se han desarrollado en los últimos años son muy confiables en el sentido de las propiedades requeridas. Para los fines que perseguimos no es necesario utilizar generadores tan sofisticados, en este trabajo generamos números aleatorios en $(0, 1)$ mediante las librerías de la IMSL [5]. Esta genera números aleatorios usando el método multiplicativo congruencial de forma

$$x_{n+1} = \alpha x_n (mod\, 2^{31} - 1).$$

Los posibles valores que toma α son 16 807, 397 204 094 y 950 706 376. El valor inicial se obtiene asignando un valor entre 0 y 2 147 483 647 en una subrutina de nombre RNSET, si se elige el valor cero o si no se inicializa con dicha subrutina, se obtendrá uno mediante el reloj del sistema.

1.3 Simulación de Variables Aleatorias

El material de esta sección nos será de gran utilidad ya que en la práctica la simulación de ciertas variables aleatorias es lo que nos permitirá dar un valor numérico a la integral. El siguiente teorema nos proporciona un procedimiento para la simulación de variables aleatorias.

Teorema 1.1. Sea U una variable aleatoria distribuida uniformemente en $(0,1)$, F una función de distribución arbitraria y X una variable aleatoria definida como:

$$X := \min\{y : F(y) \geq U\} \tag{1.2}$$

entonces la función de distribución de X es F.
Demostración. Supongamos que $\hat{F}$ es la función de distribución de X. Por demostrar que $F = \hat{F}$.

Para ésto, supongamos por el momento que se cumple que los eventos $[U \leq F(a)]$ y $[X \leq a]$, $a \in \Re$ arbitraria, son equivalentes, en el sentido que uno de ellos ocurre si y solo si el otro también ocurre. Entonces,

$$P[U \leq F(a)] = P[X \leq a], \quad a \in \Re \tag{1.3}$$

Ahora, como $U \sim \mathrm{U}(0,1)$ tenemos que $P[U \leq F(a)] = F(a)$ y por definición $P[X \leq a] = \hat{F}(a)$. Por lo tanto, (1.3) implica que $F(a) = \hat{F}(a)$, $a \in \Re$.

Para concluir, solo falta demostrar la equivalencia de los eventos. Sea $u \in R_U = (0,1)$. Si $u \leq F(a)$ entonces, $a \in \{y : F(y) \geq u\}$. De aquí, $\min\{y : F(y) \geq U\} \leq a$. Por lo tanto, $[U \leq F(a)]$ implica $[X \leq a]$.

Por otro lado, como F es no decreciente, $x \leq a$ implica que $F(x) \leq F(a)$. Además, de (1.2) tenemos que $u \leq F(x)$. Juntando estas dos desigualdades tenemos que

$$u \leq F(x) \leq F(a), \ \ a \in \Re$$

Por lo tanto, $[X \leq a]$ implica $[U \leq F(a)]$. Esto prueba la equivalencia de los eventos. ■

En vista del Teorema 1.1, podemos formular el siguiente algoritmo para generar valores de una variable aleatoria X con cierta función de distribución F mediante la generación de números aleatorios en $(0, 1)$:

Algoritmo 1.2. Generación de números aleatorios con función de distribución F.

1. Generar una sucesión $\{u_k\}_{k=1}^{n}$ de n números aleatorios.
2. Calcular $x_m = \min\{y : F(y) \geq u_m\}$, $m = 1, 2, ...$

El Algoritmo 1.2 es un generador universal en el sentido de que se pueden generar valores de cualquier variable aleatoria ya sea discreta o continua. En la siguiente gráfica se visualiza este hecho, donde a propósito se presentan dos puntos de discontinuidad.

1.3.1 Variables Aleatorias Continuas

En el caso particular de que se quieran generar variables aleatorias continuas, tenemos que

$$\min\{y : F(y) \geq U\} = F^{-1}(U)$$

donde F^{-1} es la transformación inversa de F. Esto es debido a que en este caso F es contina y creciente. Por lo tanto la expresión (1.2) toma la forma

$$X = F^{-1}(U)\,,$$

de tal manera que la expresión del paso 2, Algoritmo 1.2 la podemos sustituir por

$$x_m = F^{-1}(u_m)\,, \qquad m = 1, 2, ... \tag{1.4}$$

y así, obtenemos un algoritmo para generar valores de variables aleatorias continuas. A este algoritmo se le conoce como "Transformación Inversa", debido a la expresión (1.4).

Revisemos algunos ejemplos en los que generaremos valores de variables aleatorias con densidad muy conocidas, dado un conjunto de números aleatorios en $(0, 1)$. Algunos de estos ejemplos serán usados en el siguiente capítulo.

Ejemplo 1.3. Sea X una variable aleatoria con función de densidad $f(x) = 2x,\ 0 \leq x \leq 1$.La función de distribución está dada por

$$F(x) = \int_0^x 2t\ dt = x^2$$

y

$$F^{-1}(y) = \sqrt{y}\,, \qquad y \in (0,1).$$

Así la expresión 1.4 toma la forma:

$$x_m = \sqrt{u_m}\,, \qquad u_m \in (0,1).$$

Ejemplo 1.4. Distribución uniforme en (a,b). La función de densidad es

$$f(x) = \begin{cases} \frac{1}{b-a} & si\ a \le x \le b \\ \\ 0 & en\ otro\ caso \end{cases}$$

y la función de distribución está dada por

$$F(x) = \int_a^x \frac{1}{b-a}\,dt = \frac{x-a}{b-a}\,, \qquad a \le x \le b.$$

En este caso, $F^{-1}(y) = (b-a)y + a, \quad y \in (0,1)$, y por lo tanto, el generador (1.4) nos queda de la forma

$$x_m = (b-a)u_m + a, \qquad u_m \in (0,1).$$

Ejemplo 1.5. Distribución exponencial con parámetro λ. La función de densidad en este caso es

$$f(x) = \lambda e^{-\lambda x}, \quad \lambda > 0,\ x \ge 0.$$

La función de distribución y su inversa son:

$$F(x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}, \qquad x \geq 0$$

y

$$F^{-1}(y) = -\frac{1}{\lambda} \ln(1-y), \qquad y \in (0,1)$$

En este caso la expresión (1.4) toma la forma

$$x_m = -\frac{1}{\lambda} \ln(1-u_m), \qquad u_m \in (0,1) \tag{1.5}$$

Por otro lado, observemos que $U \sim$U$(0,1)$ implica que $1-U \sim$U$(0,1)$. Aplicando este hecho en nuestro ejemplo podemos sustituir $1-u_m$ por u_m, para cada $m \in \mathbb{N}$, en la expresión (1.5) y obtenemos el siguiente generador más simplificado de la distribución exponencial

$$x_m = -\frac{1}{\lambda} \ln u_m, \qquad u_m \in (0,1).$$

En el siguiente ejemplo generamos una distribución exponencial con rango arbitrario.

Ejemplo 1.6. Generalización de la exponencial.

Sea X una variable aleatoria con $R_x = [a, \infty)$. Si como generalización proponemos a la función

$$f(x) = \lambda e^{-\lambda(x-a)}, \qquad x \geq a.$$

es fácil ver que es una función de densidad.

La función de distribución está dada por

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)dt \\ &= \int_{a}^{x} \lambda e^{-\lambda(t-a)}dt \\ &= -e^{-\lambda(t-a)}\,|_{a}^{x} \\ &= 1 - e^{-\lambda(x-a)}\,. \end{aligned}$$

De aquí, siguiendo un procedimiento completamente análogo al Ejemplo 1.5 llegamos a que

$$x_m = -\frac{1}{\lambda}\ln u_m + a, \qquad u_m \in (0,1).$$

Una de las distribuciones de probabilidad más importantes es la normal o gaussiana. Para simular esta distribución no podemos usar el método de la transformación inversa debido a que no se cuenta con una expresión para su función de distribución. En el siguiente ejemplo se genera una variable aleatoria normal basándonos en el Teorema del Límite Central.

Ejemplo 1.7. Distribución Normal. La función de densidad normal de una variable aleatoria X se define como:

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \qquad -\infty < x < \infty,$$

donde $\mu = E[X]$ y $\sigma^2 = Var[X]$. Esto lo denotaremos como $X \sim N(\mu, \sigma^2)$. Si $\mu = 0$ y $\sigma^2 = 1$ decimos que X tiene distribución normal estandar y la denotaremos por Z $(Z \sim N(0,1))$. En este caso la función de densidad toma la forma

$$f(z) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}z^2}, \qquad -\infty < z < \infty\,.$$

Es bien conocido que la transformación de X:

$$Z = \frac{X - \mu}{\sigma}\,. \tag{1.6}$$

nos lleva a una variable aleatoria con función de probabilidad como la anterior.

Ahora, sean $U_1, U_2, ..., U_k$, k variables aleatorias independientes con distribución uniforme en $(0,1)$. Se sabe que $\mu = E[U_j] = \frac{1}{2}$ y $\sigma^2 = Var[U_j] = \frac{1}{12}$, $j = 1, 2, ..., k$.

De aquí, por el Teorema del Límite Central (ver Teorema A.18, Apéndice A), para k suficientemente grande,

$$\frac{S_k - k\mu}{\sqrt{k}\,\sigma} = \frac{\sum_{j=1}^{k} U_j - \frac{k}{2}}{\sqrt{\frac{k}{12}}} \sim N(0,1), \tag{1.7}$$

donde $S_k = U_1 + U_2 + ... + U_k$.

Por lo tanto, si queremos generar un valor de la variable aleatoria normal estandar debemos generar k números aleatorios en $(0,1)$ y usar (1.7). En general, para n valores necesitamos generar n arreglos de k números aleatorios en $(0,1)$ y hacer

$$Z_i = \frac{\sum_{j=1}^{k} U_j^{(i)} - \frac{k}{2}}{\sqrt{\frac{k}{12}}}\,, \qquad i = 1, 2, ..., n. \tag{1.8}$$

Ahora, de (1.6) tenemos que $X = \sigma Z + \mu$. De aquí, para generar valores de una variable aleatoria normal con media μ y varianza σ^2 hacemos

$$X_i = \sigma Z_i + \mu\,, \qquad i = 1, 2, ..., n,$$

con Z_i definido en (1.8).

1.3.2 Variables Aleatorias Discretas

El algoritmo de la Transformación Inversa no es aplicable para generar valores de variables aleatorias discretas debido a que la función de distribución de éstas no es continua. En este caso debemos aplicar el Algoritmo 1.2 directamente.

Supongamos que se quiere generar valores de una variable aleatoria X, con $R_x = \{0, 1, 2, ...\}$. Por definición tenemos que

$$F(x) = \sum_{k=0}^{]x[} P[X = k]$$

donde $]x[$ representa la parte entera de x.

Es fácil ver que en este caso $\min\{y : F(y) \geq u\} = n$, donde

$$F(n-1) \leq u \leq F(n) \quad \text{y } n \in R_X. \tag{1.9}$$

Esto significa que para generar los valores de X primero debemos generar la sucesión $\{u_m\}$ de números aleatorios y calcular posteriormente para cada u_m el valor de n que satisface (1.9) y hacer

$$x_m = n. \tag{1.10}$$

Ejemplo 1.8. Veamos como obtener un conjunto de números aleatorios de una variable aleatoria X tal que $R_X = \{0, 1, 2, ..., b\}$, $b \in \mathbb{N}$ y $P_n = P[X = n] = \frac{1}{b+1}$, $n = 0, 1, 2, ..., b$.

En este caso tenemos que para $n \in \mathbb{Z}$

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{b+1} & 0 \leq x < 1 \\ \frac{2}{b+1} & 1 \leq x < 2 \\ \vdots & \\ \frac{b}{b+1} & b-1 \leq x < b \\ 1 & x \geq b. \end{cases}$$

De aquí tenemos

$$F(n) = \frac{n+1}{b+1}, \qquad n = 0, 1, 2, ..., b.$$

Por la expresión (1.9)

$$\frac{n}{b+1} \leq u \leq \frac{n+1}{b+1}$$

y al resolver las desigualdades obtenemos

$$u(b+1) - 1 \leq n \leq u(b+1)$$

de donde concluimos que

$$n =]u(b+1)[$$

Por lo tanto, de (1.10) obtenemos que

$$x_m =]u_m(b+1)[$$

Ejemplo 1.9. Generalizando el ejemplo anterior, sea X una variable aleatoria con rango $R_X = \{a, a+1, a+2, ..., b\}$, es decir X toma valores entre a y b con $a, b \in \mathbb{Z}$ y tiene función de probabilidad

$$P_n = P[X = n] = \frac{1}{b-a+1}, \qquad n \in R_X.$$

La función de distribución está dada por

$$F(x) = \begin{cases} 0 & x < a \\ \frac{1}{(b-a)+1} & a \leq x \leq a+1 \\ \frac{2}{(b-a)+1} & a+1 \leq x \leq a+2 \\ \vdots & \\ 1 & x \geq b \end{cases}$$

y de aquí

$$F(n) = \frac{n-a+1}{b-a+1}, \qquad n = a, a+1, ..., b.$$

Aplicando (1.9) tenemos

$$\frac{n-a}{b-a+1} \leq u \leq \frac{n-a+1}{b-a+1}$$

y al resolver obtenemos

$$u(b-a+1) + a - 1 \leq n \leq u(b-a+1) + a,$$

entonces

$$n =]u(b-a+1)+a[\ .$$

De aquí y de (1.10) obtenemos el generador

$$x_m =]u_m(b-a+1)+a[.$$

Ejemplo 1.10. Distribución geométrica. La función de probabilidad geométrica con parámetro p se define

$$P_n = P[X=n] = pq^n, \qquad n = 0,1,2,...$$

donde $q = 1-p$. De aquí tenemos que la función de distribución toma la forma

$$F(x) = P[X \leq x] = \sum_{k=0}^{]x[} pq^k, \qquad x \geq 0.$$

Además,

$$1 - F(x) = q^{]x[+1}, \qquad x \geq 0 \tag{1.11}$$

o equivalentemente

$$F(x) = 1 - q^{]x[+1}, \qquad x \geq 0 \tag{1.12}$$

Sin pérdida de generalidad supondremos que $x \in \{0,1,2,...\}$

La expresión (1.11) se demuestra por inducción de la siguiente manera.

Si $n = 0$, tenemos que

$$1 - F(0) = 1 - p = 1 - (1-q) = q$$

Supongamos que se cumple para $n = k$, es decir $1 - F(k) = q^{k+1}$.

Entonces

$$\begin{aligned} 1 - F(k+1) &= 1 - P[X \leq k+1] \\ &= 1 - (P[X \leq k] + P[X = k+1]) \\ &= 1 - F(k) - pq^{k+1} \\ &= q^{k+1} - pq^{k+1} \\ &= q^{k+1}(1-p) \\ &= q^{k+2} \end{aligned}$$

lo que prueba que (1.11) y (1.12) son válidas.

Ahora, para generar los valores de la variable aleatoria X sustituimos la expresión (1.12) en (1.9) y tenemos

$$1 - q^n \leq u \leq 1 - q^{n+1}$$

Al resolver la desigualdad anterior obtenemos

$$n \leq \tfrac{\ln(1-u)}{\ln q} \quad \text{y} \quad n \geq \tfrac{\ln(1-u)}{\ln q} - 1$$

Con esto y usando el hecho de que si $U \sim \mathrm{U}(0,1)$ la variable aleatoria $(1-U)$ también es uniforme en $(0,1)$, obtenemos

$$\frac{\ln u}{\ln q} - 1 \leq n \leq \frac{\ln u}{\ln q}$$

Por lo tanto podemos elegir $n = \left]\frac{\ln u}{\ln q}\right[$ y generar los valores de la variable aleatoria X de la siguiente forma

$$x_m = \left]\frac{\ln u_m}{\ln q}\right[, \qquad u_m \in (0,1)$$

Ejemplo 1.11. En este caso generalizaremos los resultados del ejemplo anterior considerando una variable aleatoria geométrica que toma valores en $R_X = \{a, a+1, a+2, ...\}$, $a \in \mathbb{N}$. En este caso la función de probabilidad es

$$P_n = P[X = n] = pq^{n-a}, \qquad n = a, a+1, a+2, ...$$

y

$$\sum_{n=a}^{\infty} pq^{(n-a)} = 1$$

De manera completamente análoga al procedimiento del Ejemplo 1.10 se prueba que

$$1 - F(n) = q^{n-a+1}, \qquad n = a, a+1, a+2, ...$$

o

$$F(n) = 1 - q^{n-a+1}, \qquad n = a, a+1, a+2, ...$$

Por la expresión (1.9) tenemos

$$1 - q^{n-a} \leq u \leq 1 - q^{n-a+1}$$

lo cual implica que

$$\frac{\ln(1-u)}{\ln q} + a - 1 \leq n \leq \frac{\ln(1-u)}{\ln q} + a$$

Esta expresión también es equivalente a

$$\frac{\ln u}{\ln q} + a - 1 \leq n \leq \frac{\ln u}{\ln q} + a$$

y de aquí

$$n = \left]\frac{\ln u_m}{\ln q} + a\right[$$

Por lo tanto para generar los números aleatorios usamos la relación

$$x_m = \left]\frac{\ln u_m}{\ln q} + a\right[.$$

Capítulo 2

Integración Numérica por el Método de Monte Carlo

2.1 Introducción

El presente capítulo constituye la parte esencial del trabajo. Aquí presentaremos algoritmos que resuelven integrales del tipo

$$\int_B \cdots \int F(x_1, x_2, ..., x_d) dx_1 dx_2 ... dx_d,$$

donde F es una función de valor real definida en $\Re^d$, $d \geq 1$, y B es una región arbitraria en $\Re^d$.

El capítulo está estructurado de la siguiente manera: en la Sección 2.2 se formula el algoritmo que resuelve integrales en regiones acotadas arbitrarias; en la Sección 2.3 utilizamos el método de Monte Carlo para resolver integrales en regiones acotadas; finalmente, en las Secciones 2.4 y 2.5 analizamos cómo la teoría desarrollada nos permite calcular integrales impropias y series, respectivamente.

2.2 Formulación del Método de Monte Carlo para Integración Numérica en Regiones Acotadas

Primeramente consideraremos el caso cuando la región de integración es de la forma $B = R^d = C_1 \times C_2 \times ... \times C_d$, donde C_i, $i = 1, 2, ..., d$, es un intervalo en $\Re$. A la región B la llamaremos d-rectángulo. Posteriormente extenderemos los resultados al caso donde ésta es una región acotada arbitraria en $\Re^d$.

En este sentido, nuestro primer problema es asignarle un valor numérico a la integral

$$I := \int_{R^d} F(x)dx \tag{2.1}$$

donde $x = (x_1, x_2, ..., x_d)$, $dx = dx_1 dx_2 ... dx_d$.

Sea $X = (X_1, X_2, ..., X_d)$ un vector aleatorio con rango en $\Re^d$ tal que, $X_1, X_2, ..., X_d$ son variales aleatorias independientes cada una con distribución uniforme en $C_i \subset \Re$, $i = 1, 2, ..., d$ respectivamente. Entonces, X tiene una distribución uniforme en el d-rectángulo R^d [ver Ejemplo A.11, Apéndice A]. Esto es, la función de densidad de X (o la conjunta de $X_1, X_2, ..., X_d$) toma la forma

$$f_X(x) = \begin{cases} \frac{1}{\lambda(R^d)} & si\ x \in R^d \\ \\ 0 & en\ otro\ caso \end{cases}$$

donde $\lambda\left(R^d\right)$ representa el volumen del d-rectángulo R^d.

Ahora, como $F : \Re^d \to \Re$, tenemos que $F(X)$ es una variable aleatoria y usando la propiedad A.12.4, Apéndice A, tenemos

$$\mu_F := E\left[F(X)\right] = \int_{R^d} F(x)\frac{1}{\lambda\left(R^d\right)}dx,$$

lo que es equivalente a

$$I = \int_{R^d} F(x)dx = \lambda\left(R^d\right)\mu_F. \tag{2.2}$$

Esto significa que para estimar el valor de I, basta estimar el valor de la esperanza μ_F. Este problema ha sido ampliamente estudiado y resuelto en la teoría estadística y lo presentamos aquí adaptado a nuestro problema.

Sean $X^{(1)}, X^{(2)}, ..., X^{(n)}$, n vectores aleatorios independientes, cada uno con distribución uniforme en el d-rectángulo R^d; los cuales estan expresados de la forma $X^{(i)} = \left(X_1^{(i)}, X_2^{(i)}, ..., X_d^{(i)}\right)$, $i = 1, 2, ..., n$. Además supongamos que para cada $j = 1, 2, ..., d$ fija, las variables aleatorias $X_j^{(i)}$, $i = 1, 2, ..., n$ tiene distribución uniforme en el intervalo $C_j \subset \Re$. Esto significa que $X^{(1)}, X^{(2)}, ..., X^{(n)}$ forma una muestra aleatoria del vector X distribuido uniformemente en R^d.

Bajo estas condiciones, por la Ley Fuerte de los Grandes Números (ver Teorema A.17(b), Apéndice A) tenemos que

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} F\left(X^{(i)}\right) = \mu_F, \qquad \textit{casi seguramente} \tag{2.3}$$

y más aún, por la Ley Débil de los Grandes Números (ver Teorema A.17(a), ApéndiceA)

$$\lim_{n\to\infty} P\left[\left|\frac{1}{n}\sum_{i=1}^{n} F\left(X^{(i)}\right) - \mu_F\right| < \epsilon\right] = 1, \tag{2.4}$$

para cualquier $\epsilon > 0$.

De (2.3) y (2.4) podemos concluir que

$$\frac{1}{n}\sum_{i=1}^{n} F\left(X^{(i)}\right)$$

es un estimador razonable para la esperanza μ_F, lo que significa que un estimador para la integral I puede ser

$$I_n := \frac{\lambda\left(R^d\right)}{n} \sum_{i=1}^{n} F\left(X^{(i)}\right) \tag{2.5}$$

Inclusive, I_n es un "buen" estimador en el sentido de que es insesgado es decir, $E[I_n] = I$. Para probar esto nos apoyaremos en las propiedades A12.2-A12.4, Apéndice A

$$\begin{aligned} E\left[I_n\right] &= E\left[\frac{\lambda\left(R^d\right)}{n} \sum_{i=1}^{n} F\left(X^{(i)}\right)\right] \\ &= \frac{\lambda\left(R^d\right)}{n} E\left[\sum_{i=1}^{n} F\left(X^{(i)}\right)\right] \\ &= \frac{\lambda\left(R^d\right)}{n} \sum_{i=1}^{n} E\left[F\left(X^{(i)}\right)\right] \end{aligned}$$

De aquí, como cada vector aleatorio $X^{(i)}$, $i = 1, 2, ..., n$, tiene distribución uniforme en R^d, sus esperanzas coinciden y de hecho $E\left[F\left(X^{(i)}\right)\right] = \mu_F$. Por lo tanto

$$\begin{aligned} E\left[I_n\right] &= \frac{\lambda\left(R^d\right)}{n} \sum_{i=1}^{n} \mu_F \\ &= \lambda\left(R^d\right) \mu_F \\ &= I, \end{aligned} \tag{2.6}$$

lo cual prueba lo que se quería.

Lo anterior lo podemos resumir en el siguiente teorema:

Teorema 2.1. Sea $X^{(1)}, X^{(2)}, ..., X^{(n)}$ una muestra aleatoria con distribución uniforme en el d-rectángulo R^d, definido anteriormente. Entonces,

$$(a) \quad \lim_{n\to\infty} I_n = I \quad \textit{casi seguramente};$$

$$(b) \quad E[I_n] = I$$

donde I e I_n fueron definidos en (2.1) y (2.5), respectivamente.

Es importante observar que I_n es una variable aleatoria, por lo que necesitamos obtener una observación de I_n para poder asignarle un valor numérico a I. Es claro que el valor numérico asignado no siempre será el mismo debido a la aleatoriedad, sin embargo, por el Teorema 2.1 (a), podemos decir que entre más grande sea el tamaño de la muestra, el resultado de la observación se acerca más al valor exacto de I.

El Teorema 2.1 nos da la formulación del Método de Monte Carlo para integración numérica en d-rectángulos. Esto lo podemos extender a regiones acotadas arbitrarias $B \subset \Re^d$ de la siguiente manera.

Sea $F : \Re^d \to \Re$ y R^d un d-rectángulo tal que $B \subset R^d$. Entonces

$$I = \int_B F(x)dx = \int_{R^d} F(x)\mathrm{I}_B(x)dx \tag{2.7}$$

Es decir, una integral sobre regiones arbitrarias la podemos transformar en una integral sobre d-rectángulos. De esta manera podemos aplicar la teoría desarrollada anteriormente, en particular el Teorema 2.1. El estimador para I en este caso, toma la forma

$$\begin{aligned} I_n &:= \frac{\lambda(R^d)}{n} \sum_{i=1}^{n} F\left(X^{(i)}\right)\mathrm{I}_B\left(X^{(i)}\right) \\ &= \frac{\lambda(R^d)}{n} \sum_{\substack{i=1 \\ X^{(i)}\in B}}^{n} F\left(X^{(i)}\right) \end{aligned} \tag{2.8}$$

En la siguiente sección desarrollaremos los algoritmos que resuelven nuestro problema.

2.3 Integración en regiones acotadas

A manera de introducción revisaremos el caso más sencillo $d = 1$ que consiste en asignar un valor a la siguiente integral:

$$I = \int_a^b F(x)dx, \qquad F : \Re \to \Re$$

En este caso $\lambda\left(R^d\right) = \lambda\left(R^1\right) = b - a$, lo cual representa la longitud del intervalo de integración y, $X^{(i)}$ son las variables aleatorias distribuidas uniformemente en el intervalo (a, b). Con esto tenemos que el estimador definido en (2.5) toma la forma

$$I_n := \frac{b-a}{n} \sum_{i=1}^{n} F(X^{(i)}) . \tag{2.9}$$

Ahora debemos obtener una observación para I_n con el fin de asignar un valor numérico a I. Para este fin obtenemos una observación de las variables $X^{(1)}, X^{(2)}, ..., X^{(n)}$, generando un conjunto de n números aleatorios $x_1, x_2, ..., x_n$ en (a, b) y luego los sustituimos en (2.9).

En el Ejemplo 1.4, Capítulo 1, vimos que podemos generar números aleatorios distribuidos uniformemente en (a, b) mediante la relación

$$x^{(i)} = (b - a)u_i + a, \qquad i = 1, 2, ..., n$$

donde u_i, $i = 1, 2, ..., n$ forman un conjunto de números aleatorios en el intervalo $(0, 1)$, los cuales pueden ser generados directamente en la computadora.

Algoritmo 2.2. Estimación de $\int_a^b F(x)dx$.

1. Generar una sucesión $\{u_i\}_{i=1}^n$ de n números aleatorios.
2. Calcular $x^{(i)} = (b - a)u_i + a$, $i = 1, 2, ..., n$.
3. Calcular $F\left(x^{(i)}\right)$, $i = 1, 2, ..., n$.

4. Calcular I_n de acuerdo a (2.9)

Ahora, consideremos el caso $d \geq 2$ para regiones rectangulares. En este caso nuestro problema consiste en asignarle un valor numérico a

$$I = \int_{C_1}\int_{C_2}\cdots\int_{C_d} F(x_1, x_2, ..., x_d)dx_1dx_2...dx_d$$

donde $C_i = (a_i, b_i)$, $i = 1, 2, ..., d$ para lo cual aplicamos el estimador (2.5).

Aquí $\lambda\left(R^d\right) = (b_1 - a_1)(b_2 - a_2)\cdots(b_d - a_d)$ y $X^{(1)}, X^{(2)}, ..., X^{(n)}$ son n vectores aleatorios independientes con distribución uniforme en R^d. Por lo tanto el problema se reduce a obtener una observación de la muestra $X^{(1)}, X^{(2)}, ..., X^{(n)}$, donde $X^{(i)} = \left(X_1^{(i)}, X_2^{(i)}, ..., X_d^{(i)}\right)$, $i = 1, 2, ..., n$.

Con este propósito, recordando que para cada $j = 1, 2, ..., d$ fija, las variables aleatorias $X_j^{(i)}$, $i = 1, 2, ..., n$ tienen una distribución uniforme en $C_j = (a_j, b_j)$ y basándonos en el Ejemplo A.11, Apéndice A, necesitamos generar n números aleatorios en (a_j, b_j) los cuales representan la $j - ésima$ componente de cada uno de los n vectores. Es decir,

$$\begin{aligned}
\text{para} \quad j = 1: &\quad X_1^{(1)}, X_1^{(2)}, ..., X_1^{(n)} \sim U(a_1, b_1) \\
j = 2: &\quad X_2^{(1)}, X_2^{(2)}, ..., X_2^{(n)} \sim U(a_2, b_2) \\
&\vdots \\
j = d: &\quad X_d^{(1)}, X_d^{(2)}, ..., X_d^{(n)} \sim U(a_d, b_d)
\end{aligned}$$

La observación de la muestra de los n vectores la constituyen las columnas del arreglo anterior, es decir, $X^{(i)} = \left(X_1^{(i)}, X_2^{(i)}, ..., X_d^{(i)}\right)$, $i = 1, 2, ..., n$.

Ahora, para seleccionar los n números aleatorios distribuidos uniformemente en cada C_j, $j = 1, 2, ..., d$, utilizaremos la relación

$$x_j^{(i)} = (b_j - a_j)u_i + a_j, \quad i = 1, 2, ..., n$$

donde $\{u_i\}_{i=1}^n$ son n números aleatorios distribuidos uniformemente en el intervalo $(0,1)$.

Algoritmo 2.3. Estimación de $\int_{R^d} F(x)dx$.

1. Para cada $j = 1, 2, ..., d$, generamos una sucesión $\{u_i\}_{i=1}^n$ de n números aleatorios.
2. Calculamos $x_j^{(i)} = (b_j - a_j)u_i + a_j, \ \ i = 1, 2, ..., n.$
3. Hacemos $x^{(i)} = \left(x_1^{(i)}, x_2^{(i)}, ..., x_d^{(i)}\right), \ i = 1, 2, ..., n.$
4. Calculamos $F\left(x^{(i)}\right), \ \ i = 1, 2, ..., n.$
5. Calculamos I_n de acuerdo a (2.5).

Veamos ahora algunos ejemplos:

Ejemplo 2.4. Aproximar el valor de la integral

$$I = \int_2^3 \int_1^2 (x^2 + y^2)dxdy$$

cuyo valor exacto es $I = 8.\bar{6}$.

Se corrió el Programa 1, Apéndice B para 1,000 y 10,000 vectores aleatorios, y en cada caso obtuvimos cinco aproximaciones. Los resultados obtenidos son los siguientes:

Aproximación	1,000 *vectores*	10,000 *vectores*
1	8.699588	8.672840
2	8.659231	8.667934
3	8.647486	8.666948
4	8.631983	8.697117
5	8.668292	8.6626851

Notemos que aunque el número de vectores sea fijo, cada vez que aproximamos la integral el resultado es distinto. Esto es porque cada vez que implementamos el método los vectores que intervienen son distintos, debido a que son seleccionados aleatoriamente

Ejemplo 2.5. Aproximación del valor de la integral

$$I = \int_2^3 \int_0^1 \int_1^2 \int_0^2 \int_0^1 (xy^2 + xzr + yw^2) dx dy dz dw dr. \tag{2.10}$$

Sabemos que el valor exacto de la integral es $I = 5.75$. Con el Programa 1, Apéndice B, obtuvimos los siguientes resultados:

Aproximación	1,000 *vectores*	10,000 *vectores*
1	5.705205	5.778417
2	5.641860	5.737118
3	5.948819	5.768692
4	5.841824	5.757380
5	5.758955	5.751548

Para concluir esta sección, veremos como se resuelve el problema para una región de integración arbitraria. En vista de que este tipo de integrales se pueden transformar a integrales sobre regiones rectangulares [ver (2.7)], el algoritmo para resolverlas es similar al Algoritmo 3.4.

Algoritmo 2.6. Estimación de $\int_B F(x)dx$, $B \subset R^d \subset \Re^d$.

1. Para cada $j = 1, 2, ..., d$, generamos una sucesión $\{u_i\}_{i=1}^n$ de n números aleatorios.
2. Calculamos $x_j^{(i)} = (b_j - a_j)u_i + a_j, \ \ i = 1, 2, ..., n.$
3. Hacemos $x^{(i)} = \left(x_1^{(i)}, x_2^{(i)}, ..., x_d^{(i)}\right)$, $i = 1, 2, ..., n.$
4. Calculamos $F\left(x^{(i)}\right)\mathrm{I}_B\left(x^{(i)}\right)$, $i = 1, 2, ..., n.$

5. Calculamos I_n de acuerdo a (2.8).

Es importante observar que si queremos una buena precisión en los resultados, el rectángulo R^d debe ser el más pequeño que contenga a B.

Ejemplo 2.7. Aproximemos el volumen de un cilindro de altura 5, cuya base es el círculo de radio 0.5 y centro en $(\frac{1}{2}, \frac{1}{2})$.

Entonces, aproximaremos el valor de la integral

$$I = \iint_B 5dxdy$$

donde B es la región $\left(x - \frac{1}{2}\right)^2 + \left(y - \frac{1}{2}\right)^2 \leq \frac{1}{4}$.

El valor exacto de la integral es $I = 1.25\pi \approx 3.926991$. Consideremos el d-rectángulo $[0,1]\times[0,1]$, que contiene a la región B. Para este las aproximaciones obtenidas mediante el Programa 2, Apéndice B, son las siguientes.

Aproximación	1,000 *vectores*	10,000 *vectores*
1	4.0000	3.9270
2	3.8650	3.8705
3	3.9300	3.9315
4	4.0750	3.9255
5	3.9750	3.9360

2.4 Integrales Impropias

En esta sección estudiaremos el caso de asignar un valor numérico a integrales del tipo

$$I = \int_B F(x)dx$$

donde $F : \Re \to \Re$, y B es un intervalo no acotado de la forma (a, ∞), $a \in \Re$, o de la forma $(-\infty, \infty)$.

Sea X una variable aleatoria que toma valores en B con función de densidad $G(x)$, y sea Y la variable aleatoria definida como $Y = \frac{F(X)}{G(X)}$. Entonces, por la Proposición A.12.4 Apéndice A

$$E[Y] = \int_B \frac{F(x)}{G(x)} G(x)dx = I$$

De aquí tenemos que para estimar el valor de I basta estimar $E[Y]$.

Sean $X^{(1)}, X^{(2)}, ..., X^{(n)}$ n variables aleatorias independientes cada una con densidad G. Del desarrollo que se hizo en las secciones anteriores es fácil concluir que

$$I_n := \frac{1}{n} \sum_{i=1}^{n} \frac{F\left(X^{(i)}\right)}{G\left(X^{(i)}\right)} \tag{2.11}$$

satisface el Teorema 2.1.

Para obtener una observación del estimador (2.11) necesitamos especificar la función de densidad G. Esta elección depende del tipo de región de integración. Por ejemplo, si $B = (a, \infty)$, $a \in \Re$, la función de densidad G puede ser del tipo exponencial, digamos

$$G(x) = \begin{cases} \lambda e^{-\lambda(x-a)} & si\ x \geq a \\ \\ 0 & en\ otro\ caso \end{cases} \tag{2.12}$$

Sólo para fijar ideas, tomemos $\lambda = 1$ en (2.12).

Si $B = (-\infty, \infty)$ la función G puede ser la densidad normal con media μ y varianza σ^2, es decir,

$$G(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty \tag{2.13}$$

Específicamente, el estimador para integrales del tipo $\int_a^\infty F(x)dx$ toma la forma

$$I_n = \frac{1}{n}\sum_{i=1}^{n} \frac{F\left(X^{(i)}\right)}{e^{a-X^{(i)}}} \tag{2.14}$$

donde $X^{(1)}, X^{(2)}, ..., X^{(n)}$ forman una muestra aleatoria con densidad (2.12). Una observación de ésta muestra se puede obtener aplicando la siguiente fórmula (ver Ejemplo 1.6, Capítulo 1):

$$x^{(i)} = -\ln u_i + a, \qquad i = 1, 2, ..., n$$

donde $\{u_i\}_{i=1}^n$ es un conjunto de números aleatorios distribuidos uniformemente en $(0, 1)$.

Algoritmo 2.8. Estimación de $\int_a^\infty F(x)dx$, $a \in \Re$.

1. Generar una sucesión $\{u_i\}_{i=1}^n$ de n números aleatorios.
2. Calcular $x^{(i)} = -\ln u_i + a$, $i = 1, 2, ..., n$.
3. Calcular $\frac{F\left(x^{(i)}\right)}{e^{a-x^{(i)}}}$, $i = 1, 2, ..., n$.
4. Calcular I_n de acuerdo a (2.14).

Por otra parte, para integrales del tipo $\int_{-\infty}^\infty F(x)dx$ usamos el estimador

$$I_n = \frac{1}{n}\sum_{i=1}^{n} \frac{F\left(X^{(i)}\right)}{\frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(\frac{X^{(i)}-\mu}{\sigma}\right)^2}} \tag{2.15}$$

donde $X^{(1)}, X^{(2)}, ..., X^{(n)}$ es una muestra aleatoria con densidad (2.13). Por el Ejemplo 1.7, Capítulo 1, podemos concluir que dicha muestra se puede generar mediante la fórmula

$$x^{(i)} = \frac{\sigma \sum_{j=1}^{k} u_j^{(i)} - \frac{k}{2}}{\sqrt{\frac{k}{12}}} + \mu, \qquad i = 1, 2, ..., n \tag{2.16}$$

para k suficientemente grande y donde $\left\{u_j^{(i)}\right\}_{i=1}^{n}$ son números aleatorios distribuidos uniformemente en $(0, 1)$.

Se muestra en la literatura [ver Yakowitz (1977)] que para $k \geq 10$ se obtienen buenas aproximaciones. Para fijar ideas y para simplificar la relación (2.16) tomaremos $k = 12$. Por otro lado, la elección del parámetro μ depende de la forma de la función que deseamos integrar, mientras que σ^2 no. El valor de μ se toma aproximadamente como el punto tal que alrededor de él se encuentren los que tienen mayor imagen (centro) y tomaremos $\sigma^2 = 1$ para simplificar los cálculos. En los Ejemplos 2.15 y 2.16 se ve claramente este punto.

Algoritmo 2.9. Estimación de $\int_{-\infty}^{\infty} F(x)dx$

1. Para cada $i = 1, 2, ..., n$, generamos una sucesión $\left\{u_j^{(i)}\right\}_{j=1}^{12}$ de 12 números aleatorios.

2. Calcular $x^{(i)} = \sigma \sum_{j=1}^{12} u_j^{(i)} - 6 + \mu, \quad i = 1, 2, ..., n.$

3. Calcular $\frac{F(X^{(i)})}{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x^{(i)}-\mu}{\sigma}\right)^2}}, \quad i = 1, 2, ..., n.$

4. Calcular I_n de acuerdo a (2.15).

A continuación mostraremos algunos cálculos obtenidos con el Programa 3, Apéndice B.

Ejemplo 2.10. Aproximar el valor de la integral

$$I = \int_2^{\infty} \frac{1}{(x-1)^2} dx$$

cuyo valor exacto es $I = 1$.

Se obtuvieron las siguientes aproximaciones:

Aproximación	10,000 *nodos*	100,000 *nodos*
1	0.9236628	0.9623947
2	0.9167542	0.9813234
3	0.9078638	0.9777002
4	0.9340730	0.9518143
5	0.9131120	0.9754807

Ejemplo 2.11. Aproximar el valor de la integral

$$I = \int_0^{\infty} \frac{1}{1+x^2} dx$$

cuyo valor exacto es $I = \frac{\pi}{2} \approx 1.5707$.

Las aproximaciones obtenidas en este caso son las siguientes:

Aproximación	100,000 *nodos*
1	1.799913
2	1.603679
3	1.478176
4	1.493712
5	1.510096

En los dos ejemplos siguientes siguientes se obtuvieron resultados con el Programa 4, Apéndice B.

Ejemplo 2.12. Daremos una aproximación a la integral

$$I = \int_{-\infty}^{\infty} Sech\, x\, dx.$$

En este caso el valor exacto es $I = \pi \approx 3.141593$ y las aproximaciones obtenidas son las siguientes:

Aproximación	100,000 *nodos*	200,000 *nodos*
1	3.042625	3.032793
2	3.068993	3.031445
3	3.061886	3.051841
4	3.084252	3.042888
5	3.034239	3.034393

Ejemplo 2.13. Aproximación del valor de la integral

$$I = \int_{-\infty}^{\infty} xe^{-x^2} dx$$

cuyo valor exacto es $I = 0$.

En este caso obtuvimos las siguientes aproximaciones:

Aproximación	100,000 *nodos*	1,000,000 *nodos*
1	2.330882×10^{-3}	3.119611×10^{-4}
2	1.195181×10^{-4}	4.085115×10^{-4}
3	-1.453371×10^{-3}	-7.460424×10^{-4}
4	-1.313358×10^{-3}	-1.365496×10^{-3}
5	-3.202152×10^{-4}	-2.68412×10^{-4}

2.5 Series

En muchos problemas de la matemática aplicada surge la necesidad de calcular el valor de una suma de la forma

$$S = \sum_{i \in B} F(i), \qquad B \subset \mathbb{N}\,, \tag{2.17}$$

donde $F : \mathbb{N} \to \Re$.

La mayoría de las veces, cuando B es un conjunto no acotado, solo podemos determinar si converge o no aplicando algún criterio de convergencia. Por otro lado, si B es un conjunto acotado es importante calcular de manera más rápida el valor exacto de la serie.

En ésta sección estudiaremos la aplicación del método de Monte Carlo para aproximar el valor de la serie (2.17) principalmente cuando B es un conjunto no acotado. El caso cuando B es acotado puede ser resuelto exactamente através de un programa de computadora sin que represente grandes dificultades, sin embargo expondremos el caso como una forma de introducir el tema y para fines teóricos.

Sin pérdida de generalidad supondremos que $B = \{0, 1, 2, ..., c\},\ c < \infty$, $c \in \mathbb{N}$. Entonces, nuestro problema es aproximar la serie

$$S = \sum_{i=0}^{c} F(i)\,. \tag{2.18}$$

Sea X una variable aleatoria con rango $R_X = \{0, 1, 2, ..., c\}$ y función de probabilidad

$$P[X = i] = \frac{1}{c+1}\,, \qquad i = 0, 1, 2, ..., c. \tag{2.19}$$

Sabemos que

$$\mu_F = \sum_{i=0}^{c} F(i)\ P[X = i] = \frac{1}{c+1} \sum_{i=0}^{c} F(i)\,,$$

entonces

$$\sum_{i=0}^{c} F(i) = (c+1)\mu_F = S.$$

Es decir que para asignar un valor numérico a S basta asignar un valor a μ_F.

Por lo anterior, nuestro problema se reduce a estimar μ_F. De las secciones anteriores sabemos que un estimador insesgado de μ_F es la media muestral, por lo tanto un estimador insesgado para S es

$$S_n = \frac{c+1}{n} \sum_{i=1}^{n} F\left(X^{(i)}\right),$$

donde $X^{(1)}, X^{(2)}, ..., X^{(n)}$ es una muestra aleatoria con función de probabilidad (2.19), la cual puede ser simulada mediante la fórmula dada en el Ejemplo 1.8, Capítulo 1.

Revisemos ahora el caso cuando B es un conjunto no acotado. En general, nuestro problema ahora consiste en aproximar

$$S = \sum_{i=a}^{\infty} F(i), \qquad a \in \mathbb{Z}\ .$$

Sea X una variable aleatoria con rango $R_X = \{a, a+1, a+2, ...\}$ y función de probabilidad

$$P(i) = pq^{(i-a)}, \qquad i = a, a+1, a+2, ..., \tag{2.20}$$

y sea Y una variable aleatoria definida como $Y = \frac{F(X)}{P(X)}$. Entonces

$$E[Y] = \sum_{i=a}^{\infty} \frac{F(i)}{P(i)} P(i) = S.$$

De aquí una estimación de S la obtenemos al estimar $E[Y]$.

Sean $X^{(1)}, X^{(2)}, ..., X^{(n)}$ un conjunto de n variables aleatorias independientes con función de probabilidad $P(i)$. Sabemos que un estimador para la esperanza $E[Y]$ es la media muestral, entonces

$$S_n = \frac{1}{n} \sum_{k=1}^{n} \frac{F\left(X^{(k)}\right)}{P\left(X^{(k)}\right)}$$

Sustituyendo (2.20) en esta expresión tendremos que

$$S_n = \frac{1}{n} \sum_{k=1}^{n} \frac{F\left(X^{(k)}\right)}{pq^{\left(X^{(k)}-a\right)}} \tag{2.21}$$

es un estimador para S.

Ahora sólo falta obtener una observación de la muestra $X^{(1)}, X^{(2)}, ..., X^{(n)}$. Para ello aplicaremos la fórmula del Ejemplo 1.11, Capítulo 1:

$$x^{(k)} = \left] \frac{\ln u_k}{\ln q} + a \right[, \qquad k = 1, 2, ..., n$$

donde $\{u_k\}_{k=1}^{n}$ es un conjunto de números aleatorios en $(0,1)$.

Algoritmo 2.14. Estimación de $\sum_{i=a}^{\infty} F(i)$, $a \in \mathbb{N}$.

1. Generar una sucesión $\{u_k\}_{k=1}^{n}$ de n números aleatorios.
2. Calcular $x^{(k)} = \left] \frac{\ln u_k}{\ln q} + a \right[$, $k = 1, 2, ..., n$.
3. Calcular $\frac{F\left(x^{(k)}\right)}{pq^{\left(x^{(k)}-a\right)}}$, $k = 1, 2, ..., n$.
4. Calcular S_n de acuerdo a (2.21).

Las aproximaciones a las series mencionadas en los ejercicios siguientes, se obtuvieron eon el Programa 5, Apéndice B.

Ejemplo 2.15. Aproximaremos el valor de la serie infinita

$$S = \sum_{n=0}^{\infty} \left(\frac{1}{7}\right)^n,$$

la cual converge a $S = 1.1\bar{6}$.

En este caso obtuvimos los siguientes resultados:

Aproximación	1,000 *nodos*	10,000 *nodos*
1	1.190994	1.176143
2	1.142294	1.166444
3	1.164506	1.156145
4	1.178264	1.178736
5	1.136274	1.163490

Ejemplo 2.16. Aproximaremos el valor de la serie

$$S = \sum_{n=2}^{\infty} \frac{1}{n(\log n)^2}.$$

En este caso, sabemos que la serie converge pero desconocemos a que valor lo hace. Las aproximaciones obtenidas fueron:

Aproximación	10,000 *nodos*	100,000 *nodos*
1	9.438172	9.356206
2	9.073421	9.535509
3	9.704359	9.341381
4	9.714072	9.407254
5	9.361118	9.353600

Capítulo 3

Análisis de Error

3.1 Introducción

Siempre que resolvemos un problema por medio de algún método numérico, lo que obtenemos en realidad es una aproximación a la solución y con esto surge un nuevo problema: ¿qué tan precisa es la aproximación obtenida?, en otras palabras, ¿qué tanto nos aproximamos a la solución exacta?

El error absoluto, que se define como la diferencia entre la solución exacta y la aproximada, nos indica la precisión de la aproximación, sin embargo este no puede calcularse por razones obvias. Lo que se hace normalmente es dar una cota para el error, digamos $\varepsilon > 0$, y analizar bajo que condiciones se satisface

$$Err := |I_n - I| < \varepsilon \tag{3.1}$$

donde I_n es un estimador para I.

En el contexto de nuestro trabajo, la relación (3.1) no tiene sentido o es difícil de interpretar debido a que los estimadores I_n son variables aleatorias. Esto implica que aún cuando fijemos el tamaño de la muestra n, cada vez que se aplique el algoritmo, es casi seguro que el resultado sea distinto; más aún, sería un evento extraordinario que dos resultados coincidieran. Esto nos lleva a tener que hacer una interpretación probabilística de (3.1) y/o considerar el valor esperado de la variable aleatoria $|I_n - I|$ para obtener información

sobre el error. Esto último lo podemos obtener calculando el error cuadrático medio: $E\left[I_n - I\right]^2$. En el sentido probabilístico nuestro problema es calcular $\varepsilon > 0$ tal que

$$P\left[\ |I_n - I| < \varepsilon\right] = 1 - \alpha\,, \qquad \alpha \in (0,1). \tag{3.2}$$

A $(1-\alpha)$ se le llama nivel de confianza. La interpretación de (3.2) es que aproximadamente el $(1-\alpha)\%$ de las veces ocurre la relación (3.1).

El objetivo de este capítulo es analizar el error del método de Monte Carlo para integración numérica y está estructurado de la siguiente manera. En la Sección 3.2 calculamos explícitamente la cota de error, además de calcular el error cuadrático medio de la estimación. En esta parte veremos que la varianza de I_n juega un papel importante en el cálculo del error, entre más pequeña sea mejor es la precisión. Es por ésto que en la Sección 3.3 presentamos algunos métodos para reducir la varianza, con lo que concluimos el capítulo.

3.2 Análisis de Error

Presentaremos dos teoremas que nos darán información sobre el error. El primero es respecto al error cuadrático medio el cual puede ser interpretado como el error cuadrático más representativo.

Teorema 3.1. Sea I_n el estimador definido en (2.5), I como en (2.1) y σ_F^2 la varianza de la variable aleatoria $F(X)$, entonces

$$E\left[I_n - I\right]^2 = \frac{\left[\lambda\left(R^d\right)\right]^2}{n}\,\sigma_F^2\,, \qquad n \in \mathbb{N}\ .$$

Demostración. Primero, de (2.5) y las propiedades A.13.1-A.13.4, Apéndice A,

$$
\begin{aligned}
Var\,[I_n] &= Var\left[\frac{\lambda(R^d)}{n}\sum_{i=1}^{n}F(X^{(i)})\right] = \frac{[\lambda(R^d)]^2}{n^2}\,Var\left[\sum_{i=1}^{n}F(X^{(i)})\right] \\
&= \frac{[\lambda(R^d)]^2}{n^2}\sum_{i=1}^{n}Var\left[F(X^{(i)})\right] = \frac{[\lambda(R^d)]^2}{n^2}\,n\sigma_F^2 \qquad (3.3) \\
&= \frac{[\lambda(R^d)]^2}{n}\,\sigma_F^2
\end{aligned}
$$

Por otro lado, como $E\,[I_n] = I$ (ver 2.6), tenemos que por Definición A.13, Apéndice A,

$$E\,[I_n - I]^2 = Var(I_n) \qquad (3.4)$$

El resultado se sigue de (3.3) y (3.4). ■

Una primera conclusión que se desprende del Teorema 3.1 y la Proposición A.12.5, Apéndice A, es que el error promedio $E\,|I_n - I|$ es del orden $n^{-\frac{1}{2}}$. Además, se ve claramente que la varianza del estimador I_n o de la variable aleatoria $F(X)$ influye bastante en el cálculo del error. Para fines prácticos, este tipo de error tiene el inconveniente de que no se le puede extraer más información que la mencionada anteriormente. De entrada, uno quisiera saber cual es la precisión del resultado para un tamaño de muestra dado o el tamaño de la muestra (n) para obtener cierta precisión. El Teorema 3.2 resuelve en parte este problema basándonos en la relación (3.2).

Sea $\Phi(z) := P\,[Z \leq z]$ la función de distribución de una variable aleatoria normal estandar, denotaremos por Φ^{-1} su función inversa, es decir, $\Phi^{-1}(\beta)$, $0 \leq \beta \leq 1$, es un valor tal que $P\,[Z \leq \Phi^{-1}(\beta)] = \beta$.

En la demostración del siguiente teorema usamos el hecho de que una función de distribución, en particular Φ, satisface

$$P\,[|Z| \leq z] = 2\Phi(z) - 1 \qquad (3.5)$$

Teorema 3.2. Sea $\varepsilon > 0$. Si $P[\,|I_n - I| \leq \varepsilon] = 1 - \alpha$, para algún $\alpha \in (0,1)$, entonces para n suficientemente grande

$$\varepsilon \cong \frac{\sigma_F \; \lambda\left(R^d\right) \; \Phi^{-1}(1 - \frac{\alpha}{2})}{\sqrt{n}} \tag{3.6}$$

Demostración. Sea $\varepsilon > 0$ arbitrario y fijo. De (2.5) y (2.2), tenemos

$$\begin{aligned} P[\,|I_n - I| < \varepsilon] \;&\cong P\left[\left|\frac{\lambda(R^d)}{n}\sum_{i=1}^{n} F(X^{(i)}) - \lambda\left(R^d\right)\,\mu_F\right| < \varepsilon\right] \\ &= P\left[\left|\frac{\sum_{i=1}^{n} F(X^{(i)})}{n} - \mu_F\right| < \frac{\varepsilon}{\lambda(R^d)}\right] \\ &= P\left[\left|\frac{\sum_{i=1}^{n} F(X^{(i)}) - n\,\mu_F}{n}\right| < \frac{\varepsilon}{\lambda(R^d)}\right] \\ &= P\left[\left|\frac{\sum_{i=1}^{n} F(X^{(i)}) - n\,\mu_F}{\sqrt{n}\,\sigma_F}\right| < \frac{\varepsilon\sqrt{n}}{\sigma_F\;\lambda(R^d)}\right] \end{aligned}$$

De aquí, por el Teorema del Límite Central (Teorema A.18) y por (3.5), para n suficientemente grande tenemos que

$$P[\,|I_n - I| < \varepsilon] = P\left[\,|Z| < \frac{\varepsilon\sqrt{n}}{\sigma_F\;\lambda\,(R^d)}\right] = 2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma_F\;\lambda\,(R^d)}\right) - 1 \tag{3.7}$$

Ahora, si $P[\,|I_n - I| < \varepsilon] = 1 - \alpha$ para algún $\alpha \in (0,1)$, entonces de (3.7)

$$2\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma_F\;\lambda\,(R^d)}\right) - 1 \cong 1 - \alpha$$

lo que implica que

$$\Phi\left(\frac{\varepsilon\sqrt{n}}{\sigma_F\ (R^d)}\right) \cong \frac{2-\alpha}{2} = 1 - \frac{\alpha}{2} \tag{3.8}$$

Aplicando la función inversa Φ^{-1} en ambos lados de (3.8) obtenemos

$$\frac{\varepsilon\sqrt{n}}{\sigma_F\ \lambda\,(R^d)} \cong \Phi^{-1}(1-\frac{\alpha}{2}).$$

De aquí,

$$\varepsilon \cong \frac{\sigma_F\ \lambda\left(R^d\right)\ \Phi^{-1}(1-\frac{\alpha}{2})}{\sqrt{n}}$$

lo cual prueba lo que se quería. ■

En vista del Teorema 3.2, hemos encontrado una aproximación para el error la cual depende de la varianza σ_F^2, el tamaño de la muestra n y del nivel de confianza $1-\alpha$. Entre más grande es el tamaño de la muestra mejor será la precisión de los resultados. Es importante observar también el papel que juega el nivel de confianza. Como Φ^{-1} es una función creciente, entonces entre más grande sea $1-\alpha$ mayor será $\Phi^{-1}\left(1-\frac{\alpha}{2}\right)$. Esto significa que si queremos los resultados con un nivel de confianza muy alto para obtener cierta precisión, debemos aumentar el tamaño de la muestra.

En la práctica, para calcular ε se presentan dos dificultades. La primera es que $\sigma_F^2 = E\left\{[F(X)]^2\right\} - \mu_F^2$ (ver Definición A.13, Apéndice A), es decir, para calcular σ_F debemos calcular μ_F que es precisamente nuestro problema original. La segunda dificultad es que no se cuenta con una expresión analítica para Φ, y por lo tanto, tampoco para Φ^{-1}.

El primer problema queda resuelto si en (3.6), en vez de σ_F usamos su estimador. Se sabe de la teoría estadística que

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(F\left(X^{(i)}\right) - \bar{Y}\right)^2$$

es un estimador insesgado de σ_F^2, donde $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} F\left(X^{(i)}\right)$. De aquí, un estimador para σ_F es $S = \sqrt{S^2}$, del cual podemos obtener una observación generando una observación de la muestra adecuada $X^{(1)}, X^{(2)}, ..., X^{(n)}$.

Otro problema que se ve es lo difícil de calcular $\Phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ para $\alpha \in (0, 1)$ arbitrario, ya que sólo se cuenta con algunos valores los cuales se pueden calcular de cualquier tabla de distribución normal estandar. En nuestro caso usaremos los niveles de confianza suficientes para ilustrar los puntos más importantes: $1 - \alpha = 0.1,\ 0.5,\ 0.9,\ 0.95,\ 0.99$.

Con estos niveles de confianza obtenemos las siguientes expresiones para ε:

$1-\alpha$	ε
0.1	$\frac{0.13\sigma_F\,\lambda\left(R^d\right)}{\sqrt{n}}$
0.5	$\frac{0.68\sigma_F\,\lambda\left(R^d\right)}{\sqrt{n}}$
0.9	$\frac{1.65\sigma_F\,\lambda\left(R^d\right)}{\sqrt{n}}$
0.95	$\frac{1.96\sigma_F\,\lambda\left(R^d\right)}{\sqrt{n}}$
0.99	$\frac{2.58\sigma_F\,\lambda\left(R^d\right)}{\sqrt{n}}$

Observemos que entre más pequeño es el valor de $1-\alpha$, más pequeño será el valor de ε. Esto significa que podemos obtener una cota de error pequeña pero en la práctica poco confiable en el sentido de que sólo podemos esperar que un pequeño porcentaje de observaciones del estimador I_n satisfacerían la relación $|I_n - I| < \varepsilon$. Una interpretación análoga se puede dar si el valor de $1 - \alpha$ es grande. En este caso se tiene que ε toma valores grandes pero confiables.

Un punto importante que se rescata de todo lo anterior es que podemos dar una expresión para n cuando se ha especificado el nivel de confianza y la precisión con la que se quieren los resultados:

$$n > \left[\frac{S \lambda(R^d)\ \Phi^{-1}\left(1 - \frac{\alpha}{2}\right)}{\varepsilon} \right]^2$$

Ejemplo 3.3. Para la aproximación a la integral (2.10), calculamos el error para distintos números (n) de vectores aleatorios y para distintos niveles de confianza $(1 - \alpha)$. Los resultados son los siguientes:

$(1-\alpha)$	$n = 100$	$n = 1,000$	$n = 10,000$
0.1	0.488×10^{-1}	0.146×10^{-1}	0.473×10^{-2}
0.5	0.255	0.764×10^{-1}	1.248×10^{-1}
0.9	0.620	0.185	0.600×10^{-1}
0.95	0.737	0.220	0.714×10^{-1}
0.99	0.970	0.290	0.939×10^{-1}

Posteriormente para la misma integral calculamos el número de vectores necesarios para obtener un error dado (ε), para distintos niveles de confianza $(1 - \alpha)$ y obtuvimos:

$(1-\alpha)$	$\varepsilon = 0.1$	$\varepsilon = 0.01$	$\varepsilon = 0.001$
0.1	22	2,242	221,974
0.5	594	61,334	6,073,402
0.9	3,493	361,120	35,758,728
0.95	4,929	509,560	50,457,572
0.99	8,540	882,923	87,428,609

3.3 Métodos de Reducción de Varianza

Como se vió en la sección anterior, para reducir el error, fijando el nivel de confianza, podemos incrementar n o podemos reducir la varianza del estimador que es equivalente a reducir la varianza σ_F^2. El hecho de que la convergencia del método sea del orden $n^{-1/2}$ tiene serias consecuencias. Por ejemplo, si incrementamos el valor de n por un factor de 100, la precisión se incrementa solo en un factor de 10. Esto significa que en algunos casos el estar aumentando el valor de n para obtener cierta precisión puede ser inconveniente.

Por lo general, el problema de reducir el error se resuelve reemplazando el estimador original por uno equivalente el cual tiene varianza menor. Este procedimiento es conocido como método de Reducción de Varianza.

Un método muy sencillo pero poco conveniente para este caso es el siguiente:

Sea $G : \Re^d \to \Re$ tal que $J = \int_{R^d} G(x)\,dx < \infty$ es conocido, $\mu_{F-G} = E[F(X) - G(X)]$ y X un vector aleatorio con distribución uniforme en R^d. Poniendoahora

$$I = J + \int_{R^d} [F(x) - G(x)]\,dx = J + \lambda\left(R^d\right)\,\mu_{F-G} \tag{3.9}$$

la expresión sugiere que podemos tomar como estimador de I a

$$I'_n = J + \frac{\lambda\left(R^d\right)}{n} \sum_{i=1}^{n} \left[F\left(X^{(i)}\right) - G\left(X^{(i)}\right)\right] \tag{3.10}$$

en lugar de I_n (ver (2.5)), donde $X^{(1)}, X^{(2)}, ..., X^{(n)}$ son vectores aleatorios independientes con distribución uniforme en R^d. Es fácil ver que I'_n es un estimador insesgado de I.

Bajo este contexto el error de estimación es el que se produce cuando se estima μ_{F-G}. Por lo tanto, nuestro problema queda resuelto si logramos reducir la varianza de la variable aleatoria $F(X) - G(X)$. Considerando que

$$Var\,[F(X)-G(X)] = E\left\{[F(X)-G(X)]^2\right\} - \{E\,[F(X)-G(X)]\}^2$$
$$\leq E\left\{[F(X)-G(X)]^2\right\},$$

si elegimos a la función G tal que $|F(X)-G(X)| < \delta$, $\delta > 0$, entonces $Var\,[F(X)-G(X)] < \delta^2$. Esto significa que para poder reducir la varianza basta elegir la función G lo "más cerca" posible de la función F. Esto último es precisamente la desventaja de este método, ya que por lo regular la elección de G no resulta fácil.

A continuación presentaremos otros métodos alternativos.

3.3.1 Variables Antagónicas

Explicaremos este método en el caso cuando $d = 1$, de tal forma que nuestro problema es asignar un valor numérico a $I = \int_a^b F(x)\,dx$.

Sea $G : [a,b] \to \Re$ una función definida como

$$G(y) = \frac{1}{2}F(y) + \frac{1}{2}F(a+b-y), \qquad y \in [a,b]. \tag{3.11}$$

Observemos que $a \leq a+b-y \leq b$. Este hecho implica que si X es una variable aleatoria con distribución uniforme en (a,b), entonces la variable aleatoria $a+b-X$ tiene distribución uniforme en $[a,b]$ y por lo tanto

$$E\,[F(X)] = E\,[F(a+b-X)]\,. \tag{3.12}$$

De (3.12) y por las propiedades A.12.2-A.12.4, Apéndice A,

$$\begin{aligned}\mu_G &:= E\left[G(X)\right] \\ &= \tfrac{1}{2}E\left[F(X)\right] + \tfrac{1}{2}E\left[F(a+b-X)\right] \\ &= E\left[F(X)\right] \\ &= \mu_F\end{aligned}$$

De aquí y usando (2.2) tenemos que

$$I = (b-a)\mu_F = (b-a)\mu_G.$$

Por lo tanto, un estimador insesgado para I es:

$$\begin{aligned}\tilde{I}_n &= \tfrac{(b-a)}{n}\sum_{i=1}^{n} G\left(X^{(i)}\right) \\ &= \tfrac{(b-a)}{2n}\sum_{i=1}^{n}\left[F\left(X^{(i)}\right) + F\left(a+b-X^{(i)}\right)\right],\end{aligned}$$

donde $X^{(1)}, X^{(2)}, ..., X^{(n)}$ son variables aleatorias independientes con distribución uniforme en $[a,b]$.

Teorema 3.1. Sea F una función monótona continua con primera derivada continua. Entonces

$$Var(\tilde{I}_n) \leq \frac{1}{2}Var(I_n),$$

donde I_n es el estimador definido en (2.9).

Demostración. Es suficiente mostrar que

$$Var\left[G(X)\right] \leq Var\left[F(X)\right]. \tag{3.13}$$

Primero por (3.11) y la Definición A.13, Apéndice A, tenemos que:

$$\begin{aligned} Var[G(X)] &= E\left\{[G(X)]^2\right\} - \{E[G(X)]\}^2 \\ &= E\left\{\left[\tfrac{1}{2}F(X) + \tfrac{1}{2}F(a+b-X)\right]^2\right\} \\ &\quad - \left\{E\left[\tfrac{1}{2}F(X) + \tfrac{1}{2}F(a+b-X)\right]\right\}^2. \end{aligned}$$

Desarrollando los términos y aplicando (3.12) obtenemos:

$$\begin{aligned} Var[G(X)] &= \tfrac{1}{4}E[F^2(X)] + \tfrac{1}{4}E[F^2(a+b-X)] \\ &\quad + \tfrac{1}{2}E[F(X)\ F(a+b-X)] \\ &\quad - \tfrac{1}{4}\{E[F(X)]\}^2 - \tfrac{1}{4}\{E[F(a+b-X)]\}^2 \\ &\quad - \tfrac{1}{2}E[F(X)]\ E[F(a+b-X)] \qquad (3.14) \\ &= \tfrac{1}{2}E[F^2(X)] + \tfrac{1}{2}E[F(X)\ F(a+b-X)] \\ &\quad - \tfrac{1}{2}\mu_F^2 - \tfrac{1}{2}E[F(X)]\ E[F(a+b-X)] \\ &= \tfrac{1}{2}E[F^2(X)] + \tfrac{1}{2}E[F(X)F(a+b-X)] - \mu_F^2. \end{aligned}$$

Supongamos por el momento que

$$E[F(X)F(a+b-X)] \leq \mu_F^2. \qquad (3.15)$$

Entonces, de (3.14)

$$\begin{aligned} Var\,[G(X)] &\leq \tfrac{1}{2}E\,[F^2(X)] + \tfrac{1}{2}\mu_F^2 - \mu_F^2 \\ &= \tfrac{1}{2}E\,[F^2(X)] - \tfrac{1}{2}\mu_F^2 \\ &= \tfrac{1}{2}\,[\,E\,[F^2(X)] - \mu_F^2] \\ &= \tfrac{1}{2}Var\,[F(X)]\;, \end{aligned}$$

donde la última desigualdad se sigue de la Definición A.13, Apéndice A. Con ésto mostramos (3.13).

Para concluir solo falta mostrar la desigualdad (3.15). Si F es constante claramente se satisface la igualdad. Ahora supondremos que F es no decreciente tal que $F(b) > F(a)$. Esto implica que

$$F'(x) \geq 0, \qquad x \in [a,b]. \tag{3.16}$$

Definamos la función

$$\Phi(y) = \int_a^y F(a+b-x)\,dx - (y-a)\mu_F, \qquad y \in [a,b] \tag{3.17}$$

De (3.12) tenemos que

$$\mu_F = \frac{1}{b-a}\int_a^b F(a+b-x)\,dx,$$

lo cual implica que $\Phi(a) = \Phi(b) = 0$ y por el Teorema Fundamental del Cálculo,

$$\Phi'(y) = F(a+b-y) - \mu_F. \tag{3.18}$$

Por otro lado, como F es no decreciente, tenemos que $F(a) \leq F(x) \leq F(b)$, $x \in (a,b)$ y ésto implica que $F(a) \leq \mu_F \leq F(b)$. De aquí, $\Phi'(a) > 0$ y $\Phi'(b) < 0$. Por lo tanto, $\Phi(y) \geq 0$, $y \in [a,b]$. Usando este hecho y la relación (3.16) obtenemos:

$$\int_a^b \Phi(x)F'(x)\,dx \geq 0. \tag{3.19}$$

Integrando (3.19) por partes

$$\begin{aligned}\int_a^b \Phi(x)F'(x)\,dx &= \Phi(x)\,F(x)\Big|_a^b - \int_a^b F(x)\,\Phi'(x)\,dx \\ &= -\int_a^b F(x)\Phi'(x)dx \geq 0,\end{aligned}$$

donde la última igualdad se sigue del hecho de que $\Phi(a) = \Phi(b) = 0$. Ahora, la relación (3.19) implica

$$\int_a^b F(x)\Phi'(x)\,dx \leq 0.$$

Finalmente, sustituyendo (3.18) en esta expresión y multiplicando por $1/(b-a)$ llegamos a que

$$\begin{aligned}0 &\geq \tfrac{1}{b-a}\int_a^b F(x)\,[F(a+b-x)-\mu_F]\,dx \\ &= \tfrac{1}{b-a}\int_a^b F(x)\;F(a+b-x)\,dx - \frac{\mu_F}{b-a}\int_a^b F(x)\,dx \\ &= E\,[F(X)\,F(a+b-X)] - \mu_F^2\end{aligned}$$

Esto muestra que se cumple (3.15). La demostración en el caso cuando F es no creciente es completamente análoga. ■

Para el caso cuando $d \geq 2$, definimos una función $G : [a,b]^d \to \Re$ como:

$$G(y) := \frac{1}{2}F(y) + \frac{1}{2}F(\mathbf{a}+\mathbf{b}-y), \qquad y \in [a,b]^d,$$

donde $\mathbf{a} = (a,a,...,a)$, $\mathbf{b} = (b,b,...,b)$, $y = (y_1,y_2,...,y_d)$ y $a \leq y_i \leq b$, $i = 1,2,...,d$.

En este contexto la función G tiene propiedades análogas al caso $d = 1$. Por ejemplo, si X es un vector aleatorio con distribución uniforme en $[a,b]^d$, entonces $\mathbf{a}+\mathbf{b}-X$ tiene distribución uniforme en $[a,b]^d$ y además $\mu_G = \mu_F$. De aquí, el estimador toma la siguiente forma

$$\tilde{I}_n = \frac{\lambda\left([a,b]^d\right)}{2n} \sum_{i=1}^{n} \left[F\left(X^{(i)}\right) + F\left(\mathbf{a}+\mathbf{b}-X^{(i)}\right)\right],$$

donde $X^{(1)}, X^{(2)}, ..., X^{(n)}$ son vectores aleatorios independientes con distribución uniforme en $[a,b]^d$.

Este método lo hemos desarrollado considerando regiones de integración cuadradas, pero es fácil generalizar a regiones de integración arbitrarias usando los mismos argumentos que en(2.7).

3.3.2 Partición de Región

Aún cuando este método es muy general, en este trabajo lo restringiremos al contexto de la Sección 2.4 referente a integrales impropias, y particularmente a integrales del tipo

$$I = \int_a^\infty F(x)\,dx,$$

donde $a \in \Re$ y $F : \Re \to \Re$.

Sea X una variable aleatoria tomando valores en $[a, \infty)$ con función de densidad $G(x)$, y sea $Y := \frac{F(x)}{G(x)}$. En este caso se tiene que $E[Y] = I$ de tal manera que un estimador insesgado para I es

$$\hat{I}_n = \frac{1}{n} \sum_{i=1}^{n} \frac{F\left(X^{(i)}\right)}{G\left(X^{(i)}\right)}$$

donde $X^{(1)}, X^{(2)}, ..., X^{(n)}$ son variables aleatorias independientes con densidad G.

Claramente, tanto la varianza de $\hat{I}_n$ como la de Y depende de la elección de la función G. En la Sección 2.4 tenemos la función exponencial generalizada haciendo el papel de G, y la pregunta que surge es si esta función es la que produce menor varianza. Es decir, nuestro problema se reduce a ver qué función G minimiza $Var[Y] = Var\left[\frac{F(x)}{G(x)}\right]$.

Teorema 3.4. Sea $\mathcal{G}$ el conjunto de densidades de una variable aleatoria X con $R_X = [a, \infty)$. Entonces

$$\min\{Var[Y] : G \in \mathcal{G}\} = Var[Y_0] := \left[\int_a^{\infty} |F(x)|\, dx\right]^2 - I^2, \tag{3.20}$$

y ésto ocurre cuando

$$G(x) = \frac{|F(x)|}{\int_a^{\infty} |F(x)|\, dx}. \tag{3.21}$$

Demostración. Por definición tenemos que

$$\begin{aligned} Var[Y_0] &= E\left[\frac{F^2(X)}{G^2(X)}\right] - \left\{E\left[\frac{F(X)}{G(X)}\right]\right\}^2 \\ &= \int_a^{\infty} \frac{F^2(x)}{G(x)}\, dx - I^2, \quad G \in \mathcal{G}. \end{aligned} \tag{3.22}$$

Por demostrar que $Var[Y_0] \leq Var[Y]$, $\forall G \in \mathcal{G}$. Para ésto es suficiente mostrar

$$\left[\int_a^\infty |F(x)|\, dx\right]^2 \leq \int_a^\infty \frac{F^2(x)}{G(x)}\, dx. \tag{3.23}$$

Sea $G \in \mathcal{G}$. Por la desigualdad de Cauchy-Schwartz tenemos

$$\begin{aligned}
\left[\int_a^\infty |F(x)|\, dx\right]^2 &= \left[\int_a^\infty \frac{|F(x)|}{(G(x))^{1/2}}\, (G(x))^{1/2}\, dx\right]^2 \\
&\leq \left[\left(\int_a^\infty \frac{F^2(x)}{G(x)}\, dx\right)^{1/2} \left(\int_a^\infty G(x)\, dx\right)^{1/2}\right]^2 \\
&= \int_a^\infty \frac{F^2(x)}{G(x)}\, dx \int_a^\infty G(x)\, dx \\
&= \int_a^\infty \frac{F^2(x)}{G(x)} dx
\end{aligned}$$

Esto prueba (3.23). Finalmente, la relación (3.20) se sigue directamente de sustituir la expresión (3.21) en (3.22). ■

El Teorema 3.4 muestra que la función G que produce la mínima varianza del estimador $\hat{I}_n$ es dada por (3.21); pero observemos que para expresarla necesitamos conocer el valor de I y, en el caso en que $F(x) \cdot F(y) < 0$ para algún x y y en $[a, b]$, conocer también $\int_a^\infty |F(x)|\, dx$. Podemos salvar este problema si tomamos

$$G(x) \approx \frac{|F(x)|}{\int_a^\infty |F(x)|\, dx},$$

pero como se comentó al principio de la sección, este procedimiento podría resultar un poco difícil.

Otra alternativa para salvar el problema es la conocida como partición de región (la cual sólo la presentamos) que consiste en lo siguiente:

Sea $a^* > a$ tal que $I_1 := \int_a^{a^*} F(x)\,dx$ es conocida. Entonces

$$I = \int_a^{\infty} F(x)\,dx = I_1 + \int_{a^*}^{\infty} F(x)\,dx.$$

Considerando que G es una función de densidad en $[a, \infty)$, definimos la función de densidad truncada en (a^*, ∞) como $H(x) := G(x)/(1-P)$ donde $P := \int_a^{a^*} G(x)\,dx$. De aquí,

$$\begin{aligned} I &= I_1 + \int_{a^*}^{\infty} \frac{F(x)}{H(x)} H(x)\,dx \\ &= I_1 + E\left[\tfrac{F(X)}{H(X)}\right] = I_1 + (1-P)\,E\left[\tfrac{F(X)}{G(X)}\right]. \end{aligned}$$

Por lo tanto un estimador para I puede ser:

$$\bar{I}_n = I_1 + \frac{(1-P)}{n} \sum_{i=1}^{n} \frac{F\left(X^{(i)}\right)}{G\left(X^{(i)}\right)}$$

donde $X^{(1)}, X^{(2)}, ..., X^{(n)}$ son variables aleatorias independientes con densidad $H(x)$.

Para ejemplificar, tomemos $G(x) = e^{a-x}$ [ver (2.12)]. Es fácil mostrar, usando el método de la Transformación Inversa estudiado en el Capítulo 1, que el generador de números aleatorios con densidad H es:

$$x^{(i)} = -\ln\left[e^{a-a^*} - u_i(1-P)\right] + a, \qquad i = 1, 2, ..., n.$$

Para finalizar, enunciaremos el teorema que nos permite asegurar que con el estimador $\bar{I}_n$ reducimos la varianza y por lo tanto el error.

Teorema 3.5. $Var\left[\bar{I}_n\right] \leq (1-P)\,Var\left[\hat{I}_n\right]$.

La demostración de este teorema puede revisarse en [11] p.p.139.

Epílogo

El trabajo consistió en la formulación e implementación del método de Monte Carlo en la asignación de un valor numérico a integrales múltiples, integrales impropias y series.

Para este fin expresamos la integral en términos del valor esperado, lo que hizo necesario recurrir a la teoría de inferencia estadística para conseguir estimadores insesgados. Las observaciones del estimador las obtenemos mediante la generación de números aleatorios.

La principal ventaja del método de Monte Carlo que encontramos sobre las reglas clásicas de integración se reflejan definitivamente en el cálculo de integrales múltiples, cuando la dimensión del espacio de integración es mayor o igual a 5, lo que se refleja claramente en la cota del error, además de la innegable facilidad (gracias a poderosos generadores de números aleatorios en la computadora) en la implementación del método, independientemente de la función que tengamos en el integrando.

Sin embargo, a pesar de que pareciera que el método de Monte Carlo es la panacea para resolver problemas de éste tipo, debemos aclarar que lo que representa una ventaja potencial, a su vez oculta una desventaja. En primer lugar hablamos de una cota de error de naturaleza probabilística, lo que quiere decir que no tendremos la garantía de que la cota de error se satisface cada vez que implementemos el método, aún fijando el tamaño de la muestra. Esto nos acarrea problemas cuando requerimos de muchísima precisión. Otra desventaja que se presenta es que al tener un integrando "bueno" , ésto no se refleje en una mejor cota de error, como sucede con los métodos tradicionales del análisis numérico. Sin embargo este mismo argumento, como lo mancionamos anteriormente, viene a ser una ventaja en los casos de intgrandos difíciles de trabajar.

El hecho de que se considere a la integral como un parámetro implica que se puede aproximar mediante varios estimadores, de los cuales unos darán mejor precisión que otros. El criterio que se sigue para elegir al mejor estimador es el principio de que sea insesgado y luego que tenga mínima varianza. Esto último no es un problema trivial. Aquí presentamos solo 2 métodos de reducción de varianza, pero existen otros más eficientes que requieren de una teoría más avanzada, la cual no está al alcance de la desarrollada en el presente trabajo.

Un aspecto que vale la pena comentar es el referente a la generación de vectores aleatorios. La manera como los generamos fue ordenando de forma apropiada en bloques de tamaño d (ver Sección 2.3), una serie de números aleatorios. Existe otra manera de generarlos análoga al método de congruencias para la generación de números aleatorios. Es decir los vectores aleatorios x_i de dimensión d se generan mediante una relación del tipo [10]

$$x_i = Ax_{i-1}(mod\ m), \quad i = 1, 2, ..., n,$$

donde A es una matriz de dimensión apropiada y m un entero positivo. Evidentemente ésta es una manera más complicada y sofisticada de generarlos, pero puede ser una buena tarea futura, probar si con este método obtenemos mejores resultados.

Existen otras aplicaciones del método de Monte Carlo en muchos problemas básicos del análisis numérico, como son la solución de sistemas de ecuaciones lineales, solución de ecuaciones diferenciales parciales y de ecuaciones integrales, entre otras. En todas ellas se presentan ventajas y desventajas análogas a las que se mencionaron aquí.

Finalmente, haremos un comentario sobre la bibliografía utilizada para la realización del trabajo. La teoría básica de los números aleatorios se revisó en [5,7,10] y la de simulación de variables aleatorias en [9,13]; la fundamentación del método de Monte Carlo en [1,10,13] y el análisis de error en [10,11]. El resto de la bibliografía nos sirvió como material de apoyo.

Apéndice A

Teoría de Probabilidad

Definición A.1. Un espacio de probabilidad es una terna $(\Omega, \Im, P)$ donde Ω es un conjunto no vacío llamado espacio muestral, $\Im$ es una $\sigma-$algebra de subconjuntos de Ω y P es una función definida en $\Im$ llamada medida de probabilidad que satisface:

(i) $0 \leq P(A) \leq 1, \quad A \in \Im$;

(ii) $P(\Omega) = 1$;

(iii) Si $A_1, A_2, ... \in \Im$ es una colección numerable de eventos tal que $A_i \bigcap A_j = \emptyset$ si $i \neq j$, entonces

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Definición A.2. Una variable aleatoria X es una función medible definida en Ω tomando valores en $\Re(X : \Omega \to \Re)$.

Denotaremos por R_X al rango de la variable aleatoria.

Definición A.3. La función de distribución de una variable aleatoria X se define como:

$$F(x) := P[X \leq x], \qquad x \in \Re.$$

Propiedades de la función de distribución.

A.3.1. $F(x)$ es no decreciente, es decir, si $x_1 < x_2$, entonces $F(x_1) \leq F(x_2)$.

A.3.2. $\lim_{x \to -\infty} F(x) = 0$ y $\lim_{x \to \infty} F(x) = 1$.

A.3.3. F es continua por la derecha, es decir, $F(x) = \lim_{\substack{y \to x \\ y > x}} F(y)$.

Definición A.4. (a) Decimos que una variable aleatoria X es discreta si su rango es numerable, digamos, $R_X = \{x_1, x_2, ...\}$.
(b) La función de probabilidad de una variable aleatoria discreta se define como:

$$f(x) = P[X = x] := P[\varpi \in \Omega : X(\varpi) = x].$$

Además,

$$\sum_{x \in R_X} f(x) = 1$$

Definición A.5. Decimos que una variable aleatoria X es continua si su función de distribución es continua y tiene derivada continua exepto en un conjunto finito.

Si X es una variable aleatoria continua, entonces existe una función no-negativa f definida en $\Re$, tal que para cualquier $B \subset \Re$,

$$P[X \in B] := P\{\varpi \in \Omega : X(\varpi) \in B\} = \int_B f(x)\, dx.$$

A $f(x)$ se le llama función de densidad de X y satisface

$$\int_{-\infty}^{\infty} f(x)\,dx = 1.$$

Ejemplo A.6. La función de distribución de una variable aleatoria uniforme en (a, b) es

$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x \leq b \\ 1 & x > b \end{cases}$$

La función de densidad es:

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \textit{en otro caso} \end{cases}$$

En este caso denotaremos $X \sim U(a, b)$.

Definición A.7. Sean $X_1, X_2, ..., X_d$ variables aleatorias definidas en el mismo espacio muestral Ω. Al vector $(X_1, X_2, ..., X_d)$ se le llama vector aleatorio de dimensión d. Observemos que $X = (X_1, X_2, ..., X_d)$ puede usarse como una función definida en Ω tomando valores en $\Re^d$, $X : \Omega \to \Re^d$.

Definición A.8. Sean $X_1, X_2, ..., X_d$ variables aleatorias. Definimos la función de distribución conjunta de $X_1, X_2, ..., X_d$ o la función de distribución del vector $X = (X_1, X_2, ..., X_d)$ como:

$$F(x_1, x_2, ..., x_d) = P[X_1 \leq x_1, ..., X_d \leq x_d],$$

donde $(x_1, x_2, ..., x_d) \in \Re^d$.

Definición A.9. (a) Si $X_1, X_2, ..., X_d$ son variables aleatorias discretas, definimos a su función de probabilidad conjunta como:

$$f(x_1, x_2, ..., x_d) = P[X_1 = x_1, X_2 = x_2, ..., X_d = x_d].$$

(*b*) Decimos que $X_1, X_2, ..., X_d$ tienen distribución continua conjunta si existe una función f no negativa definida en $\Re^d$ tal que para cualquier $A \subset \Re^d$,

$$\begin{aligned} P[X \in A] &= P[(X_1, X_2, ..., X_d) \in A] \\ &= \int \cdots \int_A f(x_1, x_2, ..., x_d)\, dx_1 dx_2 \cdots dx_d. \end{aligned}$$

A f se le llama función de densidad conjunta.

Definición A.10. Decimos que las variables aleatorias $X_1, X_2, ..., X_d$ son independientes si

$$P[X_1 \in A_1, X_2 \in A_2, ..., X_d \in A_d] = \prod_{i=1}^{d} P[X_i \in A_i], \quad A_i \subset \Re,\ i = 1, 2, ..., d.$$

En general si $X^{(1)}, X^{(2)}, ..., X^{(n)}$ son vectores aleatorios de dimensión d, decimos que son independientes si

$$P\left[X^{(1)} \in A_1, X^{(2)} \in A_2, ..., X^{(n)} \in A_n\right] = \prod_{i=1}^{n} P\left[X^{(i)} \in A_i\right],$$

con $A_i \subset \Re^d,\ i = 1, 2, ..., n$.

Si $X_1, X_2, ..., X_d$ son variables aleatorias independientes, entonces,

$$F(x_1, x_2, ..., x_d) = F_{X_1}(x_1) F_{X_2}(x_2) \cdots F_{X_d}(x_d)$$

y

$$f(x_1, x_2, ..., x_d) = f_{X_1}(x_1) f_{X_2}(x_2) \cdots f_{X_d}(x_d)$$

donde F_{X_i} y f_{X_i} son la función de distribución y densidad de la variable aleatoria X_i, $i = 1, 2, ..., d$, respectivamente.

Ejemplo A.11. Sean $X_1, X_2, ..., X_d$ variables aleatorias independientes tal que $X_i \sim U(a_i, b_i)$, $i = 1, 2, ..., d$. Entonces la densidad conjunta de las variables aleatorias o lo que es lo mismo, la densidad del vector aleatorio $X = (X_1, X_2, ..., X_d)$ es

$$f(x_1, x_2, ..., x_d) = \begin{cases} \frac{1}{\lambda(R^d)} & si\ (x_1, x_2, ..., x_d) \in R^d \\ \\ 0 & en\ otro\ caso \end{cases}$$

donde $R^d := (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_d, b_d)$ y $\lambda\left(R^d\right) := \prod_{i=1}^{d} (b_i - a_i)$. En este caso decimos que X tiene distribución uniforme en R^d.

Definición A.12. (a) La esperanza o valor esperado de una variable aleatoria discreta se define como

$$\mu_X := E[X] = \sum_{x \in R_X} x\, f(x)$$

donde f es la función de probabilidad de X.
(b) Si X es una variable aleatoria continua, la esperanza se define como:

$$\mu_X = \int_{R_X} x\, f(x)\, dx,$$

donde f es la función de densidad de X.

Propiedades de la esperanza

Sean c una constante, $X, X_1, X_2, ..., X_d$, variables aleatorias y r una función de variable real.

A.12.1. $E[c] = c$

A.12.2. $E[cX] = cE[X]$

A.12.3. $E\left[\sum_{i=1}^{d} X_i\right] = \sum_{i=1}^{d} E[X_i]$

A.12.4. Si X es una variable aleatoria discreta

$$E[r(X)] = \sum_{x \in R_X} r(x) f(x)$$

donde f es la función de probabilidad; y si X es una variable aleatoria continua

$$E[r(X)] = \int_{R_X} r(x) f(x) \, dx$$

donde f es la función de densidad.

A.12.5. $(E|X|)^2 \leq E(X^2)$

Definición A.13. La varianza de una variable aleatoria X se define como:

$$\sigma_X^2 := Var(X) = E(X - E[X])^2,$$

de aquí

$$\sigma_X^2 = E\left[X^2\right] - (E[X])^2.$$

Propiedades de la varianza

Sean c una constante y $X_1, X_2, ..., X_d$ variables aleatorias

A.13.1. $Var[X] \geq 0$

A.13.2. $Var[c] = 0$

A.13.3. $Var[cX] = c^2 Var[X]$

A.13.4. Si $X_1, X_2, ..., X_d$ son independientes,

$$Var\left[\sum_{i=1}^{d} X_i\right] = \sum_{i=1}^{d} Var\left[X_i\right].$$

Definición A.14. La desviación estandar de una variable aleatoria se define como

$$\sigma = \sqrt{Var\left[X\right]}.$$

Ejemplo A.15. Si $X \sim U\left(a, b\right)$, entonces

$$\mu_X = \frac{b-a}{2}, \qquad \sigma_X^2 = \frac{(b-a)^2}{12}$$

Definición A.16. (a) decimos que $X_1, X_2, ..., X_n$ es una muestra aleatoria de tamaño n si las $X_i's$ son variables aleatorias independientes con la misma distribución.
(b) La media muestral de los $X_1, X_2, ..., X_n$ se define como

$$\bar{X}_n = \frac{X_1 + X_2 + ... + X_n}{n}$$

Teorema A.17. Supóngase que $X_1, X_2, ..., X_n$ constituyen una muestra aleatoria con media μ y sea $\bar{X}_n$ la media muestral. Entonces:
(a) (Ley Débil de los Grandes Números)

$$\lim_{n\to\infty} P\left[\left|\bar{X}_n - \mu\right| < \varepsilon\right] = 1,\ \varepsilon > 0,$$

(b) (Ley Fuerte de los Grandes Números)

$$P\left[\lim_{n\to\infty} \bar{X}_n = \mu\right] = 1.$$

Teorema A.18. Teorema del Límite Central. Si las variables aleatorias $X_1, X_2, ..., X_n$ constituyen una muestra aleatoria de tamaño n de una distribución con media μ y varianza $0 < \sigma^2 < \infty$, entonces para cualquier número fijo x,

$$\lim_{n\to\infty} P\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right] = \Phi(x),$$

donde $S_n := \sum_{i=1}^{n} X_i$ y Φ es la función de distribución normal estandar.

Apéndice B

Programas

Los programas utilizados fueron compilados con la versión 5.10 del FORTRAN y se utilizaron las librerías de la IMSL en la generación de los números aleatorios, específicamente la subrutina RNUN.

Programa 1

```
C     Programa principal y subprogramas para la aproximación de
C     integrales múltiples en regiones acotadas.

C     La lectura de valores iniciales se hace desde
C     un archivo de datos.

      DIMENSION  N_NALE(5)
      REAL A(5),B(5),FX(10000),X(10000),X01(10000)
      REAL CONF(5),INV_FI(5),ERR(5)
      EXTERNAL APROX_IN,RNUN

      INTEGER M,NALE,K_FI
      PARAMETER(M=5,NALE=10000,K_FI=5)
```

```
C     Lectura de valores iniciales.

      OPEN(7,STATUS='OLD',FILE='VALINI.DAT')
      DO I=1,M
      READ(7,*)A(I),B(I)
      ENDDO
      READ(7,*)(CONF(I),I=1,K_FI)
      READ(7,*)(INV_FI(I),I=1,K_FI)
      READ(7,*)ER_DES

C     Generación de vectores aleatorios.

      DO I=1,M
      CALL RNUN(NALE,X01)
      DO J=1,NALE
      X(I,J)=(B(I)-A(I))*X01(J)+A(I)
      ENDDO
      ENDDO

C     Cálculo de la aproximación a la integral.
      CALL APROX_IN(M,NALE,A,B,X,FX,SUMFXI,VCUBO,APROIN)

C     Cálculo del error.

      CALL ERROR(K_FI,NALE,SUMFXI,VCUBO,FX,INV_FI,
      ER_DES,ERR,N_NALE)

C     Impresión de resultados.

      OPEN(3,FILE='CON')
      WRITE(3,10)APROIN,(CONF(I),I=1,K_FI),NALE,
      (ERR(I),I=1,K_FI),ER_DES,(N_NALE(I),I=1,K_FI)
   10 FORMAT(2X,'LA APROXIMACION A LA INTEGRAL ES',F12.8,///,2X,
      'CONFIANZA (1-ALFA)',12X,5(F4.2,10X),//,2X,'ERROR CON ',I6,
      'NODOS',4X,5(E12.5,2X),//,2X,'No. DE NODOS CON ERROR',
      F6.4,2X,5(I8,6X))

      END



      SUBROUTINE APROX_IN(M,NALE,A,B,X,FX,SUMFXI,VCUBO,APROIN)
      REAL A(5),B(5),V(5),FX(10000),X(5,10000)

      REAL APROIN

      REAL FXI
      EXTERNAL FXI
      REAL SUMFXI,VCUBO
```

```
      SUMFXI=0.0
      VCUBO=1.0

C     Generación de los vectores aleatorios.

      DO J=1,NALE
      DO I=1,M
      V(I)=X(I,J)
      ENDDO

C     Evaluación de la función en cada vector aleatorio
C     y la sumatoria de f(x).

      FX(J)=FXI(V)
      SUMFXI=SUMFXI+FX(J)
      ENDDO

C     Cálculo del volumen del d-rectángulo.

      DO I=1,M
      VCUBO=VCUBO*(B(I)-A(I))
      ENDDO

C     Cálculo de la aproximación a la integral.
      APROIN=VCUBO*SUMFXI/NALE

      RETURN
      END



      SUBROUTINE ERROR(K_FI,NALE,SUMFXI,VCUBO,FX,INV_FI,ER_DES,
      ERR,N_NALE)

      DIMENSION N_NALE(5)
      REAL FX(10000),INV_FI(5),ERR(5)

      REAL SUMFXI,VCUBO,ER_DES

      REAL SUM
      SUM=0.0

C     Estimación de la desviación estándar (S).

      DO I=1,NALE
      SUM=SUM+(FX(I)-(SUMFXI/NALE))**2
      ENDDO
      S=SQRT(SUM/(NALE-1))
C     Cálculo del error dado un número de nodos y
```

```
C     del número de nodos dado un error deseado,
C     ambos para distintos niveles de confianza.

      DO I=1,K_FI
      ERR(I)=S*VCUBO*INV_FI(I)/SQRT(FLOAT(NALE))
      N_NALE(I)=(S*VCUBO*INV_FI(I)/ER_DES)**2+1
      ENDDO

      RETURN
      END



      REAL FUNCTION FXI(X)

      REAL X(5)

      FXI=X(1)*X(2)**2+X(3)*X(5)+X(4)**2*X(2)

      RETURN
      END
```

Programa 2

```
C     Programa principal y subprogramas para la aproximación de
C     integrales múltiples en regiones irregulares acotadas.

C     La lectura de valores iniciales se hace desde
C     un archivo de datos.

      DIMENSION N_NALE(5)
      REAL A(2),B(2),X(2,10000),X01(10000)

      INTEGER M,NALE
      PARAMETER(M=2,NALE=10000)

C     Lectura de valores iniciales.

      OPEN(7,STATUS='OLD',FILE='VALINI.DAT')
      DO I=1,M
      READ(7,*)A(I),B(I)
      ENDDO

C     Generación de vectores aleatorios.

      DO I=1,M
      CALL RNUN(NALE,X01)
      DO J=1,NALE
      X(I,J)=(B(I)-A(I))*X01(J)+A(I)
      ENDDO
      ENDDO

C     Cálculo de la aproximación a la integral.
      CALL APROX_IN(M,NALE,A,B,X,APROIN)

C     Cálculo del error.

      CALL ERROR(K_FI,NALE,SUMFXI,VCUBO,FX,INV_FI,ER_DES,
      ERR,N_NALE)

C     Impresión de resultados.

      OPEN(3,FILE='CON')
      WRITE(3,10)APROIN
   10 FORMAT(2X,'LA APROXIMACION A LA INTEGRAL ES',F12.8)

      END
```

```
      SUBROUTINE APROX_IN(M,NALE,A,B,X,APROIN)

      REAL A(2),B(2),V(2),FX(10000),X(2,10000)

      REAL APROIN

      REAL FXI
      EXTERNAL FXI

      REAL SUMFXI,VCUBO
      SUMFXI=0.0
      VCUBO=1.0

C     Generación de los vectores aleatorios.
      DO J=1,NALE
      DO I=1,M
      V(I)=X(I,J)
      ENDDO

C     Decidir si el vector está en la región de integración
C     y evaluar.
      IF((V(1)-0.5**2+(V(2)-0.5**2.LE.0.25)THEN
      FX=FXI(V)
      ELSE
      FX=0.0
      ENDIF
      SUMFXI=SUMFXI+FX
      ENDDO

C     Cálculo del volumen del d-rectángulo.
      DO I=1,M
      VCUBO=VCUBO*(B(I)-A(I))
      ENDDO

C     Cálculo de la aproximación a la integral.
      APROIN=VCUBO*SUMFXI/NALE

      RETURN
      END



      REAL FUNCTION FXI(X)

      REAL X(2)
      FXI=5

      RETURN
      END
```

Programa 3

```
C     Programa principal y subprogramas para aproximar el
C     valor de una integral impropia de A a infinito.

C     Los valores iniciales se leen a través de una
C     subrutina.

      REAL X01(100000)

C     Tomar valores referentes a la integral.
      CALL VALINI(A,NALE)

C     Generar números aleatorios.
      CALL RNUN(NALE,X01)

C     Cálculo de la aproximación a la integral.
      CALL XI(A,NALE,X01,APROIN)

C     Impresión de resultados.

      OPEN(3,FILE='CON')
      WRITE(3,10)APROIN
   10 FORMAT('LA APROXIMACION A LA INTEGRAL ES',F12.8)

      END



      SUBROUTINE XI(A,NALE,X01,APROIN)
      REAL X01(NALE)

      REAL A,APROIN,X

      REAL FXI

      REAL SUMFXI
      SUMFXI=0.0

C     Cálculo de números aleatorios en (a,infinito),
C     la función evaluada en ellos y su sumatoria.

      DO 10 I=1,NALE
      X=-LOG(X01(I))+A
      SUMFXI=SUMFXI+FXI(X)/EXP(A-X)
   10 CONTINUE
```

```
C     Cálculo de la aproximaión a la integral.
      APROIN=SUMFXI/NALE

      RETURN
      END



      SUBROUTINE VALINI(A,NALE)

      REAL A
      INTEGER NALE

C     Intervalo de integración (a,infinito).
      A=2.0

C     Cantidad de números aleatorios a generar.
      NALE=100000

C     Escribe la función a integrar en el subprograma FXI.

      RETURN
      END



      REAL FUNCTION FXI(T)
      REAL T

      FXI=1/(T-1)**2

      RETURN
      END
```

Programa 4

```
C     Programa principal y subprogramas para aproximar
C     integrales de -infinito a infinito.

C     La lectura de los datos se hace mediante una
C     subrutina.

      REAL X01(12)
      PARAMETER(PI=3.141593)

C     Lectura de valores iniciales
      CALL VALINI(K,NALE)

C     Cálculo de números aleatorios en (-infinito,infinito),
C     y la aproximación a la integral.

      DO I=1,NALE

C     Generar K números aleatorios en (0,1).
      CALL RNUN(K,X01)

      SUM=0.0
      DO J=1,K
      SUM=SUM+X01(J)
      ENDDO
      XI=SUM-6
      FX=(FXI(XI)/(EXP(-0.5*XI**2)/SQRT(2.0*PI)))
      SUMFXI=SUMFXI+FX
      ENDDO
      APROIN=SUMFXI/NALE
      WRITE(*,*)'LA APROXIMACION A LA INTEGRAL ES',APROIN

      END



      SUBROUTINE VALINI(K,NALE)

C     Parámetro K (mayor que 10)
      K=12

C     Números aleatorios en (-infinito,infinito)
      NALE=1000000

      RETURN
      END
```

```
REAL FUNCTION FXI(T)

REAL T

FXI=T*EXP(-(T**2))

RETURN
END
```

Programa 5

```
C     Programa principal y subprogramas para aproximar el
C     valor de una serie.

C     Los valores iniciales los toma de una subrutina.

      REAL X01(100000)
      INTEGER A
      REAL P,Q,APROSE

C     Tomar valores referentes a la serie.
      CALL VALINI(A,P,Q,NALE)

C     Generar números aleatorios en (0,1).
      CALL RNUN(NALE,X01)

C     Cálculo de la aproximación a la serie.
      CALL XI(A,P,Q,NALE,X01,APROSE)

C     Impresión de resultados.

      OPEN(3,FILE='CON')
      WRITE(3,10)APROSE
   10 FORMAT('LA APROXIMACION A LA SERIE ES',F11.8)

      END



      SUBROUTINE XI(A,P,Q,NALE,X01,APROSE)
      REAL X01(NALE)

      REAL APROSE
      INTEGER A,X

      REAL FXI

      REAL SUMFXI
      SUMFXI=0.0

C     Cálculo de naturales aleatorios en (a,infinito),
C     la función evaluada en ellos y su sumatoria.

      DO 10 I=1,NALE
      X=LOG(X01(I))/LOG(Q)+A
      SUMFXI=SUMFXI+(FXI(X)/(P*Q**(X-A)))
```

```
   10 CONTINUE

C     Cálculo de la aproximación a la serie.
      APROSE=SUMFXI/NALE

      RETURN
      END



      SUBROUTINE VALINI(A,P,Q,NALE)

      INTEGER NALE

C     Intervalo de integración (a,infinito).

      A=2

C     Cantidad de números aleatorios a generar.
      NALE=100000

C     Parámetros p y q.
      P=1.0/2.0
      Q=1.0/2.0

C     Escribe la función a integrar en el subprograma FXI.

      RETURN
      END



      REAL FUNCTION FXI(T)
      INTEGER T

      FXI=1/(T*LOG10(T)**2)

      RETURN
      END
```

Bibliografía

[1] P. J. Davis / P. Rabinowitz, *Methods of numerical integration*, Academic Press, INC (1984).

[2] T. M. R. Ellis, *Fortran 77 programming with an introduction to the Fortran 90 standard*, Addison-Wesley Publishing Company (1990).

[3] C. F. Gerald, *Análisis numérico*, Representaciones y Servicios de Ingeniería (1987).

[4] D. B. Hernández / L. J. Alvarez, *El método de Monte Carlo*, Publicación del Departamento de Matemáticas y Física de la Universidad Autónoma Metropolitana-Iztapalapa (1984).

[5] IMSL, *Stat/Library*, Vol. 3(1987).

[6] W. J. Kennedy, Jr / J. E. Gentle, *Statistical computing*, Marcel Dekker, INC (1980).

[7] D. E. Knuth, *The art of computer programming*, Vol. 2: *Seminumerical algorithms*, Addison-Wesley Publishing Company (1969).

[8] J. H. Maindonald, *statistical computation*, John Wiley & Sons (1984).

[9] T. Naylor / J. Balintey / D. Burdick / K. Chu, *Técnicas de simulación en computadoras*, Limusa (1988).

[10] H. Niederreiter, *Random numbers generation and quasi-Monte carlo methods*, Regional Conference Series in Aplied Mathematics #63, CBMS-NSF (1992).

[11] R. Y. Rubinstein, *Simulation and the Monte Carlo method*, John Wiley & Sons (1981).

[12] I. M. Sóbol, *Método de Monte Carlo*, Lecciones Populares de Matemáticas, MIR (1983).

[13] S. Yakowitz, *Computational probability and simulation*, Addison-Wesley Publishing Company (1977).