

Curso:
Métodos de Monte Carlo.
Unidad 2, Sesión 3: Estimación de volúmenes

Departamento de Investigación Operativa
Instituto de Computación, Facultad de Ingeniería
Universidad de la República, Montevideo, Uruguay

dictado semestre 1 - 2016

Contenido:

1. Introducción y definiciones.
2. Estimación de volúmenes.
3. Ejemplo.
4. Ejercicio.

Monte Carlo para estimar volúmenes e integrales

Esta unidad introduce los problemas básicos que surgen al aplicar el método de Monte Carlo para resolver el conjunto de problemas encontrados con mayor frecuencia en el área del cálculo numérico.

En su forma más simple, el problema consiste en evaluar el volumen de una región acotada en un espacio euclidiano multi-dimensional.

Una formulación más general consiste en la evaluación de la integral de una función en una región de este estilo.

El método de Monte Carlo con frecuencia es una manera competitiva de resolver este problema (más aún, muchas veces es la única forma práctica para lograrlo).

Las ventajas del método surgen cuando la forma de la región de interés hace imposible resolver el problema por métodos analíticos, y, en el caso general de la integración de funciones, cuando las propiedades de suavidad

y de variación del integrando son desconocidas o impiden la aplicación de técnicas de evaluación numérica clásicas.

Definiciones

Esta unidad se centra en los problemas relacionados a la obtención de *estimadores puntuales* y de *estimadores de intervalo* bajo forma de *intervalos de confianza*.

- *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.

Estimación de volúmenes

- Sea \mathcal{R} una región 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), supondremos que es posible transformarla para que quede incluida en \mathcal{J}^m .
- Supongamos que \mathcal{R} está definida por un conjunto de desigualdades y de relaciones implícitas entre las variables espaciales $0 \leq x_i \leq 1$, $i = 1, \dots, m$, y que la forma de \mathcal{R} es tal que el cálculo exacto de $\lambda(\mathcal{R})$ es intratable.
- Si disponemos de un procedimiento que genera una secuencia de puntos $\mathcal{X}_{m,n} = \{\mathbf{x}^{(j)} = (x_1^{(j)}, \dots, x_m^{(j)}) \in \mathcal{J}^m; j = 1, \dots, n\}$, es posible aproximar $\lambda(\mathcal{R})$ a través de $\bar{\lambda}(\mathcal{R})$ calculada por el algoritmo que sigue:

Procedimiento Volumen (region \mathcal{R} , entero n)

Entrada: region \mathcal{R} , n tamaño de la muestra

Salida: $\bar{\lambda}(\mathcal{R})$

1. $S = 0$; /* Inicialización */
2. For $j = 1, \dots, n$ do
 - 2.1 Generar $\mathbf{x}^{(j)}$ de $\mathcal{X}_{m,n}$;
 - 2.2 If $\mathbf{x}^{(j)} \in \mathcal{R}$ then $\phi(\mathbf{x}^{(j)}) = 1$ else $\phi(\mathbf{x}^{(j)}) = 0$;
 - 2.3 $S = S + \phi(\mathbf{x}^{(j)})$; /* Acumular */
3. $\bar{\lambda}(\mathcal{R}) = S/n$;

- La exactitud de $\bar{\lambda}(\mathcal{R})$ dependerá directamente de las propiedades de \mathcal{R} y de $\mathcal{X}_{m,n}$. Por ejemplo, una forma sencilla es generar n puntos de evaluación en la malla m – *dimensional*:

$$\mathcal{X}_{m,n} = \{\mathbf{x} = (x_1, \dots, x_m) : x_i = (z_i + 1/2)/k; z_i = 0, 1, \dots, k-1; i = 1, \dots, m, n = k^m\}$$

- En este caso, cada uno de los n puntos en $\mathcal{X}_{m,n}$ es el centro de un hipercubo de m dimensiones y volumen $1/k^m = 1/n$, y de esta forma S/n , el volumen total de cada uno de los S m -cubos con centros en \mathcal{R} , será una aproximación de $\lambda(\mathcal{R})$.

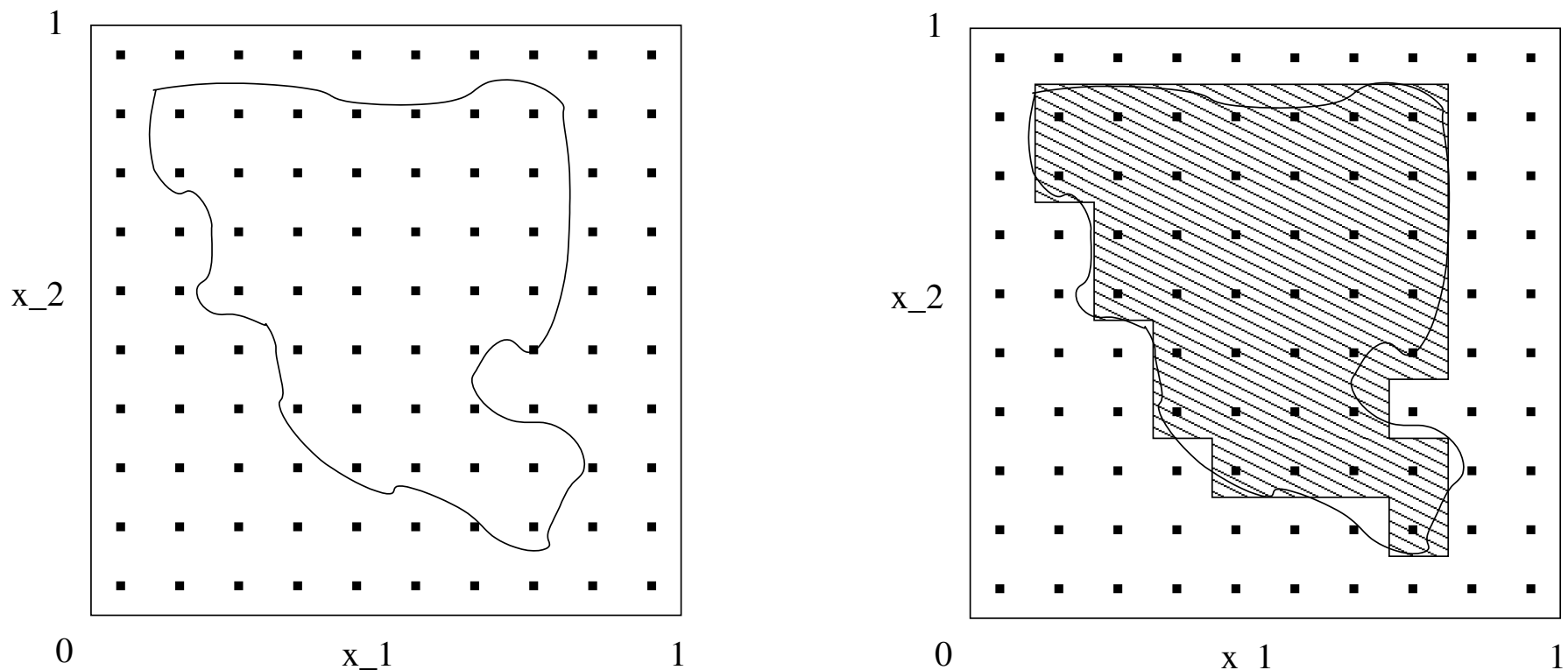


Figure 1: Ejemplo de una superficie a estimar, y la estimación a partir de la malla 2-dimensional con $k = 10$.

- La figura 1 presenta un ejemplo en 2 dimensiones (en el que el volumen es en realidad una superficie), con la malla correspondiente y la estimación que surge de la misma (figura a la derecha). Es claro que se cometen errores de dos tipos, por un lado 2-cubos (cuadrados) que están parcialmente fuera de la figura, por otro partes de la figura que no son cubiertas por ningún 2-cubo. El error total, $\bar{\lambda}(\mathcal{R}) - \lambda(\mathcal{R})$, es la diferencia de estos dos errores.
- En el caso 2-dimensional, sea $s(\mathcal{R})$ el largo de la frontera de \mathcal{R} . Entonces el error absoluto tiene la siguiente cota superior:

$$|\bar{\lambda}(\mathcal{R}) - \lambda(\mathcal{R})| \leq s(\mathcal{R})/k.$$

Este error surge de considerar que los errores pueden agruparse en un rectángulo que tiene, en el peor caso, largo $s(\mathcal{R})$, y ancho $1/k$ (el ancho de cada 2-cubo de la malla), y de usar el área de este rectángulo como cota del error en el peor caso.

- En general, si $s(\mathcal{R})$ es la “superficie” (en $m - 1$ dimensiones) de \mathcal{R} , la

misma fórmula sigue siendo válida:

$$|\bar{\lambda}(\mathcal{R}) - \lambda(\mathcal{R})| \leq s(\mathcal{R})/k.$$

Dado que $n = k^m$, podemos expresar esta misma fórmula así:

$$|\bar{\lambda}(\mathcal{R}) - \lambda(\mathcal{R})| \leq s(\mathcal{R})/n^{1/m},$$

lo que demuestra que la tasa de convergencia del error disminuye al aumentar la dimensión m del problema. En particular, para garantizar un error absoluto de a lo sumo $\epsilon \in (0, 1)$, es necesario tomar una malla (de dimensión m) con al menos $n(m, \epsilon) = \lceil [s(\mathcal{R})/\epsilon]^m \rceil$ puntos (notación: $\lceil x \rceil = \text{ceiling}(x)$, menor entero mayor o igual a x).

- Por lo tanto, si tuviéramos una secuencia de regiones de dimensión creciente e igual volumen, para obtener igual precisión ϵ necesitaríamos un número de puntos que crece exponencialmente con la dimensión m .

Estimación de volumen por Monte Carlo

- La forma más sencilla de aplicar Monte Carlo para resolver este problema consiste en usar el mismo algoritmo anterior, con la única modificación consistente en sustituir el empleo de la secuencia determinística $\mathcal{X}_{m,n} = (\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(n)})$ por una muestra aleatoria independiente $(\mathbf{X}^{(1)}, \dots, \mathbf{X}^{(n)})$ donde cada $\mathbf{X}^{(j)} = (X_1^{(j)}, \dots, X_m^{(j)})$ se sortea de acuerdo a la distribución uniforme en \mathcal{J}^m .
- Si $\mathbf{X} = X_1, \dots, X_m$ es uno de estos puntos, su p.d.f. tiene la forma siguiente (donde $x = (x_1, \dots, x_m)$):

$$f_X(x) = \begin{cases} 1 & \text{si } 0 \leq x_i \leq 1 \text{ para todo } i = 1, \dots, m, \\ 0 & \text{si no.} \end{cases}$$

Esto corresponde a tener distribuciones $U(0, 1)$ independientes para cada una de las componentes del vector \mathbf{X} ; por lo tanto, para sortear un

valor del mismo, alcanza con sortear de manera uniforme cada uno de los valores X_i .

- El siguiente pseudocódigo describe este método de Monte Carlo.

Procedimiento MonteCarlo-Volumen (region \mathcal{R} , entero n)

Entrada: region \mathcal{R} , n tamaño de la muestra,

Salida: $\bar{\lambda}(\mathcal{R})$ estimación del volumen, $V[\bar{\lambda}(\mathcal{R})]$ estimación de la varianza

1. $S = 0$; /* Inicialización */
2. For $j = 1, \dots, n$ do
 - 2.1 Sortear $\mathbf{X}^{(j)}$ con distribución uniforme en \mathcal{J}^m ;
 - 2.2 If $\mathbf{X}^{(j)} \in \mathcal{R}$ then $\phi(\mathbf{X}^{(j)}) = 1$ else $\phi(\mathbf{X}^{(j)}) = 0$;
 - 2.3 $S = S + \phi(\mathbf{X}^{(j)})$; /* Acumular */
3. $\bar{\lambda}(\mathcal{R}) = S/n$; /* Estimador puntual de $\lambda(\mathcal{R})$ */
4. $V[\bar{\lambda}(\mathcal{R})] = (S/n)(1 - S/n)/(n - 1)$; /* Estimador puntual de la varianza $\text{Var}(\bar{\lambda}(\mathcal{R}))$ del estimador puntual $\bar{\lambda}(\mathcal{R})$ */

- Se puede ver que en el pseudocódigo precedente utilizamos una fórmula para estimar la varianza distinta de la discutida en la Sesión 02. Esta nueva fórmula se basa en propiedades de v.a. de Bernoulli (v.a. discretas que toman valores 0 y 1 exclusivamente). Para el problema de estimar volúmenes (y en general para los problemas donde se acumulan v.a. Bernoulli), ambas fórmulas son equivalentes.
- Derivamos la fórmula empleada ahora. Comenzamos observando que $\phi(\mathbf{X}^{(j)})$ es una v.a. que toma dos valores, 0 o 1; la probabilidad de que $\phi(\mathbf{X}^{(j)}) = 1$ es la probabilidad de que $\mathbf{X}^{(j)} \in \mathcal{R}$, que es igual a $\lambda(\mathcal{R})$, y la probabilidad de que valga 0 es su complemento, $1 - \lambda(\mathcal{R})$.
- Por lo tanto, $\phi(\mathbf{X}^{(j)})$ es una v.a. de Bernoulli con parámetro $\lambda(\mathcal{R})$, y su varianza se puede calcular fácilmente, resultando igual a $\lambda(\mathcal{R})(1 - \lambda(\mathcal{R}))$.
- Usando la independencia entre los $\mathbf{X}^{(j)}$, y propiedades de la varianza de una combinación lineal de v.a. independientes, sabemos que el

estimador $\bar{\lambda}(\mathcal{R}) = S/n$ tendrá varianza

$$\begin{aligned}\text{Var}(\bar{\lambda}(\mathcal{R})) &= \frac{1}{n^2} \text{Var}(S) = \frac{1}{n^2} \sum_{j=1}^n \text{Var}(\phi(\mathbf{X}^{(j)})) = \\ &= \frac{1}{n^2} n \lambda(\mathcal{R})(1 - \lambda(\mathcal{R})) = \frac{1}{n} \lambda(\mathcal{R})(1 - \lambda(\mathcal{R})).\end{aligned}$$

- Aunque no conocemos el valor exacto de $\lambda(\mathcal{R})(1 - \lambda(\mathcal{R}))$, podemos estimarlo. Para esto, observamos que

$$\begin{aligned}\mathbb{E}((S/n)(1 - S/n)) &= \mathbb{E}((S/n) - S^2/n^2) = \frac{1}{n} \mathbb{E}(S) - \frac{1}{n^2} \mathbb{E}(S^2) = \\ &= \frac{1}{n} \mathbb{E}(S) - \frac{1}{n^2} (\text{Var}(S) + (\mathbb{E}(S))^2) = \lambda(\mathcal{R}) - \lambda(\mathcal{R})(1 - \lambda(\mathcal{R}))/n - \lambda(\mathcal{R})^2 = \\ &= \frac{(n - 1)\lambda(\mathcal{R}) - (n - 1)\lambda(\mathcal{R})^2}{n} = \frac{(n - 1)\lambda(\mathcal{R})(1 - \lambda(\mathcal{R}))}{n}.\end{aligned}$$

Por lo tanto,

$$\frac{\lambda(\mathcal{R})(1 - \lambda(\mathcal{R}))}{n} = \mathbb{E} \left(\frac{(S/n)(1 - S/n)}{(n - 1)} \right),$$

y entonces $(S/n)(1 - S/n)/(n - 1)$ es un estimador insesgado de la varianza.

Ejemplo

Problema: Supongamos que tenemos un círculo con centro en el punto $(x_1, x_2) = (0.3, 0.2)$ y radio 0.4, y queremos calcular la superficie de la intersección de este círculo con el cuadrante positivo (es decir, $x_1 \geq 0$, $x_2 \geq 0$).

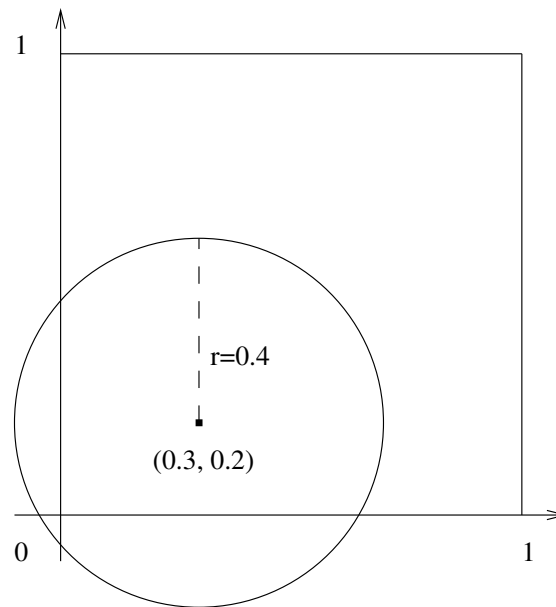


Figure 2: Ejemplo de una superficie a estimar.

Si aplicamos Monte Carlo, sortearemos una cantidad n de puntos independientes y uniformemente distribuidos en $(0, 1) \times (0, 1)$, y usaremos como estimador del área a calcular la cantidad de puntos que cayeron dentro de la misma sobre el total de puntos sorteados.

La siguiente tabla muestra el resultado de $n = 10$ sorteos independientes. Dado que de 10 sorteos sólo 3 puntos cayeron en la figura \mathcal{R} cuya área se desea calcular, nuestra estimación es $\bar{\lambda}(\mathcal{R}) = 0.30$. La estimación de la varianza es $\bar{\lambda}(\mathcal{R})(1 - \bar{\lambda}(\mathcal{R}))/ (n - 1) = (0.3)(0.7)/9 = 0.0233$, y la estimación de la desviación estándar es la raíz cuadrada de este valor, aproximadamente 0.153.

x_1	x_2	Pertenece a \mathcal{R}
0.12	0.94	0
0.85	0.92	0
0.25	0.73	0
0.59	0.43	1
0.45	0.53	1
0.44	0.91	0
0.25	0.20	1
0.09	0.57	0
0.87	0.37	0
0.13	0.73	0

Table 1: Una simulación detallada con $n = 10$ experimentos.

La figura siguiente muestra la ubicación de los 10 puntos sorteados.

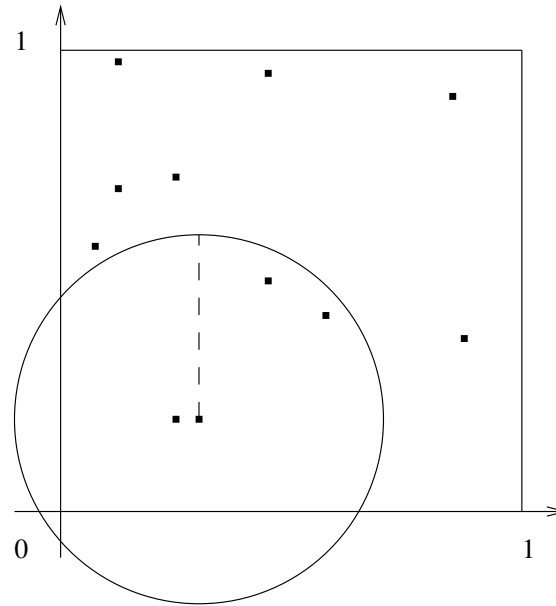


Figure 3: Ejemplo de sorteo de 10 puntos uniformemente distribuidos en $(0, 1) \times (0, 1)$.

Una implementación C de Monte Carlo para resolver este problema se encuentra disponible en la siguiente dirección:

<http://www.fing.edu.uy/inco/cursos/mmc/codigoC/circulo.c>

Dicho código genera puntos independientes y uniformemente distribuidos en $(0, 1) \times (0, 1)$, y controla si su distancia al centro del círculo es menor o mayor que el radio del mismo (lo que equivale a preguntar si están en su interior o en su exterior).

Preguntas para auto-estudio

- ¿Cuál es la diferencia entre un estimador puntual y un estimador de intervalo?
- Si tenemos una secuencia de regiones de igual volumen pero dimensión m creciente, y queremos estimar este volumen con una cierta precisión prefijada ϵ : ¿cuál es el comportamiento computacional de un método de aproximación clásico y de un método Monte Carlo en función de m ?
- Dar el pseudocódigo del Método de Monte Carlo para calcular la medida de una región arbitraria incluida dentro de $\mathcal{J}^m = [0, 1]^m$.
- Explicar cómo se calcula la varianza en el caso de experimentos basados en sumas de variables aleatorias de Bernoulli.

Ejercicio

Problema: se desea estimar el volumen de una región \mathcal{R} de $[0, 1]^6$ definida por todos los puntos de la hiper-esfera de centro $(0.4, 0.45, 0.55, 0.6, 0.5, 0.4)$ y radio 0.4, que además cumplan las restricciones siguientes: $5x_1 + 10x_2 \leq 4$; $x_3 + x_4 \leq 1$; $2x_1 - 3x_2 + x_5 + x_6 \geq 0$.

Entrega 2 - Ejercicio 3.1 (individual)

- Parte a: implementar un programa que reciba como parámetro la cantidad de replicaciones n a realizar, y emplee Monte Carlo para calcular (e imprimir) la estimación del volumen de \mathcal{R} , y la desviación estándar de este estimador. Incluir código para calcular el tiempo de cálculo empleado por el programa. Utilizar el programa con $n = 10^4$ y luego con $n = 10^6$ para estimar el volumen de \mathcal{R} . Discutir si los dos valores obtenidos parecen consistentes.
(en la sesión 5 se continuará este ejercicio).

- Parte b: como forma de validar el programa, eliminar las restricciones adicionales de desigualdad, y comparar el volumen calculado por Monte Carlo con $N = 10^6$ con el valor exacto del volumen de una hiperesfera de dimensión 6, $\pi^3 r^6 / 6$ (<http://www.sjsu.edu/faculty/watkins/ndim.htm>). Discutir también la relación de este valor con el obtenido en la parte a.

Comentario: la hiperesfera de dimensión k , centro (c_1, \dots, c_k) y radio r es el conjunto de puntos que verifican $\sum_{i=1}^k (x_i - c_i)^2 \leq r^2$.

Fecha entrega: ver el calendario en el EVA.