Modelado y Simulación Numérica

Daniel Pérez Palau

Tema 8. Simulación de Monte Carlo

Calendario

	Semana	Tema	Refuerzo	Laboratorio	Actividad
09/11/2020					
16/11/2020	1	S0 + T1			
23/11/2020	2	T2			
30/11/2020	3	T3		L1	
07/12/2020	4	T4			
14/12/2020	5	T5			L1
21/12/2020	--	Semana de repaso	R-L1		
28/12/2020	--	Semana de repaso			
04/01/2021	6	T6 + repaso			
11/01/2021	7	T6			
18/01/2021	8	T7			
25/01/2021	9	T7			AG
01/02/2021	10	T8			
08/02/2021	11	T9		L2	
15/02/2021	12	T10	R-AG1		
22/02/2021	13	T11			L2
01/03/2021	14	Sesión examen	R-L2		
08/03/2021	15	Repaso (sesión doble)			
15/03/2021	16	Semana			

Próximas sesiones
T8-> (03/02 19:00CET)

Contenidos

- Tema 1. Conceptos generales de modelado matemático y simulación
- Tema 2. Modelado matemático de sistemas físicos
- Tema 3. Sistemas físicos y sus modelos
- Tema 4.Simulación
- Tema 5. Generación de números aleatorios
- Tema 6. Generación de variables aleatorias
- Tema 7. Medidas estadísticas
- Tema 8. Simulación de Monte Carlo
- Tema 9. Conceptos y elementos de simulación con eventos
- Tema 10. Modelado y simulación de sistemas de eventos discretos
- Tema 11. *Software* para modelado matemático y simulación

Objetivos

- Conocer los conceptos elementales de la simulación de Monte Carlo

Origenes del método

- El método de Montecarlo es un método no determinista (estadístico numérico)

- Se suele usar para aproximar problemas matemáticos complejos.

- El método de Montecarlo permite obtener **soluciones aproximadas** de una gran variedad de problemas matemáticos:
  - estocásticos o
  - deterministas.

- Para utilizar este método es necesario disponer de una fuente de números aleatorios. Por lo tanto, la aparición de las computadoras benefició el desarrollo de este tipo de técnicas.

Origenes del método

- Origen: Laboratorio de los Álamos (1940) - bomba atómica

- El nombre proviene del Casino de Montecarlo, ya que la ruleta es un generador de números aleatorios simple.

- Simulación de problemas probabilísticos de hidrodinámica:
  difusión de neutrones dentro de material fisible.
  Los neutrones se mueven aleatoriamente durante la difusión.

Origenes del método

- Usualmente se emplean números generados pseudo-aleatoriamente de forma uniforme e IID (Tema 5)
- A partir de estos valores generamos entradas de variables aleatorias (Tema 6)
- Se analizan los resultados mediante técnicas de medidas estadísticas (Tema 7)

Proceso de Simulación de Monte Carlo (I)

1. seleccionar el modelo matemático a utilizar.

2. seleccionar aquellas variables cuyo comportamiento se quiere simular.
   - no deben ser deterministas, si no que deben tener un componente aleatorio, o bien deben ser complicadas de obtener analíticamente.

3. determinar la función de densidad de probabilidad teórica asociada a cada una las v.a. de 2.

Proceso de Simulación de Monte Carlo (I)

4. generar una secuencia de números pseudoaleatorios uniforme en [0, 1].

5. generar las v.a. identificadas con anterioridad.(transformada inversa, aceptación-rechazo, etc…).

6. Los valores generados de las v.a. se introducen en el modelo con el fin de obtener las salidas correspondientes.

7. Repetir el proceso hasta obtener datos suficientes para realizar análisis estadístico.
   - Para ello se utilizan las técnicas de inferencia estadística (intervalos de confianza, test de hipótesis, etc…)

Estimación de variables y tamaño de muestra

- Con el fin de determinar el tamaño muestral de las v.a. generadas (y, por tanto, de las salidas), se pueden utilizar varias técnicas.

- Las más sencilla es analizar los datos de salida y calcular estadísticos muestrales que nos indiquen si el número de muestras analizado es realmente significativo.

- Para ello se calcula la **media** y la **desviación típica** de la salida en un momento dado de la simulación.
    a. Se van **añadiendo muestras** nuevas y calculando de nuevo la media y la desviación típica de las v.a. de salida del modelo y comparándolas con los resultados obtenidos anteriormente.
    b. Cuando la **diferencia entre los estadísticos muestrales** sucesivos está por debajo de una tolerancia determinada, se selecciona el tamaño de muestra actual.

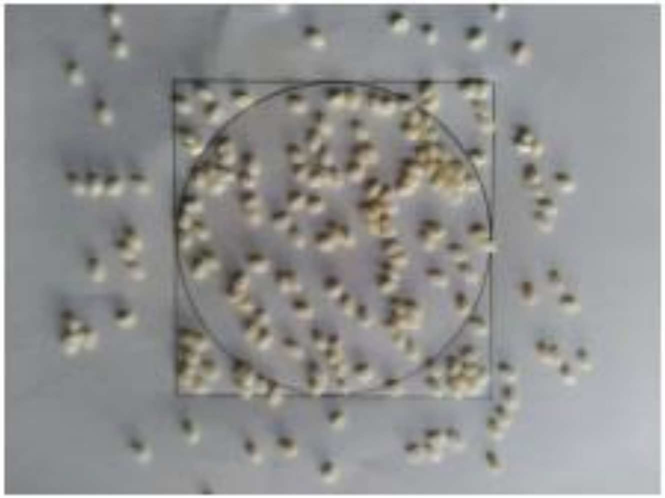
Estimación de variables y tamaño de muestra

- En el método de Montecarlo se puede demostrar que la precisión de los resultados obtenidos es del orden $\frac{1}{\sqrt{n}}$.
- Esto implica que por cada cifra decimal que se quiera ganar en precisión, es necesario aumentar el tamaño de la muestra en un factor 100.
- La Ley de los grandes números nos asegura que cuando el número de muestras tiende a infinito, la media muestral de los datos tiende en probabilidad a la media poblacional. Sin embargo no nos dice cuantos datos hay que añadir para que el error se reduzca una cantidad dada. También sabemos que el tamaño de los intervalos de confianza se reduce al aumentar el tamaño muestral.
- A continuación veremos un ejemplo que nos permitirá determinar por qué la precisión en los cálculos varía de esta manera.

Integración de Monte Carlo

- Supongamos que queremos calcular el área de un círculo de radio R, pero que hemos olvidado la fórmula. Para calcular el área podríamos inscribir dicho círculo en un cuadrado de lado 2R. Sabemos que el área de un cuadrado es el lado al cuadrado, luego el área de nuestro cuadrado será $4R^2$.

- Una forma fácil de calcular el área es cubrir el área del cuadrado y del círculo inscrito con puntos aleatorios, por ejemplo con granos de arroz.



Integración de Monte Carlo

- Es fácil darse cuenta de que el área del cuadrado es proporcional al número de granos de arroz que caen dentro del mismo. Por la misma razón, el área del círculo es proporcional al número de granos de arroz que caigan dentro de él. Por lo tanto, podemos establecer la relación de proporcionalidad siguiente:

$$\frac{A_{circ}}{A_{cuad}} = \frac{N_{circ}}{N_{cuad}}$$

- Siendo A_{circ} y A_{cuad} el área del círculo y del cuadrado respectivamente y N_{circ} y N_{cuad} el número de granos de arroz que caen dentro del círculo y dentro del cuadrado respectivamente. Aquellos que caigan fuera del cuadrado no los tendremos en cuenta.
- Contando los granos de la imagen, obtenemos que N_{circ}=109 y N_{cuad} =144. Puesto que:

$$A_{cuad} = L^2 = (2R)^2 = 4R^2$$

Integración de Monte Carlo

- Sabemos que:

$$A_{circ} = A_{cuad}\frac{N_{circ}}{N_{cuad}} = 4R^2\frac{N_{circ}}{N_{cuad}}$$

- Suponiendo que el radio del círculo es 1, el área del círculo debería ser π.

- Un rápido cálculo nos proporciona que:

$$A_{circ} = 4\frac{N_{circ}}{N_{cuad}} = 4\frac{109}{144} \approx 3{,}0278$$

- Que es una aproximación bastante buena, teniendo en cuenta el método utilizado.

Integración de Monte Carlo

- ¿Podríamos plantearnos cual es la precisión del método utilizado?. Sabemos que el error de nuestros cálculos es:

$$error = |3{,}0278 - \pi| \approx 0{,}1138$$

- También sabemos que el experimento de lanzar granos de arroz sobre el cuadrado y el círculo y determinar si caen dentro o fuera, es un experimento de tipo binomial donde todos los puntos del espacio son equiprobables (recordar que un experimento de este tipo es la ejecución de n experimentos de tipo Bernouilli, en los que el resultado sólo puede ser éxito o fracaso con una probabilidad p o 1-p respectivamente. El objetivo de la distribución binomial es determinar la probabilidad de que en esos n experimentos de tipo Bernouilli haya m éxitos).
- La probabilidad de éxito, vendrá dada por la probabilidad de que un grano de arroz caiga dentro del círculo: $p = \frac{A_{circ}}{A_{cuad}}$

Integración de Monte Carlo

- La desviación típica de la distribución de tipo binomial del problema considerado es:

$$\sigma = \sqrt{N_{cuad}\; p(1-p)}$$

- Tener en cuenta que N_{cuad} es el número de granos totales que consideramos en el experimento (los que caen fuera del cuadrado no los tenemos en cuenta.)
- Sabiendo que

$$p = \frac{A_{circ}}{A_{cuad}} \qquad\qquad A_{circ} = A_{cuad}\frac{N_{circ}}{N_{cuad}}$$

- Y sustituyendo en la desviación típica, obtenemos

$$\sigma = \sqrt{N_{cuad}\; A_{cuad}\frac{N_{circ}}{N_{cuad}A_{cuad}}(1-p)} = \sqrt{N_{circ}(1-p)} = \sqrt{N_{circ} - \frac{N_{circ}^2}{N_{cuad}}}$$

Integración de Monte Carlo

- La desviación típica de la distribución binomial nos proporciona una estimación del error absoluto respecto de la media en el cálculo del número de éxitos.
- Si dividimos la desviación típica por el número total de muestras, obtenemos el error relativo.
- Dicho error relativo para el cálculo realizado será:

$$E_r = \frac{\sigma}{N_{cuad}} = \frac{1}{N_{cuad}}\sqrt{N_{circ} - \frac{N_{circ}^2}{N_{cuad}}}$$

- Supongamos que multiplicamos el número de granos de arroz por 100. Aproximadamente, el número de granos de arroz que caen dentro y fuera del círculo se multiplicarán por 100. El nuevo error relativo será:

$$E_r' = \frac{\sigma'}{100N_{cuad}} = \frac{1}{100N_{cuad}}\sqrt{100N_{circ} - \frac{(100N_{circ})^2}{100N_{cuad}}}$$

Integración de Monte Carlo

- Dividiendo ambos errores relativos y simplificando, obtenemos que

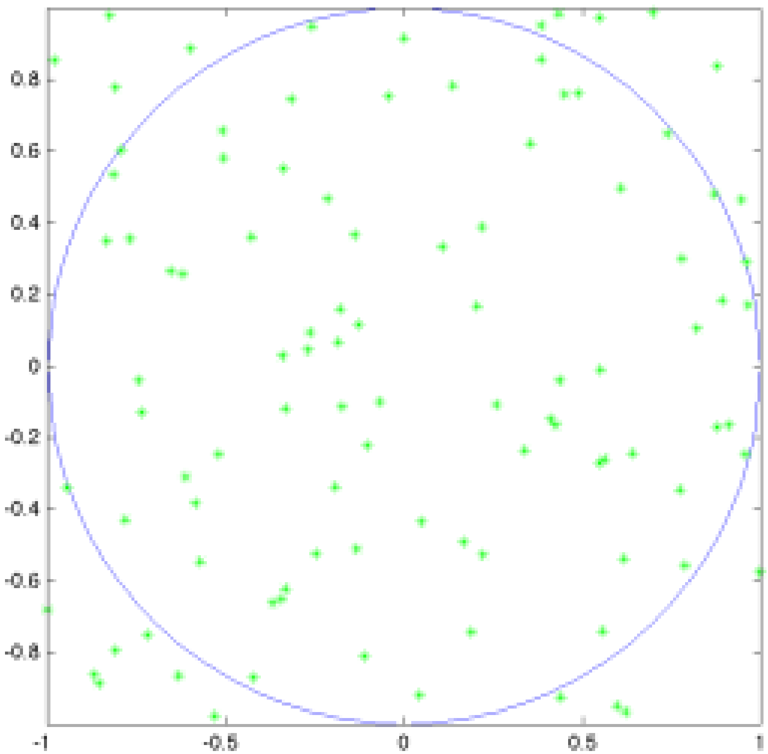
$$\frac{E'_r}{E_r} = \frac{\frac{1}{100N_{cuad}}\sqrt{100N_{circ} - \frac{(100N_{circ})^2}{100N_{cuad}}}}{\frac{1}{N_{cuad}}\sqrt{N_{circ} - \frac{N^2_{circ}}{N_{cuad}}}} = \frac{1}{10}$$

- Con lo que vemos que multiplicar el número de muestras por N, reduce el error relativo (o aumenta la precisión) en una cantidad igual a la raíz cuadrada de N.

- Con lo que la incertidumbre en el cálculo del área mediante el método de Montecarlo depende directamente de la raíz cuadrada del número de muestras.

Integración de Monte Carlo

- Ahora podemos sustituir los granos de arroz por pares de números aleatorios uniformemente distribuidos en el intervalo (-1,1). La distribución espacial de 100 de estos números junto con el cuadrado y el círculo correspondiente es la siguiente:



- Donde podemos ver que el número de puntos dentro del círculo es N_{circ}=77. Por supuesto, N_{cuad}=100.
- Sabemos entonces que el área del círculo será:

$$A_{circ} = A_{cuad}\frac{N_{circ}}{N_{cuad}} = 4\frac{77}{100} \approx 3{,}0800$$

- Pero podemos calcular una estimación del error para dicha área.

Integración de Monte Carlo

- Sabemos que el número medio de aciertos vendrá dado por la media de la distribución binomial:

$$\mu = np = N_{cuad}\frac{\pi R^2}{L^2} = N_{cuad}\frac{\pi}{4} \approx 78{,}5398$$

- El error absoluto vendrá dado por la desviación típica:

$$\sigma = \sqrt{N_{circ} - \frac{N_{circ}^2}{N_{cuad}}} = \sqrt{77 - \frac{77^2}{100}} \approx 4{,}2083$$

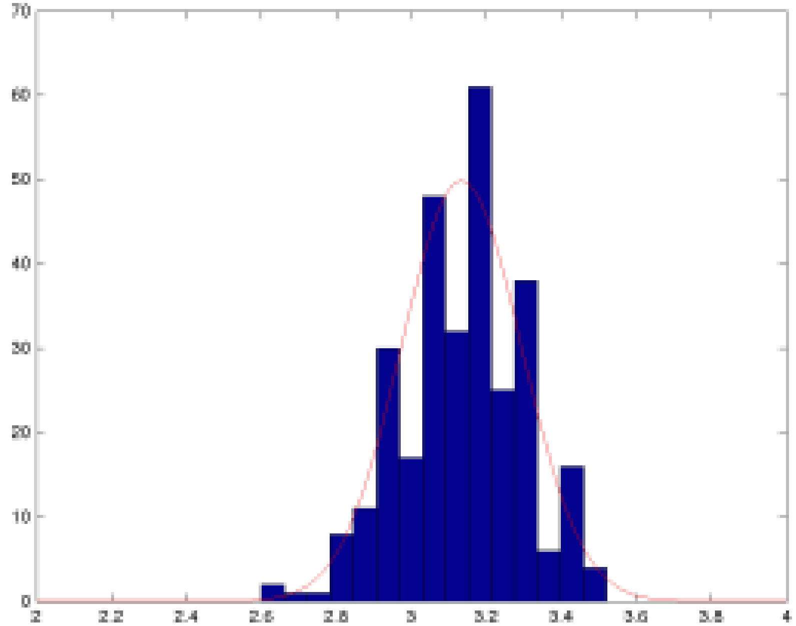
- Luego $N_{circ} \approx (78{,}5398 \pm 4{,}2083)$ y el intervalo:

$$A_{circ} = A_{cuad}\frac{N_{circ}}{N_{cuad}} \approx (78{,}5398 \pm 4{,}2083)\frac{4}{100} \approx (3{,}1416 \pm 0{,}1683)$$

- Debería contener nuestras estimaciones del área del círculo.

Integración de Monte Carlo

- Si ejecutamos 300 veces el mismo experimento y obtenemos la media de las áreas y la desv. típica de esas áreas, obtenemos el histograma siguiente:



- La curva normal se ha dibujado utilizando la media y desv. de las 300 áreas calculadas. Según el teorema de Moivre, la binomial debe tender a la normal cuando n es muy grande. Esto da pié al cálculo de un intervalo de confianza para la media de las áreas.

Otras aplicaciones

El Método de Monte Carlo también se utiliza para:

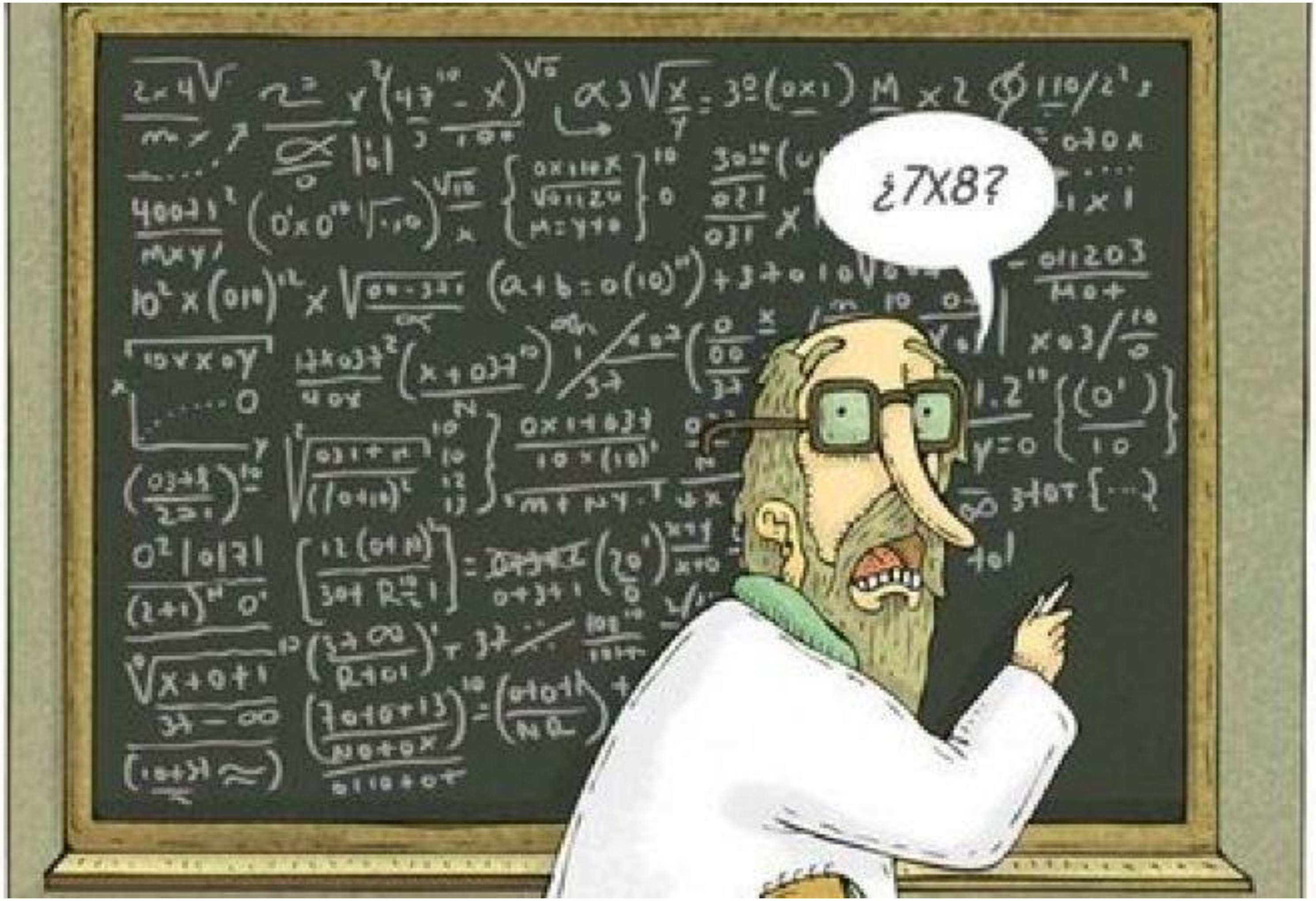
1. El cálculo de integrales multidimensionales o con límites de integración complicados.
   Para 4 o más dimensiones -> precisión elevada

2. Análisis de sensibilidad de parámetros de un modelo

3. Simulación de interacciones de Partículas

4. Resolución aproximada y gestión de la incertidumbre

¿Dudas?



unir

LA UNIVERSIDAD
EN INTERNET

www.unir.net