



14. Técnicas de simulación mediante el método de Montecarlo

- ¿Qué es la simulación?
- Proceso de simulación
- Simulación de eventos discretos
- Números aleatorios
- Análisis estadístico del resultado
- Integración con el método de Montecarlo



¿Qué es la simulación?

- Simulación = técnica que imita el comportamiento de un sistema del mundo real cuando evoluciona en el tiempo.
- Modelo de simulación = conjunto de hipótesis acerca del funcionamiento del sistema expresado como relaciones matemáticas y/o lógicas entre los elementos del sistema.
- Proceso de simulación: ejecución del modelo a través del tiempo en un ordenador para generar muestras representativas del comportamiento del sistema.
- Es una técnica para estimar la bondad de un modelo. NO ES UNA TÉCNICA DE OPTIMIZACIÓN.



¿Qué es la simulación?

- Se basa en un muestreo aleatorio: la salida de la simulación está sujeta a variaciones aleatorias y debe ser examinada utilizando pruebas de inferencia estadística.
- Un sistema es una colección de elementos que actúan e interactúan para lograr algún fin lógico.
- El estado de un sistema es el conjunto de variables necesarias para describir el estado del sistema en un momento dado.



Tipos de simulación

- Un modelo de simulación estática (simulación de Monte Carlo) es una representación de un sistema en un instante de tiempo determinado.
- Una simulación dinámica es una representación de un sistema cuando evoluciona con el tiempo.
- Un modelo de simulación determinista no contiene variables aleatorias.
- Un modelo de simulación estocástica contiene una o más variables aleatorias.



Tipos de simulación

- Modelos continuos: su comportamiento cambia continuamente con el tiempo. Se suelen representar mediante ecuaciones diferenciales que describen las interacciones entre los elementos del sistema
- Modelos discretos: su comportamiento cambia solo en instantes concretos de tiempo: eventos.



Proceso de simulación

- Enunciado explícito de los objetivos que se persiguen: preguntas que se han de contestar, hipótesis que se quiere probar, posibilidades a considerar.
- Creación del modelo y reunión de datos
- Diseñar un programa de ordenador para el modelo
- Verificar el programa
- Validar el modelo
- Utilizar el modelo para experimentar y contestar a las preguntas iniciales.
- Reunir, procesar y analizar los datos generados como soluciones del modelo y en términos de validez y confiabilidad estadística.



Elementos de una simulación

- Salidas: objetivos del estudio expresadas mediante valores numéricos.
- Entradas: valores numéricos que permiten iniciar la simulación y obtener las salidas:
 - Condiciones iniciales: valores que expresan el estado del sistema al principio de la simulación
 - Datos determinísticos: valores conocidos necesarios para realizar los cálculos que producen las salidas
 - Datos probabilísticos: cantidades cuyos valores son inciertos pero necesarios para obtener las salidas de la simulación. Los valores específicos de estos datos deben conocerse a través de una distribución de probabilidad.



Simulación de eventos discretos

- Todas las simulaciones de eventos discretos describen situaciones de colas.
- Cualquier modelo de eventos discretos está formado por una red de colas interrelacionadas.
- Los dos principales eventos son la llegada y la salida.
- Si el intervalo de tiempo entre dos eventos consecutivos es probabilístico surge la aleatoriedad.



Números aleatorios

- Métodos para generar una muestra aleatoria a partir de una distribución de probabilidad $f(t)$ a partir de números aleatorios uniformemente distribuidos en $(0,1)$
 - Método de la inversa: si f es la función de densidad de la distribución se determina $F(x)=P[y \leq x]$ que es una función creciente con valores en $[0,1]$. Si R es un valor aleatorio obtenido de la distribución uniforme en $(0,1)$, $x=F^{-1}(R)$ es el valor deseado.
 - Método de convolución: la idea fundamental es expresar la muestra deseada como la suma estadística de otras variables aleatorias fáciles de muestrear
 - Método de aceptación-rechazo: se reemplaza la función de densidad f difícil de tratar analíticamente por una aproximación más simple h .



Números aleatorios

- Método de aceptación-rechazo: se reemplaza la función de densidad f difícil de tratar analíticamente por una aproximación más simple h .
 - Función mayorante: g tal que $f(x) \leq g(x)$ para todo x
 - Función aproximada: h tal que
$$h(x) = \frac{g(x)}{\int_{-\infty}^{\infty} g(y) dy}$$
 - El método consiste:
 - Obtener una muestra $x=x_1$ a partir de h con el método de la inversa o de la convolución
 - Obtener un número aleatorio R en $(0,1)$
 - Si $R \leq \frac{f(x_1)}{g(x_1)}$ se acepta x_1 como muestra de f , si no, se rechaza x_1 y se vuelve al paso 1.
 - La eficiencia del método depende de lo bien que se ajuste g a f mientras se consiguen h analíticamente manejables.



Generación de números aleatorios

- Para generar números aleatorios en un ordenador para su uso en simulación se utilizan operaciones aritméticas: números pseudoaleatorios.
- Método de la congruencia multiplicativa: dados los parámetros u_0 (semilla), b , c y m un número pseudoaleatorio R_n se genera:
$$u_n = (bu_{n-1} + c) \bmod(m), n = 1, 2, \dots$$
$$R_n = \frac{u_n}{m}, n = 1, 2, \dots$$
- Una selección adecuada de u_0 , b , c y m puede hacer que los números pseudoaleatorios cumplan con las propiedades estadísticas de los números aleatorios



Análisis estadístico del resultado

- Una sola simulación proporciona sólo un valor de entre muchos posibles del resultado, que puede ser distinto en cada simulación.
- ¿Cuántas simulaciones son necesarias para lograr un nivel deseado de confianza en el resultado?
 - Si el resultado es la media de los valores de salida
 - Si el resultado es una proporción o probabilidad
- Las técnicas de inferencia estadística proporcionan respuesta a estas preguntas. Estas técnicas también se deben tener en cuenta a la hora de reunir los datos que sirven como entradas.



Análisis estadístico del resultado

- La confianza en el resultado se expresa como la probabilidad de que el valor real esté en un intervalo centrado en el valor estimado (intervalo de confianza).
- La determinación de los intervalos de confianza en simulación es complicado porque:
 - Las salidas no suelen ser independientes.
 - Las condiciones iniciales pueden influir en los resultados.



Análisis estadístico del resultado

- Creación de un intervalo de confianza para la media verdadera de un resultado de salida.

$$P[\bar{x} - e \leq m \leq \bar{x} + e] = 1 - \alpha$$

- Elegido un valor de α el intervalo de confianza asociado es

$$\bar{x} \pm t_{(n-1), \alpha/2} * \frac{s}{\sqrt{n}}$$

donde n = tamaño de la muestra, $\bar{x}$ = media de la muestra, s = desviación estándar de la muestra, $t_{(n-1), \alpha/2}$ = valor de la distribución t con $n-1$ grados de libertad de modo que $P(T \geq t_{(n-1), \alpha/2}) = \alpha/2$



Análisis estadístico del resultado

- Creación de un intervalo de confianza para una proporción

$$P[p - e \leq P \leq p + e] = 1 - \alpha$$

- Elegido un valor de α el intervalo de confianza asociado es

$$p \pm \left[z_{\alpha/2} * \sqrt{\frac{p(1-p)}{n}} \right]$$

donde n = tamaño de la muestra, p = proporción de la muestra, $z_{\alpha/2}$ = valor de la distribución normal estándar de modo que $P(Z \geq z_{\alpha/2}) = \alpha/2$



Integración con el método de Montecarlo

- Para calcular $I = \int_a^b f(x)dx$
- El valor medio de $f(x)$ en $[a,b]$ es $E(f) = \frac{I}{b-a}$
- Se puede calcular a partir de una muestra aleatoria de puntos $x_1, x_2, \dots, x_n$ en $[a,b]$

$$\hat{E}(f_n) = \frac{1}{n} \sum_{i=1}^n f(x_i)$$

- El valor aproximado para la integral es

$$\int_a^b f(x)dx \approx \frac{b-a}{n} [f(x_1) + \dots + f(x_n)]$$



Integración con el método de Montecarlo

- Ley de los grandes números:
Sean $x_1, x_2, \dots, x_n$ variables aleatorias con función de densidad $\mu(x)$

$$\int_{-\infty}^{\infty} m(x)dx = 1$$

Supongamos que existe $I = \int_{-\infty}^{\infty} f(x)m(x)dx$

Entonces dado $\varepsilon > 0$ se verifica

$$\lim_{n \rightarrow \infty} P[I - \varepsilon \leq \frac{1}{n}[f(x_1) + \dots + f(x_n)] \leq I + \varepsilon] = 1$$
- Ley fuerte de los grandes números

$$P\left[\lim_{n \rightarrow \infty} \frac{1}{n}[f(x_1) + \dots + f(x_n)] = I\right] = 1$$



Integración con el método de Montecarlo

- Teorema Central del Límite:

Sea $I = \int_{-\infty}^{\infty} f(x)m(x)dx$ la media de f y $s^2 = \int_{-\infty}^{\infty} (f(x) - I)^2 m(x)dx$ su varianza.

Entonces

$$P\left[\left|\frac{1}{n} \sum_{i=1}^n f(x_i) - I\right| \leq \frac{ls}{\sqrt{n}}\right] = \frac{1}{2p} \int_{-l}^l e^{-x^2/2} dx + O\left(\frac{1}{\sqrt{n}}\right)$$

En la práctica σ se aproxima por

$$V_s = \frac{1}{n} \sum_{i=1}^n f^2(x_i) - \left(\frac{1}{n} \sum_{i=1}^n f(x_i)\right)^2$$



Integración con el método de Montecarlo

- Los teoremas anteriores sólo son válidos para muestras aleatorias.
- En la práctica las muestras que se obtienen son pseudoaleatorias.
- Los números pseudoaleatorios están determinados de forma cíclica.
- Para que el tratamiento estadístico de los resultados sea válido, es necesario que los números pseudoaleatorios formen parte de ciclos suficientemente grandes.