SIMULACIÓN

Introducción

Índice

1 Introducción

La mejor forma de conocer un sistema es experimentar con él. Al observar las reacciones del sistema ante diferentes entradas se obtiene información del sistema y esto permite adoptar las acciones adecuadas. Sin embargo, cuando se trabaja con sistemas complejos la experimentación o no es posible o si lo es, es cuando menos no recomendable. Los principales motivos para no experimentar con un sistema real son:

Motivos de temporalidad.

Al estudiar las formas de aprovechamiento de un bosque las diferentes políticas de talado no admiten experimentación en un horizonte temporal razonable.

Motivos éticos

Políticas de sanidad

Motivos económicos

Suelen ser los más habituales. Disposición de máquinas, determinación del número de operarios etc.

La otra importante vía de experimentación es la creación de un modelo que represente al sistema y realizar la experimentación sobre el modelo para obtener conclusiones del sistema real.

Los modelos son de diferentes tipos: modelos físicos (maquetas), modelos matemáticos (ecuaciones diferenciales, sistemas lineales, etc.) y modelos lógico-matemáticos (programas de ordenador).

La simulación es una técnica que permite analizar y estudiar sistemas complejos. Consiste en diseñar y construir un modelo, normalmente lógico-matemático y programable, que describa la parte de interés de un sistema real y llevar a cabo experimentos sobre el modelo para aprender el comportamiento del sistema real con el fin de apoyar la toma de decisiones.

En los últimos años el uso de la simulación se ha generalizado gracias a los avances de la informática y actualmente se utiliza en casi todas las áreas de la industria y la ciencia:

- Diseño de sistemas de colas, redes de comunicación, control de inventarios, procesos químicos.
- Estudios de comportamiento del consumidor, determinación de precios, predicción económica.
- Sistemas biomédicos
- Estrategias y tácticas de guerra.
- Estimación de constantes, cálculo de integrales múltiples, inversión de matrices,

2 Ejemplos de simulación

Vamos a ver unos ejemplos donde la aplicación de la simulación presenta ventajas y a la vez se ponen de manifiesto los problemas de esta técnica.

Ejemplos

La simulación Monte Carlo utiliza el muestreo aleatorio para estimar la salida de un experimento

2.- **Problemas de probabilidad**

El caballero de Mére, Antoine Gombaud, planteó a Blaise Pascal el siguiente problema:

¿Qué es más probable, sacar al menos un seis en cuatro tiradas de un dado, o sacar al menos unos seis dobles en veinticuatro tiradas de dos dados?

Resultado teórico

Caso 1: Un dado

Llamo probabilidad de éxito a la probabilidad de sacar un seis, $\Pr(\acute{e}xito) = \frac{1}{6}$, repito el experimento cuatro veces. Sea

$$X \equiv n\acute{u}mero\ de\ \acute{e}xitos\ en\ 4\ intentos, X \sim Bin\left(n = 4, p = \frac{1}{6}\right)$$

La probabilidad pedida es,

$$\Pr(X \geq 1) = 1 - \Pr(X = 0) = 1 - \binom{4}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^4 = 0.5177469$$

Caso 2: Dos dados

Llamo probabilidad de éxito a la probabilidad de sacar uno seis doble, $\Pr(éxito) = \frac{1}{36}$, repito el experimento veinticuatro veces. Sea

$$X \equiv número\ de\ éxitos\ en\ 24\ intentos, X \sim Bin\left(n = 4, p = \frac{1}{36}\right)$$

La probabilidad pedida es,

$$\Pr(X \geq 1) = 1 - \Pr(X = 0) = 1 - \binom{24}{0}\left(\frac{1}{36}\right)^0\left(\frac{35}{36}\right)^{24} = 0.4914039$$

Resultado con simulación

Primero creamos el algoritmo "lanzar_un_dado"

```
Generar u ~ U(0,1)
  Si u <1/6, exito = TRUE m=m+1
Devuelve exito
```

Caso 1: Un dado

Algoritmo "Jugar_con_un_dado"

```
Repetir desde i = 1 hasta 4
  Si lanzar un dado = éxito, m[i] = 1
  Otrosi                     m[i] = 0
Si suma(m) > 0, ganar = 1,
Otrosi          ganar = 0.
Devolver ganar
```

Algoritmo "Jugar_con_dos_dado"

```
Repetir desde i = 1 hasta 24
  Si lanzar un dado = éxito Y lanzar un dado = éxito, m[i] = 1
  Otrosi                                             m[i] = 0
Si suma(m) > 0, ganar = 1,
Otrosi          ganar = 0.
Devolver ganar
```

Cálculo de la probabilidad de ganar con un dado

```
Repetir desde j = 1 hasta Nsim
```

```
  Si jugar_con_un_dado > 1,  n = n +1
Probabilidad = n/Nsim
```

Cálculo de la probabilidad de ganar con dos dado

```
Repetir desde j = 1 hasta Nsim
  Si jugar_con_dos_dado > 1,  n = n +1
Probabilidad = n/Nsim
```

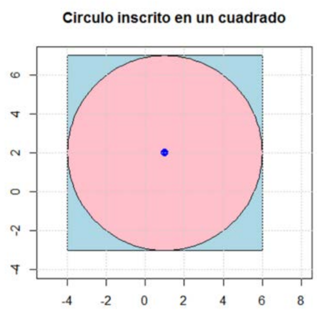
2.- **Método de Monte Carlo. Problemas de áreas**

El método de Monte Carlo se refiere a una clase de métodos estadísticos que utilizan números aleatorias para resolver problemas no probabilísticos.

En este ejemplo se desea calcular el área de un círculo de ecuación

$$(x-1)^2 + (y-2)^2 \leq 25$$

En este caso el resultado teórico es fácil de obtener: $A = \pi r^2$. El radio del circulo es de 5 unidades y por lo tanto A = 78,539816. Para obtener el área por medio de la simulación Monte Carlo encerramos el círculo dentro de un cuadrado



Suponiendo que todos los puntos del cuadrado son equiprobables, si se generan **n** puntos dentro del cuadrado la proporción de puntos que caen dentro del círculo, **m**, dará la proporción del área del cuadrado ocupada por el circulo.

El área del cuadrado es $C = 100\ u^2$, una estimación del área del círculo.

$$(\text{Área del circulo}) \approx \frac{m}{n}(\text{Área del cuadrado})$$

Un punto del cuadrado es un par ordenado (x, y), donde

$$x \in [-4, 6], y \in [-3, 7]$$

Suponiendo que u es una realización de una variable aleatoria $\mathcal{U}(0,1)$, entonces,

$x = -4 + 10 \cdot u$ es una realización de una $\mathcal{U}(-4, 6)$

$y = -3 + 10 \cdot u$ es una realización de una $\mathcal{U}(-3, 7)$.

Como obtener una secuencia de valores $\{u_i\}_{i=1}^{n}$ provenientes de una $\mathcal{U}(0,1)$ es el problema que hay que resolver.

Un modelo de algoritmo puede ser:

```
Repetir desde i hasta n
  Generar x ~ U(-4,6)
  Generar y ~ U(-3,7)
  Si (x-1)**2 + (y-2)**2 <= 25, entonces m=m+1
FinRepetir
Area=(m/n)*100
```

Algoritmo 1: **Área del círculo**

2.- **Simulación de Sucesos-Discretos**. Supongamos una estación de ITV con un único mecánico. Al llegar un automovilista es atendido inmediatamente si la estación está vacía. En caso contrario pasa a formar una línea de espera o cola. Al terminar un servicio el vehículo que primero llego pasa a ser atendido (política de servicio FIFO first in – first out).

Este tipo de situación se da cuando un servicio es atendido por un único *operario*: una tienda, una consulta médica, un banco, etc. atendido por una única persona.

Un modelo teórico que responde a esta situación es proporcionado por la teoría de colas. Según las hipótesis que se realicen respecto a cómo llegan los clientes y como se les sirve aparecen los modelos M/M/1 o el más general G/G/1.

Supongamos que los tiempos entre llegadas son independientes e idénticamente distribuidos según una distribución exponencial de parámetro $\lambda = 4$ (la primera M) y los tiempos de servicio son independientes e idénticamente distribuidas según una distribución exponencial de parámetro $\mu = 7$ (la segunda M). Se quiere conocer:

1. La probabilidad que el sistema este vacío. (P_0)
2. Número medio de personas que hay en el sistema. (L)

3. Número medio de personas que hay en la cola. (Lq)
4. Tiempo medio de permanencia en el sistema. (W)
5. Tiempo medio de permanencia en la cola. (Wq)

El método analítico proporciona las siguientes formulas,

Sea $\rho = \frac{\lambda}{\mu}$, el factor de utilización, entonces,

$$\pi_0 = 1 - \rho, L = \frac{\rho}{1-\rho}, L_q = \frac{\rho^2}{1-\rho}, W = \frac{L}{\lambda}, W_q = \frac{L_q}{\lambda}$$

estos resultados son ciertos si las distribuciones son las anteriores. Si cambiamos las hipótesis habrá que obtener otros resultados y según las condiciones esto no es siempre posible.

Empleando la simulación la hipótesis de las distribuciones no es crítica. Primero hay que desarrollar un modelo del sistema. Una primera aproximación puede ser,

```
Repetir desde i hasta n
     Generar Tiempo_entre_llegadas  ~ Exp(4)
     Tiempo_de_llegada
     Generar Tiempo_servicio ~ Exp(7)
Fin
Calcular los valores pedidos
```

Algoritmo 2

En este modelo está por determinar cómo se obtiene los valores pedidos, en general estarán en el bucle principal.

3.- **Problemas de fiabilidad**. Dado un sistema con dos componentes en serie se quiere estimar el tiempo medio antes del fallo del sistema. *Comparar los métodos analíticos, numéricos y de simulación.*

Supongamos un sistema de dos componentes (C_1 y C_2) que trabajan en serie, por ejemplo, las luces de un árbol de Navidad, unas válvulas de seguridad en un sistema de oleoductos, etc. El sistema funciona mientras funcionen las dos componentes y el sistema falla cuando lo hace una de sus componentes.

La medida del tiempo de fallo para cada componente lo recogemos en la variable aleatoria Y_i, $i = 1, 2$. De esta forma el tiempo hasta que el sistema falle es también una variable aleatoria definida a través de las variables Y_1, Y_2. Sea $T = min\,(Y_1, Y_2)$. Normalmente una de las medidas de interés es el tiempo medio hasta el fallo del sistema, $E[T]$.

El Cálculo de Probabilidades dice que,

$$E[T] = \int_{\Omega} \mathrm{t\ dF}(t)$$

el cálculo de esta integral depende del valor de $dF(t)$, donde $F(t)$ es la función de distribución de la variable aleatoria T. Para este ejemplo de simulación supondremos las siguientes hipótesis,

Y_1 variable aleatoria que mide el tiempo de fallo de C1 y sigue una $Exp(\lambda_1)$

Y_2 variable aleatoria que mide el tiempo de fallo de C1 y sigue una $Exp(\lambda_2)$

Ambas variables aleatorias son independientes. Recordemos que la función de densidad de una $Exp(\lambda)$ es:

$$f(x) = \lambda e^{-\lambda x},\ \ x \geq 0 \quad y \quad F(x) = 1 - e^{-\lambda x},\ \ x \geq 0$$

El cálculo de la $E[T]$ lo podemos realizar,

a.-

$$\begin{aligned} P(T > t) &= P(Y_1 > t, Y_2 > t) \\ &= P(Y_1 > t) \cdot P(Y_2 > t) \\ &= e^{-\lambda_1 t} \cdot e^{-\lambda_2 t} \\ &= e^{-(\lambda_1 + \lambda_2) t} \end{aligned}$$

De lo anterior, la variable aleatoria T sigue una distribución $Exp(\lambda_1 + \lambda_2)$.

Entonces,

$$E[T] = \frac{1}{\lambda_1 + \lambda_2}$$

si suponemos que $\lambda_1 = 2$, y $\lambda_2 = 2$, $E[T] = 0{,}1428571428571$.

b.- Utilizando técnicas del análisis numérico. Realizando el cambio,

$$Y_i = -\frac{\ln(1-u_i)}{\lambda_i}$$

obtenemos la integral,

$$E[T] = \iint_{[0,1]} min\left[-\frac{ln(1-u_1)}{\lambda_1}, \frac{ln(1-u_2)}{\lambda_2}\right] du_1 du_2$$

El cálculo de $E[T]$ se puede realizar por el método de los trapecios.

c.- Simulación

Creamos un modelo lógico matemático que describa el sistema y permita recoger la información necesaria para estimar $E[T]$.

Una forma de escribir el modelo puede ser,

```
TiempodeFallo=0
Desde j=1 hasta n
Desde i=1 hasta 2
     Generar Y1~Exp(2)
     Generar Y2~Exp(5)
     TiempodeFallo=TiempodeFallo+min(Y1,Y2)
FinDesde
FinDesde
Esperanza(TiempodeFallo)=TiempodeFallo/n
```

Algoritmo 3

Una vez vistos estos tres enfoques veamos sus ventajas. El método analítico es quizás el más elegante de los tres, pero requiere que las variables aleatorias sean independientes y se distribuyan según una determinada función de distribución. El empleo de análisis numérico no necesita de la hipótesis de independencia, pero, si en vez de 2 componentes fuesen **m** la aproximación sería cada vez más ineficaz. Por último, la simulación no requiere la independencia de las variables y el aumento del número de componentes no le afecta en la misma medida que a los métodos de análisis numérico.

Si lo que se modifica es la hipótesis de la distribución de los tiempos de fallo de las componentes el efecto es más acusado. Por otra parte, si la distribución del tiempo de fallo de las componentes es empírica, únicamente la simulación es aplicable.

3 Notación

Para fijar los conceptos que intervienen en un proceso de simulación daremos las definiciones más importantes.

Sistema

Un sistema es un conjunto de entidades que interaccionan con un fin.

Estado

El estado de un sistema es el conjunto de *variables necesarias* para describir la condición del sistema en un instante dado.

Sistema discreto

Un sistema es discreto si las variables de estado cambian únicamente en instantes discretos de tiempo.

Sistema continuo

Un sistema es continuo si las variables de estado cambian de forma continua a través del tiempo.

Los procesos de simulación se clasifican en *Simulación estática* (simulación de Monte Carlo), representa un sistema en un instante determinado del tiempo. *Simulación dinámica,* representa la evolución temporal de un sistema a través del tiempo. También se pueden clasificar en *simulación determinística* y *simulación estocástica*, según sea la naturaleza del modelo.

4 Ventajas e inconvenientes de la simulación

De lo visto hasta ahora se deduce que los modelos de simulación no tienen una solución como la de una ecuación diferencial, o un problema de programación lineal. El objetivo del modelo de simulación es conocer el comportamiento del sistema según unas condiciones definidas al establecer el modelo, no resolverlo.

El uso de la simulación es apropiado cuando (Shannon 1975):

a) No existe una formulación matemática que se ajuste al sistema o si existe no tiene métodos analíticos de resolución
b) Existen los modelos y los métodos, pero resulta más sencilla o económica la simulación.
c) Es recomendable experimentar con el sistema antes de construirlo por razones económicas o de seguridad.
d) No se puede experimentar con el sistema por razones prácticas o éticas.

Las desventajas que presenta la simulación básicamente son tres:

a) La construcción de un modelo de simulación puede ser caro y requerir mucho tiempo
b) El modelo creado no deja de ser un modelo particular que depende del diseñador por lo que puede no reflejar *alguna de las características relevantes del sistema y dar resultados imprecisos o falsos.*
c) El grado de fiabilidad de un modelo de simulación es difícil de saber a priori.

RECORDAR

- Cada situación requiere un análisis particular.
- La simulación no puede llegar a establecer deducciones lógicas, ya que nunca la experimentación puede probar implicaciones absolutas (Berger)

5 El proceso de simulación

En el proceso de simulación se puede distinguir varios pasos:

1. Formular el problema
2. Reunir datos y crear un modelo
3. Programar el modelo
4. Comprobar el programa
5. Validar el modelo de simulación
6. Diseñar el experimento
7. Ejecutar la simulación
8. Documentar el modelo

Al centrarnos en los aspectos más técnicos de la simulación se pueden sustituir los 8 pasos anteriores por estos 4 siguientes:

1. Obtener unas entradas de un generador de números aleatorios.
2. Transformar las entradas básicas para ser entradas al modelo.
3. Alimentar al modelo con estas entradas y recoger las salidas.
4. Recoger las estadísticas necesarias para la toma de decisiones a partir de las salidas.

este esquema se puede aplicar en la mayor parte de los casos, aunque no es rígido. Estas cuatro etapas corresponden a los pasos 6 y 7 del esquema general. En el desarrollo de los ejemplos siguientes identificaremos los cuatro pasos.

Ejemplos.

1.- Una determinada empresa tiene que decidir entre perforar un pozo de petróleo en el punto A, o en el punto B. Los estudios preliminares que tiene la empresa acerca de la profundidad a la que se encuentra el petróleo son respectivamente l_1 y l_2 (unidades de longitud) con $l_1 < l_2$. Además las características del terreno en el punto A hacen que la perforación sea mucho más lenta en el punto A que en el punto B; de forma más precisa supondremos que en el punto A se espera avanzar en la perforación a razón de x_1 unidades de longitud por jornada de trabajo, siendo la previsión, en el punto B, de x_2 unidades ($x_1 < x_2$).

La empresa espera recibir, antes de comenzar las prospecciones, un nuevo tipo de maquinaria que utilizaría en caso de recibirla. La probabilidad de que esto ocurra estima que será de 0.71. Además, la probabilidad de que se produzca una avería es de 0.14 para la nueva maquinaria en el punto A, y de 0.16 en el punto B. De la misma forma se estima que la probabilidad de avería para la maquinaria vieja es de 0.28 en el punto A y de 0.19 en el punto B.

Las averías se han tipificado en dos clases en función del tiempo necesario para realizar las reparaciones, denominándose averías grandes a aquellas cuyos tiempos medios de reparación son de 4 jornadas de trabajo y averías pequeñas a aquellas cuyo tiempo medio de reparación es de 1 jornada. Si la avería se produce, la probabilidad de que sea de las tipificadas como grandes es de 0.35.

La empresa, obviamente, trata de decidir si perforar en el punto A o en el punto B, supuesto que su objetivo es utilizar el menor tiempo posible hasta poner en funcionamiento el pozo petrolífero.

¿Qué decisión recomendarías, perforar en A o en B?

2.-A Manolo, su mejor amigo Antonio le propone el siguiente juego: lanzar una moneda al aire hasta que la diferencia entre el número de caras y el de cruces sea igual a tres. Manolo debe pagar a Antonio 100 € por cada lanzamiento que haga y recibirá 1000 € cuando la diferencia entre el número de caras y el de cruces (número de cruces - número de caras) sea igual a tres. ¿Cómo puede utilizar Manolo la simulación para comprobar cómo es su mejor amigo?