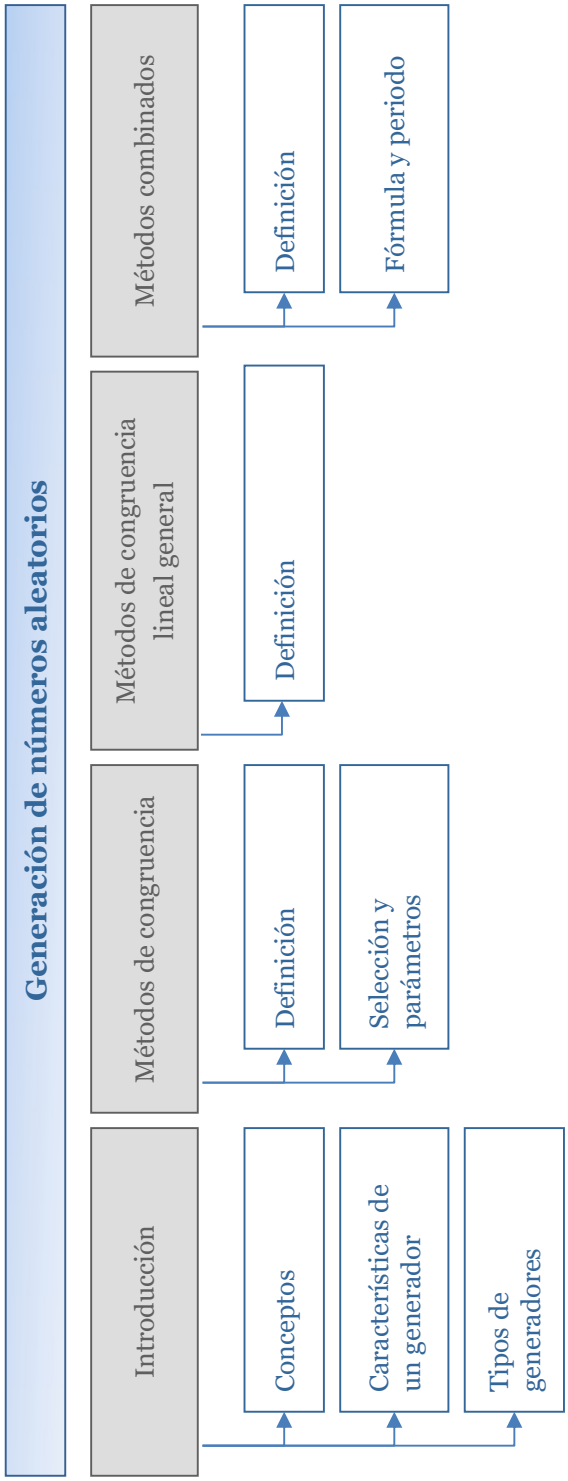
Generación de números aleatorios

TEMA 5

Esquema



 © Universidad Internacional de La Rioja (UNIR)

Ideas clave

5.1. ¿Cómo estudiar este tema?

Para estudiar este tema lee las **Ideas clave** que encontrarás a continuación.

Los números aleatorios son elementos clave para poder construir modelos de simulación que representen modelos reales y su comportamiento.

Las ideas claves de este tema son:

- Concepto de número aleatorio y pseudo-aleatorio.
- Generadores de números aleatorios y sus características.
- Tipos de generadores.
- Métodos de generación de números aleatorios.
- Métodos de congruencia: multiplicativos y mixtos.
- Métodos de congruencia lineal generalizada.
- Métodos combinados.

5.2. Introducción. Generadores de números aleatorios

Los números aleatorios se usan con distintos propósitos pero, especialmente, en los modelos matemáticos con el objeto de poder representar datos de los modelos reales que no se conocen con exactitud. Se necesita por tanto poder predecir sus valores ya sea mediante la inferencia a partir de datos estadísticos o usando modelos de simulación que permiten obtener aproximaciones mejores a la realidad y que utilizan datos que se generan de forma aleatoria. Podría decirse que el uso de números aleatorios es imprescindible en la simulación.

Se debe tener en cuenta que la aleatoriedad no es una propiedad que posee un número aislado sino que está ligada a una serie o secuencia de números.

 © Universidad Internacional de La Rioja (UNIR)

Existen **distintas formas de generar números aleatorios,** que se han ido perfeccionando con el objeto de que la generación no se vea afectada por tendencias o ciclos de comportamiento que de alguna manera restarían aleatoriedad a las secuencias de números generados. Es decir, se necesita un generador de números aleatorios que tenga, entre otras, las siguientes características:

- Se debe de poder computar de forma eficiente.
- Genera números independientes y uniformemente distribuidos.
- Las secuencias generadas no deben de contener ciclos de repetición de forma frecuente.
- El almacenamiento de los números generados debe ser el menor posible.
- Las series que se generan se deben poder reproducir, es decir, volverlas a generar.
- El periodo de generación no puede ser mayor que la cardinalidad del espacio de estados, es decir, no debe ser grande. El periodo o longitud de ciclo es la longitud de la secuencia que siempre repite el generador.
- El método de generación debe ser lo más rápido posible.

Los métodos que se utilizan para construir un generador de números aleatorios son variados y de diferente naturaleza pero todos ellos carecen de alguna de las características enumeradas.

Tipos de generadores

- **Generadores manuales.** Generalmente se usan métodos manuales en condiciones especiales y monitorizadas, como extracciones de bolas al azar. La principal ventaja es que las series obtenidas son aleatorias, pero el procedimiento es lento, no se pueden reproducir, se requiere mucho tiempo y para grandes cantidades de números, el almacenamiento necesario sería muy grande.

- **Generadores con tablas**. El uso de tablas de números generados conocidos en la que los números aleatorios están ordenados de tal forma que la probabilidad de que un número aparezca en un punto dado de la secuencia es igual a la probabilidad de que aparezca cualquier otro y las combinaciones con el mismo número de dígitos tienen la misma probabilidad de aparecer. También se pueden generar por medio de hojas de cálculo. Aun así, es un método lento y que requerirá una gran cantidad de almacenamiento y generalmente con un comportamiento muy determinista.

 © Universidad Internacional de La Rioja (UNIR)

- **Computación analógica.** Para este tipo se requiere el uso de procesos o fenómenos físicos aleatorios como puede ser el comportamiento que se observa en una corriente eléctrica. En este caso, al igual que en los métodos manuales, las series que se obtiene son aleatorias y en este caso se pueden obtener de forma bastante rápida, pero sin embargo, y como ocurría en el primero de los métodos, las series no son reproducibles.

- **Computación digital.** En este caso se requiere un ordenador sobre el que se ejecuta una función que genera los números a partir de una semilla (un número que se introduce como dato inicial para la generación). La secuencia de números depende de la semilla, es decir, se establece una relación entre la semilla y la secuencia, por lo tanto los números generados no son independientes. Esta forma de generación, sin embargo, requiere poco espacio de almacenamiento, permite la reproducción de las secuencias y genera números de una forma bastante rápida.

De todos los tipos vistos, y a pesar, de no contar con todas las características, el que se utiliza, de forma habitual y generalizada en los modelos de simulación, son los de computación digital.

De hecho, el método más fiable es el uso de programas de computación que implementan algún algoritmo con base matemáticas y determinista. Los números que se obtienen tras la ejecución del algoritmo se asemejan a los valores que toman variables uniformemente distribuidas, pero no son realmente aleatorios. Por tanto, los números generados por esta forma se conocen como números pseudo-aleatorios.

A continuación se verán distintos tipos de generadores de números pseudo-aleatorios. El objetivo de todos ellos será producir una secuencia de números uniforme e independiente que tomen valores entre 0 y 1 y con una longitud de ciclo o periodo suficientemente grande.

 © Universidad Internacional de La Rioja (UNIR)

5.3. Métodos de congruencia

En 1951, Lehmer descubre un método de generación de números aleatorios basado en las características de aleatoriedad de los restos que se obtienen en las divisiones de las potencias sucesivas de un número (Yamada, 1961).

$$x_n = a^n \ mod \ m$$

Con base en esta idea, el método define la forma de obtener una serie de números, cada uno de los cuales se obtiene como función del término anterior, de la siguiente forma:

$$x_n = (ax_{n-1} + c) \ mod \ m$$

Ecuación 1. Fórmula para calcular números por congruencia.

Donde a, m y c son números enteros positivos que se denominan multiplicador, modulo e incremento, respectivamente.

El valor inicial de la secuencia o semilla es x_0 y tanto a como c son menores que el modulo.

La longitud del periodo, que interesa que sea bastante grande, depende de la elección de los valores de los parámetros indicados y del valor de la semilla.

En realidad, en el método de *Lehmer, c* toma el valor de 0 y por eso se conoce como un método lineal multiplicativo. Posteriormente *Thomson* en el año 1958, dedujo un método lineal mixto, en el que el valor de *c* es distinto de 0.

Ambos métodos recursivos son de congruencia y se diferencian fundamentalmente en que, así como el primero genera números más rápido, en el caso del método de *Thomson*, la longitud del periodo es mayor.

 © Universidad Internacional de La Rioja (UNIR)

Selección de los parámetros

La selección inicial de Lehmer fue $a = 23$ y $m = 10^8 + 1$. Sin embargo, se analizaron distintos valores con el objeto de determinar las elecciones mejores de acuerdo con una serie de criterios que se enumeran a continuación:

- Para una computación más eficiente, el módulo m debe ser una potencia de 2.
- m debe ser grande y el periodo nunca debe ser mayor que él.
- Si c es distinto de cero, se obtendrá el periodo máximo si y sólo si:
  - m y c son primos entre sí, es decir, no tienen factores comunes que no sean 1.
  - Cada número primo que es un factor de m es también un factor de $a - 1$.
  - Si m es un múltiplo de 4, $a - 1$ debe ser un múltiplo de 4.
  - Todas las condiciones enumeradas se cumplen si $m = 2k$, $a = 1 + 4b$ y c es impar. Aquí, c, b, y k son enteros positivos.
- Si c es igual a cero, es decir, es un generador multiplicativo
  - El periodo máximo siendo $m = 2^k > 16$ es 2^{k-2} y sólo se alcanza si se cumple que x_0 es impar y $a \ mod \ 8$ es 3 o 5.
  - El periodo máximo siendo $m \neq 2^k$, un número primo es $m - 1$, si y sólo si el multiplicador a es una raíz primitiva del módulo, o lo que es lo mismo, $a^n \ mod \ m \neq 1 \ para \ n = 1, 2, \dots, m - 2$

Con este método, la sucesión de números pseudo-aleatorios se obtiene a partir de:

$$u_n = \frac{x_n}{m} \ con \ n \geq 1 \ .$$

Ecuación 2. Cálculo de números aleatorios a partir de una serie.

Ejemplo 1

Sea el generador $x_n = 3x_{n-1} \ mod \ 2^5$ y una semilla $x_0 = 1$. Según las reglas vistas anteriormente para los generadores multiplicativos, como la semilla es impar y $a \ mod \ 8 = 3$, el período máximo será $2^3 = 8$.

La secuencia de números generada será $3, 9, 27, 17, 19, 25, 11, 1, 3$ .... donde se observa que el periodo es 8.

 © Universidad Internacional de La Rioja (UNIR)

Los números aleatorios obtenidos son 0.2187, 0.28125, 0,84375, 0,53125, 0.59375, 0.78125, 0.34375, 0.03125, 0.2187…

Con la misma semilla, se considera el generador $x_n = 7x_{n-1}\ mod\ 2^5$, en este caso los números generados son $7, 17, 23, 1, 7$ … y como se ve el período no es máximo porque es 4.

Ejemplo 2

El generador es $x_n = 3x_{n-1}\ mod\ 31$ y una semilla $x_0 = 1$. El módulo no es potencia de 2. El multiplicador es raíz primitiva del módulo porque el resto de la división de las potencias de 3 entre 31 siempre es distinto de 1. En este caso el periodo es 30 y la secuencia de números generada es 1, 3, 9, 27, 19, 26, 16, 17, 20, 29, 25, 13, 8, 24, 10, 30, 28, 22, 4, 12, 5, 15, 14, 11, 2, 6, 18, 23, 7, 21, 1, …

5.4. Métodos multiplicativos

Los generadores de congruencia se pueden generalizar y definir recursiones de orden mayor. Estos métodos se definen de forma general como se muestra a continuación:

$$x_n = (a_1x_1 + a_2x_2 + \cdots + a_{n-k}x_{n-k}) mod\ m$$

Ecuación 3. Cálculo de secuencia de números mediante congruencia lineal general.

Donde m y k, son números enteros positivos y los coeficientes a_i toman sus valores entre $-(m-1)$ y $(m-1)$, donde la función que genera los números aleatorios será como ya se ha visto:

$$u_n = \frac{x_n}{m}\ con\ n \geq 1$$

Ecuación 4. Calculo números pseudo-aleatorios.

 © Universidad Internacional de La Rioja (UNIR)

5.5. Métodos mixtos

Ante la necesidad de obtener períodos más largos se puede usar la combinación de generadores lineales multiplicativos. Se trata de sumar números obtenidos de dos o más generadores.

La fórmula para obtener los números de la secuencia es

$$X_n = \left(\sum_{j=1}^{k} (-1)^{j-1} X_{i,j} \right) mod\ m_1$$

Ecuación 5. Calculo de secuencia de números mediante método combinado.

Donde las $X_{i,j}, j = 1, .., k$ son las salidas obtenidas de diferentes generadores de congruencia multiplicativa. A su vez se deberán los números pseudo-aleatorios buscados según la siguiente definición

$$U_n = \begin{cases} \dfrac{X_n}{m_1}\ si\ X_n > 0 \\ \dfrac{m_1 - 1}{m_1}\ si\ X_n = 0 \end{cases}$$

Ecuación 6. Cálculo de los números aleatorios.

El periodo máximo se calcula mediante:

$$P = \frac{(m_1 - 1)(m_2 - 1) \dots (m_k - 1)}{2^{k-1}}$$

Ecuación 7. Cálculo del periodo.

Ejemplo (Raj Jain, Washington University)

Dados los siguientes generadores de congruencia multiplicativos.

$$v_n = 157 v_{n-1}\ mod\ 32363$$
$$t_n = 146 t_{n-1}\ mod\ 31727$$
$$w_n = 142 w_{n-1}\ mod\ 31657$$

El generador combinado se define a partir de los anteriores como:

 © Universidad Internacional de La Rioja (UNIR)

$$X_n = (v_n - t_n + w_n)\ mod\ 32363$$

Y el periodo $P = \frac{32362.31726.31656}{2^2} = 8.1x10^{12}$

5.6. Referencias bibliográficas

Yamada, S. (1961). On the period of pseudo-random numbers generated by Lehmer's congruential method,. In Monographs (Vol. 3, pp. 113–123). Operations Research Soc.of Japan.

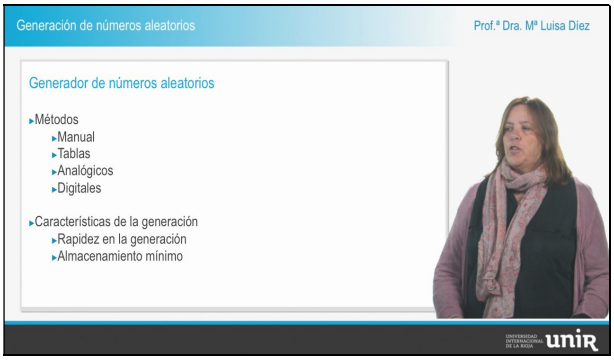
 © Universidad Internacional de La Rioja (UNIR)

Lo + recomendado

Lecciones magistrales

Generación de números aleatorios

En la siguiente lección magistral profundizaremos sobre el generador de números aleatorios.



Accede a la lección magistral a través del aula virtual

No dejes de leer…

Generación de números aleatorios

Capítulo de un libro en el que se detallan los tipos de generadores de números aleatorios, los métodos para su generación y pseudocódigos de los algoritmos que los implementan.

Accede al documento a través del aula virtual o desde la siguiente dirección web:
https://www.academia.edu/11094142/3.-Generación_de_números_aleatorios

 © Universidad Internacional de La Rioja (UNIR)

Random Number Generation

Jain, R. (2008). Random Number Generation. Saint Louis: Washington University.

Apuntes sobre generación de números aleatorios y sus generadores.

Accede al documento a través del aula virtual o desde la siguiente dirección web:
http://www.cs.wustl.edu/~jain/cse567-08/ftp/k_26rng.pdf

 © Universidad Internacional de La Rioja (UNIR)

+ Información

A fondo

The period of pseudo-random numbers generated

Artículo sobre el periodo generado por el método de congruencia de Lehmer.

Accede al artículo a través del aula virtual o desde la siguiente dirección web:
http://comjnl.oxfordjournals.org/content/19/2/173.full.pdf

Enlaces relacionados

Random Number Generators

Página Web de la Universidad de Utah que describe como trabaja un generador de números aleatorios y facilita pautas para elegir un generador y probarlo.

Accede a la página web a través del aula virtual o desde la siguiente dirección web:
http://www.math.utah.edu/~pa/Random/Random.html

 © Universidad Internacional de La Rioja (UNIR)

Bibliografía

Fritzson, P. (2011). *Introduction to Modeling and Simulation of Technical and Physical Systems with Modelica*. Nuevo York: John Wiley & Sons.

García-Dunna, E. (2012). *Simulación y análisis de sistemas con promodel*. Massachusetts: Addison-Wesley.

Shannon, R. y Johannes, J. D. (1976). Systems simulation: the art and science. *IEEE Transactions on Systems, Man and Cybernetics*. *6*(10), 723-724.

Taha, H. A. (2004). *Investigación de operaciones*. New Jersey: Pearson Educación.

Velten, K. (2009). *Mathematical Modeling and Simulation: Introduction for Scientists and Engineers*. Nueva York: Wiley.

Wouwer, A. V., Saucez, P. y Fernández, C. V. (2014). *Simulation of ODE/PDE Models with MATLAB®, OCTAVE and SCILAB: Scientific and Engineering Applications*. Nueva York: Springer.

 © Universidad Internacional de La Rioja (UNIR)

Test

1. En la generación de números aleatorios mediante:

A. Los números generados son independientes.

B. Las secuencias de números generados no se pueden reproducir.

C. Los números aleatorios están ordenados.

D. Ninguna de las anteriores es totalmente correcta.

2. En los métodos de congruencia:

A. La longitud del periodo debe ser lo más pequeña posible.

B. La longitud del periodo depende sólo del valor de la semilla.

C. La longitud del periodo depende sólo del valor de la semilla y de los valores de los parámetros elegidos.

D. Ninguna de las anteriores es verdadera.

3. En los métodos de congruencia:

A. El módulo debe ser cualquier número par.

B. El módulo debe ser menor que el periodo.

C. El módulo debe ser una potencia de dos.

D. Ninguna de las anteriores es verdadera.

4. El parámetro multiplicador de un método de congruencia:

A. Debe ser múltiplo de 4.

B. Debe cumplir que su número anterior sea múltiplo de 4, si el módulo es múltiplo de 2.

C. Debe cumplir que su número anterior sea múltiplo de 4, si el módulo es múltiplo de 4.

D. Ninguna de las anteriores es verdadera.

5. Un método multiplicativo (indique una o más opciones):

A. Es un método mixto de generación de números aleatorios.

B. Es una generalización de un método de congruencia.

C. Se puede combinar con otro para producir un método mixto.

D. Ninguna de las anteriores es verdadera.

 © Universidad Internacional de La Rioja (UNIR)

6. Un método de generación mixto:

A. Multiplica los números obtenidos de varios generadores.

B. Divide los números obtenidos de varios generadores.

C. Suma los números obtenidos de varios generadores.

D. Ninguna de las anteriores es verdadera.

7. Los periodos más largos se obtienen mediante:

A. La combinación de métodos multiplicativos.

B. La disminución del módulo de un método de congruencia.

C. El incremento del valor del multiplicador del método congruencia.

D. Ninguna de las anteriores es verdadera.

8. En un método mixto, el cálculo del periodo máximo depende de:

A. Sólo los módulos de los métodos combinados.

B. Los módulos de los métodos combinados y de los multiplicadores correspondientes.

C. Los módulos y el número de métodos combinados.

D. Ninguna es verdadera.

9. El incremento y el multiplicador de un método de congruencia:

A. Son mayores que el módulo.

B. Son menores que el módulo.

C. Son iguales que la semilla.

D. Ninguna de las anteriores es verdadera.

10. En un método multiplicativo:

A. Los multiplicadores son todos mayores que el módulo.

B. Los multiplicadores son todos menores que el módulo.

C. Los multiplicadores son todos menores que una unidad menos que el módulo.

D. Ninguna de las anteriores es verdadera.

 © Universidad Internacional de La Rioja (UNIR)