Capítulo 5

Generación de Números Aleatorios y Pseudo Aleatorios

Capítulo 5

Generación de Números Aleatorios y Pseudo Aleatorios

Los objetivos que se desean lograr en este capítulo son:

- Comprender la importancia de generar números aleatorios y pseudo aleatorios de forma rápida y confiable.
- Distinguir las características o propiedades de los números aleatorios y pseudo aleatorios.
- Conocer algunas técnicas básicas que generen números pseudo aleatorios.
- Conocer el uso de las pruebas estadísticas para los generadores de números aleatorios.
- Aplicar algunos de estos generadores de números aleatorios en casos prácticos.

5.1 Introducción

Hemos visto que un sistema puede poseer varios comportamientos no determinísticos. Cada comportamiento se convierte en una variable aleatoria. El término *variable aleatoria* se emplea para nombrar una función de valor real, definida sobre un espacio muestral asociado con los resultados de un experimento conceptual de naturaleza no determinística. El resultado particular de un experimento, o sea, el valor numérico o de la muestra de una variable aleatoria se llama *valor o número de la variable aleatoria*. Un número aleatorio es aquel cuya probabilidad de ocurrencia es igual a la de cualquier otro número de la secuencia de valores aleatorios; esto quiere decir que todos los valores o números pueden ocurrir con la misma probabilidad de ocurrencia y poseen entonces un comportamiento uniforme. Se entiende por distribución uniforme continua la probabilidad de que cada punto de un rango determinado sea el mismo, es decir, cualquiera puede ser escogido. Suponiendo que el rango posible de valores va de A a B (B > A), en cuyo caso la probabilidad de que x (variable aleatoria) caiga en un intervalo y es y / (B - A).

Una muestra tiene un comportamiento probabilístico específico (por ejemplo, normal, exponencial, etc.) que posee ciertos parámetros que la describen[1]. Al contar con los parámetros podemos entonces generar los números aleatorios de una distribución probabilística haciendo uso de un algoritmo que se traduce a un programa de computadora. Ahora bien, los números aleatorios generados deben estar uniformemente distribuidos dentro del rango establecido.

No se pierde generalidad si se supone que el rango va desde 0 a 1, porque si x_i es una secuencia de números uniformemente distribuidos en el rango de 0 a 1, (B - A) xi + A es una sucesión distribuida uniformemente en el rango de A a B. De modo que para este caso:

f(x) = 1 0 <= x >= 1
f(x) = 0 en las demás partes.

[1] Lo ideal es contar con *n* muestras y de estas conseguir los parámetros que describan el comportamiento.

Los valores de x en el intervalo unitario se llamarán valores uniformes de las variables aleatorias. Estos se representarán por números que, al menos, parezcan haberse obtenido al azar, como los valores de muestra de una población uniformemente distribuida.

El valor de la variable aleatoria se emplea como un término colectivo, que significa "números aleatorios" dados en forma de dígitos, enteros o números racionales, con un intervalo y un sistema de numeración bien definidos.

Los números aleatorios se usan en computadoras de gran escala en los problemas que implican simulación y modelación. Esto se debe a que en muchos procesos de simulación pareciera que los eventos[2] ocurrieran aleatoriamente o al azar o que implicaron atributos cuyos valores son asignados por algún cambio. En muchos casos la duración de un evento cae dentro de un rango conocido. La simulación de un evento requiere que se asigne un valor en particular.

Para ilustrar lo anterior, consideramos la simulación de una computadora de propósitos generales. Un evento que debe modelarse es la recuperación de un registro de un dispositivo de almacenamiento de acceso directo[3]. La duración de este evento puede determinarse para que caiga dentro de cierto intervalo, sin embargo, el valor actual está influenciado por el cambio de variables, tal como la posición del registro relativo a la cabeza lectora cuando se hace la petición. Esto crea una gama de posibles valores o números aleatorios.

En este capítulo se analizarán las diferentes formas de obtener estos números aleatorios y pseudo aleatorios, se describirán los métodos tradicionales y los analíticos para generarlos, al igual que las pruebas estadísticas que deben superar para ser considerados como tales, y, por último, se presentarán algunos programas que fueron diseñados utilizando los métodos analíticos que se presentan.

5.2 Definición de Números Aleatorios

Un número aleatorio es el resultado de una selección de un número de un grupo de alternativas mediante algún proceso de cambio.

Ejemplo: En el lanzamiento de un dado, el resultado será una de las seis alternativas posibles. Si al tirar el dado, el número que sale es el número cuatro (4), se considera este como el resultado aleatorio. Esto significa que un sólo evento se considera aleatorio, pues si en una docena de veces seguidas, cada vez que lanzamos un dado sale el número cuatro (4) se concluye que el dado está cargado y el resultado no se considera aleatorio.

La notación de "aleatoriedad" es relativa debido a que la aleatoriedad de un evento observado depende de su similitud con otros eventos parecidos y a la vez independencia de estos eventos.

---

[2] Tiempos de entre arribos de clientes o tiempos de servicio de Servidores.

[3] Por ejemplo, un disco magnético.

Una secuencia de números se considera aleatorio si satisface todas las pruebas estadísticas para su aleatoriedad. Estas pruebas implican la investigación de un número de elementos de una secuencia ordenada lo suficientemente grande para verificar:

- Su distribución uniforme, la cual en el lanzamiento del dado no significa que cada uno de los posibles números (1 al 6) ocurre con la misma frecuencia.
- Independencia de los elementos o números, esto significa que, aun conociendo el resultado previo, no se puede predecir el próximo resultado.

Los números verdaderamente aleatorios se generan con el apoyo de la física cuántica o del reuido atmosférico, entre otros. Existen muchas tecnologías disponibles en Internet para generar números verdaderamente aleatorios.

5.3 Definición de Números Pseudo Aleatorios

Los números pseudo aleatorios se generan partiendo de un número semilla que genera siempre los mismos números; esto quiere decir que se pueden reproducir. Esta es la diferencia principal entre un número aleatorio y uno pseudo aleatorio.

En los modelos que se llevan por computadora, existen medios para obtener números pseudo aleatorios uniformemente distribuidos, los cuales se utilizan en la generación de variables pseudo aleatorias con características deseadas. Existen computadoras que tienen subrutinas generadoras de números pseudo aleatorios como parte de su biblioteca de programas. Inclusive, muchos lenguajes de programación y de simulación, hojas electrónicas, entre otros, poseen funciones para generar diferentes tipos de números aleatorios o pseudo aleatorios.

Antes de la llegada de las computadoras, los números pseudo aleatorios eran generados por dispositivos físicos, tales como las ruletas, dados especiales o números mezclados en alguna caja que, se mostraban sacando uno o más con la mano. Esta forma era adecuada si se necesitaban pocos números. Por el gran uso que se les daba a estos números se diseñaron dispositivos electrónicos para una generación más rápida. La Rand Corporation diseñó un dispositivo que usaba un generador de pulsos eléctricos que opera un contador que cicla entre 0 y 9. Haciendo un muestreo el contador o intervalos aleatorios se puede considerar el valor como un dígito del número aleatorio deseado. Repitiendo muchas veces el proceso o corriendo varios contadores en paralelo se puede crear un número aleatorio de cualquier número deseado de dígitos. Se publicó un libro con un millón de números aleatorios generados por este dispositivo.

El uso de estos dispositivos tenía dos dificultades:

- La dificultad de proveer y mantener un dispositivo físico tal que una computadora pudiera recordar un número aleatorio siempre que lo necesitara.
- Los números generados por tales dispositivos no son reproducibles.

Se han publicado tales tablas de números aleatorios distribuidos uniformemente, producidos mediante los procesos físicos antes mencionados. La figura 5.1 es un conjunto de números aleatorios reproducidos de una sola página.

Suponiendo que se decide generar una serie de números aleatorios distribuidos aleatoriamente entre 0 y 1, divididos por 100,000; es decir que los números deben tener cuatro dígitos decimales. Se puede leer la primera columna de la izquierda de fila en fila descartando por ejemplo el último dígito, y al terminar la columna se puede utilizar la siguiente. Usando este método (utilice la figura 5.1), encontramos que los cinco primeros números aleatorios son:

0.3456	0.5514	0.8906	0.5231	0.3933

Si se necesitan números con más de cinco dígitos se pueden combinar las columnas. No es necesario comenzar con la primera columna. De hecho, para dar uso adecuado especialmente cuando se repite un cálculo con distintos conjuntos de números aleatorios, se debe elegir al azar los puntos de inicio, lo que se puede hacer utilizando un número aleatorio para decidir la página, columna y fila de principio.

Podemos hacer que las dificultades antes mencionadas no existan, generando primeramente los números y luego almacenarlas en un arreglo en la memoria de la computadora o en un disco para llamarlos cada vez que se necesiten. Existen computadoras que tienen subrutinas generadoras de números aleatorios como parte de su biblioteca de programas.

Desdichadamente este procedimiento es lento y usa un almacenamiento grande de memoria por consiguiente no resulta práctico para problemas que requieran una gran cantidad de números aleatorios.

Decimos que los números son pseudo aleatorios debido a que a pesar de que los números generados podrían pasar todas las pruebas estadísticas concernientes a su aleatoriedad y distribución, ellos son completamente determinísticos. Quiere decir que si comenzamos cada ejecución del generador con las mismas entradas (constantes y puntos de origen o semillas) obtendremos la misma secuencia de números como salida.

El término *número pseudo aleatorio* ha sido definido por Lehmer como "una noción vaga que encierra la idea de una sucesión en la cual cada término es impredecible, para la persona ajena al problema, cuyos dígitos se someten a cierto número de pruebas, tradicionales a los estadísticos, y depende en cierta forma del uso que se dará a la sucesión".

5.4 Propiedades o Características de los Números Aleatorios y Pseudo Aleatorios

Un generador de números aleatorios o pseudo aleatorios genera números que tienen las siguientes características:

- Uniformemente distribuido: Cualquier número aleatorio tiene igual probabilidad de salir. (Ya que los eventos aleatorios siguen esta distribución).
- Estadísticamente independientes: El valor de un número en una secuencia aleatoria no debe afectar el valor del próximo número. Al hacer un análisis estadístico, se utiliza el concepto de correlación, el cual indica que tan cerca se agrupan los puntos de datos. En este caso el resultado debe ser 0 (no debe existir correlación).
- Reproducibles: Permitiendo que el experimento de simulación se aplique nuevamente. Observación: Esta característica es propia de los generadores de números pseudo aleatorios solamente. En otras palabras, esta característica diferencia a los generadores aleatorios de los pseudo aleatorios.
- Sin repetición: Dentro de una longitud determinada de la sucesión de números.
- Generar números aleatorios a grandes velocidades, disminuyendo así el costo de la corrida de simulación.
- Requerir poca capacidad de almacenamiento en la memoria de la computadora, ya que los modelos de simulación generalmente se traducen en programas muy extensos.

Los criterios antes expuestos se han establecido sobre bases puramente apriorísticas[4].

5.5 Métodos para Generar Números Aleatorios o Pseudo Aleatorios

En la práctica, generalmente se requieren sucesiones de números pseudo aleatorios. En consecuencia, los métodos explicados posteriormente implican algún proceso cuasi aleatorio, en donde se generan sucesiones de números pseudo aleatorios.

Entre los métodos alternativos tenemos:

5.5.1 Métodos Manuales

Son los más sencillos, pero resultan muy lentos para su uso. Su aplicación se ha inclinado más al área pedagógica por su gran atractivo, y sencillez; entre estos tenemos: el empleo de barajas, ánforas, dados y ruletas, como también de monedas y demás artículos semejantes.

5.5.2 Método de Tablas de Números Pseudo Aleatorios

Surge como resultado de la aplicación de los métodos mencionados anteriormente, antes de ser impresos en forma de tabla. A continuación, se presenta una tabla ejemplo:

34567	30173	86551	13424	23464	57564	43352	44535	23232
55140	51455	54446	65464	56456	65646	16115	48414	44465
89064	98439	92392	29424	00022	49229	12010	09802	29212
52312	02020	28282	54832	09392	92929	20919	93093	28209
39332	93933	99001	22200	03033	29292	93939	02292	29292

[4] Con antelación.

83438	03849	09480	38045	39484	30158	79838	17342	35761
98749	29847	93287	09324	97798	79884	97498	98739	47987
94738	49387	98379	47398	79837	87492	79827	02020	02949
49038	30948	09840	79794	29954	36999	50097	87450	54884
40709	45098	49873	47947	83398	49873	48797	93093	39392
38409	74747	93048	93840	30943	25098	34787	40936	47424
97482	94839	98729	48798	98479	74982	94879	47987	98473
83749	84739	40937	40340	23840	40820	82084	83428	43711
98327	98374	87298	47982	49837	82038	83082	72097	38943
47983	49873	84739	74983	49873	47932	30293	12131	30993
34983	94803	40934	50094	93844	34098	09049	34093	09809
40384	98309	09384	98093	79762	87253	25784	23238	93409
36948	38080	35085	34209	40473	87947	79374	70287	79874
34098	09430	43084	98098	93809	83828	01430	49038	09842
93400	84980	48093	08549	98433	43493	48029	40938	42849

Figura 5.1 Tabla de Números Pseudo Aleatorios

La ventaja de este método es que tales números pueden reproducirse una vez que se tenga la tabla, pero como desventaja resulta ser un método lento, también muchas veces se requiere una mayor cantidad de números pseudo aleatorios que los que se han publicado, etc. Este tipo de tabla se puede conseguir en muchos sitios en Internet.

5.5.3 Método de Computación Análoga

En este método se aplica el uso de procesos físicos como el comportamiento de una corriente eléctrica para la producción de números aleatorios. Se dice que es uno de los métodos que conduce a verdaderos números aleatorios, pero como desventaja este utiliza sucesiones de números aleatorios no reproducibles, o sea, no son números pseudo aleatorios.

5.5.4 Método de Computación Digital

Dentro de este método existen 2 subdivisiones:

5.5.4.1 La Provisión Externa

Que trata con la grabación de tablas de números aleatorios, en un disco magnético de la computadora, a fin de tratar los números como datos de entrada para un determinado problema. Estos números también se pueden conseguir a través de Internet.

5.5.4.2 La Generación Interna por Medio de Procesos Físicos Aleatorios

Este aplica un complemento especial de la computadora digital capaz de registrar los resultados de algún proceso aleatorio y, además, reduce los resultados a sucesiones de dígitos. El defecto principal de este es que los resultados no se pueden reproducir, por lo que no es posible comprobar los cálculos efectuados (son valores puramente aleatorios).

5.6 Métodos de Generación de Números Pseudo Aleatorios

A continuación, se presentan diferentes generadores. Muchos de ellos requieren de un valor inicial X_0 que sirve para generar a partir de este, una sucesión de números pseudo aleatorios. Este valor inicial se le llama *semilla* o *punto de origen*.

5.6.1 Método del Cuadrado Medio

Este es un método propuesto por Von Neumann y Metrópolis en 1946 y fue uno de los primeros procedimientos aritméticos utilizados para generar secuencias uniformemente distribuidos. Esta técnica utiliza un número inicial o semilla que es elevado al cuadrado y cada número de una secuencia se produce tomando como siguiente número a los dígitos que están en el medio para luego elevarlo al cuadrado. El algoritmo sigue los siguientes pasos (para secuencias de números de cuatro dígitos):

1. Escoger un número de cuatro[5] dígitos (que será la semilla).
2. Elevar al cuadrado el número y agregar ceros a la izquierda si es necesario para tener un número de ocho dígitos.
3. Seleccionar los cuatros dígitos del medio como el siguiente número aleatorio a utilizarse.
4. Elevar al cuadrado el número de cuatro dígitos seleccionados en el paso 3 (agregarle ceros a la izquierda sí es necesario para tener un número de ocho dígitos).
5. Repetir los pasos 3 y 4 hasta obtener la cantidad de números aleatorios deseadas.

Ejemplo: Tomamos como semilla $X_0 = 2152$. Los números subrayados se toman como números pseudo aleatorios.

$X_0 = 2152 \quad (X_0)^2 = (2152)^2 = 04\underline{6311}04$
$X_1 = 6311 \quad (X_1)^2 = (6311)^2 = 39\underline{8287}21$
$X_2 = 8287 \quad (X_2)^2 = (8287)^2 = 68\underline{6743}69$
$X_3 = 6743 \quad (X_3)^2 = (6743)^2 = 45\underline{4680}49$
$X_4 = 4680 \quad (X_4)^2 = (4680)^2 = 21\underline{9024}00$
$X_5 = 9024$ etc.

Esta técnica es difícil de analizar, relativamente lenta y estadísticamente insatisfactoria. Las secuencias generadas tienen generalmente periodos de repetición pues la relación, entre el número inicial y el largo de las secuencias generadas antes de que empiece a repetirse el mismo periodo, es difícil de analizar anticipadamente. Otro inconveniente es que cada vez que se genera un cero, todos los números subsiguientes serán ceros. Un ejemplo sería si tomamos como semilla al número 44 (dos dígitos):

$X_0 = 44 \quad (X_0)^2 = (44)^2 = 1\underline{93}6$

[5] En sí de ***n*** dígitos.

$X_1 = 93 \qquad (X_1)^2 = (93)^2 = 8649$
$X_2 = 64 \qquad (X_2)^2 = (64)^2 = 4096$
$X_3 = 09 \qquad (X_3)^2 = (09)^2 = 0081$
$X_4 = 08 \qquad (X_4)^2 = (08)^2 = 0064$
$X_5 = 06 \qquad (X_5)^2 = (06)^2 = 0036$
$X_6 = 03 \qquad (X_6)^2 = (03)^2 = 0009$
$X_7 = 00 \qquad (X_7)^2 = (00)^2 = 0000$

Debido a las dificultades del método no se recomienda su uso, sirviendo sólo de interés histórico.

5.6.2 Métodos Congruenciales

El método es un procedimiento aritmético para la generación de una secuencia finita de números uniformemente distribuidos. Pueden usarse varias relaciones recursivas y se han desarrollados muchos métodos congruenciales. Cada uno utiliza la relación fundamental de congruencia:

Dos números enteros A y B se dicen que son congruentes al módulo **m** (m es un entero) sí y solo sí, existe un entero K de tal forma que A - B = K•m. En otras palabras, si (A-B) es divisible por **m** y si A y B dejan residuos idénticos cuando son divididos por el valor absoluto de m. Esta división se expresa así:

$A \equiv B \pmod m$ y se lee "A es congruente con B módulo m".

Ejemplo:

(A - B) / M
(10 - 4) / 3 = 2
A = 10
B = 4 10 / 3 = 3 con residuo 1
M = 3 4 / 3 = 1 con residuo 1

5.6.2.1 Método Congruencial Lineal

En este algoritmo los números sucesivos en la secuencia son generados por la relación de recursión.

$$X_{n+1} = (a\,X_n + c) \bmod m \qquad \text{para} \quad n \geq 0$$

El valor inicial X_0 se conoce como semilla, la constante **a** es el multiplicador, la constante **c** es el incremento y **m** es el módulo. Seleccionar los valores para estas constantes afectan en el largo del período de la secuencia de números aleatorios generados.

Ejemplo:

a = 2
c = 3
m = 10
$X_0 = 0$
$X_1 = (2X_0 + 3) \bmod 10 = (2* 0 + 3) \bmod 10 = 3$
$X_2 = (2X_1 + 3) \bmod 10 = (2* 3 + 3) \bmod 10 = 9$
$X_3 = (2X_2 + 3) \bmod 10 = (2* 9 + 3) \bmod 10 = 1$
$X_4 = (2X_3 + 3) \bmod 10 = (2* 1 + 3) \bmod 10 = 5$
$X_5 = (2X_4 + 3) \bmod 10 = (2* 5 + 3) \bmod 10 = 3$
$X_6 = (2X_5 + 3) \bmod 10 = (2* 3 + 3) \bmod 10 = 9$
$X_7 = (2X_6 + 3) \bmod 10 = (2* 9 + 3) \bmod 10 = 1$
$X_8 = (2X_7 + 3) \bmod 10 = (2* 1 + 3) \bmod 10 = 5$

El módulo **m** y el multiplicador **a** deben ser números positivos. Cuando $c \neq 0$ como en el ejemplo anterior el algoritmo se llama *método congruencial mixto*. Si la constante c = 0 el algoritmo se llama *método congruencial multiplicativo* y se muestra así:

$$X_{n+1} = (a\, X_n) \bmod 10 \quad \text{para} \quad n \leq 0$$

Ejemplo:

a = 2
m = 10
$X_0 = 1$
$X_1 = (2X_0) \bmod 10 = (2 * 1) \bmod 10 = 2$
$X_2 = (2X_1) \bmod 10 = (2 * 2) \bmod 10 = 4$
$X_3 = (2X_2) \bmod 10 = (2 * 4) \bmod 10 = 8$
$X_4 = (2X_3) \bmod 10 = (2 * 8) \bmod 10 = 6$
$X_5 = (2X_4) \bmod 10 = (2 * 6) \bmod 10 = 2$
$X_6 = (2X_5) \bmod 10 = (2 * 2) \bmod 10 = 4$
$X_7 = (2X_6) \bmod 10 = (2 *4) \bmod 10 = 8$
$X_8 = (2X_7) \bmod 10 = (2 * 8) \bmod 10 = 6$

Para ambos métodos, la expresión nos dice que se toma el último número aleatorio X_n, se multiplica por la constante **a** (se suma la constante c si es distinta de cero), y se toma el resultado del módulo **m** (por ejemplo: dividir $aX_n + c$ por **m** y tomar residuo como X_{n+1}). Por consiguiente, para generar una secuencia de números Xn, necesitamos un número inicial o semilla X_0, un multiplicador **a** y un módulo m. Para cualquier generador de números pseudo aleatorios, sólo un número finito de enteros distintos pueden ser generados, después la secuencia es repetitiva por sí misma. El período o largo de la secuencia (P) depende en particular de la computadora y del módulo que se escoja, mientras que las propiedades estadísticas de la secuencia generada dependen de la elección del valor inicial

o semilla y del multiplicador. La elección de a, X_0 y de **m** depende del período máximo y de un mínimo grado de correlación deseado entre los números generados.

La elección apropiada del módulo **m** depende del sistema numérico de la computadora que se usa. La elección más natural de **m** es una que iguale la capacidad de la palabra de la computadora. Para una máquina binaria, **m** es igual a: $m = 2^b$, donde b es el número de bits de la palabra de la computadora. El máximo periodo (obtenible donde **a** y X_0 se escogen apropiadamente), es entonces:

$P = 2^{b-2} = m / 4$ para sistemas binarios con $b > 2$ y
$P = (5)\ 10^{d-2} = m / 20$ para sistemas decimales con $d > 2$.

Este periodo P o el largo de la secuencia es realizado sólo si X_0 y **a** se escogen en cierta forma:

- Para el caso del sistema binario, a se selecciona así:

  $$a = 8T \pm 3$$

  donde T puede ser cualquier entero positivo y X será cualquier entero positivo impar.

- Para el caso del sistema decimal a se selecciona así:

  $$a = 200T \pm Q$$

  donde nuevamente T puede ser cualquier entero positivo y Q debe ser uno de los siguientes valores:

  (3, 11, 13, 19, 21, 27, 29, 37, 53, 59, 61, 67, 69, 77, 83 ó 91).

La semilla **X** puede ser cualquier entero positivo impar que no sea divisible ni por 2 ó 5.

Para satisfacer el requerimiento de una secuencia no correlacionada, debemos además restringir la elección del multiplicador a. Este no debe ser alguna fracción de **m** y tener 5 ó más dígitos que no contengan largas filas de ceros o unos.

Si c es distinto de cero, este debe ser un número primo relativo de m. Escogiendo a **m** como el número primo más largo y menor que 2 ^ b y **a** como raíz primitiva de m, el máximo periodo que puede obtenerse es de m / 4 hasta (m-1).

Resumiendo, el proceso de computación de números aleatorios entre 0 y 1, tenemos:

1. Escoger cualquier número menor de nueve dígitos y designarlos como valor inicial X_0. Esta semilla se escoge aleatoriamente utilizando una tabla de dígitos aleatorios.

2. Multiplicar esto por un número designado como a, por lo menos de cinco dígitos.

3. Multiplicar el producto del paso dos por una fracción o número decimal igual a 1/m. Este se usa para agilizar la operación en la computadora.

4. Escoger la parte decimal del paso tres como un número aleatorio $0 \leq X \leq 1$.

5. Quitar el punto del número obtenido en el paso cuatro y usarlo para que X se multiplique por **a** en el paso dos.

6. Repetir los pasos 2 hasta el 5 hasta obtener el número deseado de números aleatorios.

5.6.2.2 Método Congruencial Aditivo

Este método requiere de una secuencia de números X_1, X_2, X_3 , ..., X_n. Esta secuencia de números puede generarse utilizando cualquiera de las otras técnicas. La aplicación de este algoritmo produce una extensión para una secuencia X_{n+1}, X_{n+2}, X_{n+3}, ... El algoritmo se muestra así:

$$X_j = (X_{j-1} + Xj\text{-}n) \bmod m$$

La principal ventaja de esta técnica es la velocidad puesto que no es necesario realizar multiplicaciones y puede producir periodos mayores de m. Veamos un ejemplo utilizando una secuencia X_1, X_2, X_3, X_4, X_5 conocida (generada en el método congruencia multiplicativa) y m = 10.

$X_1 = 1$
$X_2 = 2$
$X_3 = 4$
$X_4 = 8$
$X_5 = 6$
$X_6 = (X_5 + X_1) \bmod 10 = (6 + 1) \bmod 10 = 7$
$X_7 = (X_6 + X_2) \bmod 10 = (7 + 2) \bmod 10 = 9$
$X_8 = (X_7 + X_3) \bmod 10 = (9 + 4) \bmod 10 = 3$
$X_9 = (X_8 + X_4) \bmod 10 = (3 + 8) \bmod 10 = 1$
$X_{10} = (X_9 + X_5) \bmod 10 = (1 + 6) \bmod 10 = 7$
$X_{11} = (X_{10} + X_6) \bmod 10 = (7 + 7) \bmod 10 = 4$
$X_{12} = (X_{11} + X_7) \bmod 10 = (4 + 9) \bmod 10 = 3$
$X_{13} = (X_{12} + X_8) \bmod 10 = (3 + 3) \bmod 10 = 6$
$X_{14} = (X_{13} + X_9) \bmod 10 = (6 + 1) \bmod 10 = 7$
$X_{15} = (X_{14} + X_{10}) \bmod 10 = (7 + 7) \bmod 10 = 4$
$X_{16} = (X_{15} + X_{11}) \bmod 10 = (4 + 4) \bmod 10 = 8$

$X_{17} = (X_{16} + X_{12}) \text{ mod } 10 = (8 + 3) \text{ mod } 10 = 1$
$X_{18} = (X_{17} + X_{13}) \text{ mod } 10 = (1 + 6) \text{ mod } 10 = 7$
$X_{19} = (X_{18} + X_{14}) \text{ mod } 10 = (7 + 7) \text{ mod } 10 = 4$
$X_{20} = (X_{19} + X_{15}) \text{ mod } 10 = (4 + 4) \text{ mod } 10 = 8$

5.6.2.3 Método Congruencial Cuadrático

Este método se usa cuando **m** es una potencia de dos. Es casi equivalente al método del cuadrado medio de doble precisión, pero con un periodo más largo. La relación recursiva para este método es:

$$X_{n+1} = (Xn(X_n + 1)) \text{ mod } m \qquad n \geq 0$$

La semilla **X** no debe satisfacer la relación $X_0 \text{ mod } 4 = 2$.

Ejemplo:

$X_0 = 2$
$X_1 = (2 * 3) \text{ mod } 16 = 6$
$X_2 = (6 * 7) \text{ mod } 16 = 10$
$X_3 = (10 * 11) \text{ mod } 16 = 14$
$X_4 = (14 * 15) \text{ mod } 16 = 2$
$X_5 = (2 * 3) \text{ mod } 16 = 6$
$X_6 = (6 * 7) \text{ mod } 16 = 10$
$X_7 = (10 * 11) \text{ mod } 16 = 14$
$X_8 = (14 * 15) \text{ mod } 16 = 2$

5.6.3 Generador de Números Pseudo Aleatorios

Esta técnica es útil para generar números pseudo aleatorios que están en el intervalo de 0 y 1.

$$X_{n+1} = < 10^{-p} \, c \, X_n >$$

Donde **< >** denota que es una parte fraccional, **p** es el número de dígitos en el número pseudo aleatorio y **c** es una constante multiplicadora que está entre $0 < c < 1$.

Este método utiliza la forma de un sistema decimal al calcular la constante **a** en el método congruencia lineal ($a = 200T \pm Q$), $c = 10^{-p}$ ($200T \pm Q$), donde T es un número entero positivo y Q es cualquier número de estos (3, 11, 13, 19, 21, 2 7, 29, 37, 53, 59, 61, 67, 69, 77, 83, ó 91).

La semilla debe ser $X_0 = 10K$, donde K es cualquier entero no divisible por 2 ó 5 tal que $0 \leq K \leq 10^p$

Ejemplo:

$X_0 = 0.33$
$T = 0$
$Q = 11$
y se desean números de dos dígitos
$c = 10^{-2} (110) = 0.11$
$X_0 = 0.33$
$X_1 = < 100 (0.11) (0.33) > = 0.63$
$X_2 = < 100 (0.11) (0.63) > = 0.93$
$X_3 = < 100 (0.11) (0.93) > = 0.23$
$X_4 = < 100 (0.11) (0.23) > = 0.53$
$X_5 = < 100 (0.11)(0.53) > = 0.83$

La principal desventaja de este método es que es lento porque se requiere de muchas multiplicaciones para generar la secuencia de números aleatorios.

5.7 Pruebas Estadísticas para los Números Pseudo Aleatorios

Los diferentes algoritmos para la generación de números pseudo aleatorios se utilizan para simular las muestras de una distribución uniforme continua. Como ya sabemos los números generados mediante algoritmos no son completamente aleatorios, y en la medida en que estos números puedan pasar las pruebas estadísticas, estos números pseudo aleatorios pueden tratarse como verdaderos números aleatorios, aunque no lo sean. En el capítulo anterior se presentaron una serie de pruebas que también se pueden utilizar para los generadores de números aleatorios y pseudo aleatorios.

5.8 Generación de Números Pseudo Aleatorios No Uniformes

Las generaciones de números aleatorios (mencionados en el punto 5.6 de este capítulo) están diseñadas para generar secuencias de números que siguen una distribución uniforme. Sin embargo, otras distribuciones teóricas como son: la distribución normal, exponencial, Poisson y Gamma, se encuentran con más frecuencia en estudios de simulación que la distribución uniforme. En muchos casos pueden encontrarse distribuciones teóricas no apropiadas y se usa una distribución empírica.

Se recomienda que primeramente se utilicen las distribuciones teóricas convencionales y si ninguna de ellas describe en forma adecuada el comportamiento del sistema o proceso, entonces se utilizará las distribuciones empíricas. Es necesario tener una técnica para generar números aleatorios que simulen la muestra de cualquier distribución. Entonces se aplica una transformación a la variable uniforme para generar números pseudo aleatorios no uniformes. En esta sección, veremos las técnicas específicas para generar valores de variables aleatorias a partir de distribuciones de probabilidad más conocidas.

Al considerar los procesos estadísticos que involucran variables continuas o discretas, pero siempre de tipo aleatorio definiremos una función F(x) llamada función de distribución

acumulativa de x (mencionada en el capítulo 3), la cual denota la probabilidad de que una variable aleatoria X, tome un valor menor o igual a x. Si la variable aleatoria es discreta, entonces x tendrá valores específicos y F(x) será una función escalonada. Si F(x) es continua en el dominio x, entonces esta función se podrá diferenciar, para lo cual se define f(x)=dF(x)/dx. La derivada f(x) recibe el nombre de función de densidad de probabilidad. La función de distribución acumulativa se puede proponer matemáticamente así:

$$F(x) = P(X \leq x) = \int^{x} f(t)dt$$

donde F(x) se define en el intervalo $0 \leq F(x) \leq 1$ y f(t) representa el valor de la función de densidad de probabilidad de la variable aleatoria X cuando X = t.

Los valores de las variables aleatorias o números pseudo aleatorios generados por los algoritmos, son muy importantes en la generación de variables aleatorias a partir de otro tipo de distribución de probabilidad. A continuación, veremos los métodos para generar valores de variables aleatorias a partir de distribuciones de probabilidad.

5.8.1 El Método de la Transformación Inversa

Esta técnica es útil para transformar una desviación estándar uniforme en cualquier otra distribución.

Si se desea generar un número pseudo aleatorio de una distribución dada por F(x) donde F satisface todas las probabilidades de una función de distribución acumulada, los pasos para esta generación son:

Generar un número aleatorio uniforme estándar usando uno de los algoritmos para generar números aleatorios entre el rango de 0 y 1. Denominamos $\underline{r}$ como el número estándar uniforme generado en el paso 1. Hacemos que F(x) = r.

Por lo tanto, para cualquier valor particular $\underline{r}$ que generemos, por ejemplo, r_0, siempre es posible encontrar el valor de x, en este caso X_0 que corresponde a $\underline{r_0}$ debido a la función inversa de F conocida como:

$X_0 = F^{-1}(r_0)$ que será la variable no uniforme deseada.

Ejemplo:

Suponga que los números aleatorios uniformes estándar 0.1021, 0.2162, 0.7621 fueron generados, y se desea transformarlos en una distribución dada por:

$$F(X) = 0, \quad X < 0$$
$$F(X) = X, \quad 0 \leq X < 1/4$$
$$F(X) = (3X+1)/7, \quad 1/4 \leq X < 2$$
$$F(X) = 1, \quad X > 2$$

La función de distribución se representa en la ilustración en la figura 2.

Entonces:

$$F^{-1}(a) = a \quad , 0 \leq a < 1/4$$
$$F^{-1}(a) = (7a-1)/3 \quad , 1/4 \leq a \leq 1$$

Así:

$$F^{-1}(0.1021) = 0.1021$$
$$F^{-1}(0.2162) = 0.2162$$
$$F^{-1}(0.7621) = 1.4449$$

Estos números son números aleatorios de la distribución dada por F(X).

Si generamos números aleatorios uniformes correspondientes a una F(X) se puede resumir el método así:

$$r = F(X) = \int^{X} f(t)dt$$

donde $F^{-1}(r)$ es una variable que tiene a f(x) como función de densidad de probabilidad.

Ejemplo:

Generar los valores de variables aleatorias con una función de densidad $f(x) = 2x$ para $0 \leq X \leq 1$.
Resolviendo tenemos:

$$r = F(X) = \int^{X} 2t\, dt$$
$$= X^2$$

La transformación inversa $F^{-1}(r)$ será:

$$X = F^{-1}(r) = \sqrt{r} \ , 0 \leq r \leq 1$$

por lo tanto, los valores de X con una función de densidad f(X) = 2X se pueden generar al determinar la raíz cuadrada de los números aleatorios r.

Ejemplo:

Generar los valores X de variables aleatorias con una función de densidad:

$$f(X) = 1/4, \quad 0 \leq X \leq 1$$
$$f(X) = 3/4, \quad 1 \leq X \leq 2$$

Resolviendo:

$$r = F(X) = \int^{X} 1/4\, dt \quad 0 \leq X \leq 1$$
$$= X/4$$

$r = F(X) = 1/4 + \int^{X} 3/4\ dt \quad 1 \leq X \leq 2$
$= 3/4X = 1/2$

Para generar un valor de X se deben calcular los valores de X para r menor que 1/4 con X = 4r; y los valores de X se determinan para r mayor igual a 1/4 con X = 4/3r + 2/3.

El método de la transformación inversa, útil particularmente cuando se tabula la función de distribución acumulativa.

La función inversa se obtiene en este caso invirtiendo los valores de la abscisa y ordenada de los puntos tabulados, (de cada uno).

Ejemplo:

Suponga una distribución que ha sido tabulada así:

X	0.0	0.5	1.0	1.5	2.0	2.5	3.0
f(X)	0.0	0.1	0.2	0.5	0.7	0.9	1.0

Suponiendo que 0.25 es un número aleatorio uniforme generado, el número aleatorio transformado correspondiente, que se obtiene es 1.0833 interpolando linealmente entre 1.0 y 1.5.

5.8.2 Generación de Números Aleatorios Uniformes No Estándar

Esta distribución uniforme no estándar se conoce también con el nombre de técnica de rechazo. Produce números aleatorios en el rango de 0 a 1. Si f(X) es una función acotada y X tiene un rango finito, $a \leq X \leq b$, se utiliza esta técnica para generar los valores de las variables aleatorias. La distribución uniforme no estándar está dada por:

$F(x) = 0 \quad , \ X < a$
$F(x) = (x-a)/(b-a), \ a \leq X \leq b$
$F(x) = 1, \ X > b$

Para producir números aleatorios distribuidos uniformemente no estándar a números aleatorios con distribución uniformemente estándar se necesita un factor de escala para los números. Si r es un número aleatorio de una distribución uniforme estándar entonces X = a + (b-a)r será el número aleatorio de una distribución uniforme no estándar con un rango de a a b. Los pasos para esta técnica son:

- Normalizar el rango de f mediante un factor de escala c tal que:
  $\leq 1 \qquad a \leq X \leq b.$
- Definir a X como una función lineal de r, o sea, X = a + (b-a)r.
- Generar parejas de números aleatorios (r_1, r_2).

- Siempre que se encuentre una pareja de números aleatorios que satisfagan la relación $r_2 \leq$ cf.[a + (b-a)r_1], dicho par será aceptado y se utilizará a X = a + (b-a)r como el valor generado de la variable aleatoria.

Si X se elige al azar dentro del rango (a, b) y en el caso de que r > c·f(X) se rechaza, la función de densidad de probabilidad de los valores de X aceptados deberá ser igual a f(X).

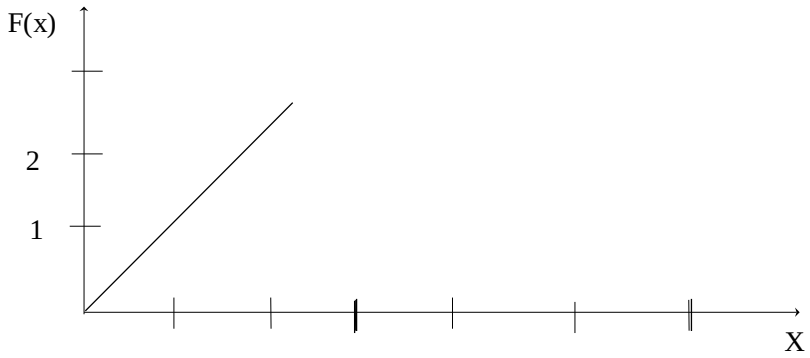
Ejemplo:

Utilizar el método del rechazo para generar valores de X de las variables aleatorias con una función de densidad: f(X) = 2X para $0 \leq X \leq 1$.

Ya que X se definió en el intervalo unitario se tiene que X = r y f(r) = 2r está definida en el intervalo $0 \leq f(r) \leq 2$. Por lo tanto, si se escala haciendo g(r) = 1/2f(r), se transforma a f(r) al intervalo unitario y g(r) = r. La figura 5.2 muestra función de densidad f(X)= 2X antes y después de escalar los números.

Los pasos seguidos en este ejemplo son:

- Generar r_1 y calcular g(r_1).
- Generar r_2 y comparados con g(r_1).
- Si $r_2 \leq$ g(r_1), se acepta a r_1 tomándolo como una X de f(X); si r_2 > g(r_1) se rechaza a r_1 y se vuelve a empezar con el paso 1.
- Este proceso se repite hasta generar n valores de x.



Antes de escalar

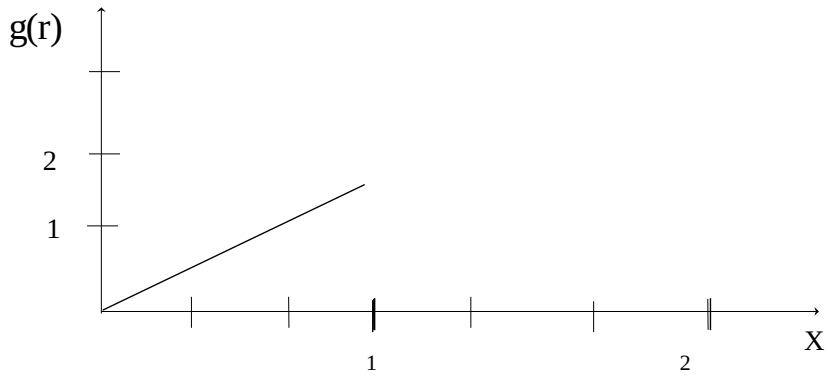


Figura 5.2 Función de Densidad

5.9 Problemas

Métodos Congruenciales

a) A = 7
B = 1
M = 1
(A - B) / M = (7 - 1) / 1 = 6
A / M = 7 / 1 = 7 A mod M= 7 mod 1 = 0
B / M = 1 / 1 = 1 A mod M = 1 mod 1 = 0

b) A = 15
B = 9
M = 6
(A - B) / M = (15 - 9) / 6 = 1
A / M = 15 / 6 = 2 A mod M = 15 mod 6 = 3
B / M = 9 / 6 = 1 B mod M = 9 mod 6 = 3

c) A = 11
B = 6
M = 5
(A - B) / M = (11 - 6) / 5 = 1
A / M = 11 / 5 = 1 A mod M = 11 mod 5 = 1
B / M = 6 / 5 = 1 B mod M = 6 mod 5 = 1

Método del Cuadrado Medio

a) $X_0 = 2418$ $(X_0)^2 = (2418)^2 = 05846724$
$X_1 = 8467$ $(X_1)^2 = (8467)^2 = 71690089$
$X_2 = 6900$ $(X_2)^2 = (6900)^2 = 47610000$
$X_3 = 6100$ $(X_3)^2 = (6100)^2 = 37210000$
$X_4 = 2100$ $(X_4)^2 = (2100)^2 = 44100000$
$X_5 = 1000$ $(X_5)^2 = (1000)^2 = 01000000$
$X_6 = 0000$

No se puede continuar iterando porque la semilla es cero.

b) $X_0 = 6214$ $(X_0)^2 = (6214)^2 = 38613796$
$X_1 = 6137$ $(X_1)^2 = (6137)^2 = 37662769$
$X_2 = 6627$ $(X_2)^2 = (6627)^2 = 43917129$
$X_3 = 9171$ $(X_3)^2 = (9171)^2 = 84107241$
$X_4 = 1072$ $(X_4)^2 = (1072)^2 = 01149184$

$X_5 = 1491 \quad (X_5)^2 = (1491)^2 = \quad 2223081$

Método Congruencial Lineal

a) $X_0 = 0$, $a = 1$, $c = 4$, $m = 6$

$X_0 = 0$
$X_1 = (X_0 + 4) \bmod 6 = (0 + 4) \bmod 6 = 4$
$X_2 = (X_1 + 4) \bmod 6 = (4 + 4) \bmod 6 = 2$
$X_3 = (X_2 + 4) \bmod 6 = (2 + 4) \bmod 6 = 0$
$X_4 = (X_3 + 4) \bmod 6 = (0 + 4) \bmod 6 = 4$
$X_5 = (X_4 + 4) \bmod 6 = (4 + 4) \bmod 6 = 2$
$X_6 = (X_5 + 4) \bmod 6 = (2 + 4) \bmod 6 = 0$

b) $X_0 = 5$, $a = 3$, $c = -1$, $m = 8$

$X_0 = 5$
$X_1 = (3X_0 - 1) \bmod 8 = (3*5 - 1) \bmod 8 = 6$
$X_2 = (3X_1 - 1) \bmod 8 = (3*6 - 1) \bmod 8 = 1$
$X_3 = (3X_2 - 1) \bmod 8 = (3*1 - 1) \bmod 8 = 2$
$X_4 = (3X_3 - 1) \bmod 8 = (3*2 - 1) \bmod 8 = 5$

c) $X_0 = 12$, $a = 2$, $c = 5$, $m = -10$

$X_0 = 12$
$X_1 = (2X_0 + 5) \bmod -10 = (2*12 + 5) \bmod -10 = 9$
$X_2 = (2X_1 + 5) \bmod -10 = (2*9 + 5) \bmod -10 = 3$
$X_3 = (2X_2 + 5) \bmod -10 = (2*3 + 5) \bmod -10 = 1$
$X_4 = (2X_3 + 5) \bmod -10 = (2*1 + 5) \bmod -10 = 7$

Método Congruencial Multiplicativo

a) $X_0 = 20$, $a = 3$, $c = 0$, $m = 13$

$X_0 = 20$
$X_1 = (3X_0) \bmod 13 = (3*20) \bmod 13 = 8$
$X_2 = (3X_1) \bmod 13 = (3*8) \bmod 13 = 11$
$X_3 = (3X_2) \bmod 13 = (3*11) \bmod 13 = 7$
$X_4 = (3X_3) \bmod 13 = (3*7) \bmod 13 = 8$

b) $X_0 = 31$, $a = 5$, $c = 0$, $m = 7$

$X_0 = 31$
$X_1 = (5X_0) \bmod 7 = (5*31) \bmod 7 = 1$
$X_2 = (5X_1) \bmod 7 = (5*1) \bmod 7 = 5$
$X_3 = (5X_2) \bmod 7 = (5*5) \bmod 7 = 4$
$X_4 = (5X_3) \bmod 7 = (5*4) \bmod 7 = 6$
$X_5 = (5X_4) \bmod 7 = (5*6) \bmod 7 = 2$

$X_6 = (5X_5) \bmod 7 = (5*2) \bmod 7 = 3$
$X_7 = (5X_6) \bmod 7 = (5*3) \bmod 7 = 1$

c) $X_0 = 43$ $X_0 = 43$
$a = 11$ $X_1 = (11X_0) \bmod 3 = (11*43) \bmod 3 = 2$
$c = 0$ $X_2 = (11X_1) \bmod 3 = (11*2) \bmod 3 = 1$
$m = 3$ $X_3 = (11X_2) \bmod 3 = (11*1) \bmod 3 = 2$

Método Congruencial Aditivo

a) Las primeras sucesiones son generadas por el método congruencia multiplicativa.

$m = 7$
$X_1 = 6$
$X_2 = 1$
$X_3 = 2$
$X_4 = 5$
$X_5 = (X_4 + X_1) \bmod 7 = (5+6) \bmod 7 = 4$
$X_6 = (X_5 + X_2) \bmod 7 = (4+1) \bmod 7 = 5$
$X_7 = (X_6 + X_3) \bmod 7 = (5+2) \bmod 7 = 0$
$X_8 = (X_7 + X_4) \bmod 7 = (0+5) \bmod 7 = 5$
$X_9 = (X_8 + X_5) \bmod 7 = (5+4) \bmod 7 = 2$
$X_{10} = (X_9 + X_6) \bmod 7 = (2+5) \bmod 7 = 0$
$X_{11} = (X_{10} + X_7) \bmod 7 = (0+0) \bmod 7 = 0$
$X_{12} = (X_{11} + X_8) \bmod 7 = (0+5) \bmod 7 = 5$
$X_{13} = (X_{12} + X_9) \bmod 7 = (5+2) \bmod 7 = 0$
$X_{14} = (X_{13} + X_{10}) \bmod 7 = (0+0) \bmod 7 = 0$
$X_{15} = (X_{14} + X_{11}) \bmod 7 = (0+0) \bmod 7 = 0$

b) Las primeras sucesiones son generadas por el método congruencia lineal

$X_1 = 1$ $m = 12$
$X_2 = 5$
$X_3 = 4$
$X_4 = 6$
$X_5 = 2$
$X_6 = 3$
$X_7 = (X_4 + X_1) \bmod 12 = (3+1) \bmod 12 = 4$
$X_8 = (X_5 + X_2) \bmod 12 = (4+5) \bmod 12 = 9$

$X_9 = (X_6 + X_3) \mod 12 = (9+4) \mod 12 = 1$
$X_{10} = (X_7 + X_4) \mod 12 = (1+6) \mod 12 = 7$
$X_{11} = (X_8 + X_5) \mod 12 = (7+2) \mod 12 = 9$
$X_{12} = (X_9 + X_6) \mod 12 = (9+3) \mod 12 = 0$
$X_{13} = (X_{10} + X_7) \mod 12 = (0+4) \mod 12 = 4$
$X_{14} = (X_{11} + X_8) \mod 12 = (4+5) \mod 12 = 9$
$X_{15} = (X_{12} + X_9) \mod 12 = (9+1) \mod 12 = 10$

c) Las primeras sucesiones son generadas por el método congruencia lineal.

$m = 5$
$X_1 = 3$
$X_2 = 9$
$X_3 = 1$
$X_4 = 5$
$X_5 = (X_4 + X_1) \mod 5 = (5 + 3) \mod 5 = 3$
$X_6 = (X_5 + X_2) \mod 5 = (3 + 9) \mod 5 = 2$
$X_7 = (X_6 + X_3) \mod 5 = (2 + 1) \mod 5 = 3$
$X_8 = (X_7 + X_4) \mod 5 = (3 + 5) \mod 5 = 3$
$X_9 = (X_8 + X_5) \mod 5 = (3+ 3) \mod 5 = 1$
$X_{10} = (X_9 + X_6) \mod 5 = (1 + 2) \mod 5 = 3$
$X_{11} = (X_{10} + X_7) \mod 5 = (3 + 3) \mod 5 = 1$
$X_{12} = (X_{11} + X_8) \mod 5 = (1 + 3) \mod 5 = 4$
$X_{13} = (X_{12} + X_9) \mod 5 = (4 + 1) \mod 5 = 0$
$X_{14} = (X_{13} + X_{10}) \mod 5 = (0 + 3) \mod 5 = 3$
$X_{15} = (X_{14} + X_{11}) \mod 5 = (3 + 1) \mod 5 = 4$

Método Congruencial Cuadrático

a) $X_0 = 5$ $\quad$ $X_0 = 5$
$m = 8$ $\quad$ $X_1 = (X_0 (X_0 + 1)) \mod 8 = (5*6) \mod 8 = 6$
$X_2 = (X_1 (X_1 + 1)) \mod 8 = (6*7) \mod 8 = 2$
$X_3 = (X_2 (X_2 + 1)) \mod 8 = (2*3) \mod 8 = 6$

b) $X_0 = 1$ $\quad$ $X_0 = 1$
$m = 16$ $\quad$ $X_1 = (X_0 (X_0 + 1)) \mod 16 = (1*2) \mod 16 = 2$
$X_2 = (X_1 (X_1 + 1)) \mod 16 = (2*3) \mod 16 = 6$
$X_3 = (X_2 (X_2 + 1)) \mod 16 = (6*7) \mod 16 = 10$
$X_4 = (X_3 (X_3 + 1)) \mod 16 = (10*11) \mod 16 = 14$
$X_5 = (X_4 (X_4 + 1)) \mod 16 = (14*15) \mod 16 = 2$

c) $X_0 = 2$ $X_0 = 2$

$m = 32$ $X_1 = (X_0(X_0+1)) \bmod 32 = (2*3) \bmod 32 = 6$

$X_2 = (X_1(X_1+1)) \bmod 32 = (6*7) \bmod 32 = 10$

$X_3 = (X_2(X_2+1)) \bmod 32 = (10*11) \bmod 32 = 14$

$X_4 = (X_3(X_3+1)) \bmod 32 = (14*15) \bmod 32 = 18$

$X_5 = (X_4(X_4+1)) \bmod 32 = (18*19) \bmod 32 = 22$

$X_6 = (X_5(X_5+1)) \bmod 32 = (22*23) \bmod 32 = 26$

$X_7 = (X_6(X_6+1)) \bmod 32 = (26*27) \bmod 32 = 30$

$X_8 = (X_7(X_7+1)) \bmod 32 = (30*31) \bmod 32 = 2$

$X_9 = (X_8(X_8+1)) \bmod 32 = (2*3) \bmod 32 = 6$

5.10 Programas

Programa P1taaa

Explicación del Programa:

1. Aspectos Generales

Este programa genera números aleatorios por el método de cuadrados centrales, determina la ocurrencia de generación de números aleatorios y llama a la subrutina P1TSUB la cual calcula el valor de Chi-cuadrado.

El período de generación de números (ocurrencia) termina cuando se repita alguno de los números generados o cuando GN llega a cero.

El programa envía un mensaje cuando se genera el cero, indicando que el período para esta semilla no tiene fin y se genera un ciclo infinito.

Las variables utilizadas son de tipo entero, real, de doble precisión dependiendo del lenguaje que se utiliza.

Se inicializan las variables y las tablas utilizadas en la generación de números para cada semilla.
Se imprimen en columna los números generados, al terminar la primera columna de impresión continúa imprimiendo una columna al lado y así sucesivamente hasta generar el último número aleatorio.

Definición de Variables:

SEED: valor de la semilla.
CANT: cantidad de semillas que se desean introducir.
TAB: tabla que almacena los números generados.
GN: guarda el valor leído de la semilla.
QPRO: variable de trabajo, almacena el cuadrado de la semilla o del número generado anterior.
DIV1: factor para la eliminación de dígitos a la derecha del cuadrado de la semilla.
DIV2: similar a la variable DIV1, sólo que se refiere a los dígitos a la izquierda que serán eliminados.
PRO: variable de trabajo que guarda el resultado de dividir QPRO entre DIV1.
AN: número generado (aritmética de punto flotante), esta variable recibe directamente el resultado del cálculo del residuo, a diferencia de la variable PN.
FGN: número generado (aritmética de punto flotante).
A : variable de trabajo (para control de impresión).
B : variable de trabajo (para control de impresión).
J : variable de trabajo (utilizada como índice).
I : variable de trabajo (utilizada como índice).
ICO: longitud del período.

IRES: variable de trabajo que se utiliza para determinar el número de líneas de impresión que ocuparán los números generados. Si su contenido no es igual a cero, se le debe sumar una unidad a la cantidad de líneas de impresión previamente establecidas; de lo contrario, la cantidad de líneas de impresión permanece igual a la ya establecida.
K : variable de trabajo que especifica la posición de la tabla en la que se debe almacenar el número que se acaba de generar.
KL: variable de trabajo (índice).
L : variable de trabajo. Se utiliza como índice para inicializar la tabla que almacena los números generados con ceros, y también contiene la cantidad de líneas que se han de imprimir de dicha tabla.
PN: Número generado (aritmética de punto flotante).

Programa P1tlco

Explicación del Programa:

1. Aspectos Generales

Este programa determina la longitud del ciclo para el generador lineal congruencial; se analizan los casos: el primero, escogiendo un valor de "c" que sea primo relativo de "m", y el segundo, escogiendo un valor de "c" que no sea primo relativo de "m".

El ciclo termina cuando se genera nuevamente el primer número generado; es por esta razón, que solamente es necesario almacenar el primer número generado hasta que este se repita.

Es necesario aclarar en este punto, que los valores de entrada: M, A, SEE1, C Y CANT son introducidos por el usuario por medio del teclado.

Definición de Variables:

A: Contiene un valor introducido por el usuario y constituye el multiplicador.
C: Incremento, que puede ser o no primo relativo de "m".
CANT: Cantidad de semillas.
I: Longitud del ciclo.
ICOM: Primer número generado.
IPRO: Variable de trabajo (almacena valores intermedios).
J: Variable de trabajo (índice del ciclo for).
M: Módulo para calcular la siguiente semilla.
SEE1: Semilla.
SEE2: Siguiente semilla.

Programa P1tasub

Explicación del Programa:

1. Aspectos Generales

Este programa calcula el valor de Chi-cuadrado; para ello, requiere que: se establezcan intervalos de clase, se calcula la frecuencia observada y la esperada, y se calcula la probabilidad de cada intervalo.

Este programa se corre mediante una opción del menú principal, pero antes se deberá correr el programa P1TAAA o el P1TADD.

La cantidad de intervalos de clase se calculó en base a la fórmula: Cantidad de intervalos = 1 + 3.3 + log(mayor valor clases generado) que permite una mejor distribución de los números generados.

La amplitud o tamaño de cada intervalo es igual a la diferencia entre el mayor y menor valor generado, dividida por la cantidad de intervalos de clase.

La columna "Frecuencia Observada" contendrá la cantidad de números generados que se encuentran dentro de cada uno de los intervalos, y, por lo tanto, su suma será igual a la cantidad de números generados (o período).

La probabilidad de cada intervalo es igual a 1/(cantidad de intervalo de clase) debido a que se hace la suposición de que la distribución es uniforme. Por conceptos de Probabilidad, se conoce que la suma de la probabilidad de cada intervalo de clase es igual a 1.

La frecuencia esperada es igual al cociente obtenido al dividir la cantidad de números generados entre la cantidad de intervalos de clase, ya que como la distribución es supuestamente uniforme, se espera que los números generados estén equitativamente distribuidos en los intervalos de clase.

Definición de Variables:

I: Variable de trabajo (índice de ciclos Do).
IAMP: Amplitud de los intervalos (aritméticos de punto fijo).
INT1: Cantidad de intervalos de clase (aritméticos de punto fijo).
J: Variable de trabajo (índice de un ciclo Do).
MAT: Matriz que contiene los intervalos de clase, las frecuencias observadas, la probabilidad de cada intervalo de clase, las frecuencias esperadas, y los términos que han de calcularse para determinar Chi-cuadrado.
MAY: Mayor número generado.
MEN: Menor número generado.
RAMP: Amplitud de los intervalos (aritmética de punto flotante).
RICO: Longitud del período (aritmética de punto flotante).
RIN: Cantidad de intervalos de clase (aritmética de punto flotante).

RINT: Cantidad de intervalos de clase (aritmética de punto flotante); a diferencia de la variable RIN, esta variable recibe directamente el resultado del cálculo de la cantidad de intervalos de clase.
SCHI: Chi-Cuadrado.
SESP: Suma de las frecuencias esperadas de todos los intervalos de clase.
SOBS: Suma de las frecuencias observadas de los intervalos de clase.
TAB: Tabla que almacena los números generados.
W: Variable de trabajo que almacena la diferencia entre las frecuencias observadas y esperadas, de cada intervalo de clase.

Programa P1tadd

Explicación del Programa:

1. Aspectos Generales

Este programa genera números aleatorios utilizando el método congruencial aditivo, determina la longitud del ciclo de generación y es indispensable para correr el programa P1TASUB para que calcule el valor de Chi-cuadrado.

El ciclo termina cuando se genera la secuencia inicial de semillas leídas.

El programa envía un mensaje que indica que el ciclo no tiene fin, cuando éste sobrepasa el valor 10,000.

La impresión de los números generados sigue una lógica similar a la del programa P1TAAA, ya que la tabla que almacena los números generados requiere ser inicializada con ceros cada vez que se lee un nuevo grupo de semillas, y también se imprimen diez números generados por línea.

Para la selección de un grupo de semillas cuyo ciclo fuese menor que m (ya que la mayor parte de los grupos de semillas tienen un ciclo mayor que m) fue necesario programar una calculadora (programable) para poder conseguir (por ensayo y error) un grupo de semillas que cumplieran con dicho requisito.

Definición de Variables:

CANT: Cantidad de grupos de semillas.
CONT1: Variable que controla las columnas en donde se desplegaran los valores en la pantalla.
CONT2: Variable que controla las filas en donde se desplegaran los valores en la pantalla.
DOB: Contiene el doble de la cantidad inicial de semillas.
I: Variable de trabajo (índice de ciclos Do).
IND1: Variable de trabajo que se utiliza para comparar los últimos números generados con la secuencia inicia de semillas para determinar si se ha cumplido el ciclo.
ISE: Cantidad inicial de semillas.
ISE1: Variable de trabajo que contiene una unidad más que la variable ISE.
ITAB: Variable de trabajo que contiene el número generado.

J: Variable de trabajo (índice del Do más extenso).
K: Longitud del ciclo.
SUM: Variable de trabajo que almacena la suma de las localidades de la tabla indicadas por los índices SUM1 y SUM2.
SUM1: Índice que indica la posición de la tabla en la que se encuentra el número generado anterior.
SUM2: Índice que corresponde a la posición equivalente a la diferencia entre la posición actual menos la cantidad inicial de semillas.
TAM: Tabla que almacena los números generados.
INDICADOR1: Rompe el ciclo más interno del programa.
INDICADOR2: Rompe el ciclo más externo del programa.
MODULO: Módulo (aritmética de punto flotante).

5.11 Resumen

Este capítulo ha presentado los conceptos más relevantes con respecto a la generación de números aleatorios y pseudo aleatorios. Estos dos, inclusive, no significan lo mismo. El primero denota un comportamiento totalmente no determinístico mientras que el segundo se genera por medio de un comportamiento determinístico. Existen diferentes técnicas y la que se utilice debe cumplir con una serie de requisitos para poder utilizarla dentro de un programa de simulación. Hoy día existen técnicas de generación de números verdaderamente aleatorios utilizando la física cuántica o el ruido atmosférico, entre otros. En muchos casos es necesario generar números pseudo aleatorios para poder repetir el experimento de un modelo de simulación.