

Simulación

Un enfoque práctico

Raúl Coss Bu

$F(X)$

1

R

$X = F^{-1}(R)$



LIMUSA

NORIEGA EDITORES

CONTENIDO

1. Introducción.	11
1.1. Definición de Simulación.	11
1.2. Etapas para realizar un estudio de Simulación.	12
1.3. Factores a considerar en el desarrollo del modelo de Simulación.	14
1.3.1. Generación de variables aleatorias no-uniformes	14
1.3.2. Lenguajes de programación.	15
1.3.3. Condiciones iniciales.	15
1.3.4. Tamaño de la muestra.	15
1.3.5. Diseño de experimentos.	16
1.4. Ventajas y desventajas en el uso de simulación.	17
1.5. Ejemplos de usos de simulación	18
2. Generación de números rectangulares.	19
2.1. Generadores congruenciales lineales.	20
2.1.1. Congruencial mixto.	20
2.1.2. Congruencial multiplicativo.	25
3. Pruebas estadísticas para los números pseudoaleatorios.	31
3.1. Prueba de los promedios.	31
3.2. Prueba de frecuencias.	33
3.3. Prueba de la distancia.	35
3.3.1. Números pseudoaleatorios considerados como dígitos.	35
3.3.2. Números pseudoaleatorios considerados como números reales.	36

3.4. Prueba de series.	38
3.5. Prueba de Kolmogorov-Smirnov.	40
3.6. Prueba del poker.	43
3.7. Prueba de las corridas.	46
3.7.1. Prueba de las corridas arriba y abajo del promedio.	46
3.7.2. Prueba de las corridas arriba y abajo.	46
4. Generación de variables aleatorias no-uniformes.	49
4.1. Método de la transformada inversa.	49
4.2. Método de rechazo.	53
4.3. Método de composición.	56
4.4. Procedimientos especiales.	60
5. Aplicaciones de simulación.	67
Ejemplo 5.1. Juego de volados.	67
Ejemplo 5.2. Camión transportador.	68
Ejemplo 5.3. Estimación de π	74
Ejemplo 5.4. Proyecto de inversión.	78
Ejemplo 5.5. Sistema de inventarios.	84
Ejemplo 5.6. Sistema de colas.	89
6. Análisis de los resultados de la simulación.	107
6.1. Métodos de estimación.	107
6.2. Simulación regenerativa.	110
7. Lenguajes de simulación.	123
7.1. Ventajas de los lenguajes de simulación.	123
7.2. Características de los lenguajes de simulación.	124
7.3. Factores a considerar en la selección de un lenguaje.	126
7.4. Clasificación de los lenguajes de simulación.	127
7.5. Introducción al GPSS.	129
Apéndice A. Números aleatorios uniformes.	154
Apéndice B. Distribución normal.	155
Apéndice C. Distribución χ^2	156
Bibliografía.	157

1

INTRODUCCION

Con el advenimiento de la computadora, una de las más importantes herramientas para analizar el diseño y operación de sistemas o procesos complejos es la simulación.

Aunque la construcción de modelos arranca desde el Renacimiento, el uso moderno de la palabra simulación data de 1940, cuando los científicos Von Neuman y Ulam que trabajaban en el proyecto Monte Carlo, durante la Segunda Guerra Mundial, resolvieron problemas de reacciones nucleares cuya solución experimental sería muy cara y el análisis matemático demasiado complicado.

Con la utilización de la computadora en los experimentos de simulación, surgieron incontables aplicaciones y con ello, una cantidad mayor de problemas teóricos y prácticos. En este libro se intenta por consiguiente, investigar y analizar cierto número de aplicaciones importantes de simulación de las áreas de economía, administración de negocios e investigación de operaciones, así como también sugerir algunos métodos alternativos para resolver algunos problemas teóricos y prácticos que surgen al efectuar simulaciones reales.

1.1. DEFINICION DE SIMULACION

Se ha empezado a utilizar la palabra simulación sin haber dado previamente una definición de ella. Por consiguiente, antes de proseguir con la discusión de este tema, sería conveniente describir algunas de las definiciones más aceptadas y difundidas de la palabra simulación. Thomas H. Naylor la define así:

Simulación es una técnica numérica para conducir experimentos en una computadora digital. Estos experimentos comprenden ciertos tipos de relaciones matemáticas y lógicas, las cuales son necesarias para describir el comportamiento y la estructura de sistemas complejos del mundo real a través de largos períodos de tiempo.

La definición anterior está en un sentido muy amplio, pues puede incluir desde una maqueta, hasta un sofisticado programa de computadora. En sentido más estricto, H. Maisel y G. Gnugnoli, definen simulación como:

Simulación es una técnica numérica para realizar experimentos en una computadora digital. Estos experimentos involucran ciertos tipos de modelos matemáticos y lógicos que describen el comportamiento de sistemas de negocios, económicos, sociales, biológicos, físicos o químicos a través de largos períodos de tiempo.

Otros estudiosos del tema como Robert E. Shannon, definen simulación como:

Simulación es el proceso de diseñar y desarrollar un modelo computarizado de un sistema o proceso y conducir experimentos con este modelo con el propósito de entender el comportamiento del sistema o evaluar varias estrategias con las cuales se puede operar el sistema.

Las definiciones anteriores no especifican si los sistemas modelados son continuos o discretos. Sin embargo, es necesario senalar que el grueso de este libro está dedicado al diseño, análisis y validación de sistemas dinámicos discretos. Algunos autores de este tema como Geoffrey Gordon en su libro *System Simulation*, tratan a fondo el Análisis y estudio de sistemas dinámicos continuos.

1.2. ETAPAS PARA REALIZAR UN ESTUDIO DE SIMULACION

Se ha escrito mucho acerca de los pasos necesarios para realizar un estudio de simulación. Sin embargo, la mayoría de los autores opinan que los pasos necesarios para llevar a cabo un experimento de simulación son:

- *Definición del sistema.* Para tener una definición exacta del sistema que se desea simular, es necesario hacer primeramente

un análisis preliminar del mismo, con el fin de determinar la interacción del sistema con otros sistemas, las restricciones del sistema, las variables que interactúan dentro del sistema y sus interrelaciones, las medidas de efectividad que se van a utilizar para definir y estudiar el sistema y los resultados que se esperan obtener del estudio.

- *Formulación del modelo.* Una vez que están definidos con exactitud los resultados que se esperan obtener del estudio, el siguiente paso es definir y construir el modelo con el cual se obtendrán los resultados deseados. En la formulación del modelo es necesario definir todas las variables que forman parte de él, sus relaciones lógicas y los diagramas de flujo que describan en forma completa al modelo.
- *Colección de datos.* Es posible que la facilidad de obtención de algunos datos o la dificultad de conseguir otros, pueda influenciar el desarrollo y formulación del modelo. Por consiguiente, es muy importante que se definan con claridad y exactitud los datos que el modelo va a requerir para producir los resultados deseados. Normalmente, la información requerida por un modelo se puede obtener de registros contables, de órdenes de trabajo, de órdenes de compra, de opiniones de expertos y si no hay otro remedio por experimentación.
- *Implementación del modelo en la computadora.* Con el modelo definido, el siguiente paso es decidir si se utiliza algún lenguaje como fortran, basic, algol, etc., o se utiliza algún paquete como GPSS, simula, simscript, etc., para procesarlo en la computadora y obtener los resultados deseados.
- *Validación.* Una de las principales etapas de un estudio de simulación es la validación. A través de esta etapa es posible detallar deficiencias en la formulación del modelo o en los datos alimentados al modelo. Las formas más comunes de validar un modelo son:
 1. La opinión de expertos sobre los resultados de la simulación.
 2. La exactitud con que se predicen datos históricos.
 3. La exactitud en la predicción del futuro.
 4. La comprobación de falla del modelo de simulación al utilizar datos que hacen fallar al sistema real.
 5. La aceptación y confianza en el modelo de la persona que hará uso de los resultados que arroje el experimento de simulación
- *Experimentación.* La experimentación con el modelo se realiza después de que éste ha sido validado. La experimentación

consiste en generar los datos deseados y en realizar análisis de sensibilidad de los índices requeridos.

- *Interpretación.* En esta etapa del estudio, se interpretan los resultados que arroja la simulación y en base a esto se toma una decisión. Es obvio que los resultados que se obtienen de un estudio de simulación ayudan a soportar decisiones del tipo semi-estructurado, es decir, la computadora en sí no toma la decisión, sino que la información que proporciona ayuda a tomar mejores decisiones y por consiguiente a sistemáticamente obtener mejores resultados.
- *Documentación.* Dos tipos de documentación son requeridos para hacer un mejor uso del modelo de simulación. La primera se refiere a la documentación de tipo técnico, es decir, a la documentación que el departamento de Procesamiento de Datos debe tener del modelo. La segunda se refiere al manual del usuario, con el cual se facilita la interacción y el uso del modelo desarrollado, a través de una terminal de computadora.

1.3. FACTORES A CONSIDERAR EN EL DESARROLLO DEL MODELO DE SIMULACION

Puesto que la simulación está basada fuertemente en la teoría de probabilidad y estadística, en matemáticas, en ciencias computacionales, etc., es conveniente decir algunas ideas de cómo intervienen estas áreas en el desarrollo y formulación del modelo de simulación.

1.3.1. Generación de variables aleatorias no-uniformes

Si el modelo de simulación es estocástico, la simulación debe ser capaz de generar variables aleatorias no-uniformes de distribuciones de probabilidad teóricas o empíricas. Lo anterior puede ser obtenido si se cuenta con un generador de números uniformes y una función que transforme estos números en valores de la distribución de probabilidad deseada. A este respecto, se han desarrollado una gran cantidad de generadores para las distribuciones de probabilidad más comunes como: La distribución normal, la distribución exponencial, la distribución poisson, la distribución erlang, la distribución binomial, la distribución gamma, la distribución beta, la distribución F, la distribución t, etc.

1.3.2. Lenguajes de programación

Las primeras etapas de un estudio de simulación se refieren a la definición del sistema a ser modelado y a la descripción del sistema en términos de relaciones lógicas de sus variables y diagramas de flujo. Sin embargo, llega el momento de describir el modelo en un lenguaje que sea aceptado por la computadora que se va a usar. En esta etapa se tienen dos cursos de acción a seguir si no se tiene nada de software sobre simulación: 1) Desarrollar el software requerido para estudios de simulación, ó 2) Comprar software (lenguajes de programación de propósito especial). Para esta alternativa es necesario analizar y evaluar varios paquetes de simulación (GPSS, GASP, etc.) antes de tomar la decisión final.

1.3.3. Condiciones iniciales

La mayoría de los modelos de simulación estocástica se corren con la idea de estudiar al sistema en una situación de estado estable. Sin embargo, la mayoría de estos modelos presentan en su etapa inicial estados transientes los cuales no son típicos del estado estable. Por consiguiente es necesario establecer claramente las alternativas o cursos de acción que existen para resolver este problema. Algunos autores piensan que la forma de atacar este problema sería a través de:

- Usar un tiempo de corrida lo suficientemente grande de modo que los períodos transientes sean relativamente insignificantes con respecto a la condición de estado estable.
- Excluir una parte apropiada de la parte inicial de la corrida.
- Utilizar simulación regenerativa.

Obviamente, de las tres alternativas presentadas, la que presenta menos desventajas es el uso de simulación regenerativa. Las otras alternativas presentan las desventajas de ser prohibitivamente excesivas en costo.

1.3.4. Tamaño de la muestra

Uno de los factores principales a considerar en un estudio de simulación es el tamaño de la muestra (número de corridas en la computadora).

La selección de un tamaño de muestra apropiado que asegure un nivel deseado de precisión y a la vez minimice el costo de operación del modelo, es un problema algo difícil pero muy importante. Puesto que la información proporcionada por el experimento de simulación sería la base para decidir con respecto a la operación del sistema real, esta información deberá ser tan exacta y precisa como sea posible o al menos el grado de imprecisión presente en la información proporcionada por el modelo debe ser conocida. Por consiguiente, es necesario que un análisis estadístico sea realizado para determinar el tamaño de muestra requerido.

El tamaño de la muestra puede ser obtenido de dos maneras:

1. Previa e independientemente de la operación del modelo, o
2. Durante la operación del modelo y basado en los resultados arrojados por el modelo. Para la última alternativa se utiliza la técnica estadística de intervalos de confianza.

1.3.5. Diseño de experimentos

El diseño de experimentos es un tópico cuya relevancia en experimentos de simulación ha sido reconocida pero raramente aplicado. El diseño de experimentos en estudios de simulación puede ser de varios tipos, dependiendo de los propósitos específicos que se hayan planteado. Existen varios tipos de análisis que pueden ser requeridos. Entre los más comunes e importantes se pueden mencionar los siguientes:

- Comparación de las medias y variancias de las alternativas analizadas.
- Determinación de la importancia y el efecto de diferentes variables en los resultados de la simulación.
- Búsqueda de los valores óptimos de un conjunto de variables.

Para realizar el primer tipo de análisis, al cual se le denomina comúnmente diseño de experimentos de un factor simple, es necesario tomar muy en cuenta el tamaño de la muestra, las condiciones iniciales y la presencia o ausencia de autocorrelación. Para el segundo tipo de análisis, existe una gran cantidad de literatura, puesto que la gran mayoría de los libros de texto de diseño de experimentos, explican o tratan el tema de análisis de variancia y técnicas de regresión como medios para evaluar la importancia y el efecto de varias variables en los resultados de operación de un sistema. Para el tercer tipo de análisis, generalmente se requiere utilizar algoritmos heurísticos de búsqueda como por ejemplo el algoritmo de Hooke y Jeeves.

1.4. VENTAJAS Y DESVENTAJAS EN EL USO DE SIMULACION

Aunque la técnica de simulación generalmente se ve como un método de último recurso, recientes avances en las metodologías de simulación y la gran disponibilidad de software que actualmente existe en el mercado, han hecho que la técnica de simulación sea una de las herramientas más ampliamente usadas en el análisis de sistemas. Además de las razones antes mencionadas, Thomas H. Naylor ha sugerido que un estudio de simulación es muy recomendable porque presenta las siguientes ventajas:

- A través de un estudio de simulación, se puede estudiar el efecto de cambios internos y externos del sistema, al hacer alteraciones en el modelo del sistema y observando los efectos de esas alteraciones en el comportamiento del sistema.
- Una observación detallada del sistema que se está simulando puede conducir a un mejor entendimiento del sistema y por consiguiente a sugerir estrategias que mejoren la operación y eficiencia del sistema.
- La técnica de simulación puede ser utilizada como un instrumento pedagógico para enseñar a estudiantes habilidades básicas en análisis estadístico, análisis teórico, etc.
- La simulación de sistemas complejos puede ayudar a entender mejor la operación del sistema, a detectar las variables más importantes que interactúan en el sistema y a entender mejor las interrelaciones entre estas variables.
- La técnica de simulación puede ser usada para experimentar con nuevas situaciones, sobre las cuales se tiene poca o ninguna información. A través de esta experimentación se puede anticipar mejor a posibles resultados no previstos.
- La técnica de simulación se puede utilizar también para entrenamiento de personal. En algunas ocasiones se puede tener una buena representación de un sistema (como por ejemplo los juegos de negocios), y entonces a través de él es posible entrenar y dar experiencia a cierto tipo de personal.
- Cuando nuevos elementos son introducidos en un sistema, la simulación puede ser usada para anticipar cuellos de botella o algún otro problema que puede surgir en el comportamiento del sistema.

A diferencia de las ventajas mencionadas, la técnica de simulación presenta el problema de requerir equipo computacional y recursos hu-

manos costosos. Además, generalmente se requiere bastante tiempo para que un modelo de simulación sea desarrollado y perfeccionado. Finalmente, es posible que la alta administración de una organización no entienda esta técnica y esto crea dificultad en vender la idea.

1.5. EJEMPLOS DE USOS DE SIMULACION

Existe una gran cantidad de áreas donde la técnica de simulación puede ser aplicada. Algunos ejemplos podrían ser los siguientes:

- *Simulación de un sistema de colas.* Con la técnica de simulación es posible estudiar y analizar sistemas de colas cuya representación matemática sería demasiado complicada de analizar. Ejemplos de estos sistemas serían aquellos donde es posible la llegada al sistema en grupo, la salida de la cola del sistema, el rehusar entrar al sistema cuando la cola es excesivamente grande, etc.
- *Simulación de un sistema de inventarios.* A través de simulación se pueden analizar más fácilmente sistemas de inventarios donde todos sus parámetros (tiempo de entrega, demanda, costo de llevar inventario, etc.), son estocásticos.
- *Simulación de un proyecto de inversión.* Existen en la práctica una gran cantidad de proyectos de inversión donde la incertidumbre con respecto a los flujos de efectivo que el proyecto genera a las tasas de interés, a las tasas de inflación, etc., hacen difícil y a veces imposible manejar analíticamente este tipo de problemas. Para este tipo de situaciones el uso de simulación es ampliamente recomendado.
- *Simulación de sistemas económicos.* La técnica de simulación puede ser utilizada para evaluar el efecto de cierto tipo de decisiones (devaluación de la moneda, el impuesto al valor agregado, etc.), en las demás variables macroeconómicas como: producto nacional bruto, balanza comercial, inflación, oferta monetaria, circulante, etc.
- *Simulación de estados financieros.* La expansión y diversificación de una organización a través de la adquisición y establecimiento de nuevas empresas, repercuten significativamente en su posición y estructura financiera. Por consiguiente, el uso de simulación permite analizar cuál de las estrategias de crecimiento son las que llevarán a la organización al logro de sus objetivos y metas de corto, mediano y largo plazos.

2

GENERACION DE NUMEROS RECTANGULARES

En todos los experimentos de simulación existe la necesidad de generar valores de variables aleatorias que representan a una cierta distribución de probabilidad. Durante un experimento de simulación, el proceso de generar un valor de la variable aleatoria de una distribución particular, puede repetirse tantas veces como se desee y tantas veces como distribuciones de probabilidad existan en el experimento de simulación. Sin embargo, es conveniente señalar que el proceso de generación de variables aleatorias no uniformes se hace a partir de la generación de números rectangulares. Por consiguiente, el objetivo de este capítulo es mostrar un panorama general de las diferentes técnicas que existen para generar números rectangulares.

La importancia de los números rectangulares (distribución uniforme) radica en su uso para la generación de variables aleatorias más complicadas que son requeridas en los experimentos de simulación. Algunos autores como Tocher, han sugerido tres formas para obtener los números rectangulares: La provisión externa, la generación interna a partir de un proceso físico al azar y la generación interna de sucesiones de dígitos por medio de una relación de recurrencia. El primer método implica tener los números aleatorios, como por ejemplo las tablas de la Rand, en una cinta magnética o en un disco y tratar a estos números como datos de entrada para el problema que se está simulando. El segundo método implica utilizar algún aditamento especial de la computadora digital capaz de registrar los resultados de un proceso aleatorio y además, reduzca esos resultados a sucesiones de dígitos. El tercer método, y uno de los más aceptados, implica la generación de estos números rectangulares a través de una relación de recurrencia.

Independientemente del proceso o procedimiento que se utilice para

la generación de los números rectangulares, estos deben de poseer ciertas características deseables que aseguren o aumenten la confiabilidad de los resultados obtenidos de la simulación. Tales características son:

1. Uniformemente distribuidos,
2. Estadísticamente independientes,
3. Reproducibles,
4. Período largo (sin repetición dentro de una longitud determinada de la sucesión),
5. Generados a través de un método rápido.
6. Generados a través de un método que no requiera mucha capacidad de almacenamiento de la computadora.

Finalmente, es necesario señalar que algunos autores califican a los números rectangulares generados a través de relaciones de recurrencia con números *pseudoaleatorios*, por ser una sucesión de dígitos generada mediante una regla puramente determinística. Sin embargo, esta objeción puede superarse, al menos parcialmente, al tomar el punto de vista un tanto pragmático de que una sucesión puede considerarse aleatoria si satisface un cierto conjunto de pruebas estadísticas de aleatoriedad.

2.1. GENERADORES CONGRUENCIALES LINEALES

Varios esquemas han sido propuestos para la generación de los números pseudoaleatorios a través de relaciones matemáticas de recurrencia. Estos números se consideran pseudoaleatorios, porque aunque pasan todas las pruebas estadísticas de aleatoriedad, ellos son de hecho completamente determinísticos. Actualmente, casi todas las computadoras incluyen en sus programas de biblioteca alguna variante de los métodos congruenciales sugeridos por Lehmer. Los dos métodos congruenciales más populares son: congruencial mixto y congruencial multiplicativo.

2.1.1. Congruencial mixto

Los generadores congruenciales lineales generan una secuencia de números pseudoaleatorios en la cual el próximo número pseudoaleatorio es determinado a partir del último número generado, es decir, el número pseudoaleatorio X_{n+1} es derivado a partir del número pseudoaleatorio X_n .

Para el caso particular del generador congruencial mixto, la relación de recurrencia es la siguiente:

$$X_{n+1} = (aX_n + c) \bmod m \quad (2.1)$$

donde:

X_0 = la semilla ($X_0 > 0$)

a = el multiplicador ($a > 0$)

c = constante aditiva ($c > 0$)

m = el módulo ($m > X_0, m > a$ y $m > c$)

Esta relación de recurrencia nos dice que X_{n+1} es el residuo de dividir $aX_n + c$ entre el módulo. Lo anterior significa que los valores posibles de X_{n+1} son 0, 1, 2, 3, ..., $m-1$, es decir, m representa el número posible de valores diferentes que pueden ser generados.

Con el propósito de ilustrar la generación de números pseudoaleatorios a través de este método, suponga que se tiene un generador en el cual los valores de sus parámetros son: $a = 5$, $c = 7$, $X_0 = 4$ y $m = 8$. Para estos valores, la secuencia de números pseudoaleatorios y números uniformes (X_{n+1}/m) son mostrados en la tabla 2-1. Como se puede apreciar en esta tabla, el período del generador es 8.

Después de haber analizado este ejemplo, podría pensarse que el período de todo generador es siempre igual a m . Sin embargo, esto es falso porque el período depende de los valores asignados a los parámetros a , c , X_0 y m , es decir, se requiere seleccionar valores adecuados para estos parámetros con el fin de que el generador tenga período completo.

TABLA 2.1. Números pseudoaleatorios del generador $X_{n+1} = (5X_n + 7) \bmod 8$.

n	X_n	$(5X_n + 7)/8$	X_{n+1}	Números uniformes
0	4	$3 + 3/8$	3	3/8
1	3	$2 + 6/8$	6	6/8
2	6	$4 + 5/8$	5	5/8
3	5	$4 + 0/8$	0	0
4	0	$0 + 7/8$	7	7/8
5	7	$5 + 2/8$	2	2/8
6	2	$2 + 1/8$	1	1/8
7	1	$1 + 4/8$	4	4/8

Para ilustrar el caso que se presenta cuando el período $< m$, suponga que se tiene un generador en el cual los valores de sus parámetros son: $a = X_0 = c = 7$ y $m = 10$. Para estos valores, la secuencia de números pseudoaleatorios y números uniformes son mostrados en la tabla 2.2. Como puede apreciarse en esta tabla, el período del generador es 4. Esto demuestra que una selección inadecuada de los valores de los parámetros del generador, puede conducirnos a obtener resultados indeseables y poco confiables del experimento de simulación.

TABLA 2.2. Números pseudoaleatorios del generador $X_{n+1} = (7X_n + 7) \bmod 10$.

n	X_n	$(7X_n + 7)/10$	X_{n+1}	Números uniformes
0	7	$5 + 6/10$	6	6/10
1	6	$4 + 9/10$	9	9/10
2	9	$7 + 0/10$	0	0
3	0	$0 + 7/10$	7	7/10

De los ejemplos anteriores, se advierte la necesidad de establecer algunas reglas que puedan ser utilizadas en la selección de los valores de los parámetros, para que el generador resultante tenga período completo. Algunas de estas reglas se mencionan a continuación:

a) Selección de m .

Existen dos opciones para seleccionar el valor apropiado del módulo:

1. Seleccionar m de modo que sea el número primo más grande posible y a que a su vez sea menor que p^d , donde p es la base del sistema (binario, decimal, hexadecimal, etc.) que se está utilizando y d es el número de bits que tiene una palabra de computadora en ese sistema. Por ejemplo, si se tiene una computadora IBM 370 que trabaja en sistema binario, entonces $p = 2$ y $d = 32$.
2. Seleccionar m como p^d . Cuando m toma este valor se facilita el cálculo del número rectangular ($U_n = X_n/m$), ya que sólo se corre el punto binario o decimal a la izquierda del número. Sin embargo, se ha comprobado que cuando el módulo toma este valor, los últimos dígitos del número pseudoaleatorio generado no se comportan en forma aleatoria.

Para ilustrar el problema que se presenta cuando se utiliza el criterio 2, suponga que se tiene un generador cuyos parámetros son: $a = 81$, $c = 89$, $X_0 = 5$ y $m = 10^2$. Para estos valores, la secuencia de números

pseudoaleatorios son mostrados en la tabla 2.3. En esta tabla se puede apreciar que el último dígito del número pseudoaleatorio tiene un período de 10. Esto significa que el último dígito puede ser determinado a partir de la siguiente relación de recurrencia:

$$Y_{n+1} = (Y_n + 9) \bmod 10$$

TABLA 2.3. Números pseudoaleatorios del generador $X_{n+1} = (81X_n + 89) \bmod 100$

n	X_n	n	X_n	n	X_n	n	X_n	n	X_n
1	94	21	74	41	54	61	34	81	14
2	03	22	83	42	63	62	43	82	23
3	32	23	12	43	92	63	72	83	52
4	81	24	61	44	41	64	21	84	01
5	50	25	30	45	10	65	90	85	70
6	39	26	19	46	99	66	79	86	59
7	48	27	28	47	08	67	88	87	68
8	77	28	57	48	37	68	17	88	97
9	26	29	06	49	86	69	66	89	46
10	95	30	75	50	55	70	35	90	15
11	84	31	64	51	44	71	24	91	04
12	93	32	73	52	53	72	33	92	13
13	22	33	02	53	82	73	62	93	42
14	71	34	51	54	31	74	11	94	91
15	40	35	20	55	00	75	80	95	60
16	29	36	09	56	89	76	69	96	49
17	38	37	18	57	98	77	78	97	58
18	67	38	47	58	27	78	07	98	87
19	16	39	96	59	76	79	56	99	36
20	85	40	65	60	45	80	25	100	05

Del ejemplo anterior, es posible generalizar una relación de recurrencia que relacione los últimos dígitos del número pseudoaleatorio generado. Si $m = p^d$, se ha encontrado que la relación de recurrencia de los últimos dígitos es la siguiente:

$$Y_{n+1,i} = X_{n+1} \bmod P^i \quad i < d \quad (2.2.)$$

donde:

- $Y_{n+1,i}$ = últimos i dígitos del número pseudoaleatorio X_{n+1} .
- i = últimos i dígitos que se están considerando. El valor de i puede ser 1, 2, 3, ..., $d - 1$.

Por ejemplo, si $i = 1$, la expresión (2.2) permite determinar el valor del último dígito del número pseudoaleatorio X_{n+1} , si $i = 2$, se determina el valor de los dos últimos dígitos y así sucesivamente.

La expresión (2.2) puede ser manipulada y expresada en función de los últimos dígitos del número pseudoaleatorio inmediato anterior, de acuerdo a lo siguiente:

$$Y_{n+1,i} = X_{n+1} \bmod p^i$$

sustituyendo en la expresión anterior la ecuación (2.1) se obtiene:

$$Y_{n+1,i} = (aX_n + c) \bmod p^d \bmod p^i$$

pero si q es la parte entera que resulta de dividir $(aX_n + c)/p^d$, entonces la expresión anterior se transforma en:

$$Y_{n+1,i} = (aX_n + c - qp^d) \bmod p^i$$

$$Y_{n+1,i} = (aX_n + c) \bmod p^i$$

$$Y_{n+1,i} = aX_n \bmod p^i + c \bmod p^i$$

y como $X \bmod m = (X \bmod m) \bmod m$, entonces la expresión anterior se transforma en:

$$Y_{n+1,i} = (aX_n \bmod p^i) \bmod p^i + c \bmod p^i$$

y haciendo uso de la identidad (2.2), la expresión anterior se transforma finalmente en:

$$Y_{n+1,i} = (aY_{n,i} + c) \bmod p^i \quad (2.3)$$

esta última expresión también tendrá periodo completo (p^i) si el generador tiene un periodo de p^d

b) Selección de a

El valor seleccionado de a debe ser entero impar, y además no debe ser divisible por 3 ó 5. Sin embargo, si queremos asegurar que el generador tenga periodo completo, el valor de a se debe de seleccionar de acuerdo al siguiente criterio:

$$(a - 1) \bmod 4 = 0 \quad \text{si } 4 \text{ es un factor de } m$$

$$(a - 1) \bmod b = 0 \quad \text{si } b \text{ es un factor primo de } m$$

Usualmente se selecciona a como $2^k + 1$ cuando se trabaja en sistema binario y $10^k + 1$ cuando se trabaja en sistema decimal. En ambos casos el valor de k debe ser mayor o igual a 2.

c) Selección de c

El valor seleccionado para este parámetro puede ser cualquier constante. Sin embargo, si se desean asegurar buenos resultados el valor de c debe ser $c \bmod 8 = 5$ si se trabaja en sistema binario y $c \bmod 200 = 21$ si se trabaja en sistema decimal. Más específicamente, el valor de c debe ser un entero impar y relativamente primo a m .

d) Selección de X_0

Para el generador congruencial mixto, se ha encontrado que el valor de la semilla es irrelevante, es decir, el valor de este parámetro resulta tener poca o ninguna influencia sobre las propiedades estadísticas de las sucesiones.

Finalmente, antes de terminar la discusión de este generador, conviene señalar que existen otras formas matemáticas de representarlo. Tales formas son las siguientes:

$$X_n = \left\{ a^n X_0 + c \left\{ \frac{a^n - 1}{a - 1} \right\} \right\} \bmod m \quad (2.4)$$

$$X_{n+k} = \left\{ a^k X_n + c \left\{ \frac{a^k - 1}{a - 1} \right\} \right\} \bmod m \quad (2.5)$$

Con la expresión (2.4) el n -ésimo número pseudoaleatorio se obtiene a partir de la semilla. Con la expresión (2.5) el $n + k$ -ésimo número pseudoaleatorio se obtiene a partir del k -ésimo número, es decir, si por ejemplo $n + k = 10$ y $k = 4$, entonces significa que el número pseudoaleatorio 10 se va a obtener a partir del número 4.

2.1.2. Congruencial multiplicativo

Al igual que el generador congruencial mixto, el generador congruencial multiplicativo determina el próximo número pseudoaleatorio a partir del último número generado, de acuerdo a la siguiente relación de recurrencia:

$$X_{n+1} = aX_n \bmod m \quad (2.6)$$

Para este generador se recomienda también seleccionar adecuadamente los valores de los parámetros a , X_0 y m , con el fin de asegurar un período máximo para las sucesiones generadas por este método. Los valores de estos parámetros dependerán del sistema en que se trabaje, es decir, estos parámetros tomarán valores distintos si se trabaja en sistema decimal, que si se trabaja en sistema binario. Por consiguiente, a continuación se describen las reglas que se recomiendan seguir para seleccionar los valores de a , X_0 y m dependiendo de si el sistema en que se trabaja es binario o decimal.

a) Sistema decimal

Si se trabaja en sistema decimal, los valores de los parámetros deben ser seleccionados de acuerdo a los siguientes criterios:

1. El valor de la semilla puede ser cualquier entero impar no divisible entre 2 ó 5 y debe ser relativamente primo a m .
2. El valor seleccionado de a debe ser obtenido de acuerdo a la siguiente identidad:

$$a = 200t \pm p$$

donde t es cualquier entero y p es cualquiera de los siguientes valores: 3, 11, 13, 19, 21, 27, 29, 37, 53, 59, 61, 67, 69, 77, 83, 91.

3. El valor seleccionado de m puede ser 10^d . Si $m = 10$ y $d \geq 5$ el período del generador es $5 \times 10^{d-2}$.

Por otra parte, si $m = 10^d$ y $d < 5$, entonces el período del generador se obtiene de acuerdo a la siguiente expresión:

$$\text{Período}^* = \text{Mínimo comun múltiplo} \{ \lambda(P_1^{d_1}), \lambda(P_2^{d_2}), \dots, \lambda(P_n^{d_n}) \} \quad (2.7)$$

donde:

$$\lambda(2) = 1, \lambda(4) = 2$$

$$\lambda(2^d) = 2^{d-2} \text{ si } d \geq 3$$

$$\lambda(p^d) = p^{d-1} (p - 1) \text{ si } p \geq 2$$

* P_i es un factor primo de m .

Con el propósito de ilustrar la obtención del período para este último caso, analicemos el siguiente generador:

$$X_{n+1} = 3X_n \bmod 100 \text{ y } X_0 = 17$$

puesto que m puede ser expresado como 10^2 o bien como $(2^2)(5^2)$, entonces el período de este generador de acuerdo a la expresión (2.7) sería:

$$\begin{aligned} \text{Período} &= \text{Mínimo común múltiplo } ((2^2), (5^2)) \\ &= \text{Mínimo común múltiplo } (2, 20) \\ &= 20 \end{aligned}$$

La tabla 2-4 muestra la secuencia de números pseudoaleatorios de este generador. Como se puede apreciar en esta tabla, el período del generador es 20.

TABLA 2.4. Números pseudoaleatorios del generador $X_{n+1} = 3X_n \bmod 100$.

n	X_n	n	X_n	n	X_n	n	X_n
1	51	6	93	11	99	16	57
2	53	7	79	12	97	17	71
3	59	8	37	13	91	18	13
4	77	9	11	14	73	19	39
5	31	10	33	15	19	20	17

b) Sistema binario

Si se trabaja en sistema binario, los valores de los parámetros deben ser seleccionados de acuerdo a los siguientes criterios:

1. El valor de la semilla puede ser cualquier entero impar relativamente primo a m .
2. El valor seleccionado de a debe ser obtenido a partir de la siguiente expresión:

$$a = 8t \pm 3$$

donde t es cualquier entero.

3. El valor seleccionado de m puede ser 2^d . Si $m = 2^d$ el período del generador es 2^{d-2} ó $m/4$. Para ilustrar la obtención del período de un generador en sistema binario, suponga que se tiene un generador en el cual los valores de sus parámetros son: $a = 5$, $X_0 = 5$ y $m = 32$. Para estos valores, la secuencia de números pseudoaleatorios son mostrados en la tabla 2.5. Como se puede apreciar en esta tabla, el período del generador es 8.

TABLA 2.5. Números pseudoaleatorios del generador
 $X_{n+1} = 5X_n \bmod 32$.

n	X_n	n	X_n
1	25	5	9
2	29	6	13
3	17	7	1
4	21	8	5

PROBLEMAS

- 2.1. Determine el período de los siguientes generadores congruenciales mixtos:

$$X_{n+1} = (8X_n + 16) \bmod 100 \text{ y } X_0 = 15$$

$$X_{n+1} = (50X_n + 17) \bmod 64 \text{ y } X_0 = 13$$

$$X_{n+1} = (5X_n + 24) \bmod 32 \text{ y } X_0 = 7$$

$$X_{n+1} = (5X_n + 21) \bmod 100 \text{ y } X_0 = 3$$

$$X_{n+1} = (9X_n + 13) \bmod 32 \text{ y } X_0 = 8$$

- 2.2. Determine el período de los siguientes generadores congruenciales multiplicativos:

$$X_{n+1} = 203 X_n \bmod 10^5 \text{ y } X_0 = 17$$

$$X_{n+1} = 211 X_n \bmod 10^8 \text{ y } X_0 = 19$$

$$X_{n+1} = 221 X_n \bmod 10^3 \text{ y } X_0 = 3$$

$$X_{n+1} = 5 X_n \bmod 64 \text{ y } X_0 = 7$$

$$X_{n+1} = 11 X_n \bmod 128 \text{ y } X_0 = 9$$

- 2.3. Defina los parámetros a , c , X_0 y m de un generador congruencial mixto que aseguren período completo.
- 2.4. Defina los parámetros a , X_0 y m de un generador congruencial multiplicativo que aseguren período completo.

- 2.5. Defina la relación de recurrencia de los últimos i dígitos de los siguientes generadores congruenciales mixtos:

$$X_{n+1} = (21X_n + 221) \bmod 100, X_0 = 7 \text{ e } i = 1$$

$$X_{n+1} = (61X_n + 421) \bmod 1000, X_0 = 11 \text{ e } i = 2.$$

3

PRUEBAS ESTADISTICAS PARA LOS NUMEROS PSEUDOALEATORIOS

Puesto que cualquier variable aleatoria no-uniforme (normal, exponencial, poisson, etc.), es obtenida a partir de números uniformes (0;1), el principal énfasis en pruebas estadísticas deberá ser con respecto al generador de números pseudoaleatorios, ya que cualquier deficiencia estadística en la distribución de la variable aleatoria no-uniforme, se deberá exclusivamente a la utilización de un deficiente generador de números pseudoaleatorios. Por consiguiente, en el presente capítulo se explican algunas de las muchas pruebas estadísticas que han sido desarrolladas para probar la aleatoriedad de los números pseudoaleatorios.

3.1. PRUEBA DE LOS PROMEDIOS

Quizá la función de densidad de probabilidad más simple es aquella que se caracteriza por ser constante en el intervalo (0;1) y cero fuera de él (ver figura 3.1.). Esta función de densidad define la distribución conocida como uniforme o rectangular. Matemáticamente, la función de densidad uniforme se define como:

$$f(x) = \begin{cases} 1 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{si } 0 > x > 1 \end{cases}$$

En esta expresión, x es una variable aleatoria definida en el intervalo (0;1). Por otra parte, la distribución acumulada $F(x)$, de una variable aleatoria x uniformemente distribuida, se puede obtener como:

$$F(x) = \int_0^x 1 dt = x$$

El valor esperado y la variancia de una variable aleatoria uniformemente distribuida están dadas por las siguientes expresiones:

$$E(x) = \int_0^1 x(1)dx = \frac{1}{2}$$

$$\text{var}(x) = \int_0^1 (x - 1/2)^2 (1)dx = 1/12$$

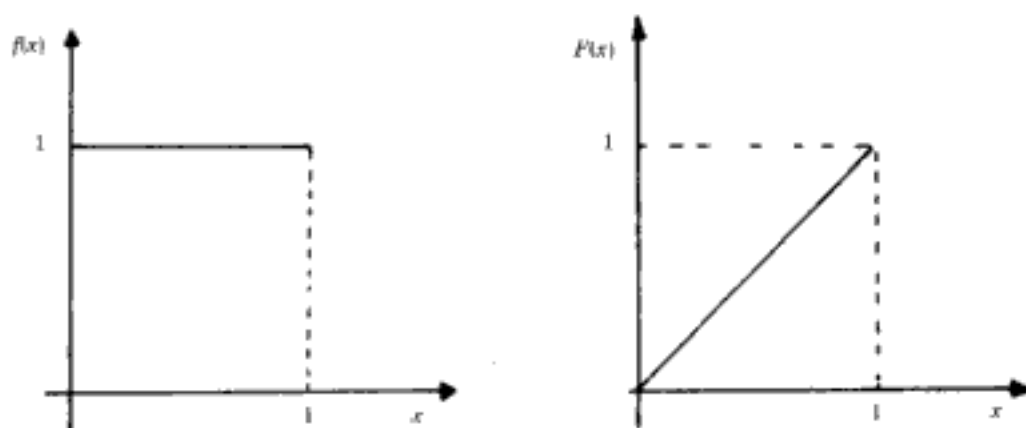


FIGURA 3.1. Distribución de probabilidad y distribución acumulada de los números uniformes (0; 1)

Conociendo los parámetros de la distribución uniforme, es posible plantear una prueba de hipótesis de promedios, con la cual se trata de probar que los números pseudoaleatorios generados provienen de un universo uniforme con media de 0.5. Más específicamente, una prueba de hipótesis de promedios puede ser planteada de la siguiente forma:

Hipótesis nula $H_0 : u = 1/2$

Hipótesis alternativa $H_1 : u \neq 1/2$

La realización de esta prueba requiere obtener una muestra de tamaño N , es decir, es necesario generar N números pseudoaleatorios. En seguida, su promedio aritmético es evaluado de acuerdo a la siguiente expresión:

$$\bar{x} = \frac{U_1 + U_2 + \dots + U_n}{N} \quad (3.1)$$

En seguida se determina el valor del estadístico Z_0 , utilizando la siguiente expresión:

$$Z_0 = \frac{(\bar{x} - 1/2) \sqrt{N}}{\sqrt{1/12}} \quad (3.2)$$

Si $|Z_0| < Z_{\alpha/2}$, entonces no se puede rechazar la hipótesis de los números pseudoaleatorios generados provienen de un universo uniforme con media de 0.5.

Para ilustrar la aplicación de esta prueba de hipótesis, considere los números pseudoaleatorios presentados en la tabla 3.1. Para estos números, la media aritmética (aplicación de la ecuación 3.1) resulta ser de 0.48234, y el estadístico Z_0 resulta ser de:

$$Z_0 = \frac{(0.48234 - 0.5) \sqrt{100}}{\sqrt{1/12}} = -0.61$$

Si se supone un nivel de significado (α) de 5%, entonces $Z_{\alpha/2}$ tiene un valor de 1.96 y puesto que $|Z_0| < 1.96$, entonces no se puede rechazar la hipótesis de que estos números pseudoaleatorios tienen una media de 0.5.

3.2. PRUEBA DE FRECUENCIAS

Probablemente una de las más importantes pruebas sobre aleatoriedad de los números pseudoaleatorios es la prueba de las frecuencias. Esta prueba consiste en dividir el intervalo (0;1) en n subintervalos (ver figura 3.2) para luego comparar para cada subintervalo la frecuencia esperada con la frecuencia observada. Si estas frecuencias son bastante parecidas, entonces la muestra proviene de una distribución uniforme. El estadístico que se usa en esta prueba es X_0^2 el cual se obtiene de acuerdo a la siguiente expresión:

$$X_0^2 = \sum_{i=1}^n \frac{(FO_i - FE_i)^2}{FE_i} \quad 3.3$$

donde:

- FO_i = Frecuencia observada del i^{th} subintervalo.
- FE_i = Frecuencia esperada del i^{th} subintervalo (N/n).
- N = Tamaño de la muestra.
- n = Número de subintervalos.

Tabla 3.1. Tabla de números pseudoaleatorios.

0.78961	0.05230	0.10699	0.55877	0.14151
0.76086	0.12079	0.27738	0.65726	0.79269
0.80548	0.82654	0.29453	0.20852	0.42989
0.58518	0.98611	0.34488	0.34358	0.11537
0.89898	0.57880	0.67621	0.05010	0.00121
0.28269	0.73059	0.70119	0.18284	0.49962
0.38618	0.76910	0.68334	0.55170	0.10850
0.79982	0.45679	0.21631	0.87616	0.55743
0.58962	0.33216	0.03185	0.61168	0.09264
0.69623	0.17028	0.05475	0.91512	0.76262
0.29931	0.30861	0.83358	0.51781	0.03272
0.57410	0.26593	0.85903	0.43308	0.35286
0.24000	0.65559	0.38507	0.90829	0.94187
0.93655	0.88809	0.81772	0.36982	0.19904
0.54325	0.62400	0.09133	0.41678	0.33954
0.58244	0.85853	0.88752	0.33729	0.15506
0.23949	0.53559	0.33381	0.49383	0.75103
0.19962	0.65002	0.74579	0.79113	0.63453
0.19147	0.40644	0.08128	0.73435	0.22724
0.22287	0.07281	0.64183	0.44267	0.72102

Este estadístico X_0^2 se compara con $X_{\alpha, (n-1)}^2$, la cual representa a una variable aleatoria chi-cuadrada con $(n-1)$ grados de libertad y un nivel de significado α . Si $X_0^2 < X_{\alpha, (n-1)}^2$, entonces no se puede rechazar la hipótesis de que la muestra proviene de una distribución uniforme.

Frecuencia esperada	$\frac{N}{n}$	$\frac{N}{n}$	$\frac{N}{n}$...	$\frac{N}{n}$	$\frac{N}{n}$
Frecuencia observada	FO_1	FO_2	FO_3	...	FO_{n-1}	FO_n
	$\frac{1}{n}$	$\frac{2}{n}$	$\frac{3}{n}$...	$\frac{n-1}{n}$	$\frac{n}{n}$

FIGURA 3.2. Frecuencia esperada y observada de cada subintervalo.

Por ejemplo, considere los números pseudoaleatorios presentados en la tabla 3.1, y además considere que el número de subintervalos es 5,

es decir, $n = 5$. Para este valor de n , la frecuencia esperada de cada subintervalo es 20, y las frecuencias observadas son:

Frecuencia esperada	20	20	20	20	20
Frecuencia observada	21	22	19	23	15
	.2	.4	.6	.8	1

Por consiguiente, aplicando la ecuación (3.3) se obtiene:

$$X_0^2 = \frac{1}{20} \left\{ (21-20)^2 + (22-20)^2 + (19-20)^2 + (23-20)^2 + (15-20)^2 \right\} = 2$$

Si se especifica un valor arbitrario de $\alpha = 0.05$, entonces $X_{0.05,4}^2 = 9.49$. Puesto que $X_0^2 < X_{0.05,4}^2$, entonces no se puede rechazar la hipótesis de que los números pseudoaleatorios presentados en la tabla 3.1 provienen de una distribución uniforme.

3.3. PRUEBA DE LA DISTANCIA

Esta prueba puede ser realizada de dos maneras: considerando a los números pseudoaleatorios generados como dígitos, o considerando como números reales.

3.3.1. Números pseudoaleatorios considerados como dígitos

Si los números pseudoaleatorios generados son considerados como dígitos, entonces la prueba consiste en contar el número de dígitos que aparece entre ocurrencias sucesivas de un mismo dígito. Por ejemplo, 58245 ilustra un hueco de tamaño 3 entre los dos cincos.

La probabilidad de cada uno de los tamaños de hueco ($i = 0, 1, 2, \dots$) se obtiene con la siguiente expresión:

$$P_i = 0.1(0.9)^i \text{ para } i = 0, 1, 2, \dots \quad (3.4)$$

Sin embargo, como teóricamente el valor del tamaño del hueco puede ser infinito, es conveniente agrupar las probabilidades para valores de i mayores o iguales a un determinado valor de n . Tal sumatoria se obtiene de acuerdo a la siguiente expresión:

$$P_{i \geq n} = \sum_{m=0}^{\infty} 0.1(0.9)^{m+n} = (0.9)^n \quad (3.5)$$

Con las ecuaciones (3.4) y (3.5) se obtienen las frecuencias esperadas para cada tamaño de hueco. La obtención de tales frecuencias se ilustra en la tabla 3.2. Si las frecuencias esperadas y observadas para cada tamaño de hueco son bastante parecidas, entonces se puede decir que los números pseudoaleatorios generados pasan la prueba de la distancia. El estadístico X_0^2 que se usa en esta prueba se obtiene como:

$$X_0^2 = \sum_{i=0}^n \frac{(FO_i - FE_i)^2}{FE_i} \quad (3.6)$$

el cual se compara con $X_{\alpha, n}^2$. Si $X_0^2 < X_{\alpha, n}^2$, entonces los números pseudoaleatorios pasan la prueba de la distancia. Es muy importante señalar que el valor seleccionado de n , debe ser tal que la suma de las frecuencias esperadas de todos los tamaños de huecos agrupados, sea mayor que 5.

Tabla 3.2. Frecuencias esperadas y observadas para los diferentes tamaños de hueco, considerando a los números pseudoaleatorios generados como dígitos.

i	P_i	FO_i	FE_i
0	0.1	FO_0	$\Sigma FO_i (0.1)$
1	0.1 (0.9)	FO_1	$\Sigma FO_i (0.1)(0.9)$
2	0.1 (0.9) ²	FO_2	$\Sigma FO_i (0.1)(0.9)^2$
.	.	.	.
.	.	.	.
.	.	.	.
i	0.1 (0.9) ^{i}	FO_i	$\Sigma FO_i (0.1)(0.9)^i$
.	.	.	.
.	.	.	.
$\geq n$	(0.9) ^{n}	FO_n	$\Sigma FO_i (0.9)^n$
Total	1.0	ΣFO_i	ΣFO_i

3.3.2. Números pseudoaleatorios considerados como números reales

Si los números pseudoaleatorios generados son considerados como números reales, entonces, para realizar esta prueba es necesario seleccionar un intervalo $(\alpha; \beta)$, el cual debe de estar contenido en el intervalo

$(0; 1)$, es decir, $0 \leq \alpha \leq \beta \leq 1$. En seguida, para cada número pseudoaleatorio generado se pregunta si es o no elemento del intervalo $(\alpha; \beta)$. Si U_j (número uniforme generado) es elemento de $(\alpha; \beta)$, U_{j+1} hasta U_{j+i} no son elementos de dicho intervalo y U_{j+i+1} vuelve a ser elemento del intervalo $(\alpha; \beta)$, entonces se tiene un hueco de tamaño i . Por ejemplo, si $\alpha = .3$ y $\beta = .5$ y los números pseudoaleatorios generados son como sigue: 0.32415, 0.22257, 0.19147, 0.75103, 0.49383, entonces el hueco es de tamaño 3.

Para el caso de considerar los números pseudoaleatorios generados como números reales, la distribución de probabilidad del tamaño del hueco es como sigue:

$$P_i = \theta (1 - \theta)^i \text{ para } i = 0, 1, 2, \dots \quad (3.7)$$

donde $\theta = \beta - \alpha$ representa la probabilidad de caer en el intervalo $(\alpha; \beta)$.

Al igual que en el inciso anterior, es conveniente agrupar las probabilidades para valores de $i \geq n$. Tal agrupación se obtiene con la siguiente expresión:

$$P_{i \geq n} = \sum_{m=0}^{\infty} \theta (1 - \theta)^{m+n} = (1 - \theta)^n \quad (3.8)$$

Con las ecuaciones (3.7) y (3.8) se obtienen las frecuencias esperadas, las cuales aparecen en la tabla 3.3. Utilizando la ecuación (3.6) y comparando el resultado obtenido con $X_{n,n}^2$, se toma la decisión de aceptar o

Tabla 3.3. Frecuencias esperadas y observadas para los diferentes tamaños de hueco, considerando a los números pseudoaleatorios generados como números reales.

i	P_i	FO_i	FE_i
0	θ	FO_0	$\Sigma FO_i (\theta)$
1	$\theta (1 - \theta)$	FO_1	$\Sigma FO_i (\theta)(1 - \theta)$
2	$\theta (1 - \theta)^2$	FO_2	$\Sigma FO_i (\theta)(1 - \theta)^2$
.	.	.	.
.	.	.	.
.	.	.	.
i	$\theta (1 - \theta)^i$	FO_i	$\Sigma FO_i (\theta)(1 - \theta)^i$
.	.	.	.
.	.	.	.
$\geq n$	$(1 - \theta)^n$	FO_n	$\Sigma FO_i (1 - \theta)^n$
Total	1.0	ΣFO_i	ΣFO_i

rechazar la prueba de la distancia. Es muy importante señalar que el valor de α y β no tienen ninguna influencia en la bondad de la prueba, y también es necesario señalar que el valor de n debe ser seleccionado de acuerdo al criterio mencionado en el inciso anterior.

Por ejemplo, si se consideran los números pseudoaleatorios presentados en la tabla 3.1, un valor de $n = 3$ y valores de $\alpha = .3$ y $\beta = .7$, entonces las frecuencias observadas y esperadas para los diferentes tamaños de hueco serían como aparecen en la tabla 3.4. Para estas frecuencias el valor del estadístico X_0^2 sería:

$$X_0^2 = \frac{(11 - 16)^2}{16} + \frac{(12 - 9.6)^2}{9.6} + \frac{(11 - 5.76)^2}{5.76} + \frac{(6 - 8.64)^2}{8.64} = 7.74$$

Si se especifica un valor arbitrario de $\alpha = 0.05$, entonces $X_{0.05,3}^2 = 7.81$. Puesto que $X_0^2 < X_{0.05,3}^2$, entonces los números pseudoaleatorios presentados en la tabla 3.1 pasan la prueba de la distancia.

Tabla 3.4. Frecuencias esperadas y observadas para los diferentes tamaños de hueco que originan los números pseudoaleatorios presentados en la tabla 3.1.

i	P_i	FO_i	FE_i
0	0.400	11	16.00
1	0.240	12	9.60
2	0.144	11	5.76
≥ 3	0.216	6	8.64
Total	1.000	40	40.00

3.4. PRUEBA DE SERIES

La prueba de series se utiliza para comprobar el grado de aleatoriedad entre números sucesivos. Usualmente esta prueba consiste en formar parejas de números, las cuales son consideradas como coordenadas en un cuadrado unitario dividido en n^2 celdas. Esta idea se puede extender a las ternas de números pseudoaleatorios que representan puntos en un cubo unitario. Sin embargo, en esta sección se analiza sólo el caso de dos dimensiones.

La prueba de series consiste en generar N números pseudoaleatorios de los cuales se forman parejas aleatorias entre U_i y U_{i+1} , es decir, si se generan 10 números, entonces las parejas aleatorias que se pueden formar serían:

$(U_1, U_2), (U_2, U_3), (U_3, U_4), (U_4, U_5), (U_5, U_6)$

$(U_6, U_7), (U_7, U_8), (U_8, U_9), (U_9, U_{10})$

En seguida, se determina la celda a que pertenece cada pareja ordenada (ver figura 3.3), con lo cual se determina la frecuencia observada de cada celda. La frecuencia esperada de cada celda se obtiene al dividir el total de parejas coordinadas $(N-1)$ por el total de celdas (n^2) . Finalmente, conocida la frecuencia observada y esperada de cada celda, se obtiene el estadístico X_0^2 como sigue:

$$X_0^2 = \frac{n^2}{N-1} \sum_{i=1}^n \sum_{j=1}^n (FO_{ij} - (N-1)/n^2)^2 \quad (3.9)$$

donde FO_{ij} es la frecuencia observada de la celda ij . Si $X_0^2 < X_{\alpha, n^2-1}^2$ entonces no se puede rechazar la hipótesis de que los números provienen de una distribución uniforme.

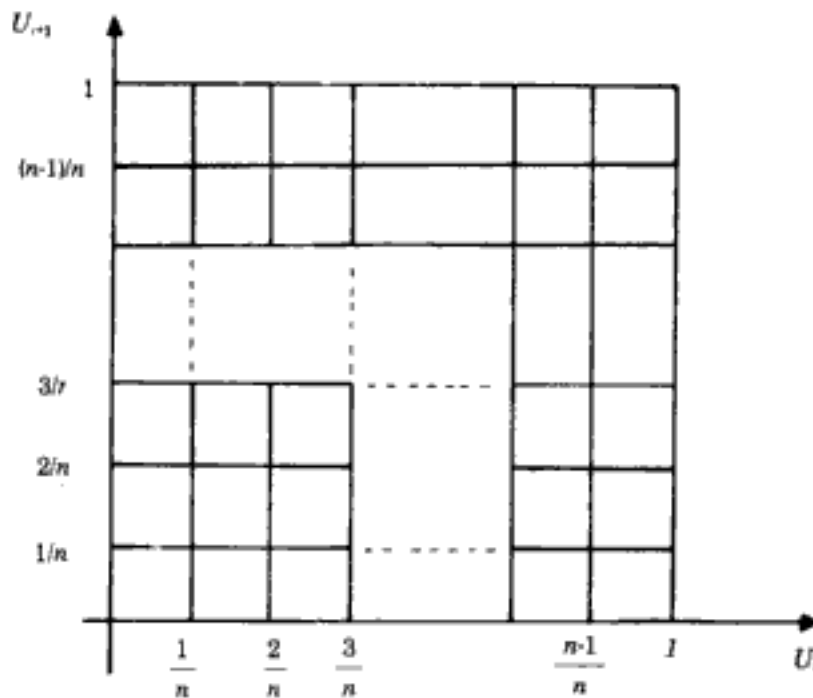


Figura 3.3. Cuadrado unitario con n^2 celdas.

Finalmente conviene mencionar que algunos autores han mezclado la prueba de series y la prueba de frecuencias en el siguiente estadístico:

$$X_0^2 = \frac{n^2}{(N-1)} \sum_{i=1}^n \sum_{j=1}^n (FO_{ij} - (N-1)/n^2) \left(\frac{n}{N} \sum_{i=1}^n (FO_i - N/n)^2 \right) \quad (3.10)$$

el cual sigue una distribución chi-cuadrada con $(n-1)^2$ grados de libertad. Si se opta por esta prueba, entonces se compara X_0^2 con $X_{\alpha, (n-1)^2}^2$. Si $X_0^2 < X_{\alpha, (n-1)^2}^2$, entonces no se puede rechazar la hipótesis de uniformidad de los números pseudoaleatorios generados.

Por ejemplo, si se aplica esta prueba a los números pseudoaleatorios presentados en la tabla 3.1, utilizando un valor de $n = 5$, entonces se obtienen las frecuencias observadas que aparecen en la figura 3.4. Para estas frecuencias, la aplicación de la ecuación (3.9) produce los siguientes resultados:

$$X_0^2 = \frac{25}{99} (5(2 - 3.96)^2 + 5(3 - 3.96)^2 + 5(4 - 3.96)^2 + 6(5 - 3.96)^2 + 4(6 - 3.96)^2) = 11.86$$

y seleccionando un valor de $\alpha = 0.05$, entonces $X_{0.05, 24}^2 = 36.4$. Puesto que $X_0^2 < X_{0.05, 24}^2$, entonces los números pseudoaleatorios presentados en la tabla 3.1 pasan la prueba de uniformidad.

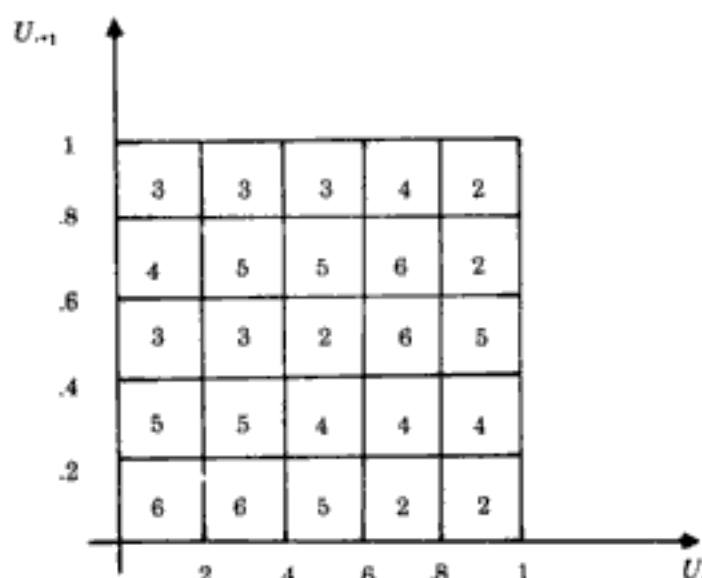


Figura 3.4. Frecuencias observadas obtenidas con los nuevos pseudoaleatorios presentados en la tabla 3.1.

3.5. PRUEBA DE KOLMOGOROV-SMIRNOV

El procedimiento de Kolmogorov-Smirnov prueba la hipótesis de que la distribución acumulada de una variable aleatoria x es $F_0(x)$. Para probar esta hipótesis, una muestra de tamaño n es obtenida de una distribución continua $F(x)$. En seguida, se determina la distribución

acumulada de la muestra, la cual se denota por $F_n(x)$. Posteriormente, $F_n(x)$ es comparada con la distribución acumulada hipotética $F_0(x)$. Si $F_n(x)$ difiere demasiado de $F_0(x)$, entonces esto es una amplia evidencia de que $F_n(x)$ no es igual a $F_0(x)$.

La aplicación de esta prueba al caso de números pseudoaleatorios uniformes, puede ser descrita en los siguientes pasos:

1. Generar n números pseudoaleatorios uniformes.
2. Ordenar dichos números en orden ascendente.
3. Calcular la distribución acumulada de los números generados con la siguiente expresión:

$$F_n(x) = \frac{i}{n} \quad (3.11)$$

donde i es la posición que ocupa el número X_i en el vector obtenido en el paso 2.

4. Calcular el estadístico Kolmogorov-Smirnov del modo siguiente (ver figura 3.5).

$$D_n = \max |F_n(X_i) - X_i| \text{ para toda } X_i \quad (3.12)$$

5. Si $D_n < d_{\alpha,n}$, entonces no se puede rechazar la hipótesis de que los números generados provienen de una distribución uniforme. La distribución de D_n ha sido tabulada (ver tabla 3.5) como una función de n y α para cuando $F_n(x) = F_0(x)$.

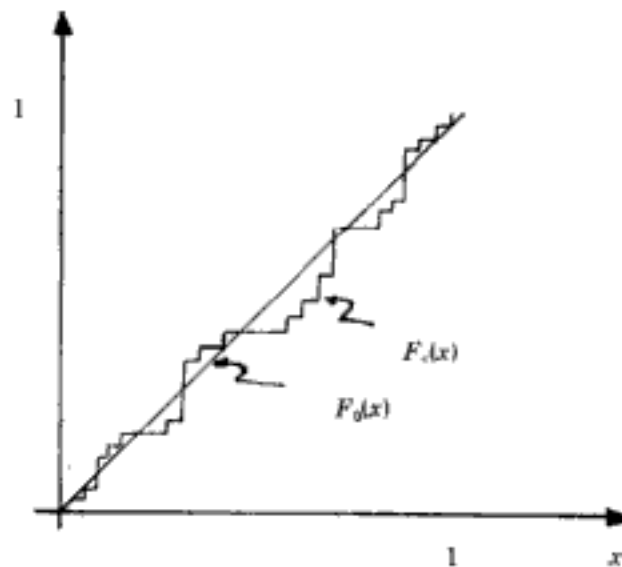


Figura 3.5. Comparación de $F_0(x)$ para cada valor de x generado.

Tabla 3.5. $d_{\alpha,n}$ para diferentes valores del nivel de significado y del tamaño de la muestra.

Tamaño de la muestra (n)	$d_{\alpha,n}$		
	$\alpha = 10\%$	$\alpha = 5\%$	$\alpha = 1\%$
1	0.950	0.975	0.995
2	0.776	0.842	0.929
3	0.642	0.708	0.829
4	0.564	0.624	0.734
5	0.510	0.563	0.669
6	0.470	0.521	0.618
7	0.438	0.486	0.577
8	0.411	0.457	0.543
9	0.388	0.432	0.514
10	0.368	0.409	0.486
11	0.352	0.391	0.468
12	0.338	0.375	0.450
13	0.352	0.361	0.433
14	0.314	0.349	0.418
15	0.304	0.338	0.404
16	0.295	0.328	0.392
17	0.286	0.318	0.381
18	0.278	0.309	0.371
19	0.272	0.301	0.363
20	0.264	0.294	0.352
25	0.240	0.264	0.317
30	0.220	0.242	0.290
35	0.210	0.230	0.270
40		0.210	0.252
50		0.188	0.226
60		0.172	0.207
70		0.160	0.192
80		0.150	0.180
90		0.141	
100		0.134	
Fórmula aproximada para $n > 100$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

Para ilustrar una aplicación de esta prueba, considere los números pseudoaleatorios presentados en la tabla 3.1. Para estos números la tabla 3.6 muestra un ordenamiento de dichos números en orden ascendente, los cuales corresponden a la distribución acumulada teórica $F_0(x)$. También, en esta tabla se muestra la distribución acumulada ex-

perimental $F_n(x)$. De acuerdo a estas distribuciones, el estadístico kolmogorov-smirnov (ecuación 3.12) resulta ser de:

$$D_n = 0.89 - 0.83358 = 0.05642$$

y seleccionando un valor de $\alpha = 0.05$, entonces $d_{0.05,100} = 0.134$ (ver tabla 3.5). Puesto que $D_n \leq d_{0.05,100}$, entonces los números pseudoaleatorios presentados en la tabla 3.1 provienen de una distribución uniforme.

3.6. PRUEBA DEL POKER

La prueba del poker examina en forma individual los dígitos del número pseudoaleatorio generado. La forma como esta prueba se realiza es tomando 5 dígitos a la vez y clasificándolos como: par, dos pares, tercia, poker, quintilla, full y todos diferentes. Lo anterior significa que los números pseudoaleatorios generados son de 5 dígitos cada uno, o bien, en caso de que el número tenga más de 5 dígitos, solamente se consideran los primeros 5. Las probabilidades para cada una de las manos de poker posibles se muestran en seguida:

$$\text{Todos diferentes} \frac{10 \times 9 \times 8 \times 7 \times 6}{10^5} = 0.30240$$

$$\text{Un par} \frac{10 \times 1 \times 9 \times 8 \times 7}{10^5} \binom{5}{2} = 0.50400$$

$$\text{Dos pares} \frac{1}{2} \frac{10 \times 1 \times 9 \times 1 \times 8}{10^5} \binom{5}{2} \binom{3}{2} = 0.10800$$

$$\text{Tercia} \frac{10 \times 1 \times 1 \times 9 \times 8}{10^5} \binom{5}{3} = 0.07200$$

$$\text{Full} \frac{10 \times 1 \times 1 \times 9 \times 1}{10^5} \binom{5}{3} \binom{2}{2} = 0.00900$$

$$\text{Poker} \frac{10 \times 1 \times 1 \times 1 \times 9}{10^5} \binom{5}{4} = 0.00450$$

$$\text{Quintilla} \frac{10 \times 1 \times 1 \times 1 \times 1}{10^5} \binom{5}{5} = 0.00010$$

Tabla 3.6 Distribución acumulada teórica experimental de los números pseudoaleatorios presentados en la tabla 3.1.

$F_0(x)$	$F_n(x)$	$F_0(x)$	$F_n(x)$	$F_0(x)$	$F_n(x)$	$F_0(x)$	$F_n(x)$	$F_0(x)$	$F_n(x)$
0.00121	0.01	0.19962	0.21	0.36982	0.41	0.58518	0.61	0.76910	0.81
0.03185	0.02	0.20852	0.22	0.38507	0.42	0.58902	0.62	0.78961	0.82
0.03272	0.03	0.21631	0.23	0.38618	0.43	0.61168	0.63	0.79113	0.83
0.05010	0.04	0.22287	0.24	0.40644	0.44	0.62400	0.64	0.79269	0.84
0.05230	0.05	0.22724	0.25	0.41678	0.45	0.63453	0.65	0.79982	0.85
0.05475	0.06	0.23949	0.26	0.42989	0.46	0.64183	0.66	0.80548	0.86
0.07281	0.07	0.24000	0.27	0.43308	0.47	0.65002	0.67	0.81772	0.87
0.08128	0.08	0.26593	0.28	0.44267	0.48	0.65559	0.68	0.82654	0.88
0.09133	0.09	0.27738	0.29	0.45679	0.49	0.65726	0.69	0.83358	0.89
0.09264	0.10	0.28269	0.30	0.49383	0.50	0.67621	0.70	0.85353	0.90
0.10699	0.11	0.29453	0.31	0.49962	0.51	0.68334	0.71	0.85903	0.91
0.10850	0.12	0.29931	0.32	0.51781	0.52	0.69623	0.72	0.87616	0.92
0.11537	0.13	0.30861	0.33	0.53559	0.53	0.70102	0.73	0.88752	0.93
0.12079	0.14	0.33216	0.34	0.54325	0.54	0.70119	0.74	0.88809	0.94
0.14151	0.15	0.33381	0.35	0.55170	0.55	0.73059	0.75	0.89890	0.95
0.15506	0.16	0.33729	0.36	0.55743	0.56	0.73435	0.76	0.90829	0.96
0.17028	0.17	0.33954	0.37	0.55877	0.57	0.74579	0.77	0.91512	0.97
0.18284	0.18	0.34358	0.38	0.57410	0.58	0.75103	0.78	0.93655	0.98
0.19147	0.19	0.34488	0.39	0.57880	0.59	0.76086	0.79	0.94187	0.99
0.19904	0.20	0.35286	0.40	0.58244	0.60	0.76262	0.80	0.98611	1.00

Con las probabilidades anteriores y con el número de números pseudoaleatorios generados, se puede calcular la frecuencia esperada de cada posible resultado, la cual al compararse con la frecuencia observada produce el estadístico:

$$X_0^2 = \sum_{i=1}^7 \frac{(FO_i - FE_i)^2}{FE_i} \quad (3.13)$$

Si $X_0^2 < X_{\alpha,6}^2$, entonces no se puede rechazar la hipótesis de que los números pseudoaleatorios provienen de una distribución uniforme.

Por ejemplo, si se aplica esta prueba a los números pseudoaleatorios presentados en la tabla 3.1, se obtienen las frecuencias observadas que se muestran en la tabla 3.7. Sin embargo, puesto que las frecuencias esperadas de full, poker y quintilla son menores que 5, entonces es necesario agrupar sus frecuencias con la frecuencia esperada de tercia. Con estas agrupaciones, el valor del estadístico X_0^2 (ecuación 3.13) resulta ser de:

$$X_0^2 = \frac{(23 - 30.24)^2}{30.24} + \frac{(58 - 50.24)^2}{50.24} + \frac{(8 - 10.80)^2}{10.80} + \frac{(11 - 8.56)^2}{8.56} = 4.35$$

y seleccionando un valor de $\alpha = 0.05$, entonces $X_{0.05,3}^2 = 7.81$. Puesto que $X_0^2 < X_{0.05,3}^2$, entonces no se puede rechazar la hipótesis de que los números pseudoaleatorios presentados en la tabla 3.1 provienen de una distribución uniforme.

TABLA 3.7. Frecuencias observadas y esperadas por cada una de las manos de poker.

	<i>Frecuencia observada</i>	<i>Frecuencia esperada</i>
Todos diferentes	23	30.24
Un par	58	50.40
Dos pares	8	10.80
Tercia	10	7.20
Full	1	0.90
Poker	0	0.45
Quintilla	0	0.10

3.7. PRUEBA DE LAS CORRIDAS

Existen dos versiones de la prueba de las corridas: la prueba de corridas arriba y abajo del promedio y la prueba de corridas arriba y abajo.

3.7.1. Prueba de corridas arriba y abajo del promedio

La prueba de corridas arriba y abajo del promedio es un caso ligeramente modificado de la prueba de la distancia en la cual $\alpha = 0$ y $\beta = 0.5$. En esta versión de la prueba de las corridas, una secuencia de números pseudoaleatorios U_1, \dots, U_n es generada. En seguida una secuencia binaria es obtenida, en la cual el i^{th} término es 0 si $U_i < 0.5$ y 1 si $U_i > 0.5$. Una vez obtenida la secuencia binaria, el siguiente paso es determinar la cantidad de veces que una misma longitud de corrida se repite (frecuencia observada de la corrida de longitud i). Una sucesión de i ceros (unos), enmarcada por unos (ceros) en los extremos, representa una corrida de longitud i . El número total esperado de corridas y el número esperado para cada tamaño de corrida, se obtienen con las siguientes expresiones:

$$E(\text{total de corridas}) = \frac{N+1}{2} \quad (3.14)$$

$$FE_i = \frac{(N-i+3)}{2^{i+1}} \quad (3.15)$$

Estas frecuencias esperadas son comparadas con las observadas a través de una distribución chi-cuadrada y una decisión sobre la aleatoriedad de los números pseudoaleatorios generados es tomada.

3.7.2. Prueba de corridas arriba y abajo

En la prueba de corridas arriba y abajo, una secuencia de números pseudoaleatorios U_1, \dots, U_n es generada, y al igual que en el inciso anterior, una secuencia binaria es obtenida, en la cual el i^{th} término es cero si $U_i < U_{i+1}$ y 1 si $U_i > U_{i+1}$. Una vez obtenida la secuencia binaria, se sigue el mismo procedimiento descrito anteriormente y se obtiene la frecuencia observada para cada tamaño de corrida. El número total esperado de corridas y el número esperado para cada tamaño de corrida, se obtienen con las siguientes expresiones:

$$E(\text{total de corridas}) = \frac{2N-1}{3} \quad (3.16)$$

$$FE_i = 2 \left(\frac{(i^2 + 3i + 1)N - (i^3 + 3i^2 - i - 4)}{(i + 3)!} \right) \text{ para } i < N - 1 \quad (3.17)$$

$$FE_{N-1} = \frac{2}{N!} \text{ para } i = N - 1 \quad (3.18)$$

Finalmente, el estadístico X_0 se determina de acuerdo a la siguiente expresión:

$$X_0^2 = \sum_{i=1}^n \frac{(FO_i - FE_i)^2}{FE_i} \quad (3.19)$$

donde n es el número de términos de la ecuación (3.19). Es importante señalar que en el cálculo del estadístico X_0^2 , la frecuencia esperada para cada tamaño de corrida debe ser mayor o igual a cinco. Si las frecuencias esperadas para corridas de tamaño grande son menores que 5, tales frecuencias se deben agrupar con las adyacentes de tal modo que la frecuencia esperada de los tamaños de corrida sea de al menos 5.

Con el propósito de ilustrar una aplicación de esta prueba, considere los números pseudoaleatorios presentados en la tabla 3.1. Para estos números, la secuencia binaria que resulta es la siguiente:

0	1	0	1	0	1	0	1	1	1	0	0	1	1	1	0	1	1	0	1
0	1	1	0	0	1	1	1	0	1	0	0	0	1	1	0	1	0	1	0
1	0	1	0	1	0	1	1	1	1	0	0	1	0	1	0	0	0	1	0
1	0	1	0	0	1	0	0	0	1	0	0	1	0	1	1	1	0	1	0
0	0	0	1	0	1	0	0	0	0	0	1	1	0	0	0	1	1	1	

A partir de esta secuencia binaria se obtienen las frecuencias observadas para cada tamaño de corrida. Estas frecuencias, así como las esperadas se muestran en la tabla 3.8. De acuerdo a estas frecuencias, el valor del estadístico X_0^2 resulta ser de:

$$X_0^2 = \frac{(41 - 41.7)^2}{41.7} + \frac{(11 - 18.1)^2}{18.1} + \frac{(11 - 6.43)^2}{6.43} = 6.04$$

y seleccionando un valor de $\alpha = 0.05$, entonces $X_{0.05,2}^2 = 5.99$. Puesto que $X_0^2 > X_{0.05,2}^2$, entonces se rechaza la hipótesis de uniformidad de los números pseudoaleatorios presentados en la tabla 3.1.

TABLA 3.8 Frecuencias esperadas y observadas para cada uno de los tamaños de corrida.

<i>Tamaño de corrida</i>	<i>Frecuencia observada</i>	<i>Frecuencia esperada</i>
1	41	41.70
2	11	18.10
3	8	5.14
4	2	1.10
5	1	0.19

PROBLEMAS

3.1. Efectuar para los siguientes números pseudoaleatorios, todas las pruebas de aleatoriedad explicadas en el capítulo.

0.03991	0.10461	0.93716	0.16894	0.98953
0.38555	0.95554	0.32886	0.59780	0.09958
0.17546	0.73704	0.92052	0.46215	0.15917
0.32643	0.52861	0.95819	0.06831	0.19640
0.69572	0.68777	0.39510	0.35905	0.85244
0.24122	0.66591	0.27699	0.06494	0.03152
0.61196	0.30231	0.92962	0.61773	0.22109
0.30532	0.21704	0.10274	0.12202	0.94205
0.03788	0.97599	0.75867	0.20717	0.82037
0.48228	0.63379	0.85783	0.47619	0.87481
0.88618	0.19161	0.41290	0.63312	0.71857
0.71299	0.23853	0.05870	0.01119	0.92784
0.27954	0.58909	0.82444	0.99005	0.04921
0.80863	0.00514	0.20247	0.81759	0.45197
0.33564	0.60780	0.48460	0.85558	0.15191
0.90899	0.75754	0.60833	0.25983	0.01291
0.78038	0.70267	0.43529	0.06318	0.38384
0.55986	0.66485	0.88722	0.56736	0.66164
0.87539	0.08823	0.94813	0.31900	0.54155
0.16818	0.60311	0.74457	0.90561	0.72848

4

GENERACION DE VARIABLES ALEATORIAS NO-UNIFORMES

En todo modelo de simulación estocástico, existen una o varias variables aleatorias interactuando. Generalmente, estas variables siguen distribuciones de probabilidad teóricas o empíricas diferentes a la distribución uniforme. Por consiguiente, para simular este tipo de variables, es necesario contar con un generador de números uniformes y una función que a través de un método específico, transforme estos números en valores de la distribución de probabilidad deseada. Existen varios procedimientos para lograr este objetivo. Entre los procedimientos más comunes y más difundidos se pueden mencionar: 1) El método de la transformada inversa, 2) El método de rechazo, 3) El método de composición y 4) Procedimientos especiales.

4.1. METODO DE LA TRANSFORMADA INVERSA

El método de la transformada inversa utiliza la distribución acumulada $F(x)$ de la distribución que se va a simular. Puesto que $F(x)$ está definida en el intervalo $(0; 1)$, se puede generar un número aleatorio uniforme R y tratar de determinar el valor de la variable aleatoria para la cual su distribución acumulada es igual a R , es decir, el valor simulado de la variable aleatoria que sigue una distribución de probabilidad $f(x)$, se determina al resolver la siguiente ecuación (ver figura 4.1):

$$F(x) = R \text{ ó } x = F^{-1}(R) \quad (4.1)$$

La dificultad principal de este método descansa en el hecho de que en algunas ocasiones es difícil encontrar la transformada inversa. Sin

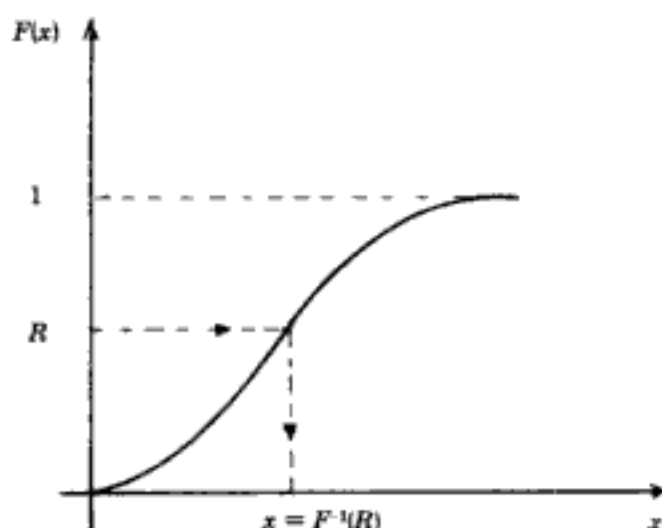


Figura 4.1. Forma gráfica del método de la transformada inversa.

embargo, si esta función inversa ya ha sido establecida, generando números aleatorios uniformes se podrán obtener valores de la variable aleatoria que sigan la distribución de probabilidad deseada.

Ejemplo 4.1. Distribución exponencial.

Se desea generar números al azar que sigan la siguiente distribución de probabilidad.

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} \quad (4.2)$$

La distribución acumulada de esta distribución es:

$$F(x) = \int_0^x \lambda e^{-\lambda t} dt = 1 - e^{-\lambda x}$$

y utilizando la ecuación (4.1), es decir, igualando la distribución acumulada con el número uniforme R , se obtiene:

$$\begin{aligned} 1 - e^{-\lambda x} &= R \\ e^{-\lambda x} &= 1 - R \end{aligned}$$

Pero si R sigue una distribución uniforme, entonces $1-R$ también sigue esta distribución. Por consiguiente:

$$\begin{aligned} e^{-\lambda x} &= R \\ x &= -\frac{1}{\lambda} \ln R \end{aligned} \quad (4.3)$$

Ejemplo 4.2. Distribución uniforme.

Se desea generar números al azar que sigan la siguiente distribución de probabilidad:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{si } a > x > b \end{cases} \quad (4.4)$$

La distribución acumulada de esta distribución es:

$$F(x) = \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a}$$

e igualando esta expresión con el número uniforme R se obtiene:

$$\frac{x-a}{b-a} = R$$

(4.5)

$$x = a + (b-a)R$$

Ejemplo 4.3. Distribución empírica.

Se desea generar números al azar que sigan la siguiente distribución de probabilidad:

$$f(x) = \begin{cases} x & \text{si } 0 \leq x \leq 1 \\ 1/2 & \text{si } 1 < x \leq 2 \end{cases} \quad (4.6)$$

La distribución acumulada de esta distribución es:

$$F(x) = \int_0^x t dt = \frac{x^2}{2} \text{ si } 0 \leq x \leq 1$$

$$\text{y} \quad F(x) = \frac{1}{2} + \int_0^x \frac{1}{2} dt = \frac{x}{2} \text{ si } 1 < x \leq 2$$

puesto que la acumulada de la función tiene un valor de 1/2 cuando $x = 1$ (ver figura 4.2), entonces el valor de la variable aleatoria x se determina de acuerdo a la siguiente expresión:

$$x = \begin{cases} \sqrt{2R} & \text{si } R \leq 1/2 \\ 2R & \text{si } R > 1/2 \end{cases} \quad (4.7)$$

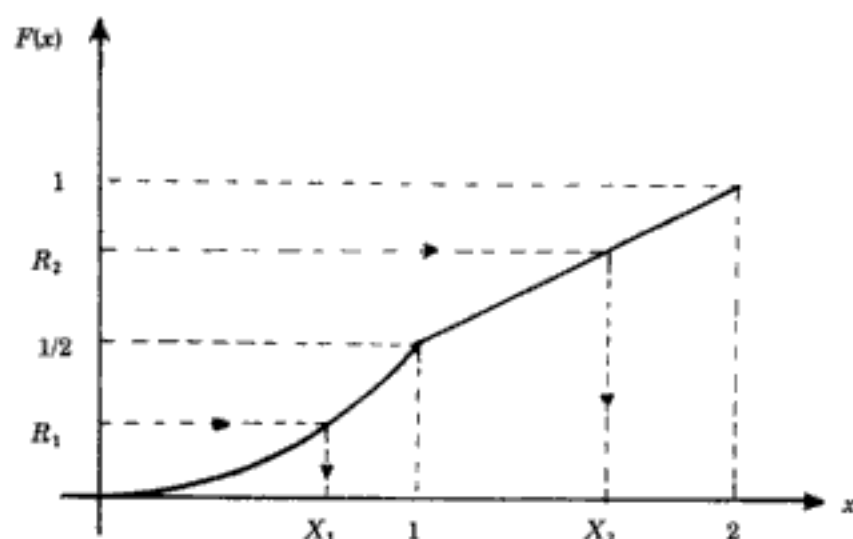


Figura 4.2. Distribución acumulada de la función (4.6).

Tabla 4.1. Distribución de probabilidad y distribución acumulada de la función $e^{-55^x}/x!$.

x	Distribución de probabilidad	Distribución acumulada
0	0.0067	0.0067
1	9.0337	0.0404
2	0.0842	0.1246
3	0.1404	0.2650
4	0.1755	0.4405
5	0.1755	0.6160
6	0.1462	0.7622
7	0.1044	0.8666
8	0.0653	0.9319
9	0.0363	0.9682
10	0.0181	0.9863
11	0.0082	0.9945
12	0.0034	0.9979
13	0.0013	0.9992
14	0.0005	0.9997
15	0.0002	0.9999

Tabla 4.2. Transformada inversa de la función $e^{-5}5^x/x!$.

Si $0.0000 \leq R < 0.0067$, entonces $x = 0$
Si $0.0067 \leq R < 0.0404$, entonces $x = 1$
Si $0.0404 \leq R < 0.1246$, entonces $x = 2$
Si $0.1246 \leq R < 0.2650$, entonces $x = 3$
Si $0.2650 \leq R < 0.4405$, entonces $x = 4$
Si $0.4405 \leq R < 0.6160$, entonces $x = 5$
Si $0.6160 \leq R < 0.7622$, entonces $x = 6$
Si $0.7622 \leq R < 0.8666$, entonces $x = 7$
Si $0.8666 \leq R < 0.9319$, entonces $x = 8$
Si $0.9319 \leq R < 0.9682$, entonces $x = 9$
Si $0.9682 \leq R < 0.9863$, entonces $x = 10$
Si $0.9863 \leq R < 0.9945$, entonces $x = 11$
Si $0.9945 \leq R < 0.9979$, entonces $x = 12$
Si $0.9979 \leq R < 0.9992$, entonces $x = 13$
Si $0.9992 \leq R < 0.9997$, entonces $x = 14$
Si $0.9997 \leq R < 0.9999$, entonces $x = 15$

Ejemplo 4.4. Distribución poisson.

Se desea generar números al azar que sigan la siguiente distribución de probabilidad:

$$f(x) = \frac{e^{-5} 5^x}{x!}, \quad x = 0, 1, 2, \dots \quad (4.8)$$

Puesto que esta distribución de probabilidad es discreta, es necesario evaluar $f(x)$ para cada valor de x , y entonces determinar la distribución acumulada $F(x)$. Tanto la distribución de probabilidad como la distribución acumulada de esta variable aparecen en la tabla 4.1. De acuerdo a esta distribución acumulada, la tabla 4.2 muestra el valor que tomará la variable aleatoria x dependiendo del intervalo al cual pertenece el número uniforme R .

4.2. METODO DE RECHAZO

Existe otro procedimiento para generar números al azar de distribuciones de probabilidad no-uniformes. A este procedimiento se le conoce

con el nombre de método de rechazo. Este método consiste primeramente en generar un valor de la variable aleatoria y en seguida probar que dicho valor simulado proviene de la distribución de probabilidad que se está analizando. Para comprender la lógica de este método, suponga que $f(x)$ (ver figura 4.3) es una distribución de probabilidad acotada y con rango finito, es decir, $a \leq x \leq b$. De acuerdo a esta función de probabilidad, la aplicación del método de rechazo implica el desarrollo de los siguientes pasos:

1. Generar dos números uniformes R_1 y R_2 .
2. Determinar el valor de la variable aleatoria x de acuerdo a la siguiente relación lineal de R_1 :

$$x = a + (b - a)R_1 \quad (4.9)$$

3. Evaluar la función de probabilidad en $x = a + (b - a)R_1$.
4. Determinar si la siguiente desigualdad se cumple:

$$R_2 \leq f[a + (b - a)R_1]/M \quad (4.10)$$

Se utiliza a $x = a + (b - a)R_1$ si la respuesta es afirmativa como un valor simulado de la variable aleatoria. De lo contrario, es necesario pasar nuevamente al paso 1 tantas veces como sea necesario.

La teoría sobre la que se apoya este método se basa en el hecho de que la probabilidad de que $R_2 \leq f(x)/M$ es exactamente $f(x)/M$. Por consiguiente, si un número es escogido al azar de acuerdo a $x = a + (b - a)R_1$ y

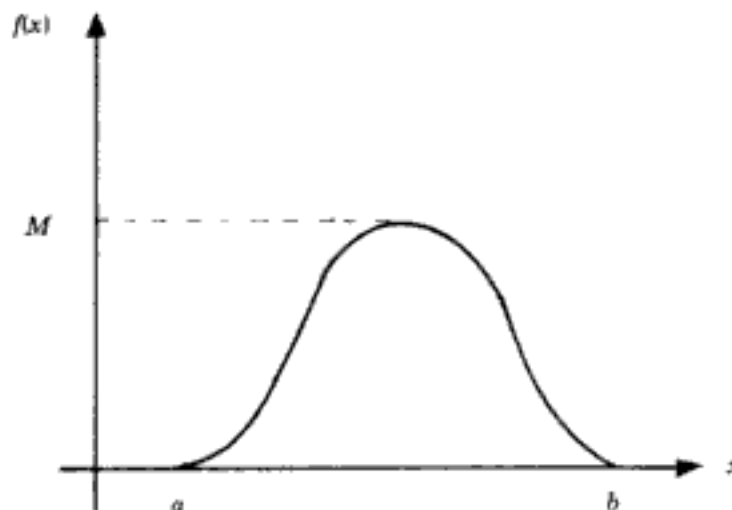


Figura 4.3. Distribución de probabilidad con rango finito y con moda M .

rechazado si $R_2 > f(x)/M$, entonces la distribución de probabilidad de las x 's aceptadas será exactamente $f(x)$. Por otra parte, conviene señalar que si todas las x 's fueran aceptadas, entonces x estaría uniformemente distribuida entre a y b .

Finalmente, es necesario mencionar que algunos autores como Tocher, han demostrado que el número esperado de intentos para que x sea aceptada como una variable aleatoria que sigue una distribución de probabilidad $f(x)$, es M . Esto significa que este método podría ser un tanto ineficiente para ciertas distribuciones de probabilidad en las cuales la moda sea grande.

Ejemplo 4.5. Distribución empírica.

Se desea generar números al azar que sigan la siguiente distribución de probabilidad:

$$f(x) = \begin{cases} 2x & \text{si } 0 \leq x \leq 1 \\ 0 & \text{si } 0 > x > 1 \end{cases} \quad (4.11)$$

Para esta función, $a = 0$, $b = 1$ y $M = 2$. Por consiguiente, aplicando los pasos descritos previamente se tiene:

1. Generar dos números uniformes R_1 y R_2 .
2. Calcular $x = R_1$.
3. Es $R_2 \leq R_1$? Si la respuesta es afirmativa, entonces $x = R_1$ es un valor simulado de la variable aleatoria. De lo contrario, se requiere regresar al paso 1 tantas veces como sea necesario.

Ejemplo 4.6. Distribución triangular.

Se desea generar números al azar que sigan la siguiente distribución de probabilidad:

$$f(x) = \begin{cases} \frac{2}{(c-a)(b-a)}(x-a) & \text{si } a \leq x \leq b \\ \frac{-2}{(c-a)(c-b)}(x-c) & \text{si } b \leq x \leq c \end{cases} \quad (4.12)$$

Para esta distribución de probabilidad, $M = 2/(c-a)$. Sin embargo, esta distribución está compuesta de dos funciones: una válida en el rango $a \leq x \leq b$ y la otra válida en $b \leq x \leq c$. Por consiguiente, los pasos necesarios para simular esta distribución por el método de rechazo serían:

1. Generar R_1 y R_2 .
2. Calcular $x = a + (c-a)R_1$.
3. Es $x < b$? Si la respuesta es afirmativa, $f(x)$ sería:

$$f(x) = \frac{2}{(c-a)(b-a)} (a + (c-a)R_1 - a) = \frac{2R_1}{(b-a)} \quad (4.13)$$

Por el contrario, si la respuesta es negativa, $f(x)$ sería:

$$f(x) = \frac{-2}{(c-a)(c-b)} (a + (c-a)R_1 - c) = \frac{2(1-R_1)}{(c-b)} \quad (4.14)$$

4. Es $R_2 < \frac{f(x)(c-a)}{2}$? Si la respuesta es afirmativa, entonces $x = a + (c-a)R_1$ se considera como un valor simulado de la variable aleatoria. De lo contrario se requiere regresar al paso 1 tantas veces como sea necesario.

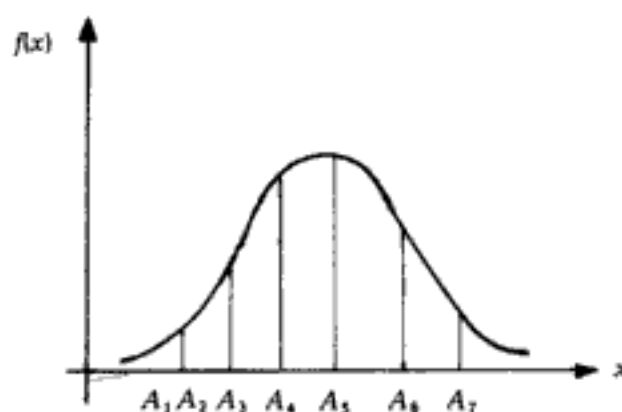
4.3. METODO DE COMPOSICION

Otro método para generar valores de variables aleatorias no-uniformes es el método de composición. Mediante este método la distribución de probabilidad $f(x)$ se expresa como una mezcla de varias distribuciones de probabilidad $f_i(x)$ seleccionadas adecuadamente.

El procedimiento para la selección de las $f_i(x)$ se basa en el objetivo de minimizar el tiempo de computación requerido para la generación de valores de la variable aleatoria analizada.

Los pasos requeridos para la aplicación de este método en la simulación de variables aleatorias no-uniformes son los siguientes:

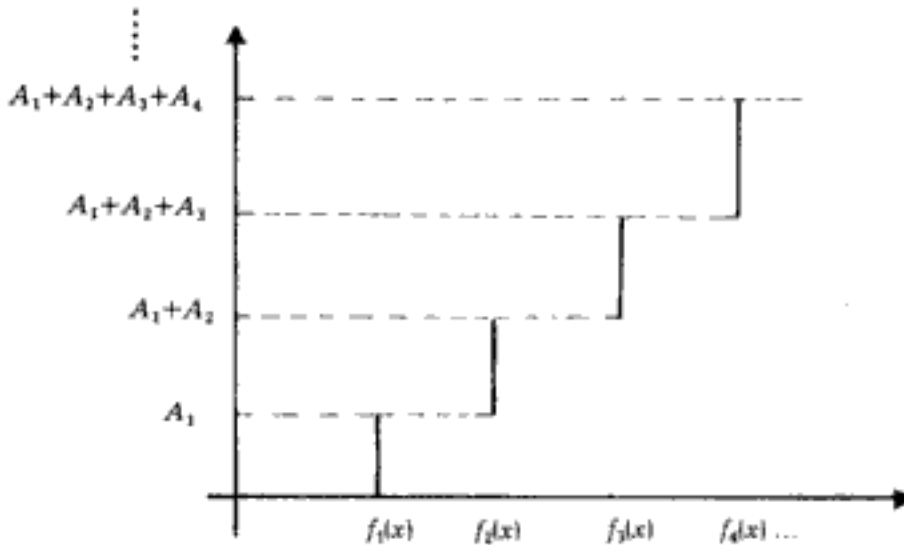
1. Dividir la distribución de probabilidad original en sub-áreas, tal como se muestra en la figura siguiente:



- Definir una distribución de probabilidad para cada sub-área.
- Expresar la distribución de probabilidad original en la forma siguiente:

$$f(x) = A_1 f_1(x) + A_2 f_2(x) + \dots + A_n f_n(x) \text{ y } \sum A_i = 1 \quad (4.15)$$

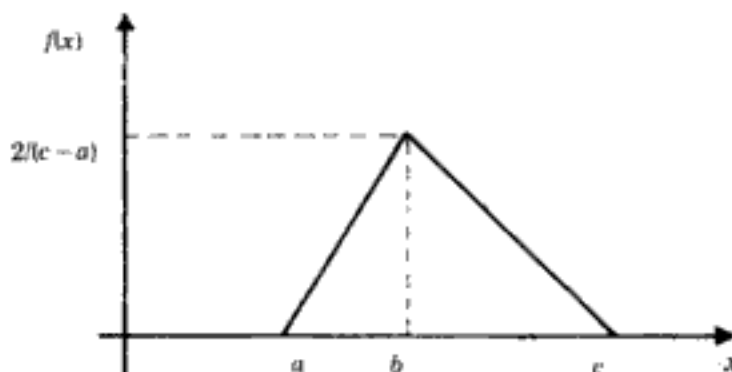
- Obtener la distribución acumulada de las áreas.



- Generar dos números uniformes R_1, R_2 .
- Seleccionar la distribución de probabilidad $f_i(x)$ con la cual se va a simular el valor de x . La selección de esta distribución se obtiene al aplicar el método de la transformada inversa, en el cual el eje Y está representado por la distribución acumulada de las áreas, y el eje X por las distribuciones $f_i(x)$. Para esta selección se utiliza el número uniforme R_1 .
- Utilizar el número uniforme R_2 para simular por el método de la transformada inversa o algún otro procedimiento especial, números al azar que sigan la distribución de probabilidad $f_i(x)$ seleccionada en el paso anterior.

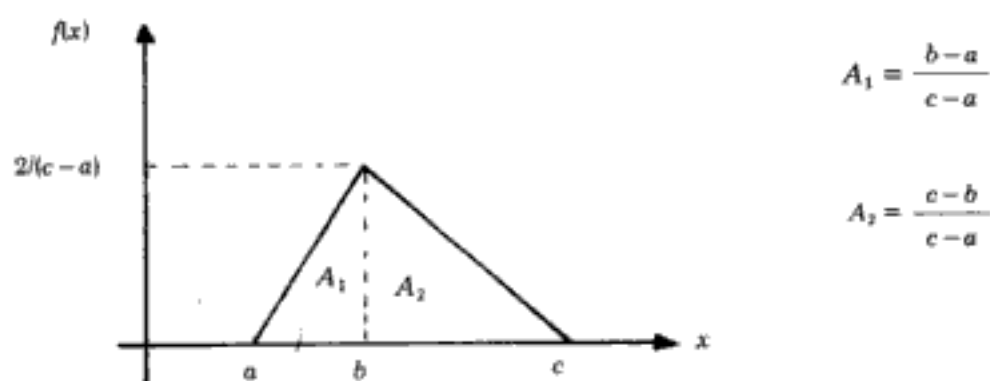
Ejemplo 4.7. Distribución triangular.

Se desea generar números al azar que sigan la siguiente distribución de probabilidad:

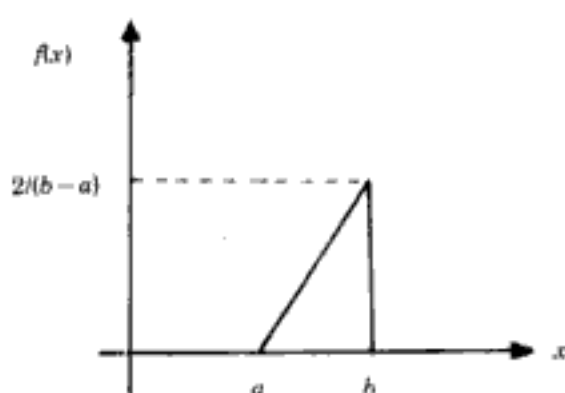


Siguiendo los pasos descritos previamente, la generación de números al azar que sigan esta distribución triangular, puede ser resumida en los siguientes pasos:

1. La distribución de probabilidad original, se va a dividir en dos áreas, tal como se muestra en la figura siguiente:

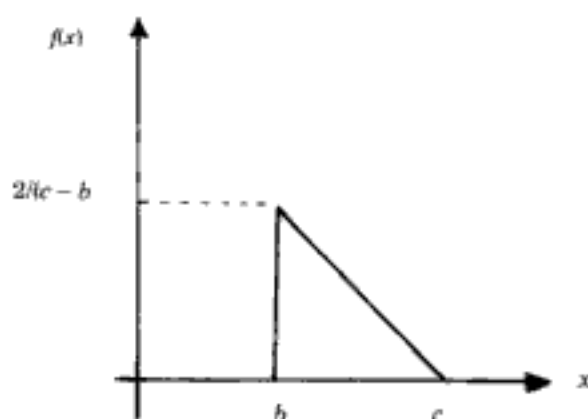


2. En seguida se determinan las distribuciones de probabilidad y distribución acumulada de las áreas definidas en el paso anterior:



$$f_1(x) = \frac{2}{(b-a)^2} (x-a) \quad (4.16)$$

$$F_1(x) = \frac{(x-a)^2}{(b-a)^2} \quad (4.17)$$



$$f_2(x) = \frac{-2}{(c-b)^2} (x-c) \quad (4.18)$$

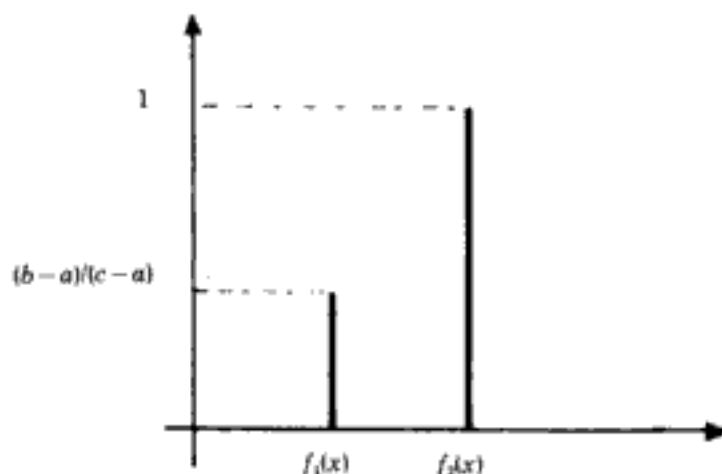
$$F_2(x) = 1 - \frac{(x-c)^2}{(c-b)^2} \quad (4.19)$$

3. Posteriormente la distribución de probabilidad original se expresa como:

$$f(x) = \frac{(b-a)}{(c-a)} \left(\frac{2}{(b-a)^2} (x-a) \right) + \frac{(c-b)}{(c-a)} \left(\frac{2}{(c-b)^2} (x-c) \right)$$

$$f(x) = \frac{2}{(c-a)(b-a)} (x-a) + \frac{2}{(c-a)(c-b)} (x-c) \quad (4.20)$$

4. Con las áreas y distribuciones $f_i(x)$ definidas en los pasos anteriores, la distribución acumulada de las áreas sería:



5. Generar dos números uniformes R_1 y R_2 .
 6. Es $R_1 < (b-a)/(c-a)$? Si la respuesta es afirmativa, entonces se simulan valores de la distribución $f_1(x)$:

$$(x-a)^2/(b-a)^2 = R_2$$

$$x = a + (b-a) \sqrt{R_2} \quad (4.21)$$

Si la respuesta es negativa, entonces se simulan valores de la distribución $f_2(x)$:

$$1 - (c-x)^2/(c-b)^2 = R_2$$

$$x = c - (c-b) \sqrt{1-R_2}$$

$$x = c - (c-b) \sqrt{R_2} \quad (4.22)$$

7. Repetir los pasos anteriores tantas veces como se desee.

4.4. PROCEDIMIENTOS ESPECIALES

Existen algunas distribuciones como la distribución erlang, la distribución normal, etc., cuya simulación a través del método de la transformada inversa sería demasiado complicado. Para éstas y algunas otras distribuciones, es posible utilizar algunas de sus propiedades para facilitar y agilizar el proceso de generación de números al azar.

Ejemplo 4.8. Distribución normal.

Se desea generar números al azar que sigan la siguiente distribución de probabilidad:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{para } -\infty < x < \infty \quad (4.23)$$

Puesto que no es posible expresar la distribución acumulada de la distribución normal en forma explícita, entonces no es posible utilizar para la generación de números al azar, el método de la transformada inversa. En lugar de este método, se puede hacer uso del teorema del límite central, el cual establece que la suma de n variables aleatorias independientes se aproxima a una distribución normal a medida que n se aproxima a infinito. Lo anterior expresado en forma de teorema sería: Si x_1, x_2, \dots, x_n es una secuencia de n variables aleatorias independientes con $E(x_i) = \mu_i$ y $\text{var}(x_i) = \sigma_i^2$ (ambas finitas) y $Y = a_1 x_1 + a_2 x_2 + \dots + a_n x_n$, entonces bajo ciertas condiciones generales:

$$Z = \frac{Y - \sum_{i=1}^n a_i \mu_i}{\sqrt{\sum_{i=1}^n a_i^2 \sigma_i^2}} \quad (4.24)$$

tiene una distribución normal estándar a medida que n se aproxima a infinito.

Si las variables que se están sumando son uniformes en el intervalo $(0; 1)$, entonces.

$$Z = \frac{\sum_{i=1}^n R_i - n/2}{\sqrt{n/12}} \quad (4.25)$$

tiene una distribución normal estándar. Puesto que la normal estándar de una variable aleatoria x distribuida normalmente se obtiene como:

$$Z = \frac{x - \mu}{\sigma} \quad (4.26)$$

entonces, la simulación de la variable aleatoria x se haría de acuerdo a la siguiente expresión.

$$x = \mu + \sigma \left(\frac{\sum_{i=1}^n R_i - n/2}{\sqrt{n/12}} \right) \quad (4.27)$$

Finalmente, se ha comprobado que utilizando un valor de $n = 12$, la confiabilidad de los valores simulados es bastante aceptable. Además, es muy poco probable obtener valores simulados de x en la región $\mu + 6\sigma < x < \mu - 6\sigma$. También, es obvio que utilizando un valor de $n = 12$, la expresión (4.27) se simplifica a:

$$x = \mu + \sigma \left(\sum_{i=1}^{12} R_i - 6 \right) \quad (4.28)$$

Ejemplo 4.9. Distribución erlang.

Se desea generar números al azar que sigan la siguiente distribución de probabilidad:

$$f(x) = \frac{(\lambda n)^n}{(n-1)!} X^{n-1} e^{-\lambda n x}, \text{ para } x \geq 0 \quad (4.29)$$

donde n y λ son parámetros positivos, y además, el valor de n está restringido a ser entero.

Ha sido demostrado por algunos matemáticos que esta distribución es justamente la suma de n variables aleatorias exponenciales cada una con valor esperado $1/\lambda$. Por consiguiente, para generar números al azar que sigan una distribución erlang, se necesita solamente sumar los valores simulados de n variables aleatorias exponenciales con media $1/\lambda$, es decir:

$$x = \sum_{i=1}^n x_i = -\frac{1}{\lambda} \sum_{i=1}^n \ln R_i = -\frac{1}{\lambda} \ln \prod_{i=1}^n R_i \quad (4.30)$$

donde las x_i 's siguen una distribución exponencial y han sido generadas por el método de la transformada inversa (ver ecuación 4.3).

Ejemplo 4.10. Distribución binomial.

Se desea generar números al azar que sigan la siguiente distribución de probabilidad:

$$f(x) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}, \text{ para } x = 0, 1, \dots, n \quad (4.31)$$

Puesto que la distribución binomial es justamente la suma de n variables aleatorias Bernoulli, entonces la generación de números al azar que sigan una distribución binomial, implica sumar los valores simulados de n variables aleatorias Bernoulli. El procedimiento para generar esta distribución puede ser descrita en los siguientes pasos:

1. Generar n números uniformes R .
2. Contar cuántos de estos números generados son menores que θ .
3. La cantidad encontrada en el paso 2, es el valor simulado de la variable aleatoria x .
4. Repetir los pasos anteriores tantas veces como se desee.

Ejemplo 4.11. Distribución poisson.

Ya se ha explicado como simular una distribución poisson por el método de la transformada inversa. Sin embargo, la distribución poisson se puede simular de otra manera si se hace uso de la relación existente entre esta distribución y la distribución exponencial. Se puede demostrar que si: 1) El número total de eventos que ocurren durante un intervalo de tiempo dado es independiente del número de eventos que ya han ocurrido previamente al inicio del intervalo y 2) La probabilidad de que un evento ocurra en el intervalo de t a $t + \Delta t$ es aproximadamente $\lambda \Delta t$ para todos los valores de t , entonces: 1) La distribución de probabilidad del tiempo entre eventos es $f(t) = \lambda e^{-\lambda t}$ y 2) La probabilidad de que ocurran x eventos durante el tiempo T es:

$$f(x)^* = \frac{e^{-\lambda} \lambda^x}{x!} \quad (4.32)$$

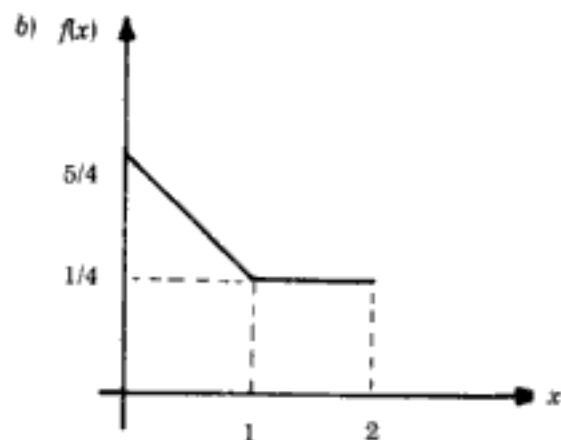
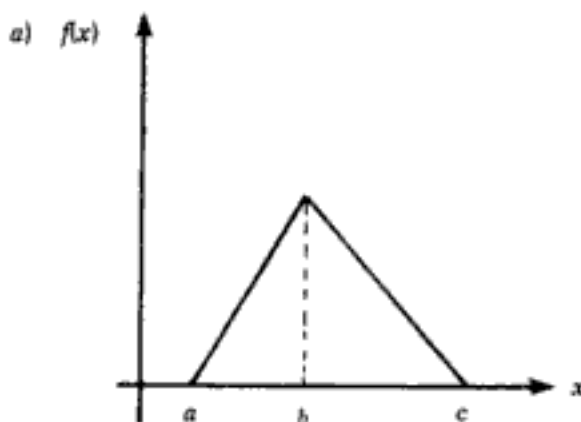
* λ es el número de eventos promedio que ocurre durante el tiempo T .

Por consiguiente, la simulación de una distribución poisson haciendo uso de la relación anterior, se haría de acuerdo a los siguientes pasos:

1. Definir el tiempo T .
2. Simular mediante el método de la transformada inversa, números al azar que sigan una distribución exponencial con media $1/\lambda$.
3. Sumar los tiempos entre eventos simulados en el paso anterior de modo que esta suma no sea mayor que T .
4. Contar de acuerdo al paso anterior, el número de eventos que ocurrieron durante el tiempo T .
5. Repetir los pasos anteriores tantas veces como se desee.

PROBLEMAS

- 4.1. Generar por el método de la transformada inversa, números al azar que sigan las siguientes distribuciones de probabilidad:

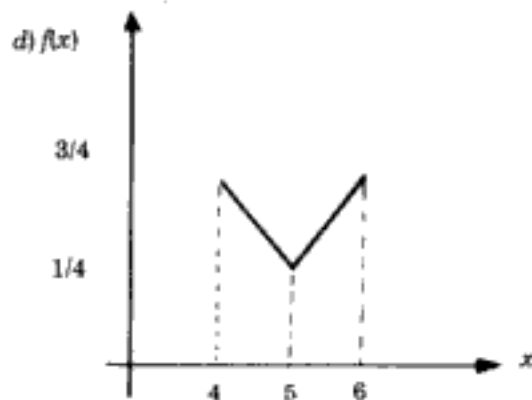
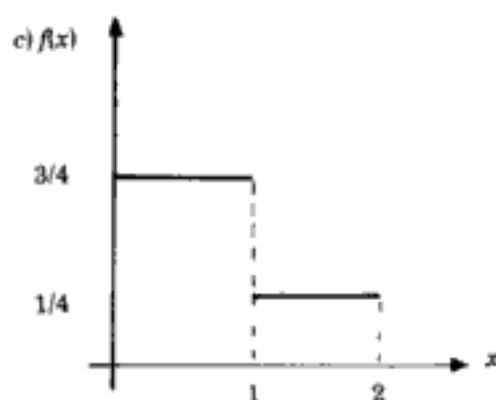
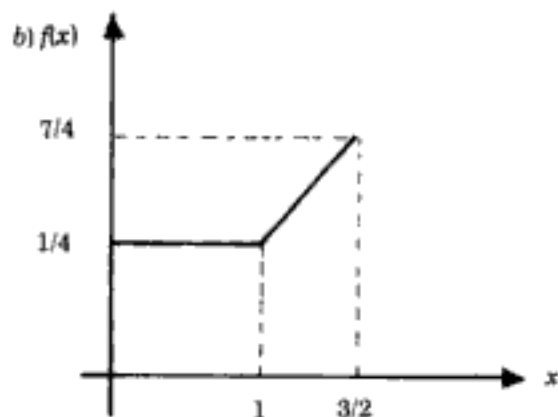
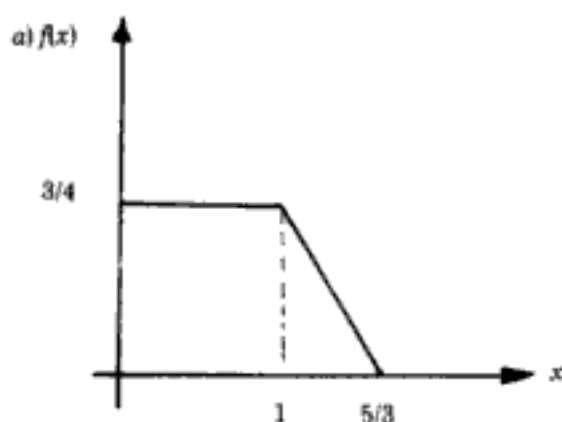


$$c) f(x) = \begin{cases} (x-3)^2/18 & \text{si } 0 \leq x \leq 6 \\ 0 & \text{si } 6 < x < \infty \end{cases}$$

$$d) f(x) = (2\lambda)^2 x e^{-2\lambda x}, x \geq 0$$

- 4.2. Generar por el método de rechazo, números al azar que sigan las siguientes distribuciones de probabilidad:

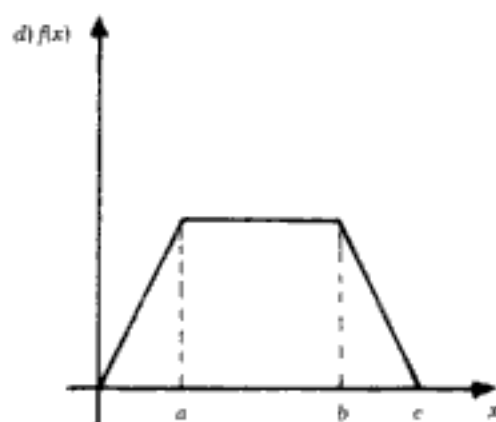
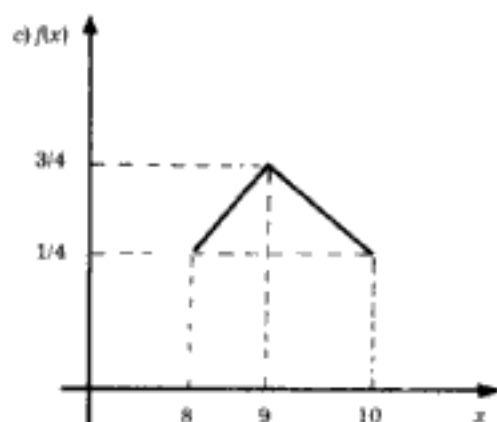
64 Generación de variables aleatorias no-uniformes



4.3. Generar por el método de composición, números al azar que sigan las siguientes distribuciones de probabilidad:

a) $f(x) = p\beta_1 e^{-\beta_1 x} + (1-p)\beta_2 e^{-\beta_2 x}$, $0 \leq x < \infty$ y $0 \leq p \leq 1$

b) $f(x) = p \binom{n}{x} \theta_1^x (1-\theta_1)^{n-x} + (1-p) \binom{n}{x} \theta_2^x (1-\theta_2)^{n-x}$, $0 \leq x \leq n$ y $0 \leq p \leq 1$



4.4. Generar utilizando procedimientos especiales, números al azar que sigan distribuciones de probabilidad:

$$a) f(x) = p q^x, x = 0, 1, 2, \dots \quad p + q = 1$$

$$b) f(x) = \frac{1}{2^{\gamma/2} \Gamma(\gamma/2)} x^{\frac{\gamma-2}{2}} e^{-\frac{x}{2}}, x \geq 0$$

$$c) f(x) = \frac{\Gamma\left(\frac{\gamma_1 + \gamma_2}{2}\right)}{\Gamma\left(\frac{\gamma_1}{2}\right) \Gamma\left(\frac{\gamma_2}{2}\right)} \left(\frac{\gamma_1}{\gamma_2}\right)^{\gamma_1/2} x^{\frac{\gamma_1}{2}-1} \left(1 + \frac{\gamma_1}{\gamma_2} x\right)^{-\frac{1}{2}(\gamma_1 + \gamma_2)}, x \geq 0$$

$$d) f(x) = \frac{\Gamma\left(\frac{\gamma+1}{2}\right)}{\sqrt{\pi\gamma} \Gamma\left(\frac{\gamma}{2}\right)} \left(1 + \frac{x^2}{\gamma}\right)^{-\frac{\gamma+1}{2}}, -\infty < x < \infty$$

5

APLICACIONES DE SIMULACION

En capítulos anteriores, se han explicado y analizado los conceptos y herramientas requeridas por un estudio de simulación. Por consiguiente, ahora ya estamos preparados para empezar a discutir el papel tan importante que los modelos de simulación juegan en el análisis y estudio de varios tipos de sistemas.

Los modelos de simulación presentan la ventaja de poder ser manipulados en diferentes formas, que serían imposibles, imprácticas y demasiado costosas si se realizaran a través de otra metodología. Por ejemplo, se puede simular la operación de un sistema sin que éste aún exista, también se puede a través del uso de simulación determinar el sitio de localización de un almacén, determinar políticas óptimas de inventarios cuando la demanda y el tiempo de entrega son estocásticos, etc. En general, se puede decir que los modelos de simulación se usan para: 1) Análisis de sistemas, 2) Diseño de sistemas, 3) Síntesis de sistemas y 4) Entrenamiento.

El presente capítulo, tiene como objetivo presentar una serie de ejemplos, a través de los cuales se podrá tener una idea de los múltiples usos y de la utilidad que se puede obtener al utilizar la técnica de simulación.

Ejemplo 5.1. Juego de volados

Existe un método viejo para jugar volados, que consiste en doblar la apuesta cada vez que se pierde. Por ejemplo, si se apuesta \$ X y se pierde, entonces, se apuesta \$ $2X$; si en esta ocasión se vuelve a perder, entonces, se apuesta \$ $4X$ y así sucesivamente. Sin embargo, si al seguir esta política sucede que la apuesta es mayor que la cantidad de que se dispone, entonces, se apuesta lo que se tiene disponible. Por el

contrario, cada vez que se gane, la apuesta será de \$X. Si la cantidad inicial disponible es de \$30, la apuesta es de \$10, la ganancia es igual a la cantidad apostada, la probabilidad de ganar en un volado es 0.5 y se quiere tener \$50, ¿cuál es la probabilidad de llegar a la meta?

Para la información anterior, la tabla 5.1 muestra los resultados de una simulación manual para este juego de volados. Como se puede observar en esta tabla, en la primera "corrida" se necesitaron dos lanzamientos para llegar a la meta de tener \$50. El primer lanzamiento de esta corrida el cual está representado por el número aleatorio 0.03991, significa que se ganó en el volado (números aleatorios menores de 0.5 representan al evento ganar, y números aleatorios mayores que 0.5 representan al evento perder), y por consiguiente, la apuesta en el siguiente volado será de \$10. Como el número aleatorio de este segundo volado es menor que 0.5, entonces, con este lanzamiento se llega a la meta de \$50.

Por otra parte, la corrida número 3 representa una situación de quiebra, es decir, esta corrida representa una situación en la que se pierde completamente la cantidad inicial disponible. En esta corrida, se pierde el primer volado, y por consiguiente, la apuesta del segundo volado es \$20. Como el número aleatorio de este lanzamiento es menor que 0.5, entonces, se gana en el segundo volado, y la apuesta del tercer volado será de \$10. Como el número aleatorio de este volado es mayor que 0.5, entonces, se pierde en el tercer volado, y la apuesta del cuarto volado será de \$20. Puesto que el número aleatorio de este volado es mayor que 0.5, entonces, se pierde en el cuarto volado, y la apuesta será de \$40. Sin embargo, la cantidad que se tiene al final del cuarto volado es de \$10. Por consiguiente, la cantidad apostada en el quinto volado será de \$10. Como esta cantidad es perdida en el último lanzamiento, entonces, se llega finalmente a una situación de quiebra.

Finalmente, en esta tabla se puede observar que de 10 corridas, en 6 ocasiones se llegó a la meta. Esto significa que la probabilidad de llegar a la meta es de 0.6. Sin embargo, es necesario señalar que esta estimación no es muy confiable por el número tan reducido de corridas que se simularon. Para tomar una decisión que garantice buenos resultados, es necesario aumentar significativamente el número de corridas simuladas. También, vale la pena mencionar que este problema sería demasiado difícil de resolver analíticamente, sin embargo, a través del uso de simulación se facilita tremendamente la estimación de la probabilidad buscada.

Ejemplo 5.2. Camión transportador

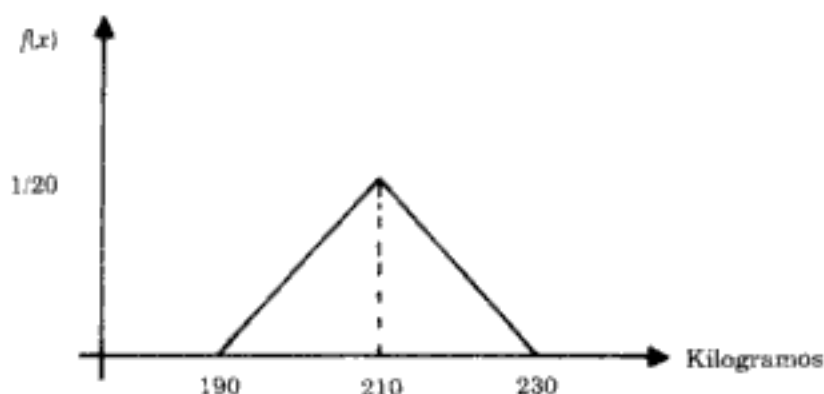
La empresa TIBASA (Fabricante de tinas de baño) tiene asignado un camión especial para el transporte de tinas terminadas. Dicho ca-

Tabla 5.1 Resultados de la simulación manual de un juego de volados

<i>Número de corrida</i>	<i>Cantidad que se tiene antes del volado</i>	<i>Apuesta</i>	<i>Número* aleatorio</i>	<i>¿Se ganó el juego?</i>	<i>Cantidad que se tiene después del volado</i>	<i>¿Se llegó a la meta?</i>
1	\$30	\$10	0.03991	sí	\$40	
	40	10	0.38555	sí	50	sí
2	30	10	0.17546	sí	40	
	40	10	0.32643	sí	50	sí
3	30	10	0.69572	no	20	
	20	20	0.24122	sí	40	
	40	10	0.61196	no	30	
	30	20	0.88722	no	10	
	10	10	0.65961	no	0	quiebra
4	30	10	0.03788	sí	40	
	40	10	0.48288	sí	50	sí
5	30	10	0.88618	no	20	
	20	20	0.71299	no	0	quiebra
6	30	10	0.27954	sí	40	
	40	10	0.80863	no	30	
	30	20	0.33564	sí	50	sí
7	30	10	0.90899	no	20	
	20	20	0.78038	no	0	quiebra
8	30	10	0.55986	no	20	
	20	20	0.87539	no	0	quiebra
9	30	10	0.16818	sí	40	
	40	10	0.34677	sí	50	sí
10	30	10	0.43305	sí	40	
	40	10	0.59747	no	30	
	30	20	0.16520	sí	50	sí

*Estos números aleatorios se obtuvieron del apéndice.

mión transporta diariamente 5 tinas. El peso de cada tina sigue la siguiente distribución de probabilidad:



Si la capacidad del camión es de 1 tonelada, ¿cuál es la probabilidad de que el peso de las tinas exceda la capacidad del camión?

La pregunta anterior, puede ser contestada en forma analítica o a través de simulación. Sin embargo, analicemos primero la obtención de esta probabilidad a través de un procedimiento analítico. El primer paso en la obtención de tal probabilidad, sería la determinación de la distribución de probabilidad $f(x)$. Esta distribución de probabilidad está definida por:

$$f(x) = \begin{cases} \frac{1}{400} (x - 190), & \text{si } 190 \leq x \leq 210 \\ -\frac{1}{400} (x - 230), & \text{si } 210 \leq x \leq 230 \end{cases}$$

En seguida, es necesario determinar la media y la variancia, las cuales pueden ser obtenidas de las siguientes expresiones:

$$E(x) = \frac{1}{400} \int_{190}^{210} (x - 190) dx - \frac{1}{400} \int_{210}^{230} (x - 230) dx = 210$$

$$VAR(x) = \frac{1}{400} \int_{190}^{210} (x - 210)^2 (x - 190) dx - \frac{1}{400} \int_{210}^{230} (x - 210)^2 (x - 230) dx = 66.67$$

Ahora, para encontrar la probabilidad de que la suma de los pesos de 5 tinas exceda la capacidad del camión, suponga que x_i representa el peso de la tina i . Por consiguiente, se desea encontrar la siguiente probabilidad:

$$\text{Prob. } (x_1 + x_2 + x_3 + x_4 + x_5 \geq 1000)$$

si ambos miembros de la desigualdad anterior se dividen por 5, se obtiene:

$$\text{Prob. } (\bar{x} \geq 200)$$

y por el teorema del límite central, sabemos que para distribuciones simétricas con valores de $n \geq 4$, \bar{x} sigue aproximadamente una distribución normal con media μ y desviación estándar σ/\sqrt{n} . Por consiguiente, los parámetros de la distribución de \bar{x} para este caso particular, serían:

$$E(\bar{x}) = 210$$

$$\text{VAR}(\bar{x}) = 13.33$$

y la probabilidad buscada sería:

$$\begin{aligned} \text{Prob. } (\bar{x} \geq 200) &= \text{Prob. } (Z \geq \frac{200 - 210}{3.65}) \\ &= \text{Prob. } (Z \geq -2.74) \\ &= 0.99693 \end{aligned}$$

La estimación de la probabilidad de que el peso de las tinas exceda la capacidad del camión, también puede ser obtenida a través del uso de simulación. Para este propósito, es necesario simular el peso de 5 tinas y compararlo con la capacidad del camión. El procedimiento anterior es necesario repetirlo tantas veces como se desee. Una vez que se hayan realizado n corridas, la estimación de esta probabilidad sería x/n , donde x representa el número de veces que el peso de las 5 tinas excedió la capacidad del camión.

Para simular el peso de una tina, es necesario utilizar alguno de los procedimientos descritos en el capítulo anterior. Suponga que para este caso particular, se va a utilizar el método de la transformada inversa. Por consiguiente, es necesario primero determinar la distribución acumulada del peso de una tina. Tal distribución se muestra a continuación:

$$F(x) = \begin{cases} \frac{1}{800} (x - 190)^2 & \text{si } 190 \leq x \leq 210 \\ 1 - \frac{1}{800} (x - 230)^2, & \text{si } 210 \leq x \leq 230 \end{cases}$$

Con la distribución acumulada definida, el procedimiento para simular el peso de una tina sería el siguiente:

1. Generar un número uniforme R .
2. Preguntar si el número uniforme R es menor que 0.5*. Si la respuesta es afirmativa, entonces:

$$x = 190 + \sqrt{800R}$$

Por el contrario, si la respuesta es negativa, entonces:

$$x = 230 - \sqrt{800(1 - R)}$$

3. Repetir los pasos anteriores tantas veces como se desee.

Aplicando el procedimiento anterior, la tabla 5.2 muestra los resultados de una simulación manual para este problema. Como se puede observar en esta tabla, en todas las corridas se excedió la capacidad del camión. Esto, a primera vista, podría indicarnos que la probabilidad de que la capacidad del camión sea excedida es de 100%. Sin embargo, lo anterior es falso, puesto que no es posible considerar como verdadero, los resultados obtenidos en un número muy reducido de corridas. Es obvio, que si el número de corridas se aumenta significativamente, el porcentaje de las veces que la capacidad del camión es excedida, debe tender a 99.69%.

En este mismo problema, es posible determinar a través de procedimientos analíticos o de simulación, la conveniencia de adquirir un nuevo camión, o seguir operando con el actual. Para tal propósito, suponga que cada vez que la capacidad del camión es excedida, una tina es enviada a través de otra compañía a un costo de \$200/tina. También, suponga que el costo promedio anual de un nuevo camión es de \$60,000. Si se trabajan 5 días a la semana y 52 semanas al año, ¿cuál de las dos alternativas mencionadas es la más atractiva?

Si se utilizan procedimientos analíticos, el costo esperado de enviar la mercancía excedente a través de la otra compañía, sería:

$$\left(\frac{\$200}{\text{tina}} \right) \left(\frac{0.99693}{\text{día}} \right) \left(\frac{5 \text{ días}}{\text{semana}} \right) \left(\frac{52 \text{ semanas}}{\text{año}} \right) = \$51,840$$

el cual resulta ser menor que el costo promedio anual de un nuevo camión. Por consiguiente, la decisión es seguir enviando los excedentes a través de la otra compañía transportista.

*La distribución acumulada toma un valor de 0.5 cuando $x = 210$.

TABLA 5.2 Resultados de la simulación manual del camión de carga

<i>Número de corrida</i>	<i>Tina</i>	<i>Número* aleatorio</i>	<i>Peso simulado de la tina</i>	<i>Peso simulado acumulado</i>	<i>¿Se excedió la capacidad del camión?</i>
1	1	0.31751	206	206	
	2	0.88491	220	426	
	3	0.30934	206	632	
	4	0.22888	204	836	
	5	0.78212	217	1,053	sí
2	1	0.70014	215	215	
	2	0.37239	207	422	
	3	0.18637	202	624	
	4	0.05327	197	821	
	5	0.95096	224	1,045	sí
3	1	0.43253	209	209	
	2	0.80854	218	427	
	3	0.80088	217	644	
	4	0.80890	218	862	
	5	0.93128	223	1,085	sí
4	1	0.41849	208	208	
	2	0.46352	209	417	
	3	0.11087	199	616	
	4	0.52701	211	827	
	5	0.57275	212	1,039	sí
5	1	0.20857	203	203	
	2	0.15633	201	404	
	3	0.92694	222	626	
	4	0.77613	217	843	
	5	0.38688	208	1,051	sí

*Estos números aleatorios se obtuvieron del apéndice.

Este mismo problema, puede ser resuelto por medio de simulación. A través de esta técnica, los pasos requeridos para determinar el costo promedio anual de enviar los excedentes (1 tina) utilizando los servicios de otra compañía, serían:

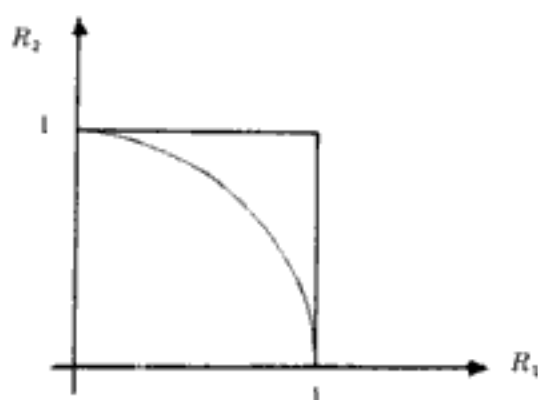
1. Simular el peso de 5 tinas.
2. ¿El peso simulado de las 5 tinas, excede la capacidad del camión? Si la respuesta es afirmativa, entonces, se acumula un

- costo de \$200. Si la respuesta es negativa, no se acumula ningún costo para ese día.
3. Repetir el paso anterior para los 260 días laborables en el año, y determinar el costo acumulado anual.
 4. Repetir los pasos anteriores para varios años.
 5. Determinar, a partir de los costos acumulados obtenidos en los pasos anteriores un costo promedio anual.
 6. Comparar el costo promedio anual obtenido en el paso anterior, con el costo promedio anual del nuevo camión. Si el costo promedio anual obtenido en el paso anterior, es menor que el costo promedio anual del nuevo camión, entonces, conviene seguir utilizando los servicios de la otra compañía transportista.

Finalmente, conviene señalar que si el número de corridas en esta simulación es el adecuado, el costo promedio anual de enviar los excedentes utilizando los servicios de la otra compañía, debe tender a \$51,840. Lo anterior significa, que la decisión de comprar un nuevo camión, o seguir operando con el camión actual, debe ser la misma independientemente del enfoque que se utilice en la solución de este problema.

Ejemplo 5.3. Estimación de π

Usando la figura que se muestra a continuación, ¿describir un procedimiento que estime el valor de π ?



En la figura anterior, es obvio que la probabilidad de que un punto pertenezca al cuarto de círculo es la siguiente:

$$\frac{\text{Area del cuarto de círculo}}{\text{Area del cuadrado}} = \frac{\pi/4}{1} = \frac{\pi}{4}$$

también, es obvio que esta probabilidad puede ser estimada mediante la relación x/n , donde x representa el número de veces que un punto simu-

lado cayó dentro del cuarto de círculo y n el número de corridas. Por consiguiente, en el límite cuando el número de corridas tienda a infinito, las dos relaciones anteriores deben ser idénticas, esto es:

$$\lim_{n \rightarrow \infty} \frac{x}{n} = \frac{\pi}{4} \text{ y } \hat{\pi} = 4 \frac{x}{n}$$

De acuerdo a la identidad anterior, el procedimiento para estimar el valor de π , sería:

1. Generar dos números uniformes R_1 y R_2 .
2. Evaluar $\sqrt{R_1^2 + R_2^2}$. Si este valor es menor que 1, entonces, el número simulado cae dentro del cuarto de círculo, y por consiguiente se incrementa el valor de x . De lo contrario, es necesario regresar al paso 1.
3. Repetir los pasos anteriores hasta que n corridas hayan sido simuladas. En este momento, el valor de π , sería: $4x/n$.

Aplicando el procedimiento anterior, la tabla 5.3 muestra los resultados de una simulación manual para este problema. Como se puede observar en esta tabla, de 25 puntos simulados, 21 cayeron adentro del cuarto de círculo. Por consiguiente, el valor estimado de π , de acuerdo a estas 25 corridas, sería:

$$\hat{\pi} = 4(21/25) = 3.36$$

Es obvio, que este valor estimado de π , difiere demasiado de su valor verdadero. Consecuentemente, si la estimación del valor de π se quiere aproximar más a su valor verdadero, entonces, es necesario aumentar significativamente el número de corridas. Por ejemplo, en este momento es posible preguntar, ¿cuántas corridas es necesario simular para que el estimador de π , difiera de su valor verdadero en una cantidad menor que 0.1, con una seguridad de 95%? Esta pregunta puede ser planteada de acuerdo a la siguiente expresión:

$$\text{Prob. } \{ |\hat{\pi} - \pi| \leq 0.1 \} = 0.95$$

la cual puede ser rearmada y expresada como:

$$\text{Prob. } \{ \pi - 0.1 \leq \hat{\pi} \leq \pi + 0.1 \} = 0.95$$

TABLA 5.3. Resultados de la simulación manual para estimar el valor de π

Número de corrida	Primer* número aleatorio	Segundo número aleatorio	$\sqrt{R_1^2 + R_2^2} \leq 1?$	Valor acumulado de X
1	0.03991	0.38555	sí	1
2	0.17546	0.32643	sí	2
3	0.69572	0.24122	sí	3
4	0.61196	0.30532	sí	4
5	0.03788	0.48228	sí	5
6	0.88618	0.71299	no	5
7	0.27954	0.80863	sí	6
8	0.33564	0.90899	sí	7
9	0.78038	0.55986	sí	8
10	0.87539	0.16818	sí	9
11	0.34677	0.45305	sí	10
12	0.59747	0.16520	sí	11
13	0.68652	0.79375	no	11
14	0.33521	0.59589	sí	12
15	0.20554	0.59404	sí	13
16	0.42614	0.34994	sí	14
17	0.99385	0.66497	no	14
18	0.48509	0.15470	sí	15
19	0.20094	0.73788	sí	16
20	0.60530	0.44372	sí	17
21	0.18611	0.58319	sí	18
22	0.61199	0.18627	sí	19
23	0.00441	0.32624	sí	20
24	0.65961	0.20288	sí	21
25	0.59362	0.99782	no	21

*Estos números aleatorios se obtuvieron del apéndice.

Sin embargo, $\hat{\pi}$ puede ser estimada usando la relación $4X/n$. Por consiguiente, sustituyendo $\hat{\pi} = 4x/n$ en la última expresión, se obtiene:

$$\text{Prob. } \left\{ \frac{\pi - 0.1}{4} \leq \frac{x}{n} \leq \frac{\pi + 0.1}{4} \right\} = 0.95$$

y puesto que x (cantidad de puntos simulados que cayeron adentro del cuarto de círculo) es una variable aleatoria que sigue una distribución binomial (cada punto simulado sigue una distribución Bernoulli), con media $n \pi/4$ y variancia $n \pi/4(1 - \pi/4)$, entonces, x/n la cual representa la fracción de éxitos, tiene la siguiente media y variancia:

$$E\left(\frac{x}{n}\right) = \frac{1}{n} E(x) = \pi/4$$

$$VAR\left(\frac{x}{n}\right) = E\left(\frac{x}{n} - \frac{\pi}{4}\right)^2 = \frac{1}{n^2} E(x - n \pi/4)^2 = \pi/4(1 - \pi/4)/n$$

Por otra parte, es obvio que cuando n es grande, tanto la distribución binomial como la fracción de éxitos (x/n), pueden ser reemplazadas por una distribución normal con los parámetros respectivos. Por ejemplo, se puede decir que cuando n es grande, x/n sigue aproximadamente una distribución normal, con media y variancia respectivamente, de $\pi/4$ y $\pi/4(1 - \pi/4)/n$. Por consiguiente, el número de corridas que es necesario realizar para que el valor estimado de π , difiera de su valor verdadero en menos de 0.1, con una seguridad de 95%, puede ser encontrado al resolver la siguiente ecuación:

$$Z_{\alpha/2} = \frac{\frac{\pi + 0.1}{4} - \frac{\pi}{4}}{\sqrt{\frac{\pi}{4} \left(1 - \frac{\pi}{4}\right) \frac{1}{n}}}$$

y despejando el valor de n , se obtiene:

$$n = \frac{4 \pi (1 - \pi/4) Z_{\alpha/2}^2}{(0.1)^2}$$

y puesto que $Z_{\alpha/2} = 1.96$, entonces el valor de n resulta ser de 1036. En esta última expresión, se puede observar que el valor de n aumenta entre mayor sea el nivel de confianza o la exactitud requerida. En conclusión, se puede decir que para que el valor estimado de π , difiera de su valor verdadero en una cantidad menor que 0.1, es necesario simular 1036 corridas.

Ejemplo 5.4. Proyecto de inversión

La compañía *X* desea incursionar en un nuevo negocio cuya inversión inicial requerida y los flujos de efectivo antes de depreciación e impuestos de los próximos cinco años siguen las siguientes distribuciones triangulares:

	<i>Estimación pesimista</i>	<i>Estimación más probable</i>	<i>Estimación optimista</i>
Activo fijo inicial	– \$100,000	– \$70,000	– \$60,000
Activo circulante inicial	– 40,000	– 30,000	– 25,000
Flujo antes de impuestos	30,000	40,000	45,000

Además, esta compañía estima que las tasas de inflación en los próximos cinco años siguen las siguientes distribuciones triangulares:

<i>Año</i>	<i>Tasas de inflación (%)</i>		
	<i>Estimación pesimista</i>	<i>Estimación más probable</i>	<i>Estimación optimista</i>
1	18	15	12
2	18	15	12
3	22	18	15
4	25	20	18
5	28	22	19

Si la tasa de impuestos es de 50%, la TREMA de 15% y la alta administración ha establecido que si $\text{Prob. } \{ \text{TIR} > \text{TREMA} \}^* \geq 0.90$, entonces los nuevos proyectos se acepten; ¿debería la compañía *X* aceptar este nuevo proyecto de inversión? (Asuna que la vida fiscal del activo fijo es de 5 años, y que el valor de rescate es un 20% del valor simulado para el activo fijo, y un 100% del valor simulado para el activo circulante.)

Este problema sería muy difícil de resolver en forma analítica, puesto que tanto los flujos de efectivo como las tasas de inflación son probabilísticas. Sin embargo, por medio de la simulación es muy sencillo establecer o desarrollar un modelo que incorpore toda la información probabilística de las diferentes variables aleatorias que intervienen en el proyecto de inversión. Específicamente, los pasos necesarios para determinar la distribución de probabilidad de la TIR y en base a ello tomar una decisión, serían:

*TIR = Tasa interna de rendimiento.

TREMA = Tasa de recuperación mínima atractiva.

1. Determinar la TIR mínima que puede resultar de la simulación. El valor de TIR mínimo resulta cuando el activo fijo inicial, el activo circulante inicial, el flujo de efectivo antes de impuestos y las tasas de inflación toman su valor pesimista. Para estos valores, la tabla 5.4 muestra los flujos de efectivo después de impuestos. Para estos flujos de efectivo, la tasa interna de rendimiento que se obtiene es de -3.1% .
2. Determinar la TIR máxima que puede resultar de la simulación. El valor de TIR máximo resulta cuando el activo fijo inicial, el activo circulante inicial, el flujo de efectivo antes de impuestos y las tasas de inflación toman su valor optimista. Para estos valores, la tabla 5.5 muestra los flujos de efectivo después de impuestos. Para estos flujos de efectivo se obtiene una tasa de rendimiento de 19.91% .
3. Dividir el intervalo $(-3.1\%; 19.91\%)$ en 20 subintervalos iguales.
4. Simular el valor del activo fijo inicial y del activo circulante inicial, los cuales representan la inversión total del proyecto.
5. Determinar el flujo de efectivo después de impuestos a pesos constantes, de acuerdo a la siguiente expresión:

$$S_t = \frac{\sum_{j=1}^t x_t \pi (1+i_j)(1-T) + (0.2AF)(T) - AC(i_j) \pi (1+i_j)}{\sum_{j=1}^t \pi (1+i_j)}$$

donde:

- S_t = Flujo de efectivo después de impuestos a pesos constantes.
 x_t = Valor simulado del flujo de efectivo antes de impuestos en el período t .
 i_j = Valor simulado de la tasa de inflación en el período j .
 T = Tasa de impuestos.
 AF = Valor simulado del activo fijo inicial.
 AC = Valor simulado del activo circulante inicial.

6. Determinar el valor de rescate a pesos constantes, utilizando la siguiente expresión:

$$VR = AC + (0.2AF)(1-T)$$

7. Calcular la tasa interna de rendimiento para estos valores simulados, de acuerdo a la siguiente expresión:

TABLA 5.4. Flujos de efectivo después de impuestos considerando que todas las variables aleatorias toman su valor pesimista.

Año	Inversión adicional en activo circulante	Flujos de efectivo antes de impuestos	Depre- ciación	Ingreso gravable	Impues- tos	Flujo de efectivo después de impuestos (pesos corrientes)	Flujo de efectivo después de impuestos (pesos constantes)
0		-\$140,000				-\$140,000	-\$140,000
1	7,200	35,400	20,000	15,400	7,200	20,500	17,373
2	8,496	41,772	20,000	21,772	10,886	22,390	16,080
3	10,365	50,962	20,000	30,962	15,481	25,116	14,785
4	12,956	63,702	20,000	43,702	21,851	28,895	13,608
5	16,584	81,539	20,000	61,539	30,769	34,185	12,577
5		163,078			27,180	135,898	50,000

Tasa interna de rendimiento = - 3.1%

TABLA 5.5. Flujos de efectivo después de impuestos considerando que todas las variables aleatorias toman su valor optimista.

Año	Inversión adicional en activo circulante	Flujo de efectivo antes de impuestos	Depre- ciación	Ingreso gravable	Impues- tos	Flujo de efectivo después de impuestos (pesos corrientes)	Flujo de efectivo después de impuestos (pesos constantes)
0		-\$85,000				-\$85,000	-\$85,000
1	3,000	50,400	12,000	38,400	19,200	28,200	25,179
2	3,360	56,448	12,000	44,448	22,224	30,864	24,605
3	3,864	64,915	12,000	52,915	26,458	34,594	23,981
4	4,560	76,600	12,000	64,600	32,300	39,740	23,346
5	5,425	91,154	12,000	79,154	39,577	46,151	22,783
5		74,949			12,154	62,795	31,000

Tasa interna de rendimiento = 19.91%

Por el contrario, si la respuesta es negativa, entonces:

$$x = c - \sqrt{(c-a)(c-b)(1-R)}$$

3. Repetir los pasos anteriores tantas veces como se desee.

Si se aplica el procedimiento descrito anteriormente, y el método de la transformada inversa para simular distribuciones triangulares, el resultado es el histograma de la TIR que se muestra en la tabla 5.6. A partir de este histograma, se obtiene la distribución acumulada de la TIR, la cual se muestra en la figura 5.1. En esta última figura se puede apreciar que la Prob. {TIR > TREMA} es prácticamente nula. Esto significa, que de acuerdo a los estándares establecidos por la alta administración, el proyecto deberá ser rechazado. Finalmente, conviene señalar que esta decisión es bastante confiable, puesto que en la tabla 5.6 se muestran los resultados de simular 1,000 veces el valor de TIR.

TABLA 5.6. Frecuencia acumulada de la TIR (%).

<i>Límite inferior del intervalo</i>	<i>Límite superior del intervalo</i>	<i>Fracción</i>	<i>Fracción acumulada</i>
- 3.10	- 1.95	0.000	0.000
- 1.95	- 0.80	0.000	0.000
- 0.80	0.35	0.000	0.000
0.35	1.50	0.000	0.000
1.50	2.65	0.003	0.003
2.65	3.80	0.019	0.022
3.80	4.96	0.082	0.104
4.96	6.11	0.141	0.245
6.11	7.26	0.189	0.434
7.26	8.41	0.193	0.627
8.41	9.56	0.184	0.811
9.56	10.71	0.111	0.922
10.71	11.86	0.057	0.979
11.86	13.01	0.015	0.994
13.01	14.16	0.006	1.000
14.16	15.31	0.000	1.000
15.31	16.46	0.000	1.000
16.46	17.61	0.000	1.000
17.61	18.76	0.000	1.000
18.76	19.91	0.000	1.000

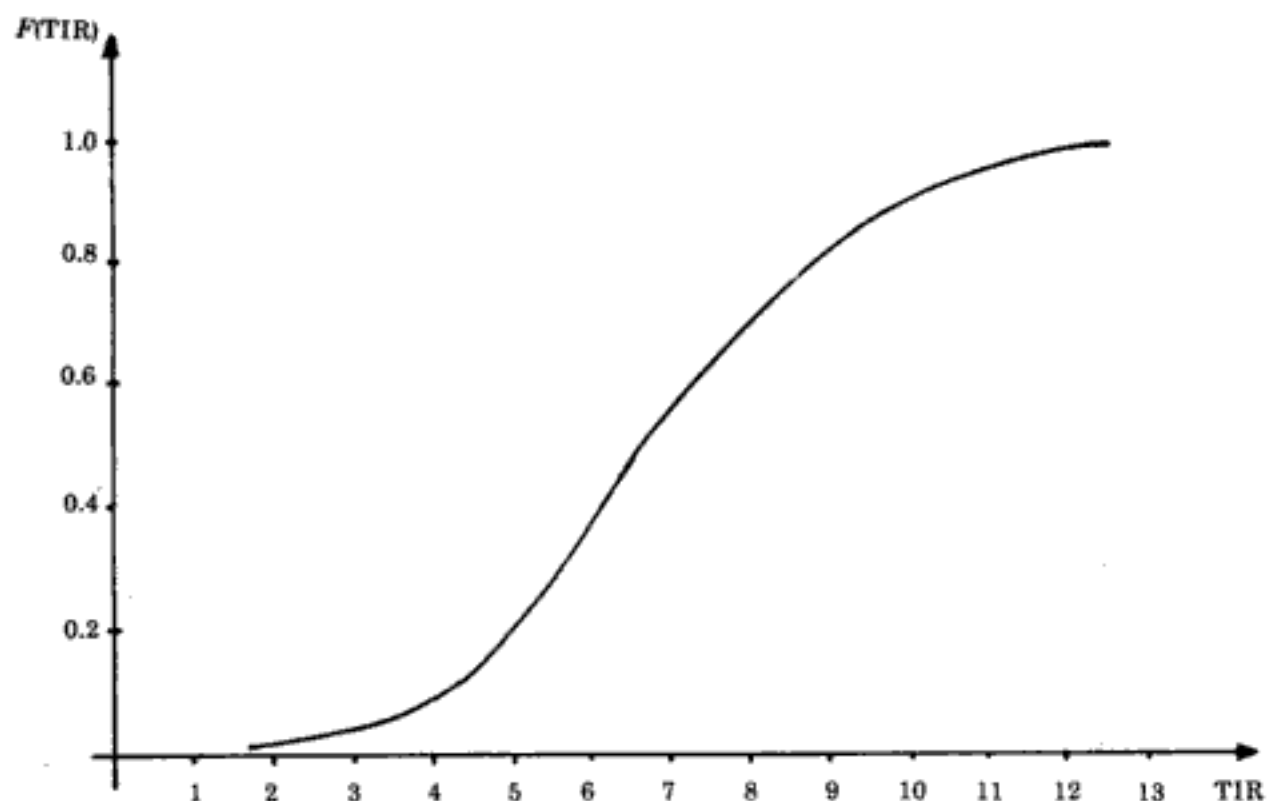


Figura 5-1 Distribución acumulada de la tasa interna de rendimiento.

Ejemplo 5.5. Sistema de inventarios

La demanda mensual de un cierto producto sigue la siguiente distribución de probabilidad empírica:

<i>Cantidad</i>	<i>Probabilidad</i>	<i>Cantidad</i>	<i>Probabilidad</i>	<i>Cantidad</i>	<i>Probabilidad</i>
35	0.010	44	0.029	53	0.065
36	0.015	45	0.035	54	0.060
37	0.020	46	0.045	55	0.050
38	0.020	47	0.060	56	0.040
39	0.022	48	0.065	57	0.030
40	0.023	49	0.070	58	0.016
41	0.025	50	0.080	59	0.015
42	0.027	51	0.075	60	0.005
43	0.028	52	0.070		

El tiempo de entrega está distribuido de acuerdo a la siguiente función de probabilidad:

Meses	1	2	3
Probabilidad	0.30	0.40	0.30

Los factores estacionales para cada uno de los meses del año son como se muestra a continuación:

<i>Mes</i>	<i>Factores estacionales</i>	<i>Mes</i>	<i>Factores estacionales</i>
1	1.20	7	0.80
2	1.00	8	0.90
3	0.90	9	1.00
4	0.80	10	1.20
5	0.80	11	1.30
6	0.70	12	1.40

La información con respecto a los costos relevantes es la siguiente:

Costo de ordenar	= \$100/orden
Costo de inventario	= \$ 20/unidad/año
Costo de faltante	= \$ 50/unidad

Si el inventario inicial se asume en 150 unidades, ¿determine la cantidad óptima a ordenar (q) y el nivel óptimo de reorden (R)?

Los sistemas de inventarios a menudo contienen varios componentes estocásticos que interactúan entre sí. Cuando estos componentes son importantes, su consideración en el modelo matemático lo hace a éste considerablemente complejo. El modelo de inventarios determinístico asume demanda conocida y constante; un tiempo de producción o de entrega conocido y fijo; una razón de producción infinita; no se permite faltante; y los costos de llevar inventario y de ordenar son parámetros que tienen un comportamiento lineal. Cuando la demanda es aleatoria, el modelo de inventario es un poco más sofisticado. Los sistemas de inventario, sin embargo, contienen más componentes estocásticos. Además de las variaciones en la demanda aleatoria, ésta puede tener un comportamiento estacional. El tiempo de entrega entre colocar y recibir una orden puede ser estocástico. Se puede satisfacer con inventarios de períodos actuales, la demanda insatisfecha de períodos previos. Los costos de ordenar, llevar inventario y de faltante pueden ser difíciles de estimar. Estos parámetros pueden también ser no-lineales. Si muchas de estas complicaciones son importantes al sistema de inventarios que se está analizando, el desarrollo de un modelo matemático que represente a este sistema podría resultar significativamente complejo. Sin embargo, un modelo de simulación, procesado con la ayuda de la computadora, podría ser más fácil, más confiable y más efectivo.

El sistema de inventarios que se analiza es *lote constante y tiempo entre pedidos variables*. Las variables de decisión para este modelo son la cantidad a ordenar q y el nivel de reorden R , las cuales minimizan los costos totales del inventario (costo de ordenar, costo de llevar inventario y costo de faltante). Por consiguiente, para evaluar el funcionamiento

del sistema de acuerdo a los valores de las variables de decisión utilizados, costos totales anuales son acumulados. Cada vez que una orden es colocada, el costo de ordenar anual es incrementado en \$100. El nivel de inventario promedio mensual es utilizado para evaluar el costo de llevar inventario*. Al final de cada mes se determina el número de unidades faltantes y el costo que esto representa. La suma de los costos anteriores, proporciona el costo total anual.

Para entender el proceso de simulación de este sistema de inventario, una simulación manual de un año de operación es presentada en la tabla 5.7, y en la figura 5.2. Los valores de las variables de decisión utilizados en esta simulación son: $q = 200$ y $R = 100$. También, para esta simulación se utilizó el método de la transformada inversa para simular las demandas (ver tabla 5.9) y los tiempos de entrega (ver tabla 5.10). El inventario inicial es de 150 unidades. La demanda simulada para el primer mes (considerando el factor de ajuste) fue de 64, lo cual reducirá el inventario al final del mes a 86. El inventario promedio del primer mes es por consiguiente $(150 + 86)/2 = 118$. Al final del primer mes, el nivel de existencias es menor que el nivel de reorden, por lo cual la primera orden es colocada. De acuerdo a la tabla 5.11, el tiempo de entrega de esta primera orden es de 1 mes. Por consiguiente, al principio del tercer mes 200 unidades se agregarán al nivel de existencias. Al final del sexto mes, una segunda orden es colocada, la cual de acuerdo a la tabla 5.11, se entregará a principios del décimo mes. Este tiempo de

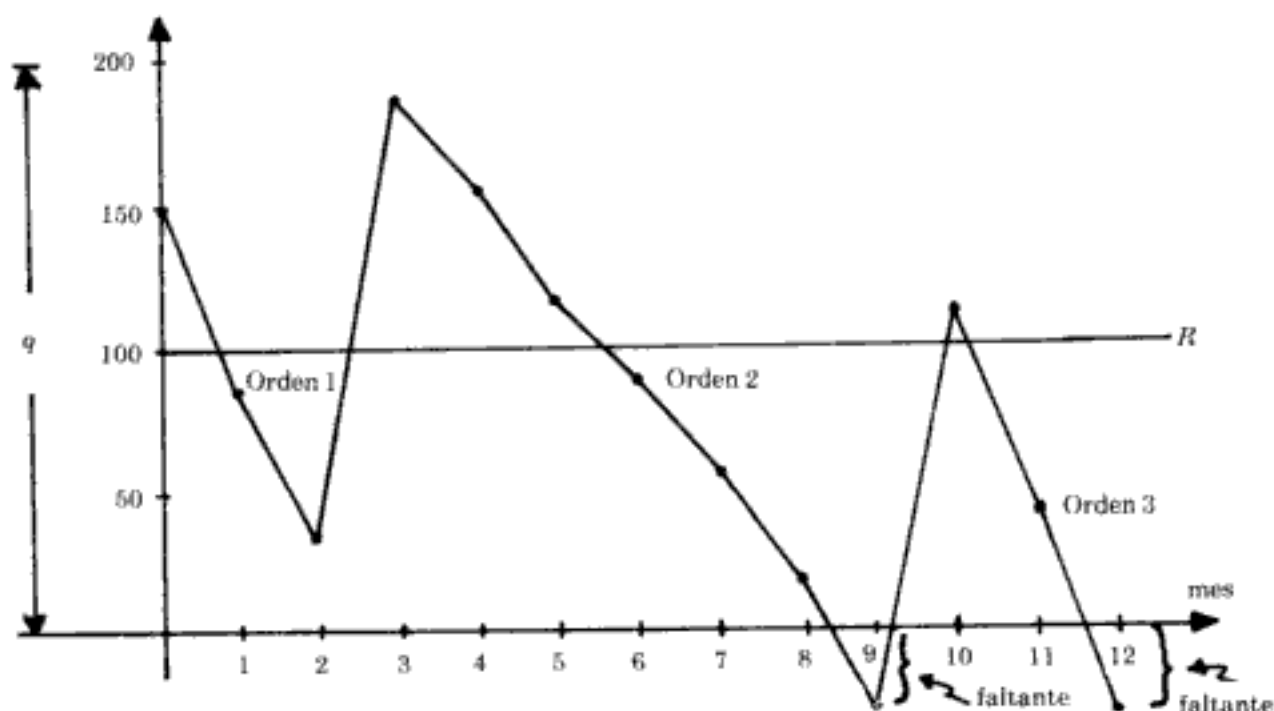


Figura 5.2. Simulación para un año de operación del sistema de inventarios.

*Se hacen ajustes para aquellos periodos que no tienen existencias suficientes para satisfacer la demanda de ese periodo.

Tabla 5.7. Simulación manual del sistema de inventarios ($q = 200$, $R = 100$).

<i>Mes</i>	<i>Inventario inicial</i>	<i>Número aleatorio</i>	<i>Demanda ajustada</i>	<i>Inventario final</i>	<i>Faltante</i>	<i>Orden</i>	<i>Inventario mensual promedio</i>
1	1 5 0	0.74022	64	86		1	118
2	8 6	0.65741	52	34			60
3	2 3 4	0.66083	47	187			211
4	1 8 7	0.08355	31	156			172
5	1 5 6	0.55121	40	116			136
6	1 1 6	0.00911	25	91		2	104
7	9 1	0.14060	34	57			74
8	5 7	0.14845	38	19			38
9	1 9	0.41839	48	0	29		4*
10	1 7 1	0.39685	58	113			142
11	1 1 3	0.74416	69	44		3	79
12	4 0	0.53152	70	0	30		11**

$$*4 = \frac{19}{2} \left(\frac{19}{48} \right)$$

$$**11 = \frac{40}{2} \left(\frac{40}{70} \right)$$

Tabla 5.8. Costos totales anuales del sistema de inventario.

<i>Costo de ordenar</i>	<i>Costo de llevar inventario</i>	<i>Costo de faltante</i>	<i>Costo total</i>
3(100) = \$300	1149(1.67) = \$1,918	59(50) = \$2,950	\$5,168

entrega de tres meses, origina en el noveno mes un faltante de 29 unidades. Consecuentemente, de las 200 unidades que llegaran al principio del décimo mes, 29 serán usadas para satisfacer la demanda que quedó insatisfecha en el período anterior. Finalmente, la tabla 5.8 muestra el costo total del sistema de inventario para un año de operación.

Si se simulan varios años más (30 por ejemplo), se podría obtener el costo total promedio anual asociado a los valores mencionados de decisión ($q = 200$, $R = 100$). Sin embargo, lo importante es aplicar una metodología que progresivamente vaya mejorando los valores de las variables de decisión hasta determinar sus valores óptimos. Para este propósito, el algoritmo de Hooke y Jeeaves* será aplicado. De acuerdo a este algoritmo, primero hay que seleccionar valores iniciales apropiados de las variables de decisión. Después, hay que obtener el costo

*Para entender los detalles de este algoritmo, ver Phillips, Ravindran, Solberg, *Operation Research Principles and Practice*, John Wiley, 1976.

total relacionado a estos valores. En seguida, se usa esta información para mejorar los valores de las variables de decisión. Esta iteración será repetida hasta que los valores óptimos de las variables de decisión hayan sido encontrados. La tabla 5.12 muestra todas las iteraciones que resultan de aplicar este algoritmo. En esta tabla se puede observar que los valores óptimos de las variables de decisión son: $q = 173$ y $R = 163$, lo cual produce un costo total promedio anual de \$3,245.

Tabla 5.9. Método de la transformada inversa para simular valores de la demanda

Si $0.000 \leq R < 0.010$, entonces $x = 35$
Si $0.010 \leq R < 0.025$, entonces $x = 36$
Si $0.025 \leq R < 0.045$, entonces $x = 37$
Si $0.045 \leq R < 0.065$, entonces $x = 38$
Si $0.065 \leq R < 0.087$, entonces $x = 39$
Si $0.087 \leq R < 0.110$, entonces $x = 40$
Si $0.110 \leq R < 0.135$, entonces $x = 41$
Si $0.135 \leq R < 0.162$, entonces $x = 42$
Si $0.162 \leq R < 0.190$, entonces $x = 43$
Si $0.190 \leq R < 0.219$, entonces $x = 44$
Si $0.219 \leq R < 0.254$, entonces $x = 45$
Si $0.254 \leq R < 0.299$, entonces $x = 46$
Si $0.299 \leq R < 0.359$, entonces $x = 47$
Si $0.359 \leq R < 0.424$, entonces $x = 48$
Si $0.424 \leq R < 0.494$, entonces $x = 49$
Si $0.494 \leq R < 0.574$, entonces $x = 50$
Si $0.574 \leq R < 0.649$, entonces $x = 51$
Si $0.649 \leq R < 0.719$, entonces $x = 52$
Si $0.719 \leq R < 0.784$, entonces $x = 53$
Si $0.784 \leq R < 0.844$, entonces $x = 54$
Si $0.844 \leq R < 0.894$, entonces $x = 55$
Si $0.894 \leq R < 0.934$, entonces $x = 56$
Si $0.934 \leq R < 0.964$, entonces $x = 57$
Si $0.964 \leq R < 0.980$, entonces $x = 58$
Si $0.980 \leq R < 0.995$, entonces $x = 59$
Si $0.995 \leq R \leq 1.000$, entonces $x = 60$

Tabla 5.10. Método de la transformada inversa para simular los tiempos de entrega

Si $0.00 \leq R < 0.30$, entonces $x = 1$ mes
Si $0.30 \leq R < 0.70$, entonces $x = 2$ meses
Si $0.70 \leq R \leq 1.00$, entonces $x = 3$ meses

camiones que están esperando al momento de que el almacén abre sus puertas, es la siguiente:

<i>Cantidad de camiones</i>	<i>Probabilidad</i>
0	0.50
1	0.25
2	0.15
3	0.10

Por otra parte, también de información pasada, se sabe que la distribución de probabilidad del tiempo entre llegadas, es la siguiente:

<i>Tiempo entre llegadas (minutos)</i>	<i>Probabilidad</i>
20	0.02
25	0.08
30	0.12
35	0.25
40	0.20
45	0.15
50	0.10
55	0.05
60	0.03

Finalmente, por medio de experimentación se han obtenido las distribuciones de probabilidad del tiempo de servicio para diferentes tamaños de equipo. Tales distribuciones se muestran a continuación:

<i>Tiempo de servicio (minutos) de tres personas</i>	<i>Probabilidad</i>	<i>Tiempo de servicio (minutos) de cuatro personas</i>	<i>Probabilidad</i>
20	0.05	15	0.05
25	0.10	20	0.15
30	0.20	25	0.20
35	0.25	30	0.20
40	0.12	35	0.15
45	0.10	40	0.12
50	0.08	45	0.08
55	0.06	50	0.04
60	0.04	55	0.01

<i>Tiempo de servicio (minutos) de cinco personas</i>	<i>Probabilidad</i>	<i>Tiempo de servicio (minutos) de seis personas</i>	<i>Probabilidad</i>
10	0.10	5	0.12
15	0.18	10	0.15
20	0.22	15	0.26
25	0.18	20	0.15
30	0.10	25	0.12
35	0.08	30	0.08
40	0.06	35	0.06
45	0.05	40	0.04
50	0.03	45	0.02

Si todo el equipo se considera como un servidor, ¿cuál es el tamaño óptimo del equipo?

El análisis de sistemas de colas frecuentemente ilustra la dificultad de construir un modelo matemático que contenga todos los elementos presentes en el sistema. Generalmente, los modelos de colas asumen que el tiempo entre llegadas y el tiempo de servicio siguen una distribución exponencial. También, estos modelos asumen que la cantidad de clientes que llegan al sistema por unidad de tiempo y la cantidad de clientes que el sistema sirve por unidad de tiempo siguen una distribución poisson. Además de las anteriores suposiciones, estos modelos generalmente asumen que: 1) La fuente que alimenta al sistema es de tamaño infinito; 2) El sistema ha alcanzado el estado estable; 3) La disciplina de servicio es primeras entradas—primeras salidas; 4) Los clientes que llegan al sistema se unen a la cola; 5) Los clientes que están en la cola, permanecerán en ella hasta que sean servidos; 6) Los clientes llegan en forma individual y 7) Las salidas del sistema son en forma individual. Estas suposiciones generalmente originan que las decisiones que se toman con base en estos modelos no sean confiables.

En muchas situaciones, sin embargo, el sistema de colas puede tener características especiales importantes. Puede tener etapas transientes y de estado estable, puede ser dependiente, o formar parte de un grupo de facilidades independientes. Puede tener prioridades de servicio, o algunos clientes que llegan al sistema cuando la cola es excesivamente grande pueden rehusar entrar al sistema. También, puede ser que la fuente que alimenta al sistema sea finita. La exclusión de tales características especiales puede distorsionar la naturaleza verdadera del sistema, por lo cual deberán ser consideradas en la modelación del mismo.

Si hay varias características especiales importantes, una modelación matemática completa resultaría difícil o casi imposible. En tales situaciones, el análisis completo del sistema puede ser logrado a través del uso de la técnica de simulación.

adelantado a las 11.40. Como el primer camión se termina de descargar a las 11.50, entonces, hasta este momento se iniciará la descarga del segundo camión. Esto significa que el tiempo de ocio del segundo camión es de 10 minutos. El tiempo de servicio para el segundo camión es de 25 minutos (número aleatorio 0.14387, que de acuerdo a la tabla 5.15, corresponde a un tiempo de servicio de 25 minutos). El proceso continúa para el resto del turno en una forma similar.

Tabla 5.13. Método de la transformada inversa para simular los camiones que esperan al empezar a operar el almacén.

Si $0.00 \leq R < 0.50$, entonces $x = 0$
Si $0.50 \leq R < 0.75$, entonces $x = 1$
Si $0.75 \leq R < 0.90$, entonces $x = 2$
Si $0.90 \leq R < 1.00$, entonces $x = 3$

Tabla 5.14. Método de la transformada inversa para simular el tiempo entre llegadas.

Si $0.00 \leq R < 0.02$, entonces $x = 20$
Si $0.02 \leq R < 0.10$, entonces $x = 25$
Si $0.10 \leq R < 0.22$, entonces $x = 30$
Si $0.22 \leq R < 0.47$, entonces $x = 35$
Si $0.47 \leq R < 0.67$, entonces $x = 40$
Si $0.67 \leq R < 0.82$, entonces $x = 45$
Si $0.82 \leq R < 0.92$, entonces $x = 50$
Si $0.92 \leq R < 0.97$, entonces $x = 55$
Si $0.97 \leq R \leq 1.00$, entonces $x = 60$

Tabla 5.15. Método de la transformada inversa para simular el tiempo de servicio (equipó de 3 personas).

Si $0.00 \leq R < 0.05$, entonces $x = 20$
Si $0.05 \leq R < 0.15$, entonces $x = 25$
Si $0.15 \leq R < 0.35$, entonces $x = 30$
Si $0.35 \leq R < 0.60$, entonces $x = 35$
Si $0.60 \leq R < 0.72$, entonces $x = 40$
Si $0.72 \leq R < 0.82$, entonces $x = 45$
Si $0.82 \leq R < 0.90$, entonces $x = 50$
Si $0.90 \leq R < 0.96$, entonces $x = 55$
Si $0.96 \leq R \leq 1.00$, entonces $x = 60$

A través de las operaciones del turno nocturno, el tiempo entre llegadas y el tiempo de servicio son obtenidos utilizando números aleatorios del apéndice. Después de cada evento, el *reloj de la simulación* es actualizado. Tan pronto como algún servicio sea terminado después de las 3 A.M., el personal dispondrá de media hora para tomar sus alimentos. Para este caso resultó de 3.15 a 3.45 A.M. Para descargar el último camión, el cual llegó a las 6.55 A.M., el personal tuvo que trabajar 10 minutos de tiempo extra. Por consiguiente, los costos totales que se incurren en este primer turno simulado son:

$$\begin{aligned}\text{Salarios} &= \text{Tiempo normal} + \text{tiempo extra.} \\ &= 3(8)(25) + 37.50(3)(1/6). \\ &= \$618.75.\end{aligned}$$

$$\begin{aligned}\text{Espera} &= \text{Operación del almacén} + \text{espera del camión.} \\ &= 500(8.67) + 100(6.42). \\ &= \$4,977\end{aligned}$$

$$\text{Costo total} = 618.75 + 4,977 = \$5,595.75$$

Tabla 5.16. Método de la transformada inversa para simular el tiempo de servicio (equipo de 4 personas).

Si $0.00 \leq R < 0.05$, entonces $x = 15$
Si $0.05 \leq R < 0.20$, entonces $x = 20$
Si $0.20 \leq R < 0.40$, entonces $x = 25$
Si $0.40 \leq R < 0.60$, entonces $x = 30$
Si $0.60 \leq R < 0.75$, entonces $x = 35$
Si $0.75 \leq R < 0.87$, entonces $x = 40$
Si $0.87 \leq R < 0.95$, entonces $x = 45$
Si $0.95 \leq R < 0.99$, entonces $x = 50$
Si $0.99 \leq R \leq 1.00$, entonces $x = 55$

Tabla 5.17. Método de la transformada inversa para simular el tiempo de servicio (equipo de 5 personas).

Si $0.00 \leq R < 0.10$, entonces $x = 10$
Si $0.10 \leq R < 0.28$, entonces $x = 15$
Si $0.28 \leq R < 0.50$, entonces $x = 20$
Si $0.50 \leq R < 0.68$, entonces $x = 25$
Si $0.68 \leq R < 0.78$, entonces $x = 30$
Si $0.78 \leq R < 0.86$, entonces $x = 35$
Si $0.86 \leq R < 0.92$, entonces $x = 40$
Si $0.92 \leq R < 0.97$, entonces $x = 45$
Si $0.97 \leq R \leq 1.00$, entonces $x = 50$

Tabla 5.18. Método de la transformada inversa para simular el tiempo de servicio (equipo de 6 personas).

Si $0.00 \leq R < 0.12$, entonces $x = 5$
Si $0.12 \leq R < 0.27$, entonces $x = 10$
Si $0.27 \leq R < 0.53$, entonces $x = 15$
Si $0.53 \leq R < 0.68$, entonces $x = 20$
Si $0.68 \leq R < 0.80$, entonces $x = 25$
Si $0.80 \leq R < 0.88$, entonces $x = 30$
Si $0.88 \leq R < 0.94$, entonces $x = 35$
Si $0.94 \leq R < 0.98$, entonces $x = 40$
Si $0.98 \leq R \leq 1.00$, entonces $x = 45$

Aunque esta es la simulación de un turno, los resultados son significativos. Los camiones esperan demasiado tiempo. Los costos de espera son mucho mayores que los salarios percibidos por el personal. A partir de estos resultados, podría pensarse que incrementando el número de personas encargadas de descargar los camiones, el costo total promedio por turno se reduciría. Veamos qué pasa si el tamaño del equipo se incrementa a 4 personas. Para este propósito, la tabla 5.20 muestra los resultados obtenidos en una simulación manual de un turno de operación. Para este nuevo tamaño de equipo, los costos totales que se incurren en el primer turno simulado son:

$$\begin{aligned}\text{Salarios} &= 4(8)(25) + 37.50(4)(1/4) \\ &= \$837.50.\end{aligned}$$

$$\begin{aligned}\text{Espera} &= 500(8.75) + 100(0.83). \\ &= \$4,458.\end{aligned}$$

$$\text{Costo total} = 837.50 + 4,458 = \$5,295.50$$

De los resultados anteriores, podría pensarse que 4 trabajadores es más económico que tres trabajadores. Sin embargo, considerar adecuadamente las variaciones en los resultados obtenidos con diferentes tamaños de equipo, requiere realizar un gran número de simulaciones.

Si se varía el tamaño del equipo y se simulan varios turnos, entonces, es posible determinar el tamaño del equipo para el cual los costos totales son mínimos. La tabla 21 muestra los resultados que se obtienen al simular la operación del almacén durante 60 turnos. Para estas simulaciones se consideraron tamaños de equipo de 3, 4, 5 y 6 personas. Como se puede apreciar en esta tabla, el tamaño óptimo del equipo es de 6 personas. Esto significa, que la forma más económica de operar el almacén es con un equipo de 6 personas.

Tabla 5.19. Simulación de las operaciones de descarga durante un turno (equipo de 3 personas).

Número aleatorio	Tiempo entre llegadas	Tiempo de llegada	Iniciación del servicio	Número aleatorio	Tiempo de servicio	Terminación del servicio	Ocio del personal	Tiempo de espera del camión	Longitud de la cola
0.48355	40	11:00	11:00	0.83761	50	11:50	-	-	1
0.98977	60	11:40	11:50	0.14387	25	12:15	-	10	1
0.06533	25	12:40	12:40	0.51321	35	1:15	25	-	0
0.45128	35	1:05	1:15	0.72472	45	2:00	-	10	1
0.15486	30	1:40	2:00	0.05466	25	2:25	-	20	1
0.19241	30	2:10	2:25	0.84609	50	3:15	-	15	1
0.15997	30	2:40	3:45	0.29735	30	4:15	-	65	1
0.67940	45	3:10	4:15	0.59076	35	4:50	-	65	2
0.90872	50	3:55	4:50	0.76355	45	5:35	-	55	2
0.58997	40	4:45	5:35	0.29549	30	6:05	-	50	2
0.68691	45	5:25	6:05	0.61958	40	6:45	-	40	2
0.73488	45	6:10	6:45	0.17267	30	7:15	-	35	1
		6:55	7:15	0.10061	25	7:40	-	20	1

Nota: En esta simulación solamente se simuló la llegada de 12 camiones, puesto que el camión número 13, de acuerdo al número aleatorio generado, llega después de las 7:00 A.M.

Tabla 5.20. Simulación de las operaciones de descarga durante un turno (equipo de 4 personas)

Número aleatorio	Tiempo entre llegadas	Tiempo de llegada	Inicio del servicio	Número aleatorio	Tiempo de servicio	Terminación del servicio	Ocio del personal	Tiempo de espera del camión	Longitud de la cola
0.68971	45	11:00	11:00	0.11403	20	11:20	-	-	1
0.18477	30	11:45	11:45	0.65622	35	12:20	25	-	0
0.14707	30	12:15	12:20	0.93997	45	1:05	-	5	1
0.83745	50	12:45	1:05	0.22567	25	1:30	-	20	1
0.16930	30	1:35	1:35	0.33361	25	2:00	5	-	0
0.20368	30	2:05	2:05	0.07126	20	2:25	5	-	0
0.41196	35	2:35	2:35	0.37480	25	3:00	10	-	0
0.66919	40	3:10	3:30	0.31678	25	3:55	-	20	1
0.35352	35	3:50	3:55	0.54131	30	4:25	-	5	1
0.79982	45	4:25	4:25	0.68416	35	5:00	-	-	0
0.46850	35	5:10	5:10	0.52326	30	5:40	10	-	0
0.69248	45	5:45	5:45	0.93460	45	6:30	5	-	0
0.04013	25	6:35	6:35	0.31792	25	7:00	5	-	0
		7:00	7:00	0.87315	45	7:45	-	-	0

Nota: También esta simulación al igual que en la tabla anterior, el último camión simulado llega después de las 7:30 A.M.

Tabla 21. Simulación por computadora del sistema de colas durante 60 turnos y considerando diferentes tamaños de equipo.

<i>Tamaño del equipo</i>	<i>Salario normal</i>	<i>Salario extra</i>	<i>Ocio del camión</i>	<i>Operación del almacén</i>	<i>Costos totales</i>
3	\$ 600	\$421	\$2,261	\$6,121	\$9,403
4	800	335	1,202	5,367	7,704
5	1,000	165	455	4,689	6,309
6	1,200	47	125	4,354	5,720

PROBLEMAS

- 5.1. El famoso juego 7-11, requiere que el jugador lance dos dados una o más veces hasta tomar la decisión de que se gana o se pierde el juego. El juego se gana si en el primer lanzamiento los dados suman 7 u 11, ó aparece un 4, 5, 6, 8, 9 ó 10 en el primer lanzamiento y la misma suma reaparece antes de que aparezca un 7. Por otra parte, el juego se pierde si en el primer lanzamiento los dados suman 2, 3 ó 12, ó aparece un 4, 5, 6, 8, 9 ó 10 en el primer lanzamiento y luego sale un 7 antes de que se repita el primer lanzamiento. Si el valor de la apuesta es de \$1, y la ganancia cada vez que se gana un juego es de \$1, ¿cuál sería la probabilidad de quiebra si la cantidad inicial disponible es de \$20? (Asuma que el juego también se termina cuando se acumulan \$50.)
- 5.2. En el famoso juego de la ruleta, existen muchas opciones para apostar. Una de ellas consiste en apostarle al color rojo o al color negro. En el tablero de la ruleta existen 10 números rojos, 10 números negros y 2 números verdes (cero y doble cero). Si un jugador apuesta a un color y el color aparece, él o ella gana la cantidad apostada. Si otro color aparece, el jugador pierde la cantidad apostada. Si el color verde aparece, la rueda de la ruleta se vuelve a girar hasta que el color rojo o negro aparezca. Si este color es el color que se apostó, el jugador no gana ni pierde. De otra forma, se pierde la cantidad apostada.

Dos jugadores usan diferentes estrategias. Un jugador simplemente apuesta \$1 al color rojo cada vez. El otro jugador empieza apostando un \$1 al color rojo. Si él gana, él apuesta otro \$1. Sin embargo, si él pierde, él apuesta \$2 la próxima vez. Si él pierde otra vez, él apuesta \$4. Este jugador puede continuar doblando la

apuesta hasta un límite de \$500.00. Si él pierde esta apuesta de \$500, él empieza apostando nuevamente \$1.

Si cada jugador inicia el juego con \$200, ¿cuál cree usted que es la mejor estrategia?

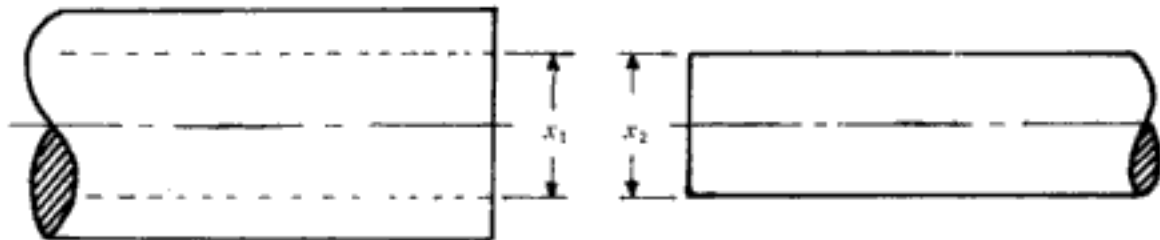
- 5.3. Una compañía de renta de autos, está tratando de determinar el número óptimo de autos a comprar. El costo promedio anual de un auto es de \$75,000. Además, esta compañía ha recopilado las siguientes probabilidades de operación:

Número de autos rentados por día	0	1	2	3	4
Probabilidad	0.10	0.10	0.25	0.30	0.25

Número de días rentados por auto	1	2	3	4
Probabilidad	0.40	0.35	0.15	0.10

Si la renta diaria por auto es de \$350, el costo de no tener un auto disponible cuando se está solicitando es de \$200, y el costo de tener un carro ocioso durante un día es de \$50, ¿cuál es la cantidad de autos que debería comprar la compañía? (Asuma que un auto que se renta por un día está disponible al día siguiente. También, asuma 365 días de operación al año.)

- 5.4. Una flecha será ensamblada en un cojinete como se muestra a continuación:



Si x_1 sigue una distribución normal con media 1.5 y variancia 0.0016, y x_2 sigue esta misma distribución con media 1.48 y variancia 0.0009, determine:

- La probabilidad de que haya interferencia.
 - El número de veces que es necesario simular el experimento, si se quiere que la probabilidad de interferencia estimada difiera de su valor verdadero en menos de 0.01, con un nivel de seguridad del 95%.
- 5.5. La demanda diaria de un cierto artículo está regida por una distribución binomial con parámetros $n = 6$ y $\theta = 1/2$. El tiempo de entrega en días es una variable aleatoria poisson con $\lambda = 3$. El costo de mantener una unidad en inventario es de \$1 por día, el costo del faltante es de \$10 por unidad, y el costo de

ordenar es de \$50 por orden. Se desea comparar dos políticas para llevar el inventario: 1) Ordenar cada 8 días hasta tener 30 artículos en inventario y 2) Ordenar hasta 30 artículos cuando el nivel del inventario sea menor o igual a 10. Si se asume que las unidades faltantes en un ciclo son surtidas por la nueva orden que arriba en el próximo ciclo, ¿cuál de las dos políticas descritas es más económica?

- 5.6. Una compañía tiene un problema de mantenimiento con cierto equipo, que contiene 4 componentes electrónicos idénticos que son la causa del mismo, el cual consiste en que los componentes fallan frecuentemente, forzando a que el equipo se desconecte mientras se hace la reposición. Lo que se ha venido haciendo es reemplazar los componentes solamente cuando se descomponen. Sin embargo, existe una nueva proposición de hacer el reemplazo de los cuatro componentes cuando falle cualesquiera de ellos, con objeto de reducir la frecuencia de desconexión del equipo.

El tiempo de vida de un componente está normalmente distribuido con media de 600 horas y desviación estándar de 100 horas. También se sabe que es necesario desconectar el equipo 1 hora si se reemplaza un componente y 2 horas si se reemplazan los 4. Un componente nuevo cuesta \$200 y se incurre en un costo de \$100 por hora cada vez que se desconecta el equipo.

Determine cuál de las dos políticas anteriores es más económica (Simule la operación del equipo durante 20,000 horas).

- 5.7. Un vendedor de revistas compra mensualmente una revista el día primero de cada mes. El costo de cada ejemplar es de \$1.50. La demanda de esta revista en los primeros 10 días del mes sigue la siguiente distribución de probabilidad:

Demanda	5	6	7	8	9	10	11
Probabilidad	0.05	0.05	0.10	0.15	0.25	0.25	0.15

Al final del décimo día, el vendedor puede regresar cualquier cantidad al proveedor, quien se las pagará a \$0.90 el ejemplar, o comprar más a \$1.20 el ejemplar. La demanda en los siguientes 20 días está dada por la siguiente distribución de probabilidad:

Demanda	4	5	6	7	8
Probabilidad	0.15	0.20	0.30	0.20	0.15

Al final del mes, el vendedor puede regresar al proveedor las revistas que le sobren, las cuales se le pagarán a \$0.60 el ejemplar. Finalmente, se asume que después de un mes ya no existe demanda por parte del público, puesto que para ese entonces ya habrá aparecido

el nuevo número de la revista. Si el precio al público es de \$2 por ejemplar, determine la política óptima de compra.

- 5.8. Una compañía desea entrar en un nuevo negocio cuya inversión inicial requerida y los ingresos netos anuales después de impuestos están distribuidos como sigue:

Inversión inicial $\sim N(\mu = 100,000; \sigma = 5,000)$
 Flujo neto del período $t \sim N(\mu = 30,000; \sigma = 3,000)$

Si la administración ha establecido que un proyecto de inversión será emprendido si $\text{Prob. } \{TIR > TREMA\} \geq 0.90$, y la TREMA es de 30%, ¿debería la compañía X aceptar este nuevo proyecto de inversión? (Asuma un horizonte de planeación de 5 años y un valor de rescate al término de este tiempo de cero).

- 5.9. Una compañía está interesada en analizar un negocio cuya inversión inicial sigue la siguiente distribución triangular:

Estimación pesimista	Estimación más probable	Estimación optimista
-130,000	-100,000	-80,000

Esta inversión tiene una vida fiscal de 5 años, y un valor de rescate al término de este período distribuido triangularmente:

Estimación pesimista	Estimación más probable	Estimación optimista
16,000	20,000	26,000

La tasa de inflación en los próximos cinco años, está regida por la siguiente distribución triangular:

Estimación pesimista	Estimación más probable	Estimación optimista
25%	20%	15%

Finalmente, asuma que los ingresos netos de los próximos cinco años siguen la siguiente distribución uniforme:

Año	1	2	3	4	5
Flujos	20,000	30,000	40,000	50,000	60,000
Probabilidad	1/5	1/5	1/5	1/5	1/5

Si la tasa de impuestos es de 50%, la TREMA de 20%, y la alta administración acepta un nuevo proyecto si $\text{Prob. } \{VPN > 0\} \geq 0.90$, ¿debería la compañía emprender este nuevo proyecto de inversión?

- 5.10. La demanda diaria y el tiempo de entrega de un cierto producto, siguen las siguientes distribuciones de probabilidad:

<i>Demanda diaria</i>	<i>Probabilidad</i>	<i>Tiempo de entrega (días)</i>	<i>Probabilidad</i>
0	0.04	1	0.25
1	0.06	2	0.50
2	0.10	3	0.20
3	0.20	4	0.05
4	0.30		
5	0.18		
6	0.08		
7	0.03		
8	0.01		

La información con respecto a los costos relevantes es la siguiente:

Costo de ordenar = \$50/orden
 Costo de inventario = \$26/unidad/año
 Costo de faltante = \$25/unidad

Si el inventario inicial es de 15 unidades, ¿determine la cantidad óptima a ordenar (q) y el nivel óptimo de reorden (R)? (Asuma que se trabajan 260 días en el año.)

- 5.11. La demanda diaria y el tiempo de entrega de un cierto producto, siguen las siguientes distribuciones de probabilidad:

<i>Demanda diaria</i>	<i>Probabilidad</i>	<i>Tiempo de entrega (días)</i>	<i>Probabilidad</i>
25	0.02	1	0.20
26	0.04	2	0.30
27	0.06	3	0.25
28	0.12	4	0.25
29	0.20		
30	0.24		
31	0.15		
32	0.10		
33	0.05		
34	0.02		

Si el producto no está disponible cuando es requerido, el cliente puede esperar la llegada de un nuevo lote por un tiempo limitado, es decir, si el cliente decide esperar 2 días y la mercancía no llega en ese tiempo, entonces, la demanda de este cliente se considera perdida. La distribución de probabilidad del tiempo que un cliente está dispuesto a esperar para que se le surta su pedido, es la siguiente:

Tiempo de espera (días) Probabilidad

0	0.40
1	0.20
2	0.15
3	0.15
4	0.10

La información con respecto a los costos relevantes es la siguiente:

Costo de ordenar = \$ 100/orden

Costo de inventario = \$ 52/unidad/año

Costo de faltante suponiendo que el cliente espera = \$20/unidad

Costo de faltante suponiendo que el cliente no espera = \$50/unidad

Si el inventario inicial es de 100 unidades, ¿determine la cantidad óptima a ordenar (q) y el nivel óptimo de reorden (R)? (Asuma que se trabajan 260 días en el año).

- 5.12. Se tiene un sistema de colas formado por dos estaciones en serie. Los clientes atendidos en la primera estación pasan en seguida a formar cola en la segunda. En la primera estación de servicio, la razón de llegadas sigue una distribución poisson con media de 20 clientes por hora, y el tiempo de servicio sigue una distribución exponencial con media de 2 minutos por persona. En la segunda estación, el tiempo de servicio está uniformemente distribuido entre 1 y 2 minutos. Para esta información, ¿cuál es el tiempo promedio en el sistema?, ¿cuál de las dos colas que se forman es mayor?
- 5.13. Un banco emplea 3 cajeros para servir a sus clientes. Los clientes arriban de acuerdo a un proceso poisson a una razón media de 40 por hora. Si un cliente encuentra todos los cajeros ocupados, entonces se incorpora a la cola que alimenta a todos los cajeros. El tiempo que dura la transacción entre un cajero y un cliente sigue una distribución uniforme entre 0 y 1 minuto. Para esta información, ¿cuál es el tiempo promedio en el sistema?, ¿cuál es la cantidad promedio de clientes en el sistema?

- 5.14. Una tienda pequeña tiene un lote de estacionamiento con 6 lugares disponibles. Los clientes llegan en forma aleatoria de acuerdo a un proceso poisson a una razón media de 10 clientes por hora, y se van inmediatamente si no existen lugares disponibles en el estacionamiento. El tiempo que un auto permanece en el estacionamiento sigue una distribución uniforme entre 10 y 30 minutos.
- ¿Qué porcentaje de los clientes es perdido por no tener más lugares disponibles?
 - ¿Cuál es la probabilidad de encontrar un lugar disponible en el estacionamiento?
 - ¿Cuál es el porcentaje promedio de espacios disponibles?
- 5.15. Debido a un aumento en las ventas, cierta compañía manufacturera necesita más espacio en su fábrica. La solución que se ha propuesto es la construcción de un nuevo depósito para almacenar los productos terminados. Este depósito estaría localizado a 30 kilómetros del lugar donde está ubicada la planta. Además, de acuerdo a este nuevo plan, se requiere que al final del día se envíe al nuevo depósito, la producción terminada.
- Por otra parte, se sabe de información pasada, que la producción diaria de esta compañía, sigue la siguiente distribución de probabilidad:

<i>Producción diaria en toneladas</i>	<i>Probabilidad</i>
50 - 55	0.10
55 - 60	0.15
60 - 65	0.30
65 - 70	0.35
75 - 80	0.08
80 - 85	0.02

También, se sabe que el tipo de camiones que se deben utilizar para trasladar esta producción, tienen una capacidad media de carga de 5 toneladas. La cantidad de viajes que se pueden realizar cada día (jornada de 8 horas), depende del tiempo de carga y de descarga, como también del tiempo que se requiere para recorrer los treinta kilómetros entre la planta y el depósito. Consecuentemente, la cantidad de producto terminado que un camión puede trasladar de la planta al depósito, es una variable aleatoria cuya distribución de probabilidad es la siguiente:

<i>Toneladas diarias trasladadas por camión</i>	<i>Probabilidad</i>
4.0 - 4.5	0.30
4.5 - 5.0	0.40
5.0 - 5.5	0.20
5.5 - 6.0	0.10

Si la cantidad diaria producida es mayor que la cantidad que puede trasladar la flotilla de camiones, el excedente es enviado a través de otra compañía transportista a un costo de \$100 por tonelada. Además, el costo promedio anual de un nuevo camión es de \$100,000. Si se trabajan 250 días en el año, ¿cuál es el número óptimo de camiones que esta compañía debe de adquirir?

- 5.16. Una cierta compañía posee un gran número de máquinas en uso. El tiempo que dura en operación cada una de estas máquinas, sigue la siguiente distribución de probabilidad:

<i>Tiempo entre descomposturas (horas)</i>	<i>Probabilidad</i>
6 - 8	0.10
8 - 10	0.15
10 - 12	0.24
12 - 14	0.26
16 - 18	0.18
18 - 20	0.07

El tiempo que un operador se tarda en reparar una máquina, sigue la siguiente distribución de probabilidad:

<i>Tiempo de reparación (horas)</i>	<i>Probabilidad</i>
2 - 4	0.15
4 - 6	0.25
6 - 8	0.30
8 - 10	0.20
10 - 12	0.10

Si el costo de tener una máquina ociosa durante una hora es de \$500, y el salario por hora para este tipo de operarios es de \$50, ¿cuántas máquinas se deben asignar a cada mecánico para que las atienda?

6

ANALISIS DE LOS RESULTADOS DE LA SIMULACION

Los resultados de una simulación evalúan la eficiencia u operación de un sistema. Sin embargo, la confiabilidad de estos resultados dependen del número de observaciones que se tienen, es decir, la confiabilidad de los resultados obtenidos en un estudio de simulación, depende de la longitud de la corrida y del número de corridas. Por consiguiente, es necesario determinar intervalos de confianza para las variables del sistema que se están analizando.

La determinación de intervalos de confianza depende de si los resultados de la simulación son independientes o están correlacionados. Por ejemplo, en la simulación de un sistema de colas, el tiempo de espera de un cliente en el sistema, depende del número de clientes que este último cliente encontró al llegar al sistema (correlación entre t_i y t_{i-1}). También, la obtención de intervalos de confianza depende de si el sistema se analiza en estado transiente o en estado estable. Por consiguiente, el objetivo de este capítulo en analizar la determinación de intervalos de confianza para las dos situaciones mencionadas, es decir, para: 1) Sistemas cuyos resultados son independientes uno del otro y 2) Sistemas transientes con alta correlación en sus resultados. Para el primer caso se utilizan los métodos tradicionales de intervalos de confianza. Para el segundo caso se utiliza simulación regenerativa.

6.1. METODOS DE ESTIMACION

Cuando los sistemas que se analizan se encuentran en estado estable, y además los valores que toman las variables que interactúan en

el sistema son independientes, entonces es posible utilizar los procedimientos tradicionales de intervalos de confianza. En la determinación de estos intervalos de confianza se utiliza el teorema del límite central. Este teorema establece que la suma de n variables aleatorias independientes, obtenidas de un universo con media μ y variancia σ^2 , es aproximadamente una distribución normal con media $n\mu$ y variancia $n\sigma^2$. También, es obvio que a partir de este teorema la distribución de:

$$\bar{x} = \frac{1}{n} X_1 + \frac{1}{n} X_2 + \dots + \frac{1}{n} X_n \quad (6.1)$$

es una distribución normal con media μ y variancia σ^2/n y por consiguiente:

$$\frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \quad (6.2)$$

sigue una distribución normal estándar con media 0 y variancia 1. De acuerdo a esta última expresión, y suponiendo un nivel de significado α , el intervalo de confianza del parámetro μ sería:

$$\text{Prob. } \left\{ -Z_{\alpha/2} \leq \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \leq Z_{\alpha/2} \right\} = 1 - \alpha$$

$$\text{Prob. } \left\{ \bar{x} - \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \leq \mu \leq \bar{x} + \frac{\sigma}{\sqrt{n}} Z_{\alpha/2} \right\} = 1 - \alpha \quad (6.3)$$

La expresión (6.3) establece que el ancho del intervalo de confianza es:

$$\frac{2\sigma Z_{\alpha/2}}{\sqrt{n}} \quad (6.4)$$

lo cual significa que el tamaño del intervalo sería más pequeño entre mayor sea el número de observaciones que se tienen de la variable aleatoria X . Por ejemplo, para reducir el intervalo de confianza a la mitad, es necesario de acuerdo a la expresión (6.4), aumentar cuatro veces el número de observaciones.

En las expresiones anteriores se asume que la variancia es conocida. Sin embargo, en muchos problemas reales la variancia generalmente

se desconoce. Para este tipo de situaciones, es posible utilizar el siguiente estimador de la variancia:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{x})^2 \quad (6.5)$$

Ahora, si esta identidad se sustituye en la expresión (6.2), lo que resulta es:

$$\frac{\bar{x} - \mu}{S/\sqrt{n}} \quad (6.6)$$

lo cual sigue una distribución t con $n-1$ grados de libertad. De acuerdo a esta nueva distribución y suponiendo un nivel de significado α , el intervalo de confianza del parámetro μ sería:

$$\text{Prob.} \left\{ \bar{x} - \frac{S}{\sqrt{n}} t_{n-1, \alpha/2} \leq \mu \leq \bar{x} + \frac{S}{\sqrt{n}} t_{n-1, \alpha/2} \right\} = 1 - \alpha \quad (6.7)$$

el cual define un ancho del intervalo de confianza de:

$$\frac{2 S t_{n-1, \alpha/2}}{\sqrt{n}} \quad (6.8)$$

Ejemplo 6.1.

Suponga que la simulación de un proyecto de inversión arrojó los resultados que se muestran en la tabla 6.1. Para estos resultados se desea determinar un intervalo de confianza del 95% ($\alpha = 5\%$) para la tasa interna de rendimiento (TIR).

Puesto que la variancia del universo de donde procede la muestra se desconoce, entonces el intervalo de confianza para la TIR, se determina a partir de la expresión (6.7), esto es:

$$13.37\% \pm \frac{6.71}{\sqrt{100}} (1.99)\%$$

$$13.37\% \pm 1.34\%$$

lo cual define el intervalo (12.03%; 14.71%) cuyo ancho es de 2.68%.

Tabla 6.1. Resultados (TIR's) de la simulación de un proyecto de inversión (%).

12	8	5	3	7	5	3	29	27	9
15	22	7	21	3	15	26	17	8	13
18	13	18	17	5	3	4	7	13	19
11	16	13	13	9	8	7	13	15	12
13	12	14	14	20	22	11	12	18	11
17	4	23	8	15	7	18	19	24	15
19	24	12	7	18	15	13	8	21	4
9	17	11	15	21	18	15	28	13	3
14	19	19	18	4	27	21	4	5	8
10	21	25	21	7	2	8	9	7	11

Finalmente, conviene señalar que si el ancho del intervalo se quisiera reducir a 1.34%, entonces se necesitaría aumentar a 400 el número de simulaciones de la TIR.

6.2. SIMULACION REGENERATIVA

La simulación de un sistema estocástico puede ser visto como un experimento estadístico, donde primero se construye el modelo del sistema, el cual captura la esencia del problema y describe su estructura fundamental en términos de relaciones causa-efecto. Enseguida, con la ayuda de una computadora se realizan experimentos y se analizan los resultados de la simulación con el propósito de hacer inferencias sobre el comportamiento del sistema. Por consiguiente, los experimentos de simulación son experimentos estadísticos, y como tales, están basados en procedimientos estadísticos profundos. En particular, las técnicas de estimación son muy útiles en estudios de simulación, pues permiten hacer inferencias estadísticas sobre las variables estocásticas analizadas en el modelo.

Recientemente, una nueva metodología estadística ha sido desarrollada para analizar los resultados de una simulación. A esta nueva técnica se le conoce con el nombre de "Simulación regenerativa". Este nuevo enfoque se basa en el hecho de que muchos sistemas estocásticos se regeneran (empiezan en el mismo punto de partida) a intervalos regulares de tiempo. Esta particularidad permite al analista observar y estudiar durante el curso de la simulación, bloques de datos independientes e idénticamente distribuidos. Más generalmente, la simulación regenerativa resuelve los problemas de: 1) ¿Cómo empezar la simulación? 2) ¿Cuándo empezar a coleccionar datos? 3) Tamaño de la corrida y núme-

ro de corridas y 4) ¿Cómo manejar los resultados cuando estos están correlacionados?

6.2.1. Simulación de un sistema de colas de un servidor

La representación de un sistema de colas de un servidor aparece en la figura 6.1. En esta figura se puede apreciar que los clientes son generados por una fuente alimentadora, la cual genera X número de clientes por unidad de tiempo. Si un cliente que llega al sistema encuentra desocupado al servidor, entonces este cliente empieza a ser inmediatamente servido y su salida del sistema empieza inmediatamente después de que su servicio ha terminado. Por el contrario, si un cliente al arribar encuentra al servidor ocupado, entonces este cliente pasa a un salón de espera donde esperará su turno para ser servido. Finalmente, los clientes que permanecen en la cola, son servidos en base a atender primero a los que llegaron primero a la cola.

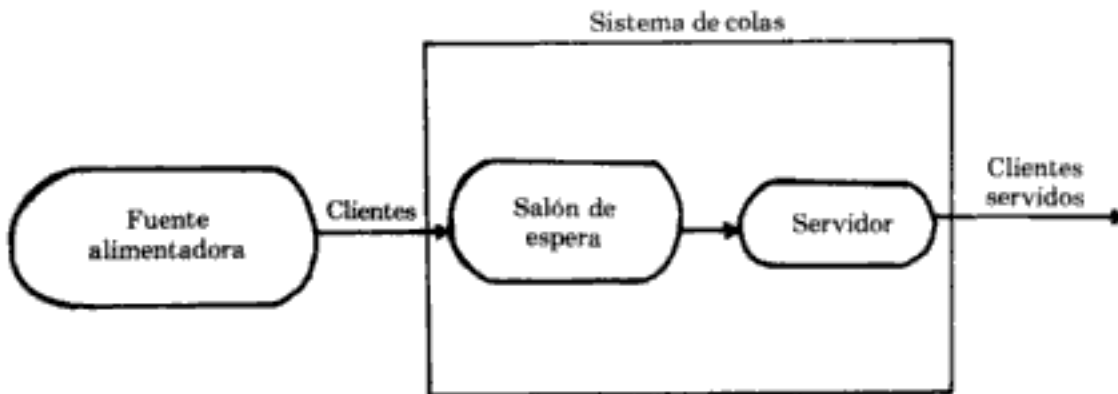


Figura 6.1. Sistema de colas de un servidor.

Para el sistema particular que se está analizando, suponga que la fuente alimentadora genera un nuevo cliente cada 2 minutos. También, suponga que el tiempo de servicio es una variable aleatoria uniformemente distribuida entre 1 y 2 minutos. Además, suponga que la forma de evaluar el funcionamiento y operación del sistema, es en base al valor esperado del tiempo que un cliente permanece en la cola $E(W)$, cuando el sistema se encuentra en una situación de estado estable. Para este sistema en particular, no existen procedimientos analíticos capaces de obtener $E(W)$. Por consiguiente, la única forma de evaluar y analizar este sistema es a través de simulación.

Por consiguiente, el objetivo de la simulación de este sistema, es analizar los resultados obtenidos con el propósito de estimar $E(W)$. Más generalmente, se desea que la estimación de $E(W)$ sea confiable, es decir, se requiere que el valor de $E(W)$ pertenezca a un intervalo con un cierto nivel de confianza, por ejemplo del 95% ($\alpha = 5\%$).

Una forma razonable de estimar $E(W)$ sería: sea W_1 el tiempo de espera en la cola del cliente 1, W_2 el tiempo de espera en la cola del cliente 2, y así sucesivamente. Por consiguiente, si durante la simulación del sistema entraron a éste N clientes, entonces el promedio muestral sería:

$$\frac{W_1 + W_2 + \dots + W_n}{N} \quad (6.9)$$

el cual representa a un estimador consistente de $E(W)$. Sin embargo, el promedio muestral sería en general un estimador sesgado de $E(W)$ debido a las condiciones iniciales. Por ejemplo, si W_1 tiene un valor de cero, entonces los tiempos de espera de los primeros clientes tienden a ser pequeños.

La forma tradicional de tratar el sesgo del estimador debido a condiciones iniciales, es correr el modelo de simulación durante algún tiempo sin coleccionar ningún dato, hasta que el sistema de colas haya alcanzado la condición de estado estable. A partir de este instante, se empieza a coleccionar la información. Por ejemplo, se puede simular el sistema para 1000 clientes y sólo considerar los tiempos de espera de los últimos 500, y entonces usar:

$$\frac{W_{501} + W_{502} + \dots + W_{1000}}{500} \quad (6.10)$$

como un estimador de $E(W)$. Sin embargo, esta solución no es muy recomendable por las siguientes razones: 1) Es muy difícil determinar la longitud del periodo de estabilización y 2) Una gran cantidad de tiempo de computadora es desperdiciado.

Otra de las desventajas que se le atribuyen al procedimiento anterior, es la alta correlación que existe entre W_k y W_{k+1} independientemente de que el sistema se encuentre o no en estado estable. Por consiguiente, si W_1, W_2, \dots, W_n no son independientes, la teoría estadística clásica sería de poca utilidad en la determinación del intervalo de confianza del tiempo promedio de espera en la cola.

Se puede observar de los comentarios anteriores, que el sesgo debido a condiciones iniciales transientes y a la alta correlación de los datos, ofrecen serios obstáculos en la determinación de $E(W)$. Por consiguiente, es necesario buscar y proponer un nuevo procedimiento que elimine y resuelva los problemas antes mencionados. Afortunadamente, sí existe tal procedimiento al cual se le llama simulación regenerativa.

Para entender la lógica de esta nueva técnica, suponga que al empezar la simulación se establece $W_1 = 0$. En seguida, se empieza a correr el modelo de simulación y valores del tiempo de espera en la cola

empiezan a ser colectados. La obtención de tales datos suponga que son tal y como se muestran en la tabla 6-2. En esta tabla se puede observar que los clientes 1, 4, 8, 9, 14, 17 y 20 fueron muy afortunados, puesto que encontraron desocupado al servidor al momento de arribar al sistema. También, en esta tabla se puede apreciar que el servidor está constantemente ocupado desde que llega el cliente 1 hasta que se va el cliente 3, en seguida se encuentra ocioso hasta que llega el cliente 4, y así sucesivamente. Si más de 20 clientes son servidos, entonces el sistema simulado mostrará este mismo comportamiento, es decir, servidor ocupado, servidor ocioso, servidor ocupado, servidor ocioso, y así sucesivamente. Por consiguiente, si al tiempo que transcurre entre empezar a servir y quedar ocioso, se le considera como un ciclo, entonces para la corta simulación mostrada en la tabla 6.2, existen 6 ciclos, donde el primer ciclo comprende a los clientes 1, 2 y 3, el segundo ciclo comprende a los clientes 4, 5, 6 y 7 y así sucesivamente.

Tabla 6.2. Tiempo de espera en la cola

No. Cliente	W		No. Cliente	W		No. Cliente	W	
1	0	} 1	8	0	} 3	14	0	} 5
2	10		9	0		15	16	
3	5		10	18		16	12	
4	0	} 2	11	20	} 4	17	0	} 6
5	25		12	30		18	25	
6	15		13	40		19	30	
7	20					20	0	

El séptimo ciclo empieza con el arribo del cliente número 20. Por consiguiente, un nuevo ciclo es iniciado cuando un cliente llega al sistema y encuentra al servidor ocioso.

Puesto que cada vez que un cliente llega al sistema y encuentra al servidor ocioso, es equivalente y está gobernado por la misma estructura probabilística que cuando llega el primer cliente al sistema, entonces sería razonable agrupar los resultados de la simulación en bloques de datos, donde el primer bloque contendrá los tiempos de espera en la cola de los clientes que forman parte del primer ciclo, el segundo bloque contendrá los tiempos de espera en la cola de los clientes que forman parte del segundo ciclo, y así sucesivamente. Para la corta simulación mostrada en la tabla 6.2, los bloques serían:

$$\begin{aligned}
 &(W_1, W_2, W_3), (W_4, W_5, W_6, W_7), (W_8), \\
 &\quad (W_9, W_{10}, W_{11}, W_{12}, W_{13}), \\
 &\quad (W_{14}, W_{15}, W_{16}), (W_{17}, W_{18}, W_{19})
 \end{aligned}$$

De este modo, puesto que cada ciclo es iniciado en las mismas con-

diciones, los bloques de datos de ciclos sucesivos son estadísticamente independientes e idénticamente distribuidos. Por consiguiente, si se define E_k como la suma de los tiempos de espera de los clientes servidos en el ciclo k , y C_k como el número de clientes servidos en el ciclo k , entonces los pares (E_1, C_1) , (E_2, C_2) , (E_3, C_3) , (E_4, C_4) , (E_5, C_5) y (E_6, C_6) son estadísticamente independientes e idénticamente distribuidos, a pesar de la alta correlación que existe entre E_k y C_k . De la tabla 6.2, los valores de (E_k, C_k) que se obtienen son los siguientes:

$$\begin{array}{ll} (E_1, C_1) = (15, 3) & (E_4, C_4) = (108, 5) \\ (E_2, C_2) = (60, 4) & (E_5, C_5) = (28, 3) \\ (E_3, C_3) = (0, 1) & (E_6, C_6) = (55, 3) \end{array}$$

Por consiguiente, la alta correlación entre W_k y W_{k+1} ha sido eliminada y los datos han sido agrupados en bloques estadísticamente independientes e idénticamente distribuidos.

De acuerdo a la agrupación anterior, y suponiendo que N clientes fueron servidos durante la simulación, a partir de la cual se formaron n bloques, la estimación del tiempo de espera en la cola se puede determinar por medio de la siguiente expresión:

$$E(W) = \frac{W_1 + W_2 + \dots + W_n}{N} = \frac{E_1 + E_2 + \dots + E_n}{C_1 + C_2 + \dots + C_n} \quad (6.11)$$

Sin embargo, la segunda parte de la expresión (6.11) puede ser escrita de la siguiente forma:

$$E(W) = \frac{(E_1 + E_2 + \dots + E_n)/n}{(C_1 + C_2 + \dots + C_n)/n} \quad (6.12)$$

Puesto que (E_1, E_2, \dots, E_n) y (C_1, C_2, \dots, C_n) son grupos de variables aleatorias independientes e idénticamente distribuidas, entonces por el teorema del límite central, la expresión (6.12) converge a:

$$\begin{array}{l} E(W) = E(E)/E(C) \\ \text{Lím} \\ n \rightarrow \infty \end{array} \quad (6.13)$$

Por consiguiente, el problema de estimar $E(W)$ se traduce a estimar $E(E)/E(C)$. Y puesto que $E(E)/E(C)$ puede ser estimado a partir de (E_1, C_1) , (E_2, C_2) , ..., (E_n, C_n) , la estadística clásica puede ser utilizada para hacer inferencias acerca el valor verdadero de $E(W)$. Además, a través de la estadística clásica se puede determinar un intervalo de confianza para $E(W)$.

el cual define un intervalo cuyo ancho es de:

$$\frac{2 S Z_{\alpha/2}}{C \sqrt{n}} \quad (6.28)$$

Finalmente, conviene mencionar que aunque la variable aleatoria $E(E)/E(C)$ sigue una distribución de probabilidad t puesto que δ^2 se desconoce, la expresión (6.27) sigue siendo válida ya que para valores de n muy grandes, la distribución t se aproxima a la distribución normal.

Ejemplo 6.2

Para los tiempos de espera en la cola mostrados en la tabla 6-2, determine un intervalo de confianza del 90% ($\alpha = 10\%$) para $E(W)$.

Para los tiempos de espera mostrados en la tabla 6-2, los bloques de datos son:

$$\begin{array}{ll} (E_1, C_1) = (15, 3) & (E_4, C_4) = (108, 5) \\ (E_2, C_2) = (60, 4) & (E_5, C_5) = (28, 3) \\ (E_3, C_3) = (0, 1) & (E_6, C_6) = (55, 3) \end{array}$$

De acuerdo a estos bloques de datos, los promedios muestrales \bar{E} y \bar{C} y el parámetro t serían:

$$\bar{E} = \frac{15+60+0+108+28+55}{6} = 44.33$$

$$\bar{C} = \frac{3+4+1+5+3+3}{6} = 3.16$$

$$\hat{t} = \frac{\bar{E}}{\bar{C}} = \frac{44.33}{3.16} = 14.03$$

Por otra parte, los valores correspondientes de S_E^2 , S_C^2 , S_{EC} y S^2 serían:

$$S_E^2 = \frac{1}{5} \left((15-44.33)^2 + (60-44.33)^2 + (0-44.33)^2 + (108-44.33)^2 + (28-44.33)^2 + (55-44.33)^2 \right)$$

$$S_E^2 = 1501.06$$

muestran a continuación, determine un intervalo de confianza del 95% para la probabilidad de interferencia.

0.30 0.22 0.25 0.33 0.30 0.38 0.29 0.31 0.32 0.34
 0.28 0.34 0.27 0.31 0.35 0.37 0.39 0.35 0.43 0.28
 0.29 0.33 0.23 0.36 0.25 0.35 0.30 0.27 0.28 0.24
 0.33 0.31 0.28 0.24 0.29 0.36 0.41 0.38 0.34 0.29
 0.35 0.34 0.24 0.29 0.24 0.33 0.38 0.36 0.38 0.32
 0.37 0.29 0.37 0.22 0.32 0.28 0.42 0.41 0.37 0.33
 0.35 0.26 0.36 0.39 0.31 0.26 0.43 0.44 0.39 0.39
 0.34 0.27 0.35 0.34 0.34 0.29 0.38 0.29 0.40 0.42
 0.28 0.31 0.31 0.33 0.35 0.34 0.39 0.30 0.42 0.31
 0.39 0.34 0.34 0.30 0.34 0.36 0.43 0.33 0.44 0.25

- 6.3. Determine un intervalo de confianza del 95% para el valor presente neto del problema 8 del capítulo anterior.
- 6.4. Para el problema 9 del capítulo anterior, simule 100 TIR's y determine un intervalo de confianza del 90% para la tasa interna de rendimiento.
- 6.5. Para los tiempos de espera en la cola mostrados a continuación, determine un intervalo de confianza del 90% para $E(W)$.

<i>No. Cliente</i>	<i>W</i>	<i>No. Cliente</i>	<i>W</i>	<i>No. Cliente</i>	<i>W</i>
1	0	11	0	21	15
2	10	12	15	22	20
3	12	13	0	23	0
4	8	14	0	24	10
5	0	15	18	25	18
6	0	16	14	26	0
7	18	17	0	27	0
8	24	18	19	28	15
9	30	19	23	29	18
10	28	20	25	30	0

- 6.6. Si se simula la llegada de 100 clientes a la primera estación de servicio del sistema de colas descrito en el problema 12 del capítulo anterior, determine un intervalo de confianza del 90% para $E(W)$.
- 6.7. Si se simula la llegada de 100 clientes al sistema de colas descrito en el problema 13 del capítulo anterior, determine un intervalo de confianza del 95% para $E(W)$.

- 6.8 Suponga que en el problema 16 del capítulo anterior, 5 máquinas son asignadas a un mecánico. Si se simula el sistema durante 1000 horas, determine un intervalo de confianza del 90% para el tiempo que una máquina permanece ociosa esperando ser reparada.
- 6.9. Si se simula la llegada de 100 camiones al sistema de colas descrito en el problema 17 capítulo anterior, determine un intervalo de confianza del 99% para $E(W)$.
- 6.10. a) La demanda diaria de un cierto artículo está regida por una distribución binomial con parámetros $n=9$ y $\theta=1/3$. El tiempo de entrega es instantáneo. El costo de mantener una unidad en inventario es de \$25 por día y el costo de faltante es de \$100 por unidad. Además, el gerente de producción ha establecido como política, colocar una orden cada vez que el inventario sea menor o igual a 3 unidades. La cantidad ordenada deberá elevar el nivel de inventario a 7 unidades. Para esta política, determine el costo promedio diario en condiciones de estado estable (Suponga un inventario inicial de 7 unidades y que un ciclo se inicia cada vez que existen 7 unidades al principio del día).
Sugerencia: El costo promedio diario para este problema se puede determinar de acuerdo a la siguiente expresión:

$$CT = 25 E \{(x - D)^+\} + 100 E \{(D - x)^+\}$$

donde:

CT = Costo total diario.

x = Nivel de inventario al final del día.

D = Demanda diaria.

y puesto que:

$$a^+ = \max(a, 0)$$

$$(a - b)^+ = a - \min(a, b)$$

entonces, la expresión anterior se transforma en:

$$CT = 25 E \{x\} + 100 E \{D\} - 125 E \{\min(x, D)\}$$

- b) Si se simulan 30 días y se obtienen los resultados que se muestran a continuación, determine un intervalo de confianza del 90% para el costo promedio diario.

7

LENGUAJES DE SIMULACION

Las primeras etapas de un estudio de simulación se refieren a la definición del sistema a ser modelado y a la descripción del sistema de términos de relaciones lógicas de sus variables y diagramas de flujo. Sin embargo, llega el momento de describir el modelo en un lenguaje que sea aceptado por la computadora que se va a usar. En esta etapa se tienen dos cursos de acción a seguir: 1) Desarrollar el software requerido para estudios de simulación, o 2) Comprar algún paquete de simulación como GPSS, GASP, simula, etc. Desafortunadamente, para la segunda alternativa existen una gran cantidad de paquetes (lenguajes de programación de propósito especial), que resulta extremadamente difícil decidir cuál paquete se ajusta mejor a una aplicación determinada. Esta situación origina que en la mayoría de las veces, la selección de un paquete depende de si el analista lo conoce, lo entiende y lo sabe aplicar. No obstante, estas dos alternativas se deben de evaluar en términos económicos antes de tomar una decisión.

7.1. VENTAJAS DE LOS LENGUAJES DE SIMULACION

El proceso evolutivo de los lenguajes de simulación ha sido largo y extenso. Empezó a finales de la década de los 50's. En un principio los lenguajes que se desarrollaron eran de propósito general. Sin embargo, poco a poco los estudiosos de este tema se dieron cuenta de la gran similitud que existía entre las diferentes situaciones o sistemas que se simulaban. Lo anterior condujo obviamente al desarrollo de lenguajes de propósito especial, los cuales en la actualidad tienen una gran de-

manda y su proceso de comercialización ha sido amplio y extenso. Entre las ventajas principales de estos lenguajes de simulación, se pueden mencionar las siguientes:

- a) Reducción en la tarea de programación. Con los lenguajes de simulación, el tiempo dedicado a la programación del modelo se reduce considerablemente. Existen algunos paquetes como GPSS, en los que con un número muy reducido de estatutos, se pueden simular sistemas que en otro lenguaje como Fortran, requerirían una gran cantidad de estatutos y subrutinas.
- b) Mejor definición del sistema. A través de los lenguajes de simulación, se facilita la tarea de definir las diferentes entidades que interactúan dentro del sistema. También, con estos lenguajes se determina con mayor facilidad las interrelaciones que existen entre las entidades que forman el sistema.
- c) Mayor flexibilidad para cambios. Con los lenguajes generales como Fortran, el proceso de cambios puede ser largo y tedioso. Sin embargo, con el uso de lenguajes de simulación, los cambios son una tarea simple y rutinaria.
- d) Mejor diferenciación de las entidades que forman el sistema. El uso de lenguajes de simulación facilita determinar o definir las características y atributos de una entidad. Con las entidades bien definidas y diferenciadas, se aumenta y mejora el entendimiento del sistema a simular.
- e) Se relacionan mejor las entidades. Con las entidades bien definidas, los lenguajes de simulación permiten relacionar mejor a cada una de estas entidades, es decir, se determina más fácilmente las relaciones que las entidades guardan entre sí y el análisis de cada una de ellas.

7.2. CARACTERÍSTICAS DE LOS LENGUAJES DE SIMULACION

Los lenguajes de simulación que actualmente existen en el mercado, tienen una serie de características propias que los diferencian de los demás. Entre estas características se pueden mencionar las siguientes:

- a) El procedimiento utilizado para generar números aleatorios uniformes.
- b) Los procedimientos o métodos utilizados para generar las variables aleatorias no-uniformes más conocidas y más usadas.

- c) La forma de adelantar el "reloj de la simulación", la cual puede ser: 1) Incrementos a tiempo fijo, o 2) Incrementos al próximo evento.
- d) El análisis estadístico de los resultados de la simulación.
- e) El formato en que los resultados de la simulación son presentados.
- f) La forma en que las inconsistencias y errores de lógica es reportada.
- g) El lenguaje en el cual el paquete está escrito, el cual puede ser: Fortran, Algol, PL/I, Assembler, etc.
- h) Los diferentes tipos de computadoras cuyo compilador es compatible con el del paquete en cuestión.

A continuación, se presentan las características principales de los lenguajes de simulación más usados.

GPSS (General Purpose Simulation System)

Persona que lo desarrolló:	Geoffrey Gordon
Versiones más conocidas:	GPSS I, GPSS II, GPSS III, GPSS/360, GPSS V.
Lenguaje del paquete:	Assembler
Reloj de la simulación:	Incrementos al próximo evento
Computadoras compatibles:	Generalmente se adapta a cualquier tipo de computadora.

SIMSCRIPT (No tiene ningún significado)

Personas que lo desarrollaron:	H. M. Markowitz, H. W. Karr y B. Hausner.
Versicnes más conocidas:	Simscript I, Simscript I.5, Simscript II, Simscript II.5, C-Simscript.
Lenguaje del paquete:	Fortran las primeras versiones, Assembler las últimas.
Reloj de la simulación:	Incrementos al próximo evento para el caso discreto, e incrementos a tiempo fijo para el caso continuo (C-Simscript).
Computadoras compatibles:	CDC 6000/7000, univac 1100, IBM 360/370, Honeywell 600/6000.

GASP (General Activity Simulation Program)

Personas que lo desarrollaron:	P. J. Kiviat y A Colher
Versiones más conocidas:	GASP II, GASP IV, GASP-PLUS-
Lenguaje del paquete:	Fortran, PL/I.
Reloj de la simulación:	Incrementos al próximo evento para el caso discreto, e incrementos a tiempo fijo para el caso continuo (GASP IV y PLUS).
Computadoras compatibles:	Cualquier computadora con compilador de Fortran o PL/I.

SLAM (Simulation Language for Alternative Modeling)

Personas que lo desarrollaron:	A. Alám B. Pritsker y Asociados
Versiones más conocidas:	SLAM fue el resultado de la fusión de varios lenguajes como GASP IV y QGERT.
Lenguaje del paquete:	Fortran IV.
Reloj de la simulación:	Incrementos al próximo evento para el caso discreto, e incrementos a tiempo fijo para el caso continuo.
Computadoras compatibles:	Cualquier computadora con compilador de Fortran

Cualesquiera de estos lenguajes tienen sus propias ventajas y desventajas y no se puede decir que un lenguaje es mejor que otro. Generalmente, entre más fácil de aprender y de usar sea un lenguaje, menor será su flexibilidad y su eficiencia. Por consiguiente, decidir qué lenguaje utilizar en una aplicación específica, no es una tarea fácil de realizar.

7.3. FACTORES A CONSIDERAR EN LA SELECCION DE UN LENGUAJE

La selección de un lenguaje de simulación generalmente está sujeta al tipo de computadora que se tiene disponible, es decir, en la mayoría de las veces ya se cuenta con cierta configuración de hardware.

re. Por consiguiente, conociendo la computadora que se va a usar, los factores a considerar en la selección del lenguaje serían:

- a) Los manuales disponibles. Es muy importante considerar la facilidad de entender e interpretar los manuales disponibles.
- b) Compilador del lenguaje. Es necesario investigar la compatibilidad del compilador del lenguaje con la computadora disponible.
- c) La documentación y diagnóstico de errores. Es conveniente analizar la forma en que el lenguaje reporta las inconsistencias y los errores de lógica.
- d) La eficiencia. Uno de los factores principales a considerar en la selección de un lenguaje es su eficiencia de operación. Dentro de la eficiencia se considera el tiempo de organizar, programar, compilar y ejecutar.
- e) Los costos involucrados. Entre los costos que origina la adquisición de un paquete se pueden mencionar: El costo de instalación, el costo de mantenimiento y actualización del software y el costo de operación.
- f) Conocimiento del lenguaje. Otro factor importante a considerar en la selección del lenguaje, es el conocimiento y dominio que de éste tengan las personas o analistas encargados de realizar los estudios de simulación.
- g) Justificación económica. Finalmente, y el más importante de todos, es la justificación económica del lenguaje de simulación. A este respecto, es conveniente señalar que la adquisición de un paquete se debe de considerar como un proyecto de inversión, donde los beneficios que se derivan de tal adquisición, deben de compensar la inversión y los gastos que origina.

7.4. CLASIFICACION DE LOS LENGUAJES DE SIMULACION

Los modelos de simulación se clasifican en dos categorías: 1) Modelos de simulación continua y 2) Modelos de simulación discreta. Los modelos de simulación continua, generalmente son apropiados cuando el analista considera al sistema bajo estudio, como un flujo continuo de información. En simulación continua, generalmente el reloj de simulación se incrementa a intervalos fijos de tiempo. Por otra parte, en modelos de simulación discreta, al analista le interesa lo que le sucede a entidades individuales del sistema. Por consiguiente, el reloj de la simulación en simulación discreta, se incrementa cada vez que ocurre un evento.

Los modelos de simulación discreta pueden desarrollarse a través de tres enfoques: 1) Enfoque de eventos, 2) Enfoque de actividades y 3) Enfoque de procesos. Un evento se define como un cambio en el estado de una entidad del sistema. Una actividad es una colección de operaciones que transforman el estado de una entidad. Un proceso es una secuencia de eventos que ocurren en un tiempo determinado. Una explicación gráfica de estos conceptos aparece en la figura 7.1.

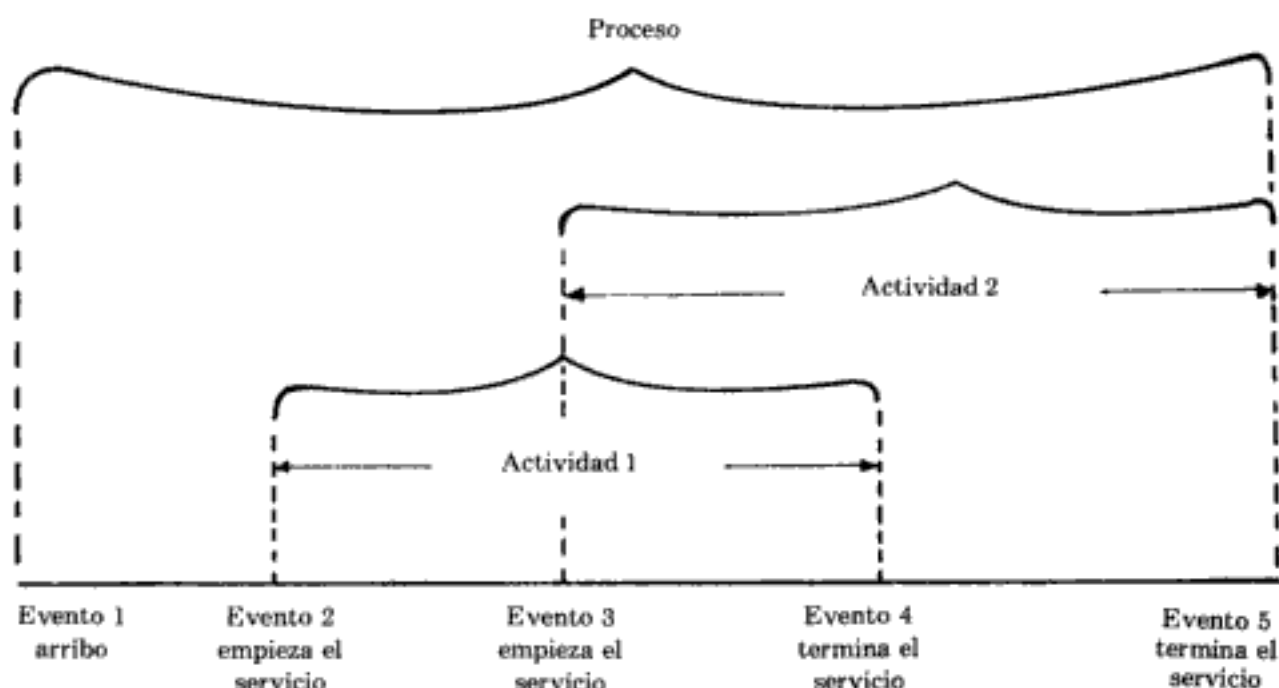


Figura 7.1 Eventos, actividades y procesos.

Además de los tres enfoques anteriores, algunos autores hacen una diferenciación entre lenguajes con orientación simbólica (GPSS) y lenguajes con orientación de estatutos de programación. Los primeros son más fáciles de aprender, pero los segundos son más flexibles. La figura 7.2, muestra una clasificación de los lenguajes de simulación. En esta figura se puede apreciar que los lenguajes simbólicos a pesar de tener un enfoque de procesos en su diseño, se consideran como una cuarta categoría para diferenciarlos de los lenguajes de orientación de estatutos de programación.

Finalmente, vale la pena mencionar que recientemente se han estado desarrollando lenguajes que permiten el análisis y evaluación de modelos discretos y modelos continuos. GASP IV y C-SIMSCRIPT fueron los primeros lenguajes de este tipo que se desarrollaron, aunque recientemente está disponible un nuevo lenguaje de este tipo llamado SLAM. Estos lenguajes permiten al analista modelar sistemas discretos, sistemas continuos y sistemas donde algunas entidades se comportan en forma discreta y otras en forma continua.

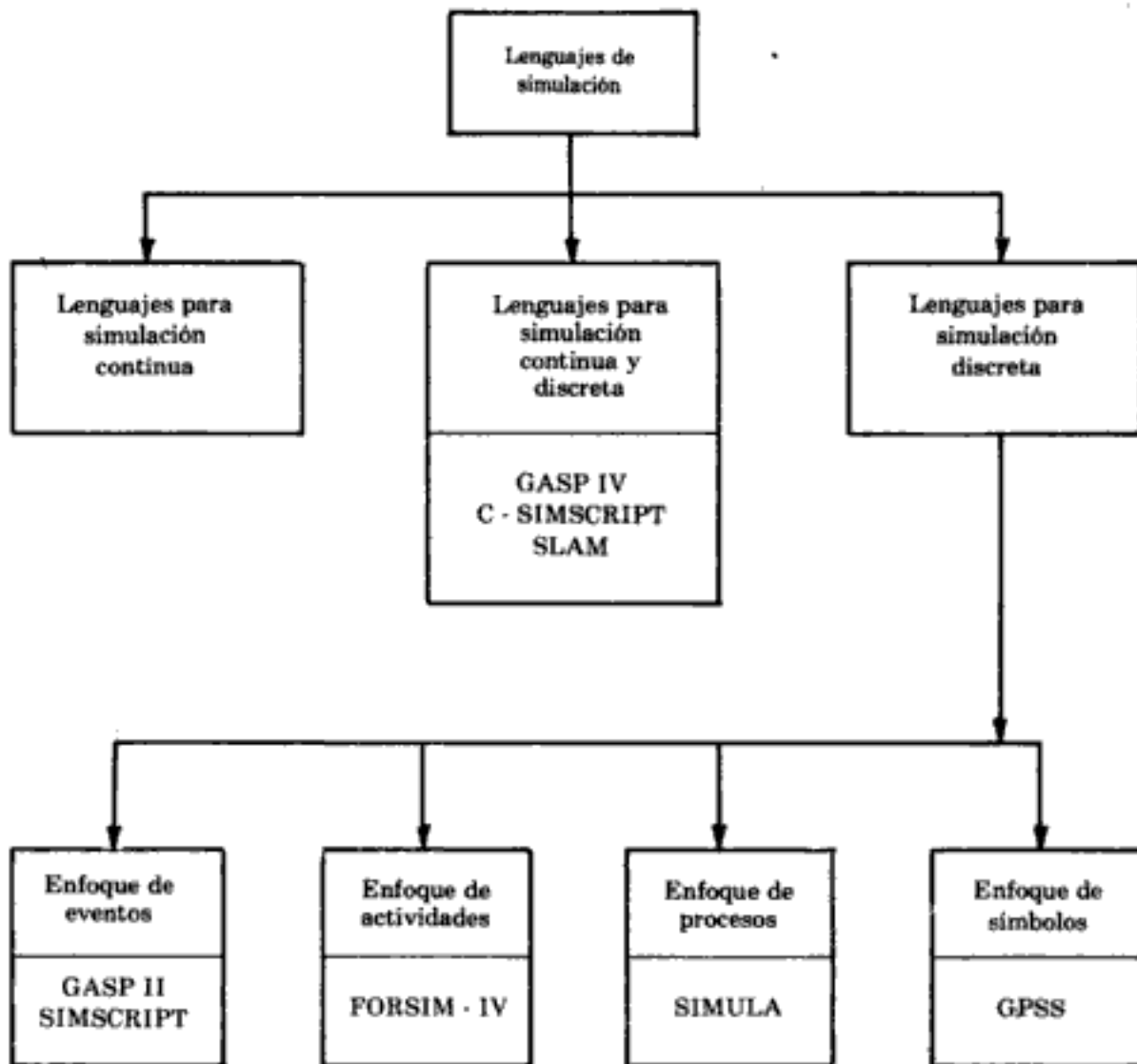


Figura 7.2 Clasificación de los lenguajes de simulación

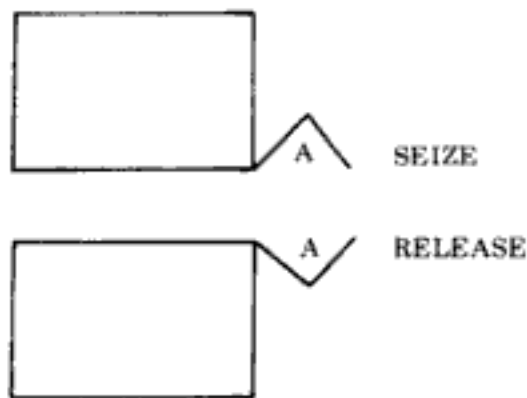
7.5. INTRODUCCION AL GPSS

Como se ha mencionado en párrafos anteriores, existen actualmente en el mercado una gran cantidad de paquetes de simulación. Sin embargo, el paquete GPSS es sin duda uno de los más conocidos y usados en empresas e instituciones educativas.* Por esta razón, en esta sección se tratará de describir a través de una serie de ejemplos, las potencialidades y los diferentes campos de aplicación en los cuales se puede aplicar este paquete (versión GPSS-V). Obviamente esta presentación no incluye toda la gama de aplicaciones que se le pueden dar a este paquete. Para tener una concepción más amplia del GPSS, se puede recurrir al libro "Simulation Using GPSS" cuyo autor es Thomas J. Schriber.

Actualmente en el ITESM se está utilizando la versión GPSS-V.

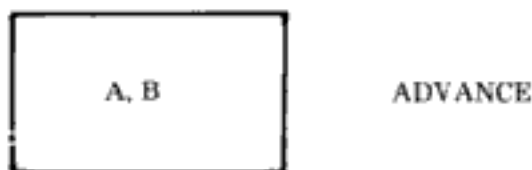
- Campo B** = Desviación con respecto a la media. Por ejemplo si el tiempo entre llegadas es uniforme entre 10 y 18 minutos, entonces, $A = 14$ y $B = 4$.
- Campo C** = Especifica el tiempo de ocurrencia de la primera transacción.
- Campo D** = Especifica el límite de transacciones que pueden ocurrir en el bloque **GENERATE**.
- Campo E** = Especifica la prioridad de las transacciones al momento de generarse.

Para representar la llegada y salida de una facilidad de servicio (1 servidor), los bloques que se utilizan son:



Donde el campo A representa la facilidad de servicio que se está utilizando. Lo anterior significa que en un sistema puede haber varias facilidades de servicio y a cada una de ellas es necesario identificarla con un nombre o un número. El primer bloque **SEIZE** indica que se inicia la entrada a la facilidad de servicio A y el bloque **RELEASE** representa la salida de dicha facilidad.

El tiempo de servicio se representa mediante el siguiente bloque:



donde:

- Campo A** = Tiempo promedio de servicio.
- Campo B** = Desviación con respecto a la media

Finalmente, para llevar un conteo de la cantidad de transacciones (clientes) que han pasado por el sistema, se utiliza el bloque:

RELATIVE CLOCK 18015
BLOK COUNTS

ABSOLUTE CLOCK 18240

BLOK	CURRENT	TOTAL	BLOCK	CURRENT	TOTAL
1	0	722	10	0	278
2	0	722	11	0	277
3	0	723	12	0	277
4	0	723	13	0	277
5	0	723	14	0	277
6	0	723	15	0	277
7	0	723	16	0	277
8	0	723	17	0	277
9	0	278	18	0	277

FACILITIES

Average utilization during

Facility	Number entries	Average time/tran	Total time	Avail time	Unavail time	Current status	Percent availability
1	1000	17.232	0.95			100.0	

QUEUES

Queue	Maximum content	Average content	Total entries	Zero entries	Percent zeros	Average time/tran	\$ Average time/tran	Current content
1	6	1.019	1001	152	15.1	19.976	23.552	1

\$ Average time/tran = average time/tran excluding zero entries

Figura 7.4 Resultados obtenidos de la simulación de la peluquería al introducir los bloques queue, depart, mark y tabulate.

Tablas

Tabla 1

<i>Entries in table</i> 723		<i>Mean argument</i> 36.116		<i>Standard deviation</i> 19.562		<i>Sum of arguments</i> 26112.0	
<i>Upper limit</i>	<i>Observed frequency</i>	<i>Percent of total</i>	<i>Cumulative porcentaje</i>	<i>Cumulative remainder</i>	<i>Mutiple of mean</i>	<i>Deviation from mean</i>	
5	0	0.00	0.0	100.0	0.138	-1.590	
15	67	09.26	9.2	90.7	0.415	-1.079	
25	192	26.55	35.8	64.1	0.692	-0.568	
35	153	21.16	56.9	43.0	0.969	-0.057	
45	123	17.91	73.9	26.0	1.245	0.454	
55	71	9.82	83.8	16.1	1.522	0.965	
65	60	8.29	92.1	7.8	1.799	1.476	
75	24	3.31	95.4	4.5	2.076	1.987	
85	9	1.24	96.6	3.3	2.353	2.498	
95	10	1.38	98.0	1.9	2.630	3.010	
105	11	1.52	99.5	0.4	2.907	3.521	
115	3	0.41	100.0	0.0	3.184	4.032	

Remaining frequencies are all zero

Figura 7.4 Continuación.

Tablas

Table 2		Tablas				
Entries in table 277		Mean argument 40.169	Standard deviation 18.062		Sum of arguments 11127.0	
Upper limit	Observed frequency	Percent of total	Cumulative percentage	Cumulative remainder	Multiple of mean	Deviation from mean
5	0	0.00	0.0	100.0	0.124	-1.947
15	0	0.00	0.0	100.0	0.373	-1.393
25	57	20.57	20.5	79.4	0.622	-0.839
35	85	30.68	51.2	48.7	0.871	-0.286
45	52	18.77	70.0	29.9	1.120	0.267
55	30	10.83	80.8	19.1	1.369	0.821
65	28	10.10	90.9	9.0	1.618	1.374
75	12	4.33	95.3	4.6	1.867	1.928
85	2	0.72	96.0	3.9	2.116	2.481
95	6	2.16	98.1	1.8	2.364	3.035
105	3	1.08	99.2	0.7	2.613	3.589
115	2	0.72	100.0	0.0	2.862	4.142

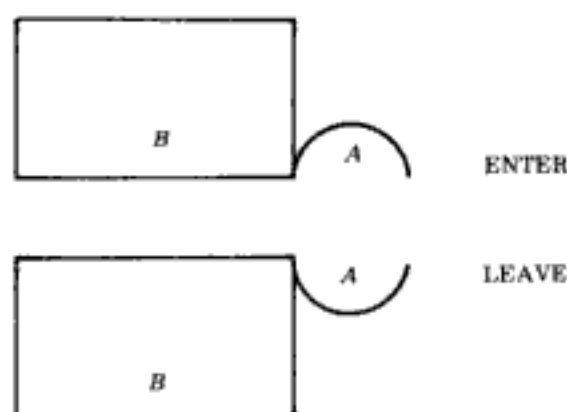
Remaining frequencies are all zero

Figura 7.4 Continuación.

Ejemplo 7.2. Centro Electrónico de Cálculo.

Al salón de terminales del Centro Electrónico de Cálculo del ITESM, arriban estudiantes de acuerdo a una distribución uniforme entre 2 y 8 minutos. En el salón existen 10 terminales. El tiempo promedio que un estudiante usa la terminal, está uniformemente distribuido entre 40 y 80 minutos. Si al llegar al salón de terminales el estudiante las encuentra todas ocupadas, entonces, existe un 50% de probabilidad de que el estudiante regrese después de 5 minutos (suponga que en esos 5 minutos el estudiante va a la cafetería a comprar golosinas y refrescos). ¿Determine al momento de que 1000 estudiantes hayan pasado por el salón de terminales, la cantidad de estudiantes que al llegar al sistema encontraron todas las terminales ocupadas y no regresaron?

Para simular este problema se requiere utilizar nuevos bloques. Por ejemplo, es necesario conocer qué bloques utilizar para representar a una facilidad de servicio con muchos servidores* (cada terminal es un servidor con un tiempo de servicio uniforme ente 40 y 80 minutos). También, se requiere conocer el bloque que representa las transferencias condicionales e incondicionales que ocurren en el sistema. Para el primer caso, los bloques que se utilizan son los siguientes:



donde:

- A = Representa el nombre o el número que se le da a la facilidad de servicio (STORAGE).
- B = Representa la cantidad de servidores que se ocupan o desocupan a la vez, al llegar o abandonar el sistema una determinada transacción.

Esto significa que en un sistema pueden existir varias facilidades de servicio con varios servidores y a cada una de ellas es necesario identificarlas con un nombre o un número. El primer bloque ENTER indica que se ini-

*Cuando se tiene un solo servidor se deben utilizar los bloques SEIZE y RELEASE.

cia la entrada a la facilidad de servicio STORAGE A y el bloque LEAVE representa la salida de dicha facilidad.

Por otra parte, para representar las transferencias condicionales e incondicionales que ocurren en un sistema, se utiliza el siguiente bloque:



Este bloque puede ser codificado de las siguientes formas:

```
TRANSFER .3, ETI 1, ETI 2
TRANSFER BOTH,, ETI 1
TRANSFER , ETI 2
```

El primer caso indica que la probabilidad de ir al bloque con etiqueta ETI 1 es de 0.70. El complemento es la probabilidad de ir al bloque con etiqueta ETI 2. El segundo caso indica que la transacción continúa en el siguiente bloque cada vez que esto sea posible, y continúa en el bloque con etiqueta ETI 1 cuando no sea posible. Si no es posible continuar en ninguno de los bloques, entonces, la transacción espera hasta que sea posible continuar en alguno de los bloques, dándole preferencia al siguiente bloque. El último caso indica una transferencia incondicional al bloque con etiqueta ETI 1.

Utilizando estos últimos bloques, la codificación de este problema es como se muestra en la figura 7.5. Note que en esta última codificación se ha introducido otro estatuto de control llamado STORAGE. En el campo de localización de este estatuto se identifica al STORAGE y en el campo A del estatuto se indica la capacidad del STORAGE. De acuerdo a esta codificación, se obtienen los resultados que se muestran en la figura 7.6. En esta figura se puede observar que cuando hayan pasado estudiantes por el C.E.C., 290 estudiantes que encontraron ocupadas las terminales al llegar, ya no regresan. También, se puede apreciar que 280 estudiantes que encontraron al llegar a todas las terminales ocupadas, sí regresaron.

Por otra parte, la información que se obtiene con respecto al salón de terminales es la siguiente: El promedio de terminales ocupadas es 9.252, la cantidad de estudiantes que entraron al salón fue de 1009, el tiempo promedio de utilización de las terminales fue de 59.311 minutos, el porcentaje de utilización de las terminales es de 92.5%, actualmente existen 9 estudiantes trabajando en las terminales y la cantidad máxima de estudiantes que hubo en el salón fue de 10.

Por último, la figura 7.6 muestra el histograma de frecuencia del

	Generate	5, 3	Llegadas al C.E.C.
OTR	Transfer	Both., LLE	Ver si hay terminales disponibles
	Enter	1	Entrar al salón
	Advance	60, 20	Ocupar la terminal
	Leave	1	Salir del salón
	Tabulate	1	Tiempo en las terminales
	Terminate	1	Se incrementa contador
LLE	Transfer	.5,, REJ	50% regresa de nuevo
	Seize	1	Entra a la cafetería
	Advance	5	Compra golosinas
	Release	1	Sale de la cafetería
	Transfer	,OTR	Regresa al C.E.C.
REJ	Terminate		Abandonan el C.E.C.
1	Storage*	10	Salón con 10 terminales
1	Table	M1,40,5,50	Histograma del tiempo en el C.E.C.
	Start	10, NP	Elimina estado transiente
	Reset		
	Start	1000	Salida de 1000 clientes
	End		

FIGURA 7.5 Codificación del problema del centro electrónico de cálculo.

tiempo de permanencia en el salón de terminales. Como se puede observar en esta figura, el tiempo promedio que un estudiante se pasa en el salón de terminales es de 61.047 minutos.

Ejemplo 7.3. Peluquería con tiempo entre llegadas y tiempo de servicio exponenciales.

Para el ejemplo 7.1, suponga que el tiempo entre llegadas de los dos tipos de clientes es exponencial con media de 65 minutos para el primer tipo de cliente y con media de 75 minutos para el segundo tipo de cliente. También, en este mismo ejemplo piense que el tiempo de servicio es exponencial con media de 15 minutos para un corte de pelo y 7 minutos para una rasurada. Por último, imagine que en la peluquería solamente existen 3 lugares disponibles para esperar el servicio.

Para resolver este problema, es necesario introducir una serie de conceptos nuevos. Por ejemplo, se requiere conocer la forma de representar dentro del programa una función determinada, así como también la forma de simular el tiempo entre llegadas y el tiempo de servicio de una función exponencial.

En GPSS es posible incluir dentro del programa una serie de funciones. Cada función se define por medio de parejas coordinadas (X, Y)

* "1 STORAGE 10" indica que el salón de terminales al cual se le designó con el número 1, tiene una capacidad de 10 terminales (10 servidores).

Las funciones que se manejan pueden ser continuas o discretas. Sin embargo, puesto que en GPSS las funciones se definen por medio de parejas coordinadas, entonces, cada vez que se trabaja con funciones continuas, se interpola linealmente entre las parejas coordinadas correspondientes. Lo anterior significa que habrá menos error entre más cercanos estén los puntos que definen la función continua. Por otra parte, no existe ninguna restricción con respecto a los valores de X y Y , a excepción de X que en algunas ocasiones representa a un número uniforme entre 0 y 1. Finalmente, conviene señalar que si se referencia a un valor de X menor que el más pequeño de las parejas coordinadas que definen la función $((X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n))$, entonces, el valor de Y que resulta es Y_1 y si se referencia un valor de X mayor que X_n , entonces se obtiene para Y un valor de Y_n .

Para hacer referencia a una función, se utiliza el siguiente estatuto de control:

ETI FUNCTION A, B

donde:

- ETI = Etiqueta en el campo de localización que identifica a la función.
- Campo A = Variable independiente que puede ser un gobernador de números uniformes o cualquier otro parámetro
- Campo B = Identifica si la función que se está usando continua o discreta. Por ejemplo, C_n representa a una función continua que está definida por n parejas coordinadas y D_n representa a una función discreta que está definida por n parejas coordinadas.

Inmediatamente después del estatuto de control que hace referencia a la función, se alimentan las parejas coordinadas que se seleccionaron para definir la función. Estas parejas se alimentan a partir de la columna 1, separando con comas los valores de X de los valores de Y , y separando con una diagonal las parejas coordinadas. También, conviene aclarar que los valores alimentados de (X, Y) no se deben pasar de la columna 71.

Puesto que el ejemplo que está analizando, la distribución del tiempo entre llegadas y el tiempo de servicio es exponencial, entonces, de acuerdo a la teoría presentada en el capítulo 4 (Método de la transformada inversa), las parejas coordinadas que se deben alimentar corresponden a la siguiente función:

$$Y = -Ln(1 - R)$$

RELATIVE CLOCK 6468			ABSOLUTE CLOCK 6585		
BLOCK COUNTS					
BLOCK	CURRENT	TOTAL	Block	CURRENT	TOTAL
1	0	1290	8	0	569
2	0	1570	9	0	279
3	0	1001	10	0	279
4	0	1001	11	0	280
5	0	1000	12	0	280
6	0	1000	13	0	290
7	0	1000			

FACILITIES

Facility	Number entries	Average time/tran	Average utilization during		
			Total time	Unavail time	Current status
1	280	4.993	0.216		100.0

STORAGES

Storage	Capacity	Average content	Entries	Average utilization during		
				Average Time/unit time	Avail time	Current content
1	10	9.252	1009	59.311	.925	9

Average utilization during		
Average content	Current content	Maximum content
9	9	10

Figura 7.6 Resultados obtenidos de la simulación del C.E.C.

TABLAS

Tabla 1

Entries in table 1000		Mean argument 61.047		Standard deviation 12.343		Sum of arguments 61048.000	
Upper limit	Observed frequency	Percent of total	Cumulative percentage	Cumulative remainder	Multiple of mean	Deviation from mean	
40	21	2.09	2.0	97.8	0.655	-1.705	
45	104	10.39	12.4	87.5	0.737	-1.300	
50	144	14.39	26.8	73.1	0.819	-0.895	
55	111	11.09	37.9	62.0	0.900	-0.489	
60	104	10.39	48.3	51.0	0.982	-0.084	
65	107	10.69	59.0	40.9	1.064	0.320	
70	131	13.09	72.1	27.8	1.146	0.725	
75	120	11.99	84.1	15.8	1.228	1.130	
80	136	13.69	97.7	2.2	1.310	1.535	
85	15	1.49	99.2	0.7	1.392	1.940	
90	4	0.39	99.6	0.3	1.474	2.345	
95	3	0.29	100.0	0.0	1.556	2.750	

Remaining frequencies are all zero

Figura 7.6 Continuación.

es decir, el valor de X se obtiene por medio de un generador de números rectangulares* y el valor de Y de acuerdo a la expresión anterior. Conociendo la media de la distribución exponencial y el valor de Y , entonces, es posible obtener un valor simulado de esa distribución, esto es:

$$\begin{array}{l} \text{Valor simulado de} \\ \text{dist. exponencial} \end{array} = \left(\begin{array}{c} \text{Medida de dist.} \\ \text{exponencial} \end{array} \right) \left(Y \right)$$

este procedimiento es muy recomendable puesto que de esta forma distribuciones exponenciales con diferentes medias pueden ser simuladas dentro del mismo programa.

Por otra parte, para simular el tiempo entre llegadas y el tiempo de servicio, es necesario introducir una variante de los bloques GENERATE y ADVANCE. Para utilizar estos bloques con distribuciones exponenciales, el campo A de estos bloques representa la media de la distribución y el campo B hace referencia a la función $Ln(1-R)$. Conociendo la media y un valor simulado de Y , es posible obtener un valor simulado para el tiempo entre llegadas o el tiempo de servicio.

Utilizando estos últimos conceptos, la codificación de este problema es como se muestra en la figura 7.7. De acuerdo a esta codificación, se obtienen los resultados que se muestran en la figura 7.8. En esta figura se puede apreciar que por el sistema pasaron 537 clientes del primer tipo y 463 clientes del segundo tipo. También en esta figura se puede observar que hubo 63 clientes que no encontraron espacio disponible al momento de arribar a la peluquería. Por otra parte, el tiempo promedio de servicio (ver FACILITIES) sin tomar en cuenta al tipo de cliente es de 17.708 minutos. Además, en esta ocasión el servidor se encuentra ocupado en 49.5% de su tiempo disponible.

Con respecto a la sala de espera, la información obtenida es la siguiente: el contenido promedio es de 0.763, al salón de espera entraron 1000 clientes, el tiempo promedio de ocupación de un lugar de la sala de espera fue de 27.305 minutos, la ocupación promedio de un lugar de la sala de espera fue de 25.4% y la cantidad máxima de clientes que hubo en el salón de espera fue de 3.

Finalmente, en la figura 7.8 se muestran los histogramas de los tiempos de permanencia en la peluquería de los dos tipos de clientes. Como se puede observar en esta figura, el primer tipo de cliente permanece en el sistema un tiempo promedio de 24.247 minutos y el segundo tipo de cliente permanece 30.85 minutos.

*Un generador de números rectangulares se presenta como RN_n donde n indica el número del generador que se está utilizando.

1 Function RN1, C24

0.0,0.0/0.1,0.104/0.2,0.222/0.3,0.355/0.4,0.509/0.5,0.69

0.6,0.915/0.7,1.2/0.75,1.38/0.8,1.6/0.84,1.83/0.88,2.12

0.9,2.3/0.92,2.52/0.94,2.81/0.95,2.99/0.96,3.2/0.97,3.5

0.98,3.9/0.99,4.6/0.995,5.3/0.998,6.2/0.999,7/0.9997,8

	Generate	65, FN1	Llegadas del 1er. tipo de cliente
	Transfer	Both,,Away	Ver si hay lugar disponible
	Enter	1	Entra a la sala de espera
	Seize	1	Se inicia el servicio
	Advance	15, FN1	Corte de pelo
	Release	1	Se termina el servicio
	Tabulate	1	Tiempo en el sis. del 1er. tipo de cliente
	Leave	1	Deja un espacio en la sala de espera
	Terminate	1	Se incrementa contador
	Generate	75, FN1	Llegada del 2do. tipo de cliente
	Transfer	Both,, Away	Ver si hay lugar disponible
	Enter	1	Entra a la sala de espera
	Seize	1	Se inicia el servicio
	Advance	7, FN1	Servicio de rasura
	Advance	15, FN1	Corte de pelo
	Release	1	Se termina el servicio
	Tabulate	2	Tiempo en el sist. 2do. tipo de cliente
	Leave	1	Deja un espacio en la sala de espera
	Terminate	1	Se incrementa contador
Away	Terminate		Los que no encuentran lugar
1	Storage	3	Tres lugares para esperar
1	Table	M1,5,10,20	Histograma de frecuencias
2	Table	M1,5,10,20	Histograma de frecuencias
	Start	10, NP	Elimina estado transiente
	Reset		
	Start		
	End	1000	Salida de 1000 clientes

Figura 7.7 Codificación del ejemplo 7.3 al considerar tiempo entre llegadas y tiempo de servicio exponencial.

TABLAS

Tabla 1

Entries in table 537		Mean argument 24.247		Standard deviation 22.937		Sum of arguments 13021.000	
Upper limit	Observed frequency	Percent of total	Cumulative percentage	Cumulative remainder	Mutiple of mean	Deviation from mean	
5	114	21.22	21.2	78.7	0.206	-0.839	
15	131	24.39	45.6	54.3	0.618	-0.403	
25	94	17.50	63.1	36.8	1.031	0.032	
35	65	12.10	75.2	24.7	1.443	0.468	
45	45	8.37	83.6	16.3	1.855	0.904	
55	27	5.02	88.6	11.3	2.268	1.340	
65	29	5.40	94.0	5.9	2.680	1.776	
75	13	2.42	96.4	3.5	3.093	2.212	
85	6	1.11	97.5	2.4	3.505	2.648	
95	6	1.11	98.6	1.3	3.917	3.084	
105	3	0.55	99.2	0.7	4.330	3.520	
115	1	0.18	99.4	0.5	4.742	3.956	
125	2	0.37	99.8	0.1	5.155	4.392	
135	1	0.18	100.0	0.0	5.567	4.828	

Remaining frequencies are all zero

Figura 7.8 Continuación.

PROBLEMAS

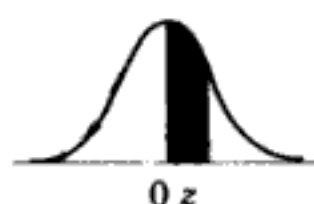
- 7.1. Cierta máquina produce partes a una razón de 1 pieza cada 15 minutos. El tiempo requerido para inspeccionar estas piezas sigue una distribución uniforme entre 8 y 14 minutos. El inspector que examina las piezas aproximadamente acepta el 90% de las piezas examinadas. Sin embargo, si una pieza es rechazada, ésta pasa a ser reprocesada. El tiempo de reproceso sigue una distribución uniforme entre 10 y 18 minutos. Si se simula la inspección de 1000 piezas, ¿Cuál es el tiempo promedio que una pieza permanece en el sistema?
- 7.2. Los estudiantes del ITESM llegan a la copiadora xerox de la biblioteca a una razón de uno cada 10 ± 15 minutos (distribución uniforme). El tiempo de servicio sigue una distribución uniforme entre 6 y 10 minutos. Sin embargo, si el estudiante encuentra en la cola más de 10 personas esperando, él decide dar un recorrido de 10 minutos por la biblioteca. Si se simula el servicio de 1000 estudiantes, determine el contenido promedio de la cola y el tiempo promedio de permanencia en el sistema.
- 7.3. Al comedor de estudiantes del ITESM llegan dos tipos de clientes. El tiempo entre llegadas de los dos tipos de clientes es uniforme entre 5 y 11 minutos. El 65% de las personas que llegan al comedor son profesores y el resto estudiantes. El tiempo requerido para servir los alimentos a un profesor es uniforme entre 2 y 6 minutos, y el tiempo de servicio para un estudiante es uniforme entre 3 y 7 minutos. Si la capacidad del comedor es de 200 personas, determine el tiempo promedio que cada tipo de cliente permanece en el sistema, al momento de que hayan pasado por el comedor 2000 personas.
- 7.4. Resolver utilizando GPSS el problema 5.12.
- 7.5. Resolver utilizando GPSS el problema 5.13.
- 7.6. Para el problema 5.14, determine mediante GPSS la cantidad de clientes que al llegar no encuentran lugar en el estacionamiento. Resuelva este mismo problema suponiendo que el estacionamiento tiene 4, 5, 7 y 8 lugares disponibles.
- 7.7. Resolver utilizando GPSS el problema 5.16.
- 7.8. Resolver utilizando GPSS el problema 5.17.
- 7.9. Los estudiantes del ITESM arriban a la biblioteca de acuerdo a una distribución exponencial con media de 3 minutos. La biblioteca consta de 4 pisos con capacidades de 250 para

APENDICES

Apéndice A. Números aleatorios uniformes.

03991	10461	93716	16894	98953	73231	39528	72484	82474	25593
38555	95554	32886	59780	09958	18065	81616	18711	53342	44276
17546	73704	92052	46215	15917	06253	07586	16120	82641	22820
32643	52861	95819	06831	19640	99413	90767	04235	13574	17200
69572	68777	39510	35905	85244	35159	40188	28193	29593	88627
24122	66591	27699	06494	03152	19121	34414	82157	86887	55087
61196	30231	92962	61773	22109	78508	63439	75363	44989	16822
30532	21704	10274	12202	94205	20380	67049	09070	93399	45547
03788	97599	75867	20717	82037	10268	79495	04146	52162	90286
48228	63379	85783	47619	87481	37220	91704	30552	04737	21031
88618	19161	41290	67312	71857	15957	48545	35247	18619	13674
71299	23853	05870	01119	92784	26340	75122	11724	74627	73707
27954	58909	82444	99005	04921	73701	92904	13141	32392	19763
80863	00514	20247	81759	45197	25332	68902	63742	78464	22501
33564	60780	48460	85558	15191	18782	94972	11598	62095	36787
90899	75754	60833	25983	01291	41349	19152	00023	12302	80783
78038	70267	43529	96318	38384	74761	36024	00867	76378	41605
55986	66485	88722	56736	66164	49431	94458	74284	05041	49807
87539	08823	94813	31900	54155	83436	54158	34243	46978	35482
16818	60311	74457	90561	72848	11834	75051	93029	47665	64382
34677	58300	74910	64345	19325	81549	60365	94653	35075	33949
45305	07521	61318	31855	14413	70951	83799	42402	56623	34442
59747	67277	76503	34513	39663	77544	32960	07405	36409	83232
16520	69676	11654	99893	02181	68161	19322	53845	57620	52606
68652	21376	92852	55866	88448	03584	11220	94747	07399	37408
79375	95220	01159	63267	10622	48391	31751	57260	68980	05339
33521	36665	55823	47641	86225	31704	88492	99382	14454	04504
59589	49067	66821	41575	49767	04037	30934	47744	07481	83828
20554	92409	96277	48257	50816	97616	22888	48893	27499	98748
59404	72059	43947	51080	43852	59693	78212	16993	45902	91386
42614	29297	01918	28316	25163	01889	70014	15021	68971	11403
34994	41374	70071	14736	65251	07629	37239	33295	18477	65622
99385	41600	11133	07586	36815	43625	18637	37509	14707	93997
66497	68646	78138	66559	64397	11692	05327	82162	83745	22567
48509	23929	27482	45476	04515	25624	95096	67946	16930	33361
15470	48355	88651	22596	83761	60873	43253	84145	20368	07126
20094	98977	74843	93413	14387	06345	80854	90279	41196	37480
73788	06533	28597	20405	51321	92246	80088	77074	66919	31678
60530	45128	74022	84617	72472	00008	80890	18002	35352	54131
44372	15486	65741	14014	05466	55306	93128	18464	79982	68416

Apéndice B. Distribución normal.



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2703	.2734	.2764	.2794	.2833	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3264	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3962	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.4990	.4990

Apéndice C. Distribución χ^2

χ^2_α	0.995	0.990	0.975	0.950	0.500	0.050	0.25	0.010	0.005
1	0.00+	0.00+	0.00+	0.00+	0.45	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	1.39	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	2.37	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	3.36	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	4.35	11.07	12.83	15.09	16.75
6	0.68	0.87	1.24	1.64	5.35	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	6.35	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	7.34	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	8.34	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	9.34	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	10.34	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	11.34	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	12.34	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	13.34	23.68	26.12	29.14	31.32
15	4.60	5.23	6.27	7.26	14.34	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	15.34	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	16.34	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	17.34	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	18.34	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	19.34	31.41	34.17	37.57	40.00
25	10.52	11.52	13.12	14.61	24.34	37.65	40.65	44.31	46.93
30	13.79	14.95	16.79	18.49	20.34	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	39.34	55.76	59.34	63.69	66.77
50	27.99	29.71	32.36	34.76	49.33	67.50	71.42	76.15	79.49
60	35.53	37.48	40.48	43.19	59.33	79.08	83.30	88.38	91.95
70	43.28	45.44	48.76	51.74	69.33	90.53	95.02	100.42	104.22
80	51.17	53.54	57.15	60.39	79.33	101.88	106.63	112.33	116.32
90	59.20	61.75	65.65	69.13	89.33	113.14	118.14	124.12	128.30
100	67.33	70.06	74.22	77.93	99.33	124.34	129.56	135.81	140.17

 ν = grados de libertad

Después del advenimiento de la computadora digital a principios de los años 50, surgieron elementos analíticos que han tenido un profundo impacto en el campo científico. Uno de dichos elementos es precisamente la simulación, cuyos usos y aplicaciones se han multiplicado en los últimos años.

Esta obra introduce al estudiante en los diferentes temas y técnicas de simulación de sistemas, y para que obtenga el máximo provecho de los conceptos y metodologías que se explican, se recomienda que tenga conocimientos de computación (lenguajes de propósito general como Fortran, Pascal, etc.) y de probabilidad y estadística.

El texto sintetiza las experiencias del autor y presenta opiniones obtenidas, tanto en el aula como en la práctica, del análisis y simulación de sistemas reales. Es un excelente libro con aplicaciones de simulación en áreas tales como Economía, Finanzas, Sistemas de Inventarios, Análisis y Evaluación de Inversiones y Sistemas de colas, entre otras.

e-mail: limusa@noriega.com.mx
www.noriega.com.mx

