

Capítulo 7

Simulación de Montecarlo

Capítulo 7

Técnica de Montecarlo

Objetivos del Capítulo:

- Conocer y entender el concepto de simulación tipo Montecarlo.
- Aplicar la simulación Montecarlo a ciertos casos prácticos.
- Conocer algunas, de las muchas aplicaciones de los números aleatorios y el significado de diferentes conceptos relacionados con el mismo.
- Resolver problemas en los que uno o más factores pueden representarse por una distribución de probabilidad.

7.1 Desarrollo Histórico

La técnica que se aplica para resolver los problemas teóricos citados puede considerarse precursora de la simulación en su uso presente. Esta recibe el nombre de Método de Montecarlo y está basada en la idea general de utilizar el muestreo para estimar un resultado deseado. El proceso de muestreo requiere que se describa el problema en estudio mediante una distribución de probabilidad adecuada de la cual se toman muestras.

El desarrollo sistemático de las Técnicas de Montecarlo se inició aproximadamente en 1943. Aunque anteriormente se habían realizado algunos intentos de aplicación semejantes a estas técnicas entre los que cabe destacar el experimento de Buffón (cálculo de π).

Históricamente, los conceptos de estas técnicas fueron planteados por John von Neumann y Stanislaw Ulam quienes propusieron un método en el cual se combinan las probabilidades de eventos separados para producir una solución aproximada del fenómeno en estudio. La idea surgió debido a los problemas que estaba causando la construcción de reactores nucleares, durante la segunda guerra mundial. A los diseñadores se les presentó el problema de la difusión de neutrones a través de diversos materiales y debido a la complejidad del mismo no se pudo dar una respuesta directa. La solución experimental resultaba muy costosa y requería mucho tiempo. Los investigadores tenían conocimiento de datos como la distancia promedio que recorre un neutrón de velocidad dada dentro de una sustancia antes de chocar con el núcleo. Conocía la probabilidad de que el neutrón fuera rechazado por el núcleo y la energía que probablemente perdería el neutrón después del choque. A pesar de tener toda esta información disponible resultó imposible plasmarla a través de un modelo matemático que permitiera analizar de manera completa el fenómeno.

Una de las primeras aplicaciones que tuvo la Técnica de Montecarlo después de la Segunda Guerra Mundial, fue la solución de integrales lineales.

La técnica matemática se ha conocido durante muchos años, pero fue revivida para el trabajo secreto en Los Álamos y se le asignó el nombre clave de Montecarlo. El Método Montecarlo fue tan exitoso que su popularidad se extendió a varios campos y el término casi se ha vuelto sinónimo de simulación en la mente de muchas personas.

Aunque nuestro interés principal en la Técnica de Muestreo Montecarlo es su utilidad en las situaciones probabilísticas de simulación, también puede usarse en ciertos problemas totalmente determinísticos que no pueden resolverse en forma analítica.

El desarrollo y proliferación de las computadoras permitió difundir esta técnica debido a que es posible aplicarla en diferentes ramas de la ciencia como la física nuclear (donde se aplicó inicialmente), en astrofísica y biología. También se aplica en la ingeniería, investigación de operaciones, en problemas tales como operaciones de teléfono, control de tráfico, inventarios, aeropuertos, etc.

7.2 Definición de la Técnica Montecarlo

Es un modelo de simulación, utilizado para analizar una decisión bajo incertidumbre, es decir, un problema en que el comportamiento de uno o más factores pueden representarse mediante una distribución de probabilidad, generando números al azar para los eventos de dicha distribución.

Su nombre fue asignado como clave de las investigaciones realizadas durante la guerra, tal como se dijo anteriormente, proviene de los posibles generadores de números al azar como lo son: Una moneda que se tira al aire, el lanzamiento de un dado, un corte de una baraja o una ruleta.

En la Técnica de Montecarlo, las experiencias artificiales o datos se generan usando algún generador de números aleatorios y la distribución de probabilidad. Los generadores de números aleatorios, pueden ser una tabla de dígitos aleatorios, una subrutina de computadora o alguna fuente de dígitos aleatorios uniformemente distribuida.

Montecarlo no desarrolla fórmulas para resolver un proceso, sino que experimenta las condiciones del problema en N corridas de prueba para ver lo que sucedería como resultado de las diferentes alternativas de decisión. El enfoque de Montecarlo, para la solución de un problema, es más bien el de estimar y no determinar con exactitud el valor de las variables del problema.

Los problemas que pueden resolverse por la técnica de Montecarlo se clasifican en 2 tipos: probabilísticos y determinísticos, lo cual depende de que el problema real esté asociado o no, a un punto proceso en el cual las variables tengan carácter probabilístico. Un ejemplo

probabilístico sería aquel que calcule un valor esperado por medio de una distribución de probabilidad y un ejemplo determinístico sería el cálculo del valor de un área bajo una curva.

Esta técnica, también conocida como *muestreo simulado*, permite introducir a un sistema datos que tienen las propiedades estadísticas de una distribución empírica o de otra clase.

La distribución de probabilidad para una muestra puede basarse en datos empíricos derivados de registros pasados, puede ser el resultado de un experimento reciente o puede ser una distribución teórica conocida **como** son: (Distribución Poisson, Normal, Gamma, entre otras).

7.3 Proceso de la Técnica Montecarlo

El concepto de la técnica es relativamente simple:

1. Tabule el dato de interés (número aleatorio), como una función de distribución de probabilidad acumulada con los valores de variables en el eje **X** (de la gráfica) y la probabilidad de 0 a 1 trazada en el eje **Y**.
2. Elija un número aleatorio entre 0 y 1 por medio de generadores de números aleatorios (puede usar una tabla).
3. Proyecte horizontalmente el punto correspondiente en el eje **Y** (para el número decimal aleatorio), hasta que la proyección de la línea interseque la curva acumulativa.
4. Proyecte la intersección hacia debajo para obtener la abscisa (**X**).
5. Anote el valor de **X** correspondiente a la intersección. Este valor de **X** se toma como un valor de la muestra.
6. Repita los pasos del 2 al 6 hasta que se genere la cantidad de variables aleatorias deseadas sin perder de vista la secuencia en que se dieron.

Para ilustrar la técnica anterior, se presenta el siguiente ejemplo:

Suponga que tenemos un sistema donde en cualquier período de 10 minutos llegan clientes a solicitar algún servicio. Las respectivas probabilidades se presentan a continuación:

Cliente	Probabilidad	Probabilidad Acumulada
0	0.40	0.40
1	0.25	0.65
2	0.20	0.85
3	0.15	1

Figura 7.1 Tabla de Probabilidades

Suponga que deseamos generar 5 periodos diferentes de tiempo. Primero graficamos la distribución de probabilidad acumulada.

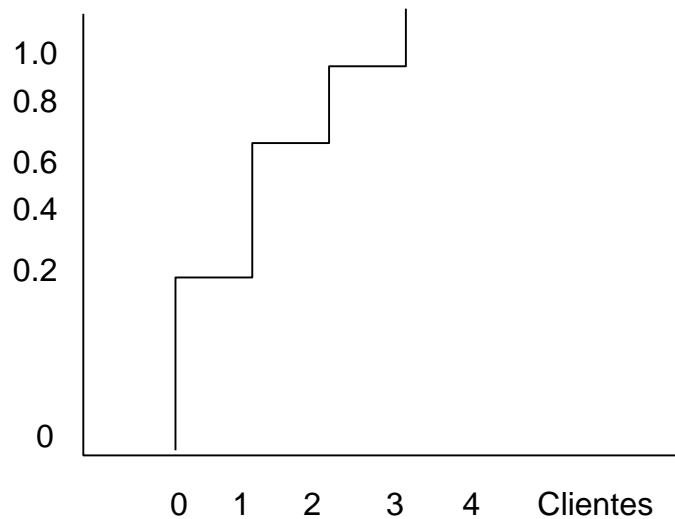


Figura 7.2 Distribución de Probabilidad Acumulada

Entonces de cualquier tabla de números aleatorios, tomando cinco números de dos dígitos y agregando un punto decimal delante de cada uno de ellos. Cada uno de estos valores se puede usar para generar el número de clientes que llegan durante un periodo en particular. Si los números aleatorios producidos fueran 09, 54, 42, 80 y 20 obtendríamos los siguientes:

Período	# Aleatorio	# de Clientes
1	0.09	0
2	0.54	1
3	0.42	1

4	0.80	2
5	0.20	0

Figura 7.3 Tabla de Números Aleatorios

Si el número aleatorio usado está en verdad distribuido uniformemente (o sea que cada valor del 00 al 99 tiene igual oportunidad de ser seleccionado) entonces cada número de los datos de interés puede llegar con la misma frecuencia relativa (en este proceso como en el mundo real). Así podemos generar una experiencia artificial (del número de clientes que llegan durante cada período de 10 minutos), que es típicamente lo que experimentamos con el sistema en el pasado.

El ejemplo anterior utiliza datos discretos. En el caso de datos continuos el proceso es el mismo con la única diferencia que la curva de la distribución acumulada es continua en vez de escalonada (Ver ejemplos al final del capítulo).

Puesto que la distribución normal es la que se encuentra con mayor frecuencia en los casos reales, estos procesos se llevan a cabo primero y se tabulan los resultados.

Para producir una variable distribuida normalmente de una población con parámetros (μ , σ) conocidos, se efectúa la siguiente fórmula:

$$\text{Variable aleatoria} = \mu (\# \text{ aleatorio}) * \sigma$$

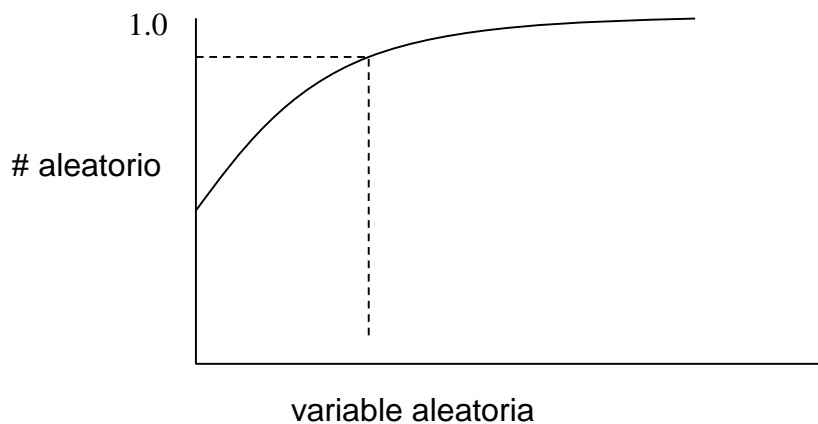


Figura 7.4 Método de Montecarlo para una Distribución Continua

Supóngase que deseamos producir una muestra de cinco valores aleatorios de una población distribuida normalmente con media = 100 y desviación estándar de 10. Tomamos 5 números aleatorios de la tabla, calculamos la ecuación $x = 100 (\# \text{ aleatorio}) * \sigma$.

Montecarlo para una distribución normal con $\mu = 100$ y $\sigma = 10$:

# Aleatorio	Cálculo	X
0.906	$100 + (0.906) * 10$	109.06
1.179	$100 + (1.179) * 10$	111.79
-1.501	$100 + (-1.501) * 10$	84.99
-0.690	$100 + (-0.690) * 10$	93.10
1.372	$100 + (1.372) * 10$	113.72

Figura 7.5 Cálculos

Los ejemplos anteriores muestran la generación manual de variables aleatorias. El uso de distribuciones teóricas es preferible sobre los datos empíricos por las siguientes razones:

1. El uso de datos empíricos no actualizados implica que todo lo que se está haciendo es simular el pasado.
2. Por lo general, resulta más eficiente usar una distribución de probabilidad teórica en términos de tiempo de computadora y de requerimientos de almacenamiento.
3. Es mucho más fácil cambiar los parámetros de un generador de distribuciones teóricas para la ejecución de pruebas de sensibilidad.

Por lo anterior, podemos concluir que un modelo será más útil si utilizamos distribuciones teóricas. Se sugiere que los datos empíricos se usen para probarlos con algunos de los métodos que permiten determinar si los mismos se ajustan a una distribución conocida para un nivel de confianza estadísticamente aceptable.

Si estos datos se ajustan a alguno de ellos entonces debe usarse la distribución teórica.

7.4 Otras Consideraciones de la Técnica

El diseño de un modelo que contiene elementos estocásticos o probabilísticos siempre da origen a la pregunta referente a si debemos usar datos empíricos directamente en el muestreo de Montecarlo o si debemos usar una de las distribuciones teóricas. Esta pregunta es muy importante y fundamental por tres razones:

Primera: El uso de datos empíricos implica que todo lo que uno está haciendo es simular el pasado. El uso de datos del año pasado duplicará en forma eficiente sólo el rendimiento de dicho año; los únicos eventos posibles son aquellos que ya ocurrieron. Una cosa es suponer que la forma básica de la distribución permanecerá inalterada a través del tiempo y otra muy diferente es suponer que la idiosincrasia de un período en particular se repetirá.

Segunda: Por lo general resulta más eficiente usar una distribución de probabilidad teórica en términos de tiempo de cómputo y de requerimientos de almacenamiento.

Tercera: Es mucho más fácil cambiar los parámetros de un generador de distribuciones teóricas para realizar las pruebas de sensibilidad o contestarse preguntas del tipo “que pasa sí”.

Como regla general, creemos que un modelo más útil tendría lugar, si pudiéramos usar las distribuciones teóricas. Por tanto, sugeriríamos que los datos empíricos se prueben mediante uno de los métodos (Pruebas Chi-cuadrada de la Bondad de Ajuste y Kolmogorov-Smirnov) para determinar si se ajustan a una distribución conocida con un nivel de confianza estadísticamente aceptado. Si lo hace, entonces debe usarse la distribución teórica.

7.5 Ejemplos de esta Técnica

7.5.1 Descompostura de una Máquina

Si se está estudiando la descompostura de una cierta máquina debido a una falla de los

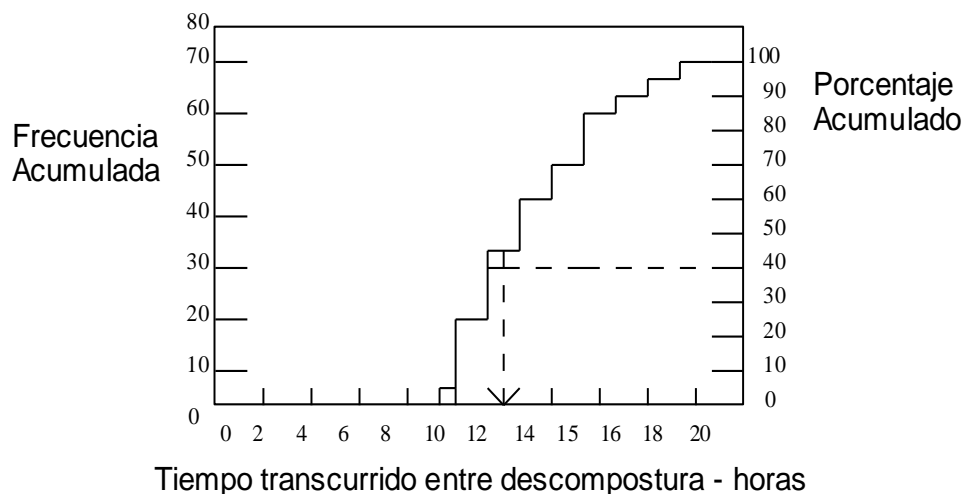


Figura 7.6 Distribución acumulada de los tiempos de descompostura

rodillos de una impresora. Los tiempos típicos de descompostura pueden simularse mediante muestreo simulado, a partir de la distribución de la vida útil de los rodillos. Se usará un ejemplo para ilustrar la forma en que se logra el muestreo simulado.

Supóngase que se trabaja con el mantenimiento de un equipo de 30 máquinas, y que se desea estimar el nivel de servicio que puede dar un mecánico. Se tienen ya, o deben obtener, los datos relativos a la frecuencia de descompostura de las máquinas y al tiempo de servicio para repararlas.

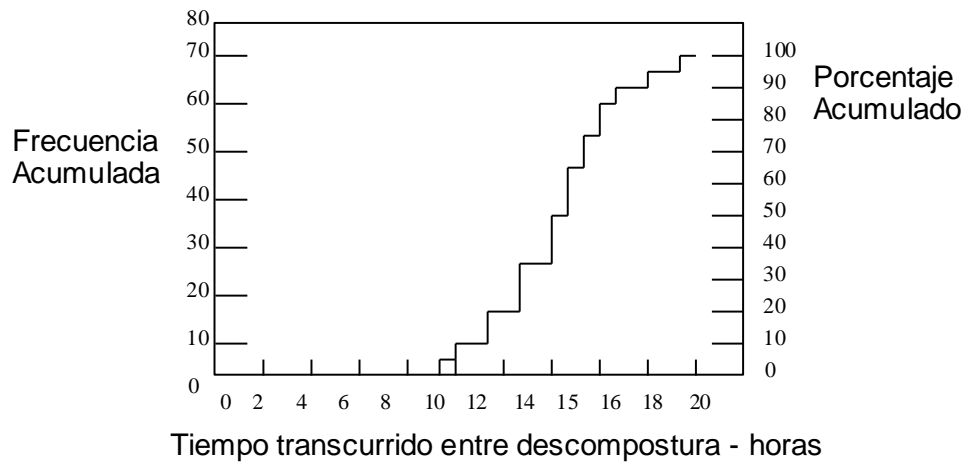


Figura 7.7 Distribución simulada de los tiempos de reparación

Si estas distribuciones siguieran ciertas funciones matemáticas, se podrían calcular directamente los resultados que aquí se determinarán mediante los métodos de simulación. El procedimiento que se aplica es el siguiente:

1. Se determinan las distribuciones de tiempo entre descompostura y tiempo de servicio. Si se dispusiera directamente de dichas distribuciones, se tendría que hacer un estudio para determinarlas; o bien, podrían tenerse registros de las descomposturas y de las reparaciones de las máquinas, a partir de los cuales se pudiesen construir tales distribuciones de las descomposturas y de los tiempos de reparación para 73 de éstos que serán la base de nuestra simulación.

Las distribuciones de frecuencia se convierten en distribuciones acumuladas de probabilidad (ver las figuras 7.6 y 7.7). Esta conversión se logra sumando las frecuencias que son menores o iguales a cada tiempo de descompostura o de reparación, y representando gráficamente los resultados. Las frecuencias acumuladas se convierten entonces a porcentajes, asignando el número 100 al valor máximo. Comenzando en el valor más bajo del tiempo de descompostura (10 horas), hay 4 ocurrencias. Se traza el valor cuatro en la gráfica acumulada para el tiempo de descompostura 10 horas. Para el tiempo de descompostura, 11 horas, hubo 10 ocurrencias; pero hubo 14 ocurrencias de 11 horas o menores, por lo que se representa el valor 14 para 11 horas. Para el tiempo de descompostura, 12 horas, se registraron 14 ocurrencias, pero hubo 28 ocurrencias de descomposturas para 12 horas o menos. Observando la figura 7.6 puede decirse que el 100% de los valores del tiempo de descompostura fueron de 19 horas o menores; el 99% fueron de 18 horas o menores, etc.

2. Se toman muestras al azar de las distribuciones acumuladas, a fin de determinar los tiempos específicos de descompostura y los tiempos de reparación necesarios al simular la operación de reparación. Esto se hace seleccionando números comprendidos

entre 0 y 100 al azar (que representa probabilidades expresadas en porcentaje). Los números pueden seleccionarse por cualquier proceso de selección al azar; por ejemplo, sacando papелitos numerados de una caja.

Los números al azar se usaron para entrar a las distribuciones acumuladas, a fin de obtener valores del tiempo en proporción a su ocurrencia en las distribuciones. En la figura 7.6 se presenta un ejemplo. El número 30, elegido al azar, se presenta para seleccionar un tiempo de descompostura de 12 horas. Puede verse ahora el propósito tras la conversión de la distribución original a una distribución acumulada. Ahora sólo puede haber un tiempo de descompostura asociado con un número al azar dado. En la distribución original resultarían dos veces a causa de la forma acompañada de la curva.

Usando números al azar para obtener valores de los tiempos de descompostura por este método (de la figura 7.6), se obtendrán valores de dichos tiempos en proporción a la probabilidad de ocurrencia indicada por la distribución original de frecuencia. De hecho, en este punto puede hacerse una tabla de números al azar, que seleccione ciertos tiempos.

Números al azar que se usaron para escoger los tiempos de descompostura y tiempos de reparación en proporción a las probabilidades de ocurrencia de las distribuciones originales			
Tiempos de descompostura		Tiempos de reparación	
1-5	10	1-3	8
6-19	11	4-7	9
20-38	12	8-18	10
39-60	13	19-40	11
61-77	14	41-59	12
78-85	15	60-75	13
86-90	16	76-86	14
91-95	17	87-93	15
96-99	18	94-97	16
100	19	98-99	17
		100	18

Figura 7.8 Tabla de tiempos de descomposturas versus tiempos de reparación

Tiempos de Descomposturas		Tiempos de Reparación	
Número al azar	Tiempo de descompostura tomado de la figura 7.6	Número al azar	Tiempo de reparación tomado de la figura 7.7

83	15	91	15
97	18	4	9
88	16	72	13
12	11	12	10
22	12	30	11
16	11	32	11
24	12	91	15
64	14	29	11
37	12	33	11
62	14	8	10
52	13	25	11
9	11	74	13
64	14	97	16
74	14	70	13
15	11	15	10
47	13	43	12
86	16	42	12
79	15	25	11
43	13	71	13
35	12	14	10

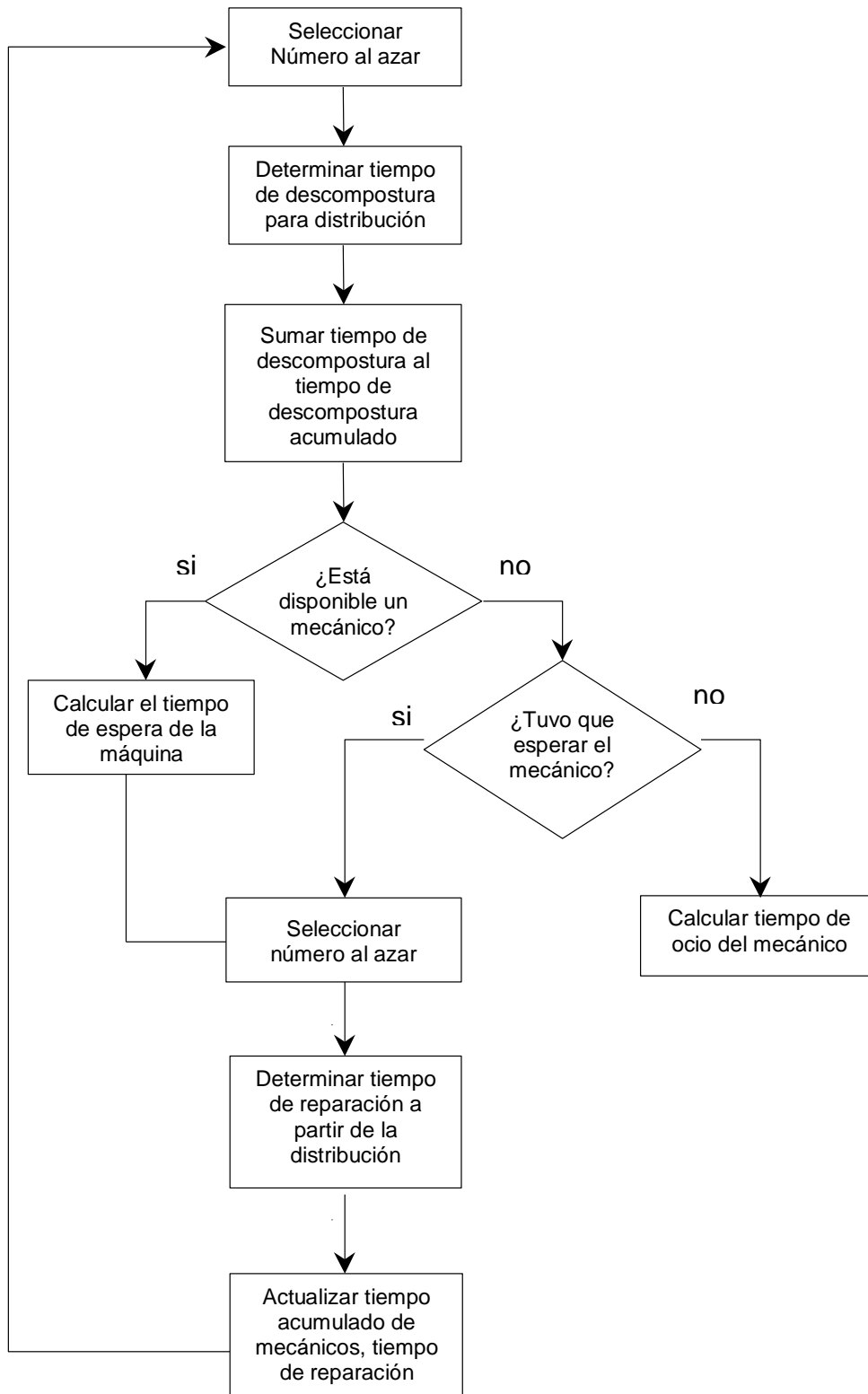


Figura 7.10 Algoritmo del Ejemplo Descompostura de una Máquina

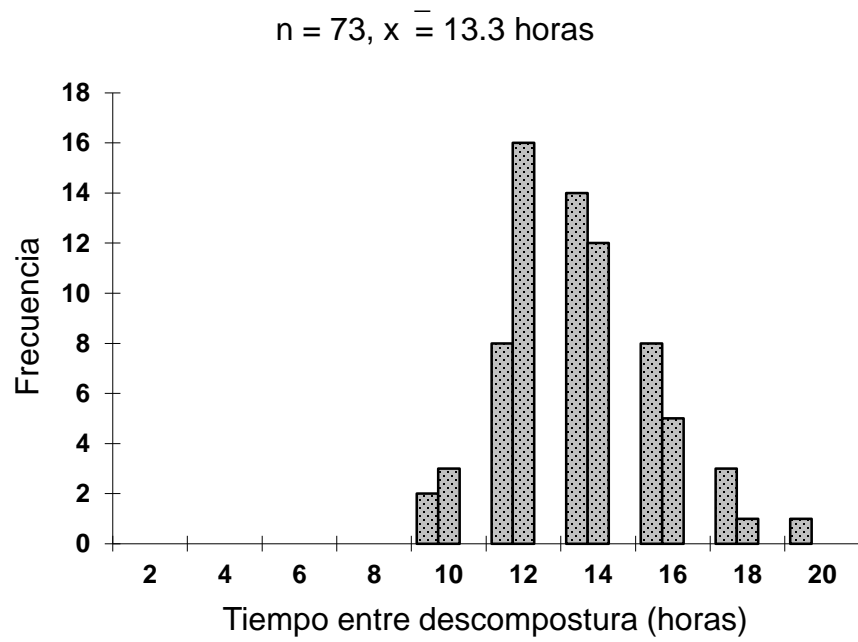


Figura 7.11 Distribución de Descomposturas

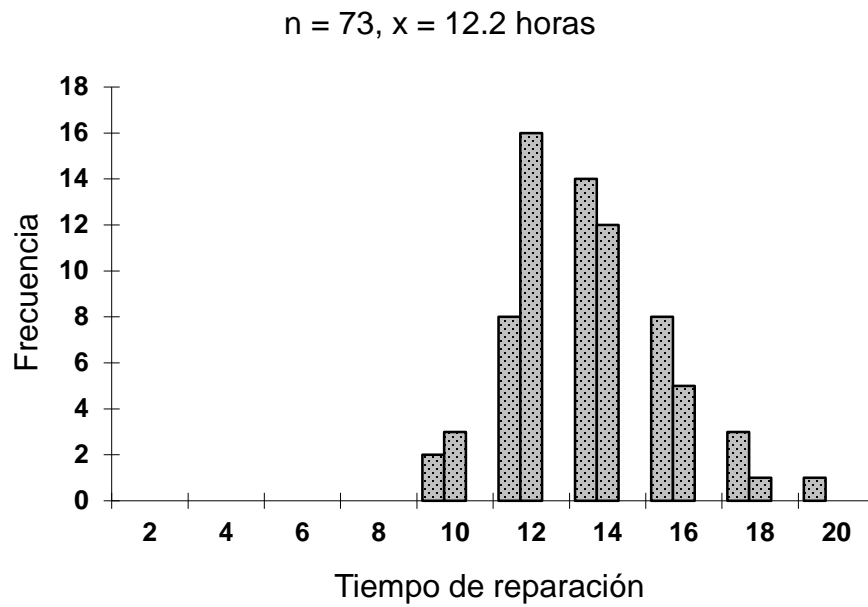


Figura 7.12 Distribución de tiempo de reparación

Tiempo de descompostura	Comienza el tiempo de reparación	Termina el tiempo de reparación	Tiempo de espera de la máquina	Tiempo ocioso del mecánico de reparación
0	0	15	0	0
18	18	27	0	3
34	34	47	0	7
45	45	57	2	0
57	57	68	0	0
68	68	79	0	0
80	80	95	0	1
94	95	106	1	0
106	106	117	0	0
120	120	130	0	3
133	133	144	0	3
144	144	157	0	0
158	158	174	0	1
172	174	187	2	0
183	187	197	4	0
196	197	209	1	0
212	212	224	0	3
227	227	238	0	3
240	240	253	0	2
252	253	263	1	0

Figura 7.13 Tabla de Tiempos Simulados de Descomposturas y de Reparación para 20 Descomposturas

Por ejemplo, leyendo en la figura 7.6, los números al azar 6 al 19 dan un tiempo de descompostura de 11 horas, etc. Esto equivale a decir que el 5% del tiempo se obtendría un valor de 10 horas, el 14% del tiempo se obtendría un tiempo de descompostura de 11 horas, etc. La figura 7.8 muestra la equivalencia de los números al azar para las figuras 7.6 y 7.7.

Tomando muestras ya sea de las distribuciones acumuladas de las figuras 7.6 y 7.7. De la figura 7.8 se obtendrán tiempos de descomposturas y tiempo de reparación en proporción a las distribuciones originales, igual que si estuvieran ocurriendo descomposturas y reparaciones reales. La figura 7.9 presenta una muestra de 20 tiempos de descompostura y de reparación determinadas de esta manera.

4. Se simula la operación real de descomposturas y reparaciones. El diagrama de flujo de la figura 7.6 muestra la estructura de lo que se desea hacer al simular la operación de reparación. Esta operación comprende la selección de un tiempo de descompostura y la determinación del hecho de que se disponga o no de un mecanismo. Si el mecánico no está disponible, la máquina debe esperar hasta que lo esté, y el tiempo de espera

puede calcularse fácilmente. Si el mecánico está disponible, hay que preguntarse ¿tuvo que esperar el mecánico? Si tuvo que esperar, se calcula el tiempo ocioso del mecánico. Si el mecánico no tuvo que esperar, se selecciona un tiempo de reparación y se procede de acuerdo con el diagrama de flujo, repitiendo el procedimiento general tantas veces como se desee, e incorporando un mecanismo para detener el procedimiento cuando se haya completado el número deseado de ciclos.

En la figura 7.13 se ilustra la simulación de la operación de reparación. En dicha tabla se han usado los tiempos de descomposturas y los tiempos de reparación seleccionados por números al azar de la Figura 7.9. Se supone que el tiempo comienza cuando se descompone la primera máquina y a partir de ese momento se acumula el tiempo de descompostura. El tiempo de reparación que tomó la primera descompostura fue de 15 horas, y como ésta es la primera ocurrencia en nuestro registro, no tuvieron que esperar ni la máquina del mecánico. La segunda descompostura ocurrió 3 horas después.

A fin de obtener el registro de la figura 7.13 se procede sumando y restando de acuerdo con las necesidades del modelo de simulación. El resumen que aparece en la parte inferior de la figura 7.13 indica que, para la muestra de 20 descomposturas, el tiempo total de espera de máquinas fue de 11 horas, y el tiempo ocioso total del mecánico fue de 26 horas. Para obtener una imagen realista, tendría que usarse una muestra mucho mayor. Utilizando los mismos datos sobre las distribuciones del tiempo de descompostura y del tiempo de reparación, el resultado de 1000 procesamientos usando un programa de computadora, fue de 15.9 % de tiempo de espera de máquinas y de 7.6% de tiempo ocioso del mecánico. Asimismo, debe suponerse que se paga al mecánico por jornada de ocho horas, cualquiera que sea la proporción de tiempo ocioso y de tiempo de servicio; sin embargo, conociendo el tiempo ocioso disponible, puede tenerse una guía para asignar al mecánico otros trabajos a desempeñar durante dicho tiempo. A fin de decidir si el servicio dado por un solo mecánico es lo suficientemente bueno o no, se tendrían que hacer ciertos cálculos relacionados con el costo del tiempo en el que no trabaja cada máquina, y probablemente corridas adicionales de simulación en las que interviniera la disponibilidad de más de un mecánico.

7.5.2 Inventario

En la descripción de un sistema por medio de un modelo, se encuentra en algunos casos, que el sistema es demasiado complicado para describirlo o que el modelo, una vez deducido, no permite una solución analítica. En estos casos, la simulación puede ser instrumento valioso para obtener la respuesta de un problema particular.

Los ejemplos que veremos a continuación nos describen sistemas de inventario por medio de modelos matemáticos y observaremos sus resultados con el propósito de analizar su comportamiento.

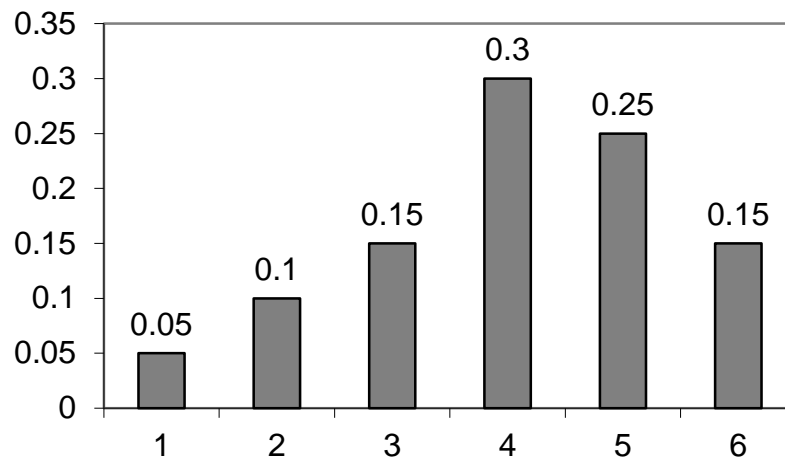


Figura 7.14 Gráfica de Distribución de la Demanda

Supongamos que la demanda diaria de un artículo particular puede expresarse mediante la distribución mostrada en la figura 7.14 y que se desea generar un patrón de demanda para diez días.

El primer paso es establecer la función de distribución acumulativa (FDA) mostrada en la figura 7.15. El siguiente paso es formar la tabla mostrada en la figura 7.16 y asignar números índice. Los números índice se asignan de manera que puedan reflejar la probabilidad de diversas demandas. El patrón de demanda para el período de 10 días mostrados en la figura 7.18 se obtiene con los números aleatorios de la figura 7.17.

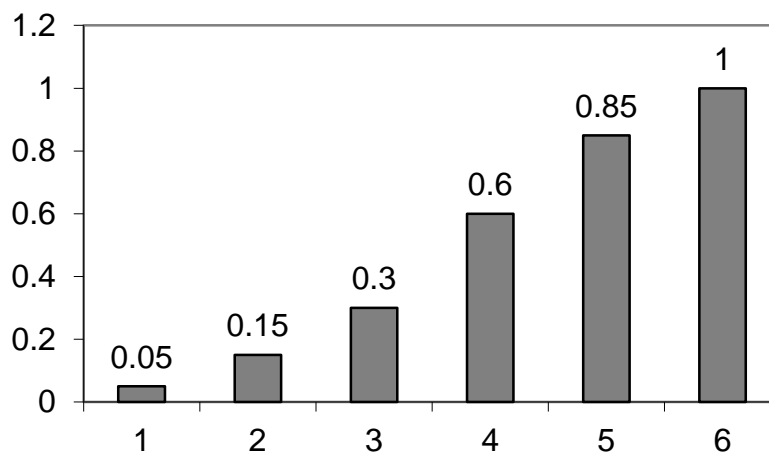


Figura 7.15 Gráfica de Distribución Acumulada de la Demanda

Demanda/Día	Números Índices
0	00-04
1	05-14
2	15-29
3	30-59
4	60-84
5	85-99

Figura 7.16 Tabla de Números Índice

14	74	24	87	07	45	26	66	26	94
----	----	----	----	----	----	----	----	----	----

Figura 7.17 Tabla de Números Aleatorios

La demanda diaria de este ejemplo se mostró como un valor entero que comienza en 0 y termina en 5, los valores intermedios están en progresión aritmética.

Los números aleatorios se obtuvieron a partir de una tabla de números aleatorios. Cuando se utiliza una Tabla de Números Aleatorios se puede comenzar en cualquier punto de la tabla y seguir en cualquier dirección.

En la figura 7.18 nos muestra el patrón de demanda para 10 días, de acuerdo con los números aleatorios y a los números índice.

Día	Demanda
1	1
2	4
3	2
4	5
5	1
6	3
7	2
8	4
9	2
10	5

Figura 7.18 Tabla Patrón de Demanda

7.5.3 Función de Efectividad

Es evidente que la simulación Montecarlo está orientada hacia el computador, ya que sin la rapidez del computador la mayoría de estos modelos no serían prácticos. Por medio de otro ejemplo podremos enfocar mejor la discusión de la simulación Montecarlo:

Suponer que la función de efectividad de algún sistema se representa por:

$$W = X + 4Y + 3Z$$

donde las variables X , Y y Z están descritas por las distribuciones de probabilidad mostradas en la figura 7.19. El problema consiste en determinar, mediante la simulación, la distribución de la función de efectividad. La figura 7.20 que nos muestra el número esperado de ocurrencias para los diferentes niveles de demanda en 10,000 ensayos.

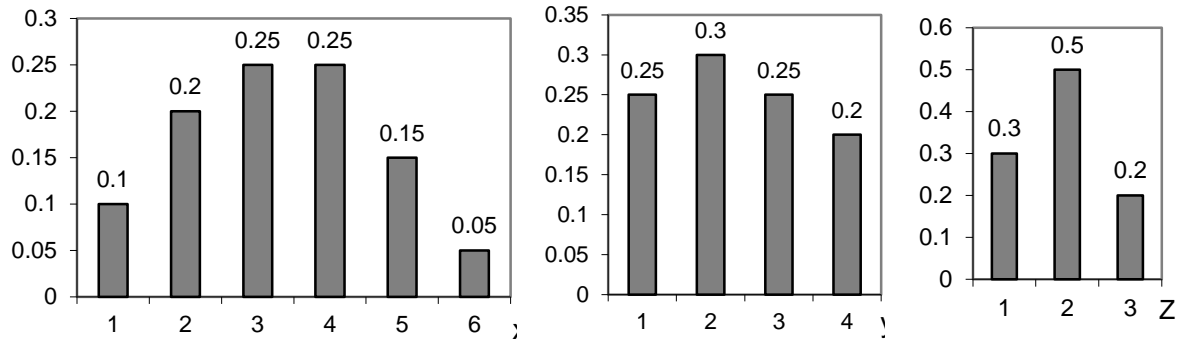


Figura 7.19 Distribución de las Variables de la Ecuación de Efectividad

Demanda/Semana	Frecuencia
0	500
1	1000
2	1500
3	3000
4	2500
5	1500

Figura 7.20 Tabla de Números Esperados de Ocurrencias para los diferentes Niveles de Demanda

Los resultados de la figura 7.21 se obtienen estableciendo las funciones de distribuciones acumuladas de cada variable y asignando números índices.

Valor de X	Nº índice	Valor de Y	Nº índice	Valor de Z	Nº índice
1	00-09	1	00-24	1	0-2
2	10-29	2	25-54	2	3-7
3	30-54	3	55-79	3	8-9
4	55-79	4	80-99		
5	80-94				
6	95-99				

Figura 7. 21 Tabla de Números Índices de las Variables de la Ecuación dada

Luego de tener los datos iniciales procedemos a la simulación de la ecuación, como lo muestra la figura 7.22:

Nº Prueba	Nº aleatorio para X	X	Nº aleatorio para Y	Y	Nº aleatorio para Z	Z	Valor de $W=X+4Y+3Z$
1	43	3	22	1	1	1	10
2	96	6	50	2	8	3	23
3	57	4	13	1	0	1	11
4	53	3	36	2	2	1	14
5	14	2	91	5	7	2	28
6	03	1	58	3	6	2	19
7	33	3	45	2	1	1	14
8	40	3	43	2	3	2	17

Figura 7.22 Tabla Simulación de la Ecuación

Lo siguiente es más sencillo, graficamos el resultado de la simulación, esto se aprecia analizando la figura 7.23, en la cual se calcula el valor w para cada prueba utilizando los valores de x , y y z obtenidos a través de muestreo aleatorio. Esto se hace para n experimentos.

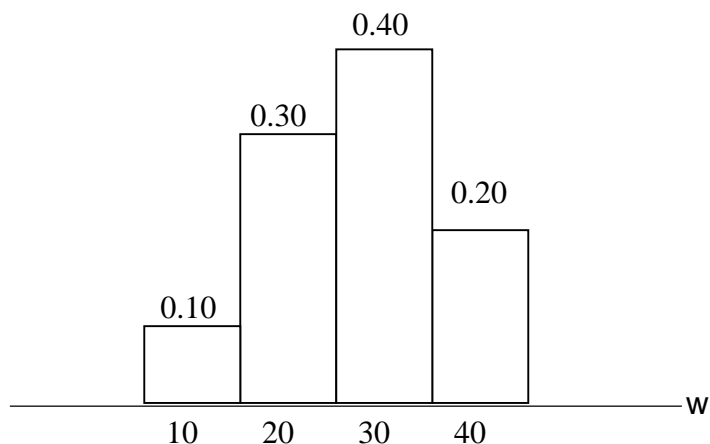


Figura 7.23 Gráfica de Distribución de w

Después de graficar los resultados de la figura 7.22 para los valores w , se puede calcular la media y la desviación típica de w , ya que esta gráfica muestra supuestamente la distribución de probabilidad de w . Además, con base en la distribución de w pueden hacerse afirmaciones probabilísticas. Por ejemplo, puede suponerse que después de muchos experimentos, una representación gráfica de la distribución de w da los resultados mostrados en la figura 7.23. Entonces pueden formularse enunciados tales como la

probabilidad de que el valor de w sea 20 es de 30% o que la probabilidad que el valor de w sea 20 o menos es 40%. Estos enunciados frecuentemente constituyen una ayuda en el proceso de toma de decisiones.

Es incorrecto utilizar el mismo conjunto de números aleatorios para cada variable. La suposición hecha en este ejemplo y en muchos otros problemas de simulación es que las variables son independientes. Por consiguiente, si se utiliza el mismo conjunto de números aleatorios para cada variable, esto puede dar lugar a un grado de correlación entre las variables, que estarían en contradicción con la suposición de independencia.

7.5.4 Cálculo del Área bajo una Curva (Valor Aproximado para la Integral)

Una de las aplicaciones de estas técnicas ha sido el cálculo de áreas y volúmenes. El procedimiento consiste en encajonar el área dentro de alguna figura geométrica conocida. Veamos el siguiente ejemplo:

Suponga que se quiere calcular el área localizada bajo la función:

$$f(x) = \text{sen} \frac{\pi x}{\pi x}$$

dentro de los límites de integración A y B , comprendidos entre -1 y 1 .

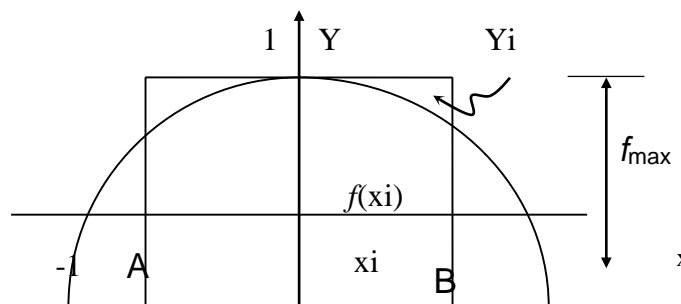


Figura 7.24 Cálculo del Área de la Función

$$f(x) = \text{sen} \frac{\pi x}{\pi x}$$

1. Primero se calcula un área conocida que contenga el área que se quiere integrar. Tomamos un rectángulo que tiene una base de longitud $B-A$ y una altura igual al valor máximo de la función el cual tiene un valor de 1 para $x = 0$.
2. Luego se puede considerar como un blanco con dos zonas sobre la cual se pueden lanzar dardos que se distribuyen de manera homogénea sobre toda el área.
3. Se lanza un dardo sobre el área y se observa como acierto en el caso de que así ocurra.
4. El proceso se repite N veces permitiéndonos aproximar el número n a la siguiente forma:

$$n = P_A * N, \text{ donde } P_A = \text{probabilidad de acertar} \quad (1)$$

Debido a la ley de los grandes números la ecuación anterior se puede expresar como una igualdad cuando N tiende al infinito.

Con base en la suposición hecha al principio de que la distribución sobre el área conocida es homogénea, podemos asociar la cantidad de aciertos (n), al área buscada y el número de lanzamientos o área rectangular de la siguiente manera:

$$AREA = P_A (\text{área rectangular}) \quad (2)$$

Si despejamos de la ecuación de la ecuación 1 la probabilidad y la sustituimos en la ecuación 2 (ecuación que relaciona las áreas), obtenemos:

$$AREA = (n/N) (\text{área rectangular})$$

El procedimiento anterior podemos desarrollarlo mediante la generación de números aleatorios.

1. Primero se genera un número aleatorio r_i en el rango de 0 a 1 que corresponderá a la abscisa y lo transformamos mediante la expresión:

$$x_i = r_i (B - A) + A$$

para que x pueda encontrarse dentro de los límites del área rectangular.

2. Al valor de x_{ile} corresponde un valor de $f(x_i)$ a partir de la función.

3. Se genera un segundo número aleatorio para obtener una ordenada. Esto resulta de la multiplicación del valor aleatorio por el valor de la máxima ordenada. (Note que por cada punto se generan dos números aleatorios).
4. Se compara el valor de $f(x_i)$ con el valor de Y_i . Si Y_i es menor o igual que $f(x_i)$, significa que el punto está dentro del área caso en el cual se cuenta un acierto, ya que el punto está por debajo del valor de la función. En caso contrario no se hace nada. El proceso se repite llevando la cuenta de los aciertos producidos.
5. El número de aciertos se divide entre el número de puntos generados para obtener el valor de **Pa**. Luego la probabilidad se multiplica por el área del rectángulo lo cual da como resultado el área buscada. El diagrama que a continuación presentamos permite resolver este problema:

A	B	Fmax	# impresiones	# Puntos simular
-1	1	1	5	20

Figura N° 25 Solución Manual

$$\text{La función} = f(x) = \text{sen} \frac{\pi x}{\pi x}$$

$$x_i = r_i(B - A) + A$$

$$\text{El área del rectángulo} = 2u^2$$

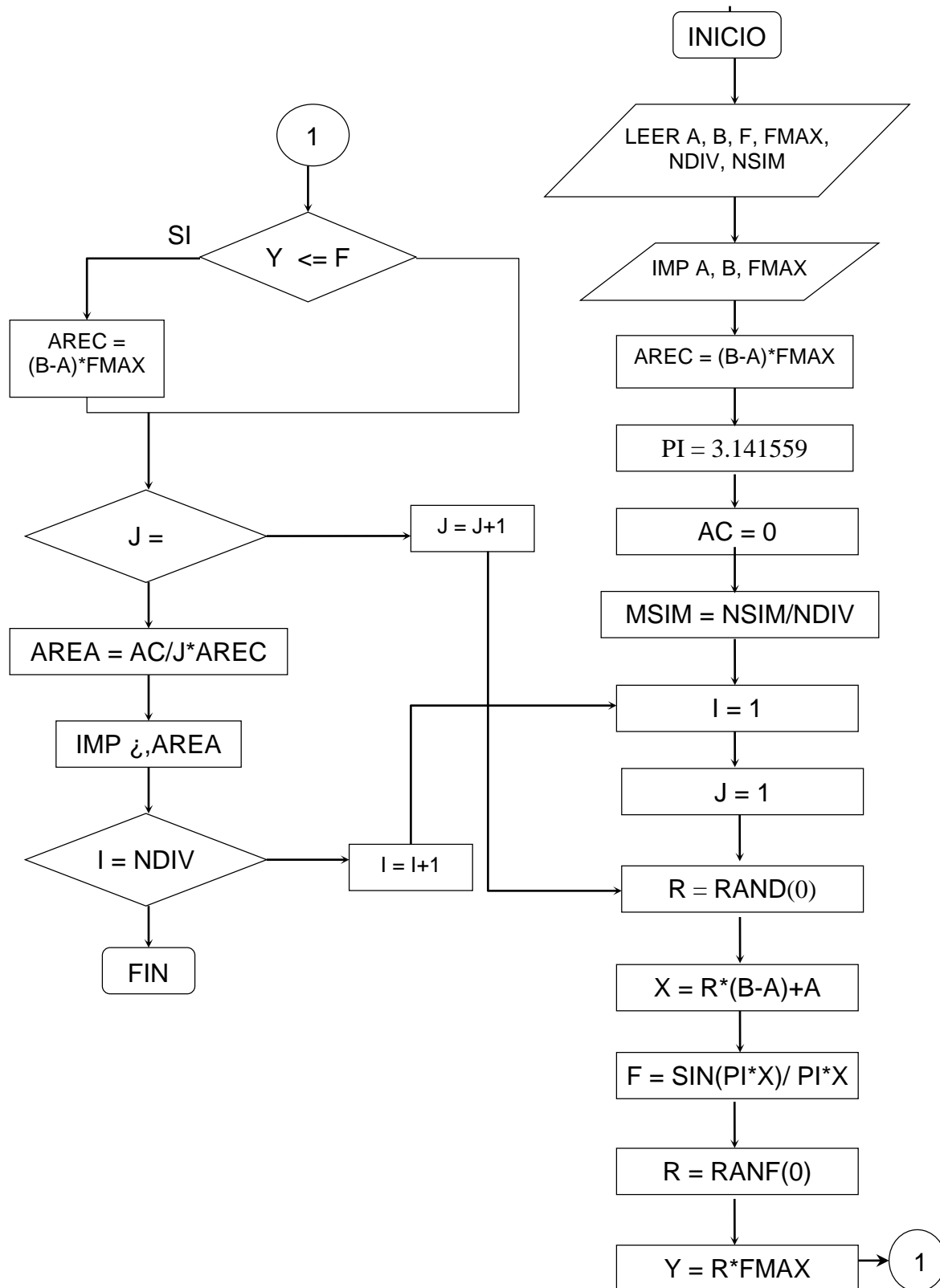
# aleatorio (r)	X_i	$f(x_i)$	# aleatorio (r_2)	Y_i $r_2(1)$	Aciertos	Área
.34	-.32	.8399	.76	.76	✓	$4/5 * (2) = 1.6$
.24	-.52	.6109	.37	.37	✓	
.23	-.54	.5848	.60	.60	X	
.38	-.24	.9079	.65	.65	✓	
.64	.28	.8759	.53	.53	✓	
.36	-.28	.8759	.70	.70	✓	$9/10 * (2) = 1.8$
.35	-.30	.8584	.14	.14	✓	
.68	.36	.8000	.18	.18	✓	
.90	.80	.2338	.18	.18	✓	
.35	-.30	.8759	.82	.82	✓	
.22	-.56	.5583	.58	.58	X	$10/15 * (2) = 1.33$
.58	0	0	.48	.48	X	
.13	-.74	.3136	.78	.78	X	
.36	-.28	.8759	.51	.51	✓	
.91	.82	.2080	.28	.28	X	
.58	.16	.9584	.74	.74	✓	
.45	.10	.9836	.74	.74	✓	
.43	-.14	.9681	.10	.10	✓	

.36	-.28	.8759	.03	.03	✓	15/20*(2) = 1.5
.46	-.08	.9895	.88	.88	✓	

Figura 7.26 Tabla de Cálculos

Como lo muestra la figura 7.26, los ganchitos (✓) muestran la cantidad de aciertos que se van dando, y cada cinco experimentos los colocamos en el área bajo la curva de la ecuación arriba mencionada; si continuamos experimentando tendiendo al infinito obtendremos el área de la función. Este tipo de situaciones es más fácil de simular con súper computadoras de alta velocidad.

Un diagrama de flujo (figura 7.27) y un programa en C mostrará la lógica de implementación de esta simulación en computadoras:



Programa en C para el Cálculo del Valor del Área bajo una Curva:

```

#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <time.h>

void main(void)
{
    int a=0, /* almacena el límite superior de la función */
        b=0, /* almacena el límite inferior de la función */
        fmax=0, /* almacena el valor máximo de la función */
        i=0, /* variable usada para los ciclos FOR */
        na=0, /* almacena el número de aciertos */
        línea=0; /* variable usada para el control de línea en impresión */
    float n, /* número de puntos a simular */
        r1=0, r2=0, /* almacenan los números aleatorios decimales */
        x=0, /* valor de  $x_i = r_1(B-A)+A$  */
        f=0, /* valor de la función  $f(x_i) = (\sin(\pi x))/(\pi x)$  */
        y=0, /* valor de  $y_i = r_2*f(\max)$  */
        pa=0, /* probabilidad de que los puntos quedaron bajo la curva */
        areabus=0; /* área bajo la curva/

    clrscr();
    gotoxy(18,2);printf("CALCULO DEL VALOR APROXIMADO DE LA INTEGRAL");
    gotoxy(16,4);printf("Introduzca el Límite Inferior de la Integral ==>");
    scanf("%d", &a);
    gotoxy(16,5);printf("Introduzca el Limite Superior de la Integral ==>");
    scanf("%d",&b);
    gotoxy(16,6);printf("Introduzca el Valor Máximo de la Función ==>");
    scanf("%d",&fmax);
    gotoxy(16,7);printf("Introduzca el Número de Puntos a Simular ==>");
    scanf("%f",&n);
    clrscr();
    gotoxy(18,2);printf("CALCULO DEL VALOR APROXIMADO DE LA INTEGRAL");
    gotoxy(16,4);printf("Límite Inferior de la Integral = %d",a);
    gotoxy(16,5);printf("Limite Superior de la Integral = %d",b);
    gotoxy(16,6);printf("Valor Máximo de la Función = %d",fmax);
    gotoxy(4, 7);printf("-----");
    gotoxy(4, 8);printf("                sen( x)");
    gotoxy(4, 9);printf("# ALEATORIO Xi=r1*(b-a)+a  f(Xi)=----- # ALEATORIO Yi=r2*f(max)");
    printf("ACIERTO");
    gotoxy(4,10);printf("    (r1)                x    (r2) ");
    gotoxy(4,11);printf("-----");
    randomize();
    for(i=1;i<=n;i++)
    {

```

```

r1 = random (100) / 100.0;
x = r1*(b-a)+a;
f = (sin(180*x))/(M_PI*x);
r2 = random (100)/ 100.0;
y = r2 * fmax;
linea+=1;
if(y<=f)
{na+=1;
 gotoxy(8,12+linea);printf("%2.2f  %2.2f  %2.2f  %2.2f  %2.2f  SI",r1,x,f,r2,y);
}
else
{
 gotoxy(8,12+linea);printf("%2.2f  %2.2f  %2.2f  %2.2f  %2.2f  NO",r1,x,f,r2,y);
}
}
if(linea>=11)
{
 getch();
 clrscr();
 gotoxy(18,2);printf("CALCULO DEL VALOR APROXIMADO DE LA INTEGRAL");
 gotoxy(16,4);printf("Límite Inferior de la Integral = %d",a);
 gotoxy(16,5);printf("Limite Superior de la Integral = %d",b);
 gotoxy(16,6);printf("Valor Máximo de la Función = %d",fmax);
 gotoxy(4, 7);printf("-----");
 gotoxy(4, 8);printf("          sen( x)");
 gotoxy(4, 9);printf("# ALEATORIO Xi=r1*(b-a)+a  f(Xi)=----- # ALEATORIO Yi=r2*f(max)");
 ACIERTO");
 gotoxy(4,10);printf("    (r1)          x    (r2) ");
 gotoxy(4,11);printf("-----");
 linea=0;}
}
pa=na/n;
areabus=pa*(b-a);
gotoxy(8, 14+linea);printf("VALOR APROXIMADO DE LA INTEGRAL = %2.4f",areabus);
 getch();
}

```

7.5.5. Cálculo de PI (π) por el Método de la Aguja de Buffón

Este ejemplo permite apreciar una vez más la ingeniosidad de la mente humana para solucionar los problemas, en este caso calcularemos el valor de la constante Pi (π) por medio de simulación utilizando el método de la Aguja de Buffón. Partimos del hecho de que p es la probabilidad de que una aguja de longitud lanzada al azar sobre una superficie plana que contiene un conjunto de líneas paralelas separadas entre sí una distancia D (con $D > L$), haga contacto con alguna de las líneas, entonces su valor está dado por la ecuación:

$$P = \frac{2L}{\pi D}$$

El caso particular de este problema se obtiene cuando en la ecuación se hace que $D = 2L$, quedando $p = 1/\pi$. De acuerdo con la teoría de probabilidades, si N es el número de veces que se lanza la aguja y n el número de aciertos, el cociente n/N tiende a la probabilidad de p cuando N tiende al infinito, de donde al sustituir en $p = 1/\pi$ se obtiene $N = n \cdot \pi$ que es la relación buscada.

Para simular este problema conviene encontrar una representación matemática (modelo) que nos permita encontrar este valor. La siguiente ilustración muestra la posición que puede tomar una aguja al caer sobre la superficie con múltiples líneas paralelas.

Se observa que la posibilidad de tocar alguna de las líneas se puede definir en términos de la distancia x , que existe entre el centro de la aguja y la línea más próxima, y en término del ángulo que forma la aguja y la línea.

Estas relaciones se pueden expresar de las siguientes formas:
Podemos observar que si:

$$x \leq \frac{L}{2} \sin \theta$$

La aguja toca la línea, y si:

$$x > \frac{L}{2} \sin \theta$$

la aguja no la toca.

De aquí se puede concluir que X y θ están limitados de la siguiente forma:

$$\begin{aligned} 0 &\leq x \leq L/2 \\ 0 &\leq \theta \leq 180^\circ \end{aligned}$$

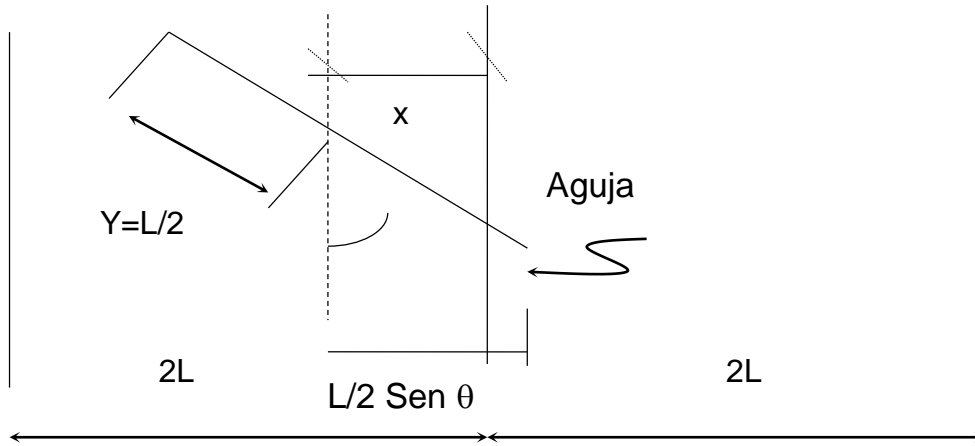


Figura 7.27 Aguja de Buffón

Una vez definida las relaciones de distancias y ángulos, proceden anunciar el procedimiento que debemos seguir para resolver este problema, estos son:

1. Generar un número aleatorio que represente la distancia del centro de la aguja a la línea. Luego multiplicarlo por $L/2$ para así obtener X .
2. Generar otro número aleatorio que represente el ángulo que forma la aguja y la línea; multiplicarlo por p para obtener θ .
3. Calcular el valor de Y .
4. Si el valor de X es menor o igual al de Y se cuenta como acierto.

El procedimiento se repite N veces y si n es el número de aciertos, la probabilidad p estará dada por n/N . Este valor permite calcular π despejando:

$$\pi = 1 / p \Rightarrow \pi = \frac{1}{\frac{n}{N}} \Rightarrow \pi = N / n$$

De esta forma calculamos el valor de π . Para observar el comportamiento en sí de la misma simulación, echemos un vistazo a la figura N° 28 que muestra una simulación manual para el cálculo de π .

Los valores iniciales son: N = número de lanzamientos = 25, L = longitud de la aguja = 3 cm, $L/2 = 1.5$

# aleatorio (r_1)	$X = L/2 * r_1$	# aleatorio(r_2)	$\theta = \pi * r_2$	$Y = L/2 * \text{sen}\theta$	Acierto $X \leq Y$
.80	1.2	.39	1.2252	1.4113	✓
.20	.3	.00	0	0	
.15	.225	.35	1.0996	1.3365	✓
.88	1.32	.04	.1257	.1881	
.98	1.47	.12	.3740	.1355	
.65	.975	.11	.3456	.5081	
.86	1.29	.23	.7226	.9920	
.73	1.095	.18	.5655	.8038	
.28	.42	.83	2.6075	.7636	✓
.60	.9	.35	1.0996	1.3365	✓
.60	.9	.50	1.5708	1.5	✓
.29	.435	.52	1.6336	1.4978	✓
.18	.27	.68	2.1363	1.2665	✓
.90	1.35	.29	.9111	1.1852	
.93	1.395	.23	.7225	.9919	
.73	1.095	.40	1.2566	1.4266	✓
.21	.399	.14	.4398	.6386	✓
.45	.675	.96	3.0159	.1880	✓
.76	1.14	.94	2.9531	.2817	✓
.96	1.44	.54	1.6965	1.4882	✓
.94	1.41	.37	1.1624	1.3766	
.53	.795	.42	1.3195	1.4529	✓
.57	.855	.22	.6912	.9562	✓
.96	1.44	.28	.8796	1.1557	
.43	.645	.07	.2199	.3272	

Figura 7.28 Tabla de Simulación Manual

El diagrama de flujo de la figura 7.29 mostrará bajo otra perspectiva la mecánica del problema:

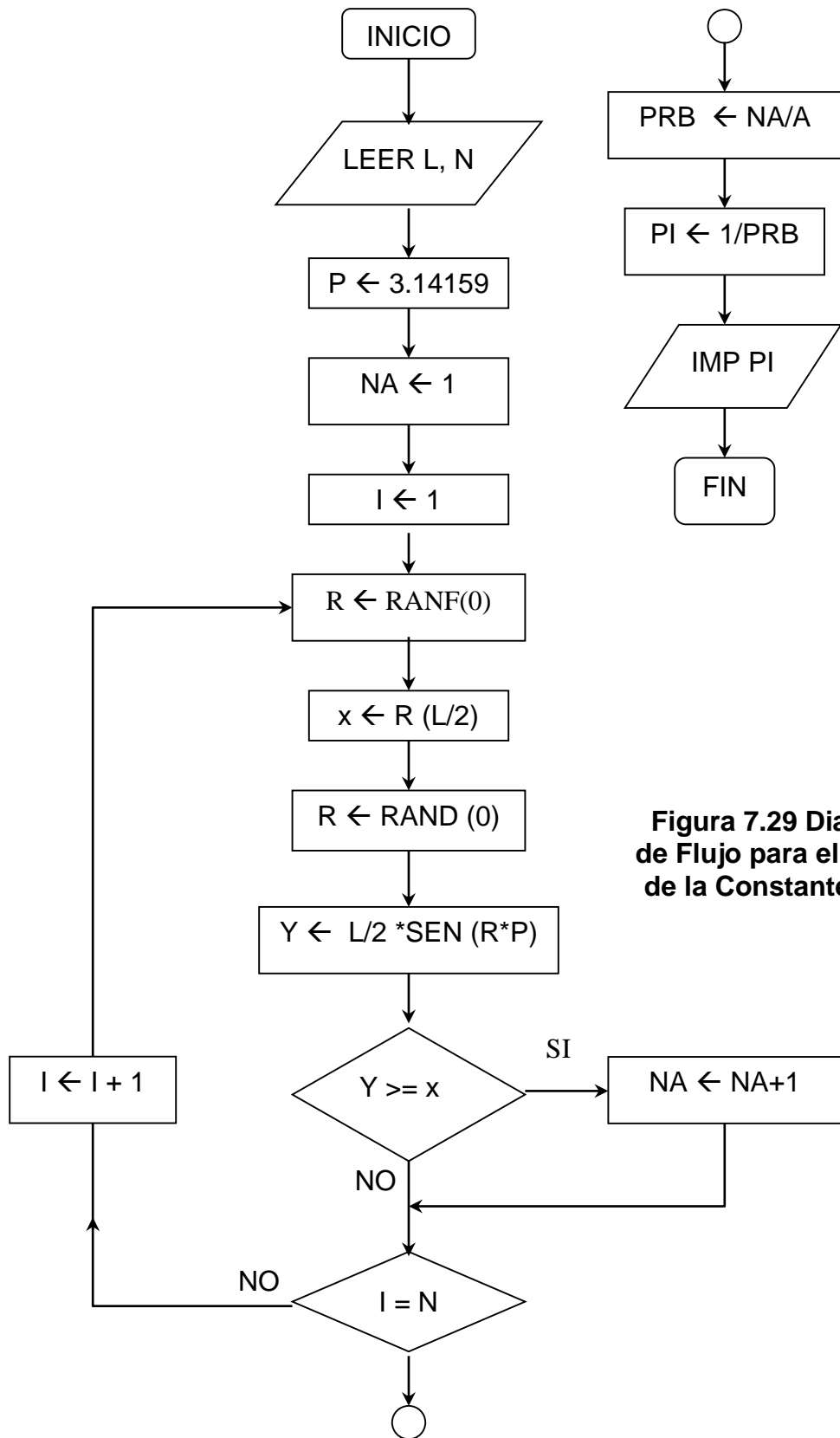


Figura 7.29 Diagrama de Flujo para el Cálculo de la Constante π (π)

Programa en C para el Cálculo de PI por la Aguja de Bufón:

```

#include <stdio.h>
#include <stdlib.h>
#include <conio.h>
#include <math.h>

void main (void)
{
    int n1, n2; /* usadas para el numero aleatorio entero */
    int i,      /* contador para los ciclos FOR */
        l,      /* longitud de la aguja */
        n,      /* números de lanzamientos */
        linea=0; /* usado para el control de las líneas en la impresión */
    float na=0.0, /*almacena el número de aciertos */
        r1, r2, /*almacenan los números aleatorios decimales */
        x,      /* */
        y,      /* */
        vpi; /* almacena el valor calculado de PI */
    clrscr();
    gotoxy(28,3);printf("CALCULO DEL VALOR DE PI");
    gotoxy(21,6);printf("Introduzca la Longitud de la Aguja (cms) ==> ");
    scanf("%d",&l);
    gotoxy(21,8);printf("Introduzca el Numero de Lanzamientos ==> ");
    scanf("%d",&n);
    gotoxy(28,3);printf("CALCULO DEL VALOR DE PI");
    clrscr();
    gotoxy(28,3);printf("CALCULO DEL VALOR DE PI");
    gotoxy(25,4);printf("Longitud de la Aguja (cms) ==> %d",l);
    gotoxy(25,5);printf("Numero de Lanzamientos ==> %d",n);
    gotoxy(2,7);printf("# ALEATORIO(r1)  x=r1(L/2) # ALEATORIO(r2)  Y=(L/2)SEN(180*r2)");
    printf("ACIERTOS");
    for(i=1;i<=n;i++)
    {
        n1=random(100);
        r1=n1/100.0;
        x=r1*(l/2.0);
        n2=random (100);
        r2=n2/100.0;
        y=(l/2.0)*sin(180*r2);
        linea+=1;
        if(x<=y)
        {
            na+=1;
            gotoxy(2,8+linea);printf("%2.2f  %2.2f  %2.2f  %2.2f  SI",r1,x,r2,y);
        }
        else

```



```
{ gotoxy(2,8+linea);printf("%2.2f    %2.2f    %2.2f    %2.2f    NO",r1,x,r2,y);
}
if(linea>15)
{getch();
clrscr();
gotoxy(28,3);printf("CALCULO DEL VALOR DE PI");
gotoxy(25,4);printf("Longitud de la Aguja (cms) ==> %d",l);
gotoxy(25,5);printf("Numero de Lanzamientos ==> %d",n);
gotoxy(2,7);printf("# ALEATORIO(r1)  x=r1(L/2) # ALEATORIO(r2)  Y=(L/2)SEN(180*r2)
ACIERTOS");
linea=0;
}
}
vpi=n/na;
gotoxy(22,10+linea);printf("VALOR APROXIMADO DE PI = %2.4f ", vpi);
getch();
}
```

7.5.6. Ejemplo 6: El Tambaleo del Borracho

A manera de ilustración consideremos el clásico tambaleo del borracho o de la caminata aleatoria (conocido también como el ebrio aleatorio).

Este tipo de simulación es utilizado en juegos de videos computarizados tales como PACMAN, TETRIS, DONKEY KONG, etc.; los cuales se basan en cálculos de posiciones de acuerdo con el comportamiento del jugador.

Supongamos que un borracho está parado en la esquina de una calle cuando decide caminar para salir de su embriagues.

Asuma que hay igual probabilidad de que vaya hacia el norte, sur, este u oeste. Si camina 10 cuadras. ¿Cuál es la probabilidad de que termine a dos cuadras de donde inició?

Utilizando el concepto de probabilidad acumulada:

Dirección	Probabilidad	Probabilidad Acumulada
Norte	.25	.25
Sur	.25	.50
Este	.25	.75
Oeste	.25	1.00

Figura 7.30 Tabla de Probabilidades Acumuladas

Designaremos una localización a cada esquina mediante coordenadas, donde X representa la dirección este-oeste y Y la dirección norte-sur. Cada vez que se mueva hacia el este incrementamos el valor de X y cada vez que se mueva hacia el oeste le restamos uno a X. Se realiza un proceso semejante con el movimiento norte-sur. Si le asignamos a la posición inicial las coordenadas (0,0) podremos establecer una relación entre cualquier punto en el que se encuentre y la posición inicial. Si al terminar el recorrido de 10 manzanas la suma de los valores X y Y es mayor que 2 entonces nuestro borracho terminó a una distancia mayor de dos cuadras del punto inicial. Podemos representar su movimiento en un plano cartesiano, o bien en una matriz.

Ya que hemos dicho que en cualquier esquina (incluyendo el inicio) existe una probabilidad igual de que se dirija en cualquier dirección, deseamos que sea una probabilidad de $\frac{1}{4}$ para que escoja cualquier dirección en particular. Por tanto, podrían sacar una serie de números

aleatorios de dos dígitos (uno para cada esquina) para decidir en qué dirección caminará enseguida, si el número obtenido va de 00 a 24 él se dirige hacia el este y sumamos 1 a x; si el número aleatorio va del 25 al 49, él camina hacia el este y sumamos 1 a y; si el número aleatorio va del 50 al 74, se dirige al norte y sumamos 1 a y; si el número aleatorio va del 75 al 99, se dirige al sur y restamos 1 de y.

El algoritmo para esta simulación es el siguiente:

1. Poner las coordenadas de la ubicación del borracho en (0, 0).
2. Obtener un número aleatorio.
3. Evaluar la dirección que toma el borracho.
4. Ajustar las coordenadas para ubicar la nueva posición del borracho.
5. Imprimir los valores de las coordenadas donde está el borracho.
6. Repetir los pasos del 2 al 5, N veces.
7. Se imprime la ubicación final del borracho.

La figura 7.33 muestra el diagrama de flujo lógico para nuestra simulación y figura 7.32 despliega los resultados de la simulación con cinco pruebas.

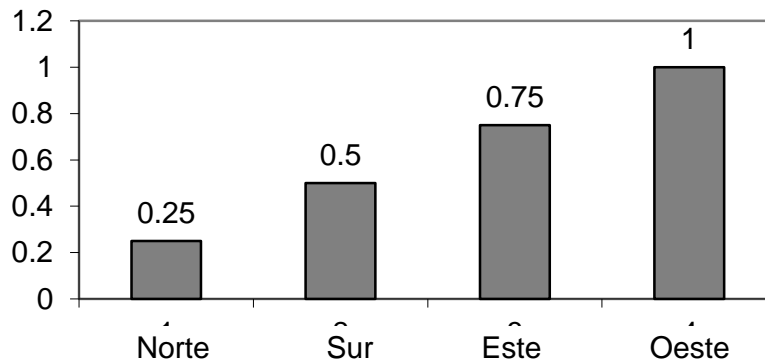


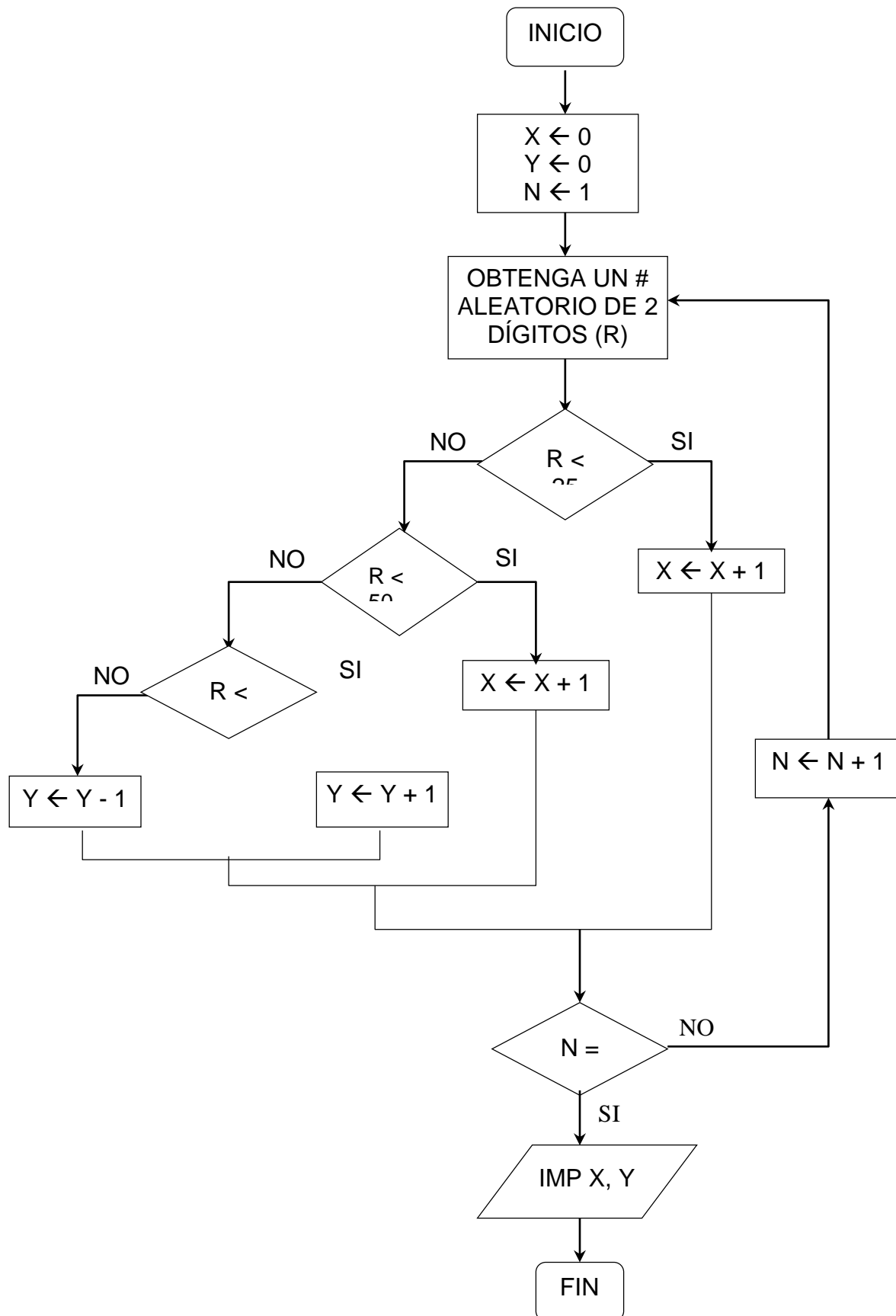
Figura 7.31 Distribución de Probabilidad Acumulada

Fin de la Esquina	Prueba 1		Prueba 2		Prueba 3		Prueba 4		Prueba 5	
	# Aleatorio	Loc.	# Aleatorio	Loc.	# Aleatorio	Loc.	# Aleatorio	Loc.	# Aleatorio	Loc.
1	73	0, 1	10	1, 0	05	1, 0	06	1, 0	64	0, 1
2	21	1, 1	89	1, -1	88	1, -1	95	1, -1	76	0, 0
3	45	0, 1	14	2, -1	10	2, -1	04	2, -1	79	0, -1
4	76	0, 0	81	2, -2	04	3, -1	67	2, 0	54	0, 0
5	96	0, -1	30	1, -2	48	2, -1	51	2, 1	28	-1, 0
6	94	0, -2	91	1, -3	19	3, -1	95	2, 0	05	0, 0
7	53	0, -1	06	2, -3	44	2, -1	73	2, 1	71	0, 1
8	57	0, 0	38	1, -3	21	3, -1	10	3, 1	75	0, 0

9	96	0, -1	79	1, -4	95	3, -2	76	3, 0	53	0, 1
10	43	-1, -1	43	0, -4	11	4, -2	30	2, 0	29	-1, 1
EXITO	SI		NO		NO		SI		SI	

Figura 7.32 Tabla de Resultados de la Simulación del Tambaleo del Borracho

Como se mencionó anteriormente esta simulación tiene una probabilidad uniforme, ya que de esta forma funciona el experimento.



Programa en C para el Tambaleo del Borracho

```

#include <stdio.h>
#include <stdlib.h>
#include <math.h>
#include <conio.h>

void main (void)
{
    int x,y,n;
    int r;

    clrscr();
    gotoxy(23,2);printf("SIMULACION DEL TAMBALEO DEL BORRACHO");
    gotoxy(10,5);printf("#    DE    CUADRAS    RECORRIDAS    #    ALEATORIO
LOCALIZACION(X,Y)");
    gotoxy(10,4);printf("-----");
    gotoxy(10,6);printf("-----");
    x=0;y=0;

    for(n=1;n<=10;n++)
    {
        r=random(100);
        if(r>=0 && r<25) y+=1;
        if(r>=25 && r<50) y-=1;
        if(r>=50 && r<75) x+=1;
        if(r>=75 && r<100) x-=1;
        gotoxy(15,7+n);printf("%d          %d          %d, %d",n,r,x,y);
    }
    if( abs(x+y)>=2 )
    {gotoxy(10,8+n);printf("EL BORRACHO TERMINO A 2 O MÁS CUADRAS DE DONDE
ESTABA INICIALMENTE ");}
    else
    {gotoxy(10,8+n);printf("EL BORRACHO TERMINO A MENOS DE 2 CUADRAS DE DONDE
ESTABA INICIALMENTE ");}
    getch();
}

```

7.5.7. Ejemplo 7: Estrategias de Servicio (Colas)¹

La técnica de Montecarlo también es utilizada para problemas de colas probabilísticas para evaluar diferentes estrategias de servicio. Por ejemplo, se supone que la distribución empírica mostrada en la figura 7.34 describe el tiempo entre llegadas a un canal simple de servicio y que el tiempo de servicio está distribuido exponencialmente. El problema consiste en determinar el tiempo medio de servicio de manera que el costo total del sistema sea mínimo, donde el costo total el sistema (CT) está dado por:

$$CT = \text{costo del tiempo de espera} + \text{costo del servicio}$$

El procedimiento de este problema consistirá en seleccionar diferentes tiempos medios de servicio y simular el sistema para cada valor seleccionado. El costo total de cada tiempo de servicio se puede calcular y evaluar.

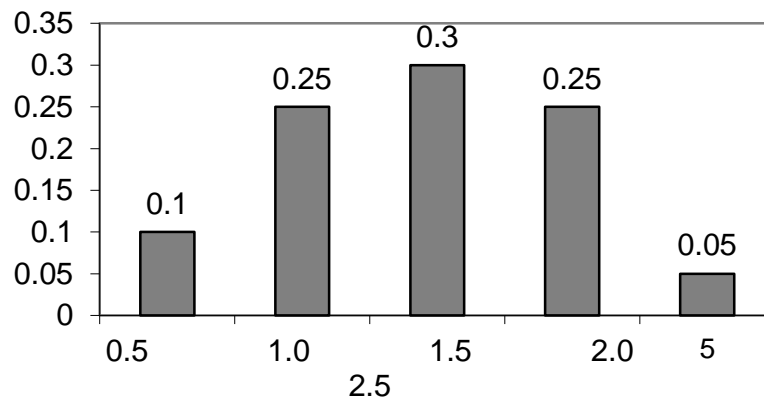


Figura 7.34 Distribución del Tiempo entre Arribos

En la figura 7.34 el tiempo medio entre llegadas es de 1.4 horas. Esto proporciona un punto de partida para seleccionar el tiempo medio de servicio, ya que este tiempo debe ser igual o menor que el tiempo medio entre llegadas. De otra manera, la cola crecería sin límites. Por tanto, se usa un tiempo medio de servicio de 1.2 horas en este ejemplo, y el costo del servicio se supone igual a B/.20.00 por hora.

De la figura 7.34 es posible sacar los números índices que la figura 7.35 muestra.

Tiempo entre Llegadas	Número Índice
0.5	00-14
1.0	15-39
1.5	40-69
2.0	70-94

¹ Del libro de Investigación de Operaciones de James Shamblin.

2.5	95-99
-----	-------

Figura 7.35 Tabla de Números Índices del Tiempo entre Llegadas

Ya que es posible considerar el tiempo de servicio como continuo, se usa la ecuación $x = -\theta \log r$ para simular los tiempos de servicio. Para simplificar se utilizan números aleatorios de un solo dígito en esta parte de la simulación, y los tiempos de servicio se aproximan a un decimal.

La figura 7.36 muestra los resultados de la simulación para ocho arribos. Como condición inicial, se supone que la primera llegada ocurre después de abrir la estación de servicio (1 hora en este caso).

Número de Llegadas	Número Aleatorio del Tiempo entre Arribo	Tiempo entre Llegadas	Número Aleatorio r del Tiempo de Servicio	Tiempo de Servicio $X = -1.2 \log r$
1	34	1.0	5	0.4
2	43	1.5	5	0.4
3	40	1.5	6	0.3
4	15	1.0	2	0.8
5	05	0.5	2	0.8
6	25	1.0	4	0.5
7	83	2.0	1	1.2
8	33	1.0	9	0.1
				4.5

Figura 7.36 Tabla de Simulación en Problemas de Cola

Procedemos a buscar lo que necesitamos, que es el costo total con la ecuación **CT = costo del tiempo de espera + costo del servicio**. El tiempo total de espera está dado por: **Tiempo total de espera = Tiempo de espera en el servicio + tiempo de espera en la cola**.

$$\text{Tiempo total de espera} = 4.5 + 0.3 + 0.2 = 5.1 \text{ horas}$$

El tiempo total para procesar los ocho arribos es de 8.8 horas. Por consiguiente, suponiendo que el costo de espera por unidad es de B/.5.00 por hora, el costo total de 8.8 horas de operación es de:

$$\text{CT} = 5 * (5.1) + 20 * (8.8) = \text{B/.201.50}$$

Entonces: **Costo promedio por hora = 201.50/8.8 = B/.22.90**

Obviamente, en un caso real se debe usar un tiempo mucho más prolongado y se debe ensayar otras condiciones diferentes de servicio. Es un ejemplo de colas bastante simple.

7.5.8 Ejemplo 8: Generando una Variable Aleatoria Exponencial²

En este ejemplo, trataremos de generar una variable aleatoria exponencial con $\lambda = 0.5$. Sea $Y = f(x)$. La figura 7.37 presenta la distribución de probabilidad de la variable x .

X	Frecuencia	Probabilidad Estimada	Probabilidad Teórica
0	7	0.035	0.03125
1	29	0.145	0.15625
2	61	0.305	0.31250
3	58	0.290	0.31250
4	37	0.185	0.15625
5	8	0.040	0.03125
total	200	1.000	1.00000

Figura 7.37 Tabla de Distribución de Probabilidad

Primeramente, generamos un número aleatorio entre 0.0 y 1.0 y hacemos e igual a ese número. A continuación, situamos este valor de Y en la figura presentada abajo y nos desplazamos horizontalmente por la gráfica, hasta interceptar la función de distribución acumulativa. En este punto, descendemos al eje x y registramos el valor de x como valor de la variable aleatoria generada.

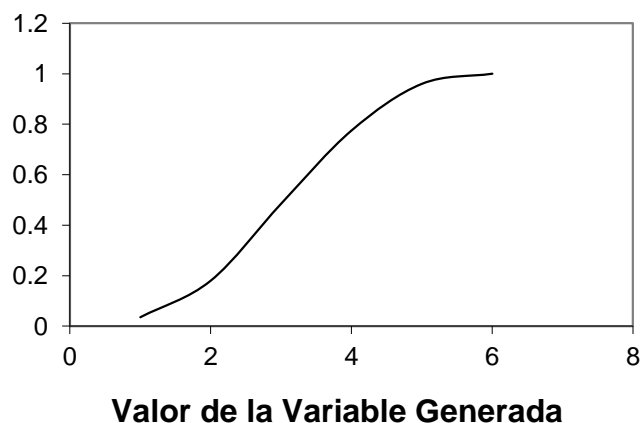


Figura 7.38 Función de Distribución de la Variable Aleatoria.

² Del libro "Análisis y Simulación de Sistemas Industriales" de J. W. Schmidt pp. 270-273

7.6. Identificación de la Muestra

Cuando trabajamos con simulaciones debemos determinar si algún elemento del sistema se comporta estocásticamente. Entonces surge el problema durante los estudios de simulación de la prueba de compatibilidad entre el conjunto de frecuencias observadas y alguna frecuencia teórica. Responderíamos a la pregunta ¿Pueden los datos observados tener una distribución de probabilidad específica? Si la frecuencia de los datos observados se compara con la frecuencia teórica, entonces la distribución teórica en nuestro modelo nos sirve para representar la población estudiada. Usualmente no podemos crear una hipótesis razonable acerca de la distribución de una variable aleatoria hasta que hayamos seleccionado y analizados los datos, (ya sean históricos o experimentales). La colección de datos se sumaria usualmente en una distribución de frecuencia, la cual presentamos en la figura N° 7.39:

Producción Semanal (X)	Frecuencia	P(X)
Menos de 46	1	0.008
46-55	1	0.008
56-65	3	0.025
66-75	7	0.058
76-85	11	0.092
86-95	21	0.175
96-105	28	0.234
106-115	16	0.134
116-125	22	0.183
126-135	7	0.058
136-145	1	0.008
146 o más	2	0.017
	120	1.000

Figura 7.39 Tabla de Producción Semanal (Valores Continuos)

Si estamos tratando con variables discretas, registramos la frecuencia de cada valor individual ocurrido. Si la variable es continua interrumpimos los rangos de valores en intervalos iguales o clases. El rango del intervalo.

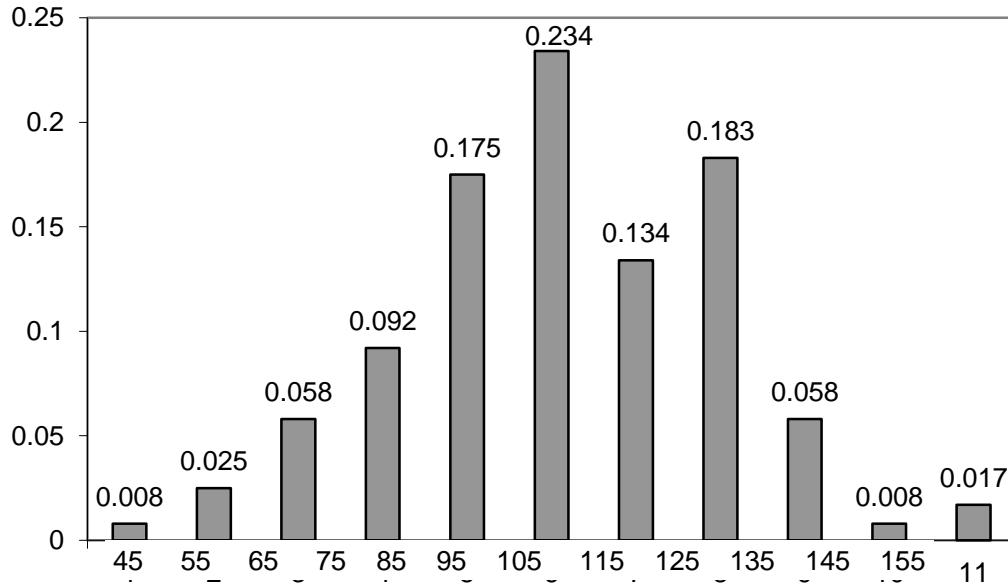


Figura 7.40: Distribución de Frecuencias de los Datos Tabulados (Datos Continuos)

Se toma usualmente entre 5 y 20 dependiendo de los datos. La frecuencia relativa de cada intervalo es la frecuencia observada en ese intervalo entre el número total de datos. Las figuras 7.39 y 7.40, antes presentadas, están basadas en datos continuos, mientras que la figura 7.41 y figura 7.42, que a continuación presentamos, están basados en datos discretos.

Después que el analista ha obtenido una distribución de frecuencia relativa, se selecciona una posible distribución de probabilidad. Esta selección se deriva de la experiencia y buen juicio del analista. Una ayuda es la comparación visual de la distribución de frecuencia observadas con las diferentes distribuciones teóricas observadas.

La comparación visual de las observaciones sólo sirve para sugerir con que la distribución podemos probarla. Esta sugerencia nunca es una justificación suficiente para aceptar como verdadera dicha distribución teórica.

N (Número de Llamadas)	Número de Intervalos de 1 Hora con N Llamadas	Frecuencia Relativa
0	315	0.619
1	142	0.279
2	40	0.078
3	9	0.018
4	2	0.004
5	1	0.002
	509	1.000

Figura 7.41 Distribución de Frecuencia de Consultas Telefónicas por Intervalos de una Hora (Datos Discretos)

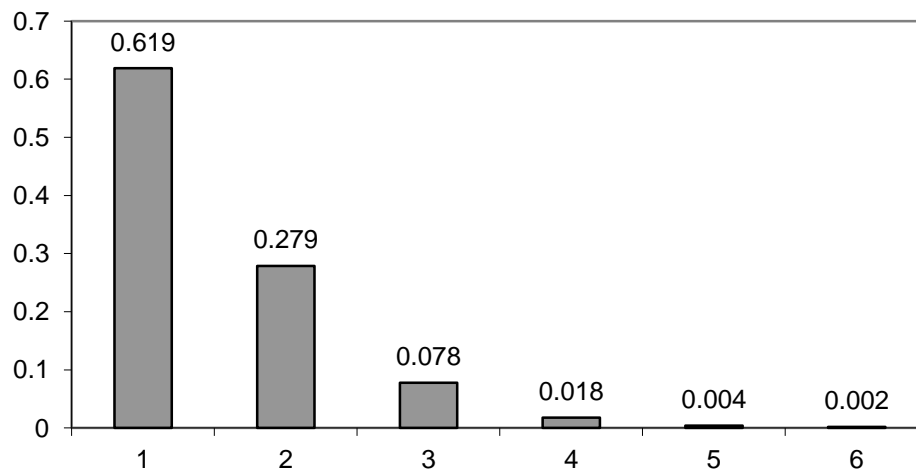
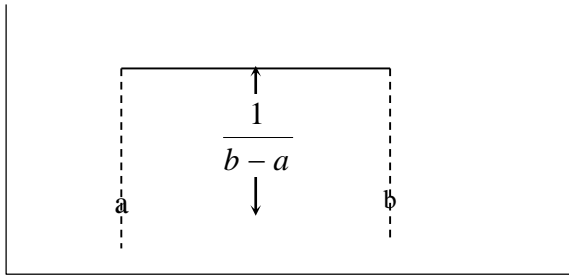
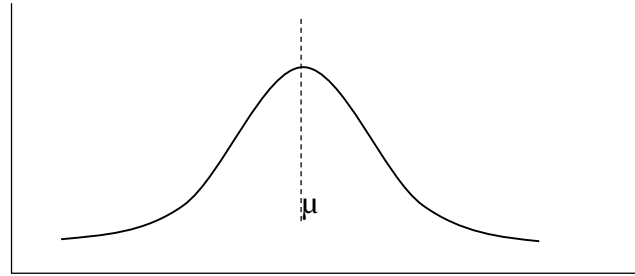
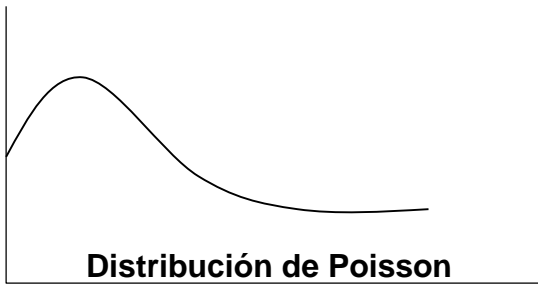
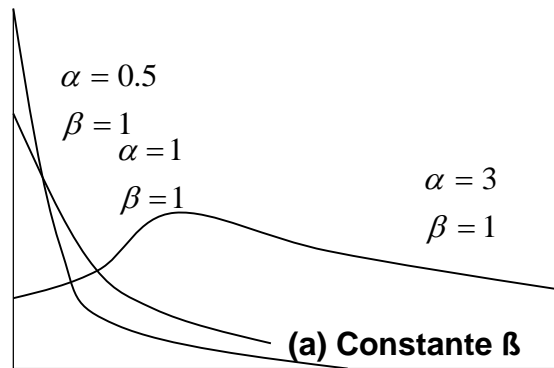
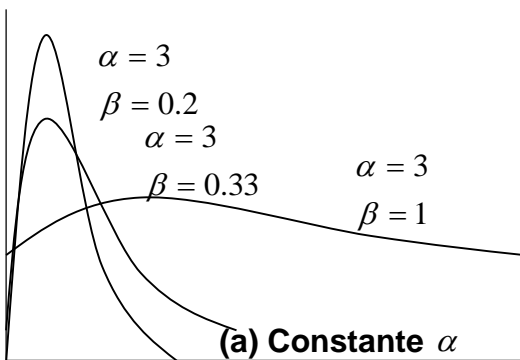


Figura 7.42 Distribución de Frecuencias de los Datos Tabulados (Datos discretos)

**Distribución Uniforme****Distribución Normal****Distribución de Poisson****Distribución Exponencial****Distribución Gamma****Figura 7.43 Distribuciones Teóricas**

Una vez que el analista ha identificado una o más distribuciones teóricas, tales como la Normal, Poisson, Gamma, etc. que él crea que se asemejan a sus datos empíricos, está obligado a determinar los parámetros de la distribución para proceder con la prueba estadística. Cuando la distribución hipotética es una función de dos parámetros, puede estimar estos parámetros de la media y la varianza de la muestra.

Para datos agrupados podemos calcular la media y la varianza como:

$$Media = \bar{x} = \frac{\sum_{i=1}^k M_i F_i}{n}$$

$$Varianza = s^2 = \frac{\sum_{i=1}^k M_i^2 F_i - n\bar{x}^2}{n-1}$$

$n = \text{Tamaño de la muestra} = \sum_{i=1}^k F_i$

$k = \text{Número de clases o intervalos}$

M_i = punto medio del i -ésimo intervalo o de la i -ésima clase si se trata de datos discretos.

Los cálculos requeridos para los datos discretos presentados anteriormente se presentan en la figura N° 44:

Mi	Fi	Mi Fi	Mi ² Fi
0	315	0	0
1	142	142	142
2	40	80	160
3	9	27	81
4	2	8	32
5	1	5	45
	509	262	440

Figura 7.44 Tabla de Cálculos para Datos Discretos

$$\bar{x} = \frac{262}{509} = 0.5147$$

$$s^2 = \frac{400 - 509(0.5147)^2}{509 - 1} = 0.6007$$

Los cálculos requeridos para los datos continuos presentados anteriormente se presentan en la figura 7.45:

M_i	F_i	M_i F_i	M_i² F_i
40.5	1	40.5	1640.25
50.5	1	50.5	2550.25
60.5	3	181.5	10,980.75
70.5	7	493.5	34,791.75
80.5	11	885.5	71,282.75
90.5	21	1,900.5	171,995.25
100.5	28	2,814.0	282,807.00
110.5	16	1,768.0	195,364.00
120.5	22	2,651.0	319,445.50
130.5	7	913.5	119,211.75
140.5	1	140.5	19,740.25
150.5	2	301.0	45,300.50
	120	12,140.0	1,275,110.00

Figura 7.45 Tabla de Cálculos para Datos Continuos

$$\bar{x} = \frac{12140}{120} = 101.17$$

$$s^2 = \frac{1275110 - 120(101.17)^2}{120 - 1} = 393.83$$

$$s = \sqrt{393.83} = 19.86$$

Una vez calculados los parámetros de cada uno de los tipos de datos (discretos y continuos) procedemos a establecer los resultados de nuestras suposiciones.

Originalmente inferimos que los datos de la figura 7.42 se podrían haber originado de una distribución de Poisson. Veremos; sabemos con base estadística, que, para una distribución de Poisson, la media = λ y también que la varianza = λ ; es decir, la media y la varianza son iguales. Como observamos en la figura 7.41, la media y la varianza no son iguales para los datos observados (es decir, $0.5147 < 0.6007$), lo cual podría llevarnos a rechazar la hipótesis de una distribución de Poisson. Sin embargo, en este caso en particular tenemos razones teóricas y prácticas para sospechar que es de tipo Poisson. Cuando es factible que ese evento tenga lugar en un intervalo como cualquier otro, y la ocurrencia de un evento no tenga efecto en lo referente a que otro ocurra, definitivamente debemos pensar que es una distribución de Poisson. Además, si existe una alta posibilidad de cero ocurrencias en cualquier intervalo y la media o número promedio de ocurrencias por intervalo es pequeño, tenemos una razón más para esperar que sea tipo Poisson. En nuestro ejemplo, datos discretos de las figuras 7.41 y 7.42, representan el número de consultas recibidas por un servicio de información técnica y todos ellos se ajustan a estos criterios. Si deseamos continuar considerando a la distribución de Poisson como una posibilidad, podríamos hacerlo dejando que equivalga al promedio de la media y de la varianza de la muestra.

$$\lambda = \frac{0.5147 + 0.6007}{2} = 0.5577$$

Programa C para el Lava Autos:

```

#include <stdio.h>
#include <stdlib.h>
#include <conio.h>

void main (void)
{

int i,j,k,n,total;
int linea, r1, r2;
char ttam[4];
int ntam[4];
clrscr();
total=0;
for(i=0;i<4;i++) ntam[i]=0;
ttam[0]='P';
ttam[1]='M';
ttam[2]='G';
ttam[3]='Y';
gotoxy(20,2);printf("SERVICIO DE LAVA AUTO");
gotoxy(15,4);printf("SIMULACION DEL NUMERO DE AUTOS QUE LLEGAN");
gotoxy(15,5);printf("Y SON ATENDIDOS EN UN PERIODO DE 12 HORAS");
gotoxy(5, 7);printf("-----");
gotoxy(5, 8);printf("# ALEATIRO # DE VEHICULOS # ALEATORIO TAMAÑO DEL");
gotoxy(5, 9);printf(" (R1)      ATENDIDOS      (R2) VEHICULO ");
gotoxy(5,10);printf("-----");
linea=0;
for(i=0;i<4;i++)
{
r1=random(100);
if(r1>=0 && r1<30) n=1;
if(r1>=30 && r1<80) n=2;
if(r1>=80 && r1<100) n=3;
gotoxy(7,12+linea);printf("%d      %d",r1,n);
for(j=0;j<n;j++)
{r2=random(100);
if(r2>=0 && r2<30) k=0;
if(r2>=30 && r2<70) k=1;
if(r2>=70 && r2<90) k=2;
if(r2>=90 && r2<100) k=3;
ntam[k]+=1;
gotoxy(35,12+linea);printf("%d      %c",r2,ttam[k]);
linea+=1;
}
}
}
gotoxy(10,13+linea);printf("TOTAL DE AUTOS ATENDIDOS DE CADA TAMAÑO");

```

```
gotoxy(10,14+linea);printf("%d PEQUEÑO(S) %d MEDIANO(S) %d GRANDE(S) %d MUY  
GRANDE(S)",ntam[0],ntam[1],ntam[2],ntam[3]);
```

```
for(i=0;i<4;i++)  
{total+=ntam[i];  
}
```

```
gotoxy(10,15+linea);printf("TOTAL DE AUTOS ATENDIDOS = %d ",total);
```

```
getch();  
}
```

7.7 Bibliografía

Bufa Elwood S., "Administración de Operaciones, La Administración de Sistemas Productivos", Editorial Limusa, México, Primera Edición, 1981.

Schmidt J. W. y Taylor R. E., "Análisis y Simulación de Sistemas Industriales", Editorial Trillas, México, 1979.

Fishman George S., "Conceptos y Métodos en la Simulación Digital de Eventos Discretos", Editorial Limusa, México, 1978.

Haber y Runyon, "Estadística General", Fondo Educativo Interamericano, México, 1972.

Freund y Wampole, "Estadística Matemática", Prentice-Hall, Hispanoamericana, México, 1987.

Taha Hamdy, "Investigación de Operaciones", Editorial Alfa y Omega, México, 1991.

Shamblin James y Steven G. P., "Investigación de Operaciones: Un Enfoque Estructurado", McGraw-Hill.

Meredith, Jack R., "Investigación de Operaciones".

Moskowitz y Wright, "Investigación de Operaciones", Prentice-Hall Hispanoamericana, México, 1985.

Shannon Robert E., "Simulación de Sistemas, Diseño, Desarrollo e Implementación", Editorial Trillas, México, 1988.

Raczynsky Stanislaw, "Simulación por Computadoras".

Coss Bú Raúl, "Simulación: Un enfoque Práctico", Editorial Limusa, México, 1985.

Meier, Newell, Pazer, "Técnica de Simulación en Administración y Economía", Editorial Trillas, México, 1975.

José A. Lozano Teruel, "Montecarlo, el azar doméstico"
<http://www.laverdad.es/panorama/ciencia141002.htm> 14/10/2002.

Carmen M. García Lobo, Francisco R. Villatoro, "Técnicas de Simulación mediante el método de Montecarlo"
http://216.239.50.100/search?q=cache?:3gb0RZS4NawC:polaris.lcc.uma.es/~villa/mmte/tema14doc.pdf+metodo+montecarlo&hl=es&lr=lang_es&ie=UTF-8 1/3/2002

Anexo

Tal como se explicó anteriormente, la Simulación Montecarlo es una técnica computarizada que combina las capacidades que tiene una computadora para generar números aleatorios al azar dentro de un rango de distribución de probabilidad, (valores que van del cero al uno incluyendo ambos, representando 0 para el cero por ciento y uno para el 100 por ciento), con conceptos estadísticos para aproximar expresiones matemáticas complejas que pueden llegar a ser costosas si se intenta llegar a los mismos resultados a través de otra metodología diferente de esta técnica. Esto lo hace a través de simulaciones sistemáticas que son propias de la técnica.

La simulación de Montecarlo es una técnica cuantitativa que consiste en crear un modelo matemático de un sistema tomando en cuenta el riesgo y la incertidumbre. Este identifica las variables cuyo comportamiento aleatorio determina la dinámica global del sistema. Una vez identificadas las variables, se realiza un análisis repetitivo sustituyendo los valores de dichas variables por valores aleatorios en una distribución de probabilidad obteniendo n cantidad de observaciones para generar muestras aleatorias y así poder analizar el sistema con un escenario determinado. Una vez obtenidos una considerable cantidad de observaciones la simulación arroja los valores de probabilidad o determina los resultados de un problema y con estos datos podemos definir cómo funciona el sistema.

La simulación de Monte Carlo proporciona una serie de ventajas sobre el análisis determinista o “estimación de un solo punto”:

- Resultados probabilísticos. Los resultados muestran no sólo lo que puede suceder, sino la probabilidad de que se de cada resultado.
- Resultados gráficos. Gracias a los datos que genera una simulación Monte Carlo, es fácil crear gráficos de diferentes resultados, decir crear un gráfico de cada resultado y ver las posibilidades de que sucedan. Esto es importante para comunicar los resultados relevantes a otras personas interesadas.
- Análisis de sensibilidad. Con sólo unos pocos resultados, en los análisis deterministas es más difícil ver las variables que más afectan el resultado. En la simulación Monte Carlo, resulta más fácil ver qué variables introducidas tienen mayor influencia sobre los resultados finales.
- Análisis de escenario. En los modelos deterministas resulta muy difícil modelar diferentes combinaciones de valores de entrada, con el fin de ver los efectos de situaciones verdaderamente diferentes. Usando la simulación Monte Carlo, los analistas pueden ver exactamente los valores que tienen cada variable cuando se producen ciertos resultados. Esto permite profundizar en los análisis.

- Correlación de variables de entrada. En la simulación Monte Carlo es posible modelar relaciones interdependientes entre diferentes variables de entrada. Esto es importante para averiguar con precisión la razón real por la que, cuando algunos factores suben, otros suben o bajan paralelamente.

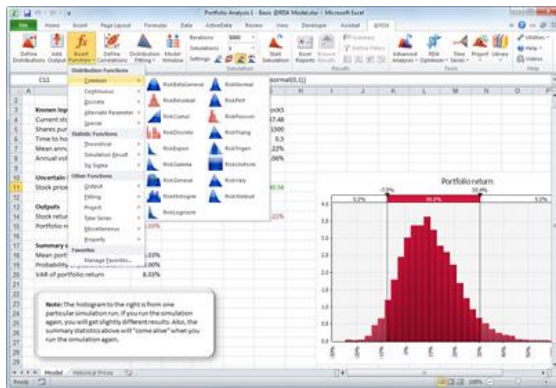
Sitios interesantes de Montecarlo

Existen varios software que utilizan la técnica de simulación de Montecarlo, entre ellos tenemos algunos:

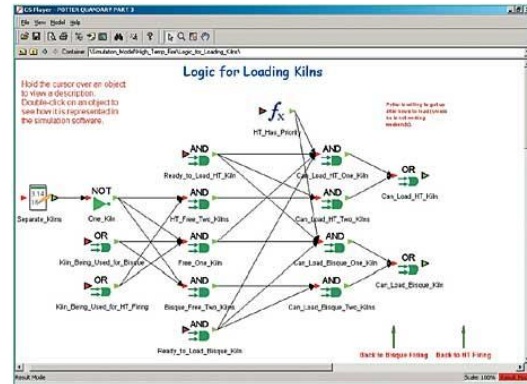
- **@Risk:** realiza análisis utilizando la técnica o simulación de Montecarlo para mostrar muchos resultados en los modelos de hojas de cálculo y dice como ellos ocurren. Computa matemática y objetivamente y rastrea muchos escenarios futuros diferentes. Esto significa que puedes considerar cuantos riesgos puedes tomar y cuantos puedes evitar, permitiéndote hacer la mejor decisión bajo incertidumbre. @Risk también ayuda a diseñar las mejores estrategias de gestor de riesgos a través de la integración de RiskOptimizer, que combina la simulación de Montecarlo con la última tecnología de solución para optimizar cualquier hoja de cálculo con valores dudosos. Utilizando algoritmos genéticos junto con funciones de @Risk, RiskOptimizer puede determinar la mejor asignación de recursos, la asignación óptima de los activos, los planes más eficientes y mucho más.



@RISK
el nuevo estándar en análisis de riesgo

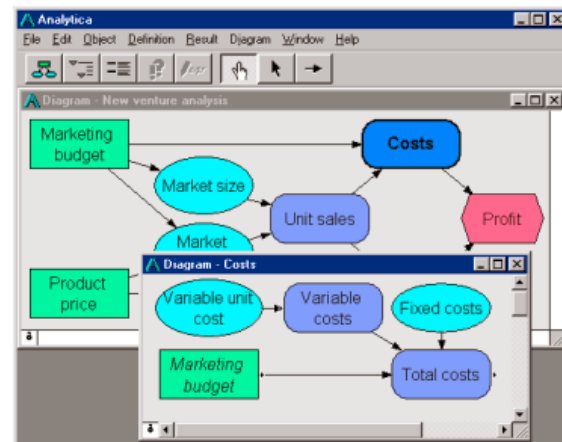
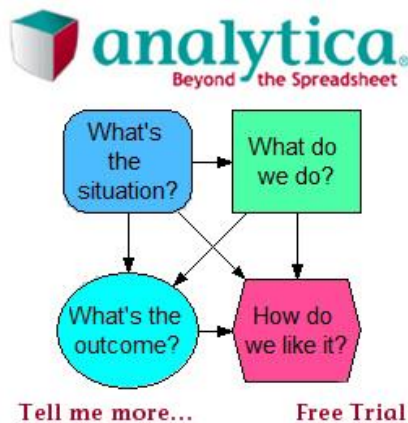


- **GoldSim:** es el primer software de solución de simulación de Montecarlo para el modelado dinámico de sistemas complejos en los negocios, ingeniería y ciencia. GoldSim soporta análisis de decisiones y riesgos simulando futuros rendimientos mientras representando cuantitativamente la incertidumbre de riesgos en todos los sistemas complejos.



- **Analytica:** fue diseñada desde el comienzo para combinar varias decisiones de tecnologías clave en un paquete totalmente integrado. Esto incluye:
 - Diagramas de influencia visual para definir y organizar modelos complejos en módulos comprensibles.
 - Arreglos inteligentes para hacer más fácil el manejo de modelos con múltiples dimensiones.
 - La simulación de Montecarlo para analizar incertidumbre y riesgo con velocidad y simplicidad.
 - Optimización, incluyendo programación lineal y no lineal.

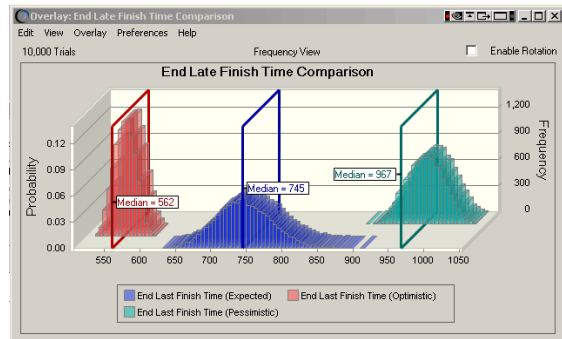
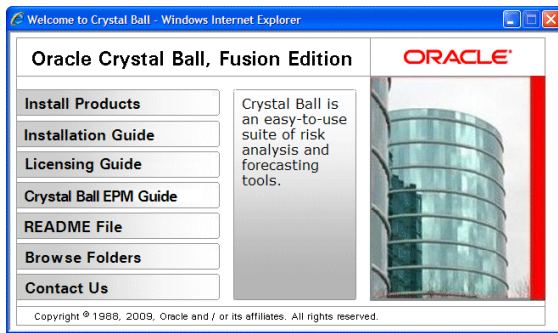
La tecnología subyacente Analytica se basa en una década de investigación por los analistas de decisiones, científicos informáticos y diseñadores de interfaz de usuario en la Universidad Carnegie Mellon, con el apoyo de la National Science Foundation.



- **Oracle Crystal Ball:** Oracle Crystal Ball es la aplicación líder basada en la hoja de cálculo para el modelado predictivo, la predicción, simulación y optimización. Te da una visión sin precedentes de los factores críticos que afectan el riesgo. Con Crystal Ball, usted puede hacer las decisiones tácticas más adecuadas para alcanzar sus objetivos y ganar una ventaja competitiva incluso en las condiciones de mercado más inciertos.



- Se basa en herramientas de modelado predictivo Monte Carlo y existentes.
- Proporciona capacidades de optimización y cálculo avanzado.
- Combina Crystal Ball de Oracle y Oracle Crystal Ball Decision Optimizer.
- Ofrece una versión con precio accesible de Crystal Ball de Oracle y el Oracle Crystal Ball Decision Optimizer.



Ejemplos Resueltos de Capítulo:

El Tambaleo del Borracho

A manera de ilustración consideremos el clásico tambaleo del borracho o de la caminata aleatoria (conocido también como el ebrio aleatorio).

Este tipo de simulación es utilizado en juegos de videos computarizados tales como PACMAN, TETRIS, DONKEY KONG, etc.; los cuales se basan en cálculos de posiciones de acuerdo con el comportamiento del jugador.

Supongamos que un borracho está parado en la esquina de una calle cuando decide caminar para salir de su embriagues.

Asuma que hay igual probabilidad de que vaya hacia el norte, sur, este u oeste. Si camina 10 cuadras. ¿Cuál es la probabilidad de que termine a dos cuadras de donde inició?

Pasos para resolver el problema.

Paso 1

Tabulamos los datos, en la primera columna colocamos las posibles direcciones que puede tomar el borracho. En la segunda columna colocamos la probabilidad, al ser 4 direcciones, y tomando en cuenta que existe igual probabilidad de que vaya hacia el norte, sur, este u oeste, entonces la probabilidad sería de 0.25 para cada dirección. En la tercera columna colocamos la probabilidad acumulada. La tabla nos queda de esta forma:

Dirección	Probabilidad	Probabilidad Acumulada
Norte	.25	.25
Sur	.25	.50
Este	.25	.75
Oeste	.25	1.00

Paso 2

Designamos una localización a cada esquina mediante coordenadas. Utilizamos la letra 'x' para representar la dirección este-oeste, y la letra 'y' para representar la dirección norte-sur.

Cada vez que se mueva hacia el este incrementamos el valor de 'x' y cada vez que se mueva hacia el oeste le restamos uno a 'x'. Se realiza un proceso semejante con el movimiento norte-sur.

Paso 3

Antes de iniciar establecemos la posición inicial del borracho con las coordenadas (0, 0).

Paso 4

Generamos números aleatorios de dos dígitos entre el 00 y el 99. Cada número podrá representar una esquina e indicará la posible dirección a la que caminará el borracho.

Paso 5

En este paso se evaluará la dirección que va a tomar el borracho, dependiendo del número aleatorio generado en el paso anterior.

Si el número obtenido va de 00 a 24 él se dirige hacia el este y sumamos 1 a x; si el número aleatorio va del 25 al 49, él camina hacia el este y sumamos 1 a y; si el número aleatorio va del 50 al 74, se dirige al norte y sumamos 1 a y; si el número aleatorio va del 75 al 99, se dirige al sur y restamos 1 de y.

Paso 6

Se ajustan las coordenadas para la nueva posición que tendrá el borracho una vez realizado el paso anterior. Se muestran los valores actuales de las coordenadas.

Paso 7

Se repiten los pasos desde el 4 hasta el 6, n veces, donde n será la cantidad de iteraciones que se desean realizar.

Paso 8

Se muestra las coordenadas de la ubicación final del borracho. Se suman los valores de las coordenadas 'x' y 'y' en valor absoluto para verificar que el resultado sea igual a 2.

De ser así se considera como “éxito” porque nos indica que el borracho, al final del recorrido de n pasos se encuentra a 2 cuadras del punto de partida.

Los resultados se plantean en la siguiente tabla:

Fin de la Esquina	Prueba 1			Prueba 2		Prueba 3		Prueba 4		Prueba 5	
	# Aleatorio	Loc.		# Aleatorio	Loc.	# Aleatorio	Loc.	# Aleatorio	Loc.	# Aleatorio	Loc.
1	73	0, 1		10	1, 0	05	1, 0	06	1, 0	64	0,1
2	21	1, 1		89	1, -1	88	1, -1	95	1, -1	76	0, 0
3	45	0, 1		14	2, -1	10	2, -1	04	2, -1	79	0, -1
4	76	0, 0		81	2, -2	04	3, -1	67	2, 0	54	0, 0
5	96	0, -1		30	1, -2	48	2, -1	51	2, 1	28	-1, 0
6	94	0, -2		91	1, -3	19	3, -1	95	2, 0	05	0, 0
7	53	0, -1		06	2, -3	44	2, -1	73	2, 1	71	0, 1
8	57	0, 0		38	1, -3	21	3, -1	10	3, 1	75	0, 0
9	96	0, -1		79	1, -4	95	3, -2	76	3, 0	53	0, 1
10	43	-1, -1		43	0, -4	11	4, -2	30	2, 0	29	-1, 1
EXITO	SI			NO		NO		SI		SI	

Conclusión:

Esta simulación tiene una distribución de probabilidad uniforme, puesto que el experimento funciona de esta forma, dado que la dirección que puede tomar el borracho en cada paso es aleatoria, y exista igual probabilidad para cualquiera de las 4 direcciones.

Cálculo de PI (π) por el Método de la Aguja de Buffón

Este ejemplo permite apreciar una vez más la ingeniosidad de la mente humana para solucionar los problemas, en este caso calcularemos el valor de la constante Pi (π) por medio de simulación utilizando el método de la Aguja de Buffón. Partimos del hecho de que p es la probabilidad de que una aguja de longitud lanzada al azar sobre una superficie plana que contiene un conjunto de líneas paralelas separadas entre sí una distancia D (con $D > L$), haga contacto con alguna de las líneas, entonces su valor está dado por la ecuación:

$$P = \frac{2L}{\pi D}$$

El caso particular de este problema se obtiene cuando en la ecuación se hace que $D = 2L$, quedando $p = 1/\pi$. De acuerdo con la teoría de probabilidades, si N es el número de veces que se lanza la aguja y n el número de aciertos, el cociente n/N tiende a la probabilidad de p cuando N tiende al infinito, de donde al sustituir en $p = 1/\pi$ se obtiene $N = n \cdot \pi$ que es la relación buscada.

Para simular este problema conviene encontrar una representación matemática (modelo) que nos permita encontrar este valor. La siguiente ilustración muestra la posición que puede tomar una aguja al caer sobre la superficie con múltiples líneas paralelas.

Se observa que la posibilidad de tocar alguna de las líneas se puede definir en términos de la distancia x , que existe entre el centro de la aguja y la línea más próxima, y en término del ángulo que forma la aguja y la línea.

Paso 1

Definir las condiciones para X , en las que se puede determinar si la aguja toca la línea o no. Si

$$x \leq \frac{L}{2} \sin \theta$$

La aguja toca la línea.

Por el contrario, si:

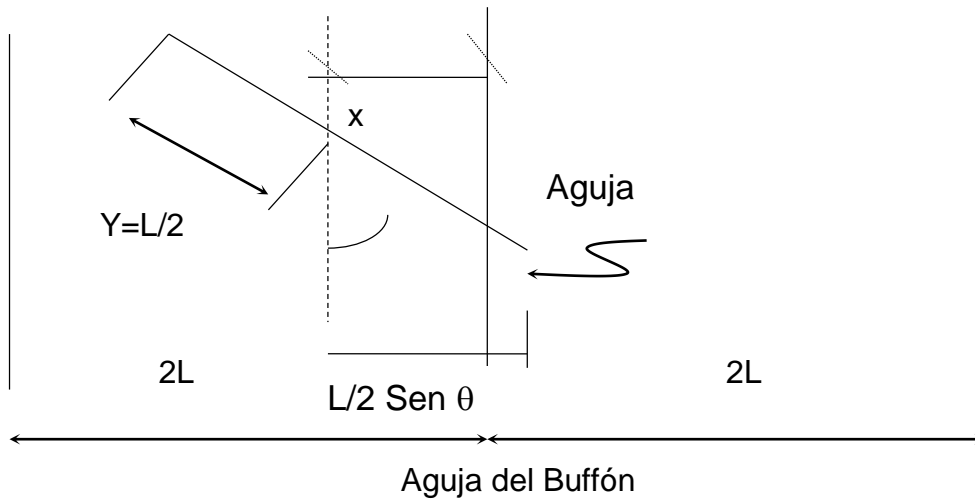
$$x > \frac{L}{2} \sin \theta$$

La aguja no la toca.

De aquí se puede concluir que X y θ están limitados de la siguiente forma:

$$0 \leq x \leq \frac{L}{2}$$

$$0 \leq \theta \leq 180^\circ$$



Paso 2

Se genera un número aleatorio que represente la distancia del centro de la aguja a la línea. Luego se multiplica por $L/2$ para así obtener X .

Paso 3

Generar otro número aleatorio que represente el ángulo que forma la aguja y la línea; multiplicarlo por p para obtener θ .

Paso 4

Calcular el valor de Y . Se calcula mediante la fórmula $Y = L/2 * \text{sen}\theta$. Donde L representa la longitud de la aguja y el valor del ángulo se toma del número aleatorio generado en el paso anterior.

Paso 5

Comparar el valor de X vs Y .

Si el valor de X es menor o igual al de Y se cuenta como acierto.

Paso 6

El procedimiento se repite N veces y si n es el número de aciertos, la probabilidad p estará dada por n/N . Este valor permite calcular π despejando:

$$\pi = 1 / p \Rightarrow \pi = \frac{I}{\frac{n}{N}} \Rightarrow \pi = N / n$$

Los valores iniciales son: N = número de lanzamientos = 25, L = longitud de la aguja = 3 cm, $L/2 = 1.5$

# Aleatorio (r_1)	$X = L/2 * r_1$	# Aleatorio (r_2)	$\theta = \pi * r_2$	$Y = L/2 * \sin\theta$	Acierto $X \leq Y$
.80	1.2	.39	1.2252	1.4113	✓
.20	.3	.00	0	0	
.15	.225	.35	1.0996	1.3365	✓
.88	1.32	.04	.1257	.1881	
.98	1.47	.12	.3740	.1355	
.65	.975	.11	.3456	.5081	
.86	1.29	.23	.7226	.9920	
.73	1.095	.18	.5655	.8038	
.28	.42	.83	2.6075	.7636	✓
.60	.9	.35	1.0996	1.3365	✓
.60	.9	.50	1.5708	1.5	✓
.29	.435	.52	1.6336	1.4978	✓
.18	.27	.68	2.1363	1.2665	✓
.90	1.35	.29	.9111	1.1852	
.93	1.395	.23	.7225	.9919	
.73	1.095	.40	1.2566	1.4266	✓
.21	.399	.14	.4398	.6386	✓
.45	.675	.96	3.0159	.1880	✓
.76	1.14	.94	2.9531	.2817	✓
.96	1.44	.54	1.6965	1.4882	✓
.94	1.41	.37	1.1624	1.3766	
.53	.795	.42	1.3195	1.4529	✓
.57	.855	.22	.6912	.9562	✓
.96	1.44	.28	.8796	1.1557	
.43	.645	.07	.2199	.3272	

Sistema de Información empresarial (EIS)

En la imagen se muestra un análisis histórico de 200 días sobre el número de consultas diarias realizadas a EIS residente en un servidor central. La tabla incluye el número de consultas diarias (0 a 5), las frecuencias absolutas (número de días que se producen 0 – 5 consultas), las frecuencias relativas ($10/200 = 0.05...$) y las frecuencias relativas acumuladas.

CONSULTAS EIS	FREC. ABS (DIAS)	FREC. RELATIVA	F.R.A
0	10	0,05	0,05
1	20	0,10	0,15
2	40	0,20	0,35
3	60	0,30	0,65
4	40	0,20	0,85
5	30	0,15	1,00
	200	1,00	

Supongamos que queremos conocer el número esperado (o medio) de consultas por día. La respuesta a esta pregunta es fácil si recurrimos a la teoría de probabilidad.

Podemos interpretar la frecuencia relativa como la probabilidad de que ocurra el suceso asociado, en este caso, la probabilidad de un determinado número de consultas (así, por ejemplo, la probabilidad que se den 3 consultas en un día sería de 0.03), por lo que la tabla anterior nos proporciona la distribución de probabilidad asociada a una variable aleatoria discreta (la variable aleatoria es el número de consultas al EIS, que sólo puede tomar valores enteros entre 0 y 5).

Por otra parte, también podemos usar simulación de Monte Carlo para estimar el número esperado de consultas diarias (en este caso se ha podido obtener el valor exacto usando teoría de probabilidad, pero ello no siempre será factible). Veamos como:

Cuando se conozca la distribución de probabilidad asociada a una variable aleatoria discreta, será posible usar la columna de frecuencias relativas acumuladas para obtener los llamados intervalos de números aleatorios asociados a cada suceso. En este caso, los intervalos obtenidos son:

- (0.00 – 0.05) para el suceso 0
- (0.05 – 0.15) para el suceso 1
- (0.15 – 0.35) para el suceso 2
- (0.35 – 0.65) para el suceso 3
- (0.65 – 0.85) para el suceso 4
- (0.85 – 1.00) para el suceso 5

Esto significa que, al generar un número pseudo-aleatorio con el ordenador (proveniente de una distribución uniforme entre 0 y 1), estaremos llevando a cabo un experimento cuyo resultado, obtenido de forma aleatoria y según la distribución de probabilidad anterior, estará asociado a un suceso. Así, por ejemplo, si el ordenador nos proporciona el número pseudo-aleatorio 0.2567, podremos suponer que ese día se han producido 2 consultas al EIS.

Asignamos pues la función ALEATORIO a una casilla (la G1 en el caso de la imagen):

G1							
=ALEATORIO()							
	A	B	C	D	E	F	G
1							0,14116748
2		CONSULTAS EIS	FREC. ABS (DIAS)	FREC. RELATIVA	F.R.A		0,49664084
3		0	10	0,05	0,05		0,90480881
4		1	20	0,10	0,15		0,98844792
5		2	40	0,20	0,35		0,91653555
6		3	60	0,30	0,65		0,94216577
7		4	40	0,20	0,85		0,41992604
8		5	30	0,15	1,00		0,02154918
9			200	1,00			0,65565457
10							0,594669
11							0,6816434

Seleccionando la celda y “arrastrando” con el ratón desde el borde inferior derecho de la misma podemos obtener un listado completo de 100 números pseudo-aleatorios:

A continuación, podemos usar la función SI de Excel para asignar un suceso a cada uno de los números pseudo-aleatorios generados.

H1										=SI(G1<E\$3; B\$3;SI(G1<E\$4;B\$4;SI(G1<E\$5;B\$5;SI(G1<E\$6;B\$6;SI(G1<E\$7;B\$7;B\$8))))))									
	A	B	C	D	E	F	G	H	I										
1							0,14116748	1	2,89										
2		CONSULTAS EIS	FREC. ABS (DIAS)	FREC. RELATIVA	F.R.A		0,49664084	3											
3		0	10	0,05	0,05		0,90480881	5											
4		1	20	0,10	0,15		0,98844792	5											
5		2	40	0,20	0,35		0,91653555	5											
6		3	60	0,30	0,65		0,94216577	5											
7		4	40	0,20	0,85		0,41992604	3											
8		5	30	0,15	1,00		0,02154918	0											
9			200	1,00			0,65565457	4											
10							0,594669	3											
11							0,6816434	4											

Si el número pseudo-aleatorio en G1 es menor que el valor de la F.R.A. en E3, entonces la consulta EIS es 1 y se coloca en H1. Luego se repite este proceso con el mismo valor en G1 y los diferentes valores de las F.R.A en E4, E5, E6, E7, E8.

Repitiendo el proceso de seleccionar y “arrastrar” obtendremos algo similar.

Finalmente, usando la función PROMEDIO será posible calcular la media de los valores de la columna H:

I1 fx =PROMEDIO(H:H)								
	A	B	C	D	E	F	G	H
1							0,14116748	1
2		CONSULTAS EIS	FREC. ABS (DIAS)	FREC. RELATIVA	F.R.A		0,49664084	3
3		0	10	0,05	0,05		0,90480881	5
4		1	20	0,10	0,15		0,98844792	5
5		2	40	0,20	0,35		0,91653555	5
6		3	60	0,30	0,65		0,94216577	5
7		4	40	0,20	0,85		0,41992604	3
8		5	30	0,15	1,00		0,02154918	0
9			200	1,00			0,65565457	4
10							0,594669	3
11							0,6816434	4

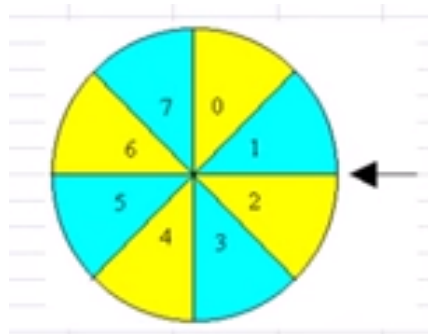
En este caso, hemos obtenido un valor estimado del número de consultas por día.

La Ruleta

Se desea simular por medio del método Montecarlo una ruleta para que salga un número esperado de 100 observaciones. Además, calcule la media, desviación estándar y la moda de los resultados obtenidos.

Solución:

Los números que saldrán van desde el 0 hasta el 7 como se ve en la figura siguiente:



Sabemos que al girar la ruleta tenemos 8 opciones y que solo puede salir un resultado de los 8 números, es decir $1/8$.

Se elabora una tabla de la Distribución de Probabilidad:

LA RULETA				
X	P(X)	F(X)	INTERVALOS	
0	0.125	0.125	0	0.125
1	0.125	0.25	0.125	0.25
2	0.125	0.375	0.25	0.375
3	0.125	0.5	0.375	0.5
4	0.125	0.625	0.5	0.625
5	0.125	0.75	0.625	0.75
6	0.125	0.875	0.75	0.875
7	0.125	1	0.875	1

Donde la columna:

- X: tiene los resultados que pueden salir al girar la ruleta.
- P(x): es la probabilidad de que salga uno de los resultados de X.
- F(X): es la probabilidad acumulada de F(X).
- INTERVALOS: es donde se encuentra ubicado los resultados según el intervalo.

Si nos damos cuenta, en la columna de probabilidad de un número $P(X)$ todas las probabilidades son iguales a 0.125, ya que la probabilidad de que salga 0 es de $(1/8)$; la probabilidad que salga 1 es de $(1/8)$; la probabilidad que salga 7 es de $(1/8)$ y así sucesivamente, porque sólo puede salir un número como resultado.

Los intervalos se determinan de la siguiente manera:

- Del 0.00 al primer valor de la columna $F(X)$, es decir 0.125.
- Del 0.125 al segundo valor de la columna $F(X)$, es decir 0.25.
- Del 0.25 al tercer valor de la columna $F(X)$, es decir 0.375.


Y así sucesivamente hasta obtener el valor 1 que nos comprueba que el proceso se hizo bien.

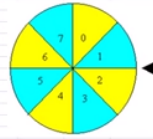
Ahora vamos a simular la fuerza con que giramos la ruleta y obtenemos un valor, usando la función de Excel ALEATORIO () en la columna I1. Luego en la columna J1 usamos la función BUSCAR () para ver en qué intervalo se encuentra el valor pseudo-aleatorio que obtuvimos y así poder determinar qué número, de forma simulada, saldría si girásemos la ruleta. La fórmula sería ésta: =BUSCAR(I1,D3:E10,A3:A10)

Debemos anclar o fijar los intervalos (D3:E10) y los resultados de X (A3:A10) con la tecla F4, ya que estos no cambiarán. El valor pseudo-aleatorio si es variable, por tanto, no se deberá anclar. La fórmula quedaría de esta manera:
=BUSCAR(I1,\$D\$3:\$E\$10,\$A\$3:\$A\$10)

Seleccionamos ambas columnas I1 Y J1 Y la “arrastramos” hacia abajo hasta obtener 100 observaciones. Luego determinamos la media, la desviación estándar y la moda de las 100 observaciones con las funciones PROMEDIO (), DESVEST () Y MODA () respectivamente.

L1					f _x						
	A	B	C	D	E	F	G	H	I	J	K
1	LA RULETA								0.66867938	5	
2	X	P(X)	F(X)	INTERVALOS					0.04440969	0	
3	0	0.125	0.125	0	0.125				0.22775241	1	
4	1	0.125	0.25	0.125	0.25				0.12040914	0	
5	2	0.125	0.375	0.25	0.375				0.93089666	7	
6	3	0.125	0.5	0.375	0.5				0.1442045	1	
7	4	0.125	0.625	0.5	0.625				0.76324673	6	
8	5	0.125	0.75	0.625	0.75				0.35154269	2	
9	6	0.125	0.875	0.75	0.875				0.0370308	0	
10	7	0.125	1	0.875	1				0.94534075	7	
11									0.12002113	0	
12	Media	3.34							0.71950339	5	
13	Desv. Est.	2.34507866							0.20209288	1	
14	moda	6							0.73272139	5	
15									0.87075087	6	
16									0.79442451	6	
17									0.46186235	3	
18									0.55923552	4	
19									0.27040245	2	
20									0.08886023	0	





Podemos ver que:

- La media = 3.34, que es una aproximación a un número o un valor "central" entre la cantidad que tenemos.
- La desviación estándar = 2.3450..., que es la medida de dispersión y nos indica cuánto pueden alejarse los valores respecto al promedio (media).
- La moda = 6, que es el valor con mayor frecuencia en la distribución de datos, es decir que el número 6 se repite más que los otros resultados obtenidos.

Y de esta manera se ha resuelto el ejemplo.

Referencias Bibliográficas

Definición de la técnica de Montecarlo:

- Capítulo 7 facilitado por el profesor, C7 Simulación de Montecarlo.
- http://es.wikipedia.org/wiki/M%C3%A9todo_de_Montecarlo
- http://www.palisade-lta.com/risk/simulacion_monte_carlo.asp
- <http://www.expansion.com/diccionario-economico/simulacion-de-monte-carlo.html>
- <http://www.uib.cat/depart/dqu/dqf/paco/docencia/QFC-T4.pdf>

Softwares que utilizan la técnica de Montecarlo:

- <http://www.palisade-lta.com/risk/>
- <http://www.goldsim.com/Home/>
- <http://www.lumina.com/>
- <http://www.oracle.com/us/products/applications/crystalball/overview/index.html>
- <http://www.oracle.com/technetwork/middleware/crystalball/overview/index.html>