VIII- Simulación

VIII.1 Introducción

La simulación es una técnica experimental de modelado de sistemas que tiene por objeto evaluar sus parámetros de comportamiento a partir de una estimación estadística basada en la realización ficticia de sus actividades. Tal realización puede lograrse si de antemano se conocen los patrones de comportamiento de cada actividad y por supuesto, se dispone también de una descripción del sistema que establece la forma en que ellas se vinculan. Con esos patrones pueden registrase muestras de la actividad total simulada del sistema y evaluar así comportamientos centrales, variabilidades, proporciones e incluso analizar tiempos, costos y beneficios. La simulación reemplaza con ventaja al modelado analítico cuando las características estocásticas del sistema y el carácter intrincado de su expresión matemática dificulta el establecimiento de fórmulas de cálculo. En ese contexto permite mayor simplicidad y similar efectividad como medio para evaluar decisiones. Por ejemplo; el mismo modelo de colas M/M/1, que permite el cálculo del tiempo esperado que cada cliente permanecerá en un sistema, puede ser representado por un modelo de simulación que siga los patrones poisson-exponencial para las entradas y salidas, con velocidades esperadas λ y μ conocidas. Bastará para ello generar simuladamente entradas y salidas de clientes del sistema y a partir de la muestra obtenida establecer una estimación $\widehat{W}$ para el tiempo esperado W. Claro que ya aprendimos que ese tiempo esperado es $W = \frac{1}{\mu-\lambda}$ pero si no contásemos con una fórmula deducida, podríamos obtener una estimación del parámetro por simulación. De tal modo, cuando por la complejidad del sistema o por cualquier otra razón, no disponemos de una fórmula analítica para un parámetro, siempre queda abierto el camino de la simulación para obtener su valor.

VIII.2 Números aleatorios

Si tenemos en una bolsa 1000 bolillas numeradas del 000 al 999, mezcladas sin arreglo a ningún patrón, la probabilidad de elegir una cualquiera de ellas es 1/1000 que resulta la misma que la de elegir cualquier otra. Además, si volvemos la bolilla extraída a la bolsa y mezclamos, la probabilidad de volver a sacarla, o de sacar cualquier otra, sigue siendo 1/1000. Esto pone de manifiesto dos características importantes de este sistema de almacenar y extraer bolillas, y por ende de los números por ellas representados. Todas tienen igual probabilidad de ser elegidas en cada sorteo y cada elección resulta independiente de toda otra anterior. Así, los números elegidos a partir de esta mecánica se denominan aleatorios. Seguramente se podrían realizar diez millones de sorteos con estas características y obtener, en base a la llamada regularidad estadística, aproximadamente una misma proporción de veces cada uno de los mil números. Es decir, tal proporción debiera ser cercana a 1/1000. En la práctica uno no anda por la vida con bolsas de 1000 bolillas encima, pero está claro que cada vez que quiere elegir un número de tres dígitos al azar, precisa hacer un sorteo que al menos emule las condiciones del sistema descripto de las bolillas en la bolsa. Ya que esto puede ser muy necesario al simular, como en seguida veremos, se han diseñado procedimientos algorítmicos, para computadoras y calculadoras, a fin de obtener números que sigan estos patrones Claro que por tratarse de un cálculo cuyos principios podemos conocer, su obtención podría no ser tan aleatoria. Sin embargo, es posible dar al resultado una característica al menos *pseudoaleatoria*, realizando el cálculo a partir de un número llamado semilla, obtenido en forma azarosa o desconocida. En ese caso nos será muy difícil reproducir el cálculo, si lo intentáramos, por lo cual a efectos prácticos actuaríamos con el número obtenido como si fuera puramente aleatorio.

El procedimiento algorítmico básico para obtener un número aleatorio es el método congruencial. Se basa, justamente en las congruencias de módulo m que estudia la aritmética. Un número x_1 es congruente con otro x_0, módulo m, si se cumple la relación $m \times q + x_1 = ax_0 + c$. Esto se corresponde con la disposición de una división por el entero m del número $ax_0 + c$, siendo q el cociente y x_1 el resto. Es decir:

$$\begin{array}{r|l} ax0 + c & m \\ \hline x1 & q \end{array}$$

Esto se suele escribir $x_1 \equiv (ax_0 + c)\ modulo\ m$

Si por ejemplo fueran $m = 8$, $a = 1$, $c = 5$ y se eligiera al azar la semilla $x_0 = 5$ se tendría

$$\begin{array}{r|l} 1x5 + 7 & 8 \\ \hline 4 & 1 \end{array}$$

siendo, $8 \times 1 + \mathbf{4} = 1 \times \mathbf{5} + 7$ o lo que resulta lo mismo, $4 \equiv (1 \times 5 + 7)\ modulo\ 8$. El azar de la elección de la semilla x_0 como un número natural entre 0 y 7, se refiere en realidad a una hipótesis de igual probabilidad e independencia. Esto puede suponerse pues, para el usuario del algoritmo, la elección resulta desconocida. La clave es que siempre está fundada en un cálculo distinto y oculto, efectuado por ejemplo a partir de la hora del reloj interno de una computadora. Así el número 5 se supondrá obtenido aleatoriamente pues, dada cualquier semilla elegida entre 0 y 7, los números congruentes con ellas aparecen solo una vez en la Tabla 1 correspondiente a la expresión $x_{n+1} \equiv (ax_n + c)\ modulo\ m$

Tabla 1

n	Semillas x_n	$1 \times x_n + 7$	q	Restos x_{n+1}
0	5	12	1	4
1	4	11	1	3
3	3	10	1	2
4	2	9	1	1
5	1	8	1	0
6	0	7	0	7
7	7	14	1	6
8	6	13	1	5

Es decir; si la experiencia se realizase en un número suficiente de oportunidades, se obtendría una proporción aproximadamente igual de veces cualquiera de los restos posibles de la división. Cada uno de ellos tendría entonces igual probabilidad de ser elegido.

La elección de $m = 8$ tampoco es casual. Se debe a que el byte tiene 8 bits y con él es posible representar los números en una computadora, aunque no nos detendremos en esto aquí. Sí, hay que aclarar que no cualquier elección de las cantidades a y c es adecuada para que todos los números entre 0 y $m - 1$ aparezcan una vez en la tabla. Si $m = 2^b$, potencia de 2, las elecciones adecuadas son $a = \{1,5,9,13 \ldots\}$ y $c = \{1,3,5,7 \ldots\}$ mientras que si es m una potencia de 10, $m = 10^d$ resultan elegibles $a = \{1,21,41,61 \ldots\}$ y $c = \{1,3,7,9,11,13,17,19 \ldots\}$

Existen, en realidad, una variedad de métodos para generar números aleatorios por medio de una computadora y, para nuestros fines, necesitaremos números que se encuentren entre los valores reales 0 y 1. Por esta razón los mil números aleatorios entre 000 y 999, por ejemplo, tendrán que considerarse como la parte decimal de tres dígitos de los números comprendidos entre 0.000 y 0.999. Estos son los números que habitualmente obtenemos con la instrucción random de muchos lenguajes de programación o con teclas en calculadoras.

VIII.3 Observaciones aleatorias

Para exponer ideas recurrimos al siguiente ejemplo. Se arrojan dos monedas y se observa el resultado. Por supuesto no todos los pares que puedan obtenerse tienen igual probabilidad de ocurrir. Cara y cara, CC, tiene probabilidad ¼ al igual que ceca y ceca, YY, mientras que cara y ceca, que se obtiene tanto en YC como en CY, tiene probabilidad ½. ¿Cómo uno podría simular la obtención del resultado con arreglo a esta probabilidad? La Tabla 2 muestra las probabilidades y su distribución acumulada.

Tabla 2

Suceso X	P(X)	$\sum P(x)$	Intervalo
CC	$1/4 = 0.250$	0.250	$0.000 \leq r \leq 0.249$
YC o CY	$1/2 = 0.500$	0.750	$0.250 \leq r \leq 0.749$
YY	$1/4 = 0.250$	1.000	$0.750 \leq r \leq 0.999$

Si se entienden 0.000 y 0.999 como suficientes aproximaciones para 0 y 1, la cuarta columna de la tabla explica entonces que números entre 0.000 y 0.999 representan a cada uno de los tres sucesos posibles. Es decir; simular la experiencia terminaría consistiendo en apretar el botón de la función random de una calculadora y observar a cuál de los tres intervalos pertenece el resultado *r* obtenido. Si por ejemplo obtuviéramos el número aleatorio $r = 0.876$, su *observación* aleatoria asociada sería YY, ceca y ceca. Este es el principio básico de la simulación y su aplicación consecuente puede llevarnos al modelado sencillo de sistemas complejos.

Obtener observaciones aleatorias, es decir sucesos simulados que respeten un patrón estadístico de ocurrencia determinado, es una tarea que puede emprenderse por distintas técnicas. La más directa consiste en invertir el sentido de los cálculos, yendo desde el valor del número aleatorio *r* hasta el valor de la observación *x*, por medio de la inversión de la función de distribución acumulada. Este método conocido como de la *transformación inversa* comprende, en la práctica real, tres etapas.

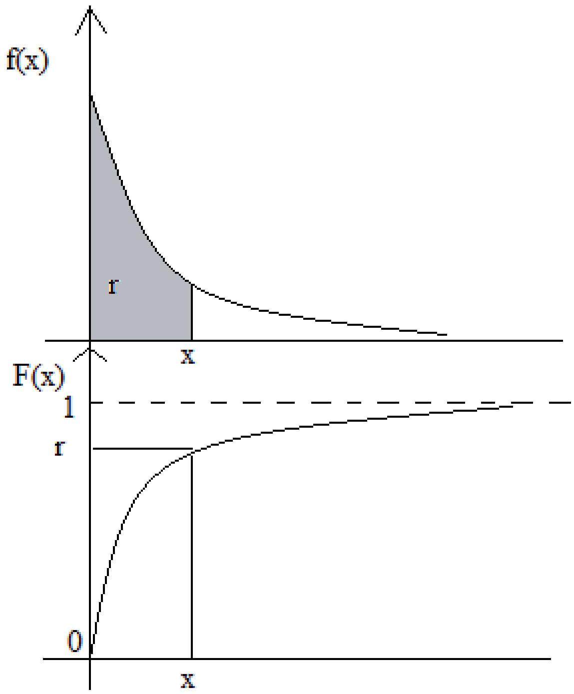
i) Hallar la función que caracteriza la actividad y comprobar que es una función de probabilidad. Usualmente esto requiere tomar muestras y realizar ajustes por medio de técnicas de cálculo numérico para encontrar la expresión del patrón estadístico. Por ejemplo, para conocer cómo se distribuyen los tiempos de atención de un cajero de supermercado habrá que tomar una muestra de tiempos reales de atención y ajustar, por medio de mínimos cuadrados, la función exponencial negativa que mejor los represente. Luego habrá que comprobar que se trata efectivamente de una función de densidad de probabilidad.

ii) Hallar la expresión de la distribución acumulada. Si la variable aleatoria fuera discreta la acumulación se produciría a través de la suma de probabilidades, pero si fuese continua, el cálculo debiera resolverse por integración. En el caso ejemplificado del cajero del supermercado habría que integrar utilizando la función de densidad de probabilidad obtenida en i).

iii) A cada suceso x, la función acumulada le hace corresponder un valor real entre 0 y 1. Es decir $F(x) = r$ con $0 \leq r \leq 1$. Este paso consiste entonces en hallar $F^{-1}(r) = x$. Es decir, invertir la distribución acumulada.

Supongamos entonces que una muestra nos permitió ajustar los tiempos de atención del cajero según la expresión $f(x) = \begin{cases} \alpha e^{-\alpha x} & si\ x \geq 0 \\ 0 & en\ otro\ caso \end{cases}$ siendo $\alpha = 30\ clientes/hora$ Deberíamos ahora realizar la secuencia de tres pasos indicada arriba. Comenzamos comprobando que se trata de una función de densidad de probabilidad. Claramente $f(x) = 30e^{-30x} = \frac{30}{e^{30x}} \geq 0$ que es la primera condición a cumplir para ser fdp. Además $\int_{-\infty}^{\infty} f(x)dx = \int_0^{\infty} 30e^{-30x}dx = 30\int_0^{\infty} e^{-30x}dx = 30\frac{1}{-30}e^{-30x}\Big|_0^{\infty} = 0 - (-1) = 1$ con lo cual queda probado el punto i) pues se trata de una función de densidad de probabilidad. En seguida armamos entonces la distribución acumulada: $F(x) = \int_0^x 30e^{-30t}dt = 30\frac{1}{-30}e^{-30t}\Big|_0^x = -e^{-30x} - (-1) = 1 - e^{-30x}$. Como es sabido y se muestra en la Gráfica 1, $F(x) = r$ adopta valores entre 0 y 1 por lo cual está completado el paso ii)

Gráfica 1



El paso iii) requiere invertir la función de distribución hallando $F^{-1}(r) = x$. Hacemos:

$r = 1 - e^{-30x}$

$e^{-30x} = 1 - r$

$\ln e^{-30x} = \ln(1 - r)$

$x = \frac{\ln(1-r)}{-30} = F^{-1}(r)$

Al asumir valores aleatoriamente elegidos del número *r*, se generan observaciones aleatorias de tiempos de atención del cajero, que en términos estadísticos seguirán una distribución exponencial con media $\frac{1}{\alpha} = \frac{1}{30 cilentes\ /hora} = 2\ minutos/cliente$

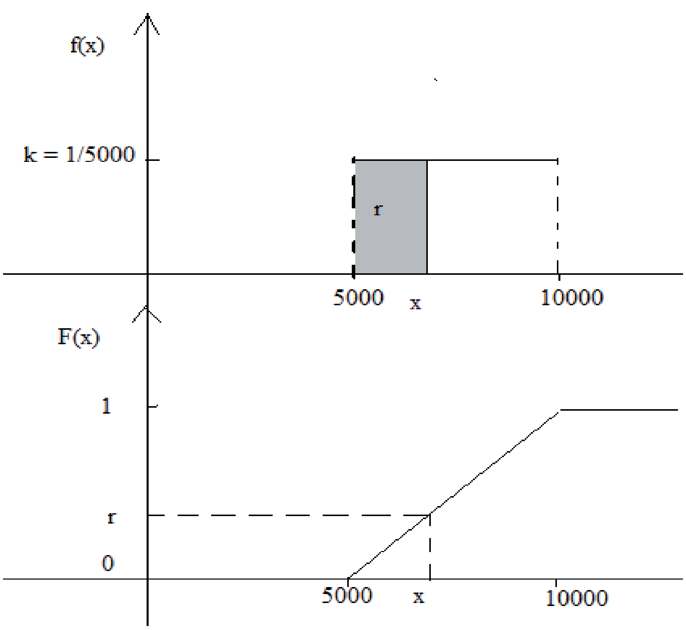
Si la variable considerada es discreta, el procedimiento para encontrar la inversa resulta similar, aunque requiere trabajar con sumatorias en vez de integrales. Pero, además, cuando no se puede expresar con una fórmula cerrada la integral de la función de densidad de probabilidad, como por ejemplo ocurre con la distribución normal, el método de la transformada inversa no puede aplicarse. En tal caso, para generar observaciones aleatorias que sigan una distribución normal de media μ y desvío estándar σ, puede usarse la fórmula de aproximación $x = \frac{\sigma}{\sqrt{n/12}} \sum_{i=1}^{n} r_i + \left(\mu - \frac{n}{2} \frac{\sigma}{\sqrt{n/12}}\right)$. Como se aprecia para una observación normal habrá que generar *n* números aleatorios lo que se corresponde con la idea de que la suma de variables aleatorias independientes tiende a una normal para un número suficientemente grande de variables. La fórmula produce buena aproximación incluso con unos pocos números aleatorios, por ejemplo, para $n = 6$ o $n = 12$.

VIII.4 Técnicas de Montecarlo

Las técnicas de Montecarlo, que deben su nombre al casino del estado homónimo, son métodos que se aplican para reducir la varianza en las muestras de modo que, a su vez, esto reduzca la cantidad de observaciones necesarias para estimar adecuadamente un parámetro.

Por ejemplo, si se conoce que cierta componente de un equipo electrónico tiene una duración que se distribuye uniformemente entre las 5000 horas y las 10000 horas de uso, se pueden simular duraciones de componentes para constituir una muestra con la que evaluar la duración promedio sin necesidad de apelar al cálculo analítico. Si se utiliza el método de la transformación inversa para obtener una expresión que genere la muestra, hay que hallar primero el valor de k que hace de densidad de probabilidad a la función $f(x) = \begin{cases} k & 5000 \leq x \leq 10000 \\ 0 & en\ otra\ parte \end{cases}$ En efecto, es fácil calcular k a partir del cumplimiento de la igualdad $1 = \int_{-\infty}^{\infty} k dx = k \int_{5000}^{10000} dx = kx|_{5000}^{10000} = k(10000 - 5000) = k\,5000$ con lo cual resulta $k = \frac{1}{5000}$. Si ahora se halla la función de distribución acumulada para el intervalo $5000 \leq x \leq 10000$, $F(x) = \int_{5000}^{x} \frac{1}{5000} dt = \frac{1}{5000} t|_{5000}^{x} = \frac{1}{5000}x - 1$. Los pasos i) y ii) del procedimiento antes descripto para hallar la transformada inversa se han completado y se los representa en la Gráfica 2

Gráfica 2



Ahora resta calcular la transformación inversa despejando x desde $r = \frac{1}{5000}x - 1$. Resulta $x = 5000(r + 1)$

Con la fórmula hallada podemos aplicar la llamada técnica de Montecarlo aproximada para calcular la media de duración de las componentes. A efecto de elegir una muestra aleatoria de diez duraciones, se toman diez números aleatorios r de tres dígitos decimales y se calculan las observaciones aleatorias correspondientes según se ve en la Tabla 2

Tabla 2

i	r_i	$x_i = 5000(r+1)$
1	0.506	7530
2	0.649	8245
3	0.941	9705
4	0.664	8320
5	0.277	6385
6	0.375	6875
7	0.637	8185
8	0.673	8365
9	0.497	7485
10	0.434	7170
		$\bar{x} = \frac{\sum x_i}{10} = 7826.5$

El ejemplo fue elegido para comparar el resultado experimental obtenido por medio de la simulación con el parámetro poblacional que en este caso puede calcularse haciendo $\mu = E(x) = \int_{5000}^{10000} x \frac{1}{5000} dx = \frac{1}{5000} \frac{x^2}{2} \Big| \begin{matrix} 10000 \\ 5000 \end{matrix} = 7500$ Como se aprecia hay una importante diferencia entre el valor real de la media y su estimación. Claramente, la ley de los grandes números nos autorizaría a continuar realizando observaciones hasta que la variabilidad de la estimación a obtener baje drásticamente y se obtenga un promedio mucho más cercano a la media poblacional real. Hay que observar que al optar por reducir la variancia de este modo se acepta incrementar grandemente el número de observaciones. Si el tamaño de las muestras es $n = 10$ el estimador $\bar{x}$ tendrá como varianza la razón $\frac{\sigma^2}{10}$ ya que, según la teoría estadística, se distribuye normalmente con media y desvío estándar de acuerdo a $\bar{x} \sim N(\mu, \frac{\sigma}{\sqrt{10}})$ Entonces para reducir en un dígito la varianza del estimador deberán agregarse 90 observaciones a las 10 ya tomadas. Si se hace esto, se dice que se aplica el método de Montecarlo aproximado o "crudo", y puede continuarse reduciendo la variabilidad incrementando en potencias de 10 el tamaño muestral.

Sin embargo, existen algunos procedimientos que permiten reducir la variabilidad de la estimación sin necesidad de incrementar la muestra. Veremos a continuación, y a título de ejemplo, uno de ellos. Se trata del llamado método de Montecarlo de números complementarios y, básicamente, consiste en armar una muestra paralela invirtiendo la suerte. Es decir, a cada número aleatorio elegido se le calcula su complemento a 1 por medio de $\acute{r} = 1 - r$ para obtener nuevas observaciones aleatorias $\acute{x}$ que compensan la suerte. En el caso del ejemplo se procede según se muestra en la Tabla 3.

Tabla 3

i	r_i	$x_i = 5000(r+1)$	$\acute{r} = 1 - r$	$\acute{x} = 5000(\acute{r} + 1)$
1	0.506	7530	0.494	7470
2	0.649	8245	0.351	6755
3	0.941	9705	0.059	5295
4	0.664	8320	0.336	6680
5	0.277	6385	0.723	8615
6	0.375	6875	0.625	8125
7	0.637	8185	0.363	6815
8	0.673	8365	0.327	6635
9	0.497	7485	0.503	7515
10	0.434	7170	0.566	7830
		$\bar{x} = \frac{\sum x_i}{10} = 7826.5$		$\bar{\acute{x}} = \frac{\sum \acute{x}_i}{10} = 7173.5$

Como era de esperar, por tratarse de una distribución uniforme, la compensación de cada observación por una complementaria debía hacer que la estimación, obtenida como el promedio de ambos promedios, fuera exacta: $\frac{\bar{x}+\bar{\acute{x}}}{2} = 7500$. Si bien no puede esperarse tal precisión para observaciones realizadas con distribuciones asimétricas, el ejemplo ilustra la efectividad del método de Montecarlo de números complementarios.

VIII.5 Modelos de simulación

Como se ha visto con anterioridad un sistema está conformado por entidades cuyos atributos pueden adquirir valores numéricos, y actividades que vinculan estas entidades y los modifican. El modelado de simulación procurará establecer la ley de probabilidad y los parámetros que caracterizan la actividad conjunta y total del sistema, basado en los patrones conocidos de cada actividad y en la forma en que cada una de ellas se vincule con las otras. Esta dinámica muchas veces puede querer estudiarse a través del tiempo, considerando los cambios probables de estado a estado. Respecto de esta simulación temporal, el valor de los parámetros de interés puede actualizarse a intervalos de tiempo fijos o en cada momento en que ocurre el siguiente evento. En ambos casos el modelo incluye también la simulación del paso del tiempo.

Para fijar ideas supongamos que en un proceso de fabricación de automóviles las carrocerías llegan a un robot de pintado siguiendo un patrón poisson- exponencial con una media $\lambda = 6\, carrocerías/hora$. El pintado se realiza en tiempos que se distribuyen en forma uniforme entre 4 y 8 minutos. Se desea calcular la cantidad esperada de carrocerías en este sistema.

Se trata de un sistema de colas, pero aquí, a diferencia del modelo MM1 que hemos estudiado antes, si bien las llegadas son del tipo poisson, los servicios se realizan en tiempos de probabilidad uniforme. Por supuesto cabría la posibilidad de deducir una fórmula para la probabilidad de estado del sistema, para su funcionamiento temporal y en definitiva para sus parámetros, uno de los cuales es la cantidad esperada de carrocerías en el sistema *L*. Sin embargo, estimar el valor del parámetro *L* por medio de una simulación puede resultar más rápido y práctico. Al considerar la relación entre las medias de poisson y de la exponencial ya estudiada, se tiene que $\frac{1}{\alpha} = E(t) = \frac{1}{E(k)} = \frac{1}{\lambda} = \frac{1}{6}\ horas/carrocería = 10\ min/carrocería$ y $\alpha = \lambda$. Por lo tanto los tiempos simulados entre llegadas de carrocerías pueden obtenerse mediante la transformación inversa de la exponencial $t_{ei} = \frac{\ln(1-r)}{-\frac{1}{10}} = -10\ln(1-r)$. De similar forma los tiempos simulados de pintado se obtienen mediante la inversión del patrón uniforme $t_{si} = 4(\acute{r}+1)$. Si se considera la actualización del valor de los parámetros del sistema cada vez que se produce un evento, la llegada de una carrocería al servicio o la culminación de un pintado, la simulación se realiza según el *criterio del evento siguiente* y el modelo queda definido por la Tabla 3 siguiente.

Tabla 3

Reloj	*r*	Tiempo hasta la siguiente entrada $t_{ei} = -10\ln(1-r)$	*r'*	Tiempo de siguiente salida $t_{si} = 4(\acute{r}+1)$.	Cantidad de carrocerías en el sistema
0.000	0.073	0.7580 (0.7580)	------	------------------------	0
0.7580	0.323	5.4260 (4.6680)	0.677	7.4460 (6.7080)	1
5.4260	0.115	6.6477 (1.2217)	------	7.4460	2
6.6477	0.722	19.4490 (12.8013)	------	7.4460	3
7.4460	------	19.4490	0.215	12.3060 (4.8600)	2
12.3060	------	19.4490	0.022	16.3940 (4.0880)	1
16.3940	------	19.4490	------	---------------------	0
19.4490	.617	30.7500 (11.3010)	0.500	25.4490 (6.0000)	1
25.4490	------	30.7500	------	---------------------	0
30.7500					

La primera columna de la Tabla 3 simula el paso del tiempo, que se actualiza cada vez que llega o sale una carrocería del sistema. Las cantidades temporales señaladas en negro son *puntos de regeneración* del sistema pues en ellos el estado actual no depende de ninguno anterior, no hay pendientes llegadas ni completamientos de servicio y hay que generar un número aleatorio por cada actividad futura. El tiempo que media entre dos puntos de regeneración consecutivos se denomina *ciclo*. En principio, para estimar la cantidad esperada de carrocerías en el sistema pueden contabilizarse los tiempos en que hubo 0, 1, 2 y 3 carrocerías y efectuar un promedio pesado.

Hubo 0 clientes en el sistema en los intervalos temporales [0.0000, 0.7580], [16.3940, 19.4490] y [25.4490, 30.7500]. Con esto durante $0.7580 + 3.0550 + 5.3010 = 9.1140$ minutos no hubo clientes. Un cliente hubo en los lapsos [0.7580, 5.4260], [12.3060, 16.3940] y [19.4490, 25.4490]. Es decir; durante 14.7560 minutos hubo 1 cliente. De forma similar calculamos que hubo 2 clientes durante 6.0870 minutos y 3 clientes durante 0.7983 minutos. Entonces la estimación del número esperado de carrocerías en el sistema en base a lo simulación del sistema realizada es:

$\hat{L} = \frac{0\times9.1140+1\times14.7560+2\times6.0870+3\times0.7983}{30.7500} = 0.95$ carrocerías

Este valor indica que seguramente el sistema permanece una parte del tiempo ocioso. Para estimar la probabilidad de que el sistema esté vacío podemos hacer $\widehat{p_0} = \frac{tiempo\ con\ 0\ clientes}{tiempo\ total} = \frac{9.1140}{30.7500} \approx 0.30$

Con todo, la cantidad de observaciones incluidas en la muestra representan un ejemplo de escritorio del procedimiento. En realidad, si se aplica la técnica de Montecarlo aproximada habrá que generar observaciones hasta que la estimación de interés se estabilice alrededor de un cierto valor. Tal cosa se logra cuando para la i-ésima observación aleatoria incorporada, se calcula la respectiva estimación $\hat{L}_i$ y se la compara con la anterior $\hat{L}_{i-1}$, por el criterio de variación absoluta $\left|\hat{L}_i - \hat{L}_{i-1}\right| < \varepsilon$, o por el más exigente criterio de variación relativa $\left|\frac{\hat{L}_i-\hat{L}_{i-1}}{\hat{L}_i}\right| < \varepsilon$, haciendo $\varepsilon > 0$ y tan pequeño como se considere necesario. Por supuesto, esto puede requerir, en sistmas complejos, realizar un gran número de observaciones para estimar sus parámetros, pero a favor del crecimiento en la rapidez de cálculo y en la capacidad de almacenamiento computacional, cada vez se torna más precisa la aplicación de esta técnica de simulación. Alternativamente cada ciclo entre puntos de regeneración puede tomarse como una muestra distinta y considerar así varias veces, no solo el estado del sistema cuando está en un régimen estable de funcionamiento, sino también aquella parte transitoria transcurrida hasta alcanzar la estabilidad. Pueden promediarse entonces las estimaciones realizadas por ciclo para obtener una nueva, más robusta en términos estadísticos.

Ejercicios resueltos

Ejercicio N° 1- Utilizar el método congruencial para generar las siguientes sucesiones de números aleatorios:

a) Una sucesión de 10 números aleatorios de un dígito tales que : $X_{n+1} \equiv (X_n + 3)(\mathrm{mod}10)$ y $X_0 = 2$

b) Una sucesión de 8 números aleatorios entre 0 y 7 tales que: $X_{n+1} \equiv (5X_n +1)\mathrm{mod}8$ y $X_0 = 1$.

c) Una sucesión de cinco números aleatorios de dos dígitos tales que: $X_{n+1} = (61X_n + 27)\ \mathrm{mod}100$ y $X_0 = 10$

a)

n	X_n	$X_n + 3$	$(X_n + 3)/10$	X_{n+1}
1	2	5	0 + 5/10	5
2	5	8	0 + 8/10	8
3	8	11	11/10 = 1+1/10	1
4	1	4	4/10 + 0	4
5	4	7	0 + 7/10	7
6	7	10	1 + 0/10	0
7	0	3	0 + 3/10	3
8	3	6	0 + 6/10	6
9	6	9	0 + 9/10	9
10	9	12	1 + 2/10	2

b)

n	X_n	$5X_n + 1$	$(5X_n + 1)/8$	X_{n+1}
1	1	6	0 + 6/10	6
2	6	31	31/8 =3 + 7/8	7
3	7	36	36/8 = 4 + 4/8	4
4	4	21	21/8 = 2 + 5/8	5
5	5	26	26/8 = 3 + 2/8	2
6	2	11	11/8 = 1 + 3/8	3
7	3	16	16/8 = 2 + 0/8	0
8	0	1	0 + 1/8	1

c)

n	X_n	$61X_n + 27$	$(61X_n + 27)/100$	X_{n+1}
1	10	637	637/100 = 6 + 37/100	37
2	37	2284	2284/100 = 22 + 84/100	84
3	84	5151	5151/100 = 51 + 51/100	51
4	51	3138	3138/100 = 31 + 38/100	38
5	38	2345	2345/100 = 23 + 45/100	45

Ejercicio N° 2- Un juego de dados requiere que el jugador lance dos dados una o más veces hasta que se llegue a una decisión de si pierde o gana. Gana si la primera tirada suma 7 u 11, o si la primera suma es 4, 5, 6, 8, 9, 10 y sale la misma suma antes de que aparezca una suma de 7. Por el contrario, pierde si el resultado de la primera tirada suma 2, 3 o 12, o si la primera suma es 4, 5, 6, 8, 9, 10 y aparece una suma de 7 antes de que la primera tirada vuelva a salir. Simule cinco jugadas de este juego para iniciar el proceso de estimar la probabilidad de ganar.

Dado => 6 caras => 1/6 $\therefore$ $0 \leq r < 0{,}167 \rightarrow$ Cara 1
$0{,}167 \leq r < 0{,}333 \rightarrow$ Cara 2
$0{,}333 \leq r < 0{,}5 \rightarrow$ Cara 3
$0{,}5 \leq r < 0{,}667 \rightarrow$ Cara 4
$0{,}667 \leq r < 0{,}833 \rightarrow$ Cara 5
$0{,}833 \leq r < \leq 1 \rightarrow$ Cara 6

n	r_1	r_2	v_1	v_2	Suma	Acción	Resultado/ Partida N°	$\hat{p}$
1	0,598	0,147	4	1	5	Sigue	---	0/1
2	0,583	0,359	4	3	7	Pierde	Perdió/1	
3	0,904	0,087	6	1	7	Gana	Ganó/2	1/2
4	0,902	0,223	6	2	8	Sigue	---	2/3
5	0,193	0,621	2	4	6	Sigue	---	
6	0,947	0,626	6	4	10	Sigue	---	
7	0,676	0,066	5	1	6	Sigue	---	
8	0,752	0,839	5	6	11	Sigue	---	
9	0,342	0,796	3	5	8	Gana	Ganó/3	
10	0,556	0,164	4	1	5	Sigue	---	3/4
11	0,713	0,718	5	5	10	Sigue	---	
12	0,561	0,041	4	1	5	Gana	Ganó/4	
13	0,365	0,802	3	5	8	Sigue	---	3/5
14	0,583	0,160	4	1	5	Sigue	---	
15	0,954	0,434	6	3	9	Sigue	---	
16	0,846	0,159	6	1	7	Pierde	Perdió/5	

El procedimiento debe continuar hasta que la estimación $\hat{p}$ se estabilice en dos dígitos decimales exactos.

Ejercicios propuestos

Ejercicio N° 3- Utilice los números aleatorios de un dígito 5, 2, 4, 9, 7 para generar observaciones aleatorias en cada una de las siguientes situaciones:

a) La tirada de una moneda legal
b) La tirada de un dado

c) El color de un semáforo que permanece el 40% del tiempo en verde , el 10% en amarillo y el resto en rojo.

Ejercicio N° 4-

a) Generar 5 observaciones aleatorias a partir de una distribución uniforme entre - 10 y 40.
b) Obtener tres observaciones aleatorias a partir de un a distribución triangular cuya función de densidad de probabilidad es:
$f(x) = 2x$ si $0 \leq x \leq 1$ y 0 en otra parte.

c) Generar tres observaciones aleatorias a partir de la función de densidad de probabilidad siguiente:
$f(x) = \frac{1}{4}(x+1)^3$ si $-1 \leq x \leq 1$ y 0 en otro caso.

Ejercicio N° 5- Generar tres observaciones aleatorias a partir de cada una de las siguientes distribuciones de probabilidad:

a) La variable aleatoria x tiene P(x=0) = ½ y si $x \neq 0$ se distribuye uniformemente entre -5 y 15.
b) La función de densidad de probabilidad es x-1 si $1 \leq x \leq 2$ y es 3-x si $2 \leq x \leq 3$ con 0 en otra parte.
c) La distribución geométrica con parámetro p = 1/3 de forma que
$P(x = k) = 1/3(2/3)^{k-1}$ si k= 1, 2, … y 0 en otro caso.

Ejercicio N° 6- Generar tres observaciones aleatorias a partir de :

a) una distribución normal de media 10 y desvío estándar 5 (utilizar n=3 para cada observación)
b) una distribución exponencial con media 4.

Ejercicio N° 7- El clima se puede considerar un sistema estocástico, por que evoluciona de una manera probabilística de un día a otro. Su pongamos que para cierto lugar este comportamiento probabilístico satisface la siguiente descripción:

a) La probabilidad de lluvia para mañana es de 0.6 si hoy llueve.
b) La probabilidad de un día despejado (sin lluvia) para mañana es de 0.8 si hoy está despejado.

Simular el comportamiento del clima durante 10 días, comenzando con un día que sigue a uno despejado.

Ejercicio N° 8- Considerar el modelo de colas M/M/1. Suponer que el promedio de arribos es de 3 cl/h y el de servicios es de 5 cl/h. Diseñar un procedimiento de simulación para estimar el tiempo promedio en cola de un cliente.

Ejercicio N° 9- Considerar la distribución de probabilidad cuya función de densidad es:

$$f(x) = \begin{cases} \dfrac{1}{x^2} & si \quad 1 \leq x \leq \infty \\ 0 & deotra \quad manera \end{cases}$$

a) Utilice el método de Monte Carlo crudo para evaluar su media.

b) Utilice el método de Monte Carlo de números complementarios con igual fin.
c) Compare los resultados obtenidos con los de la media teórica.

Ejercicio N° 10- Un producto fabricado por una compañía requiere que se perforen cilindros en un bloque metálico y que se inserten pistones en ellos. Es necesario que los pistones tengan un radio de 1.0000 cm por lo menos y, hasta donde sea posible, muy poco más grande. La distribución de probabilidad de cuánto medirá el radio del pistón tiene una función de densidad

$$f_1(x) = \begin{cases} 400e^{-400(x-1.0000)} & si \quad x \geq 1.0000 \\ 0 & de \quad otra \quad manera \end{cases}$$

De igual manera, la distribución de probabilidad de la medida de los cilindros tiene la función de densidad

$$f_2(x) = \begin{cases} 100\, si & 1.0000 \leq x \leq 1.0100 \\ 0 & de \quad otra \quad manera \end{cases}$$

Realizar un experimento simulado para estimar la probabilidad de interferencia por ser mayor el diámetro del pistón al del cilindro.

Ejercicio N° 11- Una compañía tiene un problema de mantenimiento con una compleja pieza de equipo. Este equipo contiene 4 tubos de vacío que son la causa del problema. Los tubos fallan con bastante frecuencia, forzando a que se apague el equipo mientras se reemplazan. La práctica actual es reemplazar los tubos solo cuando fallan, pero hay un propuesta para reemplazar los cuatro siempre que falle uno de ellos, para reducir la frecuencia con la que debe apagarse el equipo. El objetivo es comparar estas dos alternativas en cuanto a su costo. Los datos son los siguientes. Para cada tubo, el tiempo de operación hasta que falla tiene una distribución uniforme de 1.000 a 2.000 horas. El equipo debe apagarse durante una hora para cambiar un tubo y durante dos horas para cambiar los cuatro. El costo total asociado con la maquina apagada y reemplazo de los tubos es \$100/hora más \$20 por cada tubo nuevo.
Comenzando con los 4 tubos nuevos, simule la operación de las dos políticas durante 5000 horas de tiempo simulado.

Ejercicio N° 12- Simular la actividad de un sistema de colas con entradas y salidas tipo Poisson-Exponencial λ= 2 (entradas) μ=4(salidas). Diseñar el procedimiento para calcular el tiempo promedio de permanencia en el sistema de un cliente.

Ejercicio N° 13- Se tienen cuatro bolillas numeradas 1, 2, 3 y 4 dentro de un bolillero. Se pretende utilizarlas para simular:

a) La aparición del palo oro, espada, basto o copa al sacar una carta de un mazo de truco
b) La aparición de cara o ceca al tirar una moneda

Dar la regla de la simulación en ambos casos.

Ejercicio N° 14- Una embotelladora de gaseosas de un litro realiza la operación de colocar la tapa de metal en tiempos que se distribuyen uniformemente entre 5 y 10 segundos. Las botellas llegan al recipiente de acceso a la tapadora transportadas por una cinta y siguiendo un proceso de Poisson a razón media de 400 por hora. Construir una tabla de simulación para estimar el tiempo promedio que una botella tarda desde que entra al recipiente hasta que sale tapada de la tapadora.

Ejercicio N° 15- La empresa LIMPIAX acaba de ganar la licitación para recoger la basura en cierto distrito del conurbano bonaerense y se dispone ahora a invertir en la compra de camiones para la recolección. El pliego de la licitación establece que deberá recogerse por lo menos el 95% de la basura generada diariamente para quedar libre de quitas y multas por incumplimiento de servicio. Por esta razón la empresa quiere asegurarse de que su flota de camiones alcance para cumplir esta norma en el 90 % o más de los días. Para determinar que cantidad de camiones adquirir de acuerdo a esto, posee dos informaciones relevantes: a) cada camión posee una cantidad máxima de carga de 15 toneladas y puede hacer solo un viaje durante el horario de recolección y b) las estadísticas de la cantidad de toneladas diarias de basura recolectada en los meses pasados son:

Toneladas de basura recolectada/día	Probabilidad
10	0.10
20	0.22
30	0.25
40	0.20
50	0.12
60	0.07
70	0.04

Desarrollar un esquema de simulación que permita evaluar, de acuerdo a esta información, la cantidad de camiones necesarios para alcanzar los estándares de calidad de servicio fijados.

Ejercicio N° 16- En área céntrica los clientes llegan al cajero automático instalado según tiempos distribuidos exponencialmente a razón media de 15 cl/hora. Los tiempos de uso del cajero por parte de cada cliente se distribuyen en forma uniforme entre 3 y 5 minutos. Desarrollar un modelo de simulación para evaluar la probabilidad de que el cajero esté vacío.