

UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ENGENHARIA DE SÃO CARLOS
ÁREA DE ENGENHARIA DE PRODUÇÃO

INTRODUÇÃO À SIMULAÇÃO

João Vitor Moccellin

JULHO/1999



INTRODUÇÃO À SIMULAÇÃO

Dentro da Pesquisa Operacional, a Simulação também envolve a construção de um modelo que é, por natureza, em grande parte matemático. Em lugar de descrever diretamente o comportamento geral do sistema, o *modelo de simulação* descreve a operação do sistema em termos dos *eventos individuais* dos componentes individuais do sistema. Em particular, o sistema é subdividido em elementos cujos comportamentos possam ser preditos, pelo menos em termos das distribuições de probabilidade, para cada um dos vários possíveis estados do sistema e suas entradas. As inter-relações entre os elementos também são construídas dentro do modelo. Assim, a simulação fornece meios de dividir o trabalho de construção do modelo em partes componentes menores (onde pode ser possível formular cada uma destas partes segundo métodos descritos em outros capítulos) e, então, combinar estas partes em sua ordem natural e permitir ao computador apresentar o efeito da interação de umas com as outras. Depois de construído o modelo, ele é ativado (pela geração de dados de entrada) para simular a operação real do sistema e registrar seu comportamento agregado. Repetindo isso para as várias configurações alternativas para o projeto e políticas de operação do sistema, e comparando seus desempenhos, poderão ser identificadas as configurações mais promissoras. Por causa do erro estatístico, é impossível garantir que a configuração que resulte no melhor desempenho simulado seja, de fato, a ótima, porém seria, pelo menos, próxima da ótima, se o experimento simulado tiver sido adequadamente projetado.

Assim, a simulação é, nada mais nada menos, a técnica de fazer *experimentos amostrais* no modelo do sistema. Os experimentos são feitos no modelo, em lugar de no próprio sistema real, apenas porque o último seria muito inconveniente, dispendioso e demorado. Por outro lado, os experimentos simulados deveriam ser vistos, virtualmente, como em nada diferentes dos experimentos estatísticos comuns, de forma que eles também deveriam estar baseados em sólida teoria estatística. Mesmo que os experimentos simulados sejam, usualmente, executados num computador digital, isto apenas ocorre por causa da vasta quantidade de dados que são gerados e processados, e não por qualquer outra relação inerente.

1 - EXEMPLOS ILUSTRATIVOS

Exemplo 1

Suponhamos que nos seja dada a oportunidade de disputar um jogo em que seria jogada, repetidamente, um moeda justa, até que a *diferença* entre o número de ocorrência de cara e coroa fosse *três*. Deveríamos pagar \$ 1,00 por jogada da moeda, porém receberíamos \$ 8,00 ao final de cada rodada do jogo. Não é permitido o abandono do jogo durante uma rodada. Assim, ganharemos dinheiro se o número de jogadas necessárias for menor que oito, mas perderemos dinheiro se forem necessárias mais de oito jogadas. Como decidiríamos nossa participação ou não no jogo?

Muitas pessoas baseariam esta decisão em simulação, embora, provavelmente, não a chamassem por esse nome. (Também existe uma solução analítica para este jogo, porém não é uma solução particularmente elementar). Neste caso, a simulação

não seria nada mais que jogar o jogo muitas vezes, sozinho, até que se torne claro se vale a pena jogar. Meia hora que se perca, jogando repetidamente a moeda e registrando ganhos e perdas, pode ser suficiente.

Como este experimento simulado seria executado num computador digital? Embora o computador não possa jogar moedas, ele pode gerar números. Por isso, ele geraria (ou seria dada) uma seqüência de dígitos aleatórios, cada um dos quais correspondendo à jogada de uma moeda. (A geração de números aleatórios é discutida na seção 2). A distribuição de probabilidade para o resultado de uma jogada é que a probabilidade de dar cara é $1/2$ e a probabilidade de dar coroa é $1/2$, enquanto que existem 10 valores possíveis de um dígito aleatório, cada um tendo uma probabilidade de $1/10$. Portanto, a cinco destes valores (digamos 0, 1, 2, 3, 4) seria atribuída uma associação a cara e aos outros cinco (digamos 5, 6, 7, 8, 9) a coroa. Assim, o computador simularia a **rodada** do jogo, examinando cada dígito aleatório gerado, rotulando-o de cara ou coroa, de acordo com seu valor. Continuaria desta maneira registrando o resultado de cada jogada simulada do jogo, até quando se desejasse.

Para ilustrarmos a abordagem do computador a este experimento simulado, suponhamos que o computador tenha gerado a seguinte seqüência de dígitos:

8, 1, 3, 7, 2, 7, 1, 6, 5, 5, 7, 9, 0, 0, 3, 4, 3, 5, 6, 8, 5,
8, 9, 4, 8, 0, 4, 8, 6, 5, 3, 5, 9, 2, 5, 7, 9, 7, 2, 9, 3, 9,
8, 5, 8, 9, 2, 5, 7, 6, 9, 7, 6, 0, 7, 3, 9, 8, 2, 7, 1, 0, 3,
2, 6, 2, 7, 1, 3, 7, 0, 4, 4, 1, 8, 3, 2, 1, 3, 9, 5, 9, 0, 5,
0, 3, 8, 7, 8, 9, 5, 4, 0, 8, 3, 8, 0, 1.

Assim, denotando-se cara por H e coroa por T , a primeira jogada simulada do jogo seria $THHTHTHTTTT$, requerendo 11 jogadas simuladas de uma moeda. As jogadas simuladas subseqüentes do jogo precisaram de 5, 5, 9, 7, 7, 5, 3, 17, 5, 5, 3, 9 e 7 jogadas simuladas, respectivamente, para uma média geral de 7. Esta média amostral parece indicar que, em média, ganharíamos cerca de \$ 1,00 por rodada do jogo. Portanto, se não tivermos uma aversão relativamente alta por risco, deveríamos optar por disputar o jogo, preferivelmente um grande número de vezes. No entanto, cuidado! Um dos erros comuns no uso de simulação é que as conclusões são baseadas em amostras por demais pequenas, porque a análise estatística fora inadequada ou totalmente ausente. Neste caso, o desvio-padrão da amostra é de 3,67 de modo que o desvio-padrão estimado da média amostral é de $3,67/\sqrt{14} \approx 0,98$. Portanto, mesmo que seja suposto que a distribuição de probabilidade do número de jogadas necessárias para uma rodada do jogo seja uma distribuição normal (o que é uma suposição grosseira, porque a verdadeira distribuição é assimétrica), qualquer intervalo de confiança razoável para a verdadeira média desta distribuição se estenderia bem acima de 8. Conseqüentemente, é necessário uma amostra bem maior para se tirar uma conclusão válida a um nível razoável de significância estatística. Infelizmente, como o desvio-padrão de uma média amostral é inversamente proporcional à *raiz quadrada* do tamanho da amostra, é necessário um grande aumento no tamanho da amostra para se chegar a um aumento relativamente pequeno na precisão da estimativa da verdadeira média. Neste caso, parece que talvez um adicional de 100 rodadas simuladas do jogo seria adequado.

Acontece, então, que a verdadeira média do número de jogadas requeridas para uma rodada deste jogo é 9. Portanto, a longo prazo, realmente perderíamos cerca de \$ 1,00 por rodada do jogo.

Exemplo 2

Consideremos o modelo *básico* do servidor único da teoria das filas com entrada de Poisson e tempos de serviço exponenciais. Embora esse modelo já tenha sido resolvido analiticamente, será ilustrativo considerar como usar a simulação para estudá-lo.

Para resumirmos a operação física do sistema, os clientes que chegam entram na fila, são atendidos pelo servidor e então saem. Deste modo, é necessário que o modelo de simulação descreva e sincronize a chegada de clientes e o atendimento de clientes. Os dois métodos para manipular uma tal sincronização num computador são *incremento de tempo fixo* e *incremento para o evento seguinte*. Eles serão descritos um de cada vez.

Com o **incremento de tempo fixo** é usado o seguinte procedimento de dois passos, começando com o sistema no seu estado inicial num dado ponto no tempo. Primeiro, *avance no tempo* por uma *quantidade fixa* pequena; adicione 1 num registro que serve como o “relógio mestre” para o sistema registrar a passagem deste tempo.

Segundo, *atualize o sistema* determinando que eventos ocorreram durante essa unidade de tempo decorrida e qual é o estado do sistema resultante. Repetir estes dois passos por tantas unidades de tempo quantas desejadas.

Para o modelo da teoria das filas em consideração, apenas dois tipos de eventos podem ocorrer durante cada uma das unidades de tempo decorridas, ou seja, um ou mais clientes podem *chegar* e um ou mais clientes podem *acabar de ser atendidos*. Além disso, a probabilidade de duas ou mais chegadas ou de duas ou mais conclusões de serviços durante uma unidade de tempo será negligenciável para este modelo se a unidade de tempo for relativamente pequena. Por isso, os dois únicos eventos possíveis, durante uma tal unidade de tempo, que precisam ser investigados são a chegada de um cliente e a conclusão do serviço para um cliente. Cada um destes eventos tem uma probabilidade conhecida. Por isso, exatamente como no Exemplo 1, para simular se um evento ocorre ou não, o computador precisaria apenas gerar um número aleatório. Por exemplo, suponhamos que a probabilidade de que um cliente chegasse durante uma unidade de tempo decorrida fosse de 0,007. O computador precisaria gerar 1 dos 1.000 números de três dígitos possíveis (000, 001, ..., 999), aleatoriamente. Associando-se sete dos números possíveis (digamos, 000, ..., 006) ao evento ocorrendo, e os números restantes aos eventos não ocorrendo, o número aleatório gerado determinará o resultado simulado real para aquela unidade de tempo. Se um cliente estivesse em processo de ser atendido, o computador seria programado para usar esse mesmo método para determinar se ocorre ou não uma conclusão de serviço simulado durante a unidade de tempo decorrida, dada a probabilidade de uma tal conclusão. Entretanto, se nenhum cliente estivesse sendo servido, o computador decidiria automaticamente que nenhuma conclusão de serviço teria ocorrido durante a unidade de tempo decorrida. Para implementar isso, o computador usaria um indicador, ao qual seria dado um ou dois valores numéricos, dependendo de se o

servidor estivesse ocupado atendendo a um cliente ou não. Similarmente, seria usado um contador para registrar o número atual de clientes na fila à espera de ser servido.

Portanto, atualizar o sistema de uma unidade de tempo decorrida significa atualizar os números que deveriam ser inseridos no indicador e no contador. Ao mesmo tempo o computador registraria a informação desejada sobre o comportamento agregado do sistema durante esta unidade de tempo. Por exemplo, ele poderia registrar o *número de clientes* no sistema de fila e o *tempo de espera* de qualquer cliente que acabou de completar sua espera. Se fosse suficiente estimar apenas a média, em vez da distribuição de probabilidade de cada uma destas variáveis aleatórias, o computador simplesmente adicionaria o valor (se algum) da unidade de tempo atual, numa soma cumulativa. As médias amostrais seriam obtidas depois que a realização da simulação fosse concluída, dividindo-se estas somas pelo tempo decorrido total e pelo número total de clientes, respectivamente.

A Figura 1 apresenta o Diagrama de Blocos correspondente a esse exemplo de sistema de filas, utilizando o método do **incremento de tempo fixo**.

No diagrama são utilizadas as seguintes variáveis:

<i>Variáveis exógenas</i>	TC - tempo a decorrer entre duas chegadas consecutivas de uma unidade ao sistema;
	TS - Tempo de serviço para a unidade que entra no guichê para ser atendida
<i>Outras variáveis</i>	SOMA - somatória dos TC até a última chegada;
	TESP - tempo total de espera para as unidades na fila;
	STO - tempo total em que o guichê esta desocupado;
	FILA - comprimento da fila em cada período;
	N - número de unidades que já entraram no sistema;
	R - relógio interno;
FIM - número total de períodos em que se vai fazer a simulação.	

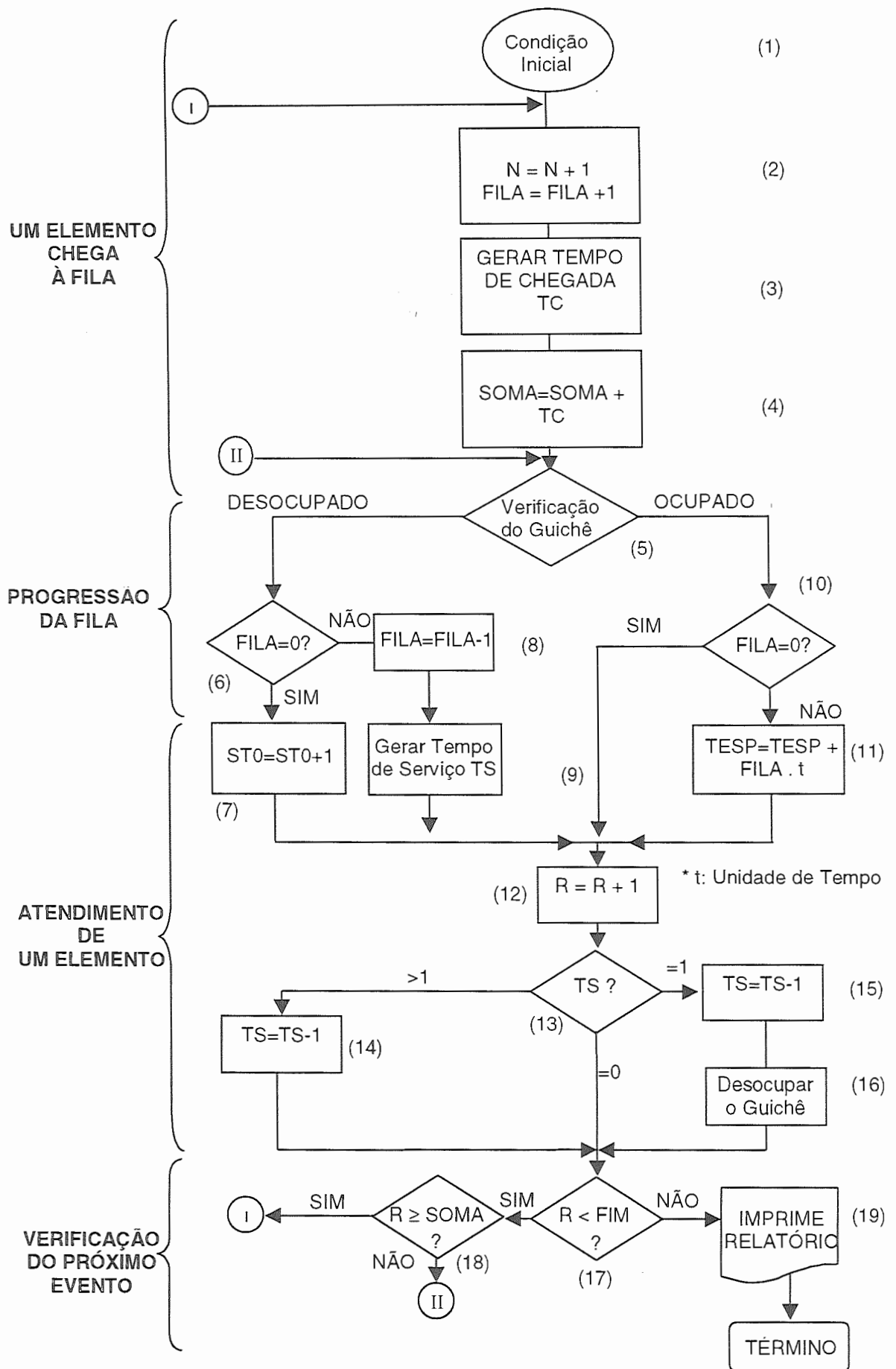


Figura 1 - Diagrama de blocos para o sistema de filas com 1 posto de serviço

O incremento para o evento seguinte difere do instrumento de tempo fixo na medida em que o relógio mestre é incrementado, a cada vez, por uma quantidade variável, em vez de por uma quantidade fixa. Conceitualmente, o procedimento do incremento para o evento seguinte é o de manter em execução o sistema simulado, sem interrupção, até que ocorra um evento, ponto em que o computador faz uma pausa momentânea para registrar a mudança no sistema. Para implementar esta idéia conceitual, o computador na verdade funciona acompanhando quando os poucos eventos simulados seguintes estiverem programados para ocorrer, pulando em tempo simulado para o primeiro destes eventos, e atualizando o sistema. Este ciclo é repetido tantas vezes quantas desejadas.

Para este exemplo, o computador precisa acompanhar dois eventos futuros, ou seja, quando o cliente seguinte chegar e quando o servidor terminar de atender seu atual cliente (se houver algum). Estes tempos são obtidos tomando-se uma observação aleatória da distribuição de probabilidade dos tempos entre chegadas e de serviços, respectivamente. Como anteriormente, o computador consegue uma tal observação aleatória gerando e usando um número aleatório. (A técnica para fazer isso será discutida na seção 2.) Portanto, cada vez que ocorra a conclusão de uma chegada ou serviço, o computador determinará primeiramente quanto tempo levará para que este evento ocorra novamente e, então, adicionará isto ao tempo atual do relógio. Se a conclusão do serviço não deixar nenhum cliente no sistema, então a geração do tempo até a conclusão do serviço seguinte será adiada até que ocorra a chegada seguinte. Esta soma é, então, armazenada num arquivo do computador. Para determinarmos que evento ocorrerá a seguir, o computador encontrará o mínimo dos tempos do relógio armazenados no arquivo. Este procedimento será ilustrado na seção 4 (Tabela 6) para um sistema de fila de um servidor único específico.

2 - FORMULAÇÃO E IMPLEMENTAÇÃO DE UM MODELO DE SIMULAÇÃO

2.1 - Construção do Modelo

O primeiro passo num estudo de simulação é desenvolver um modelo que represente o sistema a ser investigado. É evidente que isto requer que o analista se torne inteiramente familiarizado com as realidades operacionais do sistema e com os objetivos do estudo. Dado isto, o analista provavelmente irá tentar reduzir o sistema real a um diagrama de fluxo lógico. O sistema é, assim, desmembrado num conjunto de componentes, reunidos por um diagrama de fluxo mestre, onde os próprios componentes podem ser desmembrados em subcomponentes, e assim por diante. Finalmente, o sistema é decomposto num conjunto de elementos para os quais podem ser dadas as regras de operação. Estas regras de operação predizem os eventos que serão gerados pelos elementos correspondentes, talvez em termos de distribuições de probabilidade. Depois de especificados estes elementos, regras e ligações lógicas, o modelo deveria ser inteiramente testado, peça por peça. Isto pode ser feito, parcialmente, executando-se uma versão grosseira da simulação numa calculadora e verificando se cada entrada é recebida da fonte apropriada e cada resultado é aceitável para o submodelo seguinte. Entretanto, cada um dos componentes no modelo também

deveria ser testado isoladamente, para verificar se seu desempenho interno é razoavelmente consistente com a realidade.

Deveríamos enfatizar que, como qualquer modelo de pesquisa operacional, o modelo de simulação não precisa ser uma representação completamente realística do sistema real. Na verdade, parece que a maioria dos modelos de simulação peca por serem super-realistas em lugar de superidealizados.

Se o comportamento de um elemento não puder ser predito exatamente, dado o estado do sistema, será melhor usar observações aleatórias das distribuições de probabilidade envolvidas, que usar médias para simular este desempenho. Isto é verdade mesmo quando se está apenas interessado no desempenho agregado médio do sistema, porque a combinação de desempenhos médios para os elementos individuais pode resultar em alguma coisa longe da média para o sistema geral.

Uma questão que pode surgir quando da escolha de distribuições de probabilidade para o modelo é se usarem distribuições de frequência de dados históricos ou se procurar a distribuição de probabilidade teórica que melhor se ajuste a esses dados. A última alternativa é usualmente preferida, porque pareceria chegar mais perto da predição do desempenho futuro esperado do que reproduzir a idiosincrasia de um certo período do passado.

As Figuras 2 e 3 ilustram o processo que envolve a construção de um modelo de simulação.

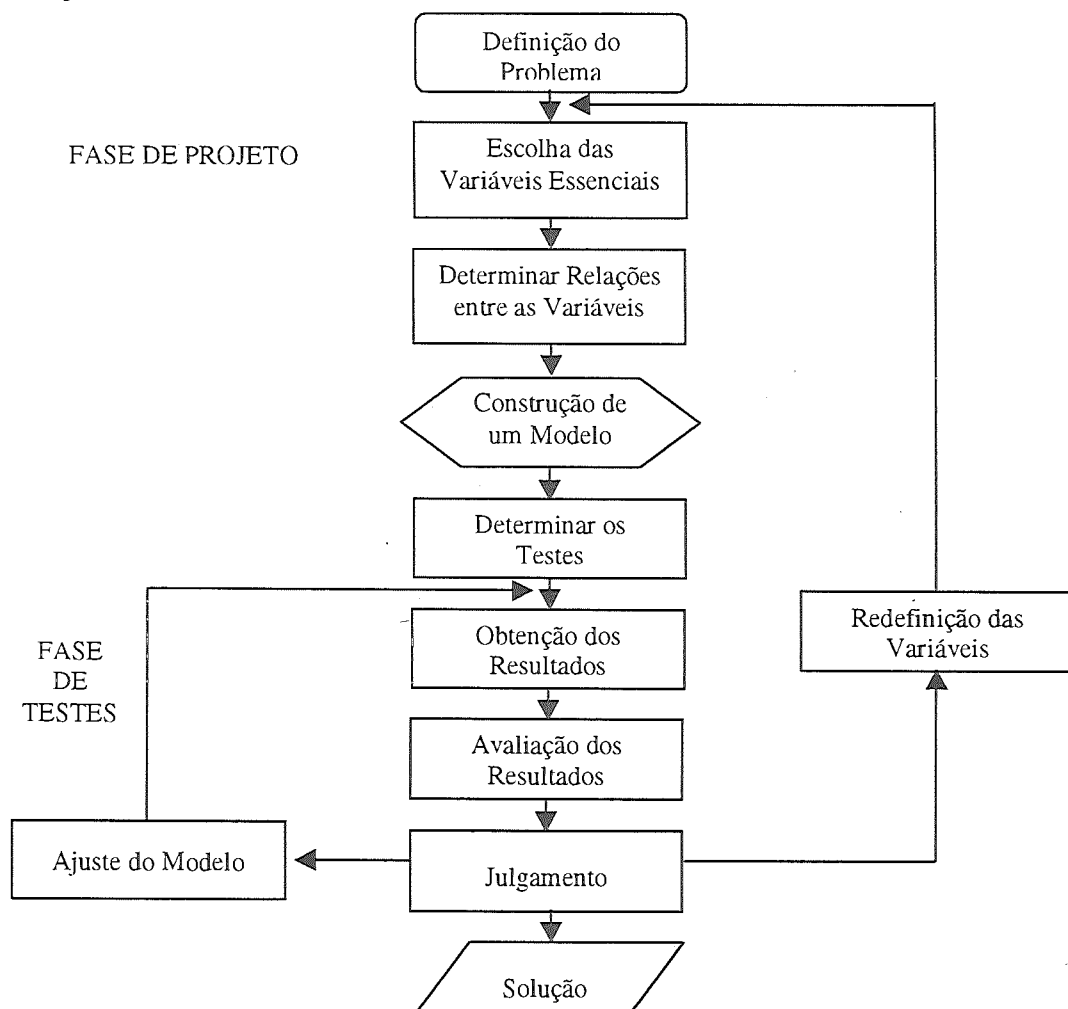


Figura 2 - Esquema da simulação

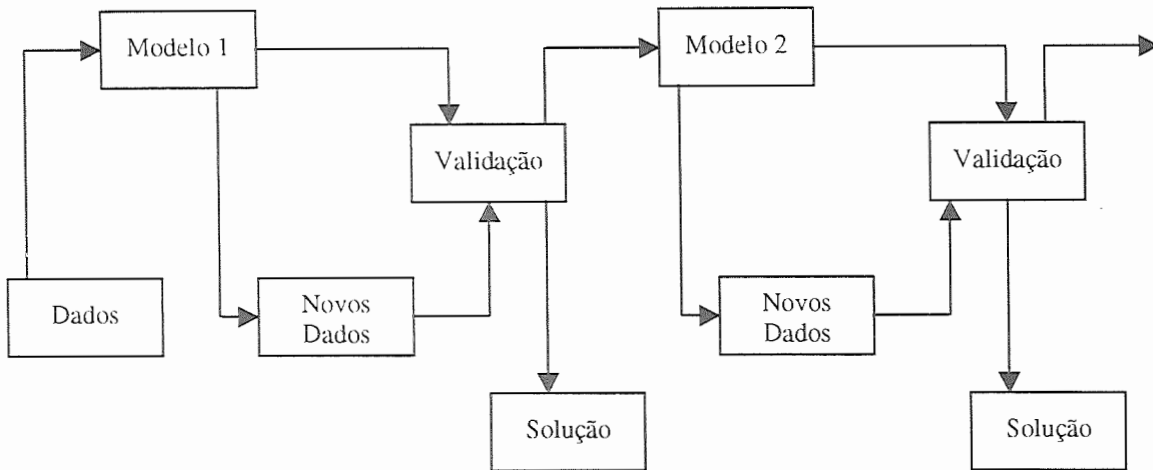


Figura 3 - Refinamento do modelo

2.2 - Geração de Números Aleatórios

A implementação de um modelo de simulação requer números aleatórios para se obterem observações aleatórias das distribuições de probabilidade. Um método para gerar tais números aleatórios é usar um dispositivo físico tal como um disco giratório ou um aleatorizador eletrônico.

Vários procedimentos estatísticos relativamente sofisticados foram propostos para testar se uma seqüência de números constitui uma amostra de números aleatórios ou não. Basicamente os requerimentos são que cada número sucessivo na seqüência tenha que ter uma probabilidade igual de assumir qualquer um dos valores possíveis, e que tenha que ser estatisticamente independente dos outros números na seqüência. Em outras palavras, os números precisam ser observações aleatórias de uma *distribuição uniforme* (discreta).

Existem diversos métodos para se fazer isso, dentre os quais os mais populares são os métodos congruenciais (aditivo, multiplicativo e misto). O *método congruencial misto* tornou-se provavelmente o mais amplamente usado, de forma que nos focalizaremos nesta abordagem.

O **método congruencial misto** gera uma *seqüência* de números aleatórios, calculando sempre o próximo número aleatório a partir do último obtido, dado um número aleatório inicial x_0 (chamado de *semente*). Em particular, ele calcula o $(n + 1)$ -ésimo número aleatório, x_{n+1} , a partir do n -ésimo número aleatório, x_n , usando a relação de recorrência

$$x_{n+1} = (ax_n + c) \text{ (módulo } m)$$

onde a , c e m são números inteiros positivos ($a < m$, $c < m$). Esta notação matemática significa que x_{n+1} é o *resto* quando $(ax_n + c)$ é dividido por m . Assim, os valores *possíveis* de x_{n+1} são $0, 1, \dots, m - 1$, de modo que m representa o número desejado de valores diferentes que poderiam ser gerados para os números aleatórios. Para

ilustrarmos, vamos supor que $m = 8$ (de modo que são desejados oito valores possíveis diferentes), $a = 5$, $c = 7$ e $x_0 = 4$. A seqüência de números aleatórios resultante é calculada na Tabela 1. (A seqüência não pode ser continuada mais adiante porque apenas começaria a repetir os números na mesma ordem.) Note-se que esta seqüência inclui cada um dos oito números possíveis, exatamente uma vez. Esta é uma propriedade desejável, porém não ocorre com algumas escolhas de a e c . (Tente-se $a = 4$, $c = 7$, $x_0 = 3$.) Felizmente, existem disponíveis regras para escolha dos valores de a e c , as quais garantirão esta propriedade. (Não existem duas restrições quanto à semente x_0 porque ela apenas afeta quando a seqüência se inicia e não na progressão dos números.)

Tabela 1 - Ilustração do método congruencial misto

n	x_n	$5x_n + 7$	$\frac{5x_n + 7}{8}$	x_{n+1}
0	4	27	$3 + \frac{3}{8}$	3
1	3	22	$2 + \frac{6}{8}$	6
2	6	37	$4 + \frac{5}{8}$	5
3	5	32	$4 + \frac{0}{8}$	0
4	0	7	$0 + \frac{7}{8}$	7
5	7	42	$5 + \frac{2}{8}$	2
6	2	17	$2 + \frac{1}{8}$	1
7	1	12	$1 + \frac{4}{8}$	4

Freqüentemente são desejados números aleatórios com um número relativamente pequeno de dígitos. Por exemplo, suponhamos que sejam desejados apenas três dígitos, de modo que os valores possíveis possam ser expressos como 000, 001, ..., 999. Num tal caso, o procedimento usual é usar $m = 2^b$ ou $m = 10^d$, de forma que um número extremamente grande de números aleatórios pode ser gerado antes que a seqüência comece a se repetir. Entretanto, exceto para propósitos de calcular o próximo número aleatório, todos, menos três dígitos de cada número aleatório, seriam descartados. Uma convenção é tomar os *últimos* três dígitos.

O *método congruencial multiplicativo* é apenas o caso especial do método congruencial misto, onde $c = 0$. O *método congruencial aditivo* também é similar, porém faz $a = 1$ e substitui c por algum número aleatório precedente, x_n , na seqüência, por exemplo, x_{n-1} (de modo que é necessária mais que uma semente para começar a calcular a seqüência).

Estritamente falando, os números gerados pelo computador não deveriam ser chamados de números aleatórios porque eles são predizíveis e reproduzíveis (o que às vezes é vantajoso). Por isso, às vezes lhes é dado o nome de *números pseudo-aleatórios*. Entretanto, o ponto importante é que eles desempenham satisfatoriamente o papel de números aleatórios na simulação, se o método usado para gerá-los for válido.

2.3 - Geração de Observações Aleatórias a partir de uma Distribuição de Probabilidade

Dada uma seqüência de números aleatórios, como se poderá gerar uma seqüência de observações aleatórias a partir de uma dada distribuição de probabilidade?

Para distribuições discretas simples, uma resposta é bastante evidente. Simplesmente se aloca os valores possíveis de um número aleatório aos vários números na distribuição de probabilidade, na proporção direta para as respectivas probabilidades daqueles números. Por exemplo, consideremos a distribuição de probabilidade do resultado de um lançamento de dois dados. É sabido que a probabilidade de se obter um 2 é $1/36$ (assim como a probabilidade de um 12), a probabilidade de um 3 é $2/36$, e assim por diante. Por isso, $1/36$ dos valores possíveis de um número aleatório deveria estar associado a obter um 2, $2/36$ dos valores com obter um 3, e assim por diante. Portanto, se forem usados números aleatórios de dois dígitos, 72 dos 100 valores seriam selecionados para serem considerados, de modo que um número aleatório seria rejeitado se assumisse qualquer um dos outros 28 valores. Então, a 2 dos 72 valores possíveis (digamos, 00 e 01) seria designada uma associação com a obtenção de um 2, a 4 deles (digamos, 02, 03, 04 e 05) seria designada a obtenção de 3, e assim por diante.

Para distribuições mais complicadas, a resposta ainda é essencialmente a mesma, embora o procedimento seja ligeiramente mais rebuscado. O *primeiro passo* é construir a função de distribuição cumulativa $F(x) = P\{X \leq x\}$, onde X é a variável aleatória envolvida. Isto pode ser feito *escrevendo-se a equação* para esta função, ou *representando graficamente a função*, ou *desenvolvendo a tabela* que dê o valor de x para valores uniformemente espaçados de $F(x)$ de 0 a 1. O *segundo passo* é gerar um número aleatório decimal entre 0 e 1. Isto é feito obtendo-se um número inteiro aleatório tendo o número desejado de dígitos (incluindo quaisquer zeros à esquerda) e, então, colocando uma vírgula decimal em frente a ele. O *passo final* é fazer $P\{X \leq x\}$ igual ao número decimal aleatório e resolver para x . Este valor de x é a observação aleatória desejada da distribuição de probabilidade. Este procedimento está ilustrado na Figura 4 para o caso em que a função de distribuição cumulativa é representada graficamente e em que acontece de o número decimal aleatório ser 0,5269.

Quando a distribuição de probabilidade dada for contínua, o procedimento acima descrito realmente aproxima esta distribuição contínua através de uma distribuição discreta cujos pontos irregularmente espaçados têm iguais probabilidades. Entretanto, isto não é particularmente grave porque a aproximação pode ser tão precisa quanto desejado usando-se um número suficientemente grande de dígitos para o número aleatório. Talvez o maior perigo seja que a aproximação venha a ser adequada em todos os pontos, exceto nas partes extremas da distribuição. Por exemplo,

suponhamos que sejam usados números aleatórios com três dígitos. Então, os valores de $P\{X \leq x\}$ que serão amostrados variariam de 0,000 a 0,999. Um refinamento ligeiro seria o de adicionar um 5 ao próximo lugar decimal de cada valor para centrar o intervalo, 0,0005 a 0,9995. Entretanto, pode acontecer que aquelas ocorrências raras em que o valor real assumido por X caia fora do intervalo permitido na simulação venham a ter um impacto crítico no sistema. Um refinamento, que retificaria isso, seria o de gerar um segundo número aleatório sempre que o primeiro fosse (para o caso dos números aleatórios de três dígitos) 000 ou 999, para seleccionar um valor de $P\{X \leq x\}$ dentro do intervalo entre 0,000000 e 0,000999, ou de 0,999000 a 0,999999.

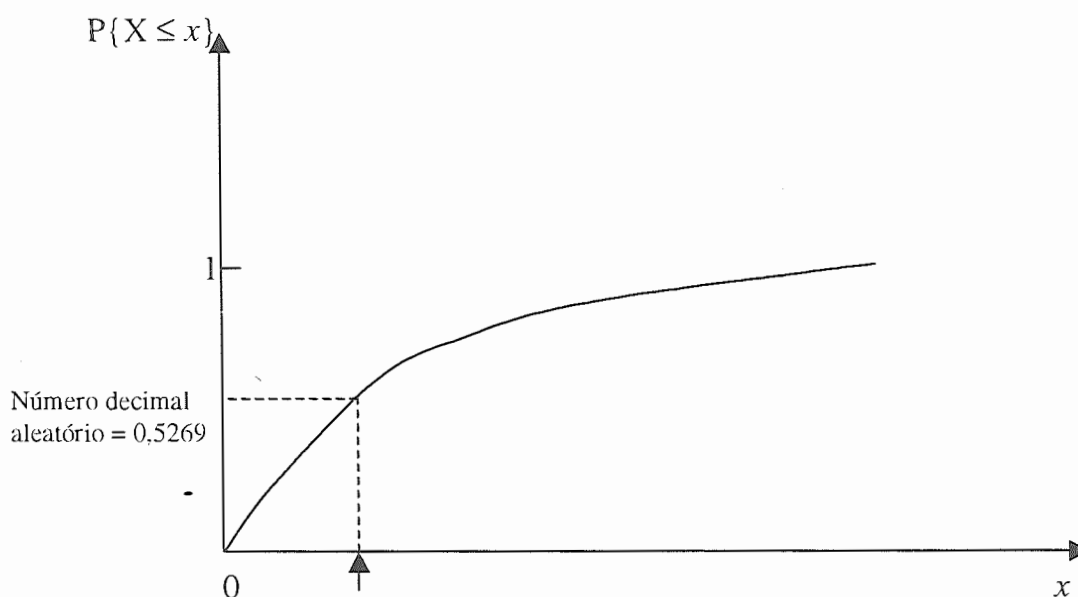


Figura 4 - Ilustração do procedimento para a obtenção de uma observação aleatória a partir de uma dada distribuição de probabilidade

A outra abordagem mencionada anteriormente requer que se *escreva a equação* para a função de distribuição cumulativa e, então, se resolva para o ponto em que esta função é igual ao número decimal aleatório. Esta abordagem conduz a uma solução explícita simples para diversas distribuições de probabilidade importantes, como é ilustrado abaixo para a distribuição exponencial. Para outros casos, seriam necessários métodos numéricos demorados para se obter a solução desta maneira. Felizmente, para certas distribuições de probabilidade importantes, foram desenvolvidas técnicas especiais para gerar eficientemente observações aleatórias a partir daquela distribuição particular. Duas das mais simples destas técnicas, as quais são aplicáveis à distribuição normal e à distribuição qui-quadrado, respectivamente, serão descritas no final desta seção.

Consideremos a *distribuição exponencial* que tenha a função de distribuição cumulativa

$$P\{X \leq x\} = 1 - e^{-\mu x}, \quad \text{para } x \geq 0,$$

onde $1/\mu$ é a média da distribuição. Aplicando a abordagem da equação discutida acima, fazamos esta função igual a um *número decimal aleatório* entre 0 e 1, denotado por r . Assim,

$$1 - e^{-\mu x} = r,$$

portanto,

$$e^{-\mu x} = 1 - r.$$

Portanto, tomando-se o logaritmo natural de ambos os lados,

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

de modo que

$$-\mu x = \ln(1 - r),$$

o qual resulta em

$$\left| x = \frac{\ln(1 - r)}{-\mu} \right|$$

como a observação aleatória desejada a partir da distribuição exponencial.

Como um número decimal aleatório entre 0 e 1 é exatamente uma observação aleatória a partir de uma *distribuição uniforme* (discreta) entre 0 e 1, $(1 - r)$ acima é, ele próprio, um número decimal aleatório. Por isso, para poupar uma subtração, é comum, na prática, simplesmente usar o número decimal aleatório *original*, r , diretamente no lugar de $(1 - r)$.

Uma outra distribuição de probabilidade importante e bastante usada em modelos de simulação é a distribuição de Poisson.

A distribuição de Poisson é caracterizada pela função

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

e representa a probabilidade de ocorrer x sucessos num certo intervalo de tempo t .

Para gerar valores com distribuição de Poisson, vamos utilizar a relação existente entre essa distribuição e a distribuição exponencial. Pela distribuição exponencial de média $1/\lambda$ são gerados intervalos de tempo t_1, t_2, \dots entre a ocorrência de 2 sucessos consecutivos e verifica-se quantos sucessos ocorreram num intervalo unitário $t = 1$ de tempo.

Esquemáticamente podemos representar essa geração como na Figura 5 que indica a escala de tempo e as ocorrências de sucessos.

No primeiro intervalo (com $t = 1$) pela geração dos intervalos de tempo $t_1, t_2, t_3,$ e t_4 houve a ocorrência de $x = 4$ sucessos, enquanto que no segundo houve $x = 3$ sucessos. Desse modo, o número x de sucessos que ocorrem no intervalo unitário de tempo tem distribuição Poisson com média λ .

Podemos gerar intervalos t_1, t_2, t_3, \dots com distribuição exponencial de média igual a $1/\lambda$ e acumular esses intervalos de tempo até que a sua soma ultrapasse 1. Desse modo, o valor x com distribuição de Poisson de média λ é determinado pela desigualdade

$$\sum_{i=1}^x t_i \leq 1 < \sum_{i=1}^{x+1} t_i$$

onde os valores t_i com distribuição exponencial são gerados pela expressão $t_i = (-1/\lambda) \ln r_i$.

A fórmula anterior pode ser reescrita na forma

$$\prod_{i=1}^x r_i \geq e^{-\lambda} > \prod_{i=1}^{x+1} r_i,$$

o que dá um método mais rápido para gerar valores x com distribuição de Poisson.

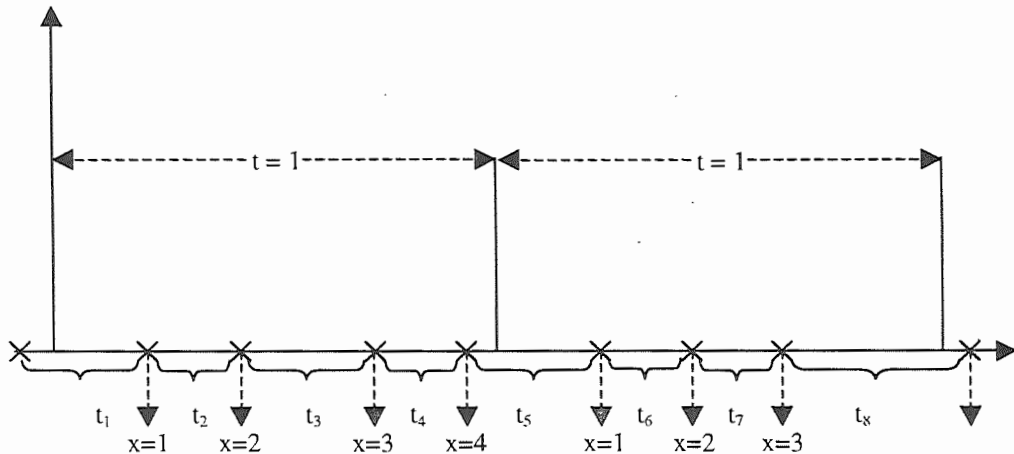


Figura 5 - Esquema de geração de eventos segundo Poisson

Uma técnica particularmente simples para gerar uma observação aleatória a partir de uma *distribuição normal* é obtida aplicando-se o teorema do limite central. Uma vez que um número decimal aleatório tem uma *distribuição uniforme* de 0 a 1, ele tem média $1/2$ e desvio padrão $1/\sqrt{12}$. Por isso esse teorema implica que a soma de n números decimais aleatórios tenha, aproximadamente, uma distribuição normal com média $n/2$ e desvio padrão $\sqrt{n/12}$. Portanto, se r_1, r_2, \dots, r_n forem uma amostra de números decimais aleatórios, então,

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

é uma observação aleatória a partir de uma distribuição aproximadamente normal com média μ e desvio padrão σ . Esta aproximação é excelente (exceto nas pontas da distribuição), mesmo com valores pequenos de n . Por isso, são usados valores de n de 5 a 10; $n = 12$ também é um valor conveniente porque elimina os termos de raiz quadrada da expressão acima.

Também foram desenvolvidas diversas outras técnicas aproximadas ou exatas para gerar observações aleatórias a partir de uma distribuição normal.

Um método simples para lidar com a *distribuição qui-quadrado* é usar o fato de que ela é obtida pela soma dos quadrados de variáveis aleatórias normais padronizadas. Assim, se y_1, y_2, \dots, y_n formarem uma amostra de observações aleatórias a partir de uma distribuição normal com média 0 e desvio-padrão 1, tal que pudesse ser obtida (aproximadamente) pela técnica acima descrita, então

$$\left(x = \sum_{i=1}^n y_i^2 \right)$$

será uma observação aleatória a partir de uma distribuição qui-quadrado com n graus de liberdade.

2.4 - Preparação de um Programa de Simulação

O propósito básico da maioria dos estudos de simulação é comparar alternativas. Por isso o programa de simulação tem que ser flexível o bastante para acomodar prontamente as alternativas que serão consideradas. Como, freqüentemente, é impossível prever exatamente que alternativas interessantes serão descobertas durante o curso do estudo, é essencial que sejam incluídas no programa flexibilidade e provisão para modificações rápidas e simples.

A maioria das instruções num programa de simulação é de operações lógicas, enquanto que o relativamente pequeno trabalho aritmético real necessário é, usualmente, de um tipo muito simples. Isto deveria ser refletido na escolha do computador e linguagem de programação a serem usados.

As considerações acima mencionadas realmente contribuíram como parte da motivação de uma importante descoberta, na arte de simulação, que ocorreu no início dos anos 60, ou seja, o desenvolvimento de *linguagens de programação de simulação*. (Por exemplo, duas dentre as primeiras que continuam a ser amplamente usadas são GPSS e SIMSCRIPT). Estas linguagens são especialmente projetadas para acelerar o tipo de programação (e reprogramação) único para simulação. Seus propósitos específicos incluem o seguinte. Um objetivo é fornecer um meio conveniente de descrever os elementos que comumente aparecem nos modelos de simulação. Um segundo é acelerar mudanças no projeto e políticas de operação do sistema que está sendo simulado, para que um grande número de configurações (incluindo algumas sugeridas durante o curso do estudo) possa ser considerado facilmente. Um outro serviço fornecido pelas linguagens de simulação é algum tipo de mecanismo interno de tempo e controle, com comandos relacionados, para dar assistência ao tipo de armazenamento de informação que é necessário quando da execução de uma rodada de simulação. Elas também são projetadas para obter, convenientemente, dados e

estatísticas sobre o comportamento agregado do sistema que está sendo simulado. Finalmente, estas linguagens fornecem procedimentos operacionais simples, tais como introduzir mudanças no modelo de simulação, inicializar o estado do modelo, alterar o tipo dos dados resultantes a serem gerados e empilhar uma série de rodadas de simulação.

Por todas essas razões, um programa de simulação quase sempre deveria ser escrito em uma destas linguagens de simulação em lugar de uma linguagem de programação geral. As enormes poupanças em tempo de programação, ordinariamente causadas pelas linguagens de simulação, usualmente compensam por qualquer pequeno aumento no tempo de uso do computador.

No final deste texto, serão citados alguns dos diversos pacotes de modelagem e linguagens de simulação existentes.

Finalmente, deveríamos enfatizar que a estratégia do estudo de simulação deveria ser estudada cuidadosamente, antes de terminar o programa de simulação. Fazer com que o computador meramente compile massas de dados, numa procura cega de alternativas atraentes, está longe de ser adequado. A simulação é, basicamente, um meio de conduzir uma investigação experimental. Por isso, exatamente como com um experimento físico, deveria ser dada uma atenção cuidadosa à construção de uma teoria de hipóteses formais a serem testadas e ao projeto habilidoso de um experimento estatístico que resultará em conclusões válidas.

2.5 - Validação do Modelo

O modelo de simulação típico consiste de um alto número de elementos, regras e ligações lógicas. Por isso, mesmo que os componentes individuais tenham sido cuidadosamente testados, diversas aproximações pequenas ainda podem se acumular em grandes distorções no resultado do modelo geral. Consequentemente, depois de escrever e tirar os erros do programa de computador, é importante testar a validade do modelo para prever, razoavelmente, o comportamento agregado do sistema que está sendo simulado.

Quando alguma forma do sistema real já tenha estado em operação, seus dados de desempenho deveriam ser comparados com os dados do resultado correspondente do modelo. Às vezes, podem ser usados testes estatísticos padrões para determinar se as diferenças nas médias, variâncias e distribuições de probabilidade que geraram os dois conjuntos de dados são estatisticamente significantes. O comportamento dos dados, dependentes do tempo, também pode ser comparado estatisticamente. Se os dados não forem passíveis de uma análise estatística, uma outra abordagem é perguntar ao pessoal familiarizado com o comportamento do sistema real se eles podem discriminar entre os dois conjuntos de dados.

Se o modelo pretender simular configurações de projetos ou políticas de operação alternativos para um sistema proposto, para o qual não existam dados reais disponíveis, talvez seja preferível conduzir um *teste de campo* para coletar alguns dados reais para comparar com o resultado do modelo. Isto poderia envolver a construção de um protótipo pequeno de alguma versão do sistema proposto e colocá-lo em operação. Outra possibilidade seria alterar temporariamente um sistema existente para que corresponda a uma das propostas.

Entretanto, os testes de campo, freqüentemente, são muito dispendiosos e demorados para serem usados. Sem nenhum dado real como um padrão de comparação, o único modo de validar o modelo geral é ter pessoas, com conhecimento de causa, que verifiquem cuidadosamente a credibilidade dos dados do resultado para uma variedade de situações. Mesmo quando não existam bases para se verificar se os dados são razoáveis para uma *única* situação, usualmente, podem-se tirar algumas conclusões quanto a como o desempenho *relativo* do sistema mudaria à medida que vários parâmetros fossem mudados. É especialmente importante convencer o *tomador de decisão* quanto à credibilidade do modelo, pois assim ele desejará usá-lo para, pelo menos, *ajudar* suas decisões. Se o modelo puder ser usado novamente, num futuro, dever-se-iam manter registros cuidadosos de suas predições e dos resultados reais, para continuar o processo de validação.

3 - PROJETO EXPERIMENTAL PARA SIMULAÇÃO

3.1 - Técnicas de Redução da Variância

Como, usualmente, é necessário um tempo considerável de computador para as execuções de simulação, é importante obtermos a maior e mais precisa quantidade de informação possível da quantidade de simulação que possa ser feita. Infelizmente, na prática, tem havido uma tendência em aplicar a simulação sem sentido crítico, sem se dar a atenção adequada à eficiência do projeto experimental. Isto tem sido verdadeiro, apesar do fato de ter havido um progresso considerável no desenvolvimento de técnicas especiais para aumentar a precisão (i.e., decrescer a variância) dos estimadores da amostra.

Estas técnicas de redução da variância são, freqüentemente, chamadas de técnicas de *Monte Carlo* (um termo aplicado, às vezes, à simulação em geral). Uma vez que elas tendem a ser bastante sofisticadas, não será possível explorá-las em profundidade aqui. Entretanto, vamos tentar dar umas pinceladas sobre estas técnicas e mostrar o grande aumento na precisão que elas, às vezes, conseguem.

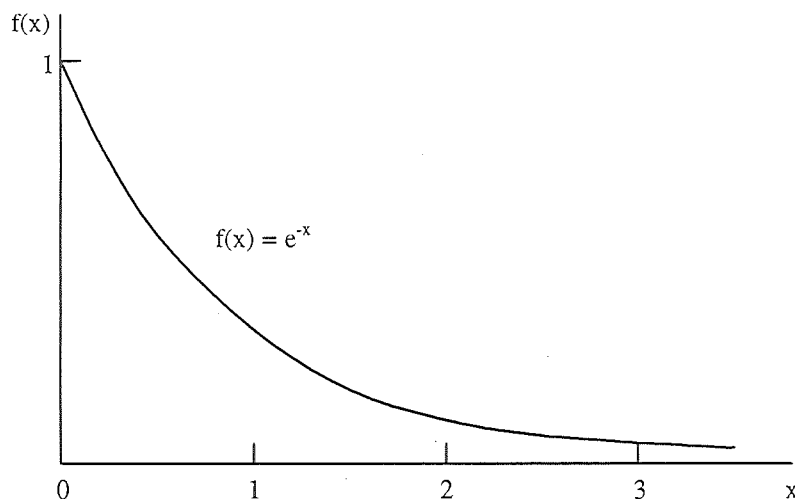


Figura 6 - Função densidade de probabilidade para o exemplo

Consideremos a *distribuição exponencial* cujo parâmetro tenha o valor de 1. Então, sua função densidade de probabilidade é $f(x) = e^{-x}$, como mostrado na Figura 6, e sua função de distribuição cumulativa é $F(x) = 1 - e^{-x}$. É sabido que a média desta distribuição é 1. Entretanto, suponhamos que isso não fosse conhecido e que seja desejado estimar esta média usando-se simulação.

A abordagem de SIMULAÇÃO DIRETA (também chamada de técnica de *Monte Carlo Simples*) envolve gerar algumas observações aleatórias a partir da distribuição exponencial que está sendo considerada e, então, usar a média destas observações para estimar a média. Como descrito na seção 2, estas observações aleatórias seriam

$$x_i = -\ln(1 - r_i) \quad \text{para } i = 1, 2, \dots, n,$$

onde r_1, r_2, \dots, r_n são números decimais aleatórios entre 0 e 1. Usando uma Tabela de Dígitos Aleatórios para se obterem 10 destes números decimais aleatórios, as observações aleatórias resultantes estão mostradas na Tabela 2. Note-se que a média amostral é 0,779, em oposição à média verdadeira de 1,000. Entretanto, como acontece de o desvio-padrão da média ser $1/\sqrt{n}$, ou $1/\sqrt{10}$ neste caso (como poderia ser estimado a partir da amostra), um erro deste tamanho ou maior ocorreria em, aproximadamente, metade das vezes. Felizmente, como o desvio-padrão de uma média amostral é sempre inversamente proporcional a \sqrt{n} , o tamanho desta amostra precisaria ser quadruplicado para reduzir este desvio padrão à metade. Estes fatos um tanto desanimadores sugerem a necessidade de outras técnicas que obtivessem tais estimativas mais precisamente e mais eficientemente.

Tabela 2 -Exemplo para a técnica de Monte Carlo Simples

i	Número aleatório r_i^*	Observação aleatória $x_i = -\ln(1 - r_i)$
1	0,495	0,684
2	0,335	0,408
3	0,791	1,568
4	0,469	0,633
5	0,279	0,328
6	0,698	1,199
7	0,013	0,014
8	0,761	1,433
9	0,290	0,343
10	0,693	1,183
Total = 7,793		
Estimativa da média = 0,779		

* Na verdade, foi adicionado 0,0005 ao valor indicado para cada um dos r_i , de modo que o intervalo dos seus valores possíveis seria de 0,0005 a 0,9995, em lugar de 0,000 a 0,999.

Uma técnica de Monte Carlo para se obter estimativas melhores é a amostragem estratificada. Existem duas falhas da abordagem de Monte Carlo Simples que são retificadas pela amostragem estratificada. Em primeiro lugar, pela própria natureza da aleatoriedade, uma amostra aleatória pode não fornecer uma representação particularmente uniforme da distribuição. Por exemplo, a amostra aleatória dada na Tabela 2 não tem nenhuma observação entre 0,014 e 0,328, mesmo que a probabilidade de que uma observação aleatória caia dentro deste intervalo seja maior que 1/4. Em segundo lugar, certas partes de uma distribuição podem ser mais críticas que outras para se obter uma estimativa precisa, porém a amostragem aleatória não dá nenhuma prioridade especial para a obtenção de observações destas partes. Por exemplo, a ponta de uma distribuição exponencial é especialmente crítica na determinação de sua média. Entretanto, a amostra aleatória na Tabela 2 não inclui nenhuma observação maior que 1,568, mesmo que pelo menos uma pequena probabilidade de valores muito maiores. Esta é a explicação básica por esta média amostral particular estar tão abaixo da média verdadeira. A amostragem estratificada contorna essas dificuldades dividindo a distribuição em partes chamadas de *estratos*, onde cada estrato seria amostrado individualmente, com amostragem desproporcionalmente pesada dos estados mais críticos.

Para ilustrarmos, vamos supor que a distribuição seja dividida em três estratos, da maneira mostrada na Tabela 3. Estes estratos foram escolhidos para corresponderem, aproximadamente, de 0 a 1, de 1 a 3, e de 3 a infinito, respectivamente. Para nos assegurarmos que as observações aleatórias geradas para cada estrato realmente caem naquela parte da distribuição, os números decimais aleatórios terão que ser convertidos à extensão indicada para $F(x)$, como mostrado na terceira coluna da Tabela 3. O número de observações a serem geradas de cada estrato é dado na Quarta coluna. Estes tamanhos amostrais estão baseados, grosso modo, um guia recomendado, de que eles sejam proporcionais ao produto *da probabilidade* de que uma observação aleatória caia dentro do estrato correspondente, vezes o desvio padrão deste estrato. A última coluna mostra, então, o peso de *amostragem* resultante para cada estrato, i.e, a *razão* entre a *proporção da amostra* (a fração da amostra total a se tirada do estrato) e a *proporção da distribuição* (a probabilidade de uma observação aleatória cair dentro do estrato). Esses pesos de amostragem refletem a importância relativa do estrato respectivo na determinação da média.

Dada a formulação da abordagem da amostragem estratificada mostrada na Tabela 3, os mesmos números aleatórios usados na Tabela 2 produzirão as observações dadas na quinta coluna da Tabela 4. Entretanto, não seria correto usar a média não ponderada destas observações para estimar a média, porque certas partes das distribuições foram amostradas mais que outras. Por isso, antes de tirar a média, as observações de cada estrato deveriam ser divididas pelo peso da amostragem para aquele estrato, para dar pesos proporcionais às diferentes partes da distribuição, como mostrado na última coluna da Tabela 4. A média *ponderada* resultante, de 0,948, dá a estimativa desejada da média.

Tabela 3 - Formulação do exemplo da amostragem estratificada

Estrato	Parte da distribuição	Nº aleatório do estrato	Tamanho da amostra	Peso da amostra
1	$0 \leq F(x) \leq 0,64$	$r'_i = 0 + 0,64 r_i$	4	$w_i = \frac{4/10}{0,64} = \frac{5}{8}$
2	$0,64 \leq F(x) \leq 0,96$	$r'_i = 0,64 + 0,32 r_i$	4	$w_i = \frac{4/10}{0,32} = \frac{5}{4}$
3	$0,96 \leq F(x) \leq 1$	$r'_i = 0,96 + 0,04 r_i$	2	$w_i = \frac{2/10}{0,04} = 5$

Tabela 4 - Exemplo para a amostra estratificada

Estrato	i	Número aleatório r_i	Nº aleatório do estrato r'_i	Observação aleatória do estrato $x_i = -\ln(1 - r'_i)$	Peso da amostra w_i	$\frac{x_i}{w_i}$
1	1	0,495	0,317	0,381	$\frac{5}{8}$	0,610
	2	0,335	0,215	0,242	$\frac{5}{8}$	0,387
	3	0,791	0,507	0,707	$\frac{5}{8}$	1,131
	4	0,469	0,300	0,357	$\frac{5}{8}$	0,571
2	5	0,279	0,729	1,306	$\frac{5}{4}$	1,045
	6	0,698	0,864	1,995	$\frac{5}{4}$	1,596
	7	0,013	0,644	1,033	$\frac{5}{4}$	0,826
	8	0,761	0,884	2,154	$\frac{5}{4}$	1,723
3	9	0,290	0,9716	3,561	5	0,712
	10	0,693	0,9877	4,398	5	0,880
Total = 9,481						
Estimativa da média = 0,948						

Este exemplo sugeriu que as técnicas de redução da variância fornecem um estimador da média muito mais preciso do que o fornecido pela simulação direta. Estes resultados, definitivamente, não foram uma coincidência. Em comparação com a simulação direta, esta técnica (inclusive diversas outras mais complicadas, não

apresentadas aqui) de fato fornecem um estimador muito mais preciso, com a mesma quantidade de tempo de computador, ou um estimador tão preciso quanto, porém com tempo de computador muito menor. Apesar do fato de que possa ser necessária uma análise adicional para incorporar uma ou mais destas técnicas ao estudo de simulação, as recompensas não deveriam ser descartadas facilmente.

Embora este exemplo tenha sido particularmente simples, muitas vezes é possível, mesmo que mais difícil, aplicar essa técnica a problemas muito mais complexos. Por exemplo, suponhamos que o objetivo do estudo de simulação seja estimar o tempo médio de espera de clientes num sistema de fila. Como tanto a distribuição de probabilidade do tempo entre chegadas quanto a distribuição de probabilidade dos tempos de serviço estão envolvidas, e como os tempos de espera consecutivos não são, estatisticamente, independentes, pode parecer que este problema esteja além da capacidade das técnicas de redução da variância. Entretanto, estas técnicas e outras podem, de fato, ser aplicadas a este tipo de problema, muito vantajosamente.

3.2 - Problemas Táticos

Muitos estudos de simulação estão interessados em sistemas de investigação que operam continuamente numa condição de estado de equilíbrio. Infelizmente, um modelo de simulação não pode ser operado desta maneira; ele tem que ser iniciado e parado. Por causa da artificialidade introduzida pelo início abrupto da operação, o desempenho do sistema simulado não se torna representativo do sistema do mundo real correspondente, até que ele tenha alcançado, essencialmente, uma continuação de estado de equilíbrio (i.e., até que a distribuição de probabilidade do estado do sistema simulado tenha, essencialmente, alcançado uma distribuição de *equilíbrio* limitante). Assim, um problema tático é como obter dados que sejam relevantes para prever o comportamento de *estado de equilíbrio* do sistema real.

A forma tradicional de lidar com este problema é rodar o modelo de simulação por algum tempo sem coletar dados, até que se creia que o sistema simulado tenha alcançado, essencialmente, uma condição de estado de equilíbrio. Infelizmente, é difícil estimar exatamente quanto tempo deverá ter este *período de estabilização*. Além disso, resultados analíticos disponíveis sugerem que é necessário um período surpreendentemente longo, de modo que terá que ser gasta uma grande quantidade de tempo improdutivo. A seção 4 apresenta uma abordagem estatística que elimina estas dificuldades.

Uma questão tática relacionada é a seleção das *condições de iniciar* do sistema simulado. A recomendação tradicional é que o sistema simulado deva ser iniciado num estado tão representativo quanto possível das condições de estado de equilíbrio, para minimizar a extensão necessária do período de estabilização. Entretanto, o objetivo subjacente ao experimento simulado é *estimar* estas condições, de modo que pode haver pouca informação disponível, de antemão, para guiar a seleção neste sentido. O procedimento na seção 4 também elimina esta dificuldade.

A maioria dos procedimentos amostrais estatísticos supõe que os dados dos resultados experimentais estejam na forma de uma coleção de observações aleatórias, distintas e estatisticamente independentes, de alguma distribuição de probabilidade

subjacente. Por outro lado, por causa da natureza dos problemas para os quais a simulação é usada, as observações a partir de um experimento simulado têm probabilidade de ser altamente correlacionadas. Por exemplo, existe uma alta correlação entre os tempos de espera de clientes consecutivos num sistema de fila. Além disso, muitas medidas de desempenho são tais que o próprio experimento simulado produz esta medida continuamente como uma função de tempo, em lugar de como uma seqüência de observações separadas. Assim, um outro problema tático é como coletar os dados de forma a contornar estas dificuldades.

Um método tradicional é executar uma série de rodadas de simulação completamente separadas e independentes, de igual extensão, e usar a medida média de desempenho para cada rodada (excluindo o período de estabilização inicial) como uma observação individual. A principal desvantagem é que cada rodada requer um período de estabilização inicial para aproximar uma condição de estado de equilíbrio, de modo que muito do tempo de simulação será improdutivo. O segundo método tradicional elimina esta desvantagem, executando as rodadas consecutivamente, usando a condição de término de uma rodada como a condição de estado de equilíbrio de início para a rodada seguinte. Em outras palavras, existe apenas uma rodada de simulação geral contínua que (exceto por um período de estabilização inicial) é dividida, para propósitos de armazenar informações, em várias partes (rodadas) iguais. A medida média de desempenho para cada parte é, então, tratada como uma observação individual. A desvantagem deste método é que ele não elimina inteiramente a correlação entre observações, muito embora possa reduzi-la consideravelmente, fazendo as partes suficientemente longas.

Novamente, essas dificuldades são eliminadas pela abordagem estatística na seção que segue.

4 - O MÉTODO REGENERATIVO DE ANÁLISE ESTATÍSTICA

O conceito básico subjacente a esta abordagem é que uma rodada de simulação para diversos sistemas pode ser dividida em uma série de *ciclos*, tais que os comportamentos do sistema durante ciclos diferentes são ambos, *estatisticamente independentes e identicamente distribuídos*. Portanto, calculando-se algumas estatísticas para resumir o comportamento em questão dentro de cada ciclo, estas estatísticas para os respectivos ciclos constituem uma série de observações independentes e identicamente distribuídas, que podem ser analisadas por procedimentos estatísticos padrões. Como o sistema se mantém percorrendo estes ciclos independentes e identicamente distribuídos quer já esteja ou não numa condição de estado de equilíbrio, estas observações são diretamente aplicáveis, desde o início, para estimar o comportamento do sistema em estado de equilíbrio.

Para que os ciclos possuam essas propriedades, cada um deles terá que *começar* no mesmo *ponto de regeneração*, isto é, no ponto em que o sistema entra, novamente, num certo estado especial, do qual a simulação pode prosseguir, sem qualquer conhecimento de sua história passada. O sistema pode ser visto como *regenerando* a si próprio neste ponto, no sentido em que a estrutura probabilística do comportamento futuro do sistema depende somente de estar neste ponto e de nada mais que tenha acontecido anteriormente. Um ciclo *termina* quando o sistema alcança, novamente, o

ponto de regeneração (quando começa o ciclo seguinte). Assim, a *extensão de um ciclo* é exatamente o tempo decorrido entre ocorrências consecutivas do ponto de regeneração, o qual é uma variável aleatória dependente da evolução do sistema.

Quando estiver sendo usado o *incremento para o evento seguinte*, um ponto de regeneração típico é um ponto em que um evento acaba de ocorrer, e que nenhum evento futuro tenha sido programado. Assim, nada precisa ser conhecido sobre a história das programações anteriores e a simulação pode começar do início, programando eventos futuros. Quando estiver sendo usado o incremento de tempo fixo, um ponto de regeneração é um ponto em que as probabilidades dos eventos possíveis ocorrendo durante a unidade de tempo seguinte não dependem de quando ocorreu qualquer evento passado, mas somente dependem do estado atual do sistema.

Nem todo sistema possui pontos de regeneração, de forma que este *método regenerativo* de coleta de dados nem sempre pode ser usado. Além disso, mesmo quando existem pontos de regeneração, aquele que é escolhido para definir os pontos de início e término dos ciclos terá que ocorrer com frequência suficiente, para que seja obtido um número substancial de ciclos dentro de uma quantidade razoável de tempo de computador. Os requerimentos teóricos para o método são que a extensão esperada do ciclo seja *finita* e que o número de ciclos vá ao infinito, se o sistema continuasse a operar indefinidamente. Por isso, tem que ser tomado algum cuidado na escolha de um ponto de regeneração apropriado.

Talvez a mais importante área de aplicação do método regenerativo tenha sido a simulação de *sistemas de filas*, incluindo *redes de filas*.

EXEMPLO. Suponhamos que seja necessário obter informações sobre o comportamento do sistema em estado de equilíbrio que possa ser formulado como um *sistema de fila com um servidor único*. Entretanto, ambos os tempos, entre chegadas e de serviço, têm uma *distribuição uniforme*, com um intervalo de 5 a 25 e de 0 a 20, respectivamente, de modo que não existem resultados analíticos disponíveis. Portanto, será usada a simulação com *incremento para o evento seguinte* para se obterem os resultados desejados.

Exceto pelas distribuições envolvidas, a abordagem geral é a mesma descrita na seção 1, para o Exemplo 2. Suponhamos que sejam usados números aleatórios de um dígito para gerar as observações aleatórias a partir das distribuições, como mostrado na Tabela 5. Começando a rodada de simulação com zero clientes no sistema, então teremos os resultados resumidos na Tabela 6 e Figura 6, onde os números aleatórios foram obtidos seqüencialmente, à medida que necessário.

Tabela 5 - Correspondência entre números aleatórios e observações aleatórias para o exemplo do sistema de fila

Número aleatório	Tempo entre chegadas	Tempo de serviço
0	6	1
1	8	3
2	10	5
3	12	7
4	14	9
5	16	11
6	18	13
7	20	15
8	22	17
9	24	19

Para este sistema, um *ponto de regeneração* é onde ocorre uma *chegada sem nenhuma* saída anterior de algum cliente. A essa altura, a estrutura probabilística, de quando ocorrerão futuras chegadas ou conclusões de serviços, é completamente independente de qualquer história passada. A única informação relevante é que o sistema acabou de entrar no estado especial de ter zero clientes e de o tempo até a próxima chegada ter alcançado zero. A rodada de simulação não teria programado com antecipação quaisquer eventos futuros, porém, agora, teria gerado ambos, o próximo tempo entre chegadas e o próximo tempo de serviço.

Os únicos outros pontos de regeneração para este sistema são quando uma chegada e uma conclusão de serviço ocorrem, simultaneamente, com um número pré-especificado de clientes no sistema. Estes pontos ocorreriam apenas raramente no sistema real porque suas distribuições para os tempos entre chegadas e de serviço são contínuas. Entretanto, eles ocorreriam com alguma frequência na rodada de simulação porque a Tabela 5 torna *discretas* as distribuições realmente usadas para gerar as observações aleatórias. Portanto, um destes pontos, de fato, poderia ser escolhido para definir um ciclo. Entretanto, o ponto de regeneração descrito no parágrafo precedente ocorre muito mais frequentemente e, portanto, é uma escolha melhor. Feita essa seleção, teremos, como os cinco primeiros ciclos completos da rodada de simulação, os apresentados na Figura 6. (Usualmente se teria um número de ciclos consideravelmente maior, considerando-se toda a rodada de simulação, para que se tenha precisão suficiente na análise estatística).

Há vários tipos de informação, sobre o comportamento do sistema em estado de equilíbrio, que podem ser obtidos a partir desta rodada de simulação, incluindo estimativas pontuais e intervalos de confiança para o número esperado de clientes no sistema, o tempo médio de espera, e assim por diante. Em cada um dos casos, apenas é necessário usar as estatísticas correspondentes aos ciclos respectivos e as extensões dos ciclos. Apresentaremos, em primeiro lugar, as expressões estatísticas gerais para o método regenerativo e depois os aplicaremos a este exemplo.

Tabela 6 - Rodada de simulação para o exemplo do sistema de fila

Tempo	Número de clientes	Números aleatórios	Próxima chegada	Conclusão do próximo serviço
0	0	9	24	...
24	1	2 e 6	34	37
34	2	4	48	37
37	1	6	48	50
48	2	4	62	50
50	1	1	62	53
53	0	...	62	...
62	1	1 e 1	70	65
65	0	...	70	...
70	1	3 e 9	82	89
82	2	1	90	89
89	1	4	90	98
90	2	1	98	98
98	2	1 e 5	106	109
106	3	6	124	109
109	2	2	124	114
114	1	1	124	117
117	0	...	124	...
124	1	5 e 6	140	137
137	0	...	140	...
140	1	9 e 3	164	147
147	0	...	164	...
164	1

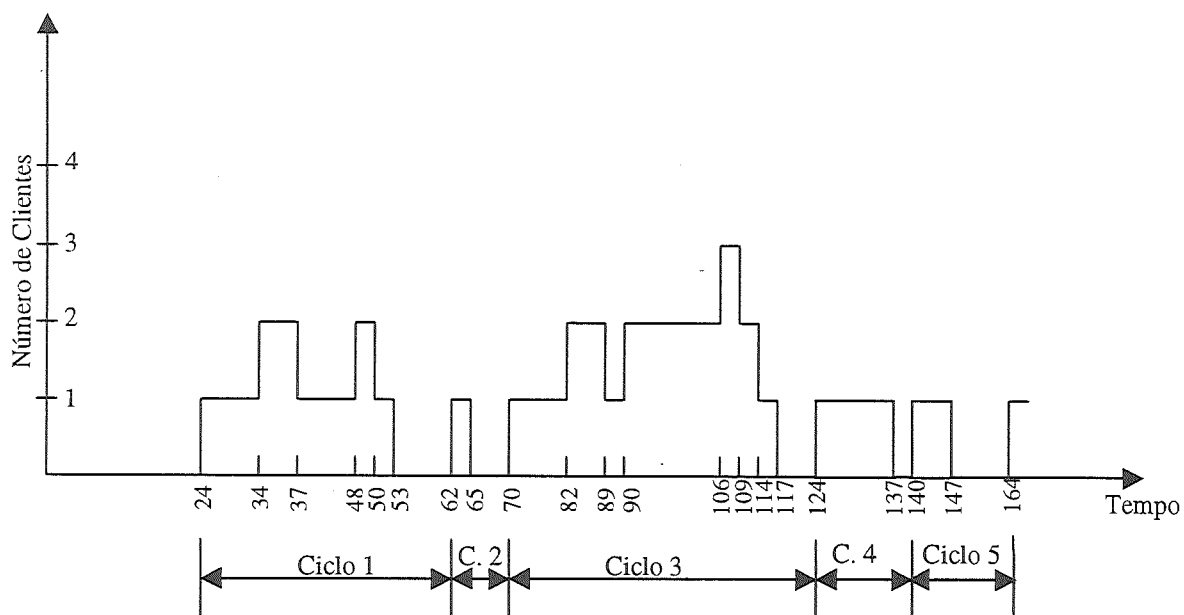


Figura 6 - Resultado da rodada de simulação para o exemplo do sistema de fila.

Fórmulas Estatísticas

Formalmente falando, o problema estatístico para o método regenerativo é obter estimativas do valor esperado de alguma variável aleatória, X , de interesse. Isto é feito calculando-se um Y estatístico para cada ciclo, tal que

$$E(X) = \frac{E(Y)}{E(Z)}$$

onde a variável aleatória Z é uma medida apropriada do *tamanho* do ciclo. Portanto, se n ciclos completos forem gerados durante a rodada de simulação, os dados coletados serão Y_1, Y_2, \dots, Y_n e Z_1, Z_2, \dots, Z_n , para os respectivos ciclos

Fazendo \bar{Y} e \bar{Z} denotarem, respectivamente, as médias amostrais para estes dois conjuntos de dados, a *estimativa pontual* correspondente de $E(X)$ seria obtida a partir da expressão

$$\text{est}\{E(X)\} = \frac{\bar{Y}}{\bar{Z}}$$

Para obtermos uma *estimativa de intervalo de confiança* de $E(X)$, é necessário primeiro calcular diversas quantidades a partir dos dados. Dentre elas se incluem as *variâncias amostrais*:

$$S_{11}^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2 = \frac{1}{n-1} \sum_{i=1}^n Y_i^2 - \frac{1}{n(n-1)} \left(\sum_{i=1}^n Y_i \right)^2,$$
$$S_{22}^2 = \frac{1}{n-1} \sum_{i=1}^n (Z_i - \bar{Z})^2 = \frac{1}{n-1} \sum_{i=1}^n Z_i^2 - \frac{1}{n(n-1)} \left(\sum_{i=1}^n Z_i \right)^2$$

e a *covariância amostral* combinada :

$$S_{12}^2 = \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})(Z_i - \bar{Z}) = \frac{1}{n-1} \sum_{i=1}^n Y_i Z_i - \frac{1}{n(n-1)} \left(\sum_{i=1}^n Y_i \right) \left(\sum_{i=1}^n Z_i \right)$$

Façamos também

$$S^2 = S_{11}^2 - 2 \left(\frac{\bar{Y}}{\bar{Z}} \right) S_{12}^2 + \left(\frac{\bar{Y}}{\bar{Z}} \right)^2 S_{22}^2.$$

Finalmente, façamos de α a constante, tal que $(1 - 2\alpha)$ seja o *coeficiente de confiança* desejado para o intervalo de confiança, e procurando K para a distribuição normal. Se n não for muito pequeno, uma boa aproximação da *estimativa de intervalo de confiança* desejado de $E(X)$ será, então, dada por

$$\frac{\bar{Y}}{\bar{Z}} - \frac{K_{\alpha} s}{\bar{Z} \sqrt{n}} \leq E(X) \leq \frac{\bar{Y}}{\bar{Z}} + \frac{K_{\alpha} s}{\bar{Z} \sqrt{n}};$$

ou seja, a probabilidade de que os pontos extremos de um intervalo, gerado desta maneira, venham a envolver o valor real de $E(X)$ é de aproximadamente $(1 - 2\alpha)$.

Aplicação das Fórmulas Estatísticas ao Exemplo

Consideremos, primeiramente, como estimar o *tempo médio de espera* para um cliente antes de começar seu serviço. Assim, a variável aleatória X , agora, representaria o tempo de espera de um cliente, excluindo serviço, de modo que

$$\bar{T}_c = E(X).$$

A informação correspondente coletada durante a rodada de simulação é o tempo de espera *real* dos clientes respectivos. Portanto, para cada ciclo, o resumo estatístico Y seria a *soma dos tempos de espera*, e o tamanho do ciclo Z seria o *número de clientes*, de modo que

$$\bar{T}_c = \frac{E(Y)}{E(Z)}.$$

Para o ciclo 1, é processado um total de três clientes, portanto $Z_1 = 3$. O primeiro cliente não sofre nenhuma espera antes de começar o serviço, o segundo espera três unidades de tempo (de 34 a 37), e o terceiro espera 2 unidades (de 48 a 50), de forma que $Y_1 = 5$. Procedendo de forma similar para os outros ciclos, os dados para o problema serão

$$\begin{aligned} Y_1 &= 5 & Z_1 &= 3 \\ Y_2 &= 0 & Z_2 &= 1 \\ Y_3 &= 34 & Z_3 &= 5 \\ Y_4 &= 0 & Z_4 &= 1 \\ Y_5 &= 0 & Z_5 &= 1 \\ \bar{Y} &= 7,8 & \bar{Z} &= 2,2 \end{aligned}$$

Por isso, a *estimativa de ponto* de \bar{T}_c é

$$\text{est} \{ \bar{T}_c \} = \frac{\bar{Y}}{\bar{Z}} = \frac{7,8}{2,2} = 3,54.$$

Para obtermos uma estimativa de intervalo de confiança de \bar{T}_c de 95 por cento, as expressões acima serão usadas, em primeiro lugar, para calcular

$$S_{11}^2 = 219,20 \quad S_{22}^2 = 3,20 \quad S_{12}^2 = 24,80 \quad S = 9,14.$$

Como $(1 - 2\alpha) = 0,95$, $\alpha = 0,025$, então $K_\alpha = 1,96$ pela tabela da Distribuição Normal. O intervalo de confiança resultante é

$$- 0,09 \leq \bar{T}_c \leq 7,19$$

isto é,

$$\bar{T}_c \leq 7,19.$$

A razão desse intervalo de confiança ser tão grande (mesmo incluindo valores negativos impossíveis) é que o número de observações amostrais (ciclos) é muito pequeno. Notemos, na fórmula geral, que a amplitude do intervalo de confiança é *inversamente proporcional à raiz quadrada* de n , de forma que, por exemplo, quadruplicando n , reduz-se a amplitude à metade (supondo-se que não haja mudança nem em S , nem em \bar{Z}). Dados valores preliminares de S e \bar{Z} , a partir de uma rodada de simulação preliminar curta (tal como a rodada acima), esta relação torna possível estimar a amplitude do intervalo de confiança que resultaria de qualquer dada escolha de n , para a rodada de simulação completa. A escolha final de n pode, então, ser feita com base na conciliação entre o tempo de computador e a precisão da análise estatística.

Suponhamos, agora, que esta rodada de simulação seja usada para estimar P_0 , a probabilidade de haver *zero* clientes no sistema. A informação correspondente obtida durante a rodada de simulação é a fração de tempo durante a qual o sistema está vazio. Por isso, o resumo estatístico Y para cada ciclo seria o *tempo total* durante o qual *nenhum* cliente estaria presente, e o tamanho de Z seria a *extensão* do ciclo, de forma que

$$P_0 = \frac{E(Y)}{E(Z)}.$$

A extensão do ciclo 1 é 38 (de 24 a 62), de modo que $Z_1 = 38$. Durante este período, o sistema está vazio de 53 a 62, de modo que $Y_1 = 9$. Procedendo desta maneira para os outros ciclos, obteríamos os seguintes dados para o problema:

$$\begin{array}{ll} Y_1 = 9 & Z_1 = 38 \\ Y_2 = 5 & Z_2 = 8 \\ Y_3 = 7 & Z_3 = 54 \\ Y_4 = 3 & Z_4 = 16 \\ Y_5 = 17 & Z_5 = 24 \\ \bar{Y} = 8,2 & \bar{Z} = 28 \end{array}$$

Assim, a *estimativa pontual* de P_0 é

$$\text{est} \{P_0\} = \frac{8,2}{28} = 0,293$$

Calculando-se

$$S_{11}^2 = 29,20 \quad S_{22}^2 = 334 \quad S_{12}^2 = 17 \quad S = 6,92$$

encontraremos uma *estimativa de intervalo de confiança* de P_0 , de 95 por cento de

$$0,076 \leq P_0 \leq 0,510.$$

(A grande extensão deste intervalo indica que seria necessária uma rodada de simulação mais longa para se obter uma estimativa de P_0 relativamente precisa).

Redefinindo-se y apropriadamente, também se poderá usar a mesma abordagem para estimar outras probabilidades envolvendo o número de clientes no sistema. Entretanto, como esse número nunca excede 3 durante essa simulação, seria necessária uma rodada muito mais longa se a probabilidade envolvesse números grandes.

Os outros valores médios básicos na teoria das filas, podem também ser estimados diretamente a partir dos resultados da rodada de simulação. Por exemplo, o *número médio de clientes esperando para serem servidos* (\bar{F}) pode ser estimado definindo-se

$$Y = \sum_{n=2}^{\infty} (n-1) T_n,$$

onde T_n é o *tempo total* que exatamente n clientes estão no sistema durante o ciclo. (Esta definição de Y , na verdade, é equivalente à definição usada quando se estima \bar{T}_c). Nesse caso, Z seria definido como quando se estimasse qualquer P_n , isto é, a *extensão* do ciclo. A estimativa pontual de \bar{F} resultante, torna-se simplesmente a estimativa pontual de \bar{T}_c multiplicada pela *taxa média de chegada* real para os ciclos completos observados.

5 - PACOTES COMPUTACIONAIS E LINGUAGENS DE SIMULAÇÃO

Como já mencionado, o uso generalizado de simulação como uma técnica de soluções de problemas de Pesquisa Operacional e ferramenta de Análise de Sistemas, deu origem a uma diversidade de linguagens especificamente projetadas para tais fins.

5.1 - Linguagens de uso geral

A grande vantagem desse tipo de linguagem é a sua utilização em uma ampla quantidade de sistemas. Por outro lado, a generalidade obriga que a formulação do modelo seja feita pelo usuário (analista).

Uma das primeiras linguagens de uso geral é a GPSS (General Purpose Simulation System), a qual é sucintamente descrita a seguir.

A base da linguagem GPSS é o uso de blocos descritivos, cada qual possuindo um significado preciso. O programador em GPSS deve escrever um diagrama usando somente os blocos descritivos da linguagem GPSS. Representa-se cada ação ou movimento de um elemento ou de uma entidade simulada por um fluxo ou por uma seqüência de tais blocos.

O exemplo, dado a seguir, trata da simulação de uma máquina que produz peças à razão de uma unidade em cada 5 minutos. Ao ser fabricada, cada peça passa por um inspetor que consome um tempo igual a 4 ± 3 minutos (ou seja tempo-médio de 4 minutos e desvio-padrão de 3 minutos) para examiná-la e rejeita aproximadamente 10% das peças. A unidade de tempo usada será o minuto. O diagrama de blocos em GPSS está apresentado na Figura 7.

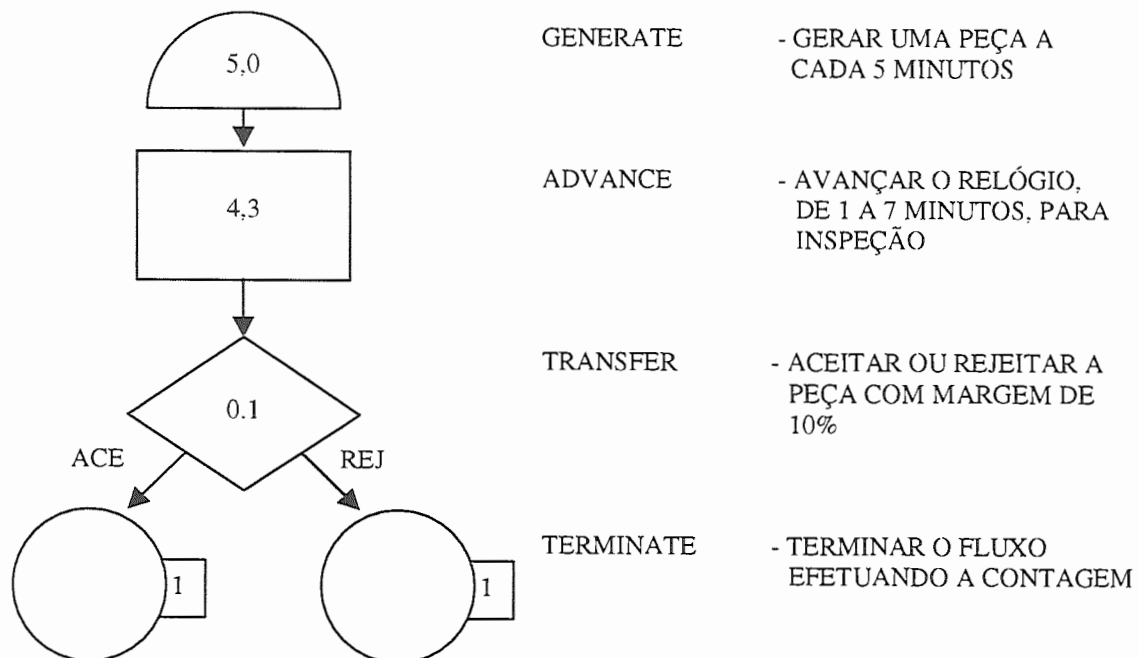


Figura 7.

A seguir descrevemos cada bloco usado no diagrama.

Bloco GENERATE -é usado para representar a geração ou a fabricação de uma peça a cada 5 minutos. A forma geral da instrução correspondente a esse bloco seria "GENERATE m,s" onde m e s indicam respectivamente a *média* e o *modificador* usados na geração. O modificador no nosso exemplo é o desvio-padrão da distribuição

usada, que no caso é zero, e a média possui o valor 5. É possível gerar através do bloco ou instrução GENERATE qualquer outro tipo de distribuição de probabilidade como por exemplo a Poisson, a Normal, etc.

Bloco ADVANCE - esse bloco, com a média 4 e modificador 3 (que nesse caso também é um desvio-padrão) representa o tempo gasto na inspeção de uma peça. Portanto, o valor do tempo gerado, que vai ser computado ao tempo de simulação, poderá estar entre 1-7 minutos. O bloco ADVANCE normalmente é o único que consome tempo de simulação e corresponderia ao processo de “avançar o relógio da simulação de k unidades de tempo”.

Bloco TRANSFER - após a inspeção a peça sofrerá um processo de seleção de modo que 90% delas prosseguem para o local ou início do fluxo chamado ACE que indica o início do fluxo das peças aceitas, enquanto que os 10% restantes prosseguem para o local chamado REJ que é o início do fluxo das peças rejeitadas. O fator ou critério de seleção que no caso é 0,1 (ou 0.1 na linguagem GPSS) será colocado dentro do bloco.

Bloco TERMINATE - tanto após a aceitação como após a rejeição, nada mais resta a simular para cada peça. Portanto, os fluxos ACE e REJ devem ser terminados usando-se o bloco TERMINATE. O valor 1 agregado a cada bloco TERMINATE indica que um contador atribuído ao bloco será incrementado de 1. Dessa maneira, ao final da simulação, teremos automaticamente calculado o número total de peças aceitas e de peças rejeitadas.

Existem cerca de 40 blocos básicos GPSS e 100 comandos. Cada bloco é pictorialmente representado por uma figura que sugere sua função.

Outras linguagens de simulação de uso geral são: SIMSCRIPT, SIMULA, RESQ, SLAM, dentre diversas outras.

5.2 - Linguagens de uso específico

Neste tipo de linguagem a formulação do modelo é construída na própria ferramenta, sendo os parâmetros especificados através de uma linguagem relacionada com o sistema modelado.

Alguns exemplos são:

PERFORMS, XL: para sistemas de computação e seus respectivos sistemas operacionais.

PET, NETWORK: para redes de computadores.

MAP/1, MAST: para sistemas de manufatura.

BETHSIM: para operações de siderurgia.

5.3 - Programas computacionais para Simulação com interface gráfica e orientados a objeto

Tais pacotes computacionais são de desenvolvimento relativamente recente e apresentam características como: interface gráfica de comunicação com o usuário, modelos orientados a objeto e capacidade de animação do modelo. Alguns dos mais conhecidos são: ARENA, PROMODEL e AUTOMOD.

A seleção de uma linguagem de simulação muitas vezes é realizada pela disponibilidade da linguagem, mais do que por suas características estruturais. Para finalizar, apresentamos abaixo alguns fatores importantes a serem considerados na comparação de linguagens de simulação.

- *Portabilidade*: Disponibilidade da linguagem em outros computadores, incluindo microcomputadores e supercomputadores.
- *Flexibilidade*: Suporte oferecido a diferentes enfoques de simulação e diferentes conceitos de modelagem.
- *Estendibilidade*: Disponibilidade de novas facilidades, incluindo capacidade gráfica e de banco de dados.
- *Tempo de Execução*: Velocidade de compilação e de execução
- *Capacidade de Memória*: Gerenciamento adequado do espaço de armazenamento, impedindo a explosão do modelo.
- *Depuração*: Facilidade de depuração, disponibilidade de sistemas de suporte, documentação adequada.
- *Considerações de Processamento*: Disponibilidade de facilidades para coleta estatística e para processamento de listas. Geração de relatórios-padrão e facilidade para geração de relatórios personalizados.
- *Considerações de Codificação*: Facilidade e clareza de codificação, isto é, grau que os códigos são auto-documentados (clareza de especificação).
- *Aprendizado*: Facilidade de aprendizado da linguagem e treinamento disponível.
- *Expressividade*: Poder de modelagem da linguagem.
- *Número de Usuários*: Uso generalizado da linguagem em institutos de pesquisa, indústrias, órgãos governamentais, universidades (lembre-se que a linguagem é também um meio de comunicação dentro de uma comunidade).
- *Facilidade de Manutenção*: Grau de suporte e melhoramentos.
- *Tempo de Vida*: Tempo de uso da linguagem.

EXERCÍCIOS

- 1) Use o *método congruencial misto* para gerar as seguintes seqüências de números aleatórios:
 - a) Uma seqüência de 10 números aleatórios de *um dígito*, tal que $x_{n+1} = (x_n + 3)$ (módulo 10) e $x_0 = 2$.
 - b) Uma seqüência de oito números aleatórios entre 0 e 7, tal que $x_{n+1} = (5x_n + 1)$ (módulo 8) e $x_0 = 2$.
 - c) Uma seqüência de cinco números aleatórios de *dois dígitos*, tal que $x_{n+1} = (61x_n + 27)$ (módulo 100) e $x_0 = 30$.
- 2) Use o *método congruencial misto* para gerar uma seqüência de cinco números aleatórios de *dois dígitos*, tal que $x_{n+1} = (41x_n + 33)$ (módulo 100) e $x_0 = 53$.
- 3) Use o *método congruencial misto* para gerar as seguintes seqüências de números aleatórios:
 - a) Uma seqüência de cinco números aleatórios entre 0 e 31, tal que $x_{n+1} = (13x_n + 15)$ (módulo 32) e $x_0 = 17$.
 - b) Uma seqüência de três números aleatórios de *três dígitos*, tal que $x_{n+1} = (201x_n + 503)$ (módulo 1.000) e $x_0 = 600$.
- 4) Use os números aleatórios de um dígito : 6, 3, 5, 0, 8 , como sementes para gerar observações aleatórias para cada uma das seguintes situações:
 - a) Jogar uma moeda justa.
 - b) Jogar um dado.
 - c) Um carro que chegue aleatoriamente a um semáforo encontrará a luz verde 40 por cento das vezes, amarela 10 por cento das vezes e vermelha 50 por cento das vezes.
- 5) Gere cinco observações aleatórias a partir de uma *distribuição uniforme* entre - 25 e + 50.
- 6) Suponha que sejam necessárias observações aleatórias a partir da *distribuição triangular* cuja função densidade de probabilidade seja:

$$f(x) = \begin{cases} 2x, & \text{se } 0 \leq x \leq 1 \\ 0, & \text{caso contrário.} \end{cases}$$

- a) Derive uma expressão para cada observação aleatória, como uma função do número decimal aleatório r .
- b) Gere cinco observações aleatórias.

7) Gere três observações aleatórias a partir de cada uma das seguintes distribuições de probabilidade:

a) A distribuição uniforme de 10 a 30.

b) A distribuição cuja função densidade de probabilidade seja

$$f(x) = \begin{cases} \frac{1}{4}(x-1)^3, & \text{se } 1 \leq x \leq 3 \\ 0, & \text{caso contrário.} \end{cases}$$

c) A distribuição cuja função densidade de probabilidade seja

$$f(x) = \begin{cases} \frac{1}{200}(x-20), & \text{se } 20 \leq x \leq 40 \\ 0, & \text{caso contrário.} \end{cases}$$

8) Gere três observações aleatórias a partir de uma *distribuição normal* com média = 50 e desvio-padrão = 50 (Use $n = 3$ para cada observação).

9) Gere quatro observações aleatórias a partir de uma *distribuição normal* com média = 0 e desvio-padrão = 1. (Use $n = 12$ para cada observação). Depois, use essas quatro observações para gerar duas observações aleatórias a partir de uma *distribuição qui-quadrado* com 2 graus de liberdade.

10) Gere duas observações aleatórias a partir de cada uma das seguintes distribuições de probabilidade:

a) A distribuição exponencial com média = 4.

b) A distribuição normal com média = 4 e desvio-padrão = $2\sqrt{2}$. (Use $n = 6$ para cada observação).

11) O tempo pode ser considerado um sistema estocástico, uma vez que ele evolui, de maneira probabilística, de um dia ao outro. Suponha que, para uma certa localidade, essa evolução probabilística satisfaça à seguinte descrição:

A probabilidade de chover amanhã é de 0,8, se chover hoje.

A probabilidade de o dia ser claro (sem chuva) amanhã é de 0,9, se ele estiver claro hoje.

Simule a evolução do tempo para 30 dias, começando do dia seguinte a um dia de chuva.

12) Um jogo de dados requer que o jogador jogue dois dados uma ou mais vezes, até que esteja decidido se ele ganhou ou perdeu. Ele ganhará se a primeira jogada resultar numa soma de 7 ou 11 ou, alternativamente, se a primeira soma for 4, 5, 6, 8, 9 ou 10 e a mesma soma se repetir antes que uma soma 7 apareça. Por outro lado, ele perderá se a primeira jogada resultar numa soma de 2, 3 ou 12 ou, alternativamente, se a primeira soma for 4, 5, 6, 8, 9 ou 10 e a soma 7 aparecer antes que a primeira soma reapareça.

- a) Simule 5 jogadas deste jogo, para começar o processo de estimar a probabilidade de ganhar.
- b) Para um grande número de rodadas do jogo, a proporção de ganhos tem aproximadamente uma distribuição normal com média = 0,493 e desvio-padrão = $0,5 / \sqrt{n}$. Use esta informação para calcular o número de jogadas simuladas que seriam necessárias para se ter uma probabilidade de pelo menos 0,95 de que a proporção de ganhos seja menor que 0,5.

13) Considere a distribuição de probabilidade cuja função densidade de probabilidade seja:

$$f(x) = \begin{cases} \frac{1}{x^2}, & \text{se } 1 \leq x \leq \infty \\ 0, & \text{caso contrário.} \end{cases}$$

O problema é realizar um experimento simulado, com a ajuda das técnicas de redução da variância, para estimar a média desta distribuição. Para que haja um padrão de comparação, derive também a média analiticamente.

Para cada um dos casos que se seguem, gere 10 observações e calcule a estimativa da média:

- a) Use o *método de Monte Carlo simples*.
- b) Use *amostragem estratificada* com três estratos:
 - $0 \leq F(x) \leq 0,6$
 - $0,6 < F(x) \leq 0,9$
 - $0,9 < F(x) \leq 1$, com 3, 3 e 4 observações, respectivamente.

14) Considere a versão do servidor único do modelo básico original da teoria das filas (chegadas de Poisson e tempo de serviço exponencial). Suponha que a taxa média de chegada seja de 20 por hora e a taxa média de serviço seja de 25 por hora, e que se deseja estimar o tempo médio de espera antes que se inicie o serviço, usando simulação.

- a) Iniciando com o sistema vazio, use o *incremento de tempo fixo* (com 1 minuto como a unidade de tempo) para executar a simulação até que dois serviços sejam concluídos.
- b) Escreva um programa de simulação para computador com *incremento de tempo fixo* e 0,1 minuto como a unidade de tempo. Use o *método regenerativo* com 100 ciclos para obter uma estimativa pontual e um intervalo de confiança de 95 por cento para o tempo médio de espera em estado de equilíbrio, antes que o serviço comece. Compare com o valor teórico.

REFERÊNCIAS BIBLIOGRÁFICAS

HILLIER, F.S. e LIEBERMAN, G. J. *Introdução à Pesquisa Operacional*. Editora da USP, São Paulo, 1988.

SHIMIZU, T. *Simulação em Computador Digital*. Editora Edgard Blücher, São Paulo, 1975.

SOARES, L. F. G. *Modelagem e Simulação Discreta de Sistemas*. IME-USP, São Paulo, 1990.