3
Revisão Bibliográfica - O Papel da Estatística na Simulação

Neste capítulo destacamos a importância da estatística nas etapas de seleção e caracterização de dados de entrada dos modelos de simulação, bem como na análise de resultados com base consistente. No APÊNDICE II.i se encontram os conceitos básicos utilizados na literatura sobre simulação, sendo recomendada a sua leitura prévia para fins de familiarização com os termos que serão citados a seguir.

3.1
Questões estratégicas

Através da simulação de um sistema, normalmente obtemos, como resultado, **variáveis de saída** do modelo, ou **indicadores**. Tais grandezas devem ser selecionadas para compor o conjunto de medidas de desempenho, durante a fase de projeto do modelo (CHWIF *et* MEDINA, 2006, cap. 6 seção 6.4, pág. 116). De acordo com os objetivos estabelecidos no estudo, a análise desses indicadores deverá permitir a elaboração de um bom diagnóstico do sistema, considerando diferentes **cenários**[17]. Cenários são combinações de valores de parâmetros e variáveis, numéricos ou não, cujas mudanças são relevantes no sistema em estudo (TORRES *et al.,* 2006, item 3.2). A utilidade da simulação é permitir rapidamente o diagnóstico da operação do sistema, considerando múltiplos cenários, cujo teste na prática seria muito difícil ou impraticável.

Assim como a escolha dos indicadores a considerar no modelo de simulação, também é fundamental a seleção das **variáveis de entrada** do modelo. A **análise de sensibilidade** preliminar relativa aos **dados de entrada** permite avaliar o seu impacto no modelo do sistema, por meio de testes sucessivos do modelo variando esses dados dentro de determinada faixa. Mediante essa análise, pode-se filtrar os dados, descartando-se variáveis que pouco afetam os indicadores

[17] Conceitos e definições mais abrangentes sobre cenários podem ser encontrados no site http://pt.wikipedia.org/wiki/Cen%C3%A1rio_(software)

PUC-Rio - Certificação Digital Nº 0612533/CA

e a performance do sistema em estudo (KELTON, SADOWSKI, *et* STURROCK, 2004, cap. 4, seção 4.5.2, pág. 154-155).

Dependendo de um exame prévio, pode-se reconhecer como determinístico ou estocástico o comportamento de algumas variáveis, e o modo mais conveniente de representá-las no modelo. Estudos do comportamento de sistemas dinâmicos envolvem **variáveis aleatórias**, na maioria dos casos. A credibilidade de um modelo, com efeito, será maior quanto mais próximo da realidade ele for. Assim como nos sistemas reais, a maioria das grandezas presentes em modelos representativos de sistemas dinâmicos possui incertezas, as quais são caracterizadas estatisticamente pela existência de uma **variância**[18]. Essas grandezas são mais bem representadas no modelo por variáveis aleatórias, cujo comportamento é descrito por suas respectivas **distribuições de probabilidade**[19].

Em simulação, cada vez que é gerado um dado de entrada do modelo para uma variável aleatória, um valor numérico, proveniente da sua distribuição de probabilidade, é atribuído a essa variável aleatória. A primeira etapa da geração desses dados passa pela geração de números aleatórios a partir de uma distribuição **uniforme** entre 0 e 1. A partir dessa etapa, podemos gerar números aleatórios provenientes de qualquer distribuição de probabilidade conhecida, como função inversa da sua respectiva **função de distribuição acumulada** (HILLIER *et* LIEBERMANN, 1974, cap. 15, seção 15.2, p. 628-631).

Surge então a questão sobre qual é a distribuição de probabilidade que melhor descreve o comportamento de uma variável aleatória cujos valores são dados de entrada do modelo de simulação. A primeira medida é a observação da operação do sistema real, e a coleta de dados da variável em questão, em quantidade suficiente para compor uma amostra de tamanho razoável para permitir análise consistente dos mesmos. Quando não é possível coletar diretamente esses dados (se o sistema não existe ainda), é necessário pesquisar quais das variáveis existentes permitem levantamento de dados necessários para atender razoavelmente às especificações do sistema, tarefa mais complexa nesse

[18] Para descrição detalhada sobre variância e desvio-padrão, consultar COSTA NETO (2002), cap 2, seção 2.3.3, pág. 24-25; e apêndice A1.2.5, pág. 233-234

[19] Para descrição completa das distribuições de probabilidade discretas e contínuas, consultar COSTA NETO, (2002), apêndices A1.3 e A1.4, pág. 234-238.

PUC-Rio - Certificação Digital Nº 0612533/CA

caso[20]. Em seguida, deve-se avaliar a validade desses dados (influência de condições passadas que não mais se aplicam, por exemplo, ou de novas condições, que não existiam no passado) para descrever o comportamento futuro da variável aleatória em questão. O uso dos dados coletados diretamente como entrada na simulação tem o inconveniente de não permitir constatar efeitos de valores críticos da variável, que não foram registrados, mas que podem ocorrer na prática. Além disso, os dados coletados podem não ser suficientes para gerar simulações em número necessário para compor uma amostra de tamanho apropriado para análise (KELTON, SADOWSKI, *et* STURROCK, 2004, cap. 4, seções 4.5.2-3, pág. 154-156).

É aconselhável também o exame de **valores atípicos** dos dados de entrada (*outliers*), seguido de uma análise de **correlação**, para verificar se a amostra é constituída de elementos **independentes** (não influenciados pelas ocorrências anteriores) e **identicamente distribuídos** (possuem a mesma distribuição de probabilidade), conforme CHWIF *et* MEDINA (2006, cap. 2 seção 2.3.2, pág. 28-32).

É preferível escolher uma distribuição de probabilidade para representar uma variável aleatória de entrada, em vez de utilizar os dados diretamente, evitando-se as restrições citadas anteriormente. Para determinar qual das distribuições de probabilidade conhecidas representa melhor essa variável aleatória, com base nos dados que foram coletados, existem, nos modernos aplicativos de simulação, utilitários específicos de análise estatística para esse fim[21], os quais fornecem histogramas e as estimativas dos respectivos parâmetros dessas distribuições, além de indicadores que medem o ajuste (aderência) da distribuição de probabilidade aos dados coletados. Com o avanço dos aplicativos de simulação, é possível efetuar essa análise para todas as distribuições **teóricas**[22] conhecidas, obtendo-se, de uma vez só, os resultados correspondentes a cada distribuição de probabilidade, classificados conforme o valor do **indicador de aderência** da distribuição para representar os dados (PRADO, 2004, cap. 22, seção 22.1, pág. 192-193).

[20] Como exemplo, para um estudo de simulação da operação de um terminal em projeto, foram utilizadas as chegadas de navios no porto existente na cidade próxima (PETROBRAS, 2006).
[21] O *software* ARENA®, por exemplo, possui o *Input Analyzer*. O *software* PROMODEL®, possui o *Stat-Fit*.
[22] As distribuições exponencial, triangular, Weibull, beta, Erlang, gama, lognormal, normal, uniforme (contínuas) e a distribuição de Poisson (discreta).

PUC-Rio - Certificação Digital Nº 0612533/CA

Os valores dos indicadores de aderência são obtidos de **testes de aderência**, que são essencialmente **testes de hipótese**[23] sobre a validade da adequação da distribuição para representar a variável aleatória cujos dados são coletados. Os mais importantes são o **teste do Qui-quadrado** e o **teste de Kolmogorov-Smirnov**, cujos fundamentos se encontram em BANKS *et al.* (2000, cap.7, seção 7.4.1, pág. 266-267 e cap.9, seções 9.4.1-4, pág. 343-351). O **erro quadrático médio**[24] é também um indicador de aderência (FREITAS FILHO, 2001, cap.5 seção 5.7.1, pág. 169-170).

É possível também eleger uma distribuição **empírica**[25] (discreta ou contínua), no caso de não se conseguir obter ajuste satisfatório de nenhuma das distribuições teóricas conhecidas aos dados de entrada analisados. A alternativa geralmente é satisfatória, contanto que a amostra não seja de tamanho reduzido. Uma das razões para a falta de ajuste das distribuições aos dados é a existência de mais de um pico de freqüência de dados na amostra (amostra **multimodal**). A recomendação, nesse caso, é a subdivisão da amostra, por conta da provável existência de perfis distintos de comportamento dos dados na amostra original. Estas e outras recomendações acerca da decisão de utilização de distribuições teóricas ou empíricas são encontradas em KELTON, SADOWSKI, *et* STURROCK (2004, cap. 4, seções 4.5.4-5, pág. 161-165). Entre elas, os autores recomendam também o procedimento para escolha da distribuição de probabilidade, no caso de não haver disponibilidade de dados para coleta. Essa é a situação quando o sistema em estudo está ainda em fase de projeto, por exemplo. Nesse caso, é muito importante uma análise de sensibilidade sobre essas variáveis de entrada, cujos dados não estão disponíveis. Visando avaliar o impacto sobre o sistema, efetuam-se simulações, atribuindo cenários respectivos para valores determinísticos da variável em questão, ou escolhe-se uma distribuição, com base na natureza da informação representada por essa variável (Tabela 3.1.1). Nesse

PUC-Rio - Certificação Digital Nº 0612533/CA

[23] Abordagem conceitual estatística sobre testes de hipóteses pode ser consultada em COSTA NETO (2002), cap. 5, seções 5.1-5.2, pág. 83-88.
[24] O erro quadrático médio é definido como o valor médio dos quadrados das diferenças entre o valor da freqüência dos dados da amostra correspondente a determinada célula do histograma e o valor da freqüência relativa da distribuição ajustada, referente à mesma célula.
[25] Uma descrição das distribuições empíricas discreta e contínua se encontra em KELTON, SADOWSKI *et* STURROCK (2004), apêndice D, pág. 622-624.

caso, uma escolha adequada dos parâmetros da distribuição de probabilidade deverá ser feita[26].

Tabela 3.1.1: Distribuições aplicáveis quando não há disponibilidade de dados

Distribuição	Parâmetros	Características	Exemplos de uso
Exponencial	Média	Variância elevada Limite inferior existe Limite superior não existe	Intervalos entre chegadas Ocorrência de defeitos (a taxa constante)
Triangular	Valores mínimo e máximo, e moda	Para distribuições simétricas ou não simétricas - possui limites inferior e superior	Tempos de serviço
Uniforme	Valores mínimo e máximo	Todos os pontos possuem a mesma probabilidade - possui limites inferior e superior	Para quando se tem pouco conhecimento sobre o processo

FONTE: KELTON, SADOWSKI *et* STURROCK, 2004

Alguns sistemas podem conter variáveis de entrada cujo comportamento é regido por um **processo não-estacionário de Poisson**[27]. Como exemplo desse comportamento, característico de incidências em horários de pico, citamos o afluxo de veículos formando congestionamento de tráfego. Geralmente a utilização de modelos desses sistemas, adotando-se processos estacionários de chegadas, fornece soluções de capacidade aquém da necessidade real.

Em uma simulação estocástica, os dados de entrada representados por variáveis aleatórias, por sua vez, originam resultados também representados por variáveis aleatórias. São variáveis dependentes, representadas por estatísticas básicas, como média, variância, etc. De modo que existirá variância associada a essas variáveis, e a precisão dos resultados será tanto **maior** quanto **menor** for essa variância. Não podemos eliminar totalmente a variância, a ponto de substituirmos uma modelagem de simulação estocástica pela determinística. Existem, entretanto, técnicas de **redução da variância** dos indicadores do modelo de simulação, sendo a mais comum, o aumento de **replicações** (corridas) da simulação[28]. As corridas devem conter resultados estatisticamente independentes entre si e identicamente distribuídos[29], devem seguir as mesmas regras, e usar os mesmos parâmetros e premissas, o que muda são os números aleatórios gerados

[26] Detalhamento sobre a técnica de escolha da distribuição de probabilidade na ausência de dados pode ser encontrado em LAW, McCOMAS *et* VINCENT, 1994, pág. 56-59.
[27] No APÊNDICE II.i.2 é explicada a formulação quantitativa do processo não-estacionário de Poisson.
[28] No APÊNDICE V se encontra a formulação matemática dessa técnica.
[29] Isto é, resultados de uma corrida não devem influenciar resultados de outra corrida.

PUC-Rio - Certificação Digital Nº 0612533/CA

para as variáveis aleatórias de entrada do modelo, em cada corrida. Uma corrida não deve iniciar retomando dados do final da corrida anterior, pois tal medida introduziria artificialmente uma **correlação** entre os resultados respectivos (KELTON, SADOWSKI, *et* STURROCK, 2004, cap.12, seção 12.4, pág. 516).

3.2
Avaliação das soluções

Existem considerações importantes sobre os dados de saída de uma simulação, as quais devem ser levadas em consideração durante a análise das soluções fornecidas pelo modelo. LAW *et* KELTON (1991, cap. 9, seção 9.1, pág 523) apontam para os grandes investimentos e tempo dedicados ao desenvolvimento da modelagem e da programação, em muitos estudos de simulação, em contraste com o pouco esforço dedicado à análise apropriada dos dados de saída da simulação. Uma prática comum é tratar os dados resultantes de uma única corrida da simulação (FREITAS FILHO, 2001, cap.1 seção 1.7), cuja duração é arbitrariamente escolhida, e proceder à formulação das estimativas do modelo a partir desses dados, interpretando assim erroneamente as características do sistema representado pelo modelo. Simulação é essencialmente um experimento estatístico conduzido em meios computacionais e, como tal, requer técnicas estatísticas apropriadas para a sua análise. A aplicação de tais técnicas requer por sua vez um número suficiente de replicações da simulação, para que os resultados da análise atinjam a precisão necessária para que as soluções do modelo de simulação permitam formular conclusões com base sólida, em termos estatísticos (CHWIF *et* MEDINA, 2006, cap. 6 seção 6.7, pág. 122).

De posse dos indicadores selecionados, dos cenários definidos, e das amostras de resultados correspondentes, é possível efetuar uma **comparação** entre as simulações correspondentes aos cenários estabelecidos, passo este crucial para subsidiar a tomada de decisão frente aos resultados obtidos. Pois é precisamente nessa fase que a técnica da simulação demonstra o seu valor, por permitir avaliar a eficácia das estratégias propostas para aplicação no sistema em estudo. Normalmente, os resultados correspondentes aos diversos cenários derivados dessas estratégias são comparados com aqueles que correspondem a um **cenário básico**, por meio do exame da variação dos principais **indicadores** que foram

PUC-Rio - Certificação Digital Nº 0612533/CA

selecionados do modelo de simulação[30]. As modificações que permitem caracterizar cenários distintos podem ser desde uma simples mudança do valor de um parâmetro do modelo até alterações na modelagem da lógica operacional do sistema.

Múltiplos cenários podem ser comparados entre si e com o cenário básico, obtendo-se por este método outras dimensões para análise do comportamento das soluções sob uma visão mais abrangente do problema. A agilidade necessária para a aplicação dessa estratégia é conseguida em virtude do avanço das técnicas presentes nos aplicativos de simulação modernos. O uso eficaz dessas técnicas dependerá da correta seleção do conjunto apropriado de indicadores, durante a fase de projeto do modelo[31].

Dois caminhos distintos se apresentam respectivamente para a solução dos modelos de simulação terminante e não-terminante. No caso de uma simulação terminante, pode-se realizar a simulação quantas vezes se desejar, e os procedimentos de análise são a medida da variância dos indicadores de performance selecionados, visando determinar os intervalos de confiança correspondentes, e verificar seu enquadramento dentro dos limites preestabelecidos. Já as simulações não-terminantes apresentam dois problemas táticos a serem previamente resolvidos. O primeiro deles é a identificação e o descarte das observações pertencentes à fase transiente da simulação, fase esta que não representa a realidade da operação do sistema[32] (FREITAS FILHO, 2001, cap.6 seções 6.7-6.7.1, pág. 217-220). Segundo esse autor, o problema da remoção da fase transiente possui quatro métodos (heurísticas) de solução:

- **Longa simulação**: consiste em efetuar uma simulação de duração longa o suficiente para garantir que a fase transiente termine antes da simulação.

---

[30] Conforme mencionado na seção 3.1.

[31] Conforme explicado no início da seção 3.1. Normalmente trata-se de variáveis do sistema, mas alguns parâmetros relevantes de controle também podem ser testados, como por exemplo, capacidades de recursos.

[32] Rigorosamente poderia ser suposta a existência de um estado “transiente” do sistema real, quando, por exemplo, observamos o seu retorno à operação, após um período de manutenção ou parada não programada. Isso não serviria de modo algum para justificar inclusão da fase transiente da simulação na análise, já que esta fase não é produto da modelagem da partida do sistema real, mas sim, apenas artificialmente resultante do início da simulação. A representação de paradas e retornos do sistema à operação acabaria ensejando a modelagem de uma simulação **terminante**.

PUC-Rio - Certificação Digital Nº 0612533/CA

A dificuldade está justamente na falta de uma indicação da duração necessária.

- **Inicialização apropriada**: inicialização das variáveis do sistema próximas das condições encontradas na fase de regime, procurando minimizar a duração da fase transiente. A dificuldade está na determinação dessas condições.

- **Truncamento**: dada uma amostra de *n* observações, efetuar a contagem *l* a partir da primeira observação, descartar a observação e recalcular os valores mínimo e máximo das observações restantes. O processo se repete até que a observação de ordem *l*+1 esteja entre o valor mínimo e o valor máximo das observações restantes.

- **Observação visual**: a partir da construção do gráfico da variação do indicador ao longo do tempo, procura-se observar qual o momento aproximado em que o comportamento do indicador se estabiliza. O exame de gráficos da **média móvel**[33] do indicador, de tamanhos diversos, facilita a identificação do momento de término da fase transiente. É o método mais indicado.

PUC-Rio - Certificação Digital Nº 0612533/CA

No caso de haver vários indicadores pertinentes ao modelo, o problema de remoção da fase transiente deverá ser repetido para cada indicador, pois a duração da fase transiente poderá ser diferente para cada um deles. Uma vez conhecidos os respectivos períodos de fase transiente, escolhe-se o maior deles, estabelecendo-o como a duração da fase transiente para a simulação inteira (CHWIF *et* MEDINA, 2006, cap. 6, seção 6.8.6, pág. 132).

O segundo problema é a determinação da duração total da simulação, a fim de permitir a obtenção da amostras de tamanho apropriado para análise dos indicadores selecionados (FREITAS FILHO, 2001, cap.6 seções 6.7 e 6.7.2, pág. 217 e 220-221). Segundo HILLIER *et* LIEBERMANN (1974, cap.15, seção 15.3, pág.641), devido à natureza dos problemas modelados, observações obtidas a partir de experimentos simulados são altamente correlacionadas[34]. No caso de sistemas terminantes, o problema fica resolvido efetuando-se replicações da

---

[33] Uma descrição dos fundamentos do cálculo da média móvel se encontra no APÊNDICE V.
[34] Um exemplo clássico é a correlação entre os tempos de espera de clientes consecutivos numa fila.

simulação. No caso dos sistemas não-terminantes, entretanto, a aplicação desse recurso poderá exigir tempo demasiado de simulação, a qual, nesse caso, é bem mais longa. Para esse tipo de problema existem métodos heurísticos específicos. Os principais são (FREITAS FILHO, 2001, cap.6 seções 6.7 e 6.7.2, pág. 221-226):

- **Replicações independentes**: consiste em efetuar várias replicações da simulação, descartando as observações pertencentes às respectivas fases transientes. A primeira dificuldade é a definição do critério de parada da simulação, já que o evento que caracteriza a parada não existe nos sistemas não-terminantes. Outro problema aparecerá se a fase transiente for muito longa, restando poucas observações da fase de regime, disponíveis para a análise.
- **Regeneração**: consiste em subdividir o período total da simulação (descartando a fase transiente) em partes com início marcado pela volta do sistema ao estado inicial (ponto de regeneração), como por exemplo, quando o sistema fica momentaneamente ocioso, e não há entidades na fila. Baseia-se na hipótese de que, em alguns sistemas, o comportamento dos indicadores não mais depende do seu comportamento anterior, após a volta do sistema a um estado equivalente ao estado inicial (regeneração do sistema). Cada subdivisão corresponde a um **ciclo de regeneração**, e a uma replicação da simulação para compor a amostra a ser analisada.
- **Loteamento**: consiste em subdividir o período total da simulação (descartando a fase transiente) em partes (lotes) que representem "replicações" da simulação, compondo a amostra a ser analisada. O tamanho do lote deverá ser suficiente, de modo que a correlação entre as médias de cada lote seja pequena, obtendo-se assim, razoável aproximação da condição de independência entre as médias dos lotes (e das "replicações" da simulação). A base do método está na hipótese de que a correlação entre as últimas observações de um lote só se propaga significativamente até as primeiras observações do lote seguinte, sendo "absorvida" quando o lote possui tamanho maior.

PUC-Rio - Certificação Digital Nº 0612533/CA

Tradicionalmente, os modelos de simulação não têm sido empregados para se avaliar uma solução ótima dos sistemas em estudo. As técnicas de otimização por programação linear têm sido utilizadas largamente neste tipo de problema. Com a modernização crescente dos programas de simulação, somada à capacidade cada vez maior de processamento dos computadores, a técnica de otimização de sistemas a partir de modelos de simulação se tornou realidade. Assim, já existem ferramentas que tornam possível efetuar a busca de uma solução "ótima", correspondente a uma combinação adequada de parâmetros do sistema, com a vantagem de se poder manter elevado grau de fidelidade aos detalhes do sistema, característica dos modelos de simulação (FREITAS FILHO, 2001, anexo I, seção A.2, pág. 299-300 e seção A.3, pág. 301-309).

3.3 Tratamento estatístico

PUC-Rio - Certificação Digital Nº 0612533/CA

Vimos que os dados de entrada pertencentes às variáveis aleatórias são submetidos a um tratamento estatístico, visando depurá-los, e em seguida representar cada variável aleatória por uma distribuição de probabilidade adequada. Como os dados de entrada de uma simulação são geralmente de natureza estocástica, originando resultados estocásticos correspondentes, existe a necessidade de aplicação de um tratamento estatístico às variáveis de saída (ou indicadores) da simulação, a fim de assegurar uma interpretação estatisticamente correta de resultados. A forma de tratamento dos valores dos indicadores obtidos do modelo de simulação é diferente daquela efetuada sobre os dados de entrada, pois trata-se de variáveis aleatórias **dependentes**[35] dos parâmetros e das variáveis de entrada do sistema.

De posse de um conjunto de replicações de simulações, contendo os mesmos parâmetros e premissas e com resultados **independentes entre si** e **identicamente distribuídos**, é possível compor uma **amostra** desses resultados, e formular estimativas dos indicadores de desempenho selecionados, para o sistema em estudo, com base na amostra obtida. A principal delas é a estimativa da média

[35] Diferentemente dos dados de entrada, nesse caso, por exemplo, não há utilidade no ajuste de distribuições de probabilidade aos valores experimentados pelos indicadores.

μ do valor do indicador[36], para a qual podemos formar um **intervalo de confiança**, com base na média (aritmética) do indicador, obtida da amostra, e no desvio-padrão do indicador, obtido da amostra. Para as simulações do tipo **terminante**[37], o intervalo de confiança é calculado pela fórmula (3.3.1) (LAW *et* KELTON, 1991, cap. 9, seção 9.4.1, pág 532-533).

$$\overline{X} \pm t_{n-1,1-\alpha/2} \, s/\sqrt{n} \qquad (3.3.1)$$

sendo: $\overline{X}$ a média do indicador obtida da amostra;

s o desvio-padrão do indicador obtido da amostra;

n o número de replicações da simulação;

α o nível de confiança adotado[38];

$t_{n-1,1-\alpha/2}$ valor crítico superior da distribuição t de Student[39] com *n-1* graus de liberdade para o nível de confiança α .

O intervalo de confiança definido acima deve ser interpretado do seguinte modo: para uma percentagem aproximadamente igual a α dos casos em que replicamos a simulação n vezes, o intervalo contém o valor verdadeiro (desconhecido) da média μ do indicador. A precisão da estimativa da média evidentemente será **maior** quanto **menor** for o intervalo de confiança obtido, o que pode ser conseguido com o aumento do número n de replicações da simulação[40] (KELTON, SADOWSKI, *et* STURROCK, 2004, cap. 2, seção 2.6.2, pág. 39-40 *et* cap.6, seção 6.3, pág. 260-263). A medida da precisão da estimativa é a **amplitude (unilateral)**[41] do intervalo de confiança, calculado para a média do indicador, conforme a fórmula acima. Normalmente se procura reduzir o valor dessa amplitude até 10% do valor da média do indicador, mediante o aumento do

[36] Teoricamente esta seria a média do indicador considerando número infinito de replicações.
[37] Simulações que operam durante determinado intervalo de tempo, ou sob condições de partida e parada, conforme definido no APÊNDICE II.i.2.
[38] Também chamado de nível de significância. Normalmente é adotado o valor 95%.
[39] Para descrição completa da distribuição t de Student, consultar COSTA NETO, (2002), cap. 3, seção 3.4.5, pág. 51-52
[40] Ver no APÊNDICE V as fórmulas pra cálculo do número de replicações necessário para redução da variância e do intervalo de confiança.
[41] Mais conhecida pelo termo “meia-largura”.

PUC-Rio - Certificação Digital Nº 0612533/CA

número de replicações da simulação (FREITAS FILHO, 2001, cap.6 seção 6.6.2, pág. 215).

Também é possível efetuar a comparação entre dois cenários, por meio de um teste de hipótese da **diferença** entre o valor do indicador em cada cenário (CHWIF *et* MEDINA, 2006, cap. 6, seção 6.9.1, pág. 140). Esse tipo de análise é particularmente útil como subsídio para tomada de decisão, por exemplo, entre aceitar ou rejeitar um investimento em ativos no sistema[42]. Nesse caso, o resultado do teste acima indicará se a alternativa de investimento proporcionará ou não uma diferença (melhoria) estatisticamente significativa no valor do indicador.

Na simulação, o estudo do impacto de variáveis de entrada sobre os indicadores do sistema é outra vertente de análise com base no tratamento estatístico de resultados. Nesse campo, a técnica conhecida como **projeto fatorial de experimentos** é muito eficiente para proporcionar uma estimativa dos principais efeitos, assim como da interação das próprias variáveis de entrada (designadas como **fatores**, no contexto de projeto fatorial de experimentos) (COX *et.* REID, 2000, Cap 5, seção 5.1, pág. 107). A busca por uma solução de custo mínimo aproximada também é possível, utilizando-se essa técnica. No CAPÍTULO 5 será ilustrada a utilização dos projetos fatoriais de experimento, nos estudos de dimensionamento da capacidade de armazenamento.

PUC-Rio - Certificação Digital Nº 0612533/CA

[42] Ou para verificar se o progresso na busca da solução de menor custo é significativo (CAPÍTULO 5).